Parcours professionnel : TAL & IA, entre recherche et ministères

Lauriane Aufrant

M2 TAL - Nanterre

5 février 2025

Lauriane Aufrant

- **2007-2014 :** classe préparatoire (MP), école d'ingénieur (X), plusieurs stages (gendarmerie, ouvrier, Systran), double diplôme (Télécom), M2 en IA/TAL (Orsay)
- **2014-2017 :** thèse de TAL au LIMSI (= LISN)
- **2018-2021 :** Direction Générale de l'Armement (Bruz, 35), chargée d'expertise en TAL puis en IA
- **2021-2023 :** Inria, chercheuse à Inria Défense & Sécurité puis à ALMAnaCH (équipe TAL / LLM)
- **2024- ..... :** Inria, cheffe de projet normalisation

2014-2017

Au commencement était la thèse

Une thèse ambitieuse

- Sur financement DGA : donc ancrée dans le réel

- **Mon sujet initial :** transfert cross-lingue pour le TAL des langues romanes
  - $\rightsquigarrow$ Et en chemin... coup de coeur pour l'analyse en dépendances
- **Mon sujet final :** le transfert cross-lingue, l'analyse en dépendances... et l'analyse en dépendances cross-lingue !

Une thèse “fondamentale appliquée”

- De nombreuses expériences, supportées par de jolies théories
- Juste une preuve de concept, mais pour un objectif très concret

  $\hookrightarrow$ Que faire avec **10-30 phrases annotées** dans une langue ?

- Et de l'interprétabilité... avant que ça existe !
- De la recherche en faisceau : exploration d'espace de recherche pour des séquences d'actions
- Des oracles dynamiques (idées proches du renforcement, imitation learning...)

Mes (vraies) conclusions de thèse

- Apprendre avec beaucoup de mauvaises données, ou avec peu de **bonnes** données ?
- Ne pas apprendre ce que l'on sait déjà $\rightsquigarrow$ **IA hybride**
- Plus-value de la **complémentarité** des sources d'information
- De l'importance de prendre du **recul** sur les méthodes
  $\hookrightarrow$ voir la "big picture", les similarités... et les synergies

La vie de thésarde

- Sans enseigner... car pas (encore) de projet académique
- Dans un domaine très mouvant
- Collaborer, échanger... la science se fait à la machine à café
- Savoir visiter la salle voisine : d'innombrables opportunités pour apprendre et découvrir !

Quitter la recherche ?

- Des hauts et des bas : compétitivité, frustrations, arbitraire...
- Les chercheurs ne sont pas forcément ceux qui cherchent
- Une formation avant tout, pas une fin en soi

  ... et puis la voie de facilité (cf. financement fléché)

2018-2021
L'expertise technique en ministère

Chargée d'expertise sur des programmes d'armement

- Travailler sur des **contrats** (code des marchés publics)... car derrière chaque signature réussie se trouve un ingénieur : recueil et analyse de besoins, spécifications techniques, suivi de projet, essais et qualifications...
  - Être à la pointe de la science, mais sans en faire partie

- Grande **diversité** des cas d'usage : texte, parole, OCR... traduction, transcription, synthèse vocale, résumé, extraction d'information, identification de langue et dialecte, recherche d'information, indexation sémantique... dans des avions, des chars, en bureautique, sur les réseaux sociaux... tous types de langues...
  - Un peu rude en sortie de thèse !

↪ Aussi un premier contact avec le monde **juridique** (& industriel)

↪ **Travail "en matriciel"** : jusqu'à 8 chefs !

IA de confiance & éthique

- Intégrer les **spécificités de l'IA** dans les pratiques contractuelles et de conduite de projet (spécification, développement, test...)
  - Groupe de travail pluridisciplinaire, collecte des écueils, et rédaction d'un **guide méthodologique**

- Un sujet **brûlant** : rapport Villani, recommandations éthiques des experts de la Commission européenne...
  - Soutien technique aux travaux sur l'**éthique de l'IA** dans les armées (dont SALA)

- Beaucoup d'efforts de **vulgarisation** pour des audiences non-techniques, dédramatiser, déraciner les mythes...

↪ Recentrage vers la **méthodologie**

↪ Premiers travaux en milieu **pluridisciplinaire**

2021-2023

La recherche appliquée, en collaboration avec l'industrie

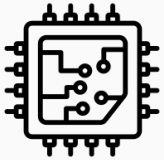
Retour vers la recherche, à Inria

- Un petit côté **“startup”**
  - ⇝ être à la fois chercheuse, développeuse web, recruteuse, chef de projet, formatrice, négociatrice, juriste...
- Dépôt de projets collaboratifs **européens et nationaux**
  - ⇝ très chronophage, longs délais, résultat aléatoire... mais un très bon exercice de planification et communication
- Partenariats **industriels**
  - ⇝ extraction d'information pour la chimie, IA pour l'aéronautique... avec des intérêts parfois divergents
- Développer **son propre projet** de recherche
  - ⇝ extraction d'information (& graphe de connaissances), avec double lien défense & méthodologie

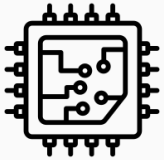
2024-...

L'Europe, la réglementation, et les politiques publiques

Quel est le point commun entre...



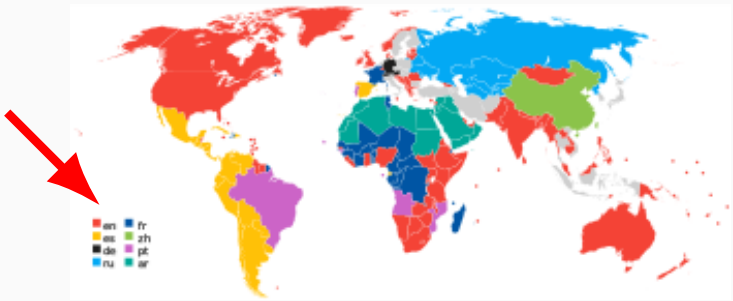
Quel est le point commun entre...



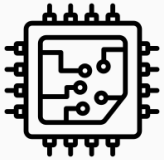
THE C PROGRAMMING LANGUAGE

mp3

```
 !"#$%&'()*+,-./0123456789:;<=>?
@ABCDEFGHIJKLMNOPQRSTUVWXYZ[\]^_
`abcdefghijklmnopqrstuvwxyz{|}~
 ¡¢£¤¥¦§¨©ª«¬-®¯°±²³´µ¶·¸¹º»¼½¾¿
ÀÁÂÃÄÅÆÇÈÉÊËÌÍÎÏÐÑÒÓÔÕÖ×ØÙÚÛÜÝÞß
àáâãäåæçèéêëìíîïðñòóôõö÷øùúûüýþÿ
```



Quel est le point commun entre...



ISO/IEC 9899

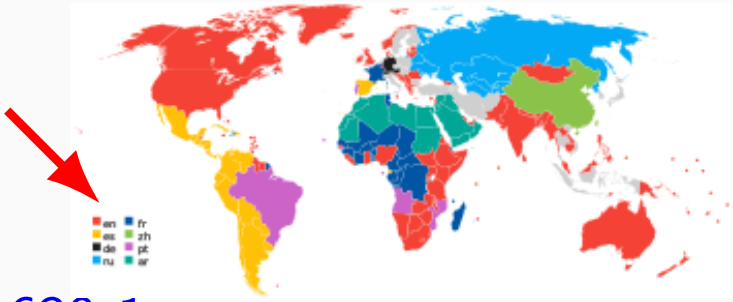
THE C PROGRAMMING LANGUAGE

mp3

ISO/IEC 11172-3

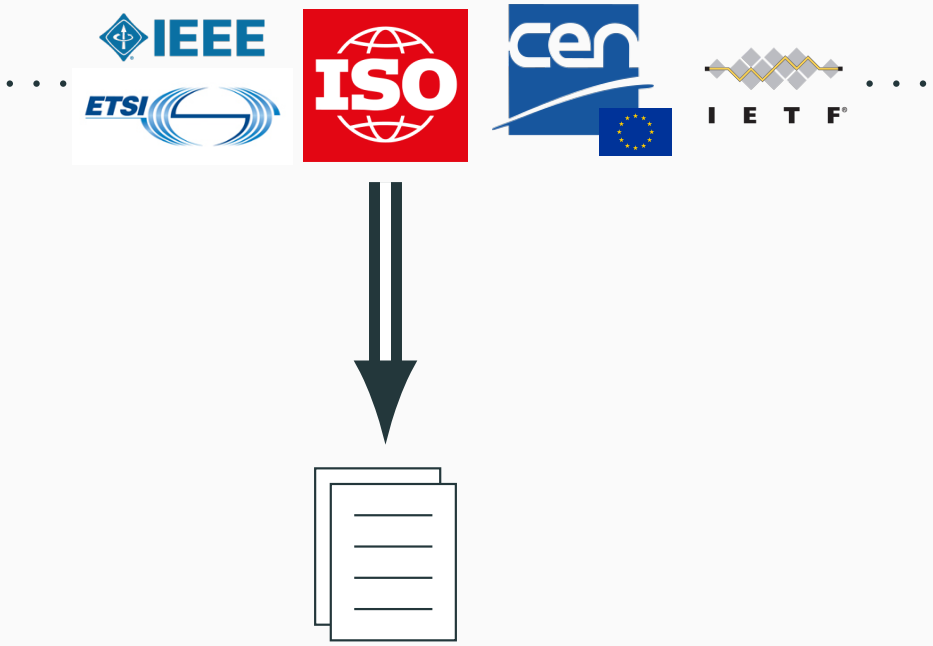
ISO/IEC 8859-1

```
 !"#$%&'()*+,-./0123456789:;<=>?
@ABCDEFGHIJKLMNOPQRSTUVWXYZ[\]^_
`abcdefghijklmnopqrstuvwxyz{|}~
 ¡¢£¤¥¦§¨©ª«¬-®¯°±²³´µ¶·¸¹º»¼½¾¿
ÀÁÂÃÄÅÆÇÈÉÊËÌÍÎÏÐÑÒÓÔÕÖ×ØÙÚÛÜÝÞß
àáâãäåæçèéêëìíîïðñòóôõö÷øùúûüýþÿ
```

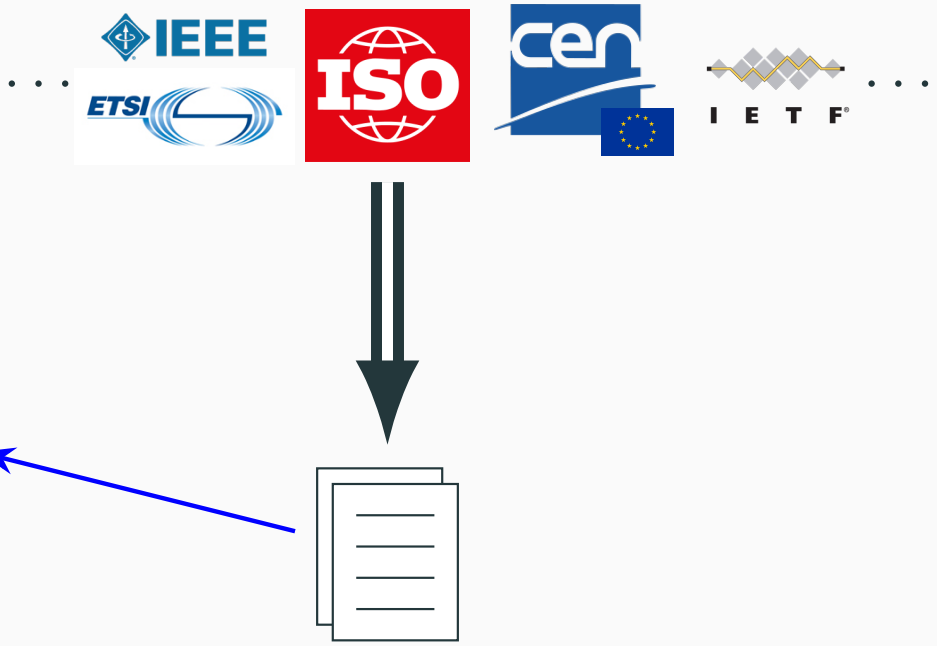


ISO 639-1

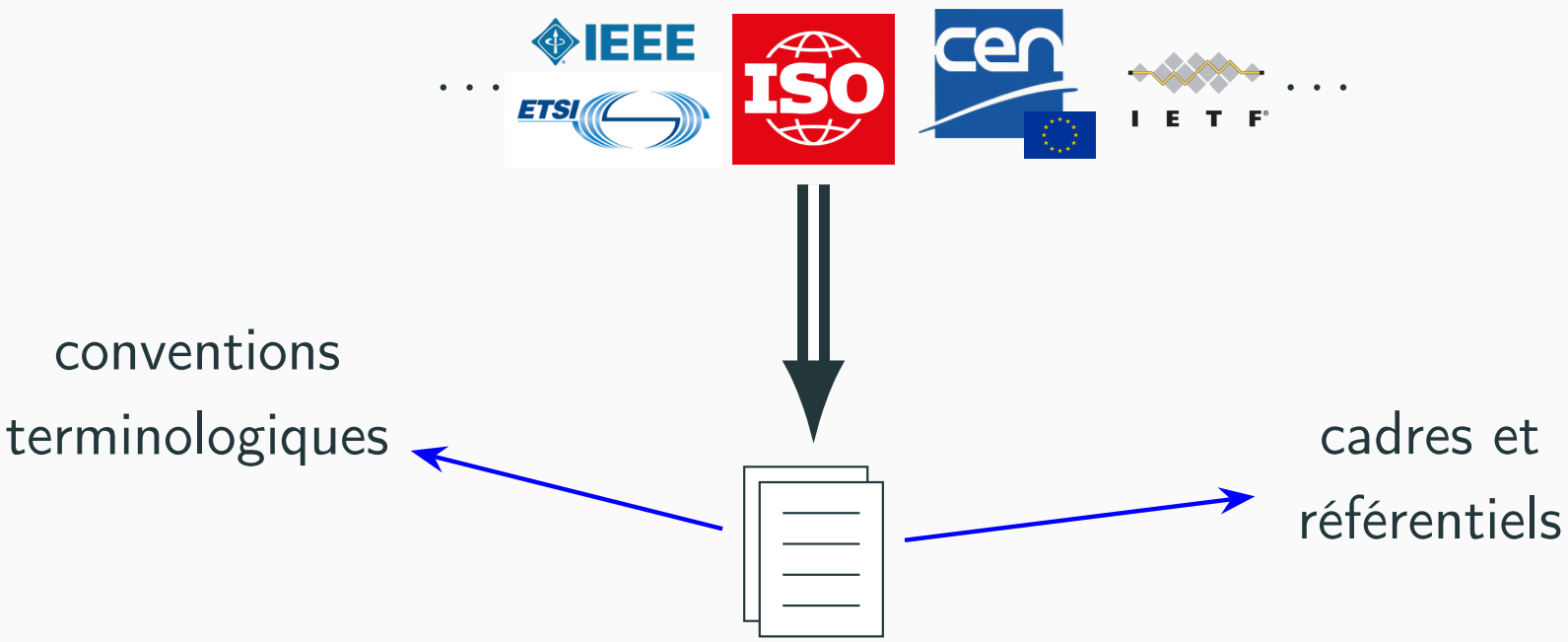
Standards



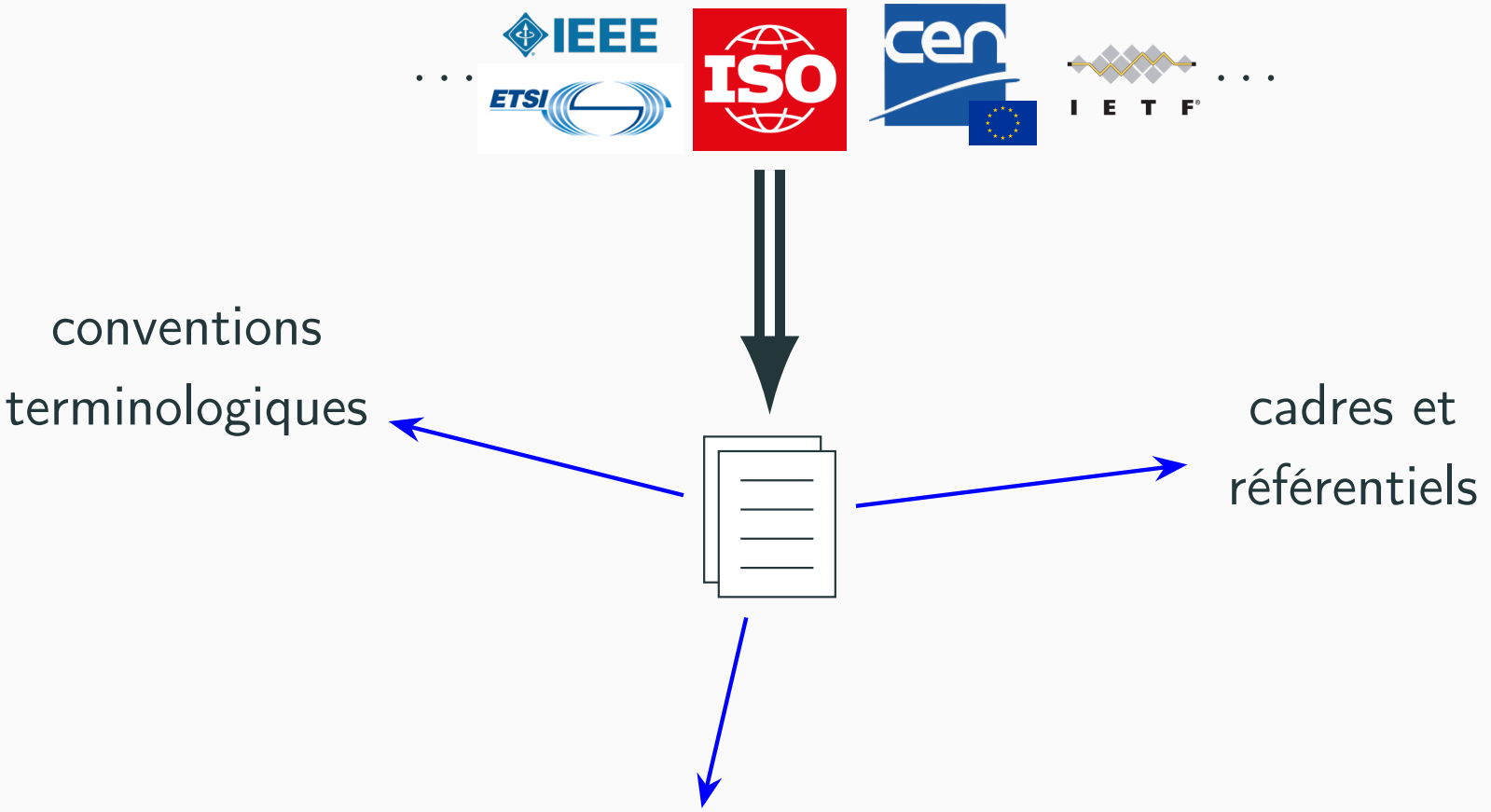
Standards



Standards


IEEE
ETSI
ISO
cen
I E T F
conventions terminologiques
cadres et référentiels

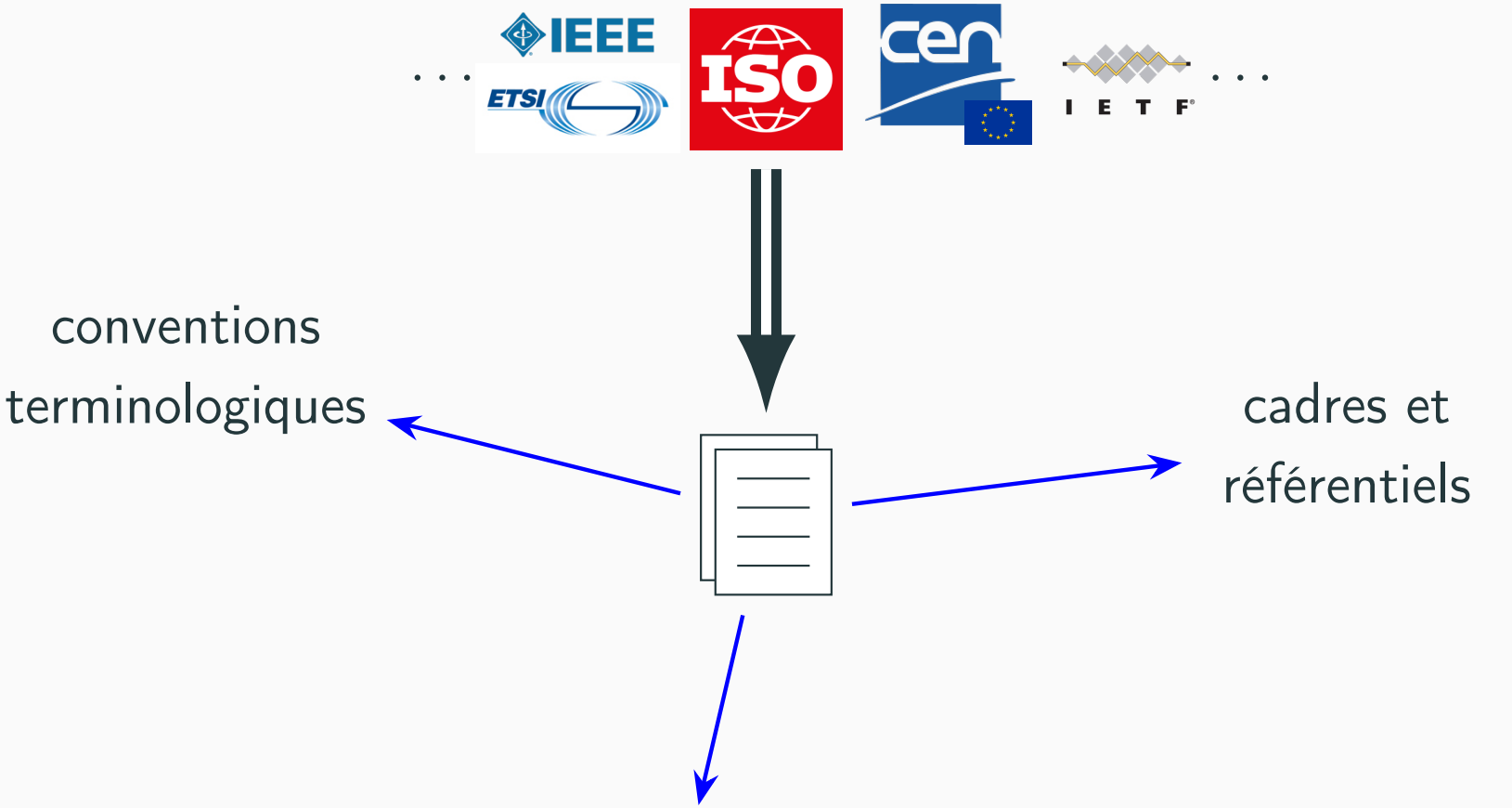
Standards

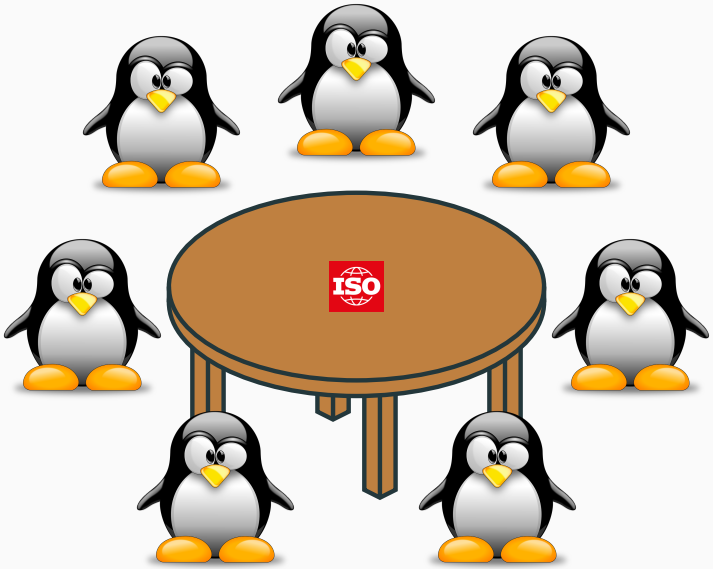


↪ pour les données, systèmes, procédures...

↪ métriques de qualité, interopérabilité (APIs, protocoles...)

Standards


IEEE
ETSI
ISO
cen
I E T F
conventions terminologiques
cadres et référentiels
spécifications techniques
↪ pour les données, systèmes, procédures...
tests de conformité
certification
↪ métriques de qualité, interopérabilité (APIs, protocoles...)


ISO


ISO


ISO

ISO/IEC 22989:2022 *Information technology – Artificial intelligence – Artificial intelligence concepts and terminology*

3.3.4
hyperparameter
characteristic of a *machine learning algorithm* (3.3.6) that affects its learning process

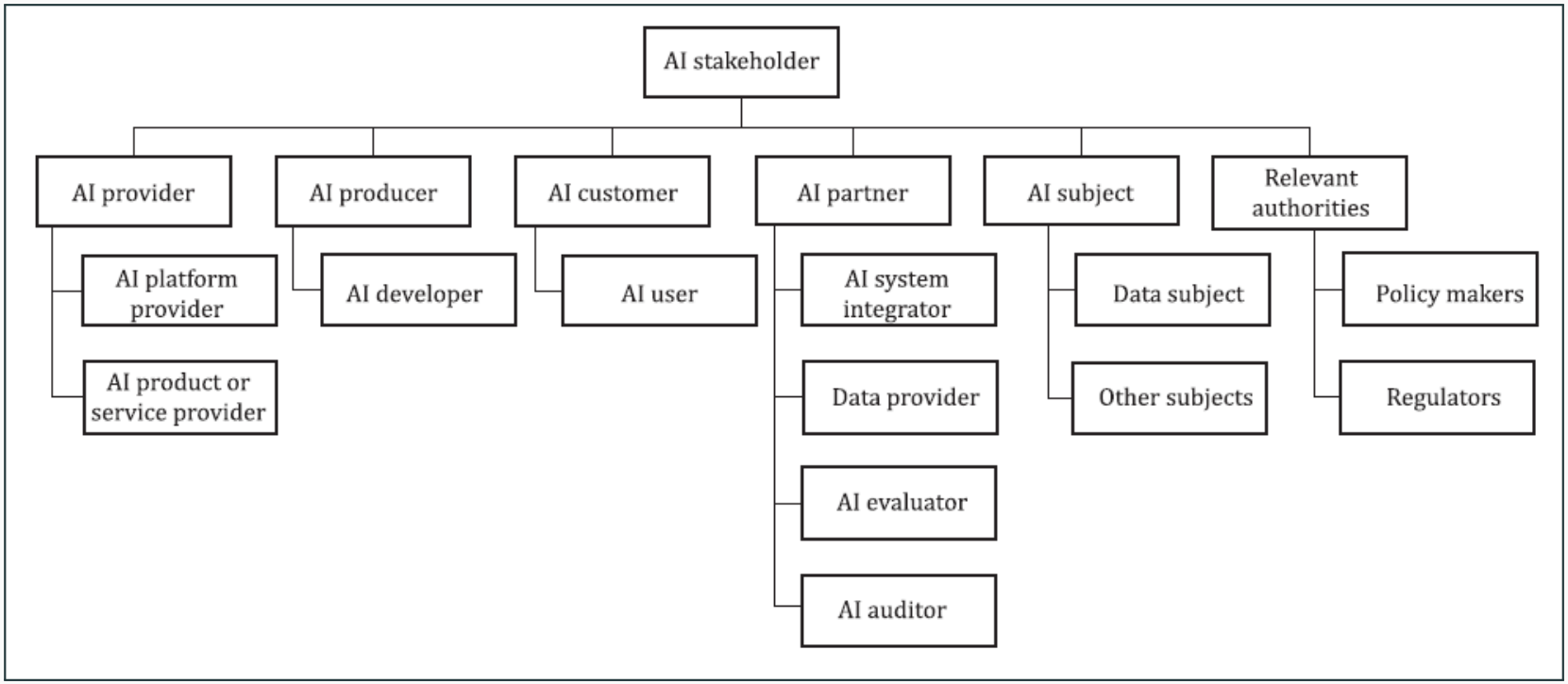
Note 1 to entry: Hyperparameters are selected prior to training and can be used in processes to help estimate model parameters.

Note 2 to entry: Examples of hyperparameters include the number of network layers, width of each layer, type of activation function, optimization method, learning rate for neural networks; the choice of kernel function in a support vector machine; number of leaves or depth of a tree; the K for K-means clustering; the maximum number of iterations of the expectation maximization algorithm; the number of Gaussians in a Gaussian mixture.

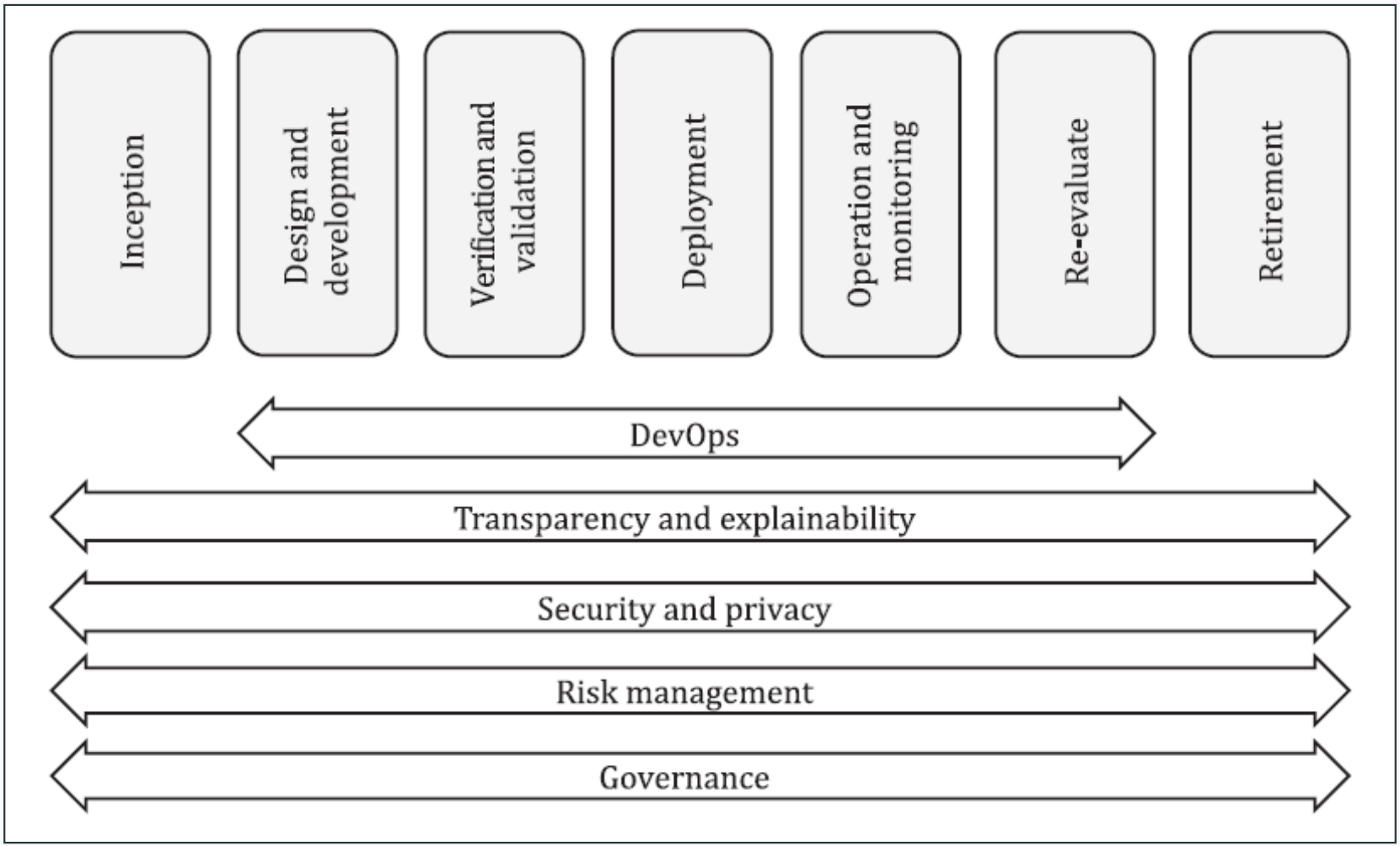
3.3.5
machine learning
ML
process of optimizing *model parameters* (3.3.8) through computational techniques, such that the *model's* (3.1.23) behaviour reflects the data or experience

3.3.6
machine learning algorithm
algorithm to determine *parameters* (3.3.8) of a *machine learning model* (3.3.7) from data according to given criteria

ISO/IEC 22989:2022 *Information technology – Artificial intelligence – Artificial intelligence concepts and terminology*



ISO/IEC 22989:2022 *Information technology – Artificial intelligence – Artificial intelligence concepts and terminology*



ISO/IEC CD TR 23281 *Artificial intelligence – Overview of AI tasks and functionalities related to NLP* (draft)

7.3.3 Speaker recognition

Speaker recognition consists in identifying the person speaking in a speech segment, by comparison with other recordings from the same person, not necessarily in the same language.

This task encompasses four distinct settings:

- Speaker clustering: Given recordings from various speakers, group all recordings from the same speaker together.
- Speaker identification: A database of speakers is available, consisting in one or more recordings for each speaker. Given a new recording from a single speaker, decide whether its speaker is in that database, and if so which one it is. This is a case of one-to-many recognition.
- Speaker verification (also called speaker authentication): Given one or more recordings from the same speaker, and another recording from a single speaker, decide whether that new recording is from the same speaker. This is a case of one-to-one recognition.
- Speaker detection: Given one or more recordings from the same speaker, and another recording (which can be from several speakers), decide whether the known speaker is present in that new recording.

ISO/IEC CD TR 23281 *Artificial intelligence – Overview of AI tasks and functionalities related to NLP* (draft)

7.4.4 Entity linking

Given a text document, an entity mention in that document and a knowledge base, entity linking consists in deciding to which entry in the knowledge base that entity corresponds. It is also known as entity disambiguation, or entity resolution.

The variant in which the entity linking system does not take the knowledge base as an input, but is designed to link entities with one knowledge base in particular, is to be reported as "knowledge base-fixed entity linking".

The term "collective entity linking" refers to the variant in which the inputs are a text collection, the set of all entity mentions in it, and a knowledge base, and the output is the set of knowledge base entries corresponding to each mention.

Entity linking differs from record linkage in that entity mentions are considered in the context of a document, whereas record linkage includes out-of-context mentions, such as database entries.

The task is defined so that any entity can be linked. Variants exist that focus on a given set of entity types, which can be identified as "type-restricted entity linking". For instance, linking can apply only to people, organizations and locations. Information on the restricted set of entity types is necessary to achieve a non-ambiguous designation of the task.

Named entity linking is a further restricted variant, which constrains both the entity types (see 7.4.3 for typical

ISO/IEC AWI 23282 *Artificial intelligence – Evaluation methods for accurate NLP systems* (draft)

6.1 BLEU

The BLEU score measures the extent to which a candidate sentence in text matches the form and content of a given reference sentence (or multiple references), accounting for terminology, phrasing, and the possibility of multiple equivalent phrasings for the same sentence.

It is defined as:

$$\mathrm{BLEU} = \mathrm{BP} \cdot \sqrt[n]{\prod_{k=1}^{n} \frac{\mathrm{TP}_{\text{n-gram}}(k)}{\mathrm{TP}_{\text{n-gram}}(k) + \mathrm{FP}_{\text{n-gram}}(k)}}$$

where $\mathrm{TP}_{\text{n-gram}}(k)$ is the number of true positives among n-grams of k tokens (with respect to one or more reference sentences), $\mathrm{FP}_{\text{n-gram}}(k)$ is the number of false positives among n-grams of k tokens, and the brevity penalty BP is defined by comparing lengths of the candidate sentence and the reference with closest length (if shorter):

$$\mathrm{BP} = e^{-max\left(0,\ \frac{L}{L_{closest-ref}}-1\right)}$$

ISO/IEC AWI 23282 *Artificial intelligence – Evaluation methods for accurate NLP systems* (draft)

The computation of BLEU can be affected by the following technical characteristics:

- Whether multiple references are used per sentence, and how many. This affects the computation of true positives and false positives. Common choices are 1, 2 and 4 references. A large number of references leads to higher BLEU scores, and more faithful evaluation.
- Whether the brevity penalty is computed separately for each sentence, or averaged over the corpus.
- The maximum n-gram length n. A common choice is 4.
- The tokenization applied to the candidate and reference sentences. For comparability, the same tokenization procedure needs to be applied in both cases. Some tokenization schemes can lead to higher or lower BLEU scores.
- Whether the computation of n-gram counts is case-sensitive (cased BLEU) or case-insensitive (uncased BLEU).
- Whether rare words are mapped to a special "unknown" token before computation.

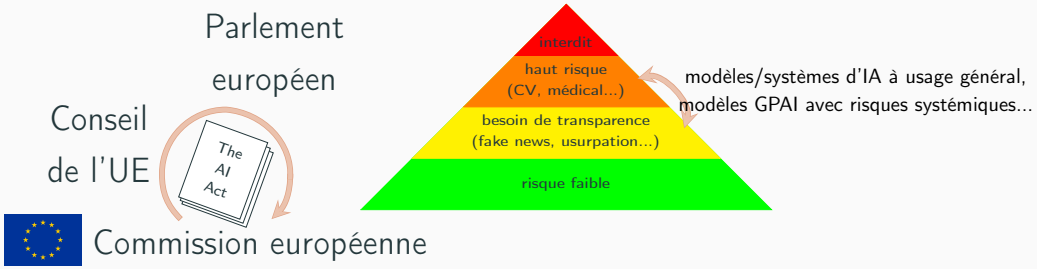
Software implementing the BLEU score shall:

Experte en standardisation pour l'IA

- Des années d'(auto-)formation (depuis 2019) avant d'être vraiment efficace : être **perséverant** avant tout !
- Entrée par la petite porte... puis cheffe d'un petit groupe... cheffe de délégation... éditrice d'un rapport, puis d'un standard... jusqu'à être **cheffe de groupe**
- Savoir s'adresser à des profils **très variés**

  $\hookrightarrow$ syndicats, consommateurs, juristes, commerciaux, ingénieurs, chercheurs...
- **Pas toujours facile** : sur Zoom avec US et Chine de 23h à 2h... lobbying, pressions...

$\hookrightarrow$ Après 2 ans de réflexions et échanges, nommée **coordinatrice** de la nouvelle feuille de route Inria sur la standardisation


Parlement
européen
Conseil
de l'UE
The
AI
Act
Commission européenne
interdit
haut risque
(CV, médical...)
besoin de transparence
(fake news, usurpation...)
risque faible
modèles/systèmes d'IA à usage général,
modèles GPAI avec risques systémiques...

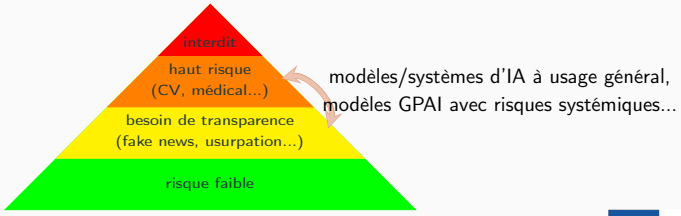
Calendrier pour les standards AI Act :

- préparatifs depuis avril 2021
- demandés dès mai 2022, pour avril 2025
- report août, mais impossible avant 2026
- entrée en vigueur : août 2026...

Parlement européen

Conseil de l'UE

Commission européenne



JTC 21

"*Intelligence artificielle*"

↪ **Requête de standardisation :**

- système de gestion du risque
- qualité des jeux de données
- journalisation
- transparence
- supervision humaine
- performance
- robustesse
- cybersecurité
- système de management de la qualité
- évaluation de la conformité

Article 40 AI Act :
norme harmonisée =
présomption de conformité

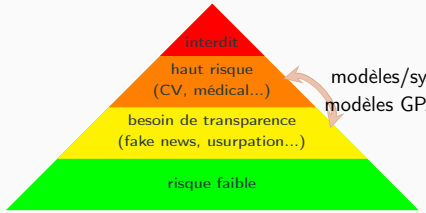
Calendrier pour les standards AI Act :

- préparatifs depuis avril 2021
- demandés dès mai 2022, pour avril 2025
- report août, mais impossible avant 2026
- entrée en vigueur : août 2026...

Parlement européen

Conseil de l'UE

Commission européenne



modèles/systèmes d'IA à usage général, modèles GPAI avec risques systémiques...

↪ **Requête de standardisation :**

- système de gestion du risque
- qualité des jeux de données
- journalisation
- transparence
- supervision humaine
- performance
- robustesse
- cybersecurité
- système de management de la qualité
- évaluation de la conformité

CEN CENELEC JTC 21

"*Intelligence artificielle*"

afnor CN IA

accord de Vienne

ISO IEC JTC 1/SC 42

"*Intelligence artificielle*"

Article 40 AI Act :
norme harmonisée =
présomption de conformité

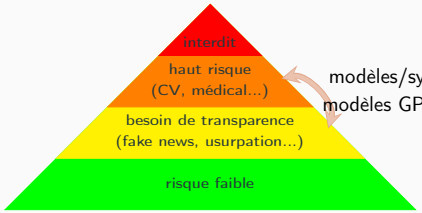
Calendrier pour les standards AI Act :

- préparatifs depuis avril 2021
- demandés dès mai 2022, pour avril 2025
- report août, mais impossible avant 2026
- entrée en vigueur : août 2026...

Parlement européen

Conseil de l'UE

Commission européenne



modèles/systèmes d'IA à usage général,
modèles GPAI avec risques systémiques...

↪ **Requête de standardisation :**

- système de gestion du risque
- qualité des jeux de données
- journalisation
- transparence
- supervision humaine
- performance
- robustesse
- cybersecurité
- système de management de la qualité
- évaluation de la conformité

CEN CENELEC JTC 21

"*Intelligence artificielle*"

afnor CN IA

accord de Vienne

accord de Vienne

ISO IEC JTC 1/SC 42

"*Intelligence artificielle*"

Joint Working Group 5

sur le TAL

ISO TC 37

"*Langue et terminologie*"

Article 40 AI Act :
norme harmonisée =
présomption de conformité

Article 15 – Accuracy, robustness and cybersecurity

1. High-risk AI systems shall be designed and developed in such a way that they **achieve an appropriate level of accuracy, robustness, and cybersecurity**, and perform consistently in those respects throughout their lifecycle.

2. To address the technical aspects of how to **measure the appropriate levels of accuracy and robustness** set out in paragraph 1 of this Article and any other relevant performance metrics, the Commission shall, in cooperation with relevant stakeholder and organisations such as metrology and benchmarking authorities, encourage as appropriate, the **development of benchmarks and measurement methodologies**.

3. The levels of accuracy and the relevant accuracy metrics of high-risk AI systems shall be declared in the accompanying instructions of use. [...]

Article 15 – Accuracy, robustness and cybersecurity

4. High-risk AI systems shall be as resilient as possible regarding **errors, faults or inconsistencies** that may occur within the system or the environment [...].

5. High-risk AI systems shall be resilient as regards to attempts by unauthorised third parties to alter their use, outputs or performance by exploiting the system vulnerabilities. [...]

The technical solutions to address AI specific vulnerabilities shall include, where appropriate, measures to prevent, detect, respond to, resolve and control for attacks trying to manipulate the training dataset ('**data poisoning**'), or pre-trained components used in training ('**model poisoning**'), inputs designed to cause the model to make a mistake ('**adversarial examples' or 'model evasion**'), **confidentiality attacks** or model flaws.

Annex IV(2) – Technical documentation

(g) the validation and testing procedures used, including information about the **validation and testing data** used and their main characteristics; **metrics** used to measure accuracy, robustness and compliance with other relevant requirements set out in Title III, Chapter 2 as well as potentially discriminatory impacts; test logs and all test reports dated and signed by the responsible persons, including with regard to pre-determined changes as referred to under point (f).

Experte en standardisation pour l'AI Act

- **Comprendre** l'IA (toute l'IA), la standardisation, les standards IA, comprendre l'AI Act... et comprendre le cadre légal des standards réglementaires

- Travail direct avec la **Commission européenne** (AI Office) : bâtir une relation de confiance, sur le long terme

- Parfois un peu le vertige ! ... mais je sais pourquoi je suis là

Stratégie nationale IA

- Rôle d'Inria dans la **stratégie IA** dès 2018, nommée "Agence de programme du numérique" en 2023 $\rightsquigarrow$ **Programme IA**
- Création de l'**INESIA** (Institut national pour l'évaluation et la sécurité de l'intelligence artificielle) : il y a quelques jours !
- Coordinatrice du projet NoLeFa : projet pilote pour un **centre d'expertise européen** pour les autorités de surveillance de marché de l'AI Act

$\hookrightarrow$ Mise à disposition Programme IA à 40%. Donc plusieurs chefs...

5 février 2025

Bilan de ce (début de) carrière

- Un fil conducteur : **l'IA**... et surtout le TAL
- De multiples allers-retours entre **recherche et ministères**
- Combiner **souveraineté et international**
- Une **pluridisciplinarité** omniprésente

Quelques principes utiles

- Être ouvert et suivre les **opportunités** : un plan c'est fait pour changer
- Nouer des **contacts**, et les garder : le TAL est un tout petit monde (même l'IA !)
- On peut utiliser sa **science** dans un rôle non-scientifique... et il en faut !

Des questions ?

first.last@inria.fr