Object Oriented and Agile Software Development

Part 3: Social and Environmental Challenges of Software Development

Christine Solnon

INSA de Lyon - 4IF - 2022/2023

Overview

1. **Trustworthy softwares**
   - Transparency and explainability
   - Fairness
   - Robustness, security and privacy
2. **Sustainability issues**

What says the law?

Article 1 of the french law "informatique et libertés, 1978":

L'informatique doit être au service de chaque citoyen.(...) Elle ne doit porter atteinte ni à l'identité humaine, ni aux droits de l'homme, ni à la vie privée, ni aux libertés individuelles ou publiques.

GDPR (General Data Protection Regulation), 2018:

- Guidelines for collecting and processing personal information from individuals
- Right not to be subject to a decision based solely on automated processing
- Right to obtain information on the logical behaviour of algorithms
- Right to contest decisions

Why studying ethical questions?

Many decisions are taken (or recommended) by softwares:

- Assess potential recidivism risk (COMPAS)
- Assign students to universities wrt wishes
- Choose informations displayed by a social network / search engine
- ...

New sensitive applications become possible:

- Transport, Justice, Medecine, ...

$\rightsquigarrow$ Frontier between recommendation and autonomous decision is narrow

Report from the CNIL on ethical challenges of AI (2017)

(...) besoin d'une réflexion collective sur un pacte social dont certains aspects essentiels (...) peuvent être remis en question dès lors que l'évolution technologique déplace la limite entre le possible et l'impossible et nécessite de redéfinir la limite entre le souhaitable et le non souhaitable.

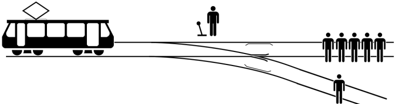


Image from Wikipedia

Some references

- CNIL : *Comment permettre à l'homme de garder la main ? les enjeux éthiques des algorithmes et de l'intelligence artificielle*, 2017
- *ACM statement on algorithmic accountability*, 2017
- C. Castelluccia, D. Le Métayer : *Understanding algorithmic decision-making : Opportunities and challenges*. European Parliamentary Research Service, 2019
- High-Level Expert Group on AI: *Ethics guidelines for trustworthy AI*, 2019
- C. Solnon : *Ethique des algorithmes*, Bulletin de la ROADEF, 2021
- T. Kirat, O. Tambou, V. Do, A. Tsoukiàs : *Fairness and Explainability in Automatic Decision-Making Systems. A challenge for computer science and law*, 2022

Overview

1. **Trustworthy softwares**
   - Transparency and explainability
   - Fairness
   - Robustness, security and privacy
2. **Sustainability issues**

What difference between transparency et explainability?

Transparency: understand the general behaviour

- Code and/or documents that describe the behaviour
- Features and origin of used Data

Explainability: understand the result of a particular run

- Identify the reasons that led to the result
- Explain these reasons in a way that is intelligible to human beings
  $\leadsto$ Exhibit counterfactuals, answer What-if questions, ...

$\leadsto$ *Explainable AI (XAI)*

Example: Admission Post-Bac (APB)

- Transparency: understand the behaviour to enter the right whishes
- Explainability: ask for explanations to contest a decision

CNIL decision (August, 30, 2017): APB does not respect the 1978 law "Informatique et Libertés"

Some challenges

Confidentiality

- Opening Data may not be compatible with privacy issues
- Opening code may not be compatible with international trade agreements (see WTO negotiations on e-commerce)

Interpretability

- Opening code and Data may not be enough to explain
- Can we explain by observing runs of a black box?

Reproducibility

- Why reproducing a run?
  For post-hoc explanations, reproduce bugs, check publish results, ...
- Why is it difficult?
  Needs complete transparency + reproducible environment

Overview

1. **Trustworthy softwares**
   - Transparency and explainability
   - Fairness
   - Robustness, security and privacy
2. **Sustainability issues**

Fairness

Softwares are often used to discriminate:

- Rank students that apply to a same university
- Select people that ask for credits
- Customise answers of recommendation services
- etc

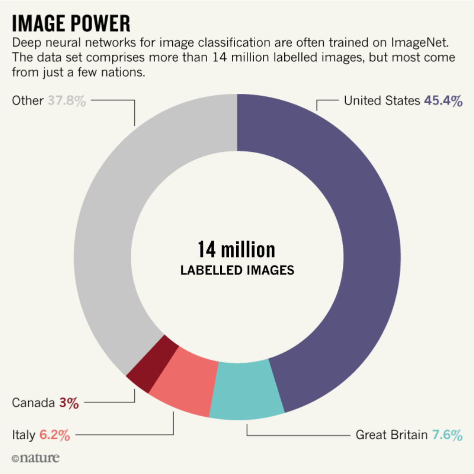
Advantages:

- The same algorithm is used for every body
- It is possible to check fairness if the algorithm is transparent
- It is possible to explain decisions if the algorithm is explainable

Risks:

- Personalization biases the perception of the world and reduces diversity of thought (especially if the algorithm is not transparent)
- How to ensure that non-relevant information is not used?
- How to ensure fairness when rules are learned?

Example 1: ImageNet



Which of these two photos represents a bride?

Training Data should fairly represent all individuals!

Zou and Schiebinger: *AI can be sexist and racist — it's time to make it fair*, in Nature 559, 324-326 (2018)

Example 2: Word Embedding (*e.g.*, Word2vec)

Basic idea of word embedding:

- Embedding of words into vectorial spaces: words that occur in similar contexts have close embeddings in the space
- May be used to infer analogies
  $\rightsquigarrow$ *king* - *man* + *woman* = ?

Training on news of *news.google.com*:

- *doctor* - *man* + *woman* =
- *computer programmer* - *man* + *woman* =

Question:

Should training Data be representative of the world as it is or as we'd like it to be?

Example 2: Word Embedding (*e.g.*, Word2vec)

Basic idea of word embedding:

- Embedding of words into vectorial spaces: words that occur in similar contexts have close embeddings in the space
- May be used to infer analogies
  $\rightsquigarrow$ *king* - *man* + *woman* = *queen*

Training on news of *news.google.com*:

- *doctor* - *man* + *woman* = ?
- *computer programmer* - *man* + *woman* = ?

Question:

Should training Data be representative of the world as it is or as we'd like it to be?

Example 2: Word Embedding (*e.g.*, Word2vec)

Basic idea of word embedding:

- Embedding of words into vectorial spaces: words that occur in similar contexts have close embeddings in the space
- May be used to infer analogies
  $\rightsquigarrow$ *king* - *man* + *woman* = *queen*

Training on news of *news.google.com*:

- *doctor* - *man* + *woman* = *nurse*
- *computer programmer* - *man* + *woman* = *homemaker*

$\rightsquigarrow$ Analogies reproduce Data biaises

Bolukbasi et al: *Man is to Computer Programmer as Woman is to Homemaker? Debiasing Word Embeddings*, in NIPS 2016

Question:

Should training Data be representative of the world as it is or as we'd like it to be?

Overview

1. **Trustworthy softwares**
   - Transparency and explainability
   - Fairness
   - Robustness, security and privacy
2. **Sustainability issues**

Robustness and security

A software should be correct wrt its specification

- Test $\leadsto$ Detect errors
- Certification $\leadsto$ Prove that a given run is correct
- Correctness $\leadsto$ Prove that all possible runs are correct

In some cases, error is part of the specification!

- The user must be aware of the error rate $\leadsto$ Transparency
- It should be possible to check the error rate $\leadsto$ Reproducibility

A software should retain its properties in the event of an attack

- It should not be possible to modify the software behaviour
- A non authorized person should not be able to interrupt a software
- ...

$\leadsto$ *Security by Design*

Privacy

Personal Data

Two kinds of Data:

- Data provided by users
- Traces of use of an application by a user

These Data must be protected...
... and should not be used without the users' consent

Anonymizing Open Data

- Make the re-identification impossible...
- ...without altering Data

Example: Cambridge Analytica

- Facebook users answer a personality test
- Their Data are used to customise political messages

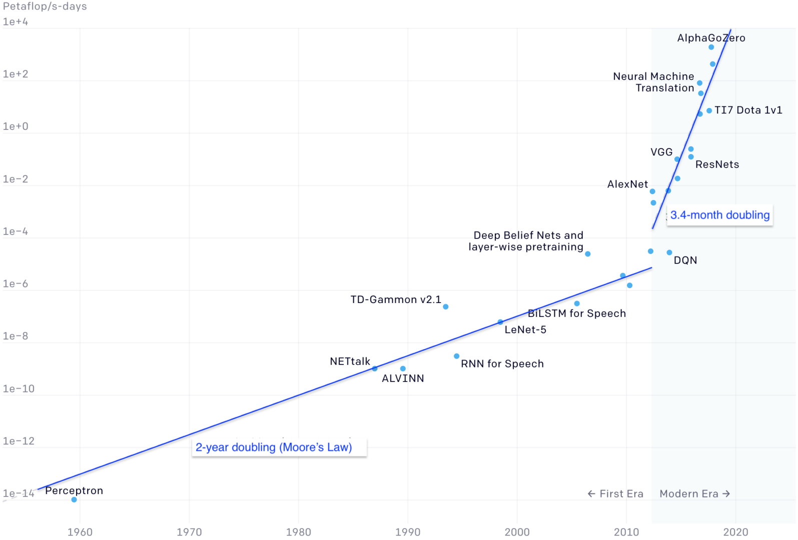
Overview

1. **Trustworthy softwares**
2. **Sustainability issues**

Some references

- K. Marquet, F. Berthoud, J. Combaz: *Introduction aux impacts environnementaux du numérique*, Bulletin de la SIF, 2019
- Shift Project: *Déployer la sobriété numérique*, 2020
- *Référentiel général d'écoconception de services numériques*, 2021
- Institut du Numérique Responsable: *Handbook of sustainable design of digital services*, 2021
- C. Bonamy, C. Boudinet, L. Bourgès, K. Dassas, L. Lefèvre, F. Vivat: *Je code : les bonnes pratiques en écoconception de service numérique à destination des développeurs de logiciels*, 2021
- J.-J. Valette: *Le numérique peut-il devenir Low-tech ?*, 2022
- Designers éthiques: *Le guide d'éco-conception de services numériques*, 2022
- ADEME: *Evaluation de l'impact environnemental du numérique en France et analyse prospective*, 2022

Example 1: Training of AIs (OpenAI)



Rebound effect:

Improving efficiency doesn't reduce consumption, but increases use

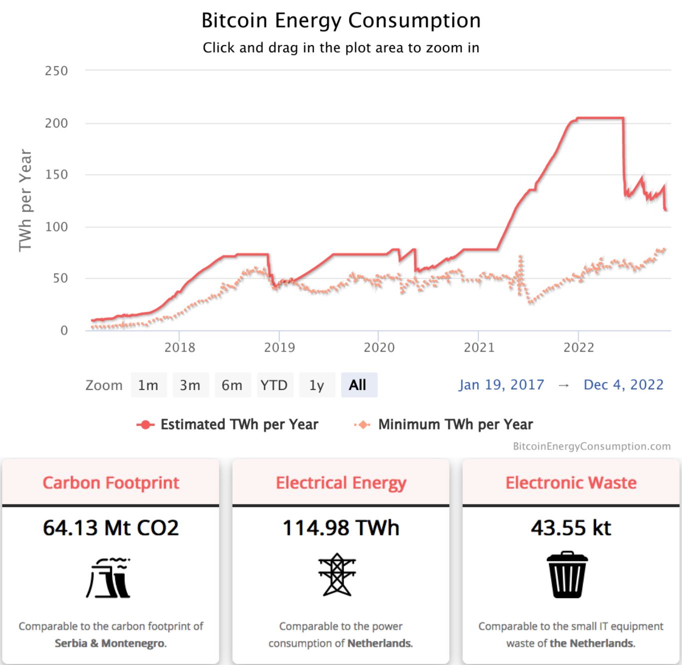
Example 1 (continued): Training vs Tuning

Consumption	CO_2e (lbs)
Air travel, 1 passenger, NY↔SF	1984
Human life, avg, 1 year	11,023
American life, avg, 1 year	36,156
Car, avg incl. fuel, 1 lifetime	126,000
Training one model (GPU)	
NLP pipeline (parsing, SRL)	39
w/ tuning & experimentation	78,468
Transformer (big)	192
w/ neural architecture search	626,155

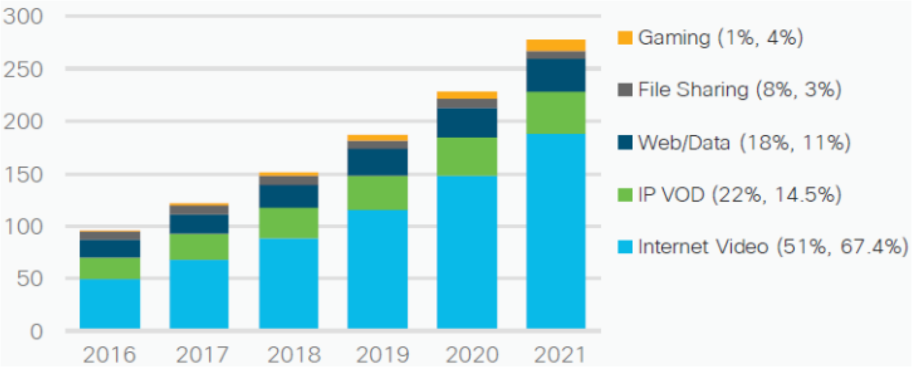
Reference:

Strubell, Ganesh, & McCallum (2019): Energy and Policy Considerations for Deep Learning in NLP

Example 2: Bitcoin (Digiconomist)



Example 3: Network traffic (The Shift Project)

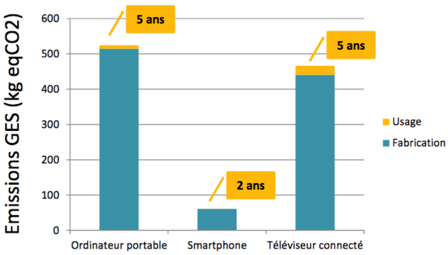


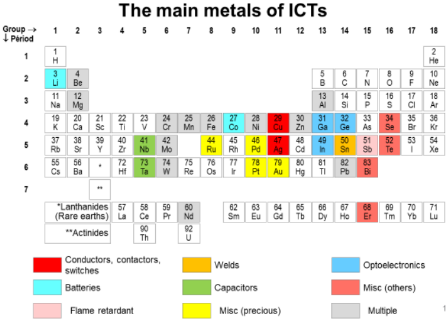
$\rightsquigarrow$ Increase of 25% per year (rebound effect, again...)

Use vs manufacturing costs (The Shift Project)

Digital services are not virtual....................They are mostly material!

- Information.......................................Sensors, Disks, etc
- Algorithms.......................................Computing centers, etc
- Human-Computer Interactions.........Terminals, End user devices, etc
- Communication..............................Network infrastructure, etc





ACV des équipements et infrastructures numériques (ADEME 2022)

Objectif : Donner une matérialité aux activités numériques

- Identifier les principaux contributeurs d'impacts
- Identifier les leviers d'amélioration les plus significatifs
- Permettre un suivi des performances dans les prochaines années
- Communiquer de manière objective
- ...

Périmètre technologique : 3 tiers

- Terminaux utilisateur (ordinateurs, écrans, box, objets connectés, ...)
- Réseau (fixe, mobile, backbone)
- Datacenters (hébergement et traitement des données)

Périmètre temporel et géographique : 2020 en France

$\leadsto$ Infrastructure à l'étranger exclue, même si utilisée en France...

ACV des équipements et infrastructures numériques (ADEME 2022)

Résultats normalisés (tableau 62) :

Nombre de personnes générant le même niveau d'impact (répartition homogène / habitants de la Terre) :

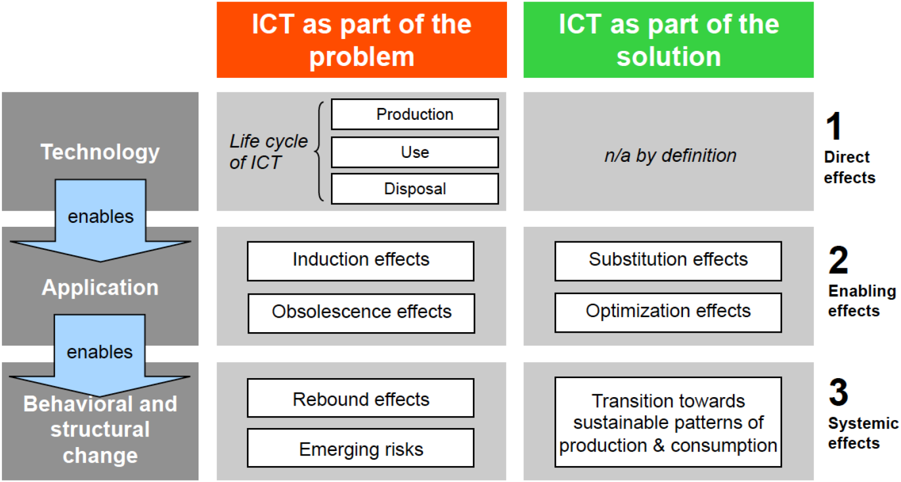
- Epuisement ressources abiotiques naturelles - éléments : 14 902 946
- Epuisement ressources abiotiques naturelles - fossiles : 12 252 844
- Ecotoxicité : 6 170 042
- Changement climatique : 2 093 266
- ...

Décomposition des impacts par tier (tableau 66) :

	Epuisement des ressources abiotiques naturelles - éléments	Epuisement des ressources abiotiques naturelles - fossiles	Acidification	Ecotoxicité	Changement climatique	Radiations ionisantes	Emissions de particules fines	Création d'ozone photochimique	MIPS	Production de déchets	Consommation d'énergie primaire	Consommation d'énergie finale (usage)
TIER 1 – Terminaux utilisateur	91,8%	68,4%	79,9%	83,0%	78,7%	67,7%	71,5%	80,0%	78,0%	86,7%	66,3%	63,6%
TIER 2 - Réseaux	4,1%	11,6%	4,9%	1,9%	5,5%	13,0%	9,8%	5,2%	6,8%	5,5%	12,4%	14,1%
TIER 3 – Centres de données	4,1%	20,0%	15,2%	15,0%	15,9%	19,3%	18,7%	14,7%	15,2%	7,9%	21,2%	22,3%

OK, but these costs should be compared to ICT benefits!

⤳ **Matrix of ICT effects (Hilty and Aebischer, 2015)**



Low tech digital services?

Specificity of computer science:

- Computers can hardly be low tech
- All other technologies strongly depend on computers

Low-technicalization of digital services

- Reduce functionalities and complexity
- Make it more user-friendly
- Limit consumptions

Low-technicalization of other technologies thanks to digital services

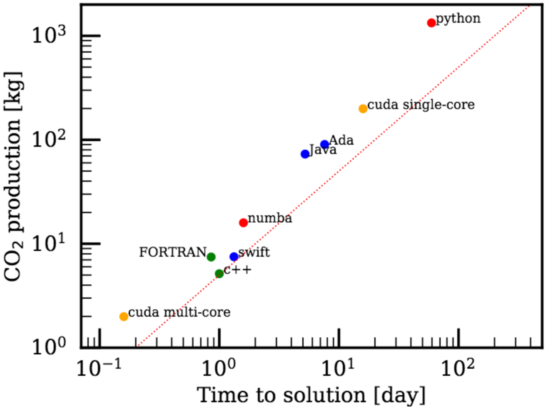
- Use ICT to produce more efficient and user-friendly objects

Beware of rebound effects!

Let's explore together!

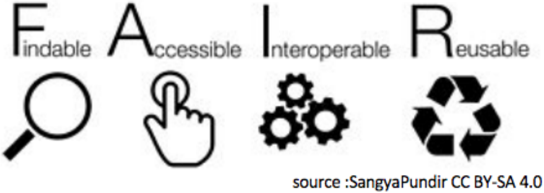
- *Référentiel général d'écoconception de services numériques*, 2021
- *Guide de Référence de Conception Responsable de Services Numériques*, 2021
- Designers éthiques: *Le guide d'éco-conception de services numériques*, 2022

Choice of a language (S. P. Swart, 2009)



- Favour compiled languages for heavy treatments
- Otherwise, favour interpreted languages to improve readability
  $\rightsquigarrow$ Use compiled libraries for heavy treatments (ex. NumPy)
- Favour parallelism
- Favour languages/libraries that won't require you to buy new computers

Data management: Make it FAIR



$\leadsto$ Avoid Data duplication and translation