ALGORITHM
WATCH

Holding algorithms accountable to protect fundamental rights

Internet Governance Forum 2016, Guadalajara, Mexico
December 7, 2016

Matthias Spielkamp | ms@algorithmwatch.org
http://algorithmwatch.org | Twitter: @algorithmwatch

Opportunities in the use of ADM

- democratise personal assistance
- increase efficiency (saving resources)
- increase safety
- more fair decisions / judgements
- less discrimination
- etc.

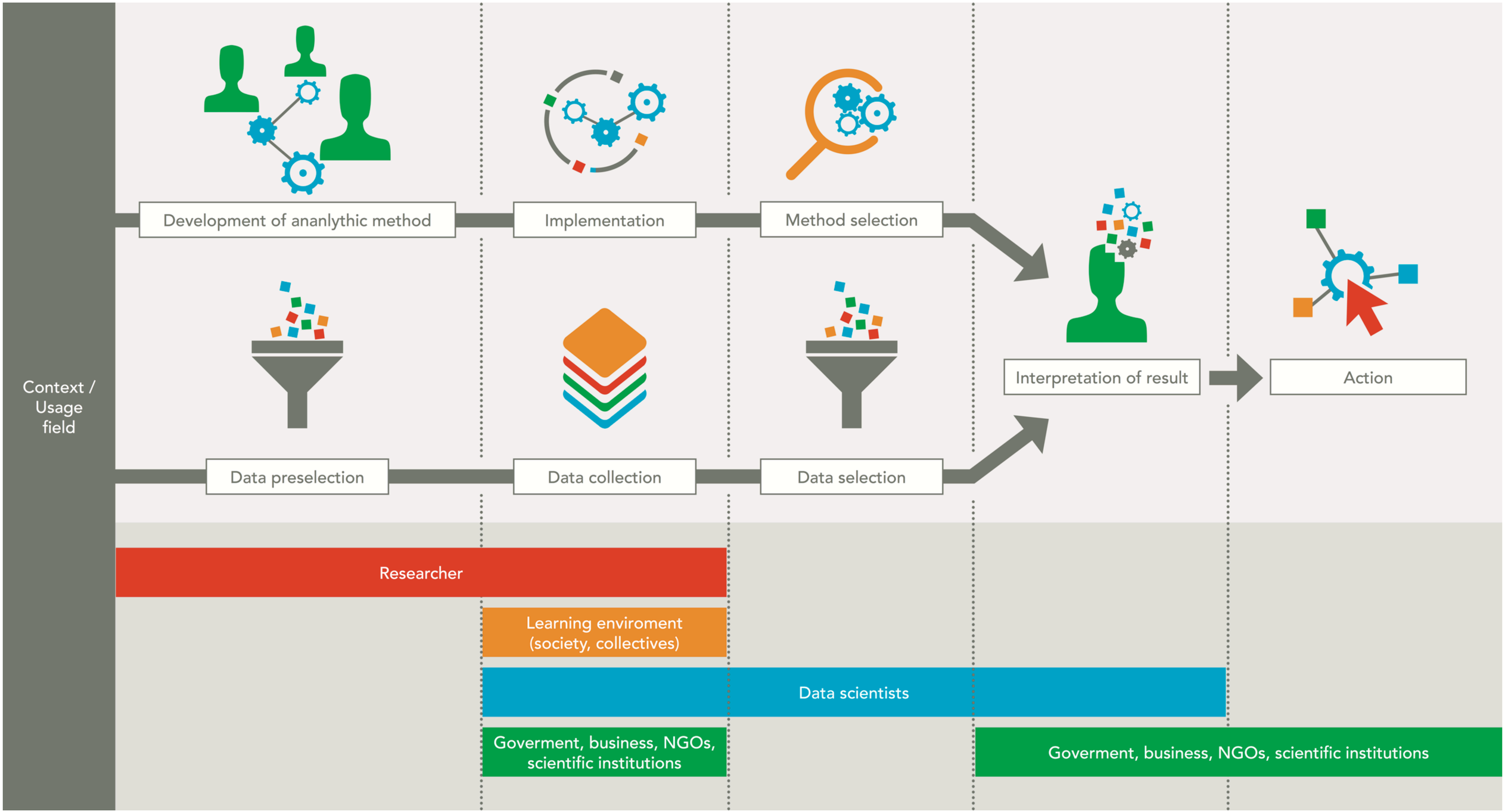
Risks in the use of ADM

- discrimination
- undesirable manipulation of individuals or collectives
- with direct impact (i.e. news consumption)
- with impact on third parties (i.e. sanction mechanisms)

What is ADM?

We call the following process algorithmic decision making (ADM):

- design procedures to gather data,
- gather data,
- design algorithms to
  - analyse the data,
  - interpret the results of this analysis based on a human-defined interpretation model,
  - and to act automatically based on the interpretation as determined in a human-defined decision making model.


Context / Usage field
Development of ananlythic method
Implementation
Method selection
Interpretation of result
Action
Data preselection
Data collection
Data selection
Researcher
Learning enviroment (society, collectives)
Data scientists
Goverment, business, NGOs, scientific institutions
Goverment, business, NGOs, scientific institutions

Issues

- What kind of scrutiny does ADM have to be submitted to?
- What objectives are meaningful, necessary and sufficient?
- Do we need to look for intelligibility, transparency, accountability?
- Can we expect any kind of control in light of self-learning systems?
- If not, what needs to be the result - a ban on ADM in cases when fundamental rights are affected?
- Would such a ban be enforceable?
- And last but not least: Who is responsible for the outcomes of ADM - the designers of the systems, the coders, the entities implementing them, the users?

Ways to investigate algorithms

- journalistic investigation and reporting
- computer science: statistical analysis, reverse engineering etc.
- law: legal analysis, strategic litigation etc.
- advocacy: campaigning
- philosophy: ethical analysis, development of criteria for evaluation, development of norms

Ways to investigate algorithms

New York City Department of Education

- value-added model (VAM) , used since 2007
- purpose: to rank about 15 percent of the teachers in the city. The model's intent is to control for individual students' previous performance or special education status and compute a score indicating a teacher's contribution to students' learning.

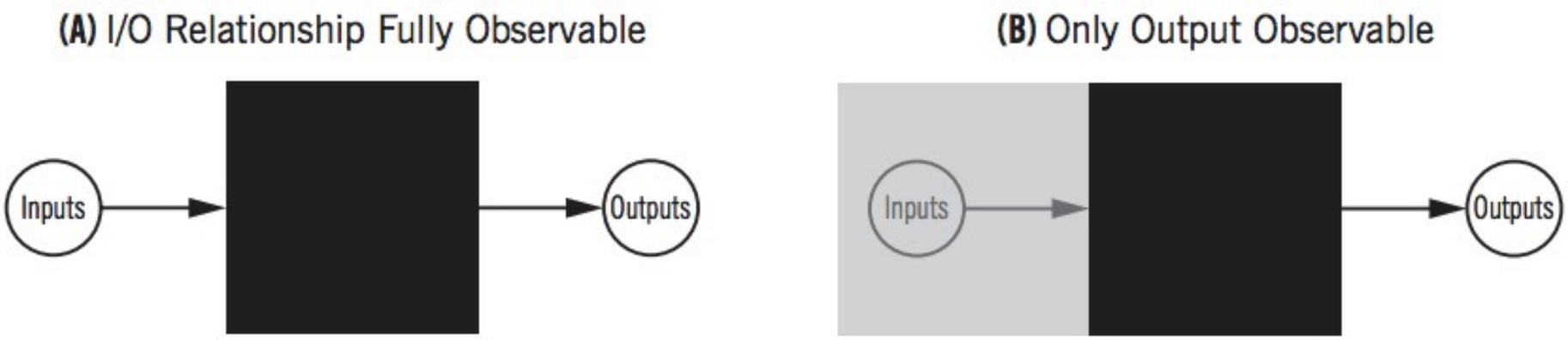
Ways to investigate algorithms

New York City Department of Education

- rankings and scores obtained through FOI request
- teacher's union: "the reports are deeply flawed, subjective measurements that were intended to be confidential."
- only correlation of 24 percent between any given teacher's scores across different pupils or classes
- suggests the output scores are very noisy and don't precisely isolate the contribution of the teacher

Ways to investigate algorithms

Reverse Engineering



Froom: Diakopoulos, Nicholas: Algorithmic Accountability Reporting: On the Investigation of Black Boxes; Tow Center for Digital Journalism, Columbia University New York, 2014, p. 14

FUTURE TENSE THE CITIZEN'S GUIDE TO THE FUTURE. DEC. 7 2015 11:48 AM

FROM SLATE, NEW AMERICA, AND ASU

Slate

Why Google Search Results Favor Democrats

It's not because the company is biased—it's more complicated.

575 23

By Daniel Trielli, Sean Mussenden, and Nicholas Diakopoulos

Why Algorithm Watch?

The ADM Manifesto

Algorithmic decision making (ADM) is a fact of life today; it will be a much bigger fact of life tomorrow. It carries enormous dangers; it holds enormous promise. The fact that most ADM procedures are black boxes to the people affected by them is not a law of nature. It must end.

Algorithm Watch Mission Statement

How do we work?

- monitor, observe and report
- explain
- network
- engage

ALGORITHM
WATCH

Thank you!
Questions?

Matthias Spielkamp | ms@algorithmwatch.org
http://algorithmwatch.org | Twitter: @algorithmwatch