

**ÉCOLES NORMALES SUPÉRIEURES
ÉCOLE NATIONALE DES PONTS ET CHAUSSÉES**

**CONCOURS D'ADMISSION SESSION 2020
FILIÈRE BCPST
COMPOSITION DE LANGUE VIVANTE ÉTRANGÈRE**

Épreuve commune aux ENS de Lyon, Paris, Paris-Saclay et à l'ENPC

Durée : 2 heures

L'utilisation des calculatrices n'est pas autorisée pour cette épreuve.

L'usage de dictionnaire est interdit.

Parmi les sujets proposés, le candidat doit traiter celui correspondant à la langue vivante étrangère qu'il a choisie lors de son inscription.

Toute copie rédigée dans une langue qui ne correspondrait pas au choix apparaissant dans le dossier d'inscription définitif du candidat sera considérée comme nulle.

Sujets proposés :

- Allemand
- Anglais
- Espagnol

ANGLAIS

I. Version (12 points)

An algorithm deployed across the United States is now known to underestimate the health needs of black patients. The algorithm uses health-care costs as a proxy for health needs. But black patients' health-care costs have historically been lower because systemic racism has impeded their access to treatment — not because they are healthier. This example illustrates how machine learning and artificial intelligence can maintain and amplify inequity. Most algorithms exploit crude correlations in data. Yet these correlations are often by-products of more salient social relationships (in the health-care example, treatment that is inaccessible is, by definition, cheaper), or chance occurrences that will not replicate. To identify and mitigate discriminatory relationships in data, we need models that capture or account for the causal pathways that give rise to them. Here we outline what is required to build models that would allow us to explore ethical issues underlying seemingly objective analyses. Only by unearthing the true causes of discrimination can we build algorithms that correct for these. Models that account for causal pathways have three advantages. These 'causal models' are: tailored to the data at hand; allow us to account for quantities that aren't observed; and address shortcomings in current concepts of fairness. A causal model represents how data are generated, and how variables might change in response to interventions. This can be shown as a graph in which each variable is a node and arrows represent the causal connections between them. (...) Having a causal model allows us to address questions related to ethics, such as does religion influence the visa process? But because many different causal models could have led to a particular observed data set, it is not generally possible to identify the right causal model from that data set alone. For example, without any extra assumptions, data generated from the causal graph described here could seem identical to those from a graph in which religion is no longer linked to visa granting. A modeler must therefore also leverage experiments and expert knowledge, and probe assumptions.

The long road to fairer algorithms, by Joshua Loftus and Matt Kusner, Nature, February 4th, 2020

II. Questions (8 points, 100 mots minimum par question)

1. "This example illustrates how machine learning and artificial intelligence can maintain and amplify inequity" Explain.
2. To what extent should we program ethics into AI?