Chaire “Law, Accountability and Social Trust in A.I.”

Céline CASTETS-RENARD

PR1 UT1/IUF – Droit du numérique, aspects éthiques et légaux de l’I.A.

Sylvain CUSSAT-BLANC

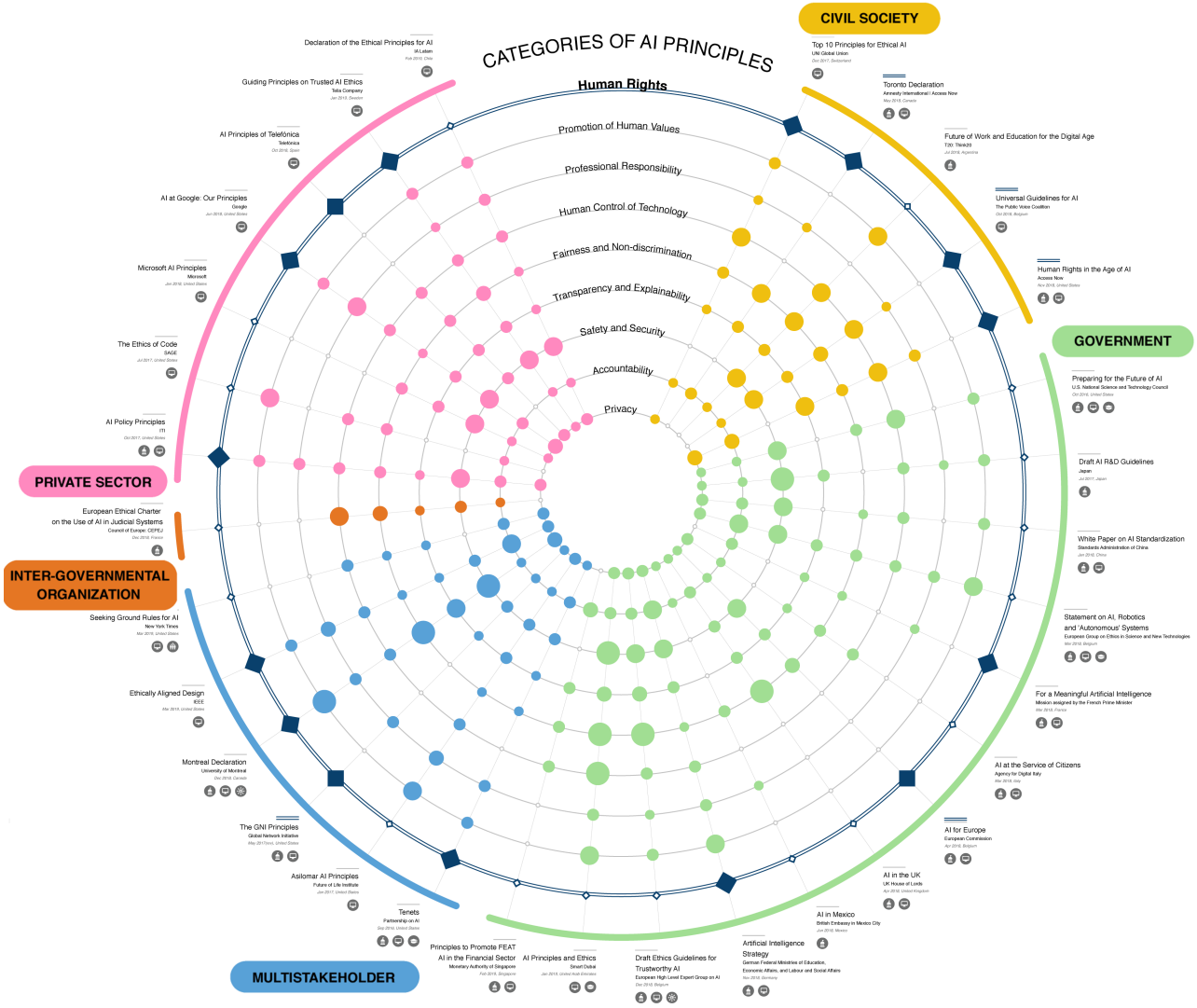
McF IRIT – Algorithmes Evolutionnaires, I.A. bio-inspirée

Laurent RISSER

IR1-HDR IMT – Interprétabilité en I.A., régularisation, applications

Ethical Principles

- Ethical principles related to fundamental rights
- A Map of Ethical and Rights-Based Approaches (Harvard, July 2019)



https://ai-hr.cyber.harvard.edu/primp-viz.html

Legal Issues

- Algorithmic decision-making: unbiased, non-discriminant, equitable and fair, transparent, explicable, accountable

- Ex. EU (GDPR, art. 22):
  - *The right not to be subject to a decision solely based on automated processing, including profiling*
- Ex. Fr (LIL, art. 10)
  - *The right to obtain a human intervention, the right to contest the decision*
  - *The right to know the rules of the processing and its main characteristics*
- Conseil Constitutionnel (May 2018):
  - *the data controller has to control the algorithmic processing and its developments to explain to the data subject the way of the processing was applied to its situation*
- Ex. NYC Bill (Dec. 2017): A local law in relation to automated decision systems used by agencies (task force)

- Sectorial regulation on autonomous vehicles, smart city, facial recognition…

Proactivity in Terms of Norms

- To contribute to the development of an ethical and legal framework to regulate the use of A.I., taking special account of bias, discrimination, equity and social justice (fairness)

- To have tools to deal with the "accountability" of actions and decisions resulting from the A.I. (transparency, accountability, auditability, “certifiability”, liability) for a social trust

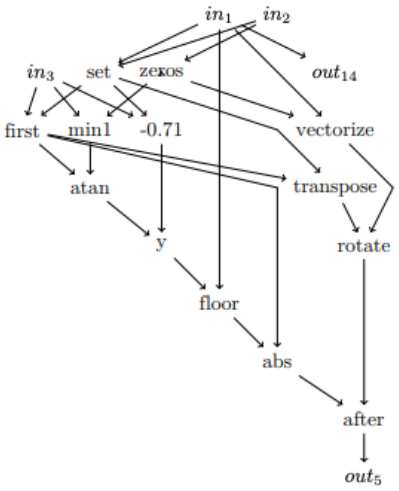
- To define effective A.I. strategies for future rules in the industry (risk limitation)

Proactivity in Terms of Tools

- Developing interpretability solutions in AI
- Application to automatic decision systems (Autonomous transportations and Smart city)
- Studying accreditation and certification methods for AI based systems

Proactivity in Terms of Tools

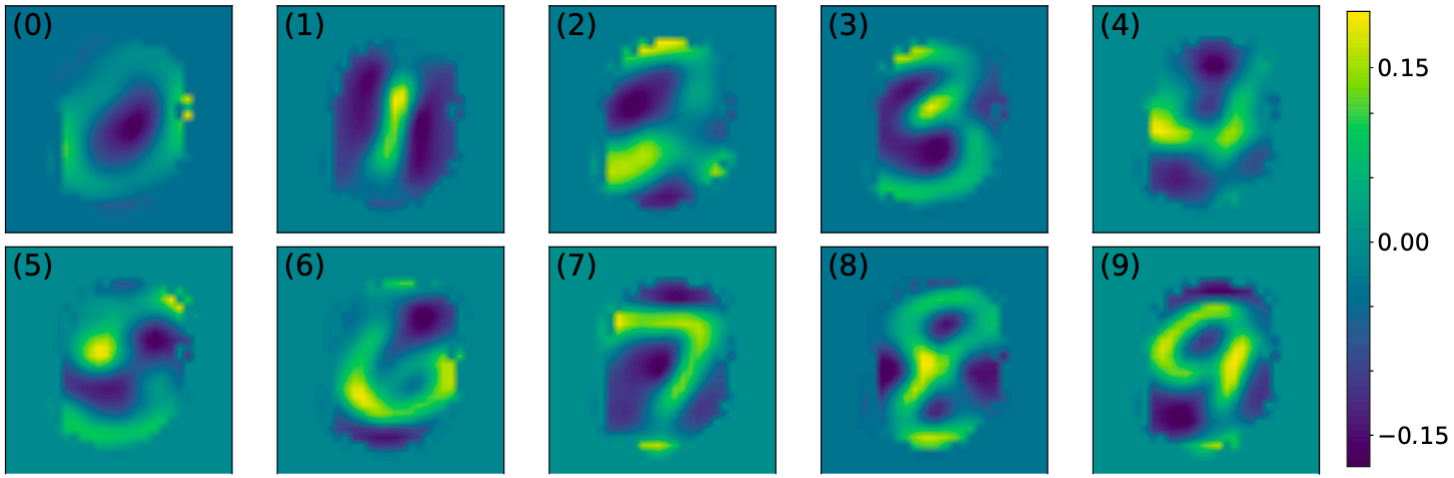
- Genetic Programming (GP):
  - Exploring usability of whitebox approaches in industrial contexts
  - Use of evolutionary computation to build:
    - Programs
    - Symbolic regressions
  - Possible applications to
    - Autonomous aircraft/cars
    - Genetic improvement (GP for debugging, code optimization, etc.)
    - Data modelling
    - Etc.



Wilson, D. G., Cussat-Blanc, S., Luga, H., & Miller, J. F. (2018, July). Evolving simple programs for playing Atari games. In *Proceedings of the Genetic and Evolutionary Computation Conference* (pp. 229-236). ACM.

Proactivity in Terms of Tools

- Interpretability of black-box decision rules



Bachoc, Gamboa, Halford, Loubes, Risser (2019) — https://arxiv.org/pdf/1810.07924.pdf

- Regularization strategies to ensure fair decision rules

$$\hat{\theta} = \arg\min_{\theta} R(\theta) + \lambda W_2^2(\mu_{\theta,0}^n, \mu_{\theta,1}^n)$$

Risser, Vincenot, Couellan, Loubes (2019) — https://arxiv.org/pdf/1908.05783.pdf

Within ANITI

IP1 (Fair Representative Data For Artificial Intelligence)
Relationships with the Chair of J.-M. Loubes on Fairness and IP2

Besse, J.-Ph., Castets-Renard, C., Garivier, A., Loubes, J.-M. (2018a), « L'IA du Quotidien peut-elle être Éthique ? : Loyauté des Algorithmes d'Apprentissage Automatique », Statistique et Société, 2018-3, vol. 6.