

The Agency of Machines

Probing the history of technology with computational linguistics

Even though machines are essentially artefacts, they often acquire anthropomorphic qualities in discourse: machines are staged as autonomous agents “capable of determining the course of events”, a discursive phenomenon the historian Leo Marx described as “hazardous” (Marx, 1997). In sentences like “Machines change the way we live”, the artefact figures as the agent of an active verb; “when used in this way” Marx argues, “[...] technology becomes hazardous to moral and political cogency of our thought.” More generally, viewing machines as active agents risks committing (deliberately or not) to a form of Technological Determinism in which machines rather than humans drive history. The abstract aims to explore aspects of historical discourse on technology by looking at some ways in which machines were portrayed as agents (obfuscating, or simply trumping, human agency) during the Industrial Revolution.

This research is a first attempt to answer the larger questions at the heart of the “Living with Machines” project¹, which aims to explore the effects of mechanisation on society, focussing on the lived experience of the Industrial Revolution. By scrutinizing *when* machines emerged as (discursive) agents, *who* propagated such depictions in *which* sources, and with *what* effect, we explicate the (often subtle) linguistic patterns, which, taken together, constitute the discursive regimes that governed how the relationship between mechanical technology and society was imagined by contemporaries.

Our sources comprise a vast collection of digitized books² and newspapers³. From a methodological point of view, we demonstrate how computational methods can be meaningfully leveraged to answer historical questions. As the concept of technology only emerged at the turn of the twentieth century, we start by manually defining a core lexicon pertaining to machines—one that captures the content and scope of the concept, formed by lexical terms from several groups (such as machines, machine parts, materials, processes, ideas, and social movements). These terms serve as “seeds” to generate a more comprehensive lexicon using data-driven techniques. To expand the lexicon algorithmically, we train language models that generate vector representations of words (such as Word2Vec (Mikolov, et al, 2013) and GloVe (Pennington et al. 2014)) and induce a lexicon from a given seed vocabulary. Several options are available here: changing the objective function, or relying on existing algorithms such as SentProp (which runs random walks over graphs starting from seed words (Hamilton et al., 2016)) and SemAxis (which uses the distance from averaged vectors (An et al., 2018)). We stress that establishing a lexicon that holistically captures the mechanical aspects of “technology” remains an iterative process, i.e. it emerges in a constant dialogue between the domain experts and the language models.

¹ <http://livingwithmachines.org>

² <https://data.bl.uk/digbks/> [Do these need expansion, with a quick description?]

³ <https://britishnewspaperarchive.co.uk>

We proceed with creating a subsample of our corpora composed of sentences that contain machine-related terms, and which is representative across the different metadata fields that accompany our data. We process each sentence with semantic role labelling (SRL) to determine the semantic roles of the different predicate arguments in the sentence.⁴ We use a state-of-the-art neural model (He et al. 2017) and complement its output with connotation frames (similar to Sap et al. 2017) to measure levels of agency and power.

Ultimately, by inspecting the semantic roles in which machines appeared, we can trace changing attitudes toward technology over time and space. Linguistically, analysing these depictions of the relationship between humans and machines helps us to understand the lived experience of the Industrial Revolution.

References

- An, Jisun, Haewoon Kwak, and Yong-Yeol Ahn. "SemAxis: A Lightweight Framework to Characterize Domain-Specific Word Semantics Beyond Sentiment." arXiv preprint arXiv:1806.05521 (2018).
- Hamilton, William L., Kevin Clark, Jure Leskovec, and Dan Jurafsky. "Inducing domain-specific sentiment lexicons from unlabeled corpora." In Proceedings of the Conference on Empirical Methods in Natural Language Processing. Conference on Empirical Methods in Natural Language Processing, vol. 2016, p. 595. NIH Public Access, 2016.
- He, Luheng, Kenton Lee, Mike Lewis, and Luke Zettlemoyer. "Deep semantic role labeling: What works and what's next." In Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pp. 473-483. 2017.
- Jurafsky, Dan, and James H. Martin. Speech and language processing, 3rd ed. 2019, available on: <https://web.stanford.edu/~jurafsky/slp3/>
- Marx, Leo. " "Technology": The Emergence of a Hazardous Concept." Social Research (1997): 965-988.
- Mikolov, Tomas, Ilya Sutskever, Kai Chen, Greg S. Corrado, and Jeff Dean. "Distributed representations of words and phrases and their compositionality." In Advances in neural information processing systems, pp. 3111-3119. 2013.

⁴ Given the sentence: "The thrashing machine has changed the employment of some of the stoutest of the peasantry, but it has directly brought into employment the labour of women and children, and thus has increased the earnings of larger families.", the agent is 'the thrashing machine', while the patients whose states undergo the change or disruption are 'the employment of some of the stoutest of the peasantry', 'the labour of women and children', and 'the earnings of larger families'.

Pennington, Jeffrey, Richard Socher, and Christopher Manning. "Glove: Global vectors for word representation." In Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP), pp. 1532-1543. 2014.

Sap, Maarten, Marcella Cindy Prasettio, Ari Holtzman, Hannah Rashkin, and Yejin Choi. "Connotation frames of power and agency in modern films." In Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, pp. 2329-2334. 2017.

Schatzberg, Eric. Technology: Critical History of a Concept. Chicago: University of Chicago, 2019.