

Contracting for Transparency: Artificial Intelligence and the Need for a Contractual Commons

Erik Stallman and Sonia Katyal

Bias in data used to build and train algorithms and machine learning models (referred to collectively as “automated decision systems”) has been identified as a central source of bias and unreliability in automated decisionmaking. As the saying goes, garbage in, garbage out. However, often there is no way to tell where the garbage came from. Calls for greater dataset transparency run into a typical legal obstacle: many datasets are proprietary, making it impossible or unlawful to audit or investigate without permission. Even when automated decision systems rely on publicly accessible data, legal and contractual uncertainty around the use of datasets, fueled in part by mixed case law, may inhibit disclosure of critical information that enables third-party investigation of those systems. This represents a significant obstacle to algorithmic transparency, but as we argue, these circumstances also represent a potentially fruitful area of transactional possibility. By taking a closer look at the contractual terms by which datasets are currently licensed, and their limits and possibilities, we can see how to improve both access and transparency. As we argue, existing contractual systems may reveal overlooked pathways toward disclosures that can overcome potential legal obstacles to certain investigative techniques, enabling meaningful third-party investigation of automated decision systems.

Our goals in this paper are both descriptive and normative. In Part I, we analyze some of the legal and practical obstacles to transparency in the use of datasets. Here, we draw a crucial distinction between announced commitments to transparency and actualized open access. As we show, despite proposals for accountability principles or model disclosures, significant barriers to auditing and independent verification exist, often due to contractual and proprietary restrictions. In the public context, despite increasingly sophisticated due process challenges and Freedom of Information Act complaints to compel disclosure, many arenas of automated decisionmaking remain largely out of reach of investigation. In the private context, contractual terms often further complicate matters, making it almost impossible to investigate or audit as a result. As we explain, even contractual silence creates a risk of liability that may dissuade the engineers behind automated decision systems from disclosing information about the data used to build them.

In Part II, developing our inquiry, we survey existing licensing terms attached to publicly and privately available datasets. As we show in our typology, there is considerable variation in the specific licensing terms or models used, producing a range of outcomes regarding the conditions placed on the use of data made available. At the same time, however, our study finds that a significant percentage of publicly available datasets include a term requiring attribution and a smaller percentage include terms imposing open-access requirements on downstream users. While transparency may not be the primary motivation for these terms, they do suggest that a norm of open access conditioned on disclosure can be productively generated in both the private and public arenas. Drawing from these examples, then, we outline several ways in which these

licensing terms (subject, of course, to limitations due to privacy and other considerations), can encourage both private and public entities to condition access to the data they generate on similar terms to enable independent research and verification.

In Part III, developing from the previous sections, we propose specific elements for model license terms that can be focused on ensuring transparency in automated decisionmaking. For example, merely disclosing that a particular dataset was used to train or build an automated decision system is likely insufficient to enable independent analysis or verification of that system. Instead, further information about other sources of data, how they are combined, whether labels are used and how variables are weighted or excluded may also be needed. A license also may need to anticipate and answer potential legal risks and uncertainties associated with third-party auditing of AI systems. In our model, we seek to address the need for some of these specifications by developing a set of terms that can resolve some of these concerns.

In Part IV, we turn to the limitations of our proposed model, and discuss practical, legal, and philosophical obstacles on licensing as a solution for transparency. As we recognize, owners of automated decision systems may balk at license terms that require disclosure of proprietary information. The fear, of course, is that such disclosures could cede a competitive advantage to others or facilitate “gaming” of automated decisionmaking or even adversarial attacks. And the owners of proprietary datasets, themselves, would have their own reasons for opposing transparency conditions, which may lead to conflicting obligations attached to different datasets used by the same AI system. And even outside of the decisionmaker and the dataset community, open access advocates might further resist a transparency mechanism that depends on recognizing proprietary and exclusive rights in datasets on philosophical grounds altogether. We look to other licensing regimes that have addressed the same difficulty by specifically accommodating limitations on exclusive rights, and we conclude that a transparency term can have a significant normative impact even when the right to enforce it is unclear or contested. Ultimately, however, we argue that much like any other form of regulation or litigation, licensing is unlikely to provide a comprehensive solution to the issue of opacity in automated decision systems. However, with a more deliberate focus on disclosures that enable third-party auditing, licensing can both help to establish a norm of transparency and provide a much-needed gap-filling function when regulation and litigation fall short.