

# De l'apprentissage machine et des masses de données éthique du numérique

Jean-Gabriel Ganascia\*

\* Lip6 et Sorbonne Université

<https://www.lip6.fr/actualite/personnes-fiche.php?ident=P34>

**Résumé.** Qu'il s'agisse de protéger l'intimité, la vie privée, l'anonymat ou la propriété, d'encourager le partage, voire de garantir l'absence de biais, les questions éthiques suscitées par la captation et l'exploitation des données sont à la fois nombreuses et anciennes. à cela s'ajoutent les effets délétères de certaines applications hâtives des systèmes prédictifs fondés sur l'utilisation de techniques d'apprentissage machine, comme la justice prédictive, l'établissement dynamique des prix ou la répartition des forces de polices. Or, l'approche de ces questions varie dans le temps et selon les cultures. De plus, certaines prescriptions morales, comme l'impartialité dans la collecte ou les principes de finalité et de proportionnalité de la CNIL, vont à l'encontre des postulats épistémologiques posés par les zéloteurs les plus enflammés des masses de données. Il s'agira là de présenter ces problématiques éthiques et de les mettre en regard des questions épistémologiques fondamentales suscitées par la science des données.

## Summary

Whether protecting privacy, anonymity or property, encourage sharing, or to guarantee the absence of bias, the ethical issues raised by the collection and use of data are both numerous and ancient. In addition, there are the deleterious effects of some hasty applications of predictive systems based on the use of machine learning techniques, such as predictive justice, dynamic pricing or the distribution of police forces. However, the approach to these questions varies over time and between cultures. In addition, certain moral prescriptions, such as impartiality in the collection or the principles of finality and proportionality of the CNIL, go against the epistemological postulates posed by the most inflamed zealots of the masses of data. This will involve presenting these ethical issues and comparing them with the fundamental epistemological questions raised by data science.

