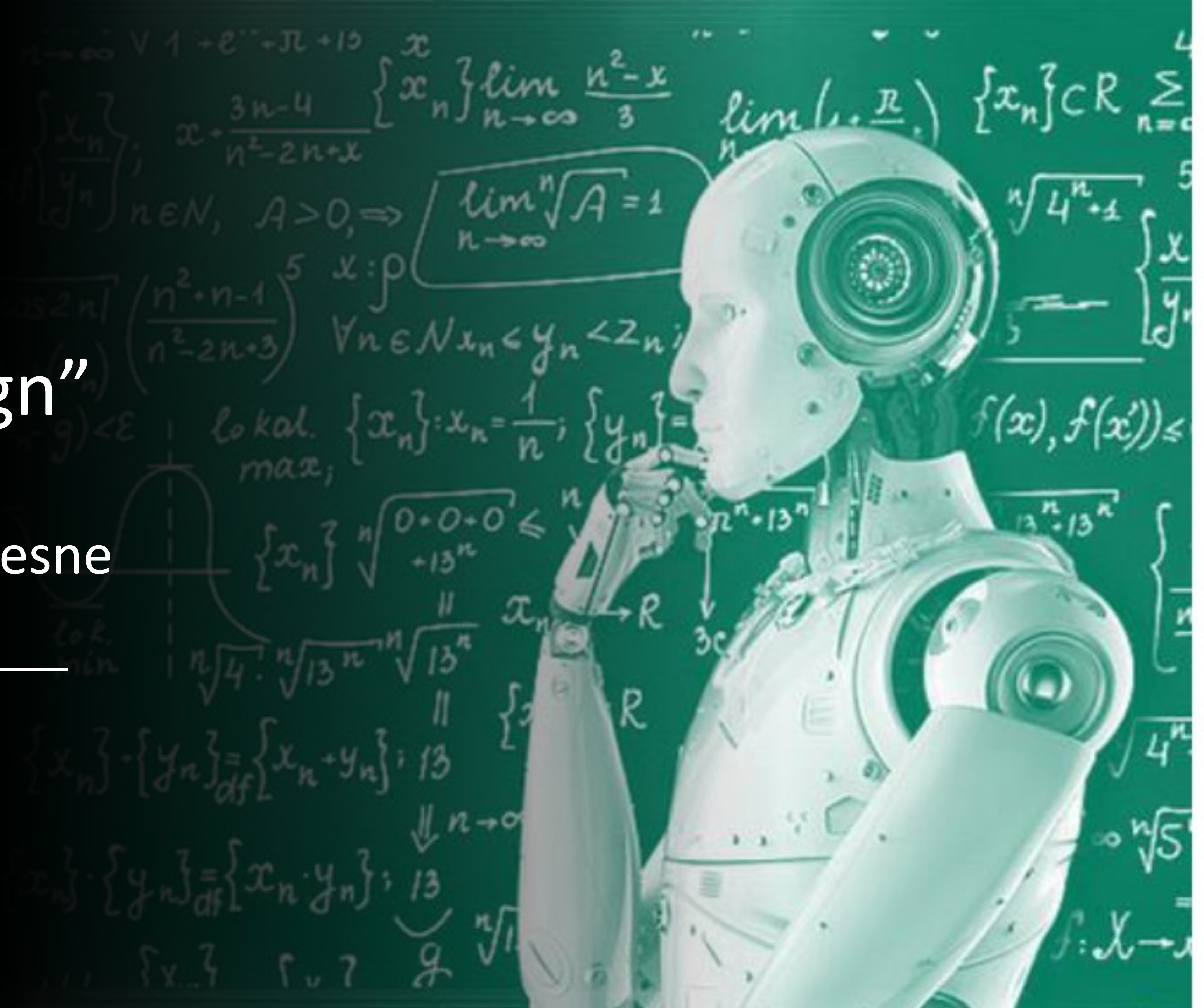
AI for health
“ethics by design”

Baptiste Couvy-Duchesne

The University of Queensland

baptisteCD

@BaptisteCouvy



Ethics recommendations

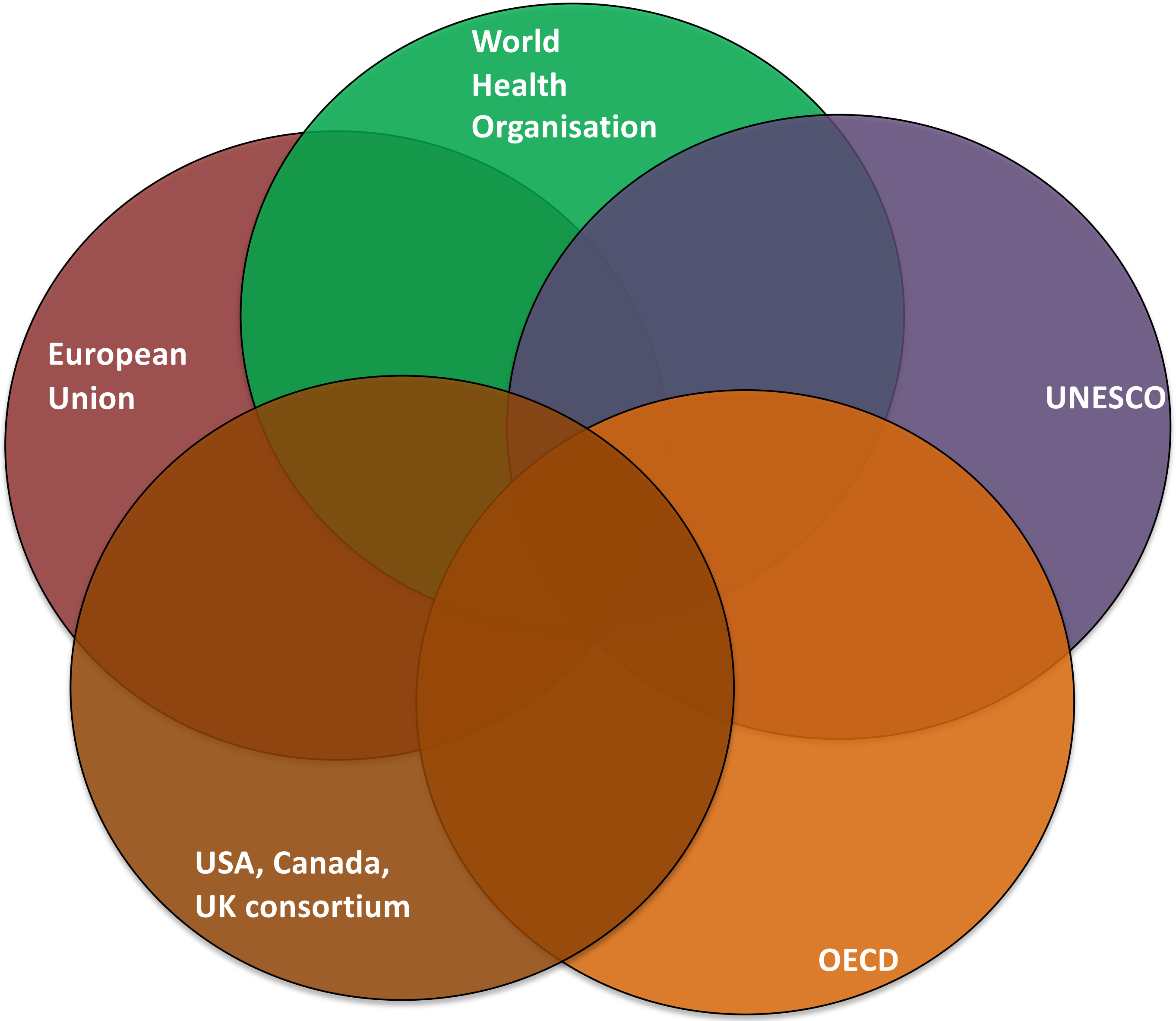
https://op.europa.eu/fr/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a

https://apps.who.int/iris/rest/bitstreams/1352854/retrieve

https://fr.unesco.org/artificial-intelligence/ethics#recommandation

https://oecd.ai/en/ai-principles

https://www.fda.gov/media/153486/download

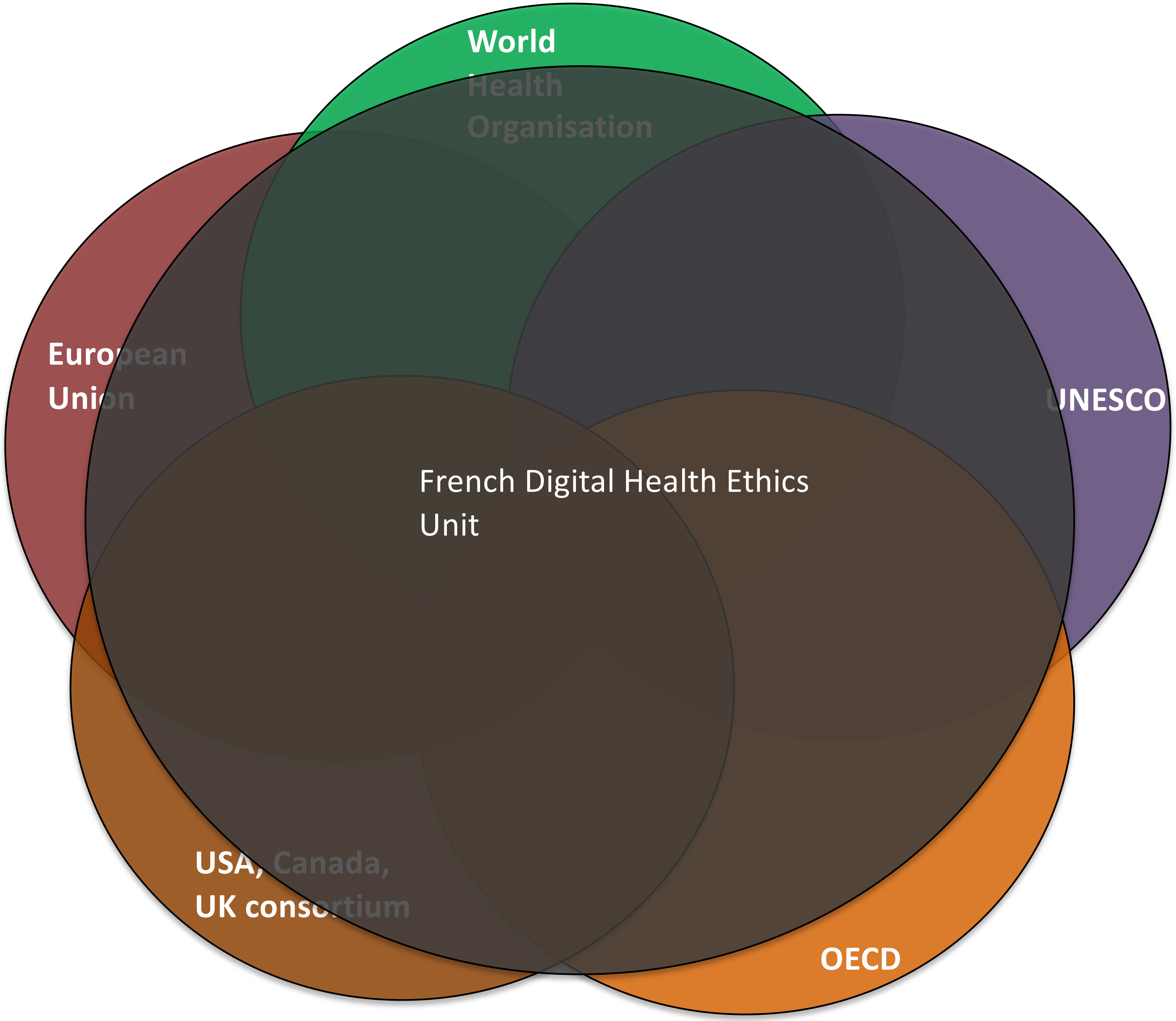


MINISTÈRE DES SOLIDARITÉS ET DE LA SANTÉ
Liberté
Égalité
Fraternité

Recommendations for good practice to integrate ethics from the development stage of Artificial Intelligence solutions in Health : Implementing "ethics by design"

Digital Health Ethics Unit of the *French Ministerial Delegation for Digital Health*
– April 2022

https://esante.gouv.fr/sites/default/files/media_entity/documents/ethic_by_design_guide_vf.pdf



Definition of AI

- AI aims at developing **devices, hardware and software**,
- **capable of implementing** a tool
- **aiming at producing the same result as** the one obtained by the cognitive mechanisms of **a human expert** engaged in a **problem-solving task**,
- for the **purpose of assisting or replacing human activities**.

Ethics by design

- By Design means:
  - By and from conception
  - Intentionally, consciously and premeditated

- Core ethical values:
  - Protection of personal data
  - Human guarantee of any high-stake decision
  - Thinking about the action, its legitimacy and its consequences

Ethics by design – steps

Framing stage
Define the purpose of the AI solution and validate the ethics of the purpose

Step 1 –
Data collection

Step 2 –
Data pre-treatment

Step 3 –
Build the algorithm

Step 4
Evaluation of the algorithm and preparation of production

Framing stage : Define the purpose of the AI solution solution and validate the ethics of the purpose

Possible use-cases of Alzheimer’s disease (risk) prediction

- Diagnostic

- Not really a clinician need
- Actual diagnosis not tedious, long process or particularity expensive
- Current diagnosis fairly accurate
- Limited treatment options

Framing stage : Define the purpose of the AI solution solution and validate the ethics of the purpose

Possible use-cases of Alzheimer's disease (risk) prediction

- Prognostic

- Limited preventative treatments
- Actionable factors (e.g. exercise, stimulation) not specific

Framing stage : Define the purpose of the AI solution solution and validate the ethics of the purpose

Possible use-cases of Alzheimer’s disease (risk) prediction

- Quantify risk
- Research

- Recruitment in clinical trials
- Measure of Alzheimer’s risk in cohorts where information not available
- Progress understanding of disorder (e.g. genetics, biology, brain markers…)

Framing stage : Define the purpose of the AI solution solution and validate the ethics of the purpose

Purpose of the AI solution.

Target users

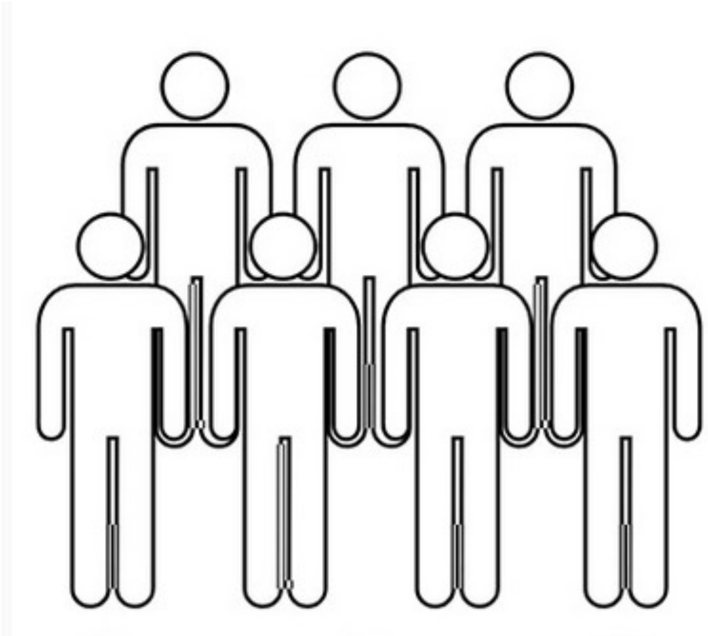
Type of learning used

The subjects for which it will be used

For example

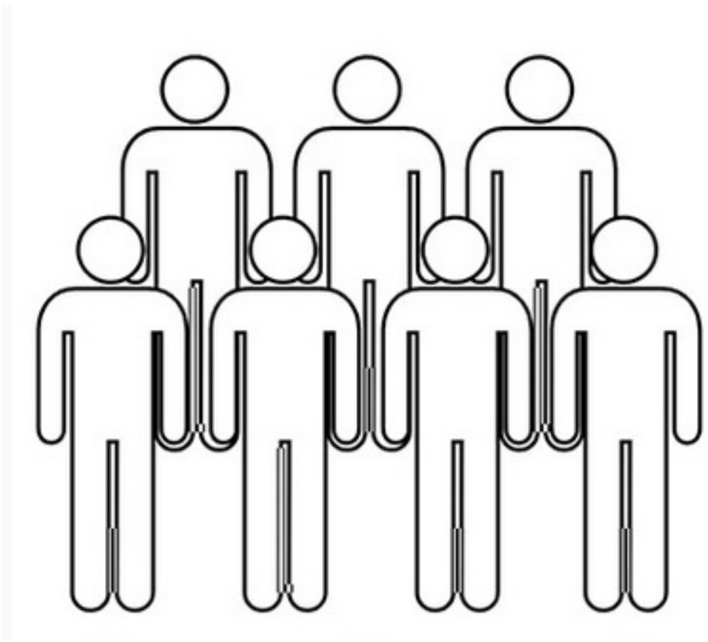
- Quantify the genetic risk of Alzheimer’s disease, to study the link between Alzheimer’s risk and Major Depressive Disorder.
- To use in research.
- We will build a linear predictor using the latest Genome Wide Association Study to date.
- We will apply the risk predictor in a local cohort of (mostly European) adults (age range 20-40) who have been screened for psychiatric disorders

Step 1 – Data collection



Informed consent	Proportionality of the data collected
Non direct re-identification of data: pseudoanonymization	Quality of the data
Representativeness of the analysis population/ target population + prevention of discrimination	User involvement
Secure data transfer + Quality of data hosting + State of the art cyber security	Measures to ensure non-ethical reuse of data

Step 1 – Data collection



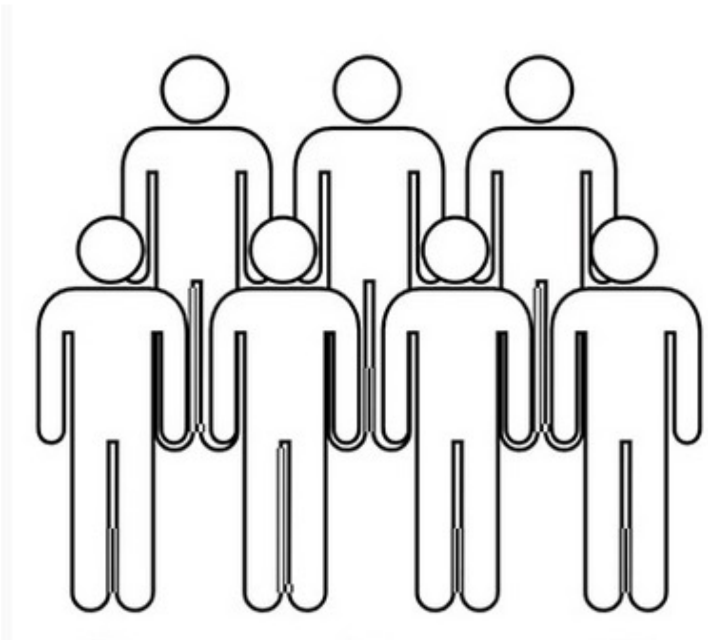
- Informed consent
- Proportionality of the data collected
- Non direct re-identification of data: pseudoanonymization
- Quality of the data
- Representativeness of the analysis population/ target population + prevention of discrimination
- User involvement
- Secure data transfer + Quality of data hosting + State of the art cyber security
- Measures to ensure non-ethical reuse of data

Required size of training sample

Power analysis

=> Sample size required to detect hypothesised association in local cohort

Step 1 – Data collection



Informed consent

Proportionality of the data collected

Non direct re-identification of data: pseudoanonymization

Quality of the data

Representativeness of the analysis population/ target population + prevention of discrimination

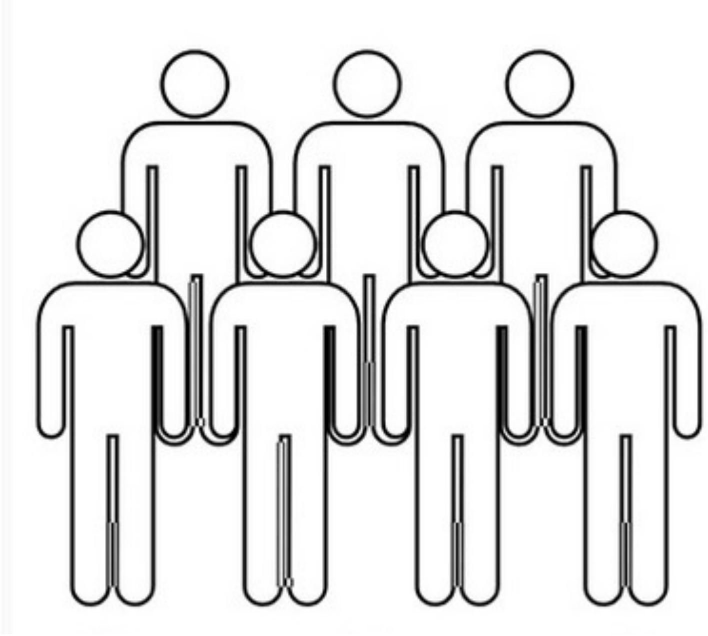
User involvement

Secure data transfer + Quality of data hosting + State of the art cyber security

Measures to ensure non-ethical reuse of data

- Recruit participants from **European** as well as **diverse ancestry present in Australia**
- Recruit **males and females**

Step 1 – Data collection



Informed consent	Proportionality of the data collected
Non direct re-identification of data: pseudoanonymization	Quality of the data
Representativeness of the analysis population/ target population + prevention of discrimination	User involvement
Secure data transfer + Quality of data hosting + State of the art cyber security	Measures to ensure non-ethical reuse of data

Meetings with people involved in project: researchers, geneticists, clinicians…

Step 2 – data pre-treatment

Handling missing data (bias reduction)

Data segregation (representativeness of the learning and evaluation sample)

Rebalancing of minority populations (bias reduction)

Expert and users involvements

Missingness in **MDD diagnosis** (bias, stigma?)

Step 2 – data pre-treatment

Handling missing data (bias reduction)

Data segregation (representativeness of the learning and evaluation sample)

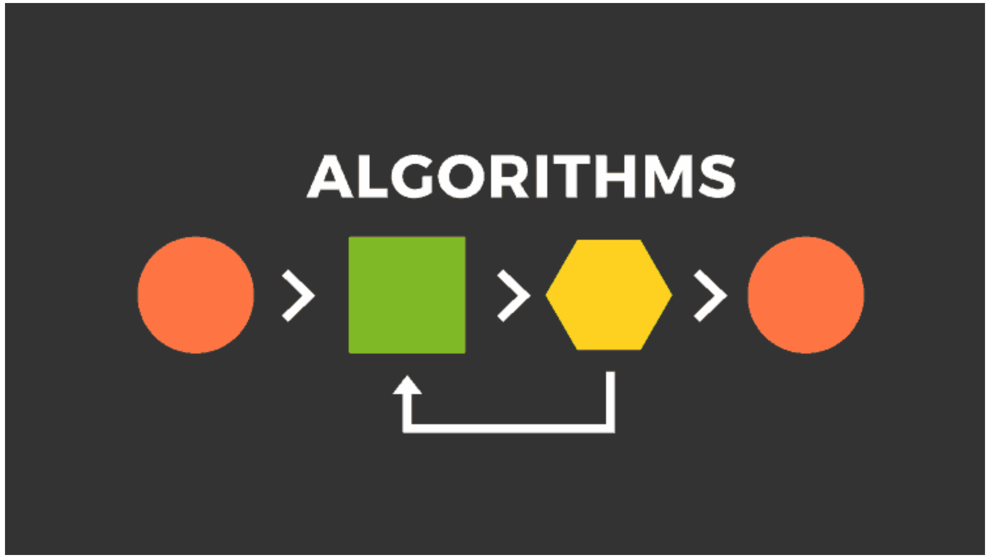
Rebalancing of minority populations (bias reduction)

Expert and users involvements

Include **covariates in** analysis (age, sex, site, MRI machine)

Oversampling in training

Step 3 – Algorithm selection



Quality policy

Transparency measures

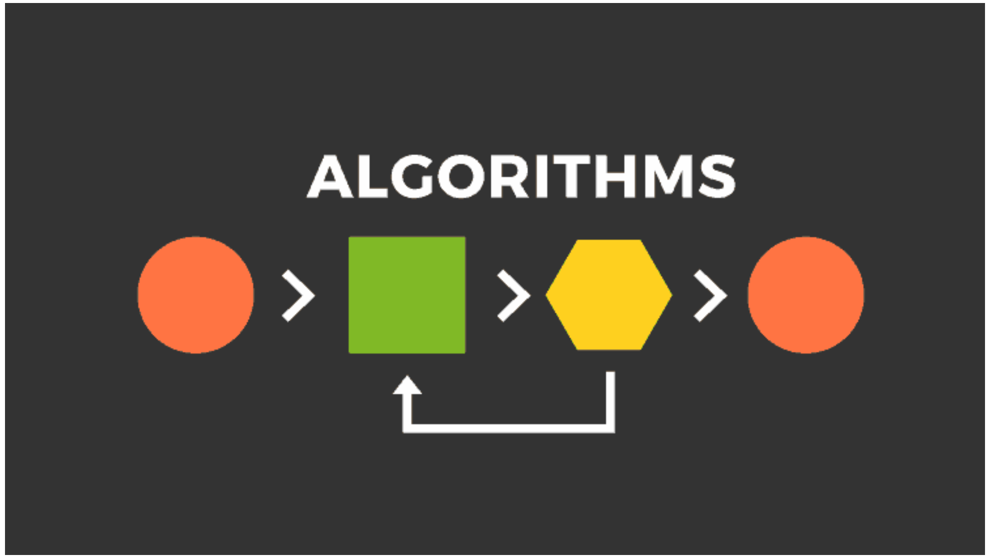
Traceability of the algorithm construction process

Explainability policy for explainable results, auditability

Use State of the art algorithm OR Pre-select **algorithms to benchmark** (to avoid overfitting test data)

Choose a **pertinent metric to evaluate algorithm**

Step 3 – Algorithm selection



Quality policy

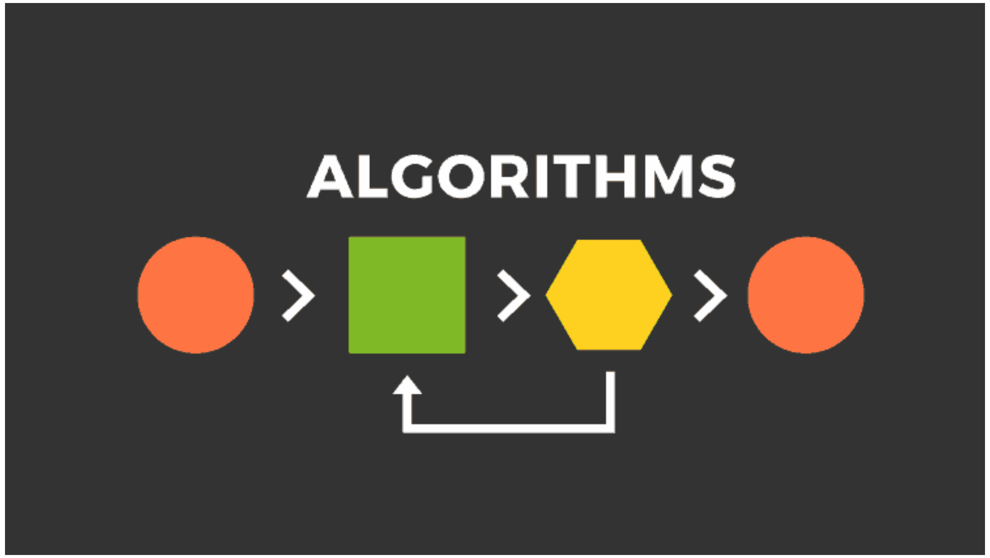
Transparency measures

Traceability of the algorithm construction process

Explainability policy for explainable results, auditability

Keep track of all decisions to help reproducibility

Step 3 – Algorithm selection



Quality policy

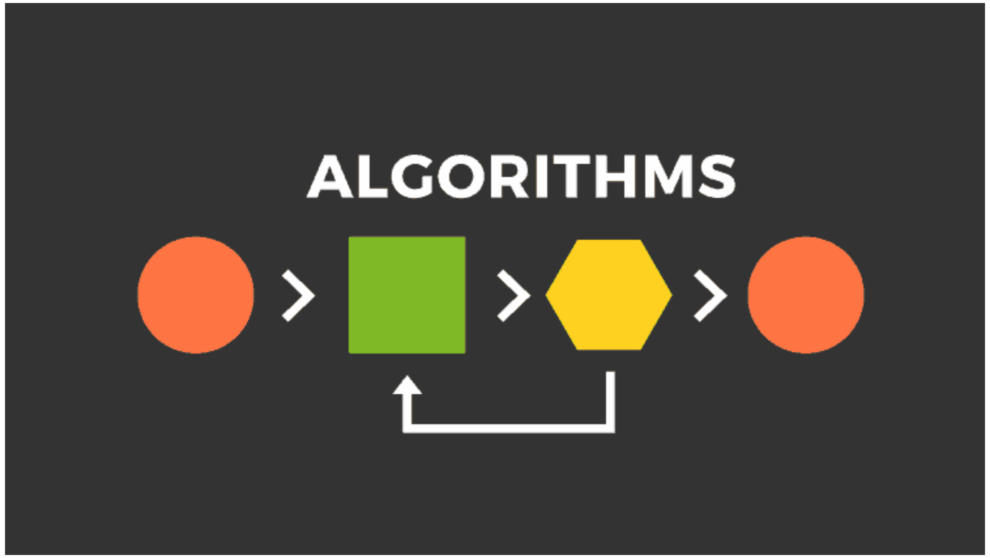
Transparency measures

Traceability of the algorithm construction process

Explainability policy for explainable results, auditability

Code stored and versioned (e.g. Github)

Step 3 – Algorithm selection



Quality policy

Transparency measures

Traceability of the algorithm construction process

Explainability policy for explainable results, auditability

Explain **why some individuals have a high estimate of genetic risk** & how risk is calculated

Step 4 - Evaluation of the algorithm before the production phase

Technical (bugs), clinical (accuracy score)	Usability
Non-discrimination	Robustness/ reproducibility
Information (fair and equal) of the users	Procedure in the event of a cyber-attack
Human guarantees	Regulatory body (audit, label)
Impact on the care pathway	Societal impacts, impacts on people or on the environment?

Step 4 - Evaluation of the algorithm before the production phase

Technical (bugs), clinical (accuracy score)	Usability
Non-discrimination	Robustness/ reproducibility
Information (fair and equal) of the users	Procedure in the event of a cyber-attack
Human guarantees	Regulatory body (audit, label)
Impact on the care pathway	Societal impacts, impacts on people or on the environment?

Evaluate algorithm performance and report results

- In each sex group

Step 4 - Evaluation of the algorithm before the production phase

- Technical (bugs), clinical (accuracy score)
- Usability
- Non-discrimination
- Robustness/ reproducibility
- Information (fair and equal) of the users
- Procedure in the event of a cyber-attack
- Human guarantees
- Regulatory body (audit, Ethique-IA label)
- Impact on the care pathway
- Societal impacts, impacts on people or on the environment?

All details of analysis in **publication / readme** that accompanies score

Key recommendations to be implemented
Responsibilities of researchers

1. Use the Ethics by design checklist
2. Limit as much as possible the potential biases
3. Define, at conception, criteria that will allow to evaluate the new tool
4. Implement necessity of human control at all stages if a medical use is targeted
5. Provide explainability as much as possible to create trust
6. Promote the development of justifications supporting the promises
7. Promote the use of clinical trials to assess the benefit-risk ratio if appropriate
8. Publish negative research results
9. Clearly state performance and limitations

Thank you

Special thanks to Pr. Ségolène Aymé – head of the Ethics Committee of the Paris Brain Institute for her contributions to this presentation.

Thank you to the funding agencies : NHMRC (CJ Martin Fellowship) and INRIA.