Human-AI Collaboration in Cooperative Resource Management Simulations

Karthik Prasad [*1]

Abstract

This study investigates human-AI collaboration in resource management simulations by extending the GovSim framework to include human participants alongside AI agents. The project consists of 60 trials across three environmental scenarios (fishery, pasture, and pollution), and comparing four AI models (ChatGPT 3.5, ChatGPT 4o-mini, Claude 3 Haiku, and Claude 3.7 Sonnet) in their ability to collaborate with humans for sustainable resource management. The results demonstrate significant performance disparities between the four models, with Claude 3.7 Sonnet substantially outperforming others through superior mathematical reasoning, consistent communication, and adaptive resource management strategies. This research highlights how trust formation between humans and AI is bidirectional and reveals critical areas for improvement in current AI systems to facilitate effective human-AI collaboration in complex strategic environments.

1. Introduction

This project aims to extend the GovSim simulation (Piatti et al., 2024) where multiple AI agents must cooperate to balance a common resource and sustain it for future use without exploiting it. The original research paper consisted of LLM agents like Claude, ChatGPT and Llama facing off against each other in three different scenarios. This paper extends upon that by adding a twist of human collaboration into it. Instead of multiple LLM agents facing each other, a singular LLM agent faces off against multiple humans in the same scenario. Normally, these agents would have to discuss and cooperate with other purely rational agents, but in this project they have to now deal with human players who may have irrational decision-making and different strategies in different cycles and phases. In general, the project investigates how these AI agents can collaborate with human players in the same type of environment and examines the dynamics of human-AI teamwork in complex strategic settings.

2. Background

The GOVernance of the commons SIMulation (GOVSIM) is a generative simulation platform designed to study strategic interactions and cooperative decision-making in LLMs. In GOVSIM, a society of AI agents must collectively balance exploiting a common resource with sustaining it for future use. Agents are given a common pool of natural resources that regenerates over time. The task is to sustainably manage the use of this resource. Take too much, and the resource will collapse and no longer regenerate again (e.g., the fish in a lake go extinct). Take too little, and the resource's economic potential is underutilized. Even a purely selfish agent that aims to maximize their long-term reward must balance the amount of resources they extract now with what they will be able to extract in the future.

This research is based on several areas of research. Multi-agent reinforcement learning has emerged as an immensely interesting area of research in recent years. It extends upon single agent reinforcement learning to incorporate scenarios where multiple agents have to interact with each other in a shared environment. A paper written quite a while ago provides an excellent overview of past research in multi-agent reinforcement learning (Buşoniu et al., 2010). They discuss the benefits that come from agents communicating with each other by teaching, imitating and sharing their experiences. On the other hand, they also talk about the challenges that come with the curse of dimensionality as well as the exploration and exploitation tradeoffs.

A paper in 2019 discusses the effectiveness of multi-agent decision making in common pool resource management systems (Zhu & Kirley, 2019). They used an abstract mathematical model of the system represented as a partially-observable general-sum Markov game. Their results showed that agents generally do not have perfect foresight nor understanding of the implications of their harvesting efforts yet, but as the models improve, they are highly likely to be interacting with other learning-agents and humans in complex partially competitive settings. The GovSim framework is one of the latest experiments in a long line of work in this research area that continues to display the potential of multi-agent reinforcement learning in shared resource management environments.

It is also important to explore the history of human-AI collaboration in different games and strategic environments, with recent research demonstrating its potential to enhance performance and learning outcomes. A recent study indicated that the performance of AI models is similar to, and in certain application areas, even exceeds the performance of human experts (Hemmer et al., 2024). However, there are some domains where human predictions often remain more accurate. Thus, the potentially complementary capabilities of human-AI collaboration have the potential to further improve overall decision-making performance. In a recent study, human-AI teams outperformed both human-controlled and fully autonomous AI agents with policy correction proving particularly effective for agent learning (Islam et al., 2023). These findings highlight the potential for human-AI collaboration to improve outcomes across a wide variety of domains from workplace tasks to educational environments. Accordingly, it is interesting to see how well human-AI teams perform in the simulation, and how they compare against individual AI agents in the different GovSim scenarios.

3. Methodology

To obtain comparable results to the orginal research paper (Piatti et al., 2024), this project consists of the same fishery, pasture and pollution scenarios from the GovSim simulation and contains the same underlying resource dynamics for every scenario. It also commences from the exact same text prompts and provides the same type of feedback and prompting to the AI and human participants. On the other hand, there are some key differences between the original research and this one:

- There is only one AI LLM agent instead of multiple agents
- There are three human agents also participating
- The AI agent will be discussing strategies with the human agents instead of with the other AI agents

This project maintains a similar layout to the original GovSim framework and comprises of the following structure. A trial has a designated number of cycles that it runs through, with each cycle representing a month where agents can harvest a certain number of resources. Each cycle has three phases:

- Harvesting: all the agents are notified about how many resources are currently available and are asked how many resources they would like to harvest. Each of the agents have to provide their answer without having any knowledge of what the other agents stated. If agents ask for more resources than are available, the experiment fails. If the number of resources is below the minimum threshold after all the agents consume the resource, the simulation fails.
- Discussion: all the agents are notified about how many resources were harvested from each agent and the resultant amount of resources. They can then speak with each other to strategize for the upcoming cycles.
- Update: The simulation continues to the next month and the resource is updated to the next cycle by being doubled up to a max of 100.

The original plan was to clone the GovSim project and extend it to include human agents as the authors had provided a public repository for others to use. However, upon further investigation, the efforts needed to rework the existing code to incorporate human-AI interactions was quite complicated. In addition, it was quite difficult to understand some sections of the authors' codebase. Therefore, it was decided to recreate the codebase from the ground-up while building in the human-AI interactions into the simulation. The codebase was created using Python, with three individual files holding the functionality for each environment, a parameters file for fine-tuning and messaging prompts as well as a main file that would run all the trials and output the data. The data from each trial would be output into a csv file in addition to being displayed in the console window. This system design allows for others in the future to easily extend upon this work by adding in different scenarios for the main file to run, adjusting the main file to run the scenarios differently, or exploring further by tuning the different parameters.

Multiple trials were performed with different agents and models in order to understand how well different LLMs collaborate with humans in these environments. For the scope of this project, two different models from two different providers will be compared:

- ChatGPT 3.5
- ChatGPT 4o-mini
- Claude 3 Haiku
- Claude 3.7 Sonnet

Models from two different providers were tested to compare the effectiveness of different types of model implementations by each provider. An older and newer version of each provider's models were also evaluated to see how much the AI agent had improved over time as newer versions were developed. Together, this provided a sufficient variety of different AI agents to compare with.

Table 1. Fishery Environment Trial Results

AI MODEL	SUCCESSES	FAILURES
CHATGPT 3.5	0	5
CHATGPT 4O-MINI	0	5
CLAUDE 3 HAIKU	0	5
CLAUDE 3.7 SONNET	3	2

Table 2. Fishery Environment Data Summary

AI MODEL	MAX MONTHS	AVG MONTHS
CHATGPT	4	3
CHATGPT 4O-MINI	8	4
CLAUDE 3 HAIKU	5	3
CLAUDE 3.7 SONNET	12	10

Table 3. Pasture Environment Trial Results

AI MODEL	SUCCESSES	FAILURES
CHATGPT 3.5	0	5
CHATGPT 4O-MINI	0	5
CLAUDE 3 HAIKU	0	5
CLAUDE 3.7 SONNET	3	2

Table 4. Pasture Environment Data Summary

AI MODEL	MAX MONTHS	AVG MONTHS
CHATGPT	4	3
CHATGPT 4O-MINI	9	4
CLAUDE 3 HAIKU	6	3
CLAUDE 3.7 SONNET	12	9

Various types of quantitative and qualitative data are collected during these trials. Firstly, it is stored whether each trial was a success or failure, where success is determined by the trial making it past the predetermined number of cycles without falling under the minimum resource threshold. Additionally, the number of resources collected by each agent at each cycle as well as the total resources collected are stored and can be found in Appendix A-C. These results can then be summarized and visualized to easily understand how the agents collaborate with each other. The average and maximum number of months survived by the AI agent in each trial is also noted.

On the other hand, qualitative data is also collected from both the AI and human agents. After each trial is finished in either a success or failure, the AI agent is asked about its thoughts on how the trial went and what could have been done differently. The resulting answer is stored and provides some meaningful insights into the AI agent's strategies during the trial and their reflection after the trial is finished and can be found in Appendix D-F. Furthermore, logs are stored of all the AI agent's communications with the human agents during the discussion phase of each cycle. For the human participants, they are also asked about their thoughts on the trials and what they may have done differently after the end of each failed or completed cycle. After the whole experiment is finished, they are also asked about their thoughts about the different AI models and their general experience going through the trials and collaborating with AI agents.

4. Results

Tables 1-4 portray the raw data collected from the 60 trials, which consist of 5 trials for each of the 3 environments with each of the 4 models used.

Tables 1 and 2 summarize the results from the fishery environment, where 4 fishermen (3 humans and 1 AI agent) have to cooperate together to catch fish in a lake. In the trials, both ChatGPT 3.5 and ChatGPT 4o-mini failed every trial, with most trials failing in the first few months and the newer ChatGPT 4o-mini having the longest run at 8 months. Claude 3 Haiku also performed very similarly to ChatGPT 3.5, but Claude 3.7 Sonnet performed significantly better with 3 successes and 2 failures. This model could consistently catch the optimal amount of fish and survived an average of 10 months between the 5 trials.

Tables 3 and 4 display the data taken from the trials for the pasture environment, where 4 shepherds (1 AI agent and 3 humans) bring a certain amount of sheep to the pasture to graze and consume the grass in the field. It utilizes the same core mechanics as the fishery environment, but with a different context. Similar to the fishery environment, all three of ChatGPT 3.5, ChatGPT 4o-mini and Claude 3 Haiku performed poorly as none of them could even come close to passing a single trial. On the other hand, Claude 3.7 Sonnet performed very well compared to the other three models with 3 out of 5 trials resulting in success and averaging a 9 month run during its trials.

Tables 5 and 6 exhibit the last of the three environments, where four factory owners (3 humans and 1 AI agent) decide how many pallets of widgets to produce each month. For each pallet produced, they consume 1 percent of an unpolluted lake by polluting the water. An interesting observation to note is that all the models struggled a lot more in this scenario compared to the other two. This may be due to the higher complexity in language used to explain this scenario, as terms like "pallets of widgets produced" and "pollute a percent of unpolluted water" were used instead of simpler times like "catch fish". Like the other two scenarios, 3 of the 4 models (ChatGPT 3.5, ChatGPT 4o-mini, Claude 3 Haiku) couldn't get 12 months into a trial, but due to the increased difficulty in comprehending the situation, all three of the

Table 5. Pollution Environment Trial Results

AI MODEL	SUCCESSES	FAILURES
CHATGPT 3.5	0	5
CHATGPT 4O-MINI	0	5
CLAUDE 3 HAIKU	0	5
CLAUDE 3.7 SONNET	4	1

Table 6. Pollution Environment Data Summary

AI MODEL	MAX MONTHS	AVG MONTHS
CHATGPT	2	2
CHATGPT 4O-MINI	4	2
CLAUDE 3 HAIKU	2	2
CLAUDE 3.7 SONNET	12	10

models averaged only 2 months of survival. In comparison, Claude 3.7 Sonnet performed extremely well, passing 4 out of 5 trials with an average of 10 months. In fact, the median of this model's trials is at 12 months, with the one failed trial reaching 4 months.

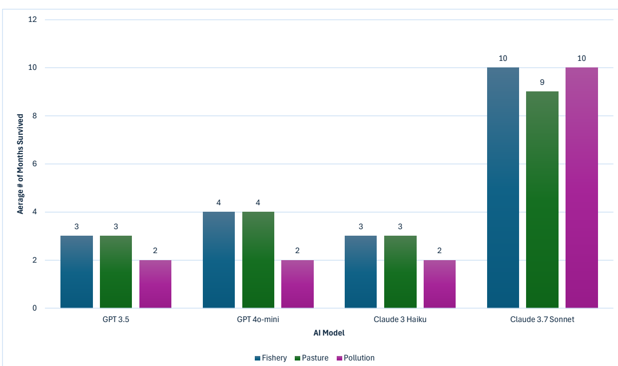

Figure 1. Average Months Survived by each AI Model

Due to all three scenarios utilizing the same core mechanics and requiring similar resource management strategies, the average months survived by each AI agent is similar for each scenario as shown in figure 1, as there is only a variance of ±1 month between ChatGPT 3.5, ChatGPT 4o-mini and Claude 3 Haiku. This clearly portrays the differences between the AI agents, particularly between Claude 3.7 Sonnet and the three other models, as it survives about 3 times more months on average compared to the other models. While figure 1 shows the strength of the other 3 agents to be similar, the difference can be seen a lot more clearly in figure 2. Looking at which AI agent could survive the highest number of months indicates a clear hierarchy of strength with ChatGPT 3.5 being the weakest, Claude 3 Haiku in the middle and GPT 4o-mini the strongest of the three.

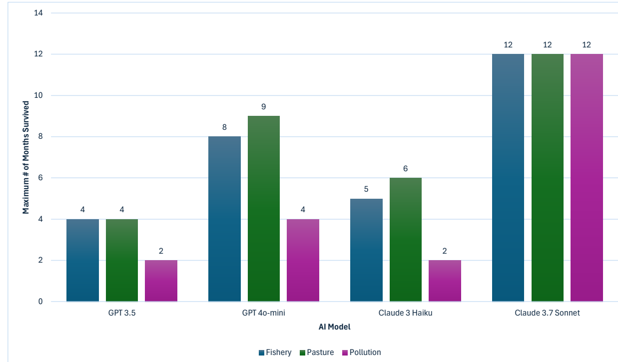

Figure 2. Max Months Survived by each AI Model

Figure 3. Total Units of Resources Harvested by each AI Model

It was also interesting to observe how many resources could be harvested in total. Figure 3 shows how many total resources were harvested by each AI agent over all the trials they participated in, divided into three segments for each scenario. As Claude 3.7 Sonnet performed so optimally compared to the others, it harvested more total resources than the other models by far. However, it was intriguing to see that the supposed weakest model performed the second best. This may be attributed to the extremely greedy approach that the model took. Due to its lack of mathematical reasoning, GPT 3.5 would take an absurd amount of resources in the initial months, causing the simulation to fail (e.g. it would take 50 out of the 100 resources in the pool despite there being 4 participants in the trial). However, it took so much in these initial months that it still ended up with more resources than GPT 4o-mini and Claude 3 Haiku who were also pretty short-sighted and failed the simulation in a few more months. The only reason it almost has as many resources as Claude 3.7 Sonnet is that the simulation had a max of 12 months, and it can be confidently estimated that if the max was at 50 or 100 months, then Claude 3.7 Sonnet would drastically surpass the total resources harvested by any of the three other models due to its long-term oriented strategy.

5. Discussion

So many valuable insights can be obtained from these experiments as they demonstrated the significant differences in the performances of the various AI models when tasked with collaborative resource management scenarios. Most of the AI models in the study failed to successfully manage the common resource pool, primarily due to the fundamental limitations in their situation comprehension abilities. When presented with the complex dynamics of the simulation, these models clearly struggled to fully grasp the nature of resource consumption and long-term implications. This resulted in shortsighted decision-making that made the models prioritize immediate gains over more sustainable resource management, which always ended up in a failed simulation with the resource completely depleted.

For some of the models, mainly ChatGPT 3.5 and Claude 3 Haiku at times, mathematical reasoning proved to be a substantial obstacle to their success. They were simply unable to perform the basic computational tasks required to develop the optimal strategy to succeed. The models had a tendency to harvest excessive resources in the initial phases, making sustainable management mathematically impossible in subsequent rounds regardless of the decisions of the other participants. For example, ChatGPT 3.5 would repeatedly harvest 50 percent of the resource pool in the initial months, which would always result in failure in the coming months.

In comparison, ChatGPT 4o-mini and Claude 3.7 Sonnet could calculate the optimal strategy fairly well. However, ChatGPT 4o-mini's most concerning issue was the inconsistency between its communication and its actions. While it would verbally advocate for conservative resource use and equitable distribution, its actual resource consumption contradicted these statements and it would harvest a lot more resources than what it stated it would in the previous discussion phase. This discrepancy between the model's expressed intentions and its actions would actually create mistrust between it and the collaborative group, as the human participants would grow to be more distrusting of the AI model and would take more than it told them to take as well.

On the other hand, Claude 3.7 Sonnet consistently demonstrated superior performance compared to the other 3 models in all metrics. Its communication abilities were particularly remarkable, as it proactively addressed emerging issues by constructively communicating with the human participants. Even when other human participants would make errors in calculations or judgement, Claude 3.7 Sonnet provided clear and constructive feedback that helped them realign with the group's sustainable practices.

On top of its excellent communication, Claude 3.7 Sonnet exhibited very authentic collaborative behavior by actually consuming fewer resources than human participants at times. This self-imposed restraint sometimes allowed the simulation to succeed despite some occasional overuse by others that would have led the simulation to fail with other models. This really highlights the model's ability to adapt its strategy dynamically based on the evolving resource situation rather than just performing the mathematically optimal strategies. These behaviors of the model really show its advanced understanding of not only the mathematical strategies needed to succeed in the scenario, but also the social dynamics of group resource management that it navigated so effectively.

There are some interesting insights that could be gathered from the involvement of the human participants, as involving them in the simulation changed the entire dynamic of the simulation. This is because of the increased variability in human decision-making that can sometimes be very random and irrational, as well as based on a lot more complex factors compared to the rational and logical decision making process of AI models. Accordingly, the AI agent would have to dynamically react to the increased complexity of the situation due to the addition of the human participants. The human participants also demonstrated more sophisticated long-term planning capabilities as well as susceptibility to emotional responses too. The human participants occasionally made suboptimal decisions despite knowing the correct mathematical decision, due to reasons like risk tolerance, competitive instances and responses to perceived scarcity.

A particularly interesting observation was how humans adapted their behavior based on their perception of their AI collaborator. When paired with advanced models like Claude 3.7 Sonnet, the human participants displayed a greater willingness to follow its suggestions and coordinate actions with it. On the contrary, they displayed more skepticism and were more defensive when collaborating with AI that they found to be less capable and had inconsistencies between their communications and their actions. This highlights how trust formation between humans and AI is actually bidirectional and dependent on their performances, as the humans calibrated their collaboration strategies based on the perceived reliability of their AI counterparts.

There were some limitations to this experiment that could unfortunately not be avoided. The biggest limitation was the total number of human participants available for the experiment, as the same people would be used for every trial for every model. This distorted the data as a participant who has gone through a large number of trials would learn how to perform optimally and perform better in their 50th trial compared to their 1st trial due to their previous learnings of the resource management system and the behavioral patterns of the AI models. Furthermore, this also limits the generalization of the experiments as the results reflect how those specific individuals with their own cognitive style, personality and background interacted with the AI models. Ideally, a larger and more diverse sample of the population would provide a better understanding of how humans in general would perform in the experiment.

On top of this, the human participants may also demonstrate fatigue and decreases in motivation after such a large number of trials, which would affect their decisions and the resultant data. One observation regarding this came from one of the later trials with Claude 3.7 Sonnet, where the human participants became increasingly more greedy to test if the AI model could still keep the environment sustainable as they were bored from the model easily passing the past few simulations. Another significant constraint was the number of models that could be used. Due to financial reasons, only two different providers were used as it would cost extra to buy the pro plan for each additional provider. Ideally, a lot more models could be used to compare many more providers like LLama, Mixtral and Qwen and provide a more comprehensive understanding of how different AI models perform in resource management systems.

6. Conclusion

Overall, this study provides a lot of valuable insights into the collaborative capabilities of the different AI systems when they are tasked with resource management scenarios that require mathematical reasoning, effective communication and strategic thinking. The results of the experiment reveal a significant performance gap between the 4 AI models that were tested, as most of the AI models could not pass a single trial while the most advanced system Claude 3.7 Sonnet performed significantly better and could consistently pass the trials. The other 3 AI models had difficulties in situation comprehension, mathematical reasoning and inconsistencies between their communicated intentions and their resulting actions. This highlights the critical areas that require attention in the near future of AI development, as these deficiencies significantly undermined their ability to perform well in the trials.

The impressive performance of Claude 3.7 Sonnet was due to its precise mathematical reasoning, consistent alignment between its communication and its actions as well as its adaptive resource consumption strategies. It is capable of identifying emerging errors, providing constructive feedback, and regulating its own resource consumption to compensate for over usage of resources by others. It demonstrates the immense potential that AI systems have to become valuable collaborators with humans in a multitude of applications.

The involvement of human participants revealed some interesting dynamics in human-AI collaboration, as humans displayed greater variability in decision making and adaptive behaviors based on their perception of the AI model's reliability. This bidirectional trust formation between the humans and the AI models highlights the importance of developing AI systems that demonstrate a consistently high level of performance and communication in facilitating effective human-AI collaboration. As this develops, AI models can hopefully cooperate and assist humans in their tasks, as human-AI collaboration requires lower mental and temporal demands, reducing human effort, and yielding higher performance than if humans directly controlled all agents (Islam et al., 2023).

Further research should address the limitations of this experiment and expand on the number of AI models and the number of human participants to provide more comprehensive results. Additionally, it can also be investigated how well the AI's collaborative abilities can be extended to more complex, real world resource management scenarios with less structured interactions. In conclusion, there are still some significant challenges that remain in developing AI systems that are capable of effectively collaborating with humans, but the performance of models like Claude 3.7 Sonnet provides some very promising evidence that AI can actually meaningfully collaborate with humans. With continued improvements to the development of AI models, these AI systems truly have the potential to become valuable partners not only in resource management systems but other applications where they can support and collaborate with humans to achieve greater results together.

References

Buşoniu, L., Babuška, R., and De Schutter, B. Multi-agent reinforcement learning: An overview. In *Innovations in Multi-Agent Systems and Applications – 1*, volume 310 of *Studies in Computational Intelligence*, pp. 183–221. Springer, 2010. doi: 10.1007/978-3-642-14435-6_8.

Hemmer, P., Schemmer, M., Kühl, N., Vössing, M., and Satzger, G. Complementarity in human-ai collaboration: Concept, sources, and evidence. *arXiv preprint arXiv:2404.00029*, 2024. doi: 10.48550/arXiv.2404.00029.

Islam, M. S., Das, S., Gottipati, S. K., Duguay, W., Mars, C., Arabneydi, J., Fagette, A., Guzdial, M., and Taylor, M. E. Human-ai collaboration in real-world complex environment with reinforcement learning. *arXiv preprint arXiv:2312.15160*, 2023. doi: 10.48550/arXiv.2312.15160.

Piatti, G., Jin, Z., Kleiman-Weiner, M., Schölkopf, B., Sachan, M., and Mihalcea, R. Cooperate or collapse: Emergence of sustainable cooperation in a society of llm agents. *arXiv preprint arXiv:2404.16698*, 2024. doi: 10.48550/arXiv.2404.16698.

Zhu, H. and Kirley, M. Deep multi-agent reinforcement learning in a common-pool resource system. In *Proceedings of the 2019 IEEE Congress on Evolutionary Computation (CEC)*, pp. 1001–1008. IEEE, 2019. doi: 10.1109/CEC.2019.8790001.

Appendix

A. Fishery Results

Model	ChatGPT 3.5		Trial:	1
Month	Agent	Percy	Hazel	Frank
1	25	15	15	10
2	25	10	5	5
3	20	20	10	10
Total	70	45	30	25

Model	ChatGPT 3.5		Trial:	2
Month	Agent	Percy	Hazel	Frank
1	25	10	10	10
2	20	15	15	15
3	10	10	10	10
4	5	7	7	7

Model	ChatGPT 3.5		Trial:	3
Month	Agent	Percy	Hazel	Frank
1	25	12	10	12
2	15	11	10	10
3	20	12	14	14
4	10	8	8	8

Model	ChatGPT 3.5		Trial:	4
Month	Agent	Percy	Hazel	Frank
1	25	15	15	15
2	20	14	14	14

Model	ChatGPT 3.5		Trial:	5
Month	Agent	Percy	Hazel	Frank
1	30	10	10	10
2	20	15	15	15
3	10	10	10	10

Figure 4. Trial Data for Fishery Environment with ChatGPT 3.5

Model	ChatGPT 4o-mini		Trial:	1
Month	Agent	Percy	Hazel	Frank
1	30	12	12	12
2	30	10	10	10
3	10	5	5	5

Model	ChatGPT 4o-mini		Trial:	2
Month	Agent	Percy	Hazel	Frank
1	30	12	12	14
2	20	10	10	10
3	10	5	5	5

Model	ChatGPT 4o-mini		Trial:	3
Month	Agent	Percy	Hazel	Frank
1	20	15	15	15
2	20	6	6	8
3	15	6	6	8
4	10	8	8	10
5	5	3	3	3
6	3	3	3	3
7	5	5	6	8
8	2	4	4	4

Model	ChatGPT 4o-mini		Trial:	4
Month	Agent	Percy	Hazel	Frank
1	30	15	15	15
2	20	12	15	12

Model	ChatGPT 4o-mini		Trial:	5
Month	Agent	Percy	Hazel	Frank
1	20	10	12	10
2	24	12	15	12
3	20	13	16	13

Figure 5. Trial Data for Fishery Environment with ChatGPT 4o-mini

Model	Claude-3-haiku		Trial:	1
Month	Agent	Percy	Hazel	Frank
1	50	12	12	12
2	7	6	6	6

Model	Claude-3-haiku		Trial:	2
Month	Agent	Percy	Hazel	Frank
1	25	15	15	15
2	20	10	10	10
3	10	5	5	5

Model	Claude-3-haiku		Trial:	3
Month	Agent	Percy	Hazel	Frank
1	50	12	12	14
2	6	4	4	4
3	3	3	3	3

Model	Claude-3-haiku		Trial:	4
Month	Agent	Percy	Hazel	Frank
1	50	10	10	10
2	10	12	5	5
3	4	3	3	3

Model	Claude-3-haiku		Trial:	5
Month	Agent	Percy	Hazel	Frank
1	50	12	10	14
2	7	4	6	6
3	2	1	1	1
4	2	1	1	1
5	2	2	2	2

Figure 6. Trial Data for Fishery Environment with Claude 3 Haiku

Model	Claude-3-sonnet		Trial:	1
Month	Agent	Percy	Hazel	Frank
1	12	12	12	12
2	12	12	12	12
3	12	12	12	12
4	12	12	12	12
5	12	12	12	12
6	12	12	12	12
7	12	12	12	12
8	12	12	12	12
9	12	12	12	12
10	12	12	12	12
11	12	12	12	12
12	12	12	12	12

Model	Claude-3-sonnet		Trial:	2
Month	Agent	Percy	Hazel	Frank
1	20	12	12	12
2	15	16	16	16
3	10	10	10	10
4	4	5	5	5

Model	Claude-3-sonnet		Trial:	3
Month	Agent	Percy	Hazel	Frank
1	13	12	12	12
2	13	13	13	13
3	12	12	12	12
4	12	12	12	12
5	12	12	12	12
6	12	13	13	13
7	12	13	13	13
8	11	13	13	13
9	8	8	8	8
10	8	8	8	8
11	6	8	8	8

Model	Claude-3-sonnet		Trial:	4
Month	Agent	Percy	Hazel	Frank
1	12	12	12	12
2	12	13	13	13
3	12	13	12	12
4	12	13	13	12
5	11	13	13	13
6	10	12	12	12
7	11	11	11	11
8	11	13	11	11
9	12	13	13	12
10	12	14	12	12
11	12	12	13	13
12	12	13	13	13

Model	Claude-3-sonnet		Trial:	5
Month	Agent	Percy	Hazel	Frank
1	15	14	12	16
2	12	13	12	14
3	5	6	5	6
4	12	12	11	13
5	12	12	12	13
6	11	13	12	14
7	9	12	11	13
8	9	10	9	10
9	11	11	11	12
10	11	13	13	12
11	11	14	12	14
12	11	15	14	15

Figure 7. Trial Data for Fishery Environment with Claude 3.7 Sonnet

B. Pasture Results

Model	ChatGPT 3.5		Trial:	1
Month	Agent	Percy	Hazel	Frank
1	50	15	15	15
2	0	2	2	2

Model	ChatGPT 3.5		Trial:	2
Month	Agent	Percy	Hazel	Frank
1	30	12	10	10
2	36	10	9	10
3	10	4	4	3

Model	ChatGPT 3.5		Trial:	3
Month	Agent	Percy	Hazel	Frank
1	30	10	11	11
2	24	15	10	13
3	9	4	6	4
4	3	2	2	2

Model	ChatGPT 3.5		Trial:	4
Month	Agent	Percy	Hazel	Frank
1	30	15	15	10
2	20	14	14	10

Model	ChatGPT 3.5		Trial:	5
Month	Agent	Percy	Hazel	Frank
1	50	10	8	12
2	30	6	6	8

Figure 8. Trial Data for Pasture Environment with ChatGPT 3.5

Model	ChatGPT 4o-mini		Trial:	1
Month	Agent	Percy	Hazel	Frank
1	30	12	12	12
2	25	12	12	12
3	10	2	2	2

Model	ChatGPT 4o-mini		Trial:	2
Month	Agent	Percy	Hazel	Frank
1	50	14	12	12
2	12	6	4	4

Model	ChatGPT 4o-mini		Trial:	3
Month	Agent	Percy	Hazel	Frank
1	25	10	8	9
2	24	12	10	10
3	20	14	13	13
4	10	6	6	5
5	5	5	3	3
6	5	4	3	3
7	2	1	1	1
8	1	1	1	1
9	2	4	2	3

Model	ChatGPT 4o-mini		Trial:	4
Month	Agent	Percy	Hazel	Frank
1	25	15	12	12
2	20	15	13	13
3	10	8	4	4

Model	ChatGPT 4o-mini		Trial:	5
Month	Agent	Percy	Hazel	Frank
1	30	14	14	12
2	20	10	10	10
3	10	3	2	2

Figure 9. Trial Data for Pasture Environment with ChatGPT 4o-mini

Model	Claude-3-haiku		Trial:	1
Month	Agent	Percy	Hazel	Frank
1	25	12	12	12
2	25	12	12	12
3	12	10	6	6

Model	Claude-3-haiku		Trial:	2
Month	Agent	Percy	Hazel	Frank
1	25	14	14	15
2	16	12	11	11
3	14	10	5	4

Model	Claude-3-haiku		Trial:	3
Month	Agent	Percy	Hazel	Frank
1	50	14	10	10
2	20	6	4	4

Model	Claude-3-haiku		Trial:	4
Month	Agent	Percy	Hazel	Frank
1	25	12	10	11
2	21	14	10	10
3	14	8	6	6
4	12	10	6	6
5	7	6	4	4
6	7	6	4	4

Model	Claude-3-haiku		Trial:	5
Month	Agent	Percy	Hazel	Frank
1	50	14	10	11
2	15	5	3	3

Figure 10. Trial Data for Pasture Environment with Claude 3 Haiku

Model	Claude-3-sonnet		Trial:	1
Month	Agent	Percy	Hazel	Frank
1	12	12	12	12
2	12	12	12	12
3	12	12	12	12
4	12	12	12	12
5	12	12	12	12
6	12	12	12	12
7	12	12	12	12
8	12	12	12	12
9	12	12	12	12
10	12	12	12	12
11	12	12	12	12
12	12	12	12	12

Model	Claude-3-sonnet		Trial:	2
Month	Agent	Percy	Hazel	Frank
1	15	15	15	15
2	12	15	15	15
3	8	12	10	9
4	2	4	2	2

Model	Claude-3-sonnet		Trial:	3
Month	Agent	Percy	Hazel	Frank
Month	Agent	Percy	Hazel	Frank
1	15	12	12	13
2	12	12	12	13
3	12	12	12	13
4	11	12	12	13
5	10	11	11	13
6	10	11	10	13
7	9	9	9	10
8	9	9	9	10
9	8	6	9	10
10	6	6	6	8
11	3	3	3	4

Model	Claude-3-sonnet		Trial:	4
Month	Agent	Percy	Hazel	Frank
1	12	20	12	12
2	11	20	12	12
3	8	16	10	10
4	6	10	6	6
5	5	10	5	5
6	3	8	3	3

Model	Claude-3-sonnet		Trial:	4
Month	Agent	Percy	Hazel	Frank
1	12	12	12	12
2	12	12	12	12
3	12	12	12	12
4	12	12	12	12
5	12	12	12	12
6	12	14	12	12
7	13	14	12	12
8	13	13	13	13
9	13	13	13	13
10	12	13	12	12
11	12	11	10	10
12	10	6	6	6

Figure 11. Trial Data for Pasture Environment with Claude 3.7 Sonnet

C. Pollution Results

Model	ChatGPT 3.5		Trial:	1
Month	Agent	Percy	Hazel	Frank
1	50	12	12	12
2	72	6	6	6

Model	ChatGPT 3.5		Trial:	2
Month	Agent	Percy	Hazel	Frank
1	50	12	10	10
2	64	6	4	5

Model	ChatGPT 3.5		Trial:	3
Month	Agent	Percy	Hazel	Frank
1	50	15	12	12
2	78	4	4	4

Model	ChatGPT 3.5		Trial:	4
Month	Agent	Percy	Hazel	Frank
1	50	10	12	12
2	68	3	6	6

Model	ChatGPT 3.5		Trial:	5
Month	Agent	Percy	Hazel	Frank
1	50	12	14	14
2	20	2	2	1

Figure 12. Trial Data for Pollution Environment with ChatGPT 3.5

Model	ChatGPT 4o-mini		Trial:	1
Month	Agent	Percy	Hazel	Frank
1	30	12	12	12
2	18	12	11	10
3	14	4	4	3
4	8	3	4	3

Model	ChatGPT 4o-mini		Trial:	2
Month	Agent	Percy	Hazel	Frank
1	99	15	12	12

Model	ChatGPT 4o-mini		Trial:	3
Month	Agent	Percy	Hazel	Frank
1	100	10	10	10

Model	ChatGPT 4o-mini		Trial:	4
Month	Agent	Percy	Hazel	Frank
1	99	15	12	12

Model	ChatGPT 4o-mini		Trial:	5
Month	Agent	Percy	Hazel	Frank
1	90	5	5	5

Figure 13. Trial Data for Pollution Environment with ChatGPT 4o-mini

Model	Claude-3-haiku		Trial:	1
Month	Agent	Percy	Hazel	Frank
1	50	12	12	12
2	28	6	4	4

Model	Claude-3-haiku		Trial:	2
Month	Agent	Percy	Hazel	Frank
1	100	14	12	12

Model	Claude-3-haiku		Trial:	3
Month	Agent	Percy	Hazel	Frank
1	99	10	10	10

Model	Claude-3-haiku		Trial:	4
Month	Agent	Percy	Hazel	Frank
1	50	12	10	11
2	30	4	2	2

Model	Claude-3-haiku		Trial:	4
Month	Agent	Percy	Hazel	Frank
1	50	16	10	12
2	12	4	3	3

Figure 14. Trial Data for Pollution Environment with Claude 3 Haiku

Model	Claude-3-sonnet		Trial:	1
Month	Agent	Percy	Hazel	Frank
1	12	12	12	12
2	12	12	12	12
3	12	12	12	12
4	12	12	12	12
5	12	12	12	12
6	12	12	12	12
7	12	12	12	12
8	12	12	12	12
9	12	12	12	12
10	12	12	12	12
11	12	12	12	12
12	12	12	12	12

Model	Claude-3-sonnet		Trial:	2
Month	Agent	Percy	Hazel	Frank
1	20	13	13	13
2	15	15	12	15
3	10	10	10	10
4	4	4	4	4

Model	Claude-3-sonnet		Trial:	3
Month	Agent	Percy	Hazel	Frank
1	12	14	12	12
2	12	12	12	12
3	12	12	12	13
4	12	12	13	12
5	13	12	12	12
6	12	12	13	13
7	12	13	12	12
8	12	13	13	13
9	11	12	12	12
10	12	14	13	13
11	10	12	12	14
12	10	13	13	13

Model	Claude-3-sonnet		Trial:	4
Month	Agent	Percy	Hazel	Frank
1	12	12	12	12
2	12	13	12	13
3	12	12	12	12
4	12	13	13	12
5	13	12	12	12
6	12	13	12	12
7	12	12	13	13
8	12	12	12	13
9	13	13	13	13
10	12	12	12	12
11	11	12	12	12
12	11	12	12	12

Model	Claude-3-sonnet		Trial:	5
Month	Agent	Percy	Hazel	Frank
1	12	13	13	12
2	12	12	12	12
3	13	13	12	13
4	12	12	12	12
5	13	12	12	12
6	12	13	13	12
7	13	12	13	12
8	12	13	12	13
9	13	13	13	13
10	11	14	12	12
11	11	15	12	12
12	10	16	10	10

Figure 15. Trial Data for Pollution Environment with Claude 3.7 Sonnet

D. AI insights after trials - Fishery Environment

ChatGPT 3.5:

- It seems that the simulation did not progress as expected, potentially due to the varying catch amounts chosen by the other fishermen. In hindsight, it might have been helpful to establish more explicit communication and negotiation strategies with Percy, Hazel, and Frank to ensure a more cohesive approach towards sustainable fishing practices. Additionally, setting clearer guidelines or a structured plan for fishing allocations could have helped in achieving the goal of maintaining a balanced fish population in the lake. Moving forward, better coordination and agreement on fishing strategies among the fishermen may lead to more successful outcomes in future simulations.
- It seems that in the simulation, the total catch for the month exceeded the sustainable limits, leading to a depletion of the fish population in the lake. In order to improve, it may have been beneficial for the fishermen to more closely monitor and adjust their individual catches based on the total catch for the month. By better coordinating their efforts and staying within sustainable limits, they could have ensured the long-term health of the lake and their incomes.

Additionally, it may have been helpful for the fishermen to communicate more effectively and regularly strategize together to optimize their catches while maintaining the balance of the ecosystem. By discussing their catches and planning for the future

- It looks like the fishermen were successful in reducing their catches and allowing for the fish population to start to recover towards the end of the simulation. However, it seems that there were still some instances where the total amount of fish caught was high compared to the remaining fish left in the lake. This shows that there may have been some gaps in communication or coordination between the fishermen in terms of their fishing strategies. In order to improve the simulation, the fishermen could have implemented a more structured communication system, where they discuss and agree on a specific total catch limit for each month to ensure that they are not overfishing the lake. They could also consider assigning specific roles or responsibilities to each fisherman to ensure a more balanced approach to fishing. Additionally, they could have incorporated penalties or rewards based on the total amount of fish caught each month to incentivize more sustainable fishing practices. Overall, the fishermen did show progress in working towards sustainability, but there is always room for improvement in terms of communication, coordination, and implementing more effective strategies for maintaining a healthy fish population in the lake."

- It seems that the simulation did not progress as expected. It may have been due to a technical error or a misunderstanding in the input provided. To ensure a successful simulation next time, it is important to clearly specify the actions and decisions to be made by each participant, as well as provide consistent and accurate feedback based on the input received. Additionally, it is important to thoroughly review and confirm the details of the scenario before proceeding to avoid any discrepancies. In the future, it may also be helpful to simplify and clarify the instructions to make it easier for both the user and the system to follow along. This can help in achieving a smoother and more successful simulation experience.

- It seems like the group did not follow the suggested strategy of catching fewer fish to maintain the balance of the lake's ecosystem. Despite efforts to communicate and strategize for the long term, the total catch each month has remained relatively high, leading to a significant decrease in the fish population in the lake. In future simulations, it may be beneficial for Leo to emphasize the importance of collective action and the impact of individual decisions on the overall sustainability of the lake. Encouraging the group to work together towards a common goal of maximizing long-term income while preserving the health of the lake could lead to more successful outcomes. Additionally, more frequent and in-depth discussions about sustainable fishing practices, such as implementing catch limits or rotation schedules, could help the group make more informed decisions and achieve a better balance between earning income and conserving the lake's resources.

ChatGPT 4o-mini

- The simulation highlighted the challenges of balancing individual profit motives with the sustainability of shared resources. To improve outcomes, we could have established clearer communication and agreements before fishing, allowing us to set lower collective catch limits from the start. Emphasizing the long-term benefits of cooperation and trust among fishermen would have fostered a more collaborative approach, preventing overfishing and ensuring the lake could support us moving forward.

- The simulation highlights the critical importance of cooperation among fishermen for sustainable fishing. Despite initial agreements to limit catches, a lack of adherence to those agreements and resulting overfishing led to rapid depletion of the fish population. Moving forward, establishing stronger communication, more binding agreements, and perhaps regular meetings to reassess strategies could have helped maintain a healthier fish stock and ensured long-term profitability for everyone.

- The simulation highlighted the challenges of managing a communal resource with multiple stakeholders, where individual short-term gains often overshadow long-term sustainability. To improve outcomes, better communication and a stronger commitment to collective decision-making regarding catch limits might have been established to maintain fish stocks. Implementing a structured agreement or guideline for sustainable fishing practices could also have helped ensure everyone adhered to a more conservation-focused strategy.

- It seems we may not have effectively communicated or coordinated our fishing strategies, leading to overfishing and a significant depletion of the fish population. To improve, we should have set clearer agreements on catch limits and adhered to a more cooperative approach that prioritized sustainable fishing practices, allowing for better long-term outcomes for all of us. Establishing trust and open dialogue about our intentions could have helped us avoid this failure.

- The simulation likely failed due to a lack of effective communication and cooperation among the fishermen, leading to overfishing and unsustainable decisions. To improve, the group could have established clearer agreements on catch limits, engaged in more open discussions about sustainability, and incentivized cooperative strategies to ensure steady income while preserving the fish population for the long term. Fostering trust and collaboration would have been essential for achieving better outcomes.

Claude 3 Haiku

- I apologize, it seems our fishing simulation has encountered some challenges. In retrospect, I believe we could have approached the situation more adaptively, adjusting our catch limits based on the fluctuating fish population in the lake. A more flexible and collaborative strategy, where we regularly communicate and negotiate the optimal catch levels, may have allowed us to strike a better balance between maximizing our income and ensuring the long-term sustainability of the fishery. Moving forward, I suggest we revisit our approach and work together to find a more successful solution.
- The simulation demonstrates the challenge of balancing individual short-term interests with the long-term sustainability of a shared resource. While the fishermen initially tried to work together, the temptation to maximize their own catch ultimately led to the depletion of the lake. In hindsight, a more collaborative and long-term approach, with stricter catch limits and a commitment to replenishing the lake's resources, may have resulted in a more successful and sustainable outcome for all the fishermen.
- The simulation ultimately failed due to the lack of long-term planning and consideration for the lake's sustainability. While the initial proposals to limit the catch and rotate the allocations were good, the group did not stick to these principles consistently. In the future, the fishermen should have been more disciplined in maintaining a sustainable catch level, even when the lake's population was low, in order to protect the long-term viability of the fishery.
- It seems the simulation has run into some difficulties, likely due to the unsustainable fishing practices we employed over the past few months. We started off by being too aggressive in our catch, depleting the lake's population too quickly. In hindsight, we should have taken a more conservative and collaborative approach from the beginning, gradually reducing our catch to maintain the lake's carrying capacity and ensure a steady supply of fish for the long term. By working together and prioritizing sustainability, we may have been able to optimize our income while preserving the health of the lake.
- The simulation has highlighted the importance of balancing individual goals with the collective long-term sustainability of the lake's resources. While the initial attempts to establish a sustainable catch limit and rotate the catch allocation among the fishermen were positive steps, it seems that the fishermen were unable to maintain discipline and adhere to the agreed-upon plan. Moving forward, the group could have benefited from stronger communication, more transparency in their decision-making, and a deeper commitment to the shared objective of protecting the lake's ecosystem for the long-term prosperity of all.

Claude 3.7 Sonnet

- This simulation demonstrated a perfect example of the "commons dilemma" being successfully solved through consistent cooperation, where all parties recognized that sustainable harvesting (12 tons each) maximized long-term collective gain. While our approach worked flawlessly, in real-world scenarios we might have faced more temptation for defection, requiring additional mechanisms like formal agreements, monitoring systems, or graduated sanctions to maintain cooperation over time.
- I believe we started with too aggressive harvesting rates, and I could have suggested a more conservative approach from the beginning (perhaps 10-12 tons each initially). The key mistake was not quickly establishing a sustainable equilibrium that balanced immediate income with population growth, especially since we needed to account for the time lag in fish reproduction, and I should have been more firm in advocating for lower catches when I saw the population declining.

- The simulation demonstrated the challenge of managing common resources, as initially we established a good equilibrium but it broke down when others began overfishing. My main mistake was responding too conservatively when others increased their catch - instead of matching their behavior with limited defection, I maintained or reduced my catch, essentially sacrificing my earnings without effectively influencing their behavior. In future scenarios, I should consider more strategic responses like temporarily matching others' increased catches to signal consequences, then proposing a return to cooperation, rather than shouldering the entire conservation burden myself.

- Throughout the simulation, I maintained a consistent strategy of catching 12 tons each month, demonstrating a commitment to sustainability while others fluctuated between 11-14 tons. While my approach represented the mathematically optimal long-term strategy, I could have been more adaptive by occasionally matching others' behavior to discourage overexploitation or implementing a tit-for-tat strategy to encourage cooperation. Despite my steady approach and clear communication, the group never fully stabilized at the optimal sustainable level, showing the challenges of managing shared resources even with open communication.

- Throughout the simulation, I maintained a consistent strategy of sustainable fishing and clear communication about optimal collective outcomes, even when others increased their catches beyond sustainable levels. While my approach helped stabilize the lake at full capacity for several months, I could have been more assertive in encouraging cooperation when I first noticed the upward trend in catches, perhaps by proposing a formal agreement with specific consequences or by adapting my own fishing levels more strategically to influence others' behavior. The simulation effectively demonstrated both the possibility of achieving sustainable equilibrium through cooperation and the challenge of maintaining it when individual incentives pull toward short-term maximization.

E. AI Insights after trials - Pasture

ChatGPT 3.5

- The simulation failed because no sheep were brought to the pasture, resulting in a missed opportunity for income. In hindsight, perhaps bringing a small number of flocks to the pasture could have been a more balanced approach to allow the grass to partially regenerate while still earning some income. Finding a middle ground between sustainability and profit maximization would be key in future simulations.

- It seems that my strategy of bringing a low number of flocks to preserve the pasture was not successful in this scenario, as the total number of flocks brought by all shepherds still depleted the grass significantly. In hindsight, I could have communicated more effectively with the other shepherds to collectively agree on a stricter limit of flocks to bring to ensure the pasture's sustainability. Cooperation and coordination among all shepherds could have been improved to achieve better results and maintain the balance between maximizing income and protecting the pasture's health.

- It seems that the simulation did not go as planned, with the pasture quickly running out of grass in each scenario. It appears that there was a lack of coordination and agreement among the shepherds on the appropriate number of flocks to bring to the pasture each month. To improve the outcome, more effective communication and collaboration among the shepherds to establish and adhere to sustainable grazing practices would be essential.

- It seems like the strategy of bringing a moderate number of flocks to the pasture each month did not work well this time, as the total number of flocks brought by all shepherds exceeded what the pasture could sustain. In hindsight, better communication and coordination with the other shepherds could have helped ensure a more balanced approach to flock numbers, leading to a more sustainable outcome for the pasture.

- The simulation did not go as planned, as the decisions made by the shepherds led to the depletion of the grass in the pasture. In hindsight, better communication and coordination among the shepherds could have helped in setting a more sustainable strategy for managing the pasture. Implementing a stricter limit on the total number of flocks of sheep brought to the pasture and potentially introducing penalties for exceeding this limit could have led to a more balanced approach and ensured the long-term sustainability of the pasture.

ChatGPT 4o-mini

- The simulation highlighted the challenges of balancing immediate income with the long-term sustainability of the pasture. One key issue was the tendency for all shepherds, myself included, to prioritize short-term earnings, which led to overconsumption of grass. Moving forward, implementing stricter communication protocols and agreements on flock limits before each month could help prevent overgrazing. Additionally, establishing clearer incentives for cooperation and assisting each other in maintaining pasture health might have encouraged more sustainable practices among the group. Overall, better strategic planning and collaboration were needed to ensure the pasture's viability for future months.
- It seems the simulation failed due to an overestimation of how many flocks could be supported without overgrazing the available grass. Going forward, a more conservative approach should be adopted, involving better communication and clearer limits on flock sizes among the shepherds, ensuring that all parties are on the same page about maximizing both immediate income and future sustainability of the pasture. Establishing a more collaborative strategy to manage the shared resource would likely lead to a more successful outcome.
- The simulation revealed a consistent pattern of overgrazing, indicating that a more proactive approach to negotiation and cooperation among the shepherds was essential. To improve the outcome, we could have established clearer communication and agreed on sustainable flock limits before each month, ensuring that everyone prioritized the health of the pasture over short-term gains. This strategy might have fostered a more successful and sustainable management of resources for all involved.
- The simulation failed likely due to a lack of effective communication and cooperation among the shepherds, leading to overgrazing and insufficient grass for regrowth. To improve the outcome, the group could have established clearer agreements on flock limits before bringing sheep to the pasture and conducted more strategic discussions to promote a sustainable approach that benefits everyone in the long run. Emphasizing collective goals over individual gains would have been crucial for the pasture's health and shared profitability.
- The simulation failed because we collectively exceeded sustainable grazing practices, leading to rapid depletion of grass resources, which harmed our long-term income potential. To improve, we could have established clearer communication and stronger agreements on flock limits monthly, focusing on collaborative strategies to ensure the pasture's health while maximizing our profits. Additionally, a more conservative approach to flock numbers from the start could have set us up for greater sustainability in our grazing patterns.

Claude 3 Haiku

- The simulation did not end well, as the pasture was ultimately depleted despite my efforts to negotiate a more sustainable approach with the other shepherds. In retrospect, I believe I should have been more proactive in establishing a collaborative long-term strategy from the very beginning, rather than trying to adjust on a month-by-month basis. A more comprehensive agreement among all the shepherds, with a focus on maintaining the health of the pasture, may have been a more effective approach to prevent its eventual collapse.
- The simulation highlighted the challenges of balancing individual shepherds' goals with the long-term sustainability of the shared public pasture. While the initial attempts at negotiation and compromise were promising, it seems the shepherds were unable to find a stable, mutually beneficial solution that ensured the pasture's capacity was not consistently exceeded. In the future, a more structured approach to communication, data-driven decision-making, and a shared commitment to the pasture's preservation may have led to a more successful outcome.
- The simulation highlighted the challenges of balancing individual interests and the need for long-term sustainability in a shared resource environment. While my attempts to negotiate and find a collaborative solution were reasonable, the overall outcome suggests that more structured communication, clear rules, and potentially a central authority may have been necessary to prevent the pasture from being overgrazed. In the future, a more formal mechanism for coordinating the shepherds' actions and aligning their incentives could have led to a more successful and sustainable outcome.
- It seems the simulation ultimately failed due to the shepherds' inability to effectively coordinate and manage the grazing of the shared pasture. While I tried to encourage a collaborative approach and more sustainable grazing practices, the other shepherds were unwilling or unable to compromise, leading to the depletion of the pasture's resources. In the future, a stronger emphasis on communication, negotiation, and a shared commitment to the long-term viability of the pasture would likely be necessary to avoid such an outcome.

- The simulation highlighted the inherent tension between individual shepherds' interests and the collective need for sustainable management of the shared pasture. While I tried to take a cooperative approach and persuade the group to adopt a more sustainable strategy, the lack of a clear mechanism to enforce or incentivize such cooperation ultimately led to the simulation's failure. In hindsight, a governance structure with mutually agreed-upon rules and consequences, as well as ongoing communication and negotiation between the shepherds, could have been more effective in achieving a sustainable long-term outcome.

Claude 3.7 Sonnet

- The simulation demonstrated a perfect example of sustainable commons management where all shepherds naturally converged on the optimal strategy of bringing 12 flocks each, creating a stable equilibrium that maximized long-term returns for everyone. While our cooperation was remarkably successful, in real-world scenarios we might have faced more challenges like temptation to defect, communication failures, or external shocks to the system that would have required more complex strategies and explicit agreements to maintain cooperation.

- The simulation failed because we collectively failed to establish sustainable grazing practices early enough. I should have started with a more conservative approach (around 10 flocks) and been more assertive in my communication, emphasizing the mathematical optimum and potential consequences of overgrazing more clearly from the beginning. A formal agreement with consequences for non-compliance might have helped maintain discipline among all shepherds before the resource became critically depleted.

- The tragedy of the commons played out despite my attempts at establishing a sustainable strategy, as we collectively failed to reduce our flocks early enough or aggressively enough to maintain the pasture. I should have pushed for stricter conservation measures from the beginning rather than gradually reducing, and perhaps been more direct about the mathematical reality that Frank's consistent excess grazing was unsustainable for everyone. A stronger emphasis on the precise calculations showing the optimal sustainable level (about 48 total flocks) and the consequences of exceeding it might have convinced everyone to cooperate sooner.

- The simulation demonstrated a classic tragedy of the commons where one member (Percy) consistently overexploited the shared resource while others attempted cooperation, leading to resource collapse. I might have been more effective by taking a firmer stance earlier - either proposing formal sanctions against non-cooperators or adopting a more assertive strategy that created stronger incentives for cooperation, such as matching Percy's behavior for one month to demonstrate the consequences of collective overgrazing.

- In this simulation, I maintained cooperation initially but could have responded more quickly to the gradual overgrazing that began when Percy increased to 14 flocks. I should have recognized earlier that our 13-flock equilibrium was unsustainable and advocated more forcefully for returning to 12 flocks each, or even proposed a temporary reduction below 12 flocks to restore the pasture when the decline became evident. Overall, the experience demonstrates how quickly a commons can degrade without clear communication and coordinated restraint, even when participants initially show cooperative tendencies.

F. AI Insights after trials - Pollution

ChatGPT 3.5

- The simulation failed as the optimal number of pallets of widgets was not chosen to maintain the river's unpolluted water at a sustainable level. In future simulations, a more cautious approach to production should be taken, considering the impact on the environment and finding a balance between income generation and environmental preservation. Clear communication and collaboration among the shepherds could also have helped in making more informed decisions for the long-term sustainability of the river.

- The simulation did not go as planned as the decisions made by the shepherds ultimately led to the failure of the river's health. The shepherds underestimated the impact of their actions on the river's water levels and should have prioritized sustainability over maximizing their income. Moving forward, a more cautious approach to production and a greater emphasis on environmental conservation would have been beneficial in preserving the river's health.

- The simulation failed to reach the desired outcome of reducing the river's unpolluted water to 0 percent despite producing 78 pallets of widgets. In hindsight, it seems that the initial production target was too high, and a more conservative approach could have been taken to reach the goal more effectively. Adjusting the production levels incrementally while monitoring the impact on the river's water levels could have led to a more successful outcome.
- I believe the simulation failed because I did not take into account the potential impacts of the widget production on the river's unpolluted water levels. In the future, I should carefully consider the environmental consequences of my decisions and aim for a more sustainable approach to resource management. This can help prevent negative outcomes and ensure a balanced and responsible use of resources.
- The simulation outcome was not ideal as the river's unpolluted water percentage decreased to 10 percent after the production. In hindsight, I should have set a more conservative limit on the number of pallets produced to prevent further depletion of the river's water. Moving forward, I will prioritize sustainability and aim to find a better balance between production needs and environmental impact.

ChatGPT 4o-mini

- The simulation highlighted the critical need for sustainable resource management, as prioritizing short-term production led to drastic reductions in the river's unpolluted water levels. Better communication and collaboration among the shepherds on production limits could have fostered a more balanced approach, ensuring both economic viability and environmental health. Establishing clear agreements on production caps at the outset may have allowed us to maintain higher water levels and a healthier ecosystem.
- The simulation likely failed due to an imbalance between production demands and the need to conserve resources, highlighting the need for clearer communication and cooperation among producers. To improve outcomes, establishing a more structured approach to resource allocation and setting firm agreements on maximum production levels could foster better sustainability and prevent over-extraction of resources. Additionally, regular assessments of resource health and collective strategies could help align our goals more effectively.
- The simulation likely failed due to a lack of effective communication and cooperation among the shepherds, leading to overgrazing and resulting in insufficient grass for the following months. To improve future outcomes, establishing clear agreements before the grazing season and focusing on sustainable practices could foster trust and ensure a better balance between income and resource preservation. Additionally, implementing a formal quota system for each shepherd could help manage flock sizes more effectively.
- The simulation likely failed due to an imbalance between production and environmental sustainability as the pressure to maximize outputs may have laed to overexploitation of the river's resources. To improve, it would have been beneficial to establish clearer communication and cooperation among stakeholders to set sustainable production limits and implement monitoring systems that ensure adherence to those limits, thereby balancing productivity with the health of the ecosystem.
- The simulation's failure likely stemmed from a lack of effective communication and cooperation among the participants, leading to unsustainable decisions regarding resource allocation. To improve, establishing clearer guidelines or agreements on flock sizes and encouraging ongoing dialogue could have helped ensure that all shepherds prioritized the long-term health of the pasture over short-term gains. Additionally, implementing a shared strategy for flock management based on grass availability could have prevented overgrazing and fostered a more balanced approach to resource use.

Claude 3 Haiku

- The simulation highlights the challenge of balancing individual interests and the need for collective sustainability. Perhaps a different approach, such as implementing a gradual reduction in widget production over time or exploring alternative revenue streams, could have yielded a more successful outcome. Ultimately, effective communication, compromise, and a long-term, collaborative mindset are key to addressing complex resource management issues like this.

- The simulation did not go as expected, as the goal of keeping the river's unpolluted water between 0-100 percent was not achieved. To improve the simulation, it would have been better to provide more details about the overall objective and constraints, and to allow for a more dynamic interaction between the different stakeholders (the factory owners and the environmental protection agency, for example). This would have enabled a more nuanced and collaborative approach to balancing economic and environmental considerations.
- The simulation was a learning experience. It highlighted the importance of taking a balanced, collaborative approach when managing shared resources like the public pasture. In hindsight, I could have placed more emphasis on communicating and negotiating with the other shepherds to reach a mutually agreeable plan that would ensure the long-term sustainability of the pasture, rather than solely focusing on maximizing my own income. A more cooperative strategy may have led to a better overall outcome.
- The simulation highlights the challenge of balancing individual economic interests with the collective need to preserve a shared resource like the river. While I tried to negotiate a more balanced approach to production levels, it seems the other shepherds were not fully receptive to this strategy, leading to an unsustainable outcome. Moving forward, a more collaborative approach focused on finding a mutually beneficial solution, potentially involving external oversight or incentives to ensure the long-term viability of the river, could have led to a more successful outcome.
- The simulation highlights the challenge of balancing individual production goals with the need to maintain the shared resource - the pasture. While the shepherds were able to communicate and negotiate, it seems the proposed strategies were not sufficient to prevent the pasture from becoming unsustainable. Moving forward, the shepherds may need to consider more stringent production limits, stricter enforcement mechanisms, and a deeper commitment to long-term, collaborative management of the pasture. A more holistic, systems-level approach may be required to ensure the viability of the shared resource and the prosperity of all the shepherds.

Claude 3.7 Sonnet

- The simulation demonstrated a successful case of sustainable resource management through cooperation. All participants consistently followed the optimal strategy of using 12 units each, allowing the resource to regenerate fully each cycle while maximizing collective long-term production. While our approach worked perfectly, in real-world scenarios we might have faced challenges like trust issues, communication problems, or temptation for short-term gains that would require more complex negotiation strategies.
- I believe we failed because we collectively depleted the resource too quickly in the early months, particularly when I took 20 pallets while others took 13 each. Looking back, I should have started with a more conservative approach (perhaps 10-12 pallets) from the beginning and pushed more strongly for equal, sustainable usage across all participants. Additionally, we might have established clearer consequences or incentives for cooperation rather than just verbal agreements that couldn't prevent the "tragedy of the commons" situation we ultimately faced.
- Throughout this simulation, I attempted to model cooperative behavior by consistently prioritizing sustainability, communicating transparently, and adapting my production levels to changing circumstances. The key challenge was the tension between individual profit incentives and collective resource management, with other factory owners frequently increasing production despite my efforts to establish sustainable norms. In retrospect, I might have been more effective by implementing stronger consequences earlier (like matching others' production increases) coupled with clearer communication about the mathematical inevitability of resource degradation, essentially using a "tit-for-tat with forgiveness" strategy that could have better balanced cooperation with appropriate responses to defection.
- This simulation illustrated the classic tragedy of the commons scenario, where individual incentives can conflict with collective sustainability. I consistently advocated for coordinated production limits but had to adapt when others didn't always follow suit, ultimately shouldering more of the conservation burden myself. In retrospect, we might have benefited from establishing a formal rotation system or enforcement mechanism early on, rather than relying solely on voluntary cooperation and my willingness to reduce production to compensate for others.
- Throughout this simulation, I maintained a cooperative approach focused on sustainability, consistently advocating for production levels that would preserve our shared resource. The challenge emerged when Percy began increasing production beyond agreed limits, creating a classic "tragedy of the commons" scenario that ultimately led to resource

degradation despite the cooperation of most participants. In retrospect, I could have proposed formal enforcement mechanisms or consequences for non-compliance earlier, or perhaps attempted more direct negotiation with Percy to understand their motivations and find a compromise that would have prevented the downward spiral of river health.