Reflective Essay: Intelligent Agent Systems

WHAT?

Throughout the Intelligent Agent Systems module, I engaged with a range of concepts that deepened my understanding of agent-based computing and its real-world applications. The module began by introducing foundational ideas around autonomous agents, such as reactivity, proactivity, and social ability, and evolved into complex applications using adaptive algorithms, agent architectures, and communication protocols. These concepts align with Russell and Norvig's (2021) core principles of intelligent behaviour and decision-making.

My final project, "A Modular Multi-Agent AI System for Anomaly Detection in Fitbit Data," synthesised this learning. I designed a multi-agent architecture where each agent had a clear, modular role—data retrieval, validation, AI analysis, reporting, and security. This system detected anomalies in health data using an Isolation Forest model with PCA for interpretability. It was the culmination of iterative learning from each weekly unit, seminar task, and formative activity. By distributing roles across agents, the project reflected real-world scalable architectures, as described in Maes' (1991) work on agent networks.

Additionally, participating in Collaborative Discussions and completing activities like constituency-based parse trees, ethical evaluations of generative AI, and industry applications in FinTech and Industry 4.0 allowed me to apply theory to modern AI challenges. These tasks reinforced the importance of combining

theoretical knowledge with reflective practice (Rolfe, Freshwater and Jasper, 2001; The University of Edinburgh, n.d.), enabling me to critically assess and adapt agent designs for practical use.

SO WHAT?

One pivotal moment in the development process came during the evaluation of our MAS architecture in Unit 6. Initial designs proposed a centralized model for data flow, but team feedback, particularly from Marwa and Abdulhakim, raised concerns about scalability and bottlenecks. In response, we collaboratively pivoted to a multi-agent architecture. This change, driven by peer critique, taught me the value of adaptive planning and iterative co-design—principles central to agile methodologies and intelligent system evolution (Tomasino, 2025). It also underscored the significance of early-stage design reviews as a safeguard against architectural rigidity in real-time systems.

Moreover, feedback on our project report pointed out the need for deeper justification of our AI model choices, particularly the selection of Isolation Forest and One-Class SVM. This critique prompted me to investigate model alternatives such as Autoencoders and DBSCAN, which I presented informally to the team. While we maintained our original choice for its real-time efficiency, this process helped me develop evaluative thinking and evidence-based reasoning—key aspects of critical reflection as advocated by Rolfe et al. (2001). It made me more confident in defending design decisions using empirical performance metrics and theoretical fit.

Another area where peer insight shifted my approach was in addressing false positives during data validation. Originally, I underestimated this risk. However, Tala’s input led us to implement a forensic logging mechanism that flagged and reviewed anomalies before final reporting. This small but meaningful change reinforced the importance of continuous peer review and reflective responsiveness—skills I now see as essential in both technical and team-oriented AI projects. It also revealed how even minor feedback, when thoughtfully considered, can elevate the overall robustness and accountability of intelligent systems.

NOW WHAT?

Looking ahead, this module has laid a strong foundation for my future in AI and intelligent systems. I now feel confident in developing modular AI architectures and using libraries like Scikit-learn and NLTK to prototype agent systems. More importantly, I understand the value of structuring agents around autonomy and communication, inspired by Maes' (1991) agent network architecture and reinforced through the concepts introduced by Russell and Norvig (2021). I now view intelligent agents as components in distributed systems that can respond and adapt to uncertainty.

The challenges I faced taught me the importance of documenting design decisions and staying reflective throughout technical work. I plan to integrate structured journaling into future projects to track emotional and cognitive progress, especially when tackling new algorithms or team-based systems. This

approach aligns with Rolfe et al.'s (2001) guidance on critical reflection as a continuous process of learning through experience.

My understanding of ethical AI has evolved, and I now approach problems with a broader perspective. I intend to pursue further reading on explainable AI (XAI) and SHAP interpretability methods, especially for health-related AI applications, where transparency is critical (Floridi, 2019). These methods can help mitigate biases and support trust in high-stakes decision-making environments.

In terms of transferable skills, this module strengthened my:

- **Time management** – by balancing readings, coding, discussions, and project work weekly.
- **Critical thinking** – by evaluating agent architectures and adaptive algorithms across sectors.
- **Technical communication** – through e-portfolio updates, annotated code, and peer feedback.
- **Ethical awareness** – through discussions of bias, privacy, and regulation in deep learning, reinforcing the need for value-sensitive design (University of Essex Online, n.d.).

As a next step, I plan to explore smart manufacturing use cases more deeply and simulate intelligent agent systems for classroom use in AI education. This would merge my passion for teaching with my growing technical expertise. I also aim to explore how intelligent agents can be designed inclusively—integrating

accessibility features and social responsibility, as advocated by Floridi (2019). This human-centric lens is now central to how I design, evaluate, and reflect on AI systems.

References:

- Bostrom, N. (2014) *Superintelligence: Paths, Dangers, Strategies*. Oxford: Oxford University Press.
- Floridi, L. (2019) *The Logic of Information: A Theory of Philosophy as Conceptual Design*. Oxford: Oxford University Press.
- Marr, B. (2024) 'The Dark Side Of AI: How Deepfakes And Disinformation Are Becoming A Billion-Dollar Business Risk', *Forbes*. Available at: https://www.forbes.com/sites/bernardmarr/2024/11/06/the-dark-side-of-ai-how-deepfakes-and-disinformation-are-becoming-a-billion-dollar-business-risk/ (Accessed: 17 April 2025).
- Rolfe, G., Freshwater, D. and Jasper, M. (2001) *Critical Reflection in Nursing and the Helping Professions: A User's Guide*. Basingstoke: Palgrave Macmillan.
- The University of Edinburgh (n.d.) *Reflection Toolkit*. Available at: https://www.ed.ac.uk/reflection/toolkit (Accessed: 17 April 2025).
- University of Essex Online (n.d.) *Short Guide to Reflective Writing*. Study Skills Hub. Available at: https://studyskills.sites.essex.ac.uk/reflective-writing (Accessed: 17 April 2025).