

電子情報通信学会

ソフトウェアインタプライズモデリング研究会 (SWIM)

# Human-Machine Teaming (HMT)

人間とAIの関係について

2024年2月16日

SC 42/WG 4 (Use cases and applications) 国際コンビーナ, SC 42国内専門委員会 幹事・WG 4小委員会主査  
学校法人中央情報学園/国立研究開発法人 産業技術総合研究所, デジタルアーキテクチャ研究センター  
丸山 文宏 (Fumihiro Maruyama)

SC 42/WG 4 (Use cases and applications) プロジェクトエディター (ISO/IEC TR 24030, ISO/IEC 5338, ISO/IEC PWI 42109)  
富士通株式会社 富士通研究所 AIトラスト研究センター

○鄭 育昌 (Yuchang Cheng)

# 目次

- HMT概要
  - Human-Machine Teaming (HMT) とは？
  - HMTの意義
- HMTのフレームワーク
  - 人間と機械の関係性
  - 各関係性タイプとユースケース (事例)
  - 関係性の組み合わせ
  - チームの進化
  - 技術課題
- まとめと今後の展望

# HMTの概要

# Human-Machine Teaming (HMT) とは？

人間と機械(AIやロボット)のチームによる問題解決  
人間と機械の両方が問題解決の役割を果たす

診断支援AI + 医者 (人間)

患者を治療する

自動運転車 + ドライバー (人間)

運転する

画像認識AI + 検査担当者 (人間)

出荷前の製品を検査する

配車 + ドライバー (人間)

ライドシェアリングを提供する

ロボット + 発見AI + 科学者 (人間)

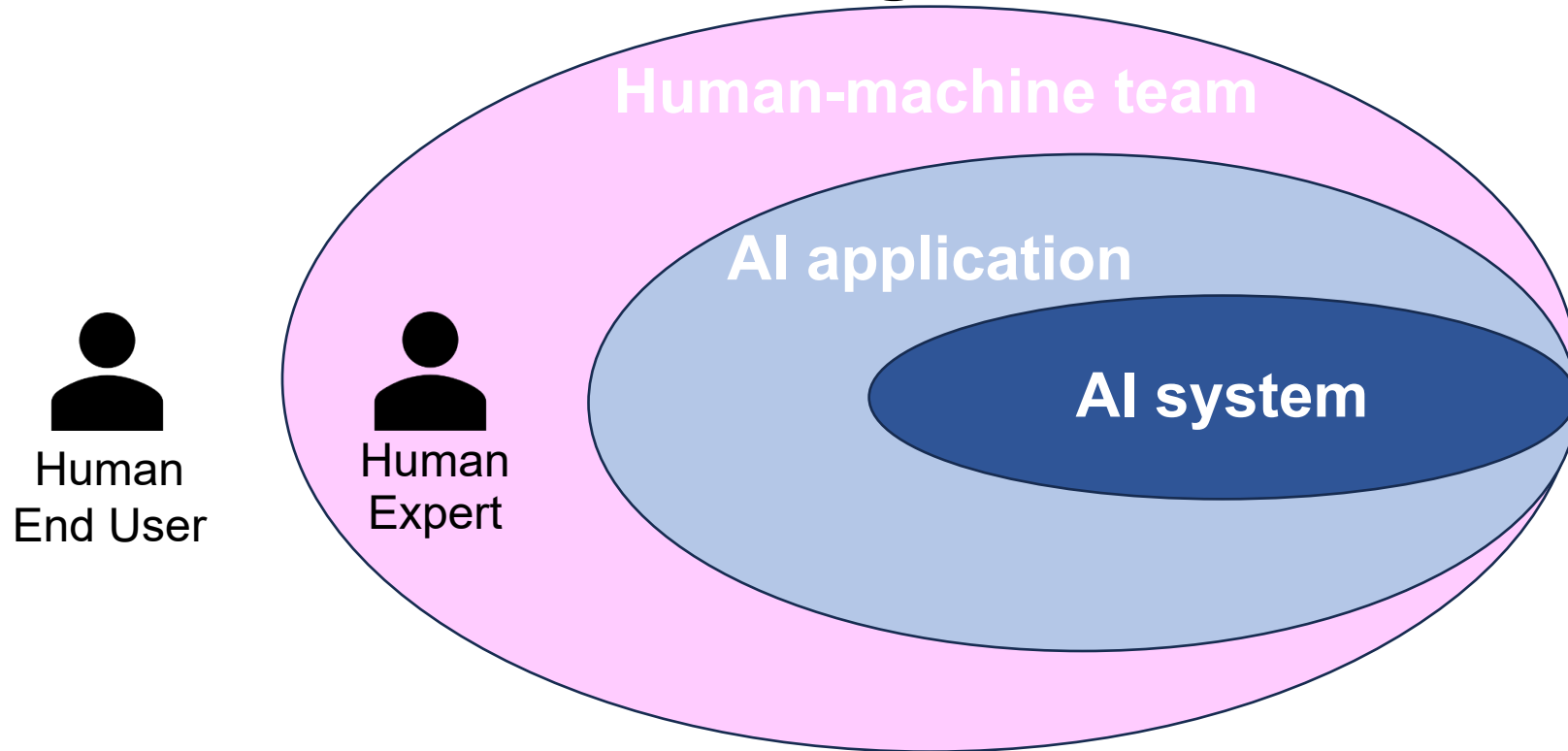
重大な科学的発見をする

パーソナルエージェント + ユーザ (人間)

日々の問題を解決する

“human-machine teaming“の定義 (ISO/IEC 22989, JIS X 22989):  
integration of human interaction with machine intelligence capabilities  
(人間と、機械の知的能力との相互作用の統合)

# Human-Machine Teamingの範囲



- Human Userがない場合（人間の専門家が同時にエンドユーザ）やHuman Expertがない場合（Human userタイプ）もあり得る
- HMTに関する規格は、AIアプリケーションに対すると同様にHuman Expertに対する要求事項も含み得る

# Human-machine teamingの意義： なぜ？

- HMTは、問題解決に関わる人間と機械（AI）の関係の全体像を提供
  - 5つの関係は、人間と機械が対称性を持つすべてのパターンをカバー
  - AIの能力が急速に進歩しているため、全体像の集中検討が必要
  - チーミングの進化も考慮が必要

# Human-machine teamingの意義： いいところ？

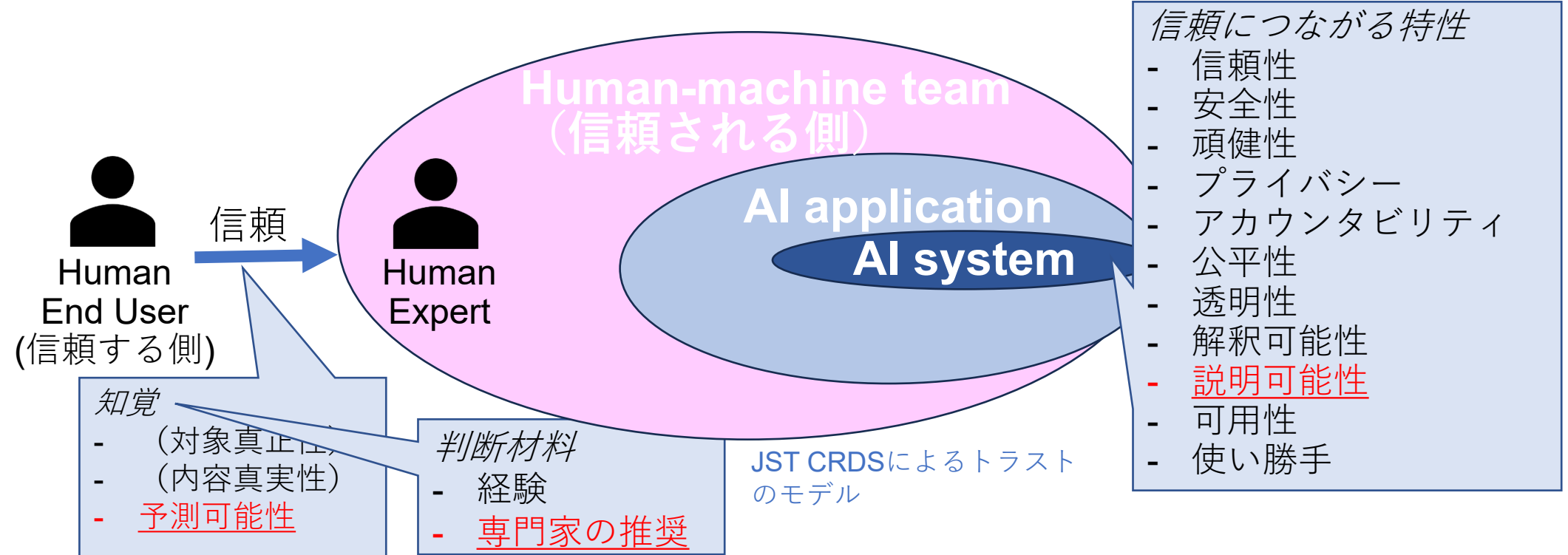
- 社会・産業面

- 人間の専門家は職を失うことはない
- 現行法制度(例えば医療)に適合する
- 人間の専門家の作業負荷によるボトルネックが解消される
- 一般の人々は(例えば、医療や法律) AIシステムの利用に、専門家よりも消極的
- HMTは人間の信頼を高める

- 技術面

- 機械知能と人間知能のベストミックスが実現できる
- 人間の専門家に高度なトレーニングの機会を提供する
- 一般人へのAIの出力の説明は(少し)簡単になる (人間の専門家の介入により)

# Human-machine teamingの意義： HMTはAIのみよりも人間の信頼を高める

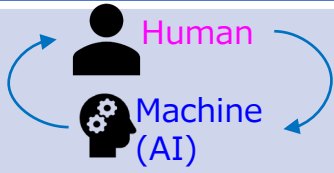

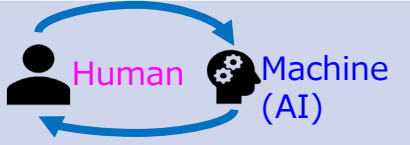




- 専門家（人）とAIのチームがユーザ（人）に与えられるもの
  - 説明可能性：チーム内の専門家に対するAIシステムの説明は一般の人に対する説明より容易で、その専門家がチームを代表してユーザ（人）に説明できる
  - 予測可能性：AIシステム単独より人とAIのチームの方が予測可能
  - 専門家の推奨：ユーザ（人）はチーム内の専門家の推奨を感じる

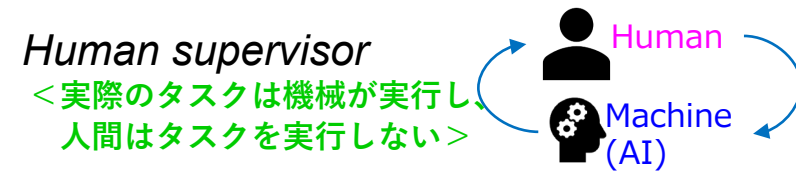


HMTのフレームワーク

# 人間と機械の関係性タイプの概要

関係性（タイプ）	実際のタスクの実行	知識の流れ
<i>Human supervisor</i> 	機械だけ	ほとんどなし
<i>Human mentor</i> 	人間も機械も実行に関わる	機械が主体 人間は一部 人間から機械への一方向
<i>Peer</i> 		両者が同程度 双方向
<i>Machine mentor</i> 		人間が主体 機械は一部 機械から人間への一方向
<i>Machine supervisor</i> 	人間だけ	ほとんどなし

# Human supervisor タイプ

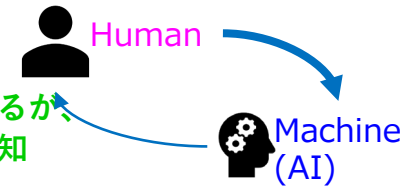


- 概要：
  - 人間は機械を監督する
  - 人間は決して実際の仕事をしない
  - 実際の作業はすべて機械が行う
- 代表的なユースケース（事例）：
  - サーチエンジンや経路検索の日常的な利用
    - 結果を見て、必要に応じて条件を変更して再実行
  - 自動運転レベル4（走行ルートなど特定条件下での完全な自動運転）
    - 車両内か遠隔で監視にあたる「特定自動運行主任者」が上司

# Human mentor タイプ 概要

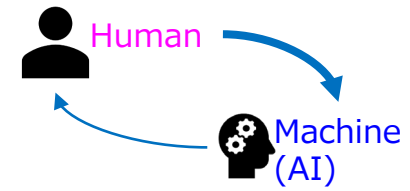
Human mentor

<実際のタスクは主に機械が実行するが、人間も一部のタスクを実行する。知識は人間から機械に流れる>



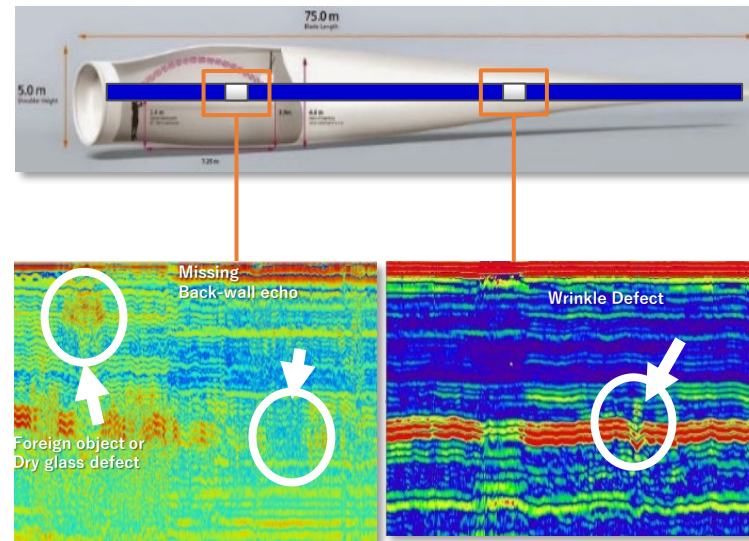
- 機械が実際のタスクを主に実行
- 人間はそのタスクを実行しようと思えば実行できるものの、主としてメンターとして機械を指導
- それまで人間の専門家が専ら行っていたタスクをAIにも実行できるように学習させたケース
- 対象とするタスクに関して人間がAIと比べてより多くの知識を持っていることが想定される
  - 対象とするタスクに関する知識は人間からAIに流れる

# Human mentor タイプの事例： 画像認識AIによる製品検査



(ISO/IEC TR 24030:2021 “Use cases” A.4の事例)

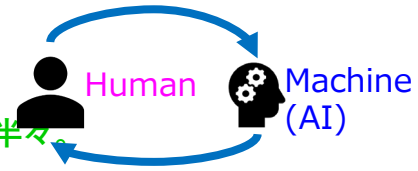
- 画像認識AIによる風力発電用ブレードの製品検査
  - 超音波スキャン画像のディープラーニングによる認識により欠陥を自動検出
  - 様々な欠陥に対して高い検出率 (95%以上)
  - 製品検査担当者はAIがフラッグを上げた部分 (< 20%) だけをチェック
  - AI導入前は一日1枚しかできなかった検査が4~5枚可能に



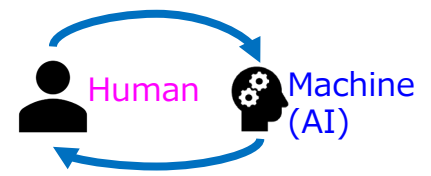
# Peerタイプ 概要

Peer

< 人間と機械が対等。  
タスクの実行は人間と機械が半々。  
知識は双方向に流れる >



- 人間とAIが同格
- どちらもタスクを実行する能力を有している
  - それぞれ異なるやり方でタスクを実行
- 条件によってどちらのパフォーマンスが優れているかが変わってくるため、状況に応じてどちらがタスクを実行するかを決める
  - 人間とAIがタスクを分担して並列に実行することもあり得る



# Peerタイプの事例： 医療診断支援AIと臨床医

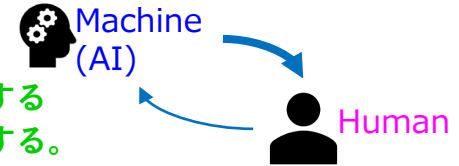
- 機械学習による医療診断支援システム
  - 救急患者の時系列データを常時モニター
  - 患者が敗血症を発症するリスクを予測
  - ハイリスクの患者が検出されたら、臨床医に警告
- 臨床医
  - 診断と治療に関する意思決定に最終的な責任を持つ
  - 枕元での患者との接触による視覚などの豊富な情報に基づいて、AIとは異なる診断プロセスを実行
  - 新たな患者への注意を喚起したり、患者の状態の急変を知らせてくれたりする、臨床医にとっての「第二の眼」としてAIを評価

# Machine mentorタイプ 概要

- 人間が主としてタスクを実行
- 一部のタスクに関しては機械が実行
- 囲碁や将棋などのゲームのように機械が完全に人間を超えていなくとも、少なくとも一部のタスクに関しては機械の能力が人間を上回っていることがこのタイプの前提

Machine mentor

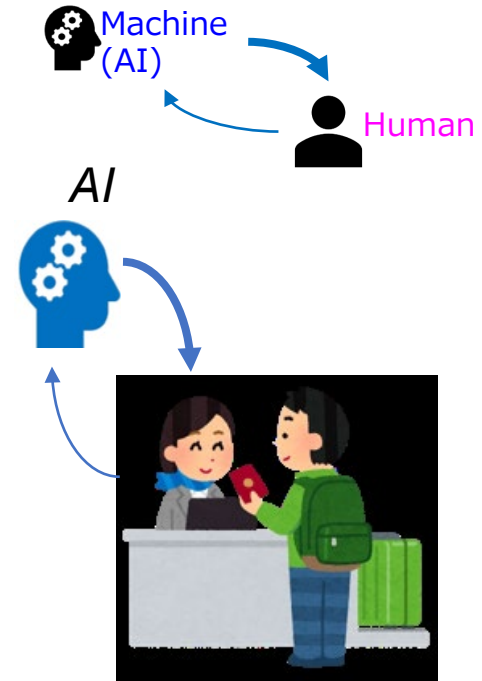
<実際のタスクは主に人間が実行するが、機械も一部のタスクを実行する。知識は機械から人間に流れる>





# Machine mentorタイプの事例： 人間のコミュニケーション支援

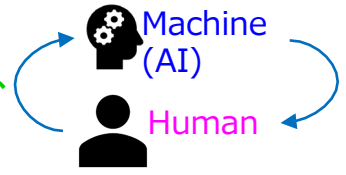
- AIは、接客の現場でサービススタッフとお客様のやり取りを観察・分析する
  - 音声テキスト変換、感情分析、表情分析
- AIはサービススタッフにサポートとトレーニングを提供する
- AIとサービススタッフがチームを組んで接客する  
(*machine mentor*)



# Machine supervisorタイプ 概要

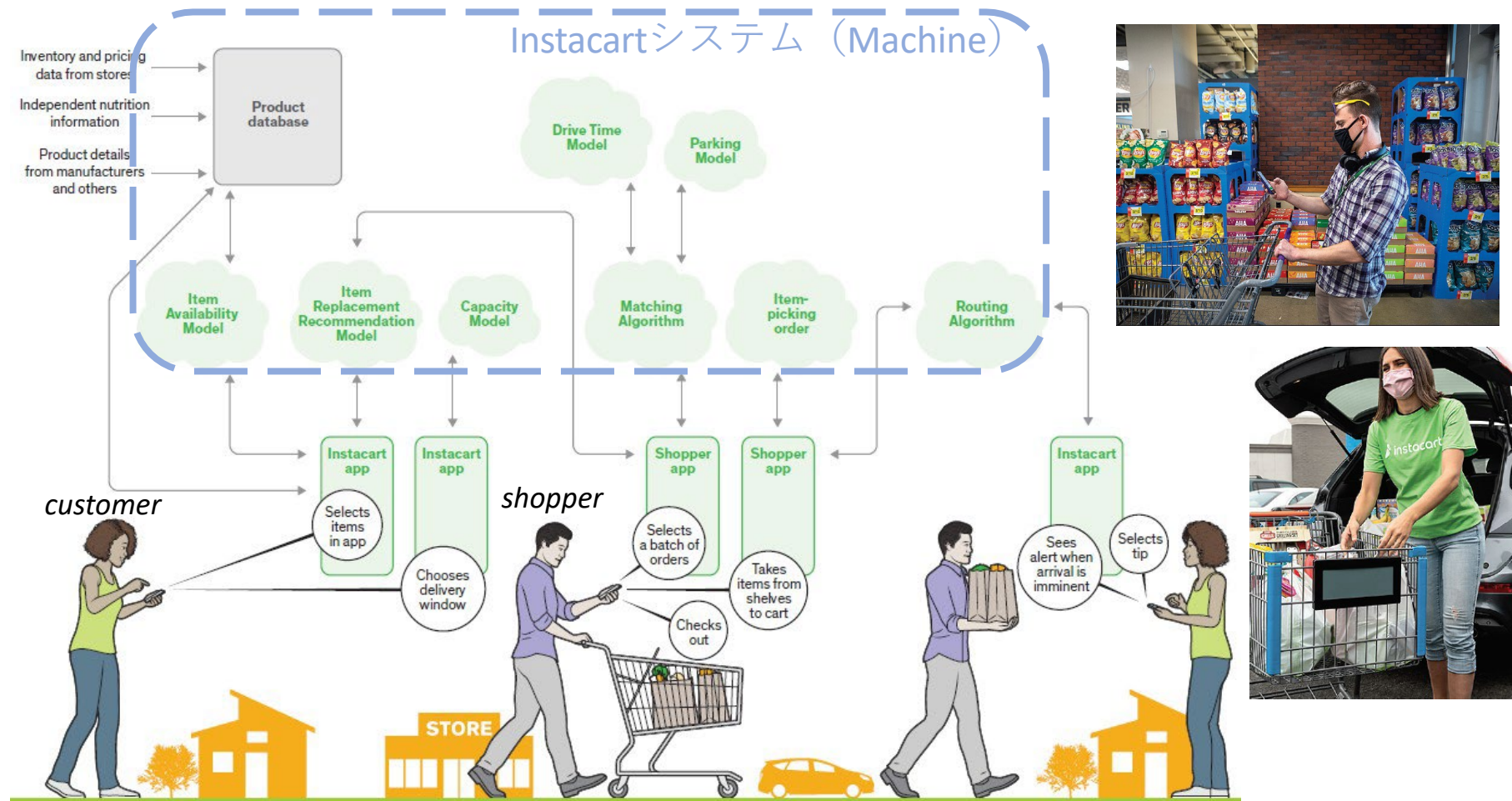
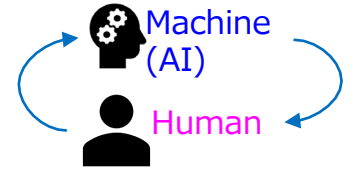
Machine supervisor

<実際のタスクは人間が実行し、  
機械はタスクを実行しない。  
知識の流れはほとんどない>



- 人が専ら実際のタスクを実行
- 上司の立場にある機械はタスクの実行には携わらない
- 代表的な例のUberでは、人間のドライバーがシステムの指示を受けて運転
  - 乗り合い (shared) ではリクエストが追加されることもあり、経路を変更してその追加リクエストの乗客をピックアップ
- タスクの割り当てが機械 (AIの機能を実装したシステム) によって行われる
- 一般にクラウドソーシングはこのタイプ

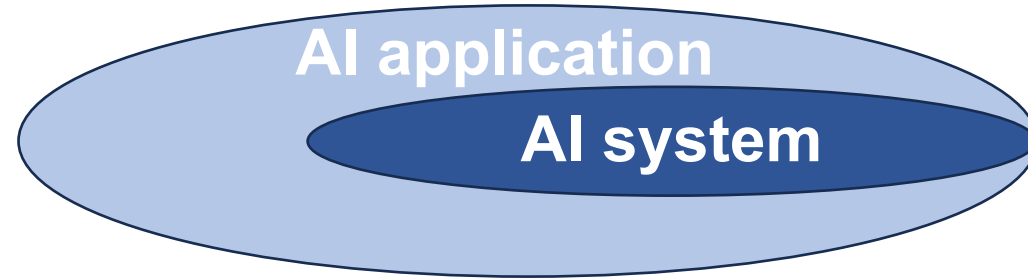
# Machine supervisorタイプの事例 Instacart（食料品買物代行サービス）



# Machine supervisorタイプの事例： Instacartサービスを実現するHMT

- Human(専門家)は：
  - 人間の買物要員
- Machineは、次のAIシステムで構成される：
  - 商品在庫予測
  - 代替品推薦
  - キャパシティ予測：
    - 将来の各配達時間帯（2時間程度）に何人の買物要員が利用可能か予測
  - 顧客からの注文と買物要員のマッチング
  - 到着時間予測
  - 買物要員がストアに到着するまでの時間を予測
  - ピッキング支援：
    - 効率的な買い回りの順序をルールベース・システムで提示
  - 配達経路生成

# Human end user タイプ



- 概要：

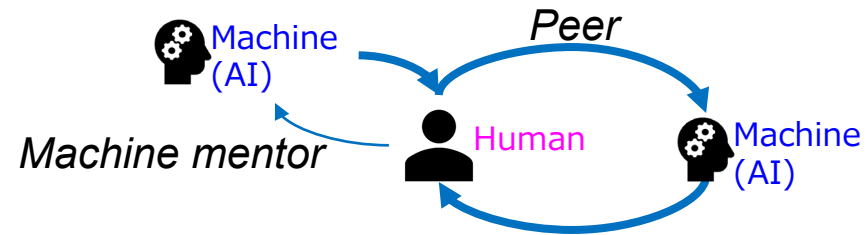
- 人間は専門家ではなく、エンドユーザーのみ ⇒ HMTではない
- AIにタスクを割り当て、AIから結果を受け取るのみ
  - 問題解決には関与しない
  - AIの出力を評価しない/できない

- 代表的な事例：

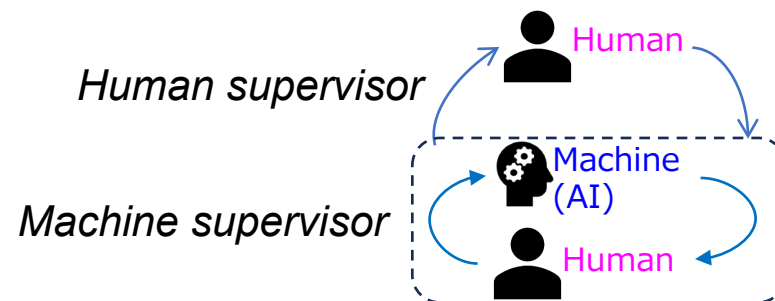
- 自動運転レベル5 (完全自動化) で自動配送を希望する乗客または荷主。
- 割り当てられたテーマのプロンプトをChatGPTに提供し、その応答を教師にそのまま送信する学生。

# 複数関係性タイプの組合せ

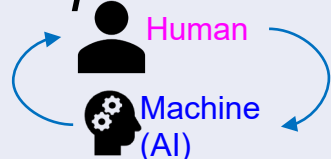

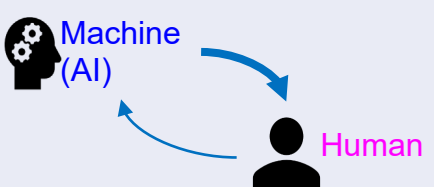
- 一部のAIソリューションでは、複数のタイプを組み合わせる必要がある
  - 人がピアタイプでタスクを実行する人 (人間またはAI) を決定すると、別のAIが導入され、人間の決定を監視し、信頼較正のために介入する (machine mentor)



- Joint型のHMTも可能
  - Machine supervisor型のHMTに対し、チームを監督する上司 (Human supervisor)がいる

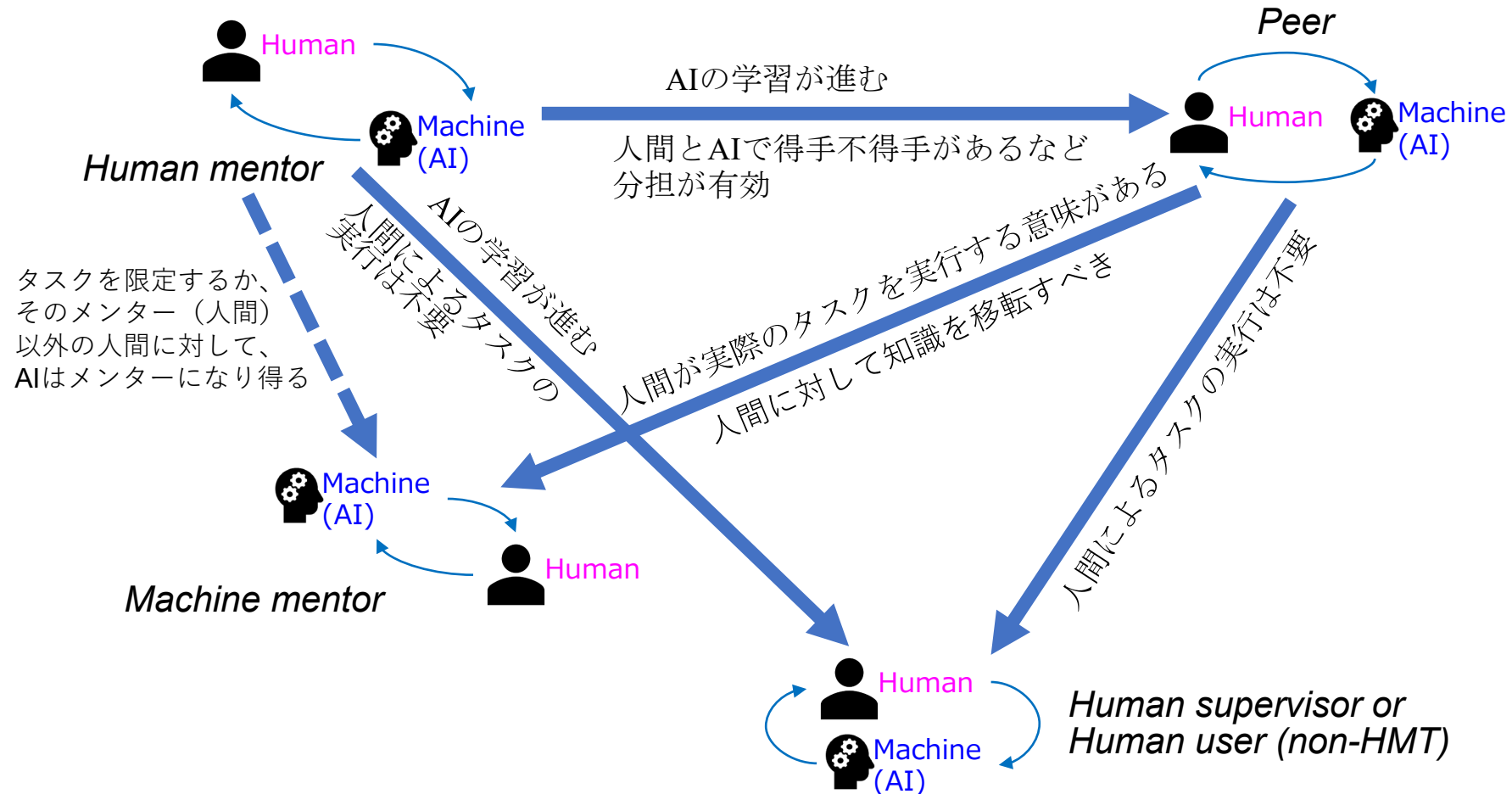


# HMTの関係性タイプの変化： 自動運転のレベルによる関係性の変化

自動運転のレベル	関係性（タイプ）
レベル5（完全運転自動化） *人間は、自動運転車の乗客や自動配達を依頼する荷主など	<i>Human end user</i>
レベル4（走行ルートなど特定条件下での完全な自動運転） *人間は車両内か遠隔で監視にあたる「特定自動運行主任者」	<i>Human supervisor</i> 
レベル3（特定条件下でシステムが運転）	<i>Peer</i> 
レベル2（自動の追い越しなど高度な運転支援）	<i>Machine mentor</i> 
レベル1（衝突被害軽減ブレーキなどの運転支援）	

# HMTの関係性タイプの変化： AIの進化による関係性変化

- AIの性能とAIへの信頼の向上に伴い、人間と機械の関係性が変化





# HMTの技術課題：関係性タイプx領域

関係性 \ 領域	Governance & Empowerment	Safety, Trust & Performance	Model evolution	Interaction
<i>Human supervisor/user</i>	<ul style="list-style-type: none"> <li>・ チーム設計</li> <li>・ 人間による監督</li> </ul>	<ul style="list-style-type: none"> <li>・ 透明性</li> <li>・ 説明可能性</li> <li>・ 評価指標</li> </ul>		<ul style="list-style-type: none"> <li>・ 機械のレポート</li> <li>・ 結果の説明</li> <li>・ 人間からのフィードバック</li> </ul>
<i>Human mentor</i>	<ul style="list-style-type: none"> <li>・ チーム設計</li> <li>・ 権限委譲</li> <li>・ エスカレーション</li> </ul>	<ul style="list-style-type: none"> <li>・ 協調安全</li> <li>・ 評価指標</li> </ul>	<ul style="list-style-type: none"> <li>・ 相互モデル</li> <li>・ 知識移転による共進化</li> </ul>	<ul style="list-style-type: none"> <li>・ ラベル/モデル</li> <li>・ エスカレーション依頼</li> </ul>
<i>Peer</i>	<ul style="list-style-type: none"> <li>・ チーム設計</li> <li>・ 動的権限委譲</li> </ul>	<ul style="list-style-type: none"> <li>・ 協調安全</li> <li>・ 信頼較正</li> <li>・ 相互の意図理解</li> <li>・ 評価指標</li> </ul>	<ul style="list-style-type: none"> <li>・ 相互モデル</li> <li>・ 知識移転による共進化</li> </ul>	<ul style="list-style-type: none"> <li>・ 相互の状態（人間の感情も含む）</li> <li>・ 権限</li> </ul>
<i>Machine mentor</i>	<ul style="list-style-type: none"> <li>・ チーム設計</li> <li>・ 人間のモチベーション向上</li> </ul>	<ul style="list-style-type: none"> <li>・ 評価指標</li> </ul>		<ul style="list-style-type: none"> <li>・ 人間のモニタリング</li> <li>・ 機械による介入/助言</li> </ul>
<i>Machine supervisor</i>	<ul style="list-style-type: none"> <li>・ チーム設計</li> <li>・ タスクの割当</li> <li>・ 人間のモチベーション向上</li> </ul>	<ul style="list-style-type: none"> <li>・ 評価指標</li> </ul>		<ul style="list-style-type: none"> <li>・ 人間のレポート</li> <li>・ 機械からのフィードバック</li> </ul>

# まとめと今後の展望

- HMTの定義と概要：
  - HMTは、人間とAIやロボットが問題解決のためにチームを組むことを指す
- 人間と機械の関係
  - 人間と機械の階層的関係、実際のタスクを誰が行うか、人間と機械の間でどのように情報が流れるかによって5つのタイプに分類
  - 5つのタイプ： Human supervisor/user; Human mentor; Peer; Machine mentor; Machine supervisor
  - AIの進歩とパフォーマンスの向上に伴い、人間-機械チームのタイプは時間とともに変化する
- HMTの技術課題：
  - 各タイプの関係には、人間と機械の間の相互作用、モデルの進化、安全性と信頼性、ガバナンスとエンパワーメントなどの領域に特有の技術的課題が存在する
- 今後の展望
  - HMTは研究テーマのみならず、AI国際標準化のテーマとしても関係規格の開発を進んでいる（ISO/IEC JTC 1/SC 42/WG 4）

ご清聴ありがとうございました

# References

- Praveen Damacharla et al., “Common Metrics to Benchmark Human-Machine Teams (HMT): A Review” IEEE Access <https://ieeexplore.ieee.org/document/8404030>
- Rodney Brooks, “A HUMAN in the Loop - AI won’t surpass human intelligence anytime soon” IEEE Spectrum, October 2021.
- MIT Center for Collective Intelligence, “Supermind Design Primer” June 2021. <https://cci.mit.edu/wp-content/uploads/2021/08/Supermind-Design-Primer-v2.pdf>
- NATIONAL ARTIFICIAL INTELLIGENCE RESEARCH AND DEVELOPMENT STRATEGIC PLAN 2023 UPDATE - A Report by the SELECT COMMITTEE ON ARTIFICIAL INTELLIGENCE of the NATIONAL SCIENCE AND TECHNOLOGY COUNCIL, May 2023. <https://www.nitrd.gov/pubs/National-Artificial-Intelligence-Research-and-Development-Strategic-Plan-2023-Update.pdf>
- EASA Artificial Intelligence Roadmap 2.0 <https://www.easa.europa.eu/en/downloads/137919/en>
- A.4 “AI solution to quickly identify defects during quality assurance process on wind turbine blades” ISO/IEC TR 24030:2021 Artificial Intelligence (AI) - Use cases.
- Katharine E. Henry et al., “Human-machine teaming is key to AI adoption: clinicians’ experiences with a deployed machine learning system” Nature Portfolio Journal Digital Medicine (2022) 5:97 ; <https://doi.org/10.1038/s41746-022-00597-7>
- Hiroaki Kitano, “Creating the Engine for Scientific Discovery: Nobel Turing Challenge as a grand challenge project in AI and Systems Biology” IJCAI-19 Invited Talk <https://www.ijcai19.org/invited-talks.html>
- H. Kagermann, Y. Nonaka (Eds.), “Revitalizing Human-Machine Interaction for the Advancement of Society” <https://en.acatech.de/publication/revitalizing-human-machine-interaction-for-the-advancement-of-society-perspectives-from-germany-and-japan/>
- K. Okamura and S. Yamada, "Empirical Evaluations of Framework for Adaptive Trust Calibration in Human-AI Cooperation," in IEEE Access, vol. 8, pp. 220335-220351, 2020, doi: 10.1109/ACCESS.2020.3042556.
- Makoto Takahashi, “Authority Transfer System to Enhance AI Human Cooperation” [https://www.jst.go.jp/mirai/jp/uploads/saitaku2019/JPMJMI19B4\\_takahashi.pdf](https://www.jst.go.jp/mirai/jp/uploads/saitaku2019/JPMJMI19B4_takahashi.pdf)
- Sharath Rao and Lily Zhang, “The Algorithms That Make Instacart Roll: How Machine Learning And Other Tech Tools Guide Your Groceries From Store To Doorstep” IEEE Spectrum, March 2021.
- Kazuki Uehara et al., “Evidence Dictionary Network Using Multiple Instance Contrastive Learning for Explainable Pathological Image Analysis” 2023 IEEE 20th International Symposium on Biomedical Imaging (ISBI).
- Yuushi Toyoda et al., “The effects of Autonomy and Task meaning in Algorithmic Management of Crowdwork” AAMAS 2020.