March 2022

Trustworthy Hybrid Team Decision-Support

André MEYER-VITALI [a,1] and Wico MULDER [b]
[a] *German Research Centre for Artificial Intelligence (DFKI)*
[b] *Netherlands Organisation for Applied Scientific Research (TNO)*

Abstract. The aim to empower human users of artificially intelligent systems becomes paramount when considering coordination in hybrid teams of humans and autonomous agents. Hereby, we consider not only one-to-one interactions, but also many-to-many situations (multiple humans and multiple agents), where we strive to make use of their complementary capabilities. Therefore, mutual awareness of each others' strengths and weaknesses is crucial for beneficial coordination. In order to address these goals, and in accordance with a hybrid theory of mind, we propose the use of trustworthy interaction patterns and epistemic orchestration with intentions and causal models.

Keywords. computer science, artificial intelligence, software engineering, design patterns, human-agent teams, hybrid AI, neuro-symbolic AI, multi-agent systems, intentions, epistemic logic, causality, simulation

1. Introduction

The aim to empower human users of artificially intelligent systems becomes paramount when considering the collaboration and competition [1] in hybrid teams of humans and autonomous agents [2]. Hereby, we consider not only one-to-one interactions, but also many-to-many situations (multiple humans and multiple agents), where we strive to make use of their complementary capabilities. Therefore, mutual awareness of each others' strengths and weaknesses is crucial for beneficial coordination. Each person and agent has individual knowledge, facilities, resources, roles, capabilities, expectations and intentions. It should be clear for each of them what to expect from each other, in order to avoid misleading anthropomorphism, and how to delegate which tasks to whom.

In order to address this aim for human empowerment, and in accordance with a hybrid theory of mind, we propose the use of hybrid team interaction patterns and epistemic orchestration. Hybrid team interaction patterns are based on previous work on modular design patterns for hybrid actors [3] [4]. Epistemic orchestration is a concept for explicit representation of intentions and causal relationships with the goal of reasoning about them for specifying and composing team architectures and their interactions.

[1]Corresponding Author: Dr. André Meyer-Vitali, Deutsches Forschungszentrum für Künstliche Intelligenz GmbH (DFKI), Stuhlsatzenhausweg 3, Saarland Informatics Campus D32, 66123 Saarbrücken, Germany; E-mail: andre.meyer-vitali@dfki.de.

Trust in hybrid teams emerges from transparent coordination and awareness of mutually shared expectations and explanations.

Hybrid team interactions for multi-party decision-making can be explored in simulated environments where agents are represented as active digital twins and humans participate either interactively or by modelling their (social) behaviour. Nevertheless, the deployment in critical trustworthy real-world use cases is the ultimate goal for hybrid intelligent systems.

As a use case the following scenario is presented: in an urban environment, smart buildings interact with each other to optimise their individual energy consumption. They interact with humans regarding their desired levels of climate comfort, shared characteristics and expected occupation. Both, humans and buildings are modelled as agents, each with their individual behaviour and intentions [5]. As such, they form temporary organisations that negotiate to meet their operational efficiency and sustainability goals.

2. Human Empowerment

Since the very beginning of humanity we have progressed with the help of ever more sophisticated tools and knowledge. Now, knowledge itself becomes a tool. Especially, if we consider the interaction among humans and artificially intelligent agents, it is essential to create and manage awareness of the diverse mutual assumptions. Agents are autonomous virtual or physical entities (software or robots), where autonomy refers to their independence from direct human control. Nevertheless, agents do not have intentions (or even consciousness [6]) on their own. Rather, these intentions and goals are defined by humans who are responsible for their agents' behaviour.

When considering the collaboration and competition in hybrid teams of humans and autonomous agents, we consider many-to-many situations where multiple humans and multiple agents form hybrid teams. The purpose of the agents is to empower humans with providing their complementary capabilities, such as fast and precise information exchange and analysis of large data sets. Agents can play many different roles, but the responsibility for decisions remains, in principle, with humans, for example by verifying, validating and approving proposals for decisions. An essential aspect of meaningful collaboration is to make mutual assumptions and expectations explicit, such that they can be used in deliberation and communication. This is a prerequisite for appropriate delegation of tasks and the accurate and concise descriptions of their underlying intentions.

2.1. Hybrid Theory of Minds

The Theory of Mind (ToM) [7] [8] refers to taking into account other people's mental states of mind when communicating and acting in a social context. Rather than reasoning only with one's own beliefs, desires, intentions, emotions, and thoughts, a person or agent with the awareness of others' states of mind can consider different and mindful acts, depending on a perceived context. This ability allows them to more easily understand, predict, and even manipulate the behaviour of others [9].

In hybrid teams, the states of mind of multiple persons and agents need to be taken into account. The states attributed to humans and agents differ significantly. For example, agents can hardly be understood to have emotions, while these are important for understanding, predicting and reacting to humans' behaviours. On the other hand, intentions - as will be explained below - are very useful for describing and explaining the behaviour of humans and agents, although in different ways. Intentions can be defined concisely in epistemic modal logic.

Also within each category, each person or agent has individual states of mind. It is unrealistic and futile to assume that everybody knows and intends the same things or that knowledge and intentions are global. However, knowledge and intentions can efficiently and effectively be acquired and shared in hybrid teams. With various roles the necessity for diverse knowledge varies, i.e. some people and agents become experts in their field and can be consulted by others whenever their domain is targeted (knowledge on demand).

2.2. Trustworthy Interaction

Instead of relying on AI systems to take over human activities, as some have predicted, it is better to focus on how humans and machines can complement each other's strengths [10]. For example, radiologists are still needed to interpret MRI images [11], but they will have to collaborate with AI systems and those systems need to support the human collaborators by providing insight into their decision-making process. Therefore, a new approach of hybrid or neuro-symbolic AI is necessary for creating trustworthiness [12] [13] [14].

Trustworthiness in interacting with artificially intelligent systems emerges from experience and as a combination of various properties, such as fairness, robustness, transparency, verification, and accuracy [15]. AI systems are trusted when we have confidence in the decisions that they take, i.e. when we understand why they are made [16], even when we disagree.

Mutual awareness of each others' strengths and weaknesses is crucial for beneficial coordination [17]. Each person and agent has individual knowledge, facilities, roles, capabilities, expectations and intentions, according to the above-mentioned hybrid theory of mind.

In hybrid teams, each participant needs to be aware of each other's intentions and capabilities. Primarily, a distinction needs to be made between biological and synthetic actors. But, of course, not all humans or agents are the same, either. It should be clear for each of them what to expect from each other, in order to avoid anthropomorphism, and how to delegate which tasks to whom [18] [19]. Avoiding misleading anthropomorphism is important to prevent bots from pretending to be humans, for example..

For creating trust, it is necessary to organise teams in ways that allow for participant's transparency of each other's roles and responsibilities - at least within the team, but preferably beyond. For this purpose, shared knowledge, intentions, and assumptions must be made explicit - a fundamental principle in software engineering (e.g., Architectural Decision Records [20] [21] [22] and Test-Driven Development [23]) that is equally applicable to organisations and other complex systems.

In a community with trustworthy interactions, it is crucial to establish and enforce social norms [24] [25] [26] [27]. Such norms can be of generic nature or valid only within certain communities or teams and specify transparently what is expected behaviour, what is allowed or forbidden and which are the consequences in case of violations. In addition, knowledge and intentions, but also norms, can change and need to be adapted in due course. Otherwise, such systems and interactions cannot be trusted any longer.

In order to address these goals, and in accordance with the hybrid theory of mind, the use of interaction patterns and epistemic orchestration is promising. Hybrid team interaction patterns are based on previous work on modular design patterns for hybrid actors and explained in more detail in chapter 3. Epistemic orchestration is a concept that combines intentions, epistemic logic and causal models as will be discussed in chapter 4.

3. Hybrid Team Interaction Patterns

Interactions in hybrid teams can be explained and documented as patterns. Below, we list a number of interaction patterns for hybrid team-building and management. These patterns show various ways how to enable collaboration or competition. In collaborative settings, which are what occurs most often, tasks are distributed (and split up in subtasks) to those who can perform them best. Access to available resources is coordinated. In competitive settings, resources will be accessed based on individual preferences and needs.

Whenever participants commit on contributing to a team's intentions they deliberately give up their autonomy to a certain degree [28]. According to mutual matches of individual intentions and team intentions, they declare their commitments. The commitments can be documented in a contract, which increases the transparency and, hence, trust in the team and its members.

3.1. Market Pattern

A market mechanism (see figure 1) is used for potential participants of a hybrid team to describe and publish their capabilities and needs. For example, a public forum serves to announce tasks to be executed. Such tasks are published by an initiator who wants to achieve some goals based on its knowledge (semantic model). Task descriptions include information about the intentions (why), (sub-)tasks (what) and required processes and resources (how). Interested agents and humans can search for and check the requirements of the tasks and, when interested and capable, can apply for performing them within a certain time frame. Known mechanisms are tuple spaces [29] [30] or blackboard architectures [31] [32] [33] [34]. The publish/subscribe pattern [35] is very similar, too. In fact, the potential participants do not communicate directly before the team is formed, but only through the market mechanism. The initiator receives bids, selects which bids to accept and informs the chosen participants that they can start to collaborate on their tasks. With the refusal, an option to allow for amended bids can be sent, including information regarding the refusal. For reasoning about the selection, the initiator can make use of the capability descriptions in the bids, which are described using a shared ontology and logic statements.

3.2. Negotiation Pattern

In contrast to the market pattern, the initiator in the negotiation pattern (see figure 2) needs to know potential participants beforehand in order to invite them to make a bid. Invitations can be passed on, however. The negotiation and selection process is identical to the aforementioned market pattern and the resulting commitments are also valid in this case.

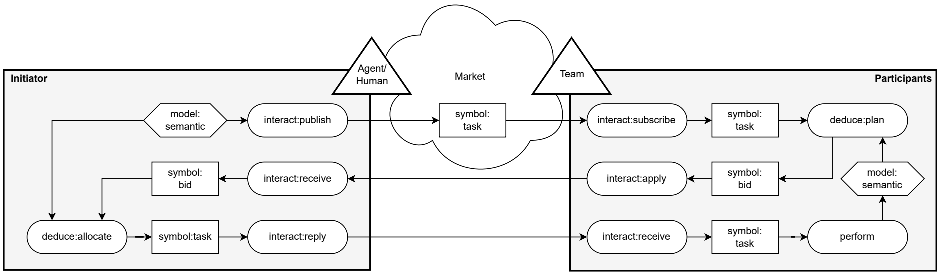

Figure 1. Market Pattern

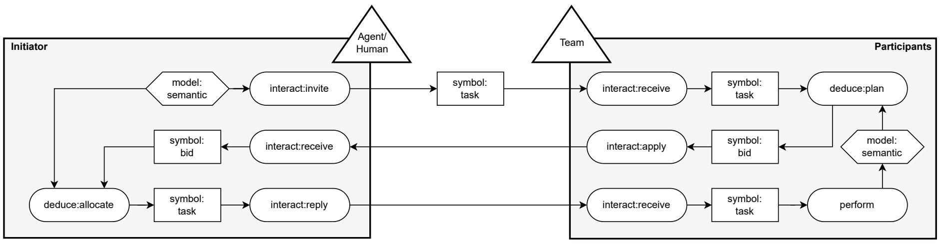

Figure 2. Negotiation Pattern

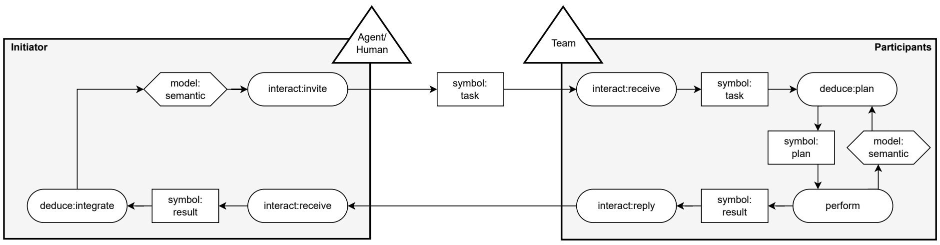

Figure 3. Delegation Pattern

3.3. Delegation Pattern

When an initiator wants a task to be performed, a team can be composed by direct interaction with performers (see figure 3). Tasks are then simply delegated (with or without consent) and resources allocated. Each participant makes a local plan for performing its task and returns the result to the initiator for integrating the sub-tasks and to update its knowledge. In such a scenario with a low level of autonomy of the participants it is still important to explicitly document the task distribution and roles in contracts in order to achieve a high level of transparency and trust.

3.4. Competition Pattern

As mentioned above, not all scenarios are based on mutual agreements. Instead of negotiating and agreeing on shared intentions, humans and agents can also compete for resources among each other (see figure 4). The competition can occur among individuals,

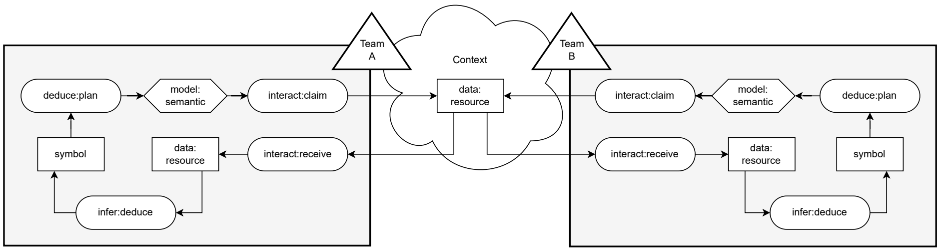

Figure 4. Competition Pattern

but also among teams. Resources are always limited and it is obvious that even a team with the best intentions and mutual agreements cannot claim all of them unilaterally. Therefore, some form of competition emerges within and among teams.

4. Epistemic Orchestration

This chapter introduces ideas for a concept of epistemic orchestration that allows for designing trustworthy AI systems by leveraging explicit models of knowledge, intentions and causal relationships. This concept is work in progress.

The aim and expectation of epistemic orchestration is to enable explicit description and communication of mutual knowledge, intentions, causes and effects in hybrid distributed systems (i.e., complex systems of human and artificial actors in an open world). Without explicit representations such mutual understanding is not guaranteed. However, the representations can also be learned and shared.

Epistemic orchestration combines intentions, epistemic logic and causal models. Rather than programming in a procedural or declarative manner, we propose to program in an intentional manner [5], where intentions describe desired outcomes, but not how to achieve them nor the exact results.

While the current ideas are only a preliminary formulation of what will be developed further, a realistic use case is described in chapter 5 as an experimental playground.

In short, what is to be achieved is to define what each party knows (using epistemic logic), what they want to know or achieve (their intentions as desired future states in the world) and how they are related (by applying causal relationships) - in a multi-party setting where participants can (but are not obliged to) collaborate.

Some pieces of the puzzle are outlined below, but how exactly they fit together to form the big picture is yet to be determined.

Epistemic logic enables the representation of knowledge of agents. It is based on modal logic that makes statements about the necessity and possibility of phenomena. Epistemic logic adds a knowledge operator for each agent to represent the individually known statements: $K_A\phi_1 \wedge K_B\phi_2 \wedge \ldots \wedge K_n\phi_n$. Because the state and knowledge of the world changes, dynamic epistemic logic allows for updating the knowledge about statements and adds actions (with pre- and post-conditions) [36] [37].

In one approach, intentions are discovered by observing the actions of agents from, on the one hand, their knowledge of the actions at their disposal and, on the other hand, the actions they have finally carried out [38]. Therefore, knowledge about the knowledge of others is essential even when it is known not to be true: $K_A K_B K_n \phi$. In this case, it is assumed that other agents' actions and their success in a given world are an indication of their intentionality.

Rather than discovering intentions by observation, in [39] epistemic frames are introduced that represent intentions, in this case in games. Intentions are, thus, explicitly constructed and epistemic frames serve for determining strategies of actions. In another context where agents collaborate, rather than compete with each other, such frames could be shared among agents in a team.

A further ingredient for achieving epistemic orchestration is to use epistemic operators in a causal language [40]. This would allow for connecting intentions to causes and effects via dynamic epistemic logic that is, on the one hand, representing intentions and knowledge, and on the other hand, modelling causal reasoning and learning causal relationships.

By continuing in this direction and integrating the above-mentioned approaches, it should be possible to learn and reason about individual intentional and causal relationships [41]. In a hybrid team setting, this would allow for transparent and trustworthy specification and sharing of individual and group intentions. Once this would be achieved, dynamic task distributions could allow for efficient and explainable (semi-) auto-configurations of modular neuro-symbolic AI systems with a programming language of intentions and interaction patterns. It would, by choice, be semi-automatic, because of the importance to keep humans in the loop of multi-party decision-making. Complex and critical AI systems could be developed according to the principle of Trust by Design.

5. Use Case: Talking Buildings

Hybrid team interactions for multi-party decision-making can be explored in simulated environments where agents are represented as active digital twins and humans participate either interactively or by modelling their (social) behaviour. However, critical real-life applications are the ultimate goal for purposeful hybrid settings in the real world. Our work is applied in the context of urban environments where smart buildings interact with each other to optimise their energy consumption and reduce their CO_2 footprint. They share global goals in managing energy peak load on the power grid.

Both humans and buildings are modelled as agents, each with their individual behaviour and intentions. They form temporary organisations in which the agents collaborate in order to further improve their operational efficiency and meet their global sustainability goals. Sensors in the building measure the actual temperature inside the building, energy consumption and, in some cases, details about its occupation. The derived models are used to control the climate systems of the building. The learning mechanisms involve supervised and unsupervised forms of classification as well as forms of multi-agent reinforcement learning. In the project we study the interaction between the buildings as well as the interaction between the buildings and humans. Furthermore, in order to address the intentions on global impact and to meet the scalability goals, the interaction and algorithmic challenges are accompanied by the need for data sharing infrastructures that fulfil the requirements on interoperability, security and sovereignty challenges.

Describing the structure of hybrid systems in terms of interaction patterns and visualising them by means of boxologies helps to better understand their underlying complex mechanisms. They facilitate designers and engineers in their communication about the design, foster its modularity and safeguard the levels of interoperability. This is important since actors in hybrid teams may be engineered by different organisations. Furthermore, actors in hybrid systems must have a shared understanding of the terminology, contracts, and team-level processes. The patterns and corresponding boxologies support in this.

The buildings share information on their individually learned energy consumption models, patterns in their occupation, local weather conditions, and parameters about their construction or surrounding environment. This is characterised in terms of *market patterns*.

Feedback on the level of perceived comfort and constraints on desired temperatures, energy consumption and model parameters can be defined in terms of *negotiation patterns*. Via their agents the buildings mutually share awareness of their heating curves. In this way, they can expect and anticipate each others' behaviours to a) further optimise their own local consumption model and b) to avoid peak loads on the electricity grid.

Intentions can be shared as common policies, such as the optimisation of energy consumption, efficient interactions with the surrounding power-grid or serving related goals on, e.g., efficient usage of the buildings, or optimising the logistics or safety around the buildings.

The project, named 'Talking Buildings' is an applied research setting in the field of collaborative learning in social contexts. Stakeholders are manifold, e.g. its residents, employees, building owners, construction engineers, energy companies and, last but not least, policy makers. New forms of living and working and their consequences on urban rhythms [42] can be explored. As a result, urban environments become active participants in the endeavour for increased sustainability and citizens' satisfaction in tackling the challenges of the post-pandemic society and climate change, among others.

6. Conclusions and Outlook

In this paper, we have discussed the context and requirements of hybrid team decision support, where multiple humans and agents cooperate. Several design patterns were explained that can be used to model trustworthy and transparent team interaction, both collaboratively and competitively. An outlook was given to the concept of epistemic orchestration. Of course, this will require more work to become useful, but the expectation is to allow for creating future AI systems built on Trust by Design. A use case about talking buildings in an urban context was explained. This is a playing field where the previously mentioned concepts will allow for experimentation and, finally, deployment of trustworthy AI systems based on mutual understanding of intentions in hybrid teams.

References

1. Castañer X and Oliveira N. Collaboration, Coordination, and Cooperation Among Organizations: Establishing the Distinctive Meanings of These Terms Through a Systematic Literature Review. Journal of Management. 2020 Jul 1; 46. Publisher: SAGE Publications Inc:965–1001. DOI: 10.1177/0149206320901565. Available from: https://doi.org/10.1177/0149206320901565 [Accessed on: 2022 Mar 11]

2. Ramchurn SD, Stein S, and Jennings NR. Trustworthy human-AI partnerships. iScience. 2021 Aug 20; 24:102891. DOI: 10.1016/j.isci.2021.102891. Available from: https://www.sciencedirect.com/science/article/pii/S2589004221008592 [Accessed on: 2022 Mar 10]
3. Bekkum Mv, Boer Md, Harmelen Fv, Meyer-Vitali A, and Teije At. Modular design patterns for hybrid learning and reasoning systems. Applied Intelligence. 2021. DOI: 10.1007/s10489-021-02394-3. Available from: https://link.springer.com/epdf/10.1007/s10489-021-02394-3 [Accessed on: 2021 Jun 18]
4. Meyer-Vitali A, Mulder W, and Boer MHT de. Modular Design Patterns for Hybrid Actors. *Cooperative AI Workshop*. Conference on Neural Information Processing Systems. Vol. 2021. NeurIPS. 2021 Dec 14. arXiv: 2109.09331. Available from: http://arxiv.org/abs/2109.09331 [Accessed on: 2021 Nov 17]
5. Dennett DC. The Intentional Stance. Cambridge, MA, USA: A Bradford Book, 1987 Oct 16. 400 pp.
6. Dennett DC. Consciousness Explained. 1st edition. Boston: Back Bay Books, 1992 Oct 20. 528 pp.
7. Premack D and Woodruff G. Does the chimpanzee have a theory of mind? Behavioral and Brain Sciences. 1978 Dec; 1. Publisher: Cambridge University Press:515–26. DOI: 10.1017/S0140525X00076512. Available from: https://www.cambridge.org/core/journals/behavioral-and-brain-sciences/article/does-the-chimpanzee-have-a-theory-of-mind/1E96B02CD9850016B7C93BC6D2FEF1D0 [Accessed on: 2022 Mar 2]
8. Verbrugge R and Mol L. Learning to Apply Theory of Mind. Journal of Logic, Language and Information. 2008 Oct 1; 17:489–511. DOI: 10.1007/s10849-008-9067-4. Available from: https://doi.org/10.1007/s10849-008-9067-4 [Accessed on: 2022 Mar 2]
9. Verbrugge R. Testing and Training Theory of Mind for Hybrid Human-agent Environments. *Proceedings of the 12th International Conference on Agents and Artificial Intelligence, ICAART 2020, Volume 1, Valletta, Malta, February 22-24, 2020*. Ed. by Rocha AP, Steels L, and Herik HJvd. SCITEPRESS, 2020 :11. Available from: https://vimeo.com/396473042
10. Marcus G. Deep Learning Is Hitting a Wall. Nautilus — Science Connected. 2022 Mar 10. Available from: https://nautil.us/deep-learning-is-hitting-a-wall-14467/ [Accessed on: 2022 Mar 10]
11. Chan S and Siegel EL. Will machine learning end the viability of radiology as a thriving medical specialty? The British Journal of Radiology. 2019 Feb; 92:20180416. DOI: 10.1259/bjr.20180416
12. Marcus G. The Next Decade in AI: Four Steps Towards Robust Artificial Intelligence. arXiv:200206177 [cs]. 2020 Feb 19. arXiv: 2002.06177. Available from: http://arxiv.org/abs/2002.06177 [Accessed on: 2021 Nov 27]
13. Marcus G and Davis E. Rebooting AI: Building Artificial Intelligence We Can Trust. Vintage, 2019 Sep 10. 290 pp. Available from: http://rebooting.ai/

14. Akata Z, Balliet D, Rijke Md, Dignum F, Dignum V, Eiben G, Fokkens A, Grossi D, Hindriks K, Hoos H, Hung H, Jonker C, Monz C, Neerincx M, Oliehoek F, Prakken H, Schlobach S, Gaag Lvd, Harmelen Fv, Hoof Hv, Riemsdijk Bv, Wynsberghe Av, Verbrugge R, Verheij B, Vossen P, and Welling M. A Research Agenda for Hybrid Intelligence: Augmenting Human Intellect With Collaborative, Adaptive, Responsible, and Explainable Artificial Intelligence. Computer. 2020 Aug 1; 53. Publisher: IEEE Computer Society:18–28. DOI: 10.1109/MC.2020.2996587. Available from: https://www.computer.org/csdl/magazine/co/2020/08/09153877/1lUB5gL2CnS [Accessed on: 2022 Feb 2]
15. Harbers M, Verbrugge R, Sierra C, and Debenham J. The Examination of an Information-Based Approach to Trust. *Coordination, Organizations, Institutions, and Norms in Agent Systems III*. Ed. by Sichman JS, Padget J, Ossowski S, and Noriega P. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 2008 :71–82. DOI: 10.1007/978-3-540-79003-7_6
16. Rudin C. Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead. arXiv:181110154 [cs, stat]. 2019 Sep 21. arXiv: 1811.10154. Available from: http://arxiv.org/abs/1811.10154 [Accessed on: 2022 Mar 9]
17. Dunin-Keplicz BM and Verbrugge R. Teamwork in Multi-Agent Systems: A Formal Approach. 1st. Wiley Publishing, 2010. 244 pp.
18. Dignum V, Dignum F, and Meyer JJ. An agent-mediated approach to the support of knowledge sharing in organizations. The Knowledge Engineering Review. 2004 Jun; 19. Publisher: Cambridge University Press:147–74. DOI: 10.1017/S0269888904000244. Available from: https://www.cambridge.org/core/journals/knowledge-engineering-review/article/abs/an-agentmediated-approach-to-the-support-of-knowledge-sharing-in-organizations/B8E27F2C2BB97291BB739F168F0D6C35 [Accessed on: 2022 Mar 4]
19. Poslad S. Using Multi-agent Systems to Specify Safe and Secure Services for Virtual Organisations. 2009 Sep 30:258–73. DOI: 10.1007/978-3-642-04879-1_18
20. Nygard M. Documenting Architecture Decisions. Cognitect.com. 500. Available from: https://www.cognitect.com/blog/2011/11/15/documenting-architecture-decisions [Accessed on: 2022 Mar 7]
21. Henderson JP. Architecture decision record (ADR). original-date: 2016-10-18T03:43:44Z. 2022 Mar 7. Available from: https://github.com/joelparkerhenderson/architecture-decision-record [Accessed on: 2022 Mar 7]
22. Chappen E, Hofschneider R, King M, and Ballard T. Architecture Decision Records: Helpful now, invaluable later. 18F Consultancy, U.S. Government. 18F is a technology and design consultancy for the U.S. Government, inside the government. 2021 Jul 6. Available from: https://18f.gsa.gov/2021/07/06/architecture_decision_records_helpful_now_invaluable_later/ [Accessed on: 2022 Mar 10]
23. Beck K. Test Driven Development: By Example. 1st edition. Boston: Addison-Wesley Professional, 2002 Nov 8. 240 pp.

24. Haynes C, Luck M, McBurney P, Mahmoud S, Vítek T, and Miles S. Engineering the emergence of norms: a review. The Knowledge Engineering Review. 2017; 32. Publisher: Cambridge University Press. DOI: `10.1017/S0269888917000169`. Available from: `https://www.cambridge.org/core/journals/knowledge-engineering-review/article/abs/engineering-the-emergence-of-norms-a-review/8761517851A1FFA22BEEEA47085009D8` [Accessed on: 2022 Mar 9]
25. Emelin D, Bras RL, Hwang JD, Forbes M, and Choi Y. Moral Stories: Situated Reasoning about Norms, Intents, Actions, and their Consequences. arXiv:201215738 [cs]. 2020 Dec 31. arXiv: `2012.15738`. Available from: `http://arxiv.org/abs/2012.15738` [Accessed on: 2022 Mar 1]
26. Jiang L, Hwang JD, Bhagavatula C, Bras RL, Forbes M, Borchardt J, Liang J, Etzioni O, Sap M, and Choi Y. Delphi: Towards Machine Ethics and Norms. arXiv:211007574 [cs]. 2021 Oct 14. arXiv: `2110.07574`. Available from: `http://arxiv.org/abs/2110.07574` [Accessed on: 2022 Mar 1]
27. Gâteau B. Normative Multi-Agent Organizations: Modeling, Support and Control, Draft Version. 2007 Dec 1
28. Vecht B van der, Meyer AP, Neef M, Dignum F, and Meyer JJC. Influence-Based Autonomy Levels in Agent Decision-Making. *Coordination, Organizations, Institutions, and Norms in Agent Systems II*. Ed. by Noriega P, Vázquez-Salceda J, Boella G, Boissier O, Dignum V, Fornara N, and Matson E. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 2007 :322–37. DOI: `10.1007/978-3-540-74459-7_21`
29. Gelernter D and Carriero N. Coordination languages and their significance. Communications of the ACM. 1992 Feb 1; 35:97–107. DOI: `10.1145/129630.129635`. Available from: `https://doi.org/10.1145/129630.129635` [Accessed on: 2022 Mar 9]
30. Wells G. A Tuple Space Web Service for Distributed Programming. *PDPTA*. 2006. DOI: `10.5220/0001517000930100`
31. Hayes-Roth B and Hayes-Roth. A blackboard architecture for control. Artificial Intelligence. 1985; 26:251–321. DOI: `10.1016/0004-3702(85)90063-3`
32. Corkill DD. Blackboard Architectures and Applications. *Design Alternatives for Parallel and Distributed Blackboard Systems*. Academic Press, 1989 :pp. 99–136
33. Jagannathan V, Dodhiawala R, and Baum LS, eds. Blackboard architectures and applications. Perspectives in artificial intelligence v. 3. Boston: Academic Press, 1989. 524 pp.
34. Craig I. Blackboard Systems. Norwood, NJ: Ablex, 1995
35. Gamma E, Helm R, Johnson R, Vlissides J, and Booch G. Design Patterns: Elements of Reusable Object-Oriented Software. 1st edition. Reading, Mass: Addison-Wesley Professional, 1994 Nov 10. 416 pp.
36. Ditmarsch H van, Hoek W van der, and Kooi B. Dynamic epistemic logic. Vol. Synthese Library Series. Berlin: Springer, 2007. DOI: `10.1007/978-1-4020-5839-4`
37. Ditmarsch H, Ruan J, and Verbrugge R. Sum and Product in Dynamic Epistemic Logic. J Log Comput. 2008 Aug 1; 18:563–88. DOI: `10.1093/logcom/exm081`

38. Chetcuti-Sperandio N, Goudyme A, Lagrue S, and Lima Td. First Steps for Determining Agent Intention in Dynamic Epistemic Logic. 12th International Conference on Agents and Artificial Intelligence (ICAART 2020). SCITEPRESS - Science, Technology Publications ; SCITEPRESS - Science, and Technology Publications, 2020 Feb 22:725. DOI: 10.5220/0008991207170724. Available from: https://hal.archives-ouvertes.fr/hal-02946618 [Accessed on: 2022 Feb 23]
39. Roy O. A Dynamic-Epistemic Hybrid Logic for Intentions and Information Changes in Strategic Games. Synthese. 2009; 171. Publisher: Springer:291–320. Available from: https://www.jstor.org/stable/40495999 [Accessed on: 2022 Feb 23]
40. Barbero F, Schulz K, Smets S, Velázquez-Quesada FR, and Xie K. Thinking About Causation: A Causal Language with Epistemic Operators. arXiv:201016217 [cs]. 2020 Oct 30. arXiv: 2010.16217. Available from: http://arxiv.org/abs/2010.16217 [Accessed on: 2022 Feb 18]
41. Mueller S and Pearl J. Personalized Decision Making - A Conceptual Introduction. Technical Report R-513. 2022 Mar 10. Available from: http://causality.cs.ucla.edu/blog/index.php/2021/04/29/personalized-decision-making/ [Accessed on: 2022 Mar 11]
42. Nevejan C, Sefkatli P, and Cunningham S. City Rhythm. 2018 May 1