An Approach to Categorize the Ethical Considerations of Modern Algorithms

Joseph B. Cofer, III

Research Mentor: Marshall Bierson, Department of Philosophy

Abstract

A classificatory schema with which to assess the morally salient features of algorithms will be presented. To assess this schema, this project will show how such classifications can capture the moral features of algorithms used by the military as well as social media companies.

By proving how the classificatory schema can provide insight into these different fields, it helps establish the broad applicability of the schema, and allows us to triangulate on the more general ethical questions raised by algorithms (rather than allowing ourselves to get lost in details of a particular sector).

Algorithms integrated into the military tend to focus on more tangible results which may be expressed through the implementation of unmanned technology. The primary focus of social media algorithms, however, is to retain engagement and earn profit.

The intersection of said algorithms can be found in the desire to gain data for external leverage upon others.

The United States military has an incentive to acquire data on individuals who may be a threat, and plan accordingly for battle. Social media organizations utilize algorithms that incorporate users' engagement and interests to leverage their data to advertising agencies, thereby increasing profits through the providing of more bespoke advertisements.

Through the comparison of seemingly contrasting sectors, we seek to analyze how how different forms of algorithms may be utilized to achieve similar results. Via an analysis of different strains of algorithms across these two fields, we underpin the intentions of algorithms through overarching themes of leverage and access to information.

Classificatory Schema Illustration

- The overarching purposes of algorithms in social media differs from United States military as algorithms in social media are implemented to encourage profit via an increase in user engagement, whereas the purpose of algorithms in the United States military is largely to ensure the strategic advantage in relation to national security and interests.

- Algorithms in both sectors have a variety of additional intentions and interactions as well.
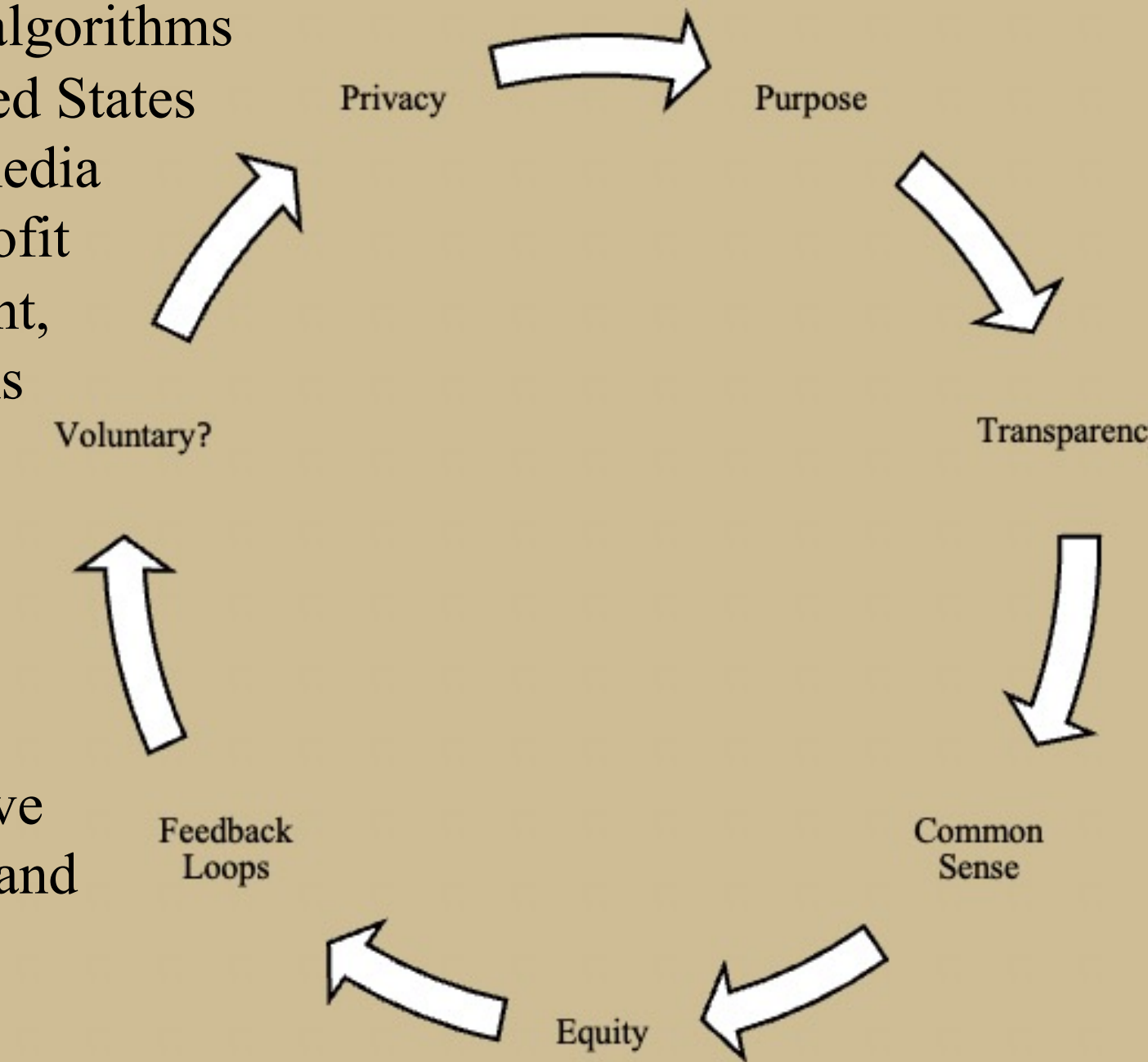
Classificatory Considerations

Intentions behind the implementation of algorithms differ between 'social media' and the 'United States military,' yet they undergo similar classifiers, as introduced below:

Beneficiaries: Which actors are being influenced by the implementation of algorithms?

Purpose: What is the overarching intention of the algorithm?

Transparency: Are the algorithms operations and intentions transparent and made aware to the defined actors?

Common Sense: Does the algorithm posses an element of common sense? (goal of numerous organizations via technology such as *deep learning*)

Quality/Equity: Are there any biases in the information being portrayed to the defined actors?

Feedback Loop: Is the information being portrayed to the defined actors consistently reshown due to a lack of discerning new, more accurate information?

Voluntary: Are the defined actors voluntary engaging in such use of algorithms?

Privacy: Are the defined actors aware/in control of their data and can limit information exposure?

Methodology

The investigation of available documents will assist in providing context in understanding how ethical implications of algorithms present across varying sectors. The fields of 'social media' and the 'United States military' were sampled to encourage a consideration of similar algorithmic methodology in seemingly diverse sectors. It becomes essential to understand the purpose behind implementing algorithms and their fluid nature for considerations of future ethical implications to arise. Several considerations arose as classifiers in developing such a schema.

The US Military Illustration Example

The United States Military:

- *Transparency:* "Traceable" (NATO). The Department of Defense (DoD) will continue a dialogue with the general public and human rights organizations to ensure constructive transparent process (Hartif & VanHoose). Can be difficult, especially as transparency may contradict ambitions of national security.

- *Common Sense:* Current lack of common sense, however the DoD has implemented the 'Machine Common Sense' program to accelerate research in this field, focusing on America's 'Multi-modal Open World Grounded Learning and Interface,' providing the United States with greater access to more in-depth analytics of all regions of the world (Mayanak).

- *Equity:* The DoD has taken steps to minimize biases in the implementation of algorithms (Mayanak). On the contrary, China's facial recognition software profiles the minority Uighur group leading to 'automated racism' (Mozur).

- *Feedback Loops:* "Governable" (NATO). The DoD intends to design algorithms with the ability to deactivate systems, ensuring human control (Mayanak). Without human authority, it can be assumed that catastrophic events could happen instantaneously.

- *Voluntary:* The idea of warfare where a foreign nation brings this conflict upon themselves persists. From a domestic perspective, China has constructed a modern surveillance state designed to wield authority over its people via text messages, internet searches, etc. (Mozur), as can be proven in the pro-Democracy Hong Kong protests and Muslim/Uighur genocide.

- *Privacy:* Similar sentiments as 'transparency.' Privacy has proven to be a false hope via China's enforcement upon their population, as stated above, and as there was an American effort to collect Americans' data for analyzation with data techniques (Mayanak).

Social Media Illustration Explanation

Social Media:

- *Transparency:* Algorithms fluctuate whether they are *supervised* machine learning (more transparency) or *unsupervised machine learning:* leading to more convoluted algorithms such as Deep Neural Networks. Algorithms are kept a secret as are a profitable asset.

- *Common Sense:* Currently social media algorithms do not possess this ability (ScienceDaily), rather algorithms create an output via training from determined inputs. If this ability were possessed, concerns regarding the human control of algorithms would likely arise.

- *Equity:* Due to a lack in common sense, algorithms simply reinforce our existing biases (Knowledge). However, algorithms also learn from past experience, thus relying on aged societal issues such as various forms of discrimination (Brown). Thus, when algorithms experience diverse content, they may typically suppress the content (Brown), leading to a technique dubbed 'shadowbanning.'

- *Feedback Loops:* Due to a lack in common sense and determining of equity within algorithms, similar recommendations are often shown, thus reinforcing already populous content (Swathi) (trends).

- *Voluntary:* While the user legally accepts the terms and conditions, a study from Harvard University found that engaging in social media triggers same portion of brain that ignites when consuming an addictive substance, and that "five to ten percent of Americans meet the criteria for social media addiction today" (Social Media Addiction).

- *Privacy:* Social media algorithms actively track a user's location from where messages are sent (Thompson) while social media companies profit from advertisers' willingness to pay for targeted data, such as age, sex, relationships, interests, etc. (Privacy).

Ethical Understandings of Algorithms

- Algorithms have evolved to become integrated aspects of our lives, stressing the importance of understanding their ethical implications in a diversity of scenarios. The sectors of 'social media' and the 'United States military' exist in distinct portions of society – typically unassociated with one another - where overarching ambitions reflect said separation in their differences, as introduced below:

- *Social media:* Algorithms exist within social media platforms for various intentions however they rely on the theme of enhancing profits from tailored advertisements through the continuous engagement of users (Joy). It may be determined that attention has become a commodity as advertisers can apply learned data pertaining to the user's interests with an intention of predicting behaviors and influencing such behavior for a monetary return (University). This is proven in Facebook's (the largest social media company in the world) revenue: in the third quarter of 2021 Facebook's total revenue summed to be $29,010,000,000, 97.5% of which came from advertising alone (Facebook). While this behavior persists, ethical implications of such algorithms on users should be considered.

- *The United States military:* While the overarching ambition of the military is to advance domestic ambitions, algorithms within this context serve a variety of distinct functions such as determining information of tangible objects (NATO) to having an advantage over our adversaries via the development of artificial intelligence that will set the tone of the United States and expectations of her allies (Johnson). The latter seems to be more provoking in rapid militaristic algorithmic evolution, particularly in relation to authoritarian powers- particularly China-proven as American heads of military understand the existence of the United States in a technological competition (Johnson). This ambition has been translated into the creation of the 'Defense Innovation Board,' which has developed five guiding principles for machine learning in the military: responsible, equitable, traceable, reliable, and governable (NATO), thus proving the dynamic relationship between algorithms and their ethical applications.

References

1. Brown, A. (2021, November 3). *Understanding the technical and societal relationship between Shadowbanning and algorithmic bias*. Forbes. Retrieved February 10, 2022, from https://www.forbes.com/sites/anniebrown/2021/10/27/understanding-the-technical-and-societal-relationship-between-shadowbanning-and-algorithmic-bias/
2. *Esper says artificial intelligence will change the battlefield*. U.S. Department of Defense. (n.d.). Retrieved February 10, 2022, from https://www.defense.gov/News/News-Stories/Article/2340972/esper-says-artificial-intelligence-will-change-the-battlefield/
3. *Facebook reports third quarter 2021 results*. Meta. (n.d.). Retrieved February 10, 2022, from https://investor.fb.com/investor-news/press-release-details/2021/Facebook-Reports-Third-Quarter-2021-Results/default.aspx
4. Hartig, L., & VanHoose, K. (2021, April 13). *Solving one of the hardest problems of military AI: Trust*. Defense One. Retrieved February 10, 2022, from https://www.defenseone.com/ideas/2019/04/solving-one-hardest-problems-military-ai-trust/155959/
5. Johnson, K. (2019, October 31). *Defense Innovation Board unveils AI Ethics principles for the Pentagon*. VentureBeat. Retrieved February 10, 2022, from https://venturebeat.com/2019/10/31/defense-innovation-board-unveils-ai-ethics-principles-for-the-pentagon/
6. Joy, A. (2021, February 18). *The attention economy: Where the customer becomes the product*. Business Today Online Journal. Retrieved February 10, 2022, from https://journal.businesstoday.org/bt-online/2021/the-attention-economy-asher-joy
7. *Knowledge is power: Fighting misinformation, disinformation, and junk news: Search engine and Social Media bias*. Research Guides. (n.d.). Retrieved February 10, 2022, from https://libguides.wustl.edu/c.php?g=955803&p=6897829
8. Mayank Kejriwal Research Assistant Professor of Industrial & Systems Engineering. (2021, October 14). *An AI expert explains why it's hard to give computers something you take for granted: Common sense*. The Conversation. Retrieved February 10, 2022, from https://theconversation.com/an-ai-expert-explains-why-its-hard-to-give-computers-something-you-take-for-granted-common-sense-165600
9. Mozur, P. (2019, April 14). *One month, 500,000 face scans: How China is using A.I. to profile a minority*. The New York Times. Retrieved February 10, 2022, from https://www.nytimes.com/2019/04/14/technology/china-surveillance-artificial-intelligence-racial-profiling.html
10. NATO. (n.d.). *Possibilities and challenges for artificial intelligence ...* Retrieved February 10, 2022, from https://www.researchgate.net/profile/Johan------Schubert/publication/326774966_Possibilities_and_Challenges_for_Artificial_Intelligence_in_Military_Applications/links/ 5b62d8140f7e9bc79a75979c/Possibilities-and-Challenges-for-Artificial-Intelligence-in-Military-Applications.pdf
11. *Privacy violations using microtargeted ads: A case study*. IEEE Xplore. (n.d.). Retrieved February 10, 2022, from https://ieeexplore.ieee.org/abstract/document/5693335
12. ScienceDaily. (2020, November 18). *New test reveals AI still lacks common sense*. ScienceDaily. Retrieved February 10, 2022, from https://www.sciencedaily.com/releases/2020/11/201118141702.htm
13. *Social Media Addiction*. Addiction Center. (2021, December 17). Retrieved February 10, 2022, from https://www.addictioncenter.com/drugs/social-media-addiction/
14. Swathi Meenakshi Sadagopan Munk Fellow. (2020, August 10). *Feedback loops and echo chambers: How algorithms amplify viewpoints*. The Conversation. Retrieved February 10, 2022, from https://theconversation.com/feedback-loops-and-echo-chambers-how-algorithms-amplify-viewpoints-107935
15. Thompson, C. (2015, May 28). *Social media apps are tracking your location in shocking detail*. Business Insider. Retrieved February 10, 2022, from https://www.businessinsider.com/three-ways-social-media-is-tracking-you-2015-5
16. University, E. F.-V. L., Fosch-Villaronga, E., University, L., University, A. P. C. S., Poulsen, A., University, C. S., Roger A. Søraa Norwegian University of Science and Technology (NTNU), Søraa, R. A., Norwegian University of Science and Technology (NTNU), University, B. C. L., Custers, B., University, A. S., Google, Tacoma, U. of W., University, S., & Metrics, O. M. V. A. (2021, June 1). *Gendering algorithms in social media*. ACM SIGKDD Explorations Newsletter. Retrieved February 10, 2022, from https://dl.acm.org/doi/abs/10.1145/3468507.3468512