Industry Corner
Debdeep Paul

In this "Industry Corner" column, we interview Dr. Antoine Fagette, the Product Design Authority for Collaborative Autonomy at Thales Digital Solutions. Dr. Fagette received his Master's degree in 2010 from École Centrale de Paris (France) with a major in Advanced Systems and Robotics. During his studies, he developed a keen interest for autonomous and intelligent systems while working for a year at Sagem Defense and Security (Safran) on the UAV program. After graduating, he was hired by Thales as a research engineer and relocated to Singapore to complete his PhD thesis in Computer Vision with University Pierre and Marie Curie hosted at the French-Singaporean research lab IPAL. He defended in June 2014 and took the lead of the Center of Excellence for Smart/Safe Cities and Maritime Security at the Thales Research and Technology lab in Singapore. Since 2018, he has been with Thales in Montreal (Canada) where he has been in charge of the R&T roadmap pertaining to Collaborative Autonomous Systems. In 2024, he took on the role of Product Design Authority for Collaborative Autonomy to support the Thales roadmap to deliver mature technological solutions in that domain.

In this interview, Dr. Fagette will give us a view of the emerging field of collaborative human-autonomy teaming, provide some open research problems, and how academia-university partnerships can help tackle these challenges. He concludes with advice for young researchers wishing to enter the field and what topics/skills are needed. We hope you enjoy the interview!

---

1. Please give us a bit of background about yourself and how you ended up at Thales.

My entire academic and professional journey has been geared towards robotics and autonomous systems, with a preference for use cases in harsh environments such as space, underwater or defence. I graduated in 2010 from École Centrale de Paris, in France, with a major in Robotics and Advanced Systems. During my studies, I spent a gap year at Sagem Defense and Security on the UAV program. At that time, drones had not flooded the civilian market (with all the risks and opportunities that implies) and the armed forces were just starting to use them in operations. I had always been attracted to new technologies but this was a real eye-opener on Innovation as a career path. In 2010, I did my end of study internship at Thales. In France, Thales is one of the top employers, especially for engineers interested in complex systems deployed on critical missions and embedding cutting-edge technology. Moreover, Thales is investing significantly in Research & Technology, which allows us to create some of that cutting-edge technology. On top of that, Thales has corporate research labs in Europe, Asia and North America, which was key for me, as I wanted to be able to move and live around the world. So, after that end-of-study internship I applied for a position with Thales and three months later I was in Singapore, starting as a research engineer and company-sponsored PhD student. I spent seven years there, working on projects to enable systems to learn, understand and decide autonomously, while experiencing the arrival of Deep Learning (and the revival of Artificial Intelligence, at last!). In 2018, as Thales was opening its center of expertise in AI in Montreal, I moved to Canada where I expanded my research activities on Autonomy to Collaborative Autonomy and Human-Autonomy Teaming (HAT).

2. Please give us a bit of background on Thales and its role in developing human-autonomy or human-AI teaming/collaboration systems across different sectors. How has it evolved over the last few years with the boom in AI?

Thales is a global technology company that operates in various sectors, including defence and security, aerospace and space, digital identity and cybersecurity, integrating and delivering mission critical systems to its customers and end-users. One can think of the avionics on board an aircraft to command for a safe flight and monitor various functions such as its navigation or flight management operations, a Command and Control (C2) system enabling coordination, communication, decision-making and control of military operations, or an Identity and Access Management system to protect, control, and monitor access to a critical infrastructure or network.

Even before the introduction of Artificial Intelligence (AI) and having the possibility to enable first levels of autonomy into complex systems, Thales has been working on human-system interaction. We have always kept in mind that we are building critical systems so that, despite their complexity, they can be trusted by their end-users, namely aircraft pilots, C2 operators, security officers and personnel accessing secured premises. For that purpose, Thales has invested significant efforts in Research & Development (R&D) activities centered around the human such as Human Factors to create better Human Machine Interfaces (HMIs), Cognition to better understand the human decision patterns and Human Sensing to capture the physiological state of an operator in a given situation and allow the system to behave accordingly.

The recent boom in AI, in particular due to the significant progress made in Machine Learning (ML) thanks to the advent of Deep Learning (DL) has brought evolutions on how we develop human-AI collaboration and ultimately human-autonomy teaming at three levels. First, it allowed us to introduce ML-based techniques in human-centric fields such as Cognition and Human Sensing. Combining these new data-driven algorithms with the more traditional model-based approaches, unleashes the full potential of hybrid AI to boost the performance of those technological components thus improving the synergy between the human operator and its system. On the second front, we saw the emergence and maturing of techniques and tools that are key to enable AI-based systems collaboration. Here I am thinking for example of Multi-Agent Reinforcement Learning (MARL), which allows multiple agents to learn how to tackle a given task as a team. In addition, one would be blind not to have noticed the rise of Large Language Models (LLMs) which are enabling a machine to build a representation of the data, information and knowledge it is subjected to, and to exploit it for enhanced human-machine dialog. Currently, Thales is involved in the research community to contribute on those techniques and tools, understanding how to integrate them in our solutions but also how humans should be involved throughout the whole lifecycle of those AI components, from design and training up to test and deployment.

Which leads to the third evolution, pertaining to autonomous systems and human-autonomy teaming. Because of its adaptability and its capability to evolve, an autonomous system cannot be considered merely as a tool or a static component. An autonomous system truly has the potential of becoming a teammate for the humans and the other autonomous systems it is working with. Much as for humans, that relationship and the trust that it implies will need to be built and constantly adapted throughout time and throughout joint trainings, joint drills and joint deployments. Working together through all those phases, humans and autonomous systems will learn together and from each other, potentially building synergies that would not be achievable with traditional systems.

This has a tremendous impact on the role Thales has in developing human-autonomy teaming. An impact at the technical level but also at the strategic level, at the business level and at the ethical level. By ethical choice, Thales' AI solutions maintain the human in the loop, hence we cannot develop, position, sell or deploy autonomous systems without enforcing this commitment.

3. Within human-autonomy/AI teaming and collaboration, what are some open research problems that new students could focus on and try to help solve?

Human-autonomy teaming is a field of research that evolves particularly quickly and that is quite rich in terms of challenges. The list of open research problems below is not provided with any specific order of importance or priority, and above all, it does not pretend to be exhaustive:

- **Mutual understanding**: the notion of mutual understanding needs to be tackled from an external and internal standpoint. As the representation of the world varies from one human to another, the possibility to achieve a similar understanding of a situation by a human and by an autonomous system seems quite a challenge. Developing shared mental models would certainly lay interesting foundations to build on. On top of understanding a situation similarly, humans and autonomous systems need to understand each other's intents to address that situation efficiently and without hindering each other's decisions and actions.
- **Trust and transparency**: autonomous systems face a major trust challenge from humans. Trust is not something that can be bought, it needs to be built. And building that trust has to start very early in the process of creation and deployment of an autonomous systems, i.e., during training. However, humans are quite irrational in terms of not evaluating mistakes in the same way as successes. It is hard for us to gain trust, and it is easy to lose it.
- **Asynchronicity**: machines and humans do not apprehend time the same way. This poses serious challenges, in particular to orchestrate joint trainings. While autonomous systems will be able to discretize their perception of the situation, the actions they take and the consequences of those actions, humans are rather more continuous. Therefore, the sharing of experience, key for multi-agent learning processes, is rather hectic at this time.
- **Dynamic team formation and task allocation**: the configuration of the team and the characteristics of the teammates (knowledge and experience for humans, algorithms and models for autonomous systems) will have a strong impact on the outcome of the mission to complete. Moreover, assembling the dream team is not sufficient, it is also crucial to let the right teammate execute the right task. And all this needs to be done dynamically in order to remain relevant even as the situation unfolds.
- **Heterogeneous teams**: ultimately, teams composed of humans and autonomous systems are heterogeneous, notably because of the significant difference between a human and an autonomous

system. But also because two autonomous systems can be different, whether be it in terms of functional capabilities or in terms of algorithmic choices to realize those functions. Currently, training a Multi-Agent System (MAS) e.g., through MARL, requires the trained function and the chosen algorithm to be the same throughout all agents. A research avenue that we are currently exploring is the concept that we have coined HeMARL (Heterogeneous MARL).

- **Human factors and interface design**: collaboration and teaming relies a lot on the capability to exchange information. At the moment, the level of interaction that is possible between two humans is far more advanced, efficient and refined than what is possible between a human and an autonomous system. We need to explore methods to design effective user interfaces and interaction mechanisms that improve communication and information exchange between humans and autonomous systems. Leveraging user experience to investigate intuitive and user-friendly displays, haptic feedback, natural language dialog, and multimodal interfaces is crucial.
- **Cyber and AI attacks prevention**: for an autonomous system to collaborate and team-up, it needs to communicate and exchange information with external entities. The moment there is a continuous interaction with a human, we need to minimize the risk and detect any potential cyber or AI attack, leveraging if possible the team setup. If under attack, we need to understand how that affects the autonomous system and the team's dynamics, isolate the system under attack and mitigate the attack's impact on the system and on the team. Ultimately, we need to put in place the proper mechanisms to recover and re-integrate the autonomous system to the team.
- **Ethical and social implications**: human-autonomy teaming will have important ethical and social implications. Those implications need to be addressed, in particular with respect to bias, fairness, accountability, privacy, trust and user adoption.

4. What role do you see academia playing in trying to help address these concerns? Is there room for academia-industry partnerships in this domain?

The role that academia can play to contribute and address the challenges discussed previously is multi-faceted. Academia has an essential role in educating and training the next generation of researchers, engineers, and practitioners in the field of human-autonomy teaming. This includes developing specialized curriculum, workshops, and training programs to promote expertise and awareness. Furthermore, thanks to the structure of universities, which gather several faculties under one organization, academia can facilitate collaboration between fields of expertise such as psychology, computer science, robotics, ethics, neuroscience, cognition, and human factors. This interdisciplinary potential can foster comprehensive approaches to untangle complex issues associated with human-autonomy teaming. Additionally, academia are expected to conduct research, experimental studies and prototyping at a more fundamental level in order to build strong knowledge foundations on which further applied research and technological developments can rely. Academia can also contribute to the collection and analysis of diverse datasets related to human-autonomy teaming. By conducting studies and experiments with human participants, researchers can generate valuable data to validate hypotheses, assess simulation models of system performance, and gain greater validated understanding of human behavior and decision-making in collaborative settings. Last, on the front of adoption and ethics, academia can help with the development of policies, guidelines, and ethical frameworks for the responsible deployment and operation of human-autonomy teaming. This involves examining legal and policy implications, privacy and security concerns, and societal impact assessments.

Whether there is room for industry-academia partnership is not even questionable. First, pragmatically, because there is co-funding available. If I take the example of Canada, the government, at the federal and provincial levels, is supporting a number of co-funding programs that are encouraging industries and academia to work together. Second, because academia-industry partnerships are a great facilitator for technology transfers, from academic research to industry applied R&D and applications. For their part, the industry partners have the necessary resources to translate academic discoveries into commercial products, services, solutions, processes, practices or methods, benefiting the whole ecosystem. Last because academia and industries are not at the place on the innovation chain. While the academia focuses more on low Technological Readiness Levels (TRLs), the industry is dedicated to maturing what comes out of the academia and increasing the TRL up to making products.

5. What advice would you offer young researchers entering the field of human-autonomy teaming, from the latest technical expertise that is/will be required, to the non-technical skills sought today by hiring managers at Thales and elsewhere?

Artificial Intelligence and Autonomy are research domains that change and evolve very quickly, and quite abruptly sometimes. Therefore, my first advice would be to stay updated on the latest advancements in the cross-knowledge and technologies contributing to human-autonomy teaming. You need to remain familiar with state-of-the-art tools, frameworks, and methodologies, which means that you need to know them and know how to use them. As a consequence to this fast-paced evolution, always be prepared to re-invent yourself, to pivot and to be surprised (if not taken aback).

Human-autonomy teaming is a field of research that requires a blend of technical skills from various disciplines. Hence, I would think it is important to cultivate a strong interdisciplinary background, combining fields like computer science, psychology, cognition, engineering, robotics, neuroscience, cybersecurity and ethics. This will enable you to approach problems from multiple perspectives and collaborate effectively with experts in different domains. It will also help you integrate considerations from all those fields in the field of expertise you are developing.

In human-autonomy teaming, there is the human component that is rather key and challenging to deal with. I can only emphasize the importance of human-centered design principles. Understanding the user needs, cognitive capabilities, and behavioral patterns will help ensure that the interaction between humans and autonomous systems is intuitive, efficient, safe and adopted. Something that really helped me in the past few years was to familiarize myself with the Design discipline and in particular user experience design. Find good designers, work with them, and learn from them.

Effective communication, both oral and written, is essential for interdisciplinary teamwork. Develop strong presentation skills to convey complex ideas clearly and concisely. Practice collaborating with researchers from different backgrounds, as teamwork and collaboration are often success factors in achieving more complex knowledge in the field. If you think of it, you are trying to foster collaboration and teaming. It would be quite strange no to apply that to the management of your research.

My last advice would be to cultivate your ethical awareness. You need to demonstrate a strong understanding of the ethical considerations related to human-autonomy teaming. Be conscious of the potential societal impact, privacy concerns, biases, fairness and adoption issues that may emerge when designing, deploying and operating autonomous systems. It is important that you stay informed about ethical trends and guidelines, and contribute to the development of responsible practices.

6. Any last words of advice?

AI-enabled autonomous systems truly bring a change of paradigm in the relationship between human operators and their systems. Currently, it is still difficult for an end-user to grasp the full extent of the potential that they could achieve with their autonomous systems if properly trained and operated. As opposed to the procurement of more traditional systems, end-users should not expect to receive out-of-the-box fully trained autonomous systems, without ever taking part to their training. Without this crucial involvement of the end-user, the system cannot adapt its behavior to its human teammate as its behavior model is frozen. More importantly, the end-user's experience, knowledge and know-how are not taught to the autonomous system. In other words, the operator and the system are not learning to work together. Furthermore, as the behavior model of the autonomous system is frozen, it cannot learn from data collected during actual deployment or even on context evolutions specific to the end-user. As the autonomous system cannot evolve while its environment changes and adapts, it becomes obsolete without ever having had the possibility to achieve an optimal level of performance.

Furthermore, besides raising their level of awareness, the end-users also need to be able to act. They need to be equipped with the right tools and competencies to allow their systems to continue training and adapting, seamlessly for the end-user. Having access to the right data collection and data management systems on the one hand and to the adequate training environments and training platforms on the other is critical. As for the competencies, there are those that we know and master nowadays. However, they are well suited in our academic and industrial contexts, but they need to be extended to the customer and end-user contexts. It is foreseeable that new competencies and job families will appear in the near future to fulfil needs at the end-user level that are still to be uncovered and clearly defined.