Governing Algorithmic Systems with Impact Assessments: Six Observations

Elizabeth Anne Watkins, ewatkins@datasociety.net | Emanuel Moss, emanuel@datasociety.net
Jacob Metcalf, jake.metcalf@datasociety.net | Ranjit Singh, ranjit@datasociety.net | Madeleine Clare Elish, mcelish@google.com

DATA&
SOCIETY

What Constitutes an Impact is Non-Obvious: Defining what counts as impact of an Algorithmic Decisions Systems (ADSs) is shaped by social, economic, and political power, and is an ethical choice, in which one set of concerns, events, or communities is prioritized over another.

Different Types of Impact Comes into Focus Depending on When an Assessment Occurs: *Ex ante* assessments are based on existing information like prior use cases, measurements of system behavior in testing environments, or narrative records of how the system was designed and developed. *Ex post* assessments determine causality of the consequences of an already deployed system or speculating outcomes in an imagined world where the system was not deployed.

Impact Assessments Structure Interactions in and between Institutions: Impact Assessments (IAs) bring different organizations into relationships with each other. The exercise of institutional power through IAs can lead to questions about: who exerts power over whom, checks and balances on the exercise of this power, what are the sources of resistance, and how abuses of power can be ameliorated.

Public Participation in an Assessment Process is Not Synonymous with Accountability to the Public: Different types of IA mobilize different kinds of representation and participation from the public through comment periods, focus groups, rapid assessments, or open meetings. Encouraging public participation without critical reflection on how they may also produce exploitative and extractive forms of involvement can jeopardize public trust.

Assessing Impacts Does Not Necessarily Mean Addressing Harms: Impacts are evaluative constructs—i.e., imperfect proxies—for harms. Any rigorous IA should be able to facilitate robust, engaged, and transparent decision-making around what the tradeoffs are between potential harms and likely benefits.

Impact Assessments Ask Us How the World Might Be Otherwise: IAs draw attention to design choices and consequences, and create room for considering alternatives in development cycles. IAs can provide a lever of influence for communities to highlight overlooked or unforeseen sets of causes and effects.

These six observations resonate with the central ethical consequence of machine learning: *the models at the core of an ADS are proposals for how the world ought to be.* **Who gets to enter this space, whose interests are formally considered, which expertise is treated as relevant, and when the counterfactuals are considered are all sites of contestation and social power.**