

# THE DANGERS OF BLACKBOXED AI IN COURT

Andrew D. Selbst, UCLA School of Law

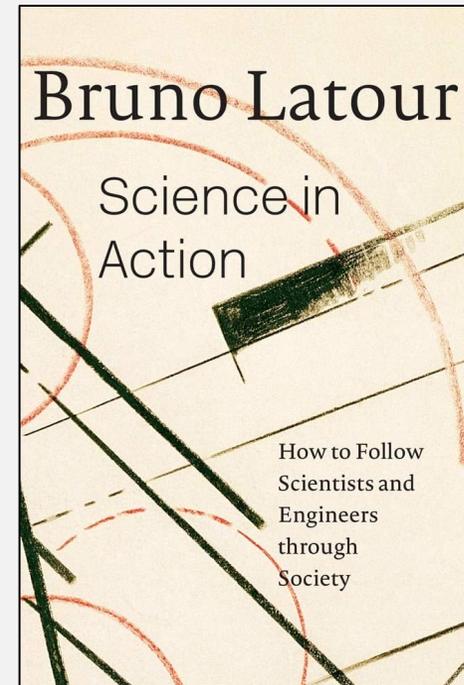
AI and Civil Justice Conference, Berkeley Law  
April 12, 2024

Black Boxes



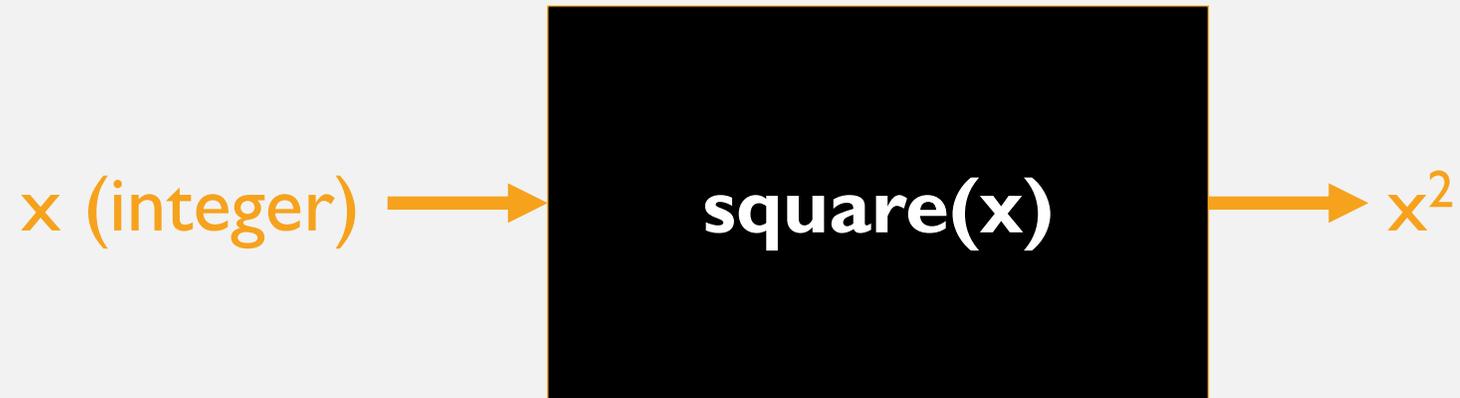
Data + Algorithm = ?

Blackboxing

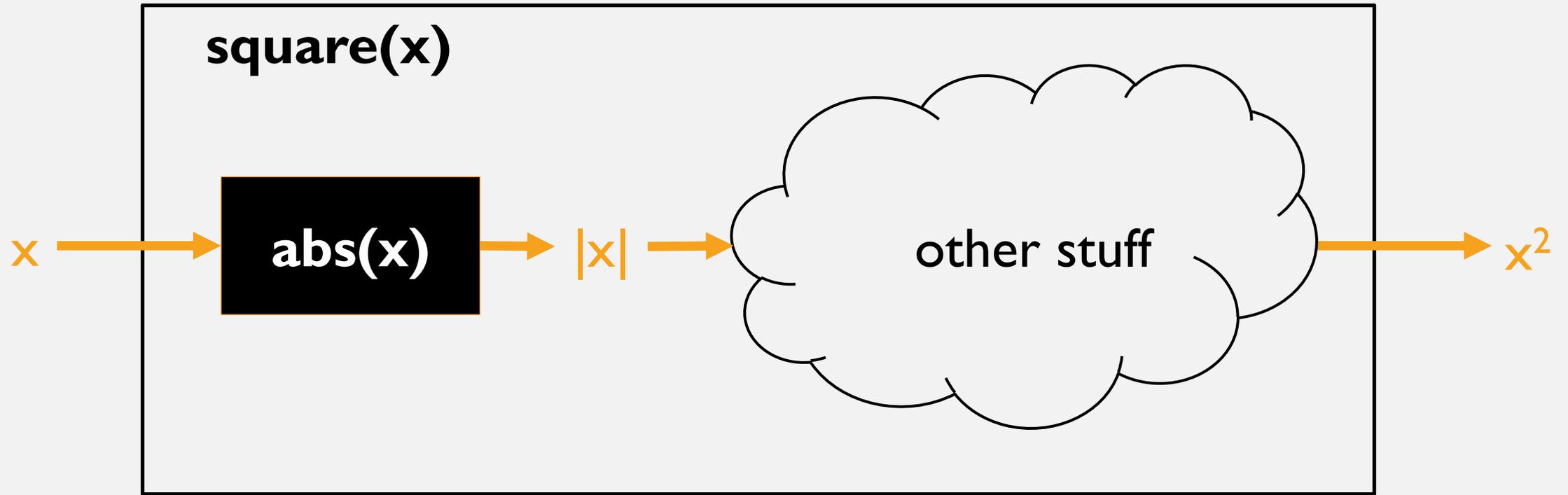


Courts must resurface and critique  
technical design choices in order to  
do their work correctly

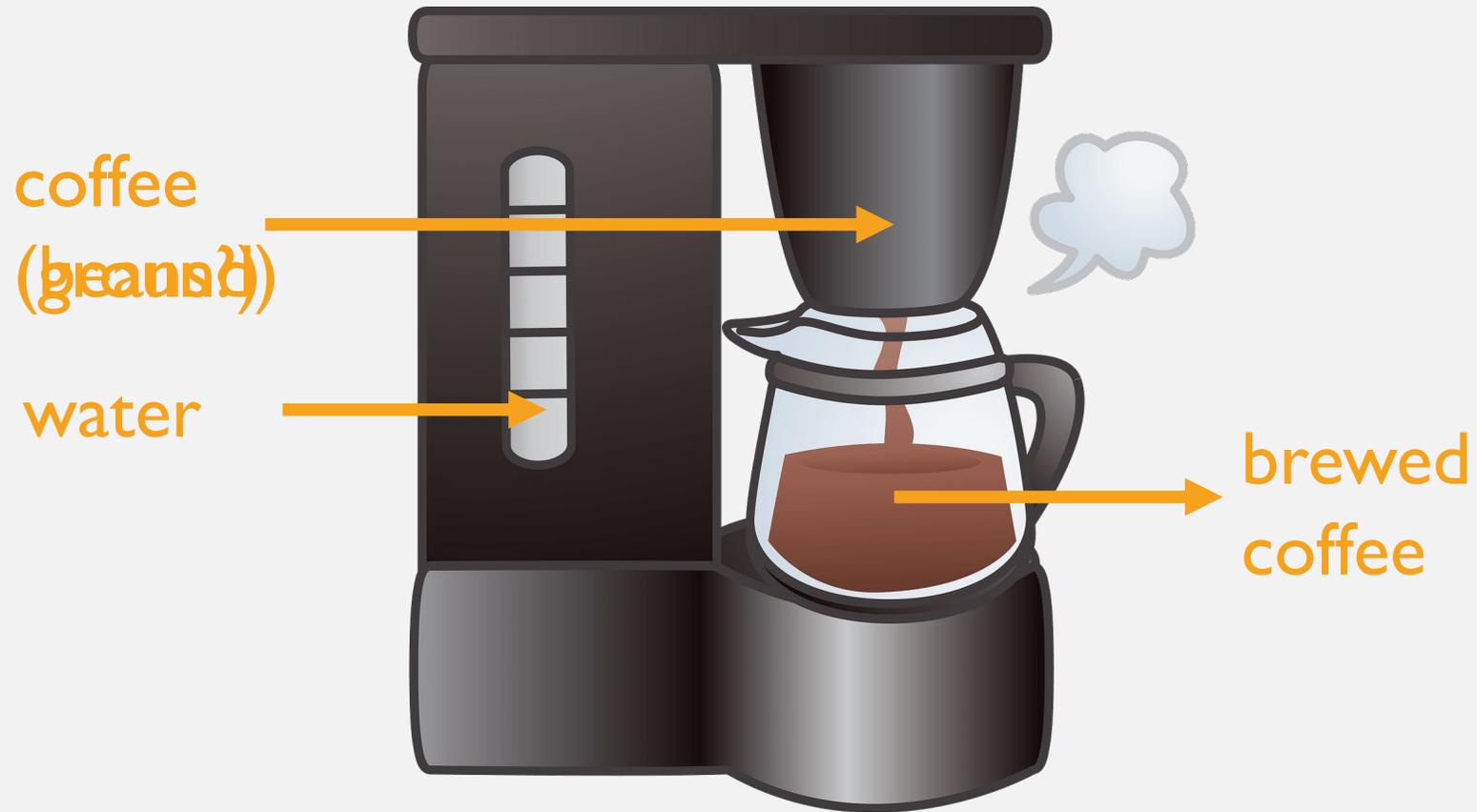
# ABSTRACTION AND ITS CONSEQUENCES



```
1) y = x*x;  
   output y  
  
2) y = 0;  
   for i = 1 to abs(x):  
       y = y + abs(x);  
   output y
```



```
define square(x){  
  y = 0;  
  for i = 1 to abs(x):  
    y = y + abs(x);  
  output y}
```



coffee grinder  
(separate)



# BLACK BOXES AND BLACKBOXING

- Abstraction boundaries define the object
- Abstraction boundaries delineate responsibility
- After abstraction choices are made, they become invisible

# WHY COURTS SHOULD CARE

## ENGINEERING VALUES

- Efficiency
- Portability
- Profit
- Context

## LEGAL VALUES

- Fairness/Due Process
- Justice
- Accountability
- Efficiency

A surprising amount of AI on the  
market is broken.

**Table 1: Failure Taxonomy**

|                          |  |
|--------------------------|--|
| Impossible Tasks         | Conceptually Impossible<br>Practically Impossible                                    |
| Engineering Failures     | Design Failures<br>Implementation Failures<br>Missing Safety Features                |
| Post-Deployment Failures | Robustness Issues<br>Failure under Adversarial Attacks<br>Unanticipated Interactions |
| Communication Failures   | Falsified or Overstated Capabilities<br>Misrepresented Capabilities                  |

COURTS DO DECONSTRUCT  
DESIGN...SOMETIMES

# A MATTER OF SALIENCE

- Some types of cases where court already do interrogate technical design:
  - Products liability
  - Copyright retransmission cases
  - Copyright and software

## PRODUCTS LIABILITY: RIDER V. TOWNSHIP OF FREEHOLD

- Plaintiff died when utility pole “intruded into the passenger compartment”
- P’s expert provided evidence of “technically feasible, practical and safer alternative” to the design of the car
- Court discusses tech arrangements in high level of detail
  - “The front and rear subframes ... were not connected to one another, creat[ing] a structural gap that allowed penetration”
  - “The floor pan ... offered no real ‘structural resistance’”
  - ...

## COPYRIGHT RETRANSMISSION CASES

- *Cablevision* (2d. Cir. 2008): Remote Service DVR not a copyright problem because it was 1) a private performance & 2) initiated by user
- Aereo, BarryDriller, FilmOn built technology designed around this holding
  - Courts couldn't understand that or rule on it without unpacking the design
- Cases became battles of abstractions: Where did Aereo's "device" end for the purpose of the law?
- Supreme Court holding implies that courts need not consider the tech
  - Unfortunate result, but lower courts showed that it's entirely possible to do so

## LESSONS FROM EXISTING CASES

- Courts already examine technological design where the question appears salient
- Courts can and regularly do make calls as to the correct levels of abstraction/generalality
- Courts can use typical evidentiary techniques to unearth design

# HOW DECONSTRUCTION HELPS IN NON-TECHNOLOGICAL CASES

# HYPOTHETICAL: DISCRIMINATORY HIRING

- Employer buys off the shelf “fair-ML” software to predict the best candidates
  - Trained on data with demographics different than the employer’s.
  - Difference causes the fairness criterion to be invalid, and software to be discriminatory.
- Abstraction: Machine inputs employment data, business objectives, outputs a version of “fair” ranking that accounts for inequities.
  - Implicit claim: User is not responsible for the internals or to know about them
  - Court’s challenge: Does this satisfy business necessity/alternative employment practice?
    - Easiest way to get a less discriminatory system is to train properly. But without technical expertise, can employers ever be liable?
    - Hard to ever hold employer liable without requiring them to know something about the internals, breaking the abstraction.

# CONCLUSIONS

- Courts need to interrogate technical design whenever technology is involved in the case
- Courts may be bad at this because they are not technical experts, but:
  - They cannot avoid it
  - They already do this (with admittedly mixed success) in cases where it is more obvious that they need to
- Courts can supplement their legal expertise with expert testimony