



# ARTIFICIAL INTELLIGENCE, GOVERNMENT DECISIONS AND LEGITIMATE AUTHORITY

Doyen **Hugo Cyr**

Founding member of LegalIA and HumanIA ([www.humania.uqam.ca](http://www.humania.uqam.ca) )

**UQÀM**

Faculté de science politique  
et de droit

Université du Québec à Montréal

**A theoretical authority...**  
...gives reasons to believe something *is*.



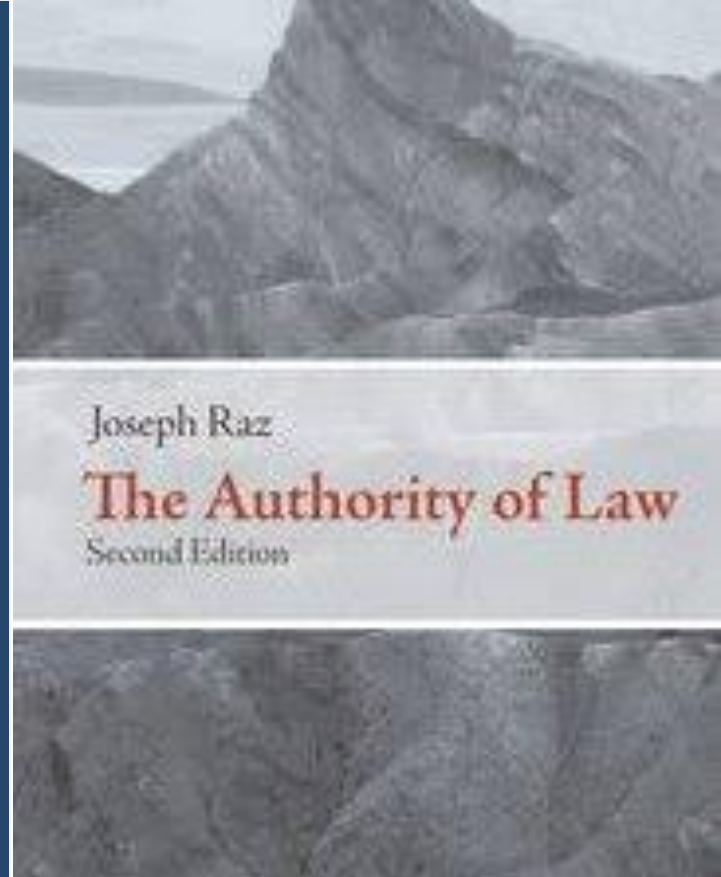
## A practical authority...

... gives reasons for actions, for what *ought to be done*



“Service conception of authority”

based on what Raz calls the “normal justification thesis” (NJT)



## Normal Justification Thesis

*One is likely better to comply with reasons which apply to him if he accepts the directives of the alleged authority as authoritatively binding and tries to follow them, rather than by trying to follow the reasons which apply to him directly.*

## Dependence Thesis

Authority's directives are meant to be *mainly* based on the balance of underlying reasons that are applicable to the subject

## Pre-emption Thesis

Once the authority's directive has been enacted, it is meant to displace the underlying balance of reasons and to exclude some or all other reasons for action.

**DIRECTIVES GIVEN BY AUTHORITY ARE  
THUS UNDERSTOOD AS...**

***USEFUL HEURISTIC DEVICES***

**Increases chances of  
successfully acting for the  
right reasons**

For example:

**Solving coordination  
problems**



- Saves **time** and **energy** in avoiding a full deliberation process





For the sake of the argument, let's take the NJT at its face value and explore how it plays out with algorithmic decision support tools...



# 1. PROBLEM WITH THE PRE-EMPTION THESIS

Lack of transparency of a decision based on an automated process

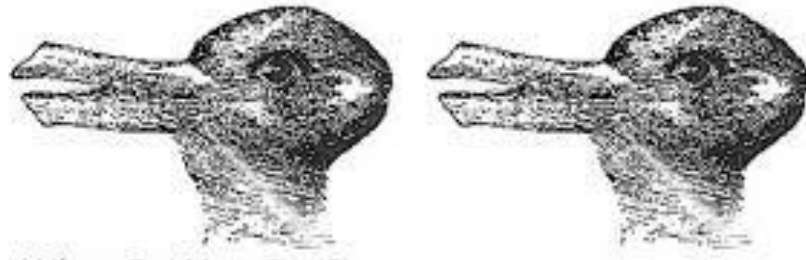
“Ah! But for an authority, one can go behind it and check for the relevant reasons”

# Procedural fairness may require **transparency** and **explainability** of algorithmically-generated directives

- Evidentiary issues
- Intelligibility issues

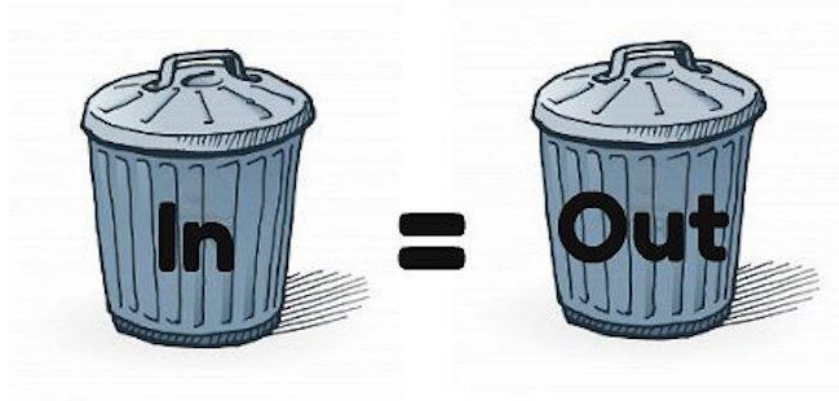
# 2. PROBLEMS WITH THE DEPENDANCE THESIS

## A) Legal characterization of the relevant facts...



*What Do You See?  
-Ducks, Rabbits or Both*

## B) Wrong rules are coded, or machine learning based on biased dataset



# What is the system trained to optimise?





MACHINE BIAS



# Bias in Criminal Risk Scores Is Mathematically Inevitable, Researchers Say

ProPublica's analysis of bias against black defendants in criminal risk scores has prompted research showing that the disparity can be addressed — if the algorithms focus on the fairness of outcomes.

by **Julia Angwin** and **Jeff Larson**, Dec. 30, 2016, 4:44 p.m. EST

## Quote from the article:

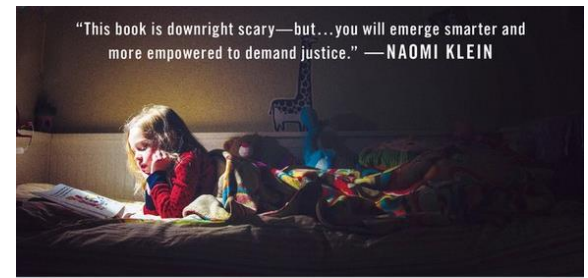
When he first heard about the COMPAS debate, Jon Kleinberg, a computer science professor at Cornell University, hoped he could figure out a way to reduce false findings while keeping the positive predictive value intact. “We thought, can we fix it?” he said.

But after he, his graduate student Manish Raghavan and Harvard economics professor Sendhil Mullainathan downloaded and crunched ProPublica’s data, they realized that the problem was not resolvable. A risk score, they found, could either be equally predictive or equally wrong for all races — but not both.



# Predictions based on missing data

- Ex.: Use of public assistance programs VS private insurances



## AUTOMATING INEQUALITY

HOW HIGH-TECH TOOLS PROFILE,  
POLICE, AND PUNISH THE POOR



AI & Legitimate Authority

**UQÀM**

Faculté de science politique  
et de droit

Université du Québec à Montréal

## C) Spurious correlations

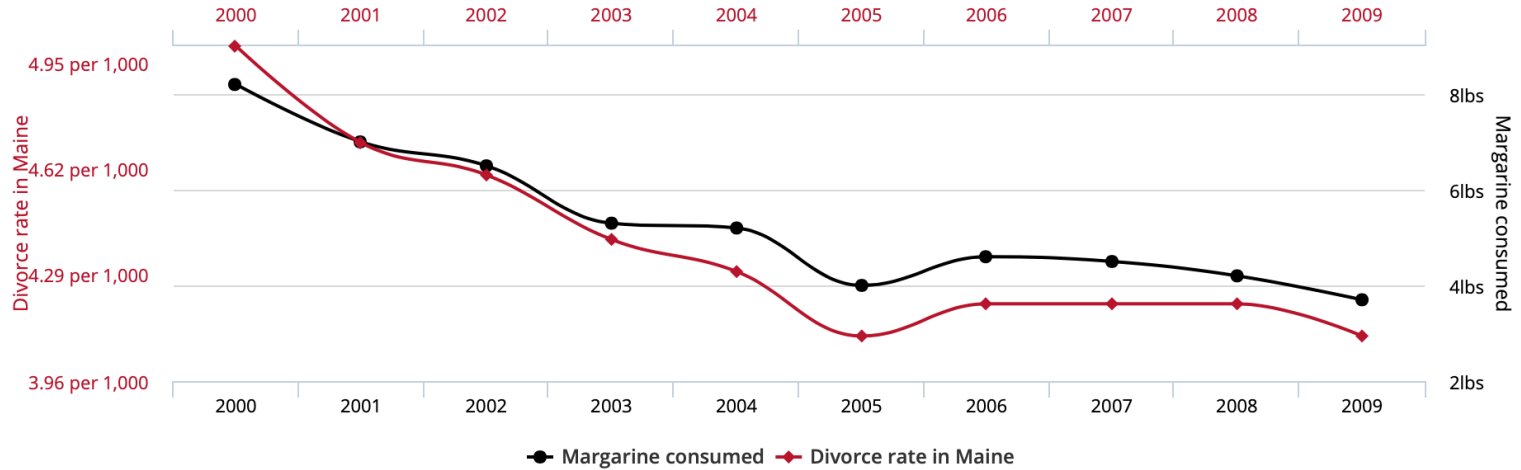
### Distinction between *causation* and *correlation*

(see <http://tylervigen.com/spurious-correlations> for an interesting collection)

# Divorce rate in Maine correlates with Per capita consumption of margarine



Correlation: 99.26% (r=0.992558)



Data sources: National Vital Statistics Reports and U.S. Department of Agriculture

tylervigen.com

This may be compounded over time by

- The principle according to which alike cases must be treated alike
- *Ad populum* fallacy
- Pressure to conform

## D) Indirect discrimination – adverse impact of decisions



AI & Legitimate Authority

**UQAM**

Faculté de science politique  
et de droit

Université du Québec à Montréal

# 3. Arguments Against the NJT



# A) Degradation of human capacities required for autonomy

# Automation may breed

- automation biases
- loss of situational awareness
- complacency
- skill degradation

# B) *Existential* Arguments Against the NJT

For certain decisions, what matters is less that they are *right*, than they are...

... *our* decisions (“self-government”)

... arrived at *by* a *responsible* moral agent

... the result of deliberations that took the *meaning* of the issues seriously *as opposed* to be treated as merely operational questions

Such cases impose constraints on *WHO* or *WHAT* can be a legitimate authority in those circumstances

- Agent part of the collective self
- Moral agent, and/or
- Meaning-making agent

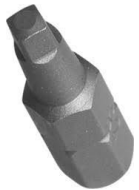
**FUNCTIONAL AUTHORITIES** (Instrumental  
rationality)

VS.

**EXISTENTIAL AUTHORITIES** (Existential logics)

Is an issue a functional or an existential one?  
Not one or the other by “nature”, but rather by  
*ascription*

Robertson



vs.

Philips



# 'You've got to use a Robertson'

CATHERINE DAWSON MARCH

TORONTO

PUBLISHED JUNE 15, 2002

UPDATED APRIL 17, 2018

0 COMMENTS



Toques? Politeness? Strong beer? Forget all that. It's a simple piece of hardware that sets us apart from Americans.

The Robertson square-drive: a screw topped with a square hole ready to receive its matching hand driver. A thing of beauty for Canadian woodworkers because that recessed square won't let you slip out. The screwdriver fits so snugly, you can torque this baby upside-down with one hand. It's the best screw around, and most Americans have never heard of it. (They're stuck fumbling with the slot- or Phillips star-head fasteners that, like any cheap screw, strip far too easily.)

Canadians who grew up using the 94-year-old invention likely haven't heard much about Peter Lymburner Robertson, the screw's Ontario inventor, or the story of how North America became a continent of two screw solitudes.

<https://www.youtube.com/watch?v=pp3if2nl3K4>

AI & Legitimate Authority

UQÀM

Faculté de science politique  
et de droit

Université du Québec à Montréal



# Conclusion

1. **Need to determine right level of automation**
  - **Nature of the variables involved (functional analysis)**
  - **Impact of automation and cognitive biases**
  - **Meaning of the decisions involved**
2. **Need for proper Algorithmic Impact Assessment**  
<https://www.canada.ca/en/government/system/digital-government/modern-emerging-technologies/responsible-use-ai/algorithmic-impact-assessment.html>



# HumanIA.uqam.ca

AI & Legitimate Authority

**UQÀM**

**Faculté de science politique  
et de droit**

Université du Québec à Montréal