



Science des données pour evidence based decision making

Centre de compétences en science des données (DSCC)

Un centre d'excellence en science des données : au service des administrations publiques suisses

Pauline Maury-Larivière

Data Scientist

Section DSAI



Competence Network for Artificial Intelligence
Kompetenznetzwerk für künstliche Intelligenz
Réseau de compétences en intelligence artificielle
Rete di competenze per l'intelligenza artificiale



Agenda

- Centre de compétences en science de données de l'OFS
- Inférence causale et contrefactuels
- Un exemple avec l'égalité entre les sexes





DSCC : mandat et vision

Mandat

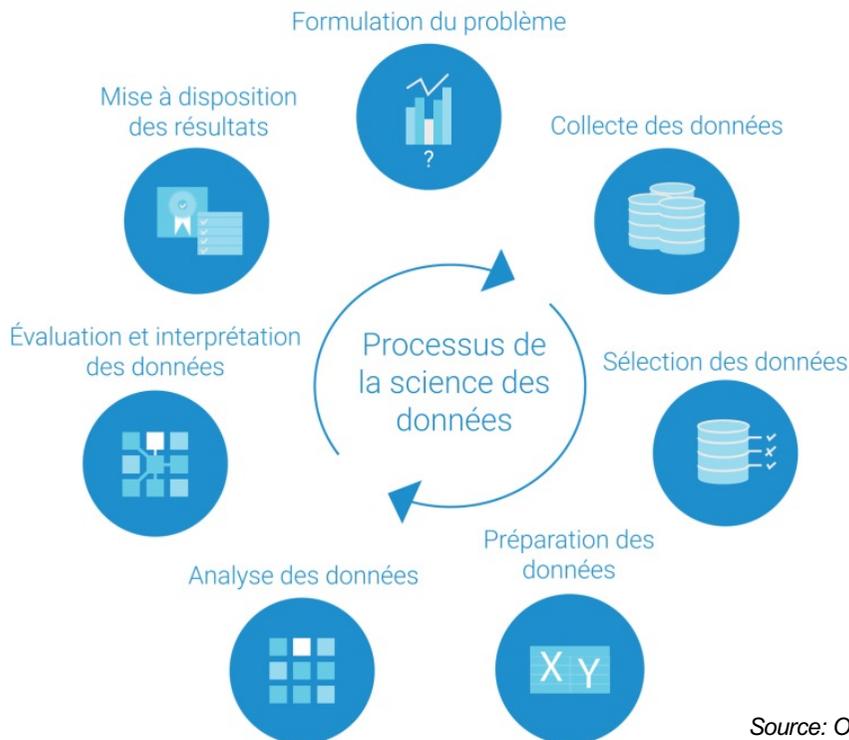
«Le Centre de compétences en science des données propose des services dans le domaine de l'innovation en matière de données et encourage le développement des connaissances au sein de l'administration fédérale, ainsi que les échanges avec les administrations cantonales et communales.»

Vision

Nous recourons à la science des données et développons des compétences pour le bien commun dans toute la Suisse (for public good).



Comment fonctionne la science des données?



Nos compétences englobent tout le processus propre à la science des données et la réalisation de produits basés sur les données.

Source: OFS



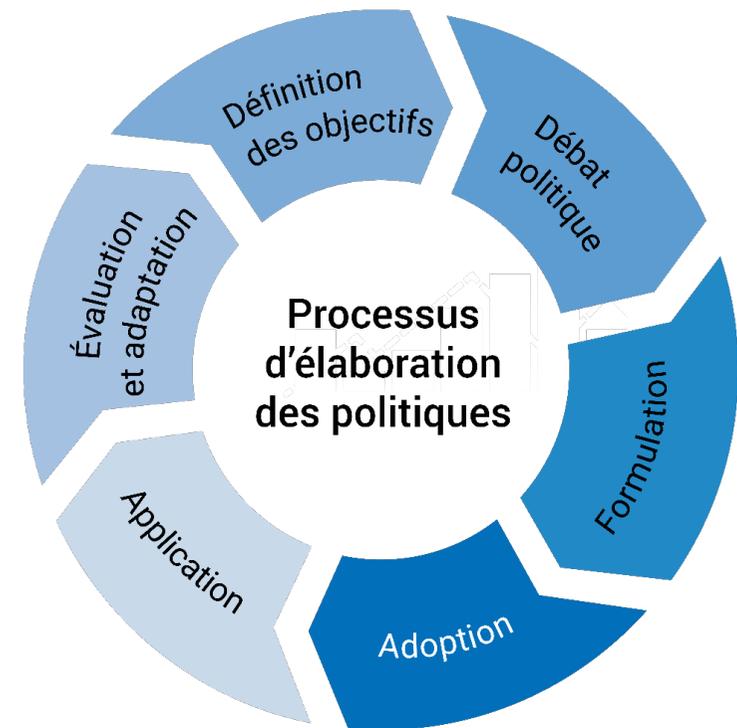
Élaboration de politiques sur la base de données probantes

Associées à des connaissances spécifiques, les données peuvent améliorer et accélérer le processus de décision politique.

Objectif: fournir un appui aux responsables politiques en analysant les politiques élaborées et en évaluant les conséquences d'autres scénarios politiques.

Le DSCC met des connaissances à disposition dans les domaines suivants:

- Inférence causale
- Évaluation des risques
- Utilisation de mégadonnées issues de sources non traditionnelles



Source: Zukunftskunst.eu

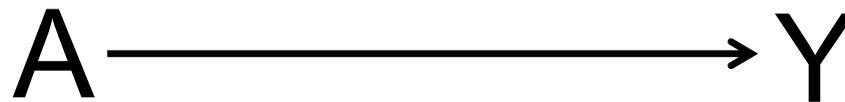


Qu'est-ce que l'inférence causale?

“l'identification et l'estimation des effets de causalité”

c.à.d. quel est l'effet de A sur Y?

On peut représenter cela sous forme de «Directed Acyclic Graph»
DAG (graphe acyclique dirigé)



Les méthodes d'inférence causale peuvent être utilisées pour estimer des contrefactuels.

*Hernán MA, Robins JM (2020). Causal Inference: What If. Boca Raton: Chapman & Hall/CRC



Qu'est-ce qu'un contrefactuel?

Répond à la question: Et si...?

Les voyages dans le temps pourraient nous permettre d'observer ce qui se produirait si notre intervention d'intérêt (A) était modifiée.

En l'absence de voyages dans le temps, nous estimons des contrefactuels à l'aide d'expériences ou en rendant les données observationnelles aussi proches que possible des expériences.



Source: Wikipedia [DeLorean time machine - Wikipedia](#)



The Potential Outcomes Framework

Permet la construction de mondes possibles proches (mondes contrefactuels) où la variable / l'intervention d'intérêt a été modifiée et prend une valeur différente:

Par exemple le genre, la politique publique etc.



Source: [Exploring the Shape of Our Universe—and the Multiverse | Tufts Now](#)

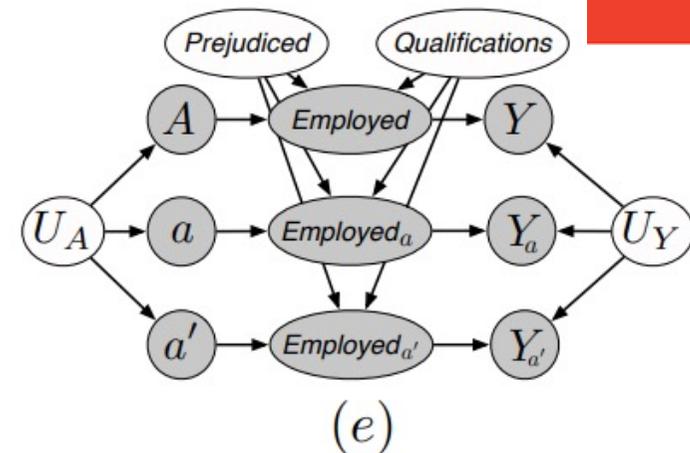


The Potential Outcomes Framework

Le «twin network» à droite permet, par exemple, de répondre à la question:

«Si la demande de prêt avait été faite par une femme, le prêt aurait-il été approuvé ?»

L'architecture d'un «twin network» peut être construite à l'aide d'un réseau neuronal.



Kusner, M. J., Loftus, J., Russell, C., & Silva, R. (2017). Counterfactual fairness. *Advances in neural information processing systems*, 30.



Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Confederation

Bundesamt für Statistik BFS
Office fédéral de la statistique OFS
Ufficio federale di statistica UST
Uffizi federal da statistica UST
Federal Statistical Office FSO



Data Science
Competence Center DSCC

