Title: Artificial Consciousness Raises Novel Concerns for Data Privacy

Over the past decade, numerous advocates and authors have raised concerns about consumer data privacy related to usage in big data applications, including Large Language Models (LLMs) and other deep learning systems. However, if we take seriously the possibility that an AI system might develop artificial consciousness (AC), I contend that this possibility creates a new type of concern about data privacy that has not previously been focused on.

Up until now, ethical objections to data harvesting and use in the field of AI research and development have generally been framed in terms of an individual right to privacy grounded in control over one's image and information. The theory is that an individual should have control over where information about them, derived from their personal actions, flows. The individual has a right to decide to what purposes this information can be put, if any. This contrasts with a theory of privacy that puts the locus of harm on nonconsensual knowledge rather than loss of control. On the knowledge theory of privacy, I am only harmed–my right to privacy is only violated–if someone comes to know some fact about me that I did not choose to reveal. A control theory of privacy is broader–I can have my right to privacy violated even if no one comes to know anything about me without my consent. To be clear, I broadly agree with this interpretation as applied to currently existing practices and technologies. This understanding of objections to data harvesting for use in big data makes sense in light of current practices, which tend to indiscriminately gather information about millions of individuals, aggregate this information, (semi-)anonymize this information, and then put it to use as part of training sets for deep learning models or as aggregated statistical information for predictive modeling. This control framework makes sense as a basis for objection even in light of the anonymized, aggregated use of many individuals' data: institutions harvesting and using data cannot defend their usage on the basis that "no one sees *your* information, so your right to privacy is not violated"--the basis of the claim is not knowledge, but control.

However, if we take seriously the possibility of artificial consciousness emerging in a deep learning context, I believe even the knowledge theory of control would be violated. If a deep learning system gained some level of consciousness–if it had genuine awareness of the content of the data it processed, rather than merely processing it–then individuals could object to their privacy being violated not merely in the sense that they were not granted adequate control over where their information flowed, but that "someone" they had not granted access to had gained specific knowledge about them–in this case, the someone in question would be an artificial agent.

In my presentation, I hope to analyze what this possibility would look like: in what sense could an artificially conscious agent be said to know facts about individual people? What form would this knowledge be likely to take? How could this potentially differ from other cases of ordinary knowledge of facts about people that take place between human agents, and would these differences matter for an ethical objection based on privacy violation concerns? I do not profess any insight into how an artificial agent might gain consciousness or what exact criteria such an agent would have to meet in order to count as conscious–rather, for the sake of this question, I will assume a hypothetical case in which we can reasonably agree such an artificially conscious agent *does* exist and then examine what new concerns for privacy flow from this possibility.