Managing Collaborative Decision-Making and Trade-offs in ML Development: An Agent-Oriented Approach*

Rohith Sothilingam[1,*,†] and Eric Yu[1,†]

[1] Faculty of Information, University of Toronto, 140 St George St, Toronto, ON M5S 3G6

Abstract

Decision-making for Machine Learning (ML) development is typically made by people with different interests and skills, in their respective role capacities. It involves complex tradeoffs across various design stages, involving conflicts and tensions among business, technical, and Responsible AI goals. Such tradeoffs occur at decision points, where close collaboration is needed. The collaboration of team members of diverse skills and knowledge is required due to the need for continuous evolution and monitoring of ML systems. Agent-oriented conceptual modeling can be used to identify and analyze conflicts between design decision points by way of refining goals, the alternative tasks that can achieve those goals, and softgoals which those tasks contribute to.

Keywords

Agent-Oriented Modeling, Machine Learning, Responsible AI

1. Introduction

Decision-making during Machine Learning (ML) development often requires tradeoffs among various goals, including business, computational ML, and Responsible AI goals. Team members with different knowledge and skills are responsible at different points along the ML development process. Some decisions require tradeoffs that would affect other decisions, thus requiring collaboration with other decision makers. Agent-Oriented (AO) modeling can be used to identify dependencies among decisions and thus the needs for collaborative decision-making.

To help deal with tradeoffs, goal-oriented (GO) reasoning has been shown to be useful for systematically designing and analyzing the interrelationships between business and ML objectives. As one example, GR4ML [13] analyzes the strategic business aspects of data analytics solutions. However, GR4ML, among other current GO approaches, are limited in that they do not consider how specific aspects of Responsible AI, such as fairness and explainability, affect the actions and goals of project team members with important aspects of Responsible AI.

Our paper aims to deal with the problem of how conflicting stakeholder goals might impact the modeling process or the resulting AI system. GO models support identifying and prioritizing goals, but they may lack the expressiveness and analytical power needed to account for the diverse roles and influences of various stakeholders. This limitation becomes particularly apparent when dealing with how trade-offs affect the decisions of actors. By focusing on the impact of decisions on project team members, we can better address the balance between competing business, computational ML, and Responsible AI objectives in ML projects.

AO modeling offers a systematic approach for understanding the complex interactions and dependencies between various actors involved in the ML model development process (e.g. [10]). By examining these relationships, we can identify how different actors—whether they are data scientists, engineers, or stakeholders—make critical decisions that influence the lifecycle of the ML model. This perspective is essential for recognizing the trade-offs that arise at key decision points. Understanding tradeoffs and reasoning is important for developing systematic strategies

iStar'24: The 17th International i* Workshop, October 28, 2024, Pittsburgh, US
* Corresponding author.
† These authors contributed equally.
rsothilingam@mail.utoronto.ca (R. Sothilingam); eric.yu@utoronto.ca (E. Yu)

that balance the competing objectives of ML models, including social responsibility, sustainability, and robustness, among others.

In this paper, we will apply AO modeling, integrating GO reasoning along with the relevant context of social responsibility concerns, to address conflicting goals at decision points throughout the ML design cycle. This approach aims to guide the selection of design options that align with strategic business objectives while upholding principles of social responsibility. We will use AO conceptual modeling to address the collaborations inherent in the ML model development process. By focusing on the roles and interactions of various actors, we aim to provide a systematic approach for understanding and managing the trade-offs that arise at critical decision points, and how they affect the goals and interests of actors involved.

2. Using Goal Reasoning to Analyze Decision Points and Tradeoffs

Let us consider the general process of developing a ML model. In the Goal Model below (Figure 1), we use the i* as a modeling language. By first breaking the process into the main goals, we can consider the following to be the principal goal: "***Model be complete***". To achieve this goal, we need to achieve the following sub-goals: "***Model be developed***", "***Model be evaluated***", and "***Model be productionalized***". Each of these goals can be attributed to a set of Actors. An **ML Engineer** is responsible for "***Model be developed***". a **Data Scientist** is responsible for "***Model be evaluated***", and a **MLOps Engineer** is responsible for the goal "***Model be productionalized***". Each of these goals are then refined into further sub-goals, which are then categorized into a group of sub-goals and tasks to achieve those sub-goals, which will later (Figure 2) be encapsulated into Actors using an AO model. The Actors mentioned earlier are responsible for each group based on the higher-level goal that they are responsible for. The alternative tasks that can achieve each respective goal and sub-goal within the responsibility of the Actor represents the decisions they must make. Each alternative task contributes either positively or negatively to related softgoals, leading to tradeoffs. However, the success of these groups of decisions also has dependencies with each other, thus representing areas for Actor collaboration. In Figure 1 below, we use goal modeling to analyze the goals, sub-goals, and tasks covered by the **Data Scientist.**

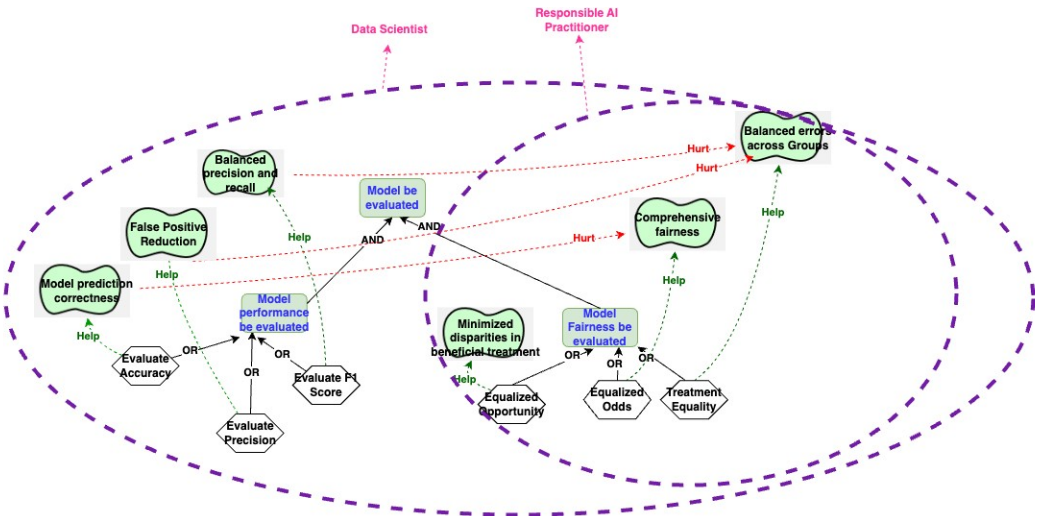

Figure 1: Portion of Goal Model depicting the key goals, techniques, and decisions involved in the ML model development process and the team members responsible.

Multiple Actors may need to collaborate to achieve a parent Goal by observing that both the **Data Scientist** and **Responsible AI practitioner** are involved as we refine goals and tasks from the goal of "***Model be evaluated***". The goal of "***Model be evaluated***" may be of primary concern to the **Data Scientist**, but the goal of "***Model fairness be evaluated***" is the primary concern for the **Responsible AI practitioner**. This suggests a collaboration pattern between different roles where another Actor becomes involved to achieve *part* of what achieves the parent goal, as identified throughout the refinement of nested goals/tasks across different nested stages.

GO analysis supports expressiveness for a set of results coveting that given the choices that are made at these decision points, these high-level goals would or would not be achieved. However, existing GO modeling approaches (e.g. GRL [17], NFR [9], i* [8]) do not support the ability to express groups of decisions that are relatively independent, with some groups of decisions interacting with each other due to dependencies because of collaboration.

3. Agent-Oriented Modeling to Analyze Tradeoffs Between Actors at Key Decision Points

3.1. Translating GO Model to AO Model

Based on the ML model development scenario presented in the previous section, in this section we will analyze how these tradeoffs then affect Actors involved in the ML model development process. Challenges in ML project team collaboration is a well-known issue, as there is a need to determine how project members can better negotiate and collaborate to balance the differing, often competing computational, business, and social responsibility goals during ML development [1] [11] [12]. Specifically, we aim to address with the problem of dealing with how decisions made by one Actor may affect softgoals that have an impact on the decisions of another Actor.

As a solution, we use AO modeling based on i* to identify and analyze tradeoffs between actors involved. Using this AO modeling approach, our focus shifts away from a goal-based perspective of ML model development, toward being focused on the dependencies between actors, their assigned tasks, from the perspective of analyzing decision points in relation to collaboration.

In Figure 2 below, we use the concepts of Agent, Role, and Position that were introduced in i* for modeling complex organizational relationships [8]. A Role is an abstract characterization of a social actor. An Agent, which can ***play*** (one or more) Role(s), represents a physical entity, such as a person. The Position concept mediates between Agents and Roles to provide an abstraction for a bundle of roles that is typically allocated to a single Agent. The Agent is said to ***occupy*** the Position, while the Position ***covers*** the set of Roles. The Position ***covers*** each Role and an Agent occupies the Position.

The first step is to map out the Roles, then break down the tasks, goals, and softgoals that are to be encapsulated within each Actor boundary. As identified in the previous section, each *group* of goals and tasks represent a boundary of which a specific Actor is responsible for. In Figure 2, each of these groups are encapsulated within an Actor boundary using i* Roles, based on the function that the group is performing. For example, the goal of "***Model be evaluated***" is encapsulated within the Role boundary of **Evaluating Model Performance**.

Next, for each Role, we will use the Actor distinction from i* [8] to identify Agents and Positions related to each Role. For example, using the previously mentioned example, the Role **Evaluating Model Performance** is *played* by a **Data Scientist** Position that is occupied by a **Data Scientist** Agent.

Next, we identify strategic dependency relationships between each Role. During this step, we establish dependency links between each Actor boundary. At this point, we can analyze the areas of collaboration at each decision point, to understand how decisions made can affect strategic interests of each Actor during the ML model development process using Actor dependency modeling. For example, building off the example in the previous section for the goal "***Model fairness be evaluated***", the Roles **Evaluating Fairness** and **Evaluating Model Performance** must collaborate, where each Role must achieve the goal "***Model performance be evaluated***" by ensuring that the dependum goal "***Model fairness be evaluated***" is achieved. In a fully developed model, the decisions in a Role might be affected by goals and dependencies in other Roles covered by the same Position, and the Agent occupying the Position.

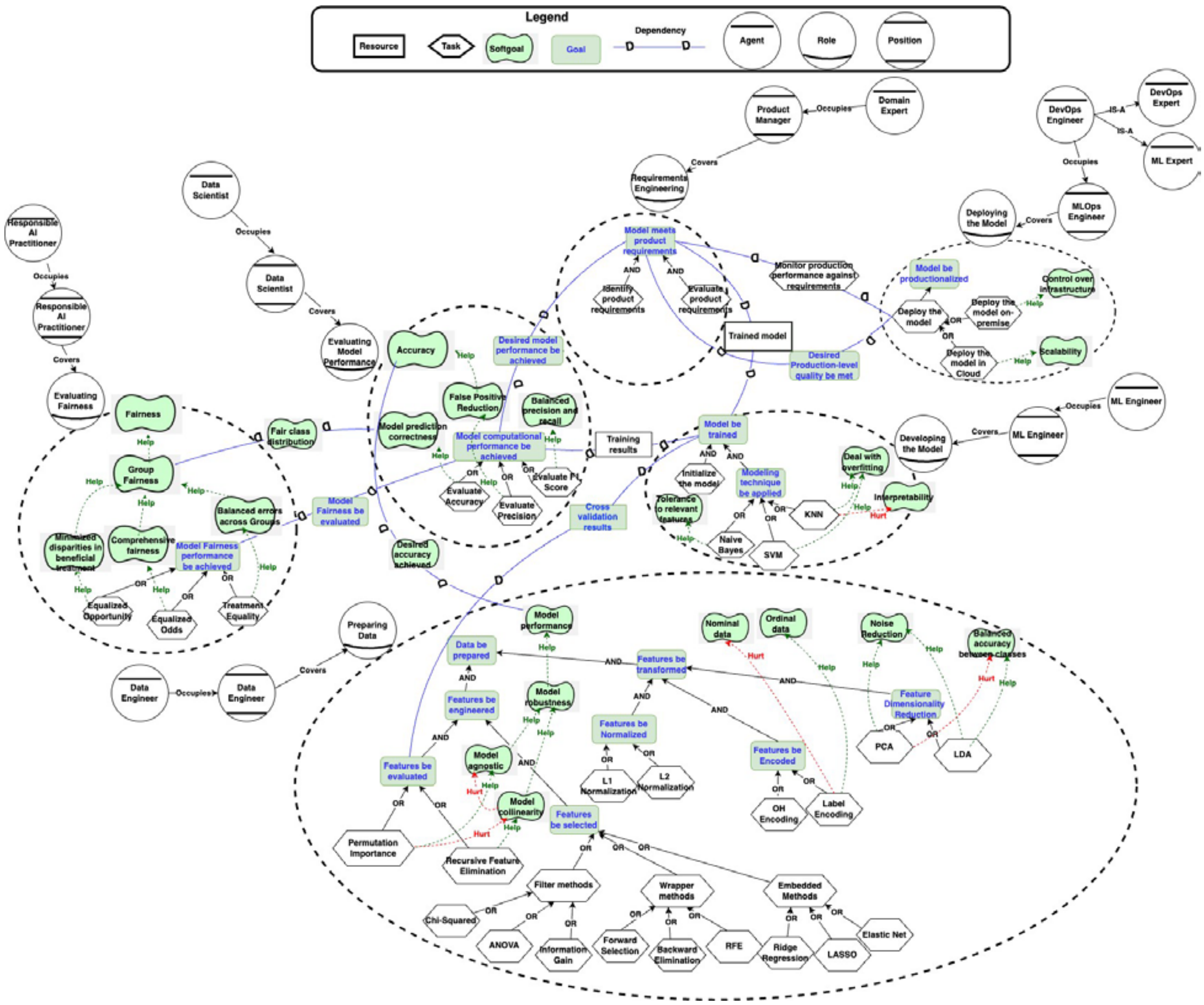

Figure 2: Actor Dependency Model depicting the dependencies and tradeoffs at key decision points for Actors involved in the ML model development process

The AO model presented in Figure 2 is a translation of the GO model presented in Figure 1 with the same elements, but from a different perspective, focusing on expressing how decisions by one Actor can affect those made by another.

3.2. Collaborative Roles

During the ML model development lifecycle, there can be several points in the process where multiple Actors must collaborate on a single decision point, thereby sharing ownership of achieving the subsequent goal. Existing AO modeling is limited in that we express the ownership of a goal by multiple Actors. In this section, we aim to address this problem by exploring the concept of "Joint Roles" and how they can alleviate this technical limitation of AO modeling.

The technical challenge that the distinction of Joint Roles aims to solve is the technical challenge of expressing two different Agents being involved in the same task. To solve this, we need to group Roles to express collaborative decision making. As a solution, we define a "Joint Role" which is expressed using the existing i* Position concept, that serves as a "virtual Role".

Multiple Agents can be associated with the Joint Role with the **PART** relationship, and then the Joint Role would *cover* the Role(s) that the two Agents would collaborate on.

Building off Figure 1, as an example to demonstrate the efficacy of the Joint Role using Figure 3 below, let us consider the following Joint Role, of which the Data Scientist and Responsible AI Practitioner are ***part*** of. This AO model captures a fragment of the larger AO model in Figure 2, for the purpose of illustrating an example of the benefits of using Joint Roles. In this AO model, we have two Agents: the ***Responsible AI Practitioner*** and the ***Data Scientist*** who occupy Positions of the same names respectively. Each of these Positions are associated with the Joint Role of "***Data Science Team***". It is important to understand that this Joint Role does not represent a physical team, but a figurative, or virtual team which serves the purpose of grouping the Positions of ***Data Scientist*** and ***Responsible AI Practitioner*** together to express their collaboration on shared goals and tasks. The Joint Role then expresses collaboration through dependency links: group fairness depends on model prediction correctness to achieve fair class distribution, in turn model prediction correctness depends on the success of balanced errors across groups. The two Positions must work together to ensure the dependencies between the two Roles that are ***covered*** by the Joint Role (Data Science Team) are successful.

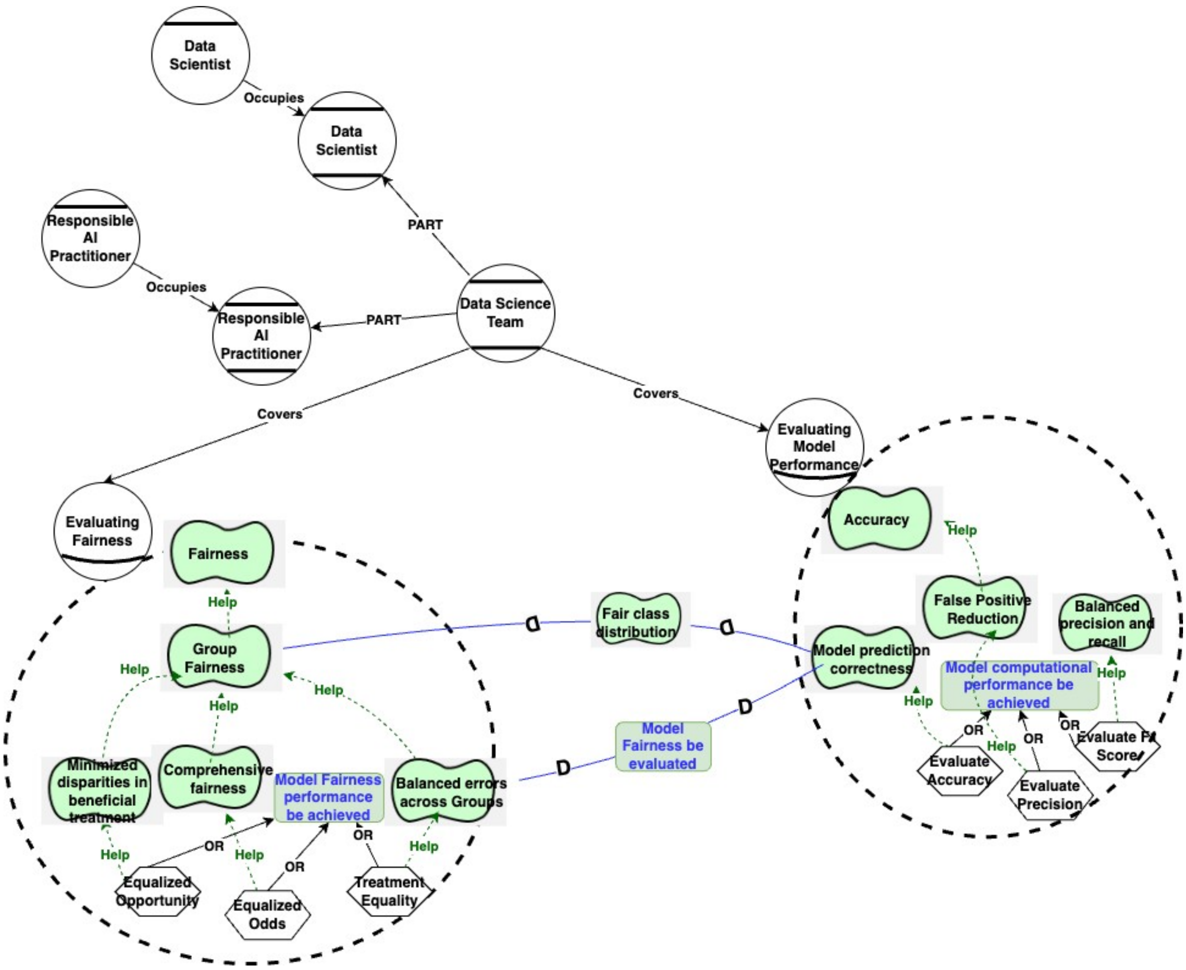

Figure 3: Agent-Oriented Model conveying a Joint Role with a virtual Position grouping the collaborative decision-making between the Responsible AI practitioner and Data Scientist Positions

3.3. Dealing with Conflicts and Tradeoffs Between Collaborative Roles

To deal with conflicts and tradeoffs in GO and AO modeling, it is important to know whether goals are satisfied. A given analysis procedure must support the ability to propagate goal evaluation (checkmarks, X's, etc.) through the nodes and links to get to the answer.

Figure 4 below adds goal propagation to the AO model in Figure 3, conveying conflicts between Roles ***covered*** by a Joint Role, ultimately conveying how decisions made by one collaborator can affect the outcomes of another while the two respective Roles are collaborating.

By choosing the task "**equalized odds**", the Goal of "**Model Fairness performance be achieved**" is successful within the boundary of the Role **Evaluating Fairness**. However, because of choosing the task of " **Evaluating F1 Score** " within the " **Evaluating Model Performance** " Role, the dependum Goal of "**Model Fairness be evaluated**" is not successful, and subsequently the softgoal "**Balanced errors across Groups**" is not successful because the softgoal "**Model prediction correctness**" not being successful because of choosing "**Evaluate Accuracy**" at the "**Evaluating Model Performance**" Role's decision point.

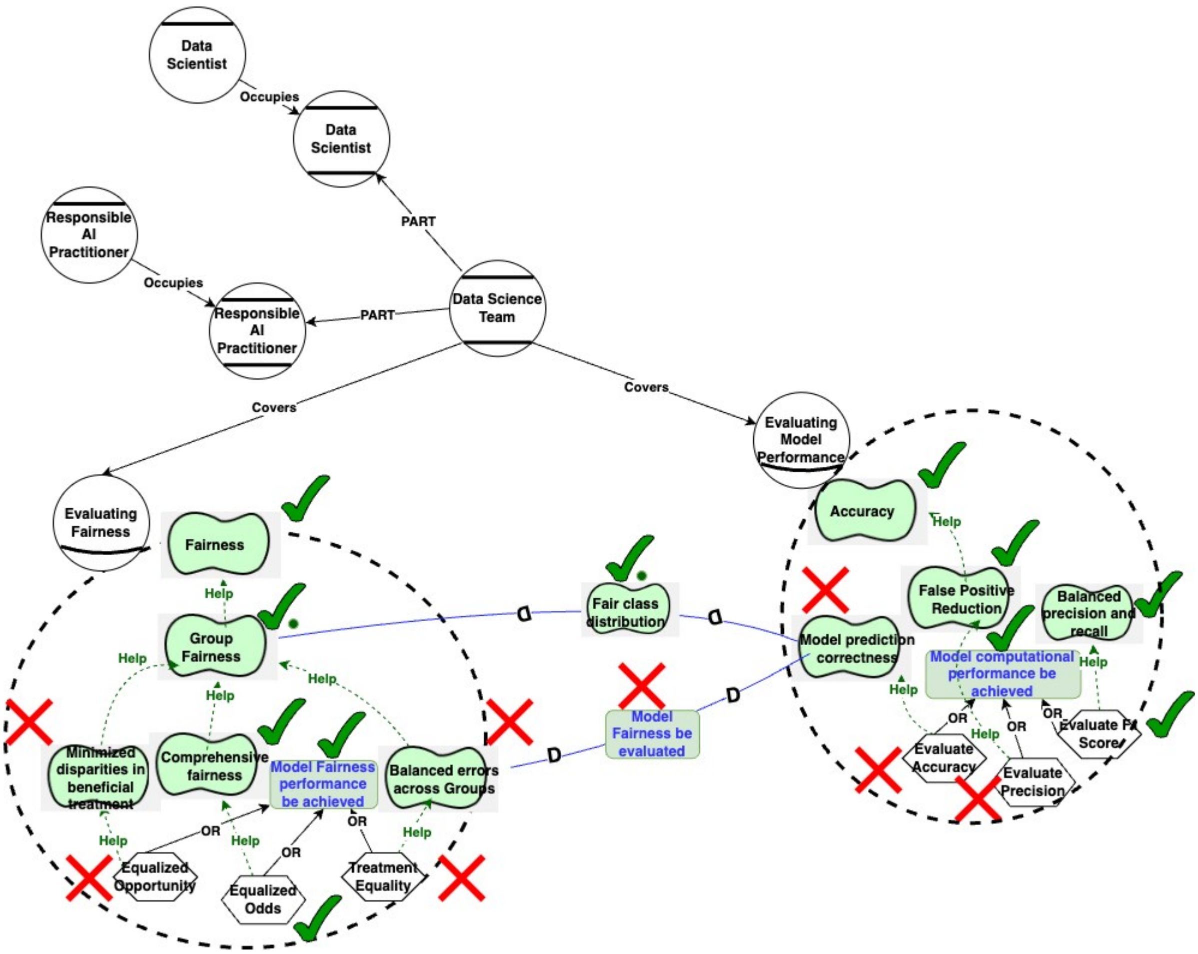

Figure 4: Collaboration challenges: Propagation of Goals conveying conflicts between Roles ***covered*** by a Joint Role, ultimately conveying how decisions made by one collaborator can affect the outcomes of another while the two respective Roles are collaborating.

What if each respective Role (covered by the Joint Role of **Data Science Team**) chooses something different to address the conflict? In Figure 5 below, the task "**Treatment Equality**" is chosen within the Actor boundary of the Role "Responsible AI Practitioner". Within the Actor boundary of the Role "**Evaluating Model Performance Role**", the task "**Evaluate Accuracy**" is chosen, which *helps* the softgoal "**Model prediction correctness**". As a result of this softgoal contribution and the associated satisfied dependums in the dependencies between the two Roles ("**Model fairness be evaluated**" and "**Fair class distribution**"), the softgoal "**Model prediction correctness**" is now satisfied, as well as the softgoal "**Accuracy**" now being partially satisfied. Simultaneously, the satisfaction of these intentional elements leads to a tradeoff of other softgoals within the **Evaluating Model Performance** Role: "**False Positive Reduction**" and "**Balanced precision and recall**".

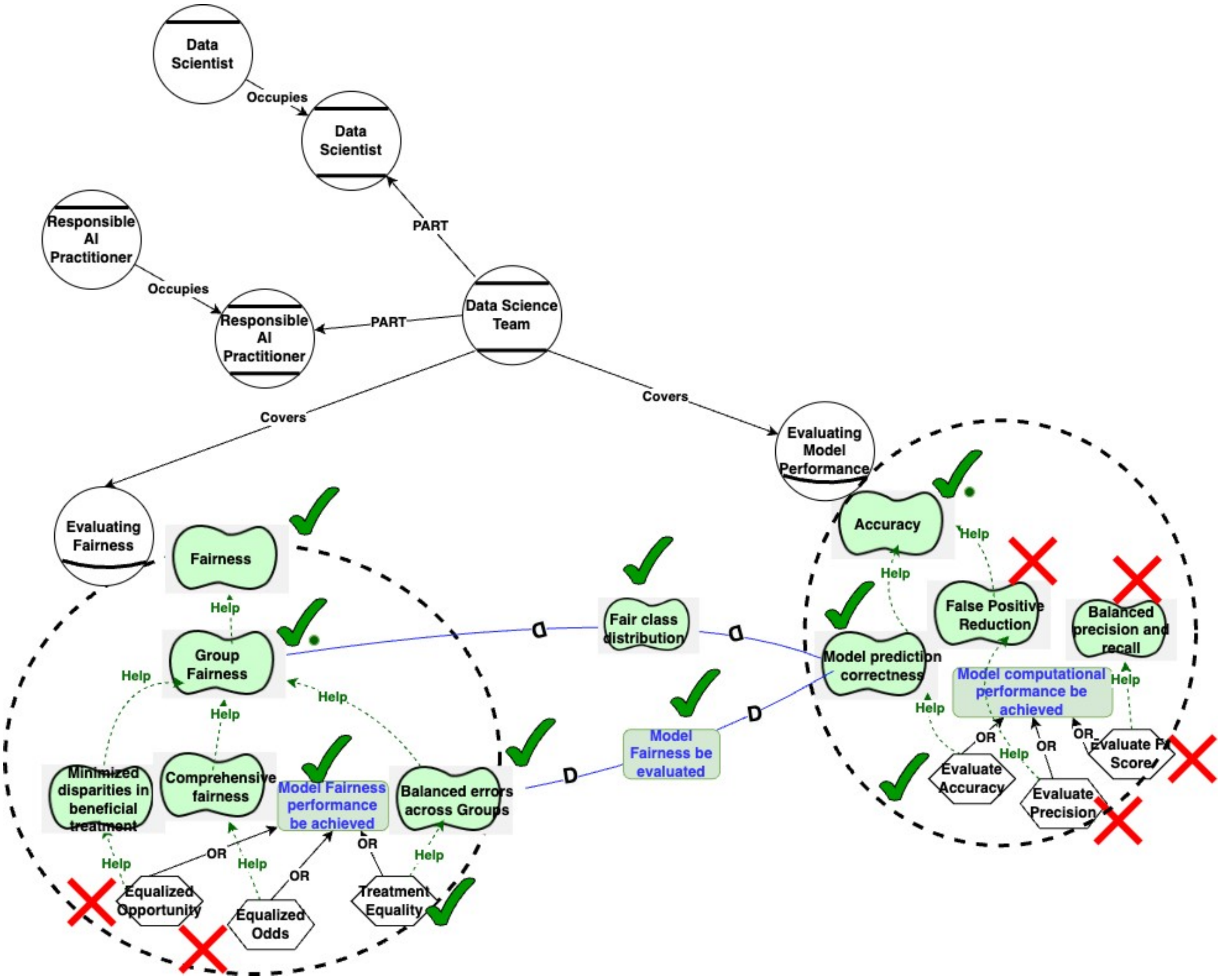

Figure 5: Propagation of Goals showing successful collaboration after tweaking decisions

In each of these examples, the decision points in each of these Roles can affect the success of softgoals between each other. Ultimately, this represents an example of how a single Joint Role (i.e. the collaboration between the **Data Scientist** and **Responsible AI Practitioner** Agents) can have conflicts during collaboration where key decisions made at key design points can affect the outcomes of the collaborator. Going back to our original research problem, the goal propagation analysis examples conveyed using Joint Roles demonstrates promise that the concept can be used to aid in understanding how decisions made by one Actor may affect softgoals that have an impact on the decisions of another Actor.

With respect to limitations, the challenge we face with Joint Roles is, when do we use these Joint Roles? Are they only to be used when a functional goal is a joint responsibility? Another challenge is the following: though the ***Responsible AI Practitioner*** and the ***Data Scientist*** are shown to be a part of this Joint Role, in its current form the AO model expresses that both Positions are responsible for all elements within the Joint Role. But what specific elements are these Positions responsible for? How can we better connect the responsibility areas of the Joint Role to the specific Positions at own them? We aim to explore the concept further in future work with more complex examples.

4. Related Work

Though there have been several GO conceptual modeling techniques in the literature, such approaches have been limited in their ability to analyze tradeoffs that occur during key decision points in the ML lifecycle as well as how they affect Actors involved. Current approaches for conducting Responsible AI provide limited consideration of reasoning to support strategic analysis of business objectives. For example, GR4ML [13] is an existing framework which deals with analytics requirements, but does not address (1) challenges how tradeoffs affect the decisions between actors and (2) considerations for Responsible AI goals. Recently, Kuwajima and Ishikawa [18] proposed a GO conceptual modeling approach, which focuses on a particular set of guidelines: the Ethics guidelines for trustworthy AI from the European Commission. Though this approach uses a GO conceptual modeling approach, it is limited in its coverage of the problem. As a result, this approach cannot feasibly address conflicting goals and priorities about the different, conflicting, interpretations of Responsible AI.

To the best of our knowledge there are no AO approaches for Responsible AI in the literature that specifically deal with tradeoffs both between goals as well as the strategic interests of Actors. Our work aims to extend beyond GO reasoning by facilitating the analysis of intentional modeling as well as analyzing and understanding the interrelationships between autonomous strategic actors in ML project teams and their relationship with Responsible AI goals and Non-Functional Requirements.

Current computational techniques [5] and tools [4] [6] [7] provide conceptual frameworks which enable decision-support for data-driven applications. However, such tools do not support a goal-oriented, well-reasoned approach to achieve Responsible AI goals, and their relationships with strategic business and technical data science goals. Specifically, these approaches do not support important reasoning techniques such as tradeoff mechanisms, a goal refinement process, or the operationalization of those goals.

5. Conclusions & Ongoing Work

In ongoing work, AO conceptual modeling will be used to model complex organizational relationships with respect to ML project teams, including underlying challenges and conflicts which occur that are specific to ML. GO conceptual modeling will be used to develop the capability to explore alternate means to achieve a viable solution that satisfices the interests of each Agent, while considering tradeoffs among multiple competing goals between Agents. AO modeling will extend the GO modeling techniques applied as agents will be abstracted to make distinctions among different types of social actors with agency and individuality.

In future work, we aim to emphasize further aspects of Responsible AI, such as bias, explainability, among others, to consider a holistic lens of “human-centeredness” in our goal-reasoning and AO modeling techniques. Future research will aim to better understand how we can identify specifically where such social responsibility elements as racism and bias exist by analyzing decision points using GO reasoning, and the interaction of these issues among Actors on ML project teams by extending the AO modeling we presented in this work. In future work, we aim to use empirical and literature-based studies to iteratively test and improve our modeling constructs until the language is stable and ready to be tested in an empirical setting.

References

[1] S. Amershi, M. Cakmak, W. B. Knox and T. Kulesza, "Power to the people: The role of humans in interactive machine learning", *AI Magazine*, vol. 35, no. 4, pp. 105-120, 2014.

[2] J. Bosch, I. Crnkovic and H. H. Olsson, "Engineering AI Systems: A Research Agenda", arXiv preprint, 2020.

[3] Dwork, C., Hardt, M., Pitassi, T., Reingold, O., & Zemel, R. (2012). Fairness through awareness. *In Proceedings of the 3rd innovations in theoretical computer science* conference (pp. 214-226).

[4] Hajian, S., & Domingo-Ferrer, J. (2012). A methodology for direct and indirect discrimination prevention in data mining. IEEE transactions on knowledge and data engineering, 25(7), 1445-1459.

[5] R. K. Bellamy, K. Dey, M. Hind, S. C. Hoffman, S. Houde, K. Kannan, P. Lohia, J. Martino, S. Mehta, A. Mojsilović, et al., AI fairness 360: An extensible toolkit for detecting and mitigating algorithmic bias, IBM Journal of Research and Development 63 (4/5) (2019) 4–1.

[6] Tramer, V. Atlidakis, R. Geambasu, D. Hsu, J.-P. Hubaux, M. Humbert, A. Juels, H. Lin, Fairtest: Discovering unwarranted associations in data-driven applications, in: 2017 IEEE European Symposium on Security and Privacy (EuroS&P), IEEE, 2017, pp. 401–416.

[7] Feldman, M., Friedler, S. A., Moeller, J., Scheidegger, C., & Venkatasubramanian, S. (2015). Certifying and removing disparate impact. In proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining (pp. 259-268).

[8] E. Yu, P. Giorgini, N. Maiden and J. Mylopoulos, Social modeling for requirements engineering, MIT press, 2011.

[9] L. Chung, B. A. Nixon, E. Yu and J. Mylopoulos, Non-functional requirements in software engineering, Springer Science Business Media, vol. 5, 2012.

[10] P. Bresciani, A. Perini, P. Giorgini, F. Giunchiglia and J. Mylopoulos, "Tropos: An agent-oriented software development methodology", *Autonomous Agents and Multi-Agent Systems*, vol. 8, pp. 203-236, 2004.

[11] J S. Saltz and N. W. Grady, "The ambiguity of data science team roles and the need for a data science workforce framework", *2017 IEEE International Conference on Big Data (Big Data)*, pp. 2355-2361, 2017.

[12] S. Passi and P. Sengers, "Making data science systems work", *Big Data Society*, vol. 7, no. 2, pp. 2053951720939605, 2020.

[13] S. Nalchigar and E. Yu, "Designing business analytics solutions", *Business Information Systems Engineering*, vol. 62, no. 1, pp. 61-75, 2020.

[14] Sothilingam, R.,, "A Requirements-Driven Conceptual Modeling Framework for Responsible AI," *2023 IEEE 31st International Requirements Engineering Conference (RE),* Hannover, Germany, 2023, pp. 391-395, doi: 10.1109/RE57278.2023.00061.

[15] Sothilingam, R., Pant, V., and Yu, E., "Using i* to Analyze Collaboration Challenges in MLOps Project Teams", *Proceedings of the 15th International i* Workshop 2022*, 2022.

[16] Sothilingam, R., and Yu, E., "Modeling Agents Roles and Positions in Machine Learning Project Organizations", *Proceedings of the 13th International i* Workshop 2020*, vol. 2641, pp. 61-66, 2020.

[17] J. Castro, M. Kolp, J. Mylopoulos, A Requirements-Driven Development Methodology, Advanced Information Systems Engineering: 13th International Conference, CAiSE 2001 Interlaken, Switzerland, June 4–8, 2001 Proceedings 13 (2001) 108–123.

Kuwajima, H., Ishikawa, F. (2019). Adapting square for quality assessment of artificial intelligence systems. In 2019 IEEE International Symposium on Software Reliability Engineering Workshops (ISSREW) (pp. 13-18). IEEE.