

Ethique artificielle et assurance – un bref historique

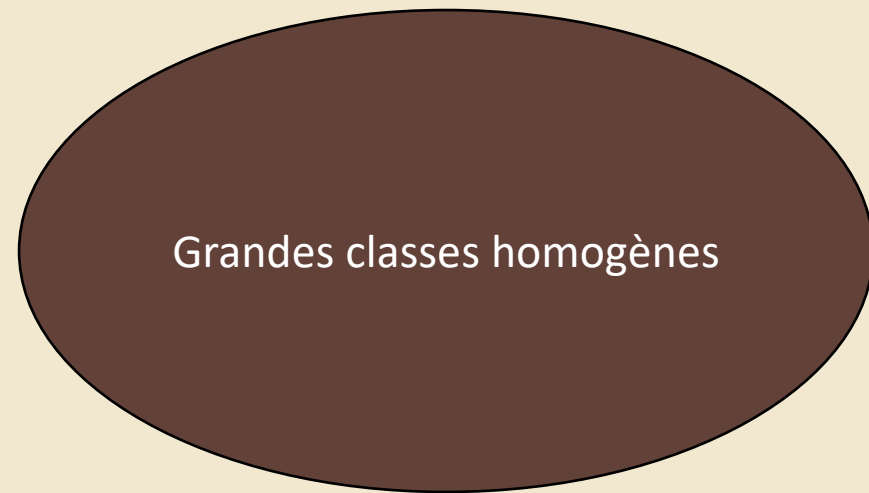
CHAIRE

PARI

PROGRAMME DE RECHERCHE
SUR L'APPREHENSION DES RISQUES
ET DES INCERTITUDES

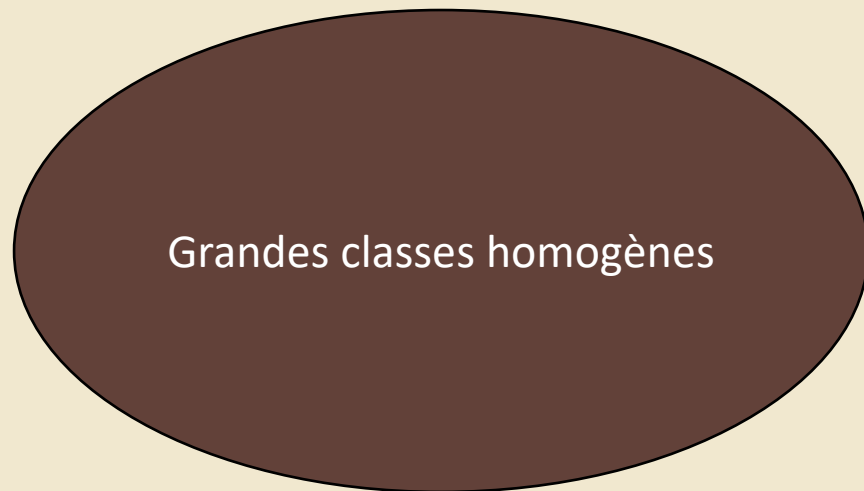
Laurence Barry & Arthur Charpentier,
Colloque Francophone 07/10/2022

L'équité actuarielle au 20e siècle



1960

L'équité actuarielle au 20e siècle



The endeavours of the rate making actuary can now be represented as follows:

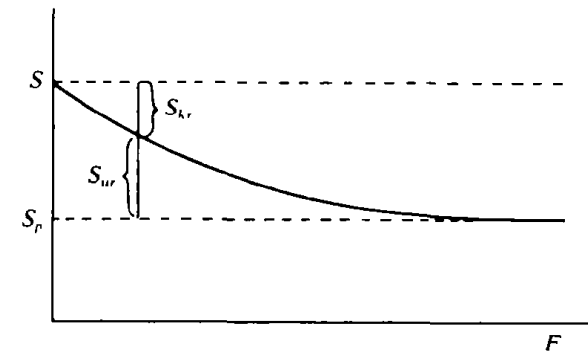


FIGURE 1 Solidanty shared among insureds

De Wit & Van Eeghen 1984

1960

L'équité actuarielle au 20e siècle

“Les moins chanceux sont
couverts par tous les autres”

“Les bons risques ne paient
pas pour les mauvais”

1960

CHAIRE

PARI

PROGRAMME DE RECHERCHE
SUR L'APPREHENSION DES RISQUES
ET DES INCERTITUDES

L'équité actuarielle au 20e siècle

Equité comme justice

Equité comme justesse

1960

Discrimination (statistique) et biais (discrimination sociale)



Art de quantifier



Art d'interpréter

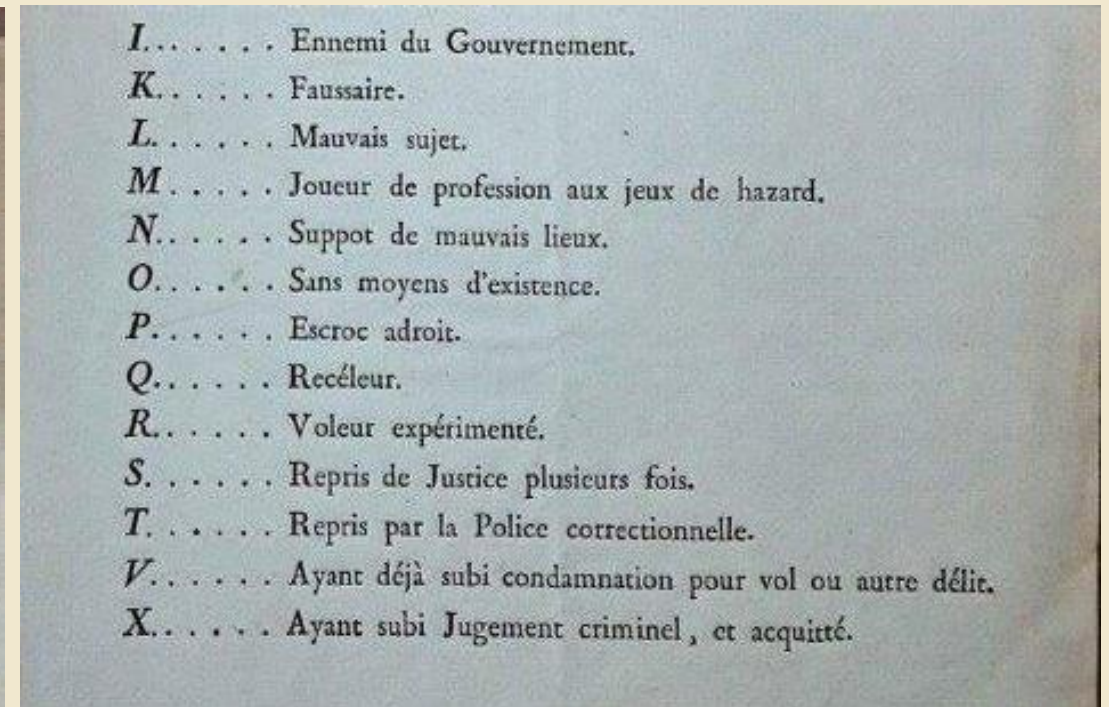
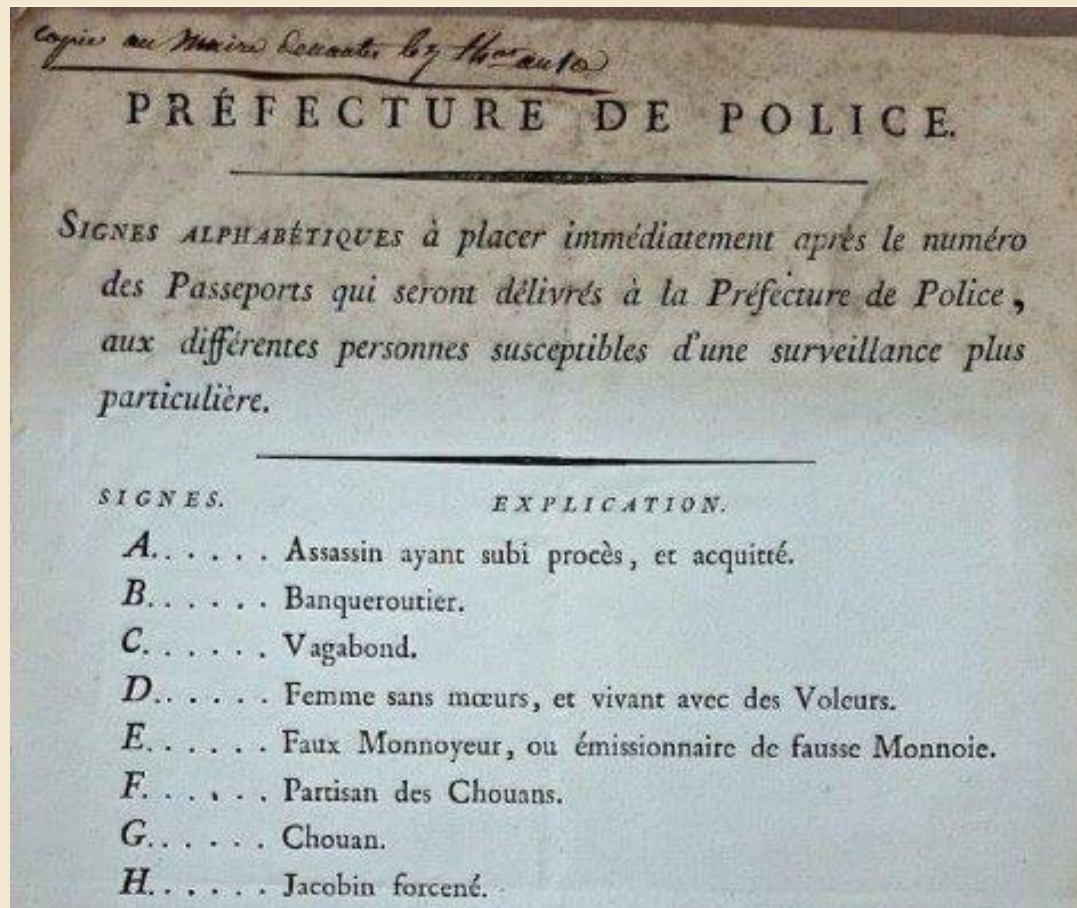
CHAIRE

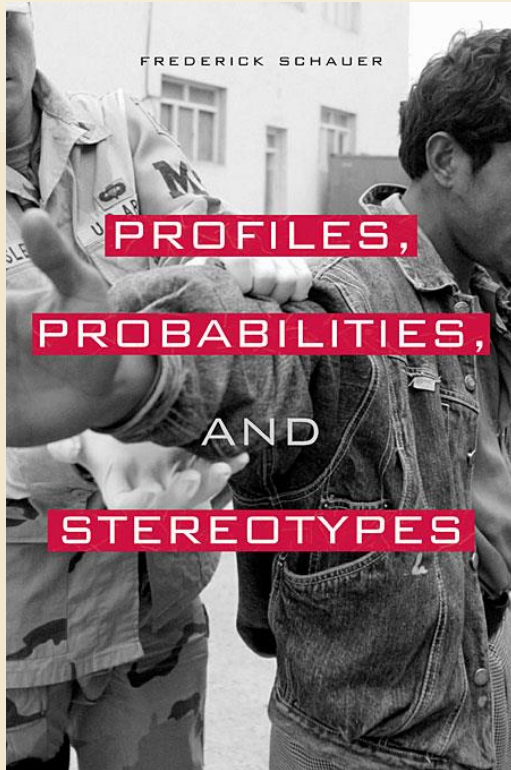
PARI

PROGRAMME DE RECHERCHE
SUR L'APPREHENSION DES RISQUES
ET DES INCERTITUDES

Discrimination (statistique) et biais (discrimination sociale)

Liste des individus dangereux, 1801/02





Pas de fondement statistique

“Generalizations based on a person's astrological sign, for example, are **pure prejudice.**”

Corrélation

“But **others have a statistical basis**, when the probability of having a character y knowing x is significantly different from the case where nothing is known.”

CHAIRE

PARI

PROGRAMME DE RECHERCHE
SUR L'APPREHENSION DES RISQUES
ET DES INCERTITUDES

Corrélations?



Trompeuse: coincidence,
mediateur, confondante...

“Vrai” lien causal

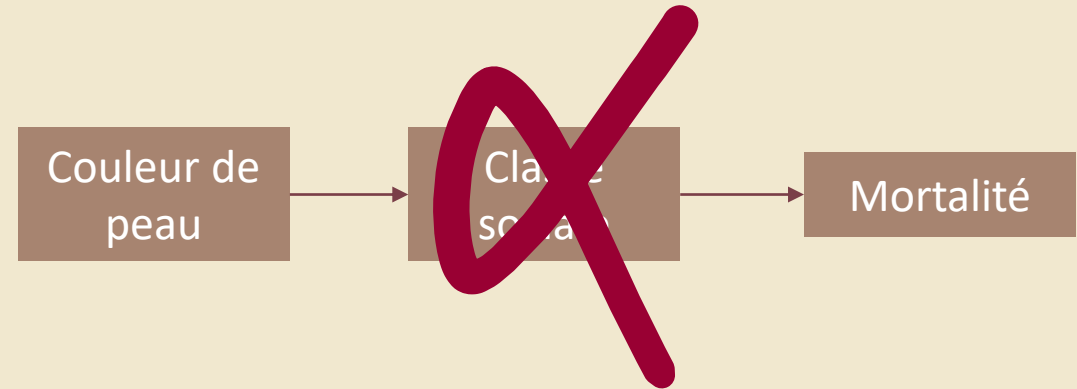
CHAIRE

PARI

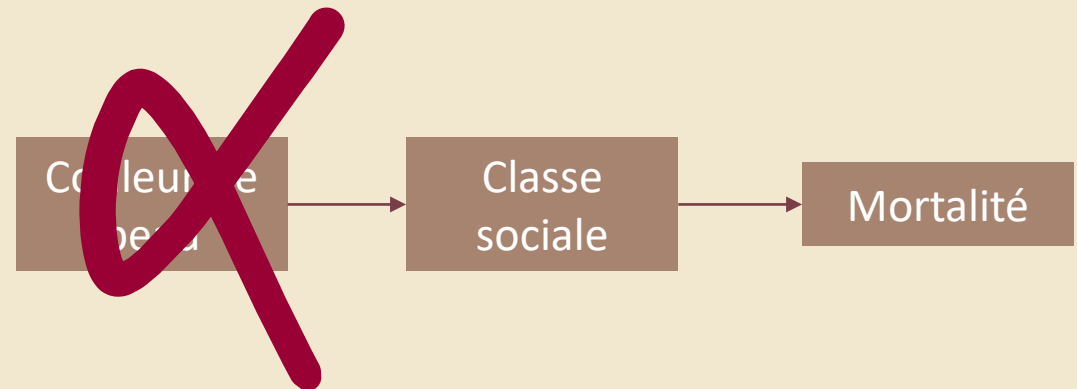
PROGRAMME DE RECHERCHE
SUR L'APPREHENSION DES RISQUES
ET DES INCERTITUDES

Correlations?

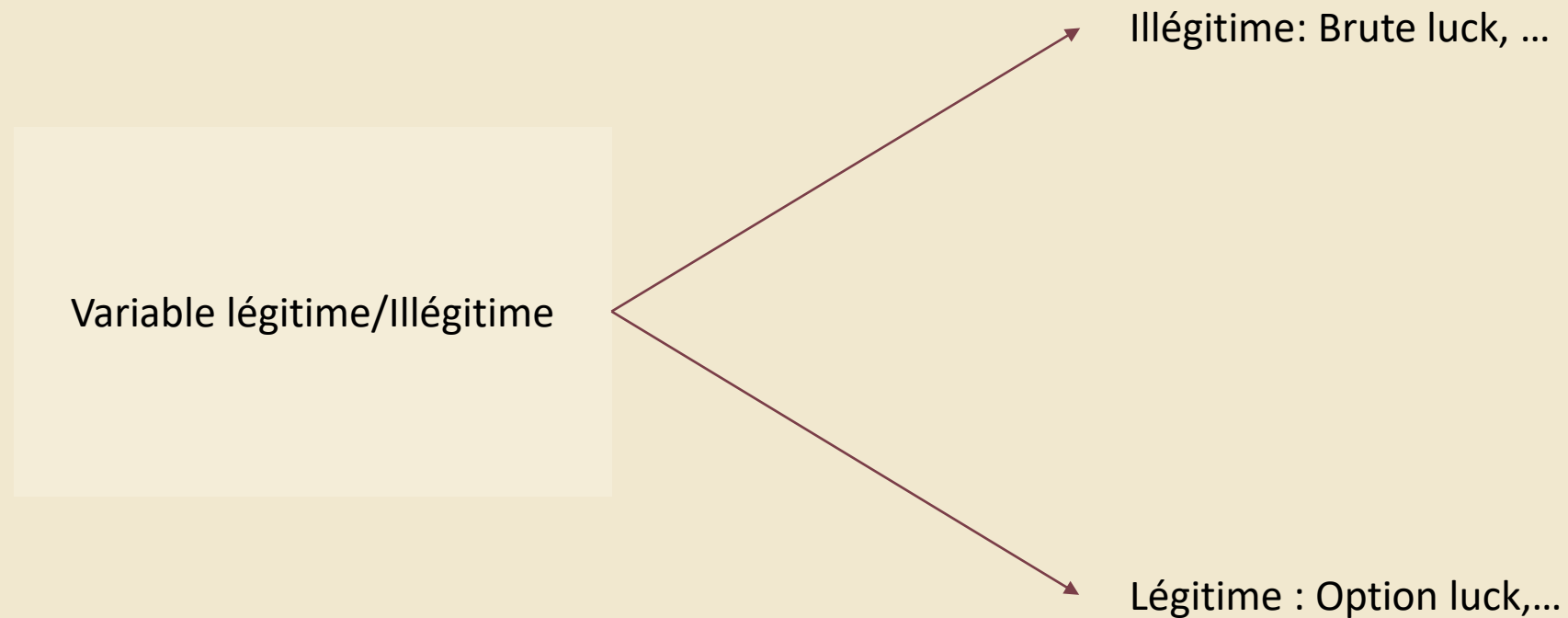
Trompeuse: coincidence,
mediateur, confondante...



“Vrai” lien causal



“Vrai” lien causal?



Une typologie des biais (dans la situation statistique classique)



Type 1: variables sans rapport avec la réalité que l'on essaie de décrire



Type 2: variables corrélées mais non causales



Type 3: variables causales mais illégitimes

Apprentissage machine (et données massives) :
quoi de neuf?

Apprentissage machine (et données massives) : quoi de neuf?

Vieux biais

L'art de quantifier



"the key aspect of deep learning is that these layers of features are not designed by human engineers: they are learned from data using a general-purpose learning procedure"

Lecun, Bengio Hinton 2015

CHAIRE

PARI

PROGRAMME DE RECHERCHE
SUR L'APPREHENSION DES RISQUES
ET DES INCERTITUDES

Type 1: capter la réalité avec les données massives?

 Biais d'échantillon

 Les données reproduisent les biais sociaux

Type 2: trouver la causalité?

Vieux biais

L'art d'expliquer

Scientists are trained to recognize that correlation is not causation, that no conclusions should be drawn simply on the basis of correlation between X and Y (it could just be a coincidence). Instead, you must understand the underlying mechanisms that connect the two. Once you have a model, you can connect the data sets with confidence. Data without a model is just noise. But *faced with massive data, this approach to science - hypothesize, model, test - is becoming obsolete*

Anderson 2008

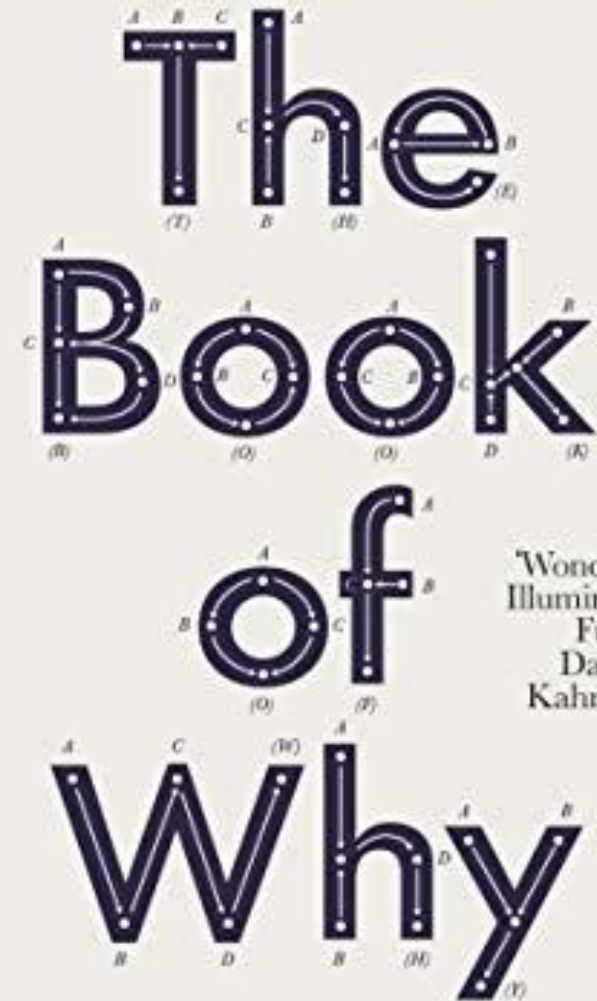
Interprétabilité vs. Performance

The Great AI Debate - NIPS2017



Data vs. Causality?

Judea Pearl
& Dana Mackenzie



The New Science
of Cause and Effect

Type 3: variable illégitime/protégée

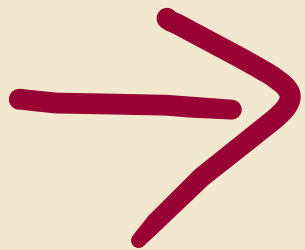
An AI deprived of information about a person's genetic test results or obvious proxies for this information (like family history) will use other information-ranging from TV viewing habits to spending habits to geolocational data-to proxy for the directly predictive information contained within the genetic test results.

Prince et Schwarcz 2019, 1274

Type 3: variable illégitime/protégée

An AI deprived of information about a person's genetic test results or obvious proxies for this information (like family history) will use other information-ranging from TV viewing habits to spending habits to geolocational data-to proxy for the directly predictive information contained within the genetic test results.

Prince et Schwarcz 2019, 1274



Avoir accès et se servir des variables protégées pour contrôler les biais du modèle

Conclusion: éthique artificielle, le neuf et le vieux

Merci!

