



Government Use of AI: Costs and Benefits of Transparency

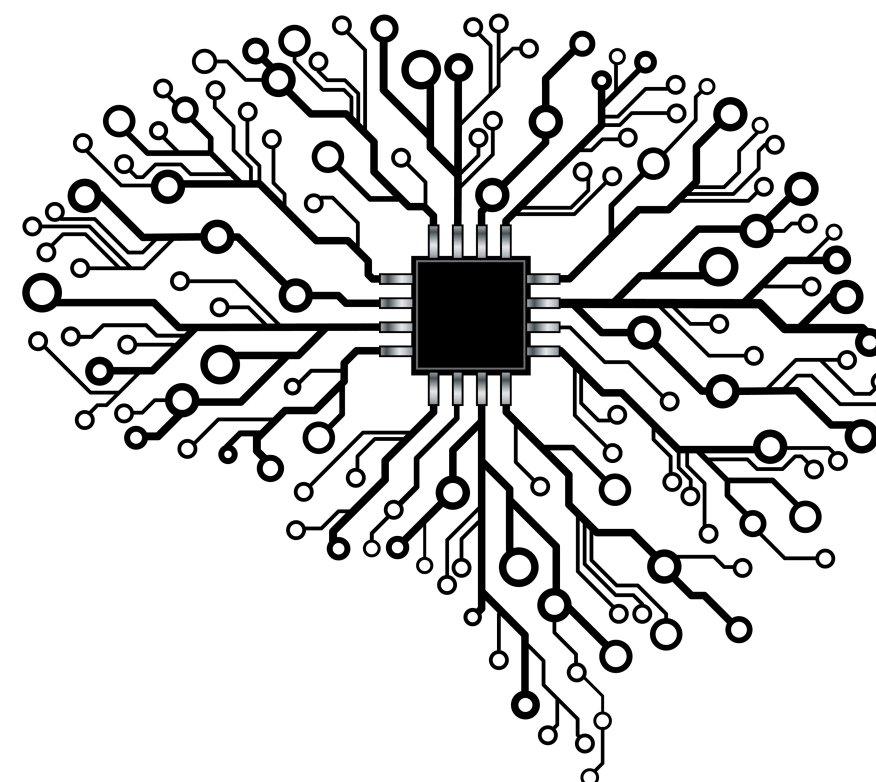
Peter Story, Carnegie Mellon University

Governments Are Using AI to Inform Their Decision-Making

AI is being adopted by all levels of government in the United States:

- COMPAS to inform sentencing decisions [1]
- Facial recognition for airport security
- License plate scanners to give parking tickets

When the government's decision-making is informed by AI, AI systems can have a large impact on our lives. **Despite this, transparency is often lacking.** Many AI systems are propriety, and so details about the systems are hidden from public scrutiny. Greater transparency is needed to combat **wasteful government spending, biased decision-making, and a lack of due process.**



Relevant Dimensions of Transparency

Transparency on each of these dimensions helps experts understand an AI system.

- **Open or closed evaluation methodology:** Different AI systems can be compared when they share a common prediction task, data for training, and data for testing.
- **Open or closed design:** Knowledge about an AI system's design allows experts to determine whether the system's operation is plausible and to identify potential weaknesses. Details about data preprocessing and model training are particularly important.
- **Open or closed access to the system:** With access to an AI system, experts can provide different inputs and evaluate the system's outputs. This facilitates checking for different types of bias.
- **Open or closed access to source code:** With access to a system's source code, experts can identify bugs and nefarious behavior, and estimate the cost of maintaining the system.
- **Degree of explainability:** For example, the output of systems based on decision trees and rule lists can be understood more easily than the output of black-box systems.

Potential Benefits of Transparency

Fiscal Responsibility

Transparency is necessary for the government to make informed decisions about which systems to acquire.

Equity

AI systems are susceptible to bias. Experts must check for inappropriate bias in AI systems.

Due Process

When AI systems inform government decisions against people, some explanation is owed to those affected [2].

Potential Costs of Transparency

Intellectual Property

To protect the intellectual property of the developers of AI systems, some forms of transparency may not be offered to the public.

Malicious Use of Systems

Knowing how a system makes decisions can facilitate malicious attacks. However, identifying weaknesses makes it possible to improve systems.

Lower Predictive Performance

In some cases, black-box systems (e.g., deep learning) can improve performance over more explainable systems (e.g., decision trees).

How to Balance Costs and Benefits?

Prior to adopting an AI system, the government should rely on the judgement of domain and technical experts to evaluate the systems under consideration. These experts can operate under NDAs. They should:

1. Describe the task to be performed, establish a common evaluation methodology, and train simple models to establish a performance baseline. This information should be provided to AI systems developers.
2. Evaluate the systems submitted for consideration for inappropriate bias and other undesirable behaviors.
3. Evaluate the appropriateness of each system's degree of explainability and consider the degree of transparency a system's developers are willing to grant to the public.
4. Draft a report of their evaluation and potentially recommend a system for adoption.

References

- [1] Ed Yong. A Popular Algorithm Is No Better at Predicting Crimes Than Random People. The Atlantic, 2018.
- [2] Henry J Friendly. Some Kind of Hearing. University of Pennsylvania Law Review, 1974.