De-Individuation of the Modern Subject in the Age of Artificial Intelligence. The Case of Self-Driving Cars and Algorithms for Decision-Making.

Fabio Iapaolo
Urban and Regional Development Ph.D. Programme
Inter-university Department of Regional and Urban Studies and Planning

This research project interrogates shifting conceptions of human subjectivity and evolving forms of technical agency in view of recent advances in the field of Artificial Intelligence. Linking their work to possibilities for positive social, political and environmental change, a significant number of scholars from the posthumanities and the new materialism has attempted to decentre the human subject, alongside the qualities traditionally associated with human exceptionalism, by acknowledging the agency of the nonhuman, including technology. This body of work, however, has fallen short of convincingly accounting for the embodied and embedded agential capabilities of contemporary computational media.

To fill this gap, and calling for the reintroduction of materiality alongside technicality as key dimensions for a critical understanding of the spatialized effects of technology, this research project brings forth the critique of the liberal humanist subject by investigating the integration of AI technologies into existing social and spatial systems. Inter-disciplinary in orientation and theoretical in scope, this research project seeks to reassess theoretical discussions of the complex entanglement of nonhuman agency with (post)human subjectivity entertained within critical theory through a technically-aware investigation of a particular AI system: self-driving cars.

Considered an essential preliminary activity for conducting the case study, this work begins by problematizing anthropocentric notions of AI manifesting both in the early research as well as in the imagination of AI in popular culture. Contrary to widespread claims for AI displacing the human subject, it is argued that, in reality, the liberal autonomous self has long been central to the design and imagination of intelligent machines. Drawing on posthumanist/feminist studies of technology, this study thus provides an in-depth operational discussion of key concepts like intelligence and autonomy. By unveiling the operational logics of the dominant paradigm of AI to date, namely, machine learning, it thus clarifies the ambiguous conceptual overlap between autonomy and automation.

Working at the intersection of social and computer science perspectives on technology, and using as theoretical framework the cognitive assemblage devised by posthumanist scholar Katherine Hayles, this research project thus delivers the first in-depth, technically-aware analysis of self-driving cars by unpacking their material functioning, inner logics, and systemic complexity. By mapping out the multiple agents concurrently affecting the vehicle's behaviour, it shows that driving

decisions always result from layered (multi-located and multi-temporal) interactions between the human and technology, hence neither can be said to be operating within fully autonomous realms. Countering both claims for human autonomous agency, and dominant views of AI as autonomous technologies to which decision-making power is delegated, it is argued that the most paradigmatic aspect of contemporary automated systems is the unprecedented level of complex imbrication and dynamic entwinedness between human culture, technics, and the environment.

This work ends with a discussion of city-specific spatiality and political materiality through an in-depth investigation of the pre-emptive logics, and present limits, of machine vision and cognition in relationship to urban variegated form and sociality. While countering dominant techno-deterministic interpretations of social innovation and spatial transformation, this work offers insight into a post-anthropocentric understanding of AI—namely, not as an abstract property susceptible of replication within discrete machines, but rather as a distributed property emerging through material interactions occurring among a multiplicity of embodied agents (human, nonhuman, and technological) within/with their sociotechnical environments.