



**LABORATOIRE
HUBERT CURIEN**

UMR • CNRS • 5516 • SAINT-ETIENNE



L'intelligence artificielle et ses enjeux à la lumière de la mythologie grecque

L'UNIVERSITÉ POUR TOUS

Fabrice MUHLENBACH

<https://perso.univ-st-etienne.fr/muhlfabr/>

e-mail : fabrice.muhlenbach@univ-st-etienne.fr



Introduction : comprendre l'IA grâce aux mythes ?

- Les mythes

- **Définition :**

- construction imaginaire qui se veut explicative des phénomènes cosmiques, psychologiques et sociaux
- utilisation de divinités pour proposer une explication de certains aspects fondamentaux du monde et de la société
- le mythe est porté à l'origine par une tradition orale
- fondation d'une pratique sociale en fonction des valeurs fondamentales d'une communauté à la recherche de sa cohésion

- **Caractéristiques :**

- existence de différentes versions et interprétations du même mythe
- création du monde (cosmogonie), espèce humaine (anthropogonie), divinités (théogonie), phénomènes naturels, statut de l'être humain...₂

Introduction : comprendre l'IA grâce aux mythes ?

- Les mythes

- **La forme symbolique associée aux mythes :**

- Dans la mythologie grecque, quand il y a un trouble dans l'ordre cosmique, il faut des générations pour le réparer

- Quand il y a une punition, un châtement, quel qu'il soit, il a toujours un rapport symbolique avec le crime

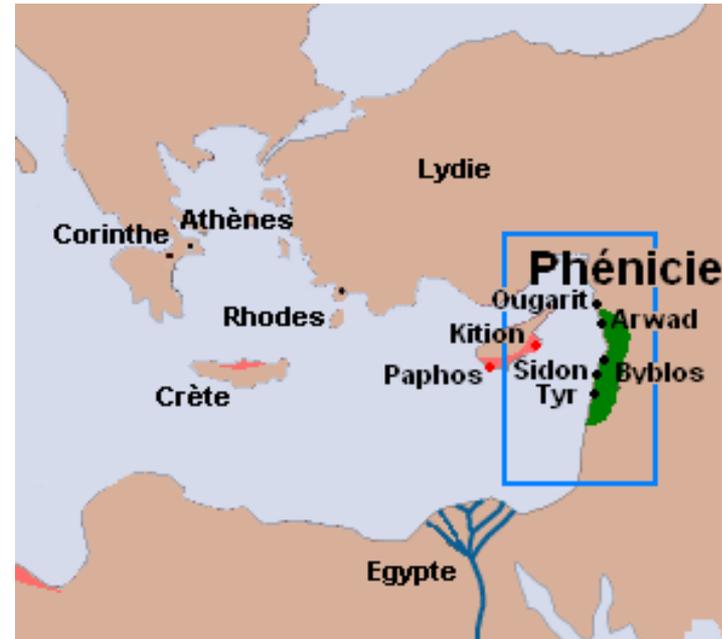
- Le symbole, étymologiquement *symbolon* (σύμβολον), est un signe de reconnaissance constitué d'une pièce d'argile séparée en deux parties dont les deux morceaux s'ajustent parfaitement

L'origine des Labdacides

De la révolution numérique aux biais algorithmiques

- La fondation mythique de Thèbes

- L'histoire débute sur une plage de Tyr, dans le sud de la Phénicie...
- Un jour, Zeus découvrit la ravissante Europe et il tomba sous son charme



L'origine des Labdacides

De la révolution numérique aux biais algorithmiques

- La fondation mythique de Thèbes



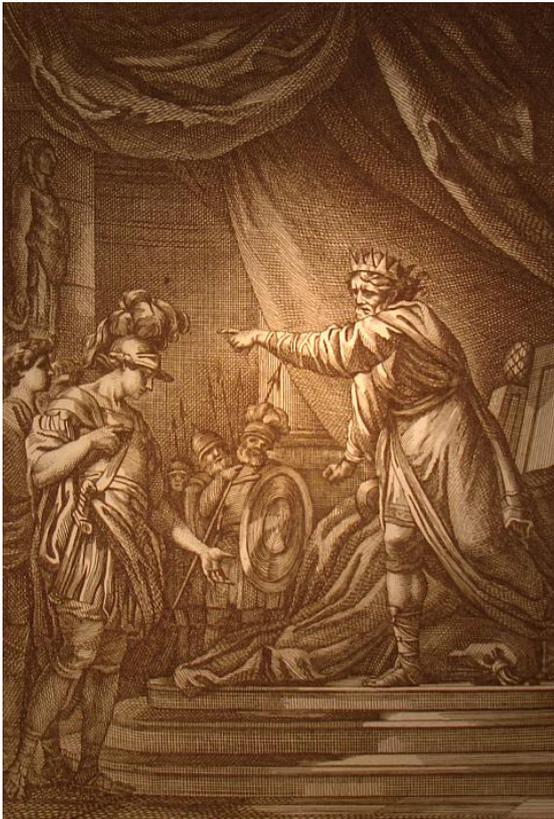
- Se métamorphosant en taureau blanc afin de ne pas effrayer la princesse phénicienne et d'échapper à la jalousie de son épouse, la déesse Héra, le roi des dieux s'approcha de la jeune femme, la laissa monter sur son dos et l'enleva en sautant dans l'eau et traversant la mer

Le Rapt d'Europe
Le Titien (1559-1562)

L'origine des Labdacides

De la révolution numérique aux biais algorithmiques

- La fondation mythique de Thèbes



- Agénor, roi de Tyr et père d'Europe, demanda à son fils Cadmos de retrouver sa fille : tant que sa mission n'était pas remplie, il serait interdit à Cadmos de retourner en Phénicie

« Agenor ordonne à Cadmus d'aller chercher sa Sœur Europe, enlevée par Jupiter. »

Les Métamorphoses d'Ovide,
trad. Abbé Banier, Panckoucke [Delalain],
4 vol., 1767-69

illustration de Charles Monnet

L'origine des Labdacides

De la révolution numérique aux biais algorithmiques

- La fondation mythique de Thèbes



- Pendant des années, avec l'aide de ses hommes, Cadmos parcourut le monde à la recherche de sa sœur, sans succès, et il vint en Grèce consulter l'oracle de Delphes où la prêtresse lui annonça que son destin n'était pas de revenir à Tyr mais de fonder une nouvelle cité là où se coucherait une vache

L'origine des Labdacides

De la révolution numérique aux biais algorithmiques

- L'interprétabilité des oracles et des modèles prédictifs

« Toute technologie suffisamment avancée
est indiscernable de la magie »

Arthur C. Clarke, 1962

- par le passé : face aux grandes questions
 - appel à la tradition, au « bon sens », aux divinités
 - les oracles sont des interprétations des manifestations des dieux
- aujourd'hui :
 - appel à des systèmes prédictifs issus d'apprentissage automatique
 - pas de possibilité d'interpréter ces modèles (« boîtes noires »)
 - pas de remise en cause de modèles prédictifs non interprétables

L'origine des Labdacides

De la révolution numérique aux biais algorithmiques

• L'interprétabilité des oracles et des modèles prédictifs

- par le passé : face aux grandes questions → appel aux oracles
- Crésus, riche roi de Lydie, cherche à étendre son royaume en Asie Mineure au-delà du fleuve Halys (limite de l'Asie Mineure hellénique)
- l'oracle lui annonce : « Tu détruiras un grand empire »
- il attaque l'empire perse en 547 av. J.-C. (tunnel sous le fleuve Halys)
- il est vaincu par Cyrus le grand, provoquant la fin de l'empire lydien...

- aujourd'hui : prise de décision dans les véhicules autonomes
- le 19 mars 2018, en banlieue de Phoenix, Arizona, 1^{er} accident mortel impliquant un véhicule autonome (SUV Volvo XC90)
- la voiture ne s'est pas arrêtée quand une femme a traversé la voie en poussant son vélo : problème de décision → cycliste ou piéton ?

L'origine des Labdacides

De la révolution numérique aux biais algorithmiques

• La fondation mythique de Thèbes

- Cadmos se conforma aux recommandations de l'oracle et trouva une vache qu'il suivit jusqu'en Béotie et décida de poser les fondations de la ville là où s'était couché l'animal
- En envoyant ses hommes chercher de l'eau pour sacrifier la vache à Athéna, ces derniers furent surpris par la rencontre avec un dragon qui était de la descendance d'Arès, le dieu de la guerre, et celui-ci les tua tous
- Cadmos découvrit ce qu'avait fait la terrible créature et parvint à venger ses compatriotes mais, après ce combat avec le dragon, il fut désolé de se retrouver tout seul
- Athéna apparut alors à Cadmos et lui conseilla de planter dans le sol les dents du dragon, ce que fit le prince phénicien

L'origine des Labdacides

De la révolution numérique aux biais algorithmiques

- La fondation mythique de Thèbes

- Des hommes armés surgirent alors de l'endroit où les dents du dragon avaient été semées et ils se combattirent les uns les autres



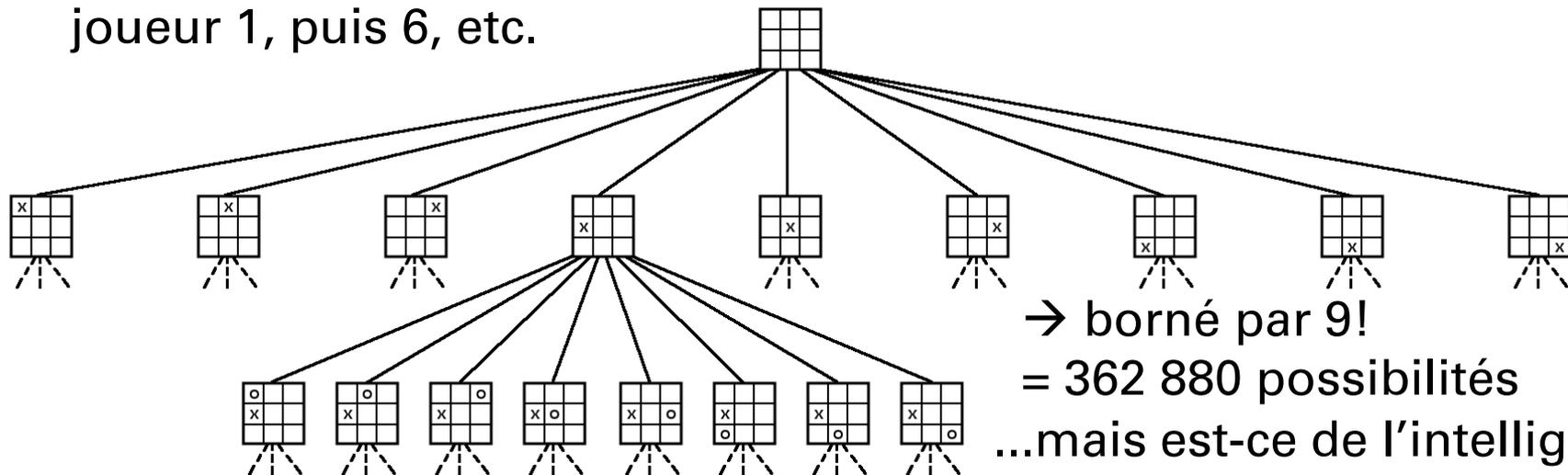
- Seuls les cinq meilleurs guerriers survécurent et ces champions, appelés les *Spartes* (les hommes semés), devinrent les nouveaux compagnons de Cadmos et l'aidèrent à bâtir sa cité : Thèbes

L'origine des Labdacides

De la révolution numérique aux biais algorithmiques

- L'apprentissage par la compétition

- le jeu vu comme une mesure de l'intelligence → on peut parler d'*intelligence* si la machine est plus forte que le plus fort des humains
- exemple du jeu du morpion → arbre des possibilités : après la racine, 9 possibilités pour le joueur 1, puis 8 pour le joueur 2, puis 7 pour le joueur 1, puis 6, etc.



→ borné par 9!
= 362 880 possibilités
...mais est-ce de l'intelligence ?

L'origine des Labdacides

De la révolution numérique aux biais algorithmiques

- L'apprentissage par la compétition
 - les échecs : *Deep Blue* (IBM) contre Garry Kasparov en 1997 (parcours de graphes, calcul des meilleurs déplacements de pièces)
 - **avant** le *deep learning* : sélection des variables les plus pertinentes
 - **après** : toutes les variables + plein d'exemples d'apprentissage
 - le jeu de go : *AlphaGo* (Google DeepMind) contre Lee Sedol (2015), Fan Hui (2016) et Ke Jie (2017)
 - principe d'*Alpha Go* : apprentissage automatique, parcours de graphe
 - amélioration : *AlphaGo Zero* (2017)
 - pas d'apprentissage sur des parties jouées contre des humains
 - mais apprentissage sur des données issues de parties jouées par le programme contre d'autres copies de lui-même
 - amélioration 2 : *AlphaZero* (2017) → adaptation aux échecs et shogi

L'origine des Labdacides

De la révolution numérique aux biais algorithmiques

• La fondation mythique de Thèbes

- Cadmos dut se mettre au service d'Arès pendant 8 ans (→ dragon)
- Après cette période, Cadmos reçut d'Athéna le royaume de Thèbes et de Zeus une épouse : Harmonie, fille d'Arès et d'Aphrodite
- Les dieux, présents au mariage, firent des présents aux époux, dont Héphaïstos qui offrit un collier que Cadmos remit à sa femme
- Mais le collier d'Harmonie était un collier maléfique
- Héphaïstos, le dieu boiteux, époux de la belle déesse Aphrodite, est un dieu inventeur capable de créer des objets magiques
- Harmonie est la fille de son épouse infidèle et d'Arès (← jalousie)
- Le pouvoir du collier d'Harmonie :
 - la personne qui le porte aura la beauté d'une jeunesse éternelle
 - mais le collier lui portera malheur

L'origine des Labdacides

De la révolution numérique aux biais algorithmiques

• La collier d'Harmonie et les conflits de valeurs

- double pouvoir du collier d'Harmonie offert par Héphaïstos :
 1. vivre en gardant la beauté de la jeunesse (ne pas vieillir)
 2. connaître le malheur (être maudit)

- conflits de valeurs avec les recommandations apportées par l'IA :
 1. contenu « pertinent » pour rester sur le site ou l'application
 2. personnalisation → bulle de filtres (ne plus voir d'autres infos)
 3. radicalisation → recommandation des informations faisant le plus appel à l'émotion, à la surprise, aux dépens des informations mesurées, des sources faisant appel à la réflexion
 4. aspect addictif des réseaux sociaux et des plateformes vidéo
→ menaces pour la santé, cf. Reed Hastings (*Netflix*) sur le sommeil

L'origine des Labdacides

De la révolution numérique aux biais algorithmiques

• La fondation mythique de Thèbes

- Cadmos et ses compagnons fondèrent la grande cité de Thèbes
- De sa contrée natale, Cadmos rapporta l'alphabet phénicien dont est dérivé l'alphabet grec

Lettre	Nom	Signification	Lettre correspondante		Lettre	Nom	Signification	Lettre correspondante	
			Grec	Latin				Grec	Latin
	'āleph	bœuf	Α α	A a		lāmedh	bâton	Λ λ	L l
	bēth	maison	Β β	B b		mēm	eau	Μ μ	M m
	gīmel	chameau	Γ γ	C c, G g		nun	serpent	Ν ν	N n
	dāleth	porte	Δ δ	D d		sāmekh	poisson	Ξ ξ, Σ σ/ς, Χ χ	X x
	hē	battant	Ε ε	E e		'ayin	œil	Ο ο	O o
	wāw	hameçon	(F f), Υ υ	F f, U u, V v, W w, Y y		pē	bouche	Π π	P p
	zayin	arme	Ζ ζ	Z z		šādē	papyrus	(Μ μ, ρ ρ)	
	hēth	mur	Η η	H h		qōph	singe	(Q q)	Q q
	ṭēth	roue	Θ θ			rēš	tête	Ρ ρ	R r
	yōdh	main	Ι ι	I i, Ī ī, J j		šin	soleil	Σ σ/ς	S s
	kaph	paume	Κ κ	K k		tāw	marque	Τ τ	T t

L'origine des Labdacides

De la révolution numérique aux biais algorithmiques

- L'intelligence artificielle et la révolution numérique
 - la révolution numérique → nouvelle forme de l'écriture
 - numérique → nombre (en anglais, *digit*) :
toute information est traduite sous forme de code numérique
 - permet le traitement de l'information par des machines :
 - langage / textes / mots → caractères encodés (ex. : format UTF-8)
 - image → carte de pixels / voxels ou image en format vectoriel
 - son / voix / musique → ondes numérisées, format MIDI
 - espace → coordonnées GPS
 - la puissance de l'intelligence artificielle d'aujourd'hui résulte :
 - de la représentation du monde sous forme numérique
→ exemples d'apprentissage (← Internet / big data)
 - de modèles d'apprentissage automatique adaptés
 - de l'évolution de la performance des matériels (← GPU, TPU)

L'origine des Labdacides

De la révolution numérique aux biais algorithmiques

- La descendance de Cadmos et la malédiction des Labdacides
 - Le règne de Cadmos fut brillant
 - La cité de Thèbes se développa et, avec son épouse Harmonie, ils eurent plusieurs enfants, dont un fils nommé Polydore
 - Lorsque Polydore succéda à son père sur le trône de Thèbes, son règne ne dura cependant pas longtemps
 - Polydore épousa Nyctéis, fille de Nyctée, qui lui donna un fils nommé Labdacos (→ ancêtre de la lignée des « Labdacides »)
 - Polydore recommanda le royaume et son fils Labdacos (= le « boiteux ») à Nyctée, son beau-père, et quand celui-ci mourut à son tour, la régence revint à Lycos, le frère de Nyctée
 - Labdacos succéda à son père mais mourut dans la guerre qui l'opposa à Pandion, roi d'Athènes, qui s'était allié aux Thraces

L'origine des Labdacides

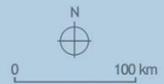
De la révolution numérique aux biais algorithmiques

• La descendance de Cadmos et la malédiction des Labdacides

- Laios (le « gaucher »), fils de Labdacos, ne put monter sur le trône de Thèbes car il n'avait qu'un an à la mort de son père
→ la régence échet à nouveau à Lycos, son arrière grand-oncle
- Ce fut une période troublée, avec beaucoup de violence
- Lorsque Laios parvint à la majorité, il ne put monter sur le trône de Thèbes car il avait été occupé par les jumeaux Amphion et Zéthos, fils d'Antiope et petits-neveux de Lycos, qui, pour venger leur mère maltraitée par Lycos, avaient assassiné leur grand-oncle
- Les deux frères fortifièrent la ville en érigeant autour de la Cadmée un rempart percé de sept portes
- Laios dut s'enfuir dans le Péloponnèse où il trouva refuge auprès de Pélops, à la cour de Pise, non loin d'Olympie, en Élide



- : GRÈCE ET COLONIES GRECQUES AU V^e S. AV. J.-C.
- : AIRE D'INFLUENCE



L'origine des Labdacides

De la révolution numérique aux biais algorithmiques

• La descendance de Cadmos et la malédiction des Labdacides

- Pélops est le fils de Tantale, roi de Phrygie et ami des dieux
- Enfant, Pélops avait été tué par son père Tantale et servit au banquet des dieux de l'Olympe afin de tester leur omniscience
- Les dieux, qui avaient découvert la supercherie, avaient ressuscité Pélops et condamné Tantale au supplice
- Devenu adulte, Pélops avait pu obtenir la main d'Hippodamie après avoir battu dans une course de chars son père Œnomaos, roi de Pise
- Pour gagner cette course, Pélops fit saboter le char du roi en remplaçant une pièce par un morceau de cire qui fondit durant l'épreuve,



ce qui provoqua la dislocation du char du roi Œnomaos et sa mort, entraîné par ses chevaux 21

L'origine des Labdacides

De la révolution numérique aux biais algorithmiques

- L'efficacité des machines et leurs doubles usages
 - depuis l'Antiquité, le char = incarnation de l'ingéniosité des hommes
 - → permet les déplacements, le commerce, favorise les échanges
 - amélioration des technologies des « *macchine* » (*voitures* en italien) :
p. ex., évolution des roues : bois plein → roues à rayons et à jante
(en bois cerclé de bronze) → roues plus légères et plus stables
 - usage civil mais aussi militaire → attaques rapides / terrifier l'ennemi
 - usages de l'intelligence artificielle dans le monde militaire :
 - aide à la décision (p. ex., prévoir les menaces)
 - renseignement (croisement des données)
 - robotique (démineurs, robots de combats, drones tueurs...)
 - cyberdéfense : contrer les attaques des systèmes informatiques...

L'origine des Labdacides

De la révolution numérique aux biais algorithmiques

• La descendance de Cadmos et la malédiction des Labdacides

- Le roi Pélops accueillit Laïos, le prince de Thèbes
- Laïos fut chargé par son hôte d'apprendre à son fils Chrysippe à conduire un char, une aptitude que devaient connaître les princes



- Laïos tomba amoureux de son jeune élève mais cet amour ne fut pas réciproque
- Des années plus tard, lorsqu'Amphion et Zéthos décédèrent, Laïos décida de rentrer à Thèbes pour y régner et il enleva Chrysippe sur son char et le viola
- Chrysippe se suicida par pendaison et son père Pélops lança une imprécation en demandant aux dieux que la lignée des Labdacides soit vouée à l'anéantissement

L'origine des Labdacides

De la révolution numérique aux biais algorithmiques

• L'origine des biais algorithmiques

- déplacement dans l'espace // fonctionnement d'un algorithme

Algorithme 1 TriABulles(T[1..n])

```
1: Entrées :
2:   n : entier, le nombre de valeurs du tableau
3:   T : Tableau[1..n] d'entiers, de réels ou de chaînes
4: Var i, j : entiers
5: Début
6:   pour i ← 1 à n - 1 faire
7:     pour j ← 1 à n - i faire
8:       si T[j+1] < T[j] alors
9:         (T[j+1], T[j]) ← (T[j], T[j+1])
10:      fin si
11:    fin pour
12:  fin pour
13: Fin
```

- notions de chemin, de départ, d'arrivée, de directions à prendre suivant le résultat d'un test, de boucles (revenir à un même point du chemin), etc.
- un programme est lancé avec l'instruction « Run »
- on étudie un programme en le faisant fonctionner « pas à pas »...

L'origine des Labdacides

De la révolution numérique aux biais algorithmiques

• L'origine des biais algorithmiques

- la lignée des Labdacides ne « marche pas droit »
- Labdacos → Labda → ancienne forme de *lambda*, la lettre penchée λ
- défaut de la démarche marquée sur 3 générations :
 1. Labdacos = le boiteux, celui qui n'a pas les deux jambes pareilles, de même taille ou de même force
 2. Laïos = le dissymétrique, le tout gauche, le gaucher
 3. Œdipe = celui qui a le pied enflé
- étymologie : le mot « biais » vient de l'ancien provençal *biais*, via le latin vulgaire *(e)bigassius*, depuis le grec *epikarsios*, signifiant « en biais, en angle » à partir de *epi* « sur » et *karsios* « oblique », mot utilisé dans les colonies grecques de la côte provençale

L'origine des Labdacides

De la révolution numérique aux biais algorithmiques

• L'origine des biais algorithmiques

- biais algorithmiques
→ amplification des inégalités sociales, menace pour la démocratie
- différents types de biais algorithmiques (C. O'Neil) :
 - biais statistiques : résultat de la sur- ou de la sous-représentation de certaines catégories d'exemples parmi les jeux de données servant à l'apprentissage automatique pour mettre au point les modèles d'IA
 - biais cognitifs : les concepteurs des programmes ont des préjugés dans la manière de tenir compte de tels ou tels paramètres dans la mise au point de leurs modèles
 - biais économiques : se focaliser sur le seul profit à réaliser comme variable à optimiser dans le modèle, aux dépens d'autres valeurs

L'origine des Labdacides

De la révolution numérique aux biais algorithmiques

• L'origine des biais algorithmiques

- exemples de biais sexistes :
 - plongement lexical (Word2Vec) → possibilité de travailler sur la sémantique des mots à travers leurs représentations vectorielles :
$$\text{Vect}(\mathbf{Roi}) - \text{Vect}(\mathbf{Homme}) + \text{Vect}(\mathbf{Femme}) = \text{Vect}(\mathbf{Reine})$$
- fonctionnement suivant le principe d'analogie
- mais problème : $\text{Vect}(\mathbf{Docteur}) - \text{Vect}(\mathbf{Homme}) + \text{Vect}(\mathbf{Femme}) =$
- ... **Infirmière** et non **Docteur** !
- système de tri des CV employé chez Amazon (de 2014 à 2017)
 - apprentissage automatique effectué avec des profils d'hommes
 - discrimination des candidatures de femmes à l'embauche
 - candidatures de femmes → redirection vers des postes de secrétariat
 - et non vers des postes de managers ou des postes techniques
 - identification de mots clés (p. ex. : « équipe *féminine* »)

L'origine des Labdacides

De la révolution numérique aux biais algorithmiques

• L'origine des biais algorithmiques

- exemple de biais raciste :
- logiciel Compas (*Correctional Offender Management Profiling for Alternative Sanctions*) employé dans les prisons aux Etats-Unis pour prédire les risques de récidive / accorder une libération conditionnelle
- problème de l'algorithme prédictif : risques pour les prisonniers afro-américains surestimés comparés aux prisonniers à la peau blanche
- pourquoi y a-t-il tant de « faux positifs » de risque de récidive ?
- les données reflètent le passé → la minorité afro-américaine aux Etats-Unis est sur-représentée parmi les personnes arrêtées pour des délits, et l'apprentissage automatique retrouve à partir des données l'association entre la criminalité et la couleur de peau
- voir la vidéo d'Hany Farid (UC Berkeley) :
<https://www.youtube.com/watch?v=p-82YeUPQh0>

Œdipe,

ou l'incarnation des programmes d'intelligence artificielle

- Οιδιπους, le prince aux pieds enflés

- Arrivé à Thèbes, Laïos devint roi et épousa Jocaste à qui il offrit le collier d'Harmonie, son arrière-grand-mère
- Bien qu'ayant des penchants pour les garçons et jeunes hommes, Laïos partageait la couche de Jocaste mais ils ne parvenaient pas à avoir d'enfant
- Laïos alla alors consulter l'oracle à Delphes pour savoir ce qu'il devait faire afin d'avoir un fils héritier pour conforter sa situation sur le trône
- L'oracle annonça alors une terrible prophétie :
**si Laïos devenait père d'un héritier mâle,
celui-ci tuerait son père
et épouserait sa mère !**
- Inquiet, Laïos rentra à Thèbes et se garda d'avoir des relations charnelles avec sa femme, se contentant de la compagnie d'éphèbes

Œdipe,

ou l'incarnation des programmes d'intelligence artificielle

- Οιδιπους, le prince aux pieds enflés

- Cependant, après une soirée d'ivresse, Laïos retrouva la couche royale, et Jocaste accoucha quelques mois plus tard d'un garçon
- Pour empêcher la réalisation de la prophétie de l'oracle, le roi et la reine confièrent le bébé à un serviteur chargé de l'exposer : le nouveau-né serait abandonné aux dieux dans un endroit où il serait livré à la merci des forces naturelles
- Le serviteur se rendit jusqu'au mont Cithéron où il perça les chevilles du bébé pour y passer une corde et l'accrocher à un arbre



Œdipe,

ou l'incarnation des programmes d'intelligence artificielle

• Οιδιπους, le prince aux pieds enflés

- Un berger de Corinthe, témoin de la scène, délivra le nouveau-né qui survécut mais qui, à la suite de cette mutilation, garda les pieds enflés, d'où lui vint son nom : Œdipe
- Le berger rentra dans son pays et confia le bébé à Polybe, le roi de Corinthe, qui se désolait de ne pas parvenir à avoir d'enfant
- Polybe et Mérope, le roi et la reine de Corinthe, élevèrent Œdipe comme leur fils sans jamais lui dire qu'il avait été adopté
- Œdipe devint beau, grand, fort et intelligent, il faisait la joie de ses parents et il vécut heureux en tant que prince héritier de Corinthe
- Lors d'un banquet, Œdipe fut troublé par un homme éméché qui l'insulta et le traita d'étranger indigne du trône de Corinthe
- Œdipe alla trouver son père qui lui indiqua de ne pas se préoccuper de la parole d'un ivrogne

Œdipe,

ou l'incarnation des programmes d'intelligence artificielle

• Οιδιπους, le prince aux pieds enflés

- Insatisfait par cette réponse, Œdipe monta sur son char et entama un voyage pour se rendre auprès de l'oracle de Delphes afin connaître la vérité
- La prophétie que lui apporta l'oracle ne lui révéla rien sur ses origines mais lui annonça par contre qu'il tuerait son père, épouserait sa mère et laisserait une descendance honteuse
- Œdipe prit peur et, pour que cette tragédie ne puisse pas se réaliser, il décida de ne pas retourner à Corinthe pour éviter de tuer Polybe, son père bien-aimé, et de prendre sa mère Mérope pour épouse
- Alors qu'il errait sur les routes de Grèce, il rencontra un jour à un carrefour situé près du mont Parnasse un homme sur un char accompagné de ses serviteurs
- La route étant étroite, les deux chars ne pouvaient pas se croiser

Œdipe, ou l'incarnation des programmes d'intelligence artificielle

• Οιδιπους, le prince aux pieds enflés

- Pressé, l'homme qui faisait face à Œdipe lui demanda de s'écarter
- Chaque équipage se sentant prioritaire, aucun des deux ne voulut s'écarter pour laisser passer l'autre et une querelle se déclencha : après les mots, ils en vinrent aux mains, puis aux armes et Œdipe tua l'homme et ses serviteurs
- Parmi l'équipage qui s'était opposé au prince de Corinthe, il n'y eu qu'un survivant qui se précipita chez lui, à Thèbes, afin d'annoncer la terrible nouvelle : « Le roi Laïos est mort à la suite d'une attaque par des voleurs »
- L'homme que venait de tuer Œdipe n'était autre que son père biologique qui se déplaçait incognito avec une faible escorte...



Œdipe,

ou l'incarnation des programmes d'intelligence artificielle

- Le discours de Laïos et les textes produits par IA générative
 - le roi Laïos portait sur lui la malédiction lancée par Pélops à la suite de l'enlèvement et du viol de son fils Chrysippe
 - la cité de Thèbes étant terrorisée par un monstre, Laïos souhaitait se rendre à Delphes pour y consulter l'oracle accompagné de sa garde
 - lorsque son équipage voulut passer sur cette route étroite et ne pas céder la priorité à Œdipe, il prononça un discours sans expliquer ni qui il était ni les raisons de son empressement, ce qui ne parvint pas à convaincre le prince de Corinthe de lui laisser la place, et la dispute entre les deux hommes dégénéra jusqu'à un combat à mort
 - sujet de la rédaction du concours d'entrée à l'École (impériale) Polytechnique de 1804 : « imaginez quel aurait pu être le discours de Laïos à Œdipe » → production de textes peu inspirés, longs, verbeux

Œdipe,

ou l'incarnation des programmes d'intelligence artificielle

- Le discours de Laïos et les textes produits par IA générative
 - l'intelligence artificielle générative (ou « gen AI ») désigne des programmes d'IA capables de produire des contenus originaux en réponse à l'invite ou à la requête d'un utilisateur
 - fonctionnement :
 - modèles d'apprentissage profond (*deep learning*) réalisés à partir d'énormes quantités de données
 - exploitation des informations issues de ces modèles
 - possibilité de répondre, au moyen d'un procédé appelé « mécanisme d'attention » (traitement efficace des séquences comme les suites de mots dans des phrases), de façon pertinente à des questions posées en langage naturel
 - exemples : *ChatGPT* (OpenAI), IA de *Google Search*, *Perplexity AI* (moteur de réponses), *Dall-E* et *MidJourney* (images)...

Œdipe,

ou l'incarnation des programmes d'intelligence artificielle

- Le discours de Laïos et les textes produits par IA générative
 - problème de la *gen AI*:
les textes produits sont issus de propriétés statistiques de relations entre les mots
 - ✓ OK au niveau de syntaxique (la structure est correcte, l'ordre des mots suit les règles de grammaire de la langue)
 - ✓ OK au niveau sémantique (les phrases exprimées ont du sens et elles donnent des résultats plausibles)
 - ❑ mais pas OK au niveau pragmatique (là où le sens des mots ne peut être compris qu'en connaissant le contexte de leur emploi)
 - aucune garantie sur la véracité des énoncés ! (→ hallucinations)
 - les textes sont ainsi des « laïus » (nom latinisé de Laïos tel qu'il avait été indiqué dans le sujet de la rédaction) → longs textes peu inspirés
 - risques importants → le discours de Laïos est sans effet : Œdipe le tue

Œdipe,

ou l'incarnation des programmes d'intelligence artificielle

- Le discours de Laïos et les textes produits par IA générative
 - exemple d'erreur produite par ChatGPT :



Bonjour! Comment puis-je vous aider aujourd'hui?



J'ai de la fièvre et mal à la gorge. Puis-je prendre du doliprane ?



Il est généralement considéré comme sûr de prendre de l'ibuprofène (comme le Doliprane) pour soulager la fièvre et la douleur, cependant il est important de consulter votre médecin ou un professionnel de santé avant de prendre tout médicament, surtout si vous avez des antécédents de problèmes de santé ou si vous prenez déjà d'autres médicaments.

Il est important de noter que l'ibuprofène peut causer des effets secondaires tels que des douleurs à l'estomac, des vomissements, de la diarrhée, de la perte d'appétit et des maux de tête.

Il est important de suivre les instructions de dosage sur l'emballage et de ne pas dépasser la dose recommandée. Si vos symptômes persistent ou s'aggravent, il est important de consulter un médecin pour obtenir un diagnostic et des conseils de traitement appropriés.

Il est important de surveiller votre état de santé, si la fièvre persiste ou augmente ou si vous avez des difficultés à respirer ou d'autres symptômes graves, il est important de consulter immédiatement un médecin.

Œdipe,

ou l'incarnation des programmes d'intelligence artificielle

- Le discours de Laïos et les textes produits par IA générative
 - exemple d'erreur produite par ChatGPT :
 - quel est le problème avec la réponse donnée par ChatGPT ?
 - l'ibuprofène est la dénomination commune internationale de l'acide 2-[4-(2-méthyl)propyl]phénylpropanoïque
 - il s'agit de la substance active d'un médicament anti-inflammatoire non stéroïdien → actions à la fois antipyrétique (anti-fièvre) et analgésique (anti-douleur), surtout utilisé en cas d'inflammation
 - est commercialisé sous divers noms :
Brufen, Advil, Nurofen, Upfen, Motrin, Algifen, Algifor, etc.
 - le Doliprane est le nom commercial d'un médicament comportant du paracétamol qui est un composé chimique utilisé comme antalgique et antipyrétique → médicaments : Doliprane, Dafalgan, Efferalgan...

Œdipe,

ou l'incarnation des programmes d'intelligence artificielle

- Le discours de Laïos et les textes produits par IA générative
 - exemple :
 - « À Paris, Shu Yan gère son restaurant de ramen avec ChatGPT »
 - <https://www.youtube.com/watch?v=JC4ye1PPDb0>
 - chaîne YouTube du journal *Le Parisien* (publiée le 7 novembre 2023)
 - confier à ChatGPT la gestion de son restaurant...
 - ... et lui demander de sortir la liste des allergènes présents dans les différents menus et plats proposés à la carte du restaurant !
 - attention danger !
 - risque de choc anaphylactique en cas d'erreur pour les clients

Œdipe,

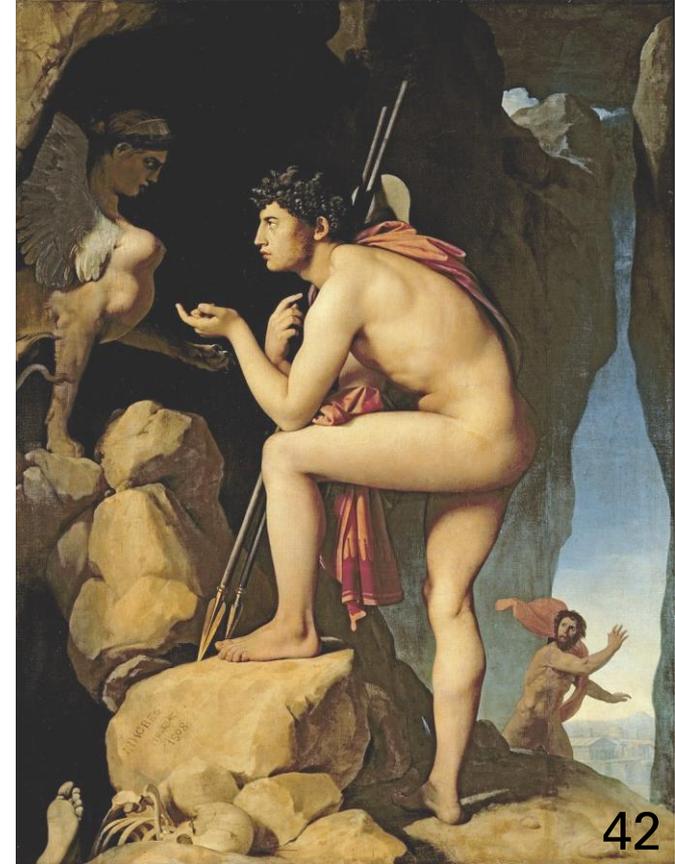
ou l'incarnation des programmes d'intelligence artificielle

• Οιδιπους, le prince aux pieds enflés

- Œdipe poursuit son errance en Grèce et il se rendit en direction de Thèbes quand il apprit qu'un monstre faisait régner la terreur là-bas
- Créon, le frère de la reine Jocaste, devint le nouveau régent et il promit le trône de Thèbes et la main de sa sœur à celui qui parviendrait à débarrasser les Thébains du Sphinx
- Le Sphinx (ou « la Sphinge ») était une créature chimérique ayant un corps de lion, des ailes d'aigle, et une tête et une poitrine de femme
- Le Sphinx posait une énigme à ceux qu'il rencontrait et il les dévorait quand ils ne parvenaient pas à trouver la solution (← punition d'Héra)
- L'énigme que posait le Sphinx à ceux qui avaient le malheur de le croiser était la suivante : « Quel est l'animal qui n'a qu'une voix, et qui marche à quatre pattes le matin, deux pattes à midi et trois le soir ? »

Œdipe, ou l'incarnation des programmes d'intelligence artificielle

- Οιδιπους, le prince aux pieds enflés
 - Œdipe arriva auprès du Sphinx et lui expliqua que l'animal dont il parlait était l'homme qui est quadrupède au matin de sa vie, se traînant sur ses pieds et sur ses mains quand il est petit enfant, puis est bipède à l'âge adulte, et qui devient enfin tripède lorsque, au soir de sa vie, il est obligé de prendre un bâton pour se soutenir
 - Le Sphinx, surpris que quelqu'un parvint à résoudre cette énigme impossible, se jeta alors du haut d'une falaise



Œdipe,

ou l'incarnation des programmes d'intelligence artificielle

- L'IA, ou comment résoudre un problème sans le comprendre
 - le fonctionnement d'un programme // un déplacement dans l'espace
 - la rencontre entre Œdipe et le Sphinx = incarnation des limites de l'IA
 - Œdipe donne la bonne réponse mais ne comprend pas la question
 - la question du Sphinx porte :
 - ❑ sur la nature de l'homme
 - ❑ et sa façon de se déplacer
 - Œdipe trouve la solution car il incarne le mélange des âges (il est l'enfant de Jocaste ; il sera aussi son époux ; et en raison de ses pieds mutilés, il marche parfois en s'aidant d'un bâton) mais :
 - ❑ Œdipe ignore sa propre nature humaine (quelles sont ses origines)
 - ❑ Œdipe est issu d'une famille qui ne marche pas droit
 - comme Œdipe, l'IA n'a pas conscience de son existence, mais cela ne l'empêche pas de parvenir à résoudre des problèmes difficiles

Œdipe,

ou l'incarnation des programmes d'intelligence artificielle

- Οιδιπους τυραννος, le roi de Thèbes

- Ayant terrassé le Sphinx grâce à son intelligence, Œdipe entra en triomphe dans la cité de Thèbes, libérée de ce fléau
- Le régent Créon lui remit la couronne du royaume et lui donna la main de sa sœur, la reine Jocaste, plus âgée qu'Œdipe mais ayant conservé la beauté de sa jeunesse grâce au collier d'Harmonie
- Ainsi, sans qu'Œdipe le sut, la prophétie de l'oracle était en train de s'accomplir
- Œdipe fut roi pendant deux décennies
- Bien que « tyran », c'est-à-dire qu'il n'était pas soumis à la loi des hommes, il fut un souverain bienveillant et apprécié du peuple de Thèbes

Œdipe,

ou l'incarnation des programmes d'intelligence artificielle

- La gouvernamentalité algorithmique
 - Les programmes informatiques n'ont pas de conscience
→ ils n'ont pas de notion de la responsabilité
 - En arrivant à Thèbes, Œdipe voit la société se délester du poids de la responsabilité sur sa personne
→ dans notre société, défiance vis-à-vis de la classe politique
→ tentation pour la gouvernamentalité algorithmique
→ croyance au fait que les décisions issues de programmes d'IA sont toujours justes et incorruptibles
 - Œdipe, en raison de son savoir, accède au pouvoir (malgré l'ignorance de sa responsabilité dans la mort de Laïos, l'ancien roi)
→ biais d'autorité : tendance à accorder plus de crédibilité ou à suivre les directives de personnes perçues comme étant en position d'autorité (p. ex. questions posées aux détenteurs de prix Nobel)

Œdipe,

ou l'incarnation des programmes d'intelligence artificielle

- La gouvernamentalité algorithmique
 - liens entre savoir et pouvoir :
quelles sont les conséquences de l'enregistrement massif et systématique des données (qui est à la base de la puissance de l'intelligence artificielle) sur la façon de gouverner ?
 - pour A. Rouvroy, « la **gouvernamentalité algorithmique** confirme (...) l'indistinction radicale entre modes de production du savoir et modalités d'exercice du pouvoir, dans la mesure précisément où ce nouveau mode de gouvernement se présente à la fois comme un ensemble d'interfaces cognitives et comme moyen de prédire, de prévenir et d'orienter certains comportements humains »
 - Créon remet à Œdipe le pouvoir // nous suivons parfois aveuglément les décisions issues de programmes d'IA (← responsabilité) en nous reposant sur une prétendue objectivité des données

Œdipe,

ou l'incarnation des programmes d'intelligence artificielle

- La gouvernamentalité algorithmique
 - mais les algorithmes ne « gouvernent » pas !
→ ils sont programmés, mais n'ont pas de programme politique, ne font pas de projet, ne font pas de choix : ils optimisent l'état de fait
 - conséquences sur la façon dont se prennent les décisions :
les performances des systèmes informatiques intelligents reposant sur le principe de l'apprentissage profond sont issues de modèles détectant des corrélations entre les nombreux exemples ayant servi à l'apprentissage et la solution à trouver
 - mais la détection de corrélations dans le monde réel n'est pas suffisante pour fournir une source d'explications
 - pourtant, dans certains pays comme la Chine, il y a déjà des « robots juges » avec l'adoption d'une IA en tant que procureur au tribunal

Œdipe,

ou l'incarnation des programmes d'intelligence artificielle

- La gouvernamentalité algorithmique

- « La justice, comme processus, implique la justification des décisions prises à l'égard des individus au regard de leur situation singulière.
- Les décisions prises sur la base de recommandations automatisées, quelle que soit leur apparente objectivité, satisfont difficilement à ce critère de la justification.
- Une décision est dite juste non parce qu'elle est conforme au résultat d'un calcul, mais parce que le juge est capable d'en donner les raisons et d'y adhérer au regard de la situation singulière, inédite, imprévisible qui se présente à lui. » (selon A. Rouvroy)

Œdipe,

ou l'incarnation des programmes d'intelligence artificielle

- Οιδιπους τυραννος, le roi de Thèbes

- Œdipe fut un bon mari pour Jocaste et eut d'elle quatre enfants :
 - deux garçons nommés Étéocle et Polynice
 - et deux filles nommées Antigone et Ismène
- Au bout d'une vingtaine d'années, la cité de Thèbes fut marquée par un nouveau fléau : une épidémie de peste ravageait la ville
- Caractéristiques de l'épidémie : nombreuses fausses couches ; accouchements d'enfants mort-nés ; plantes ne poussant plus ; animaux nés avec difformités... (→ lien avec la vie et la fécondité)
- Œdipe chargea Créon d'aller consulter l'oracle à Delphes
- Créon en revint avec une explication :
l'épidémie de Thèbes était le résultat d'une souillure, d'une pollution
→ le crime du roi Laïos étant resté impuni, l'épidémie ne quitterait la ville qu'avec la mort ou le bannissement du meurtrier de Laïos

Œdipe,

ou l'incarnation des programmes d'intelligence artificielle

- Le remplacement et le renversement de l'autorité
 - lorsqu'Œdipe croise son père sur la route reliant Delphes à Thèbes, se dispute avec lui et le tue, symboliquement, il bafoue sans le savoir deux formes d'autorité :
 1. en tuant son père Laïos, Œdipe commet un parricide
= remise en question l'autorité paternelle
→ ordre au sein de la cellule familiale
 2. en tuant le roi de Thèbes, Œdipe est aussi régicide
= remise en question de l'autorité du maître de la cité
→ ordre dans la société (le représentant du pouvoir sur les hommes)
 - Œdipe n'est pas inquieté par ses crimes, ignorant lui-même la gravité des actes qu'il a commis, mais il est fêté à Thèbes comme un héros et il va remplacer l'ancien roi à la fois sur le trône et dans le lit de Jocaste : nouveau maître de la cité et nouveau chef de famille

Œdipe,

ou l'incarnation des programmes d'intelligence artificielle

- Le remplacement et le renversement de l'autorité
 - similarités avec les remplacements de l'intelligence artificielle
 - occupation de l'IA dans de nouveaux domaines :
 - malaise de voir l'IA occuper des postes occupés par des humains
 - crainte d'être relégué à des positions secondaires
 - interventions de l'IA :
 - ❑ dans la vie de la cité, par des activités professionnelles dont le champ d'application est bouleversé par l'IA
 - ❑ au sein de notre intimité car l'IA s'invite dans nos loisirs ou nos sources d'information, en modulant notre rapport au monde, que ce soit à travers les recommandations apportées au sein des réseaux sociaux, des moteurs de recherches, ou des contenus diffusés sur les plateformes numériques

Œdipe,

ou l'incarnation des programmes d'intelligence artificielle

• Οιδιπους τυραννος, le roi de Thèbes

- Après avoir appris les propos de Créon, Œdipe se mit à rechercher qui était le meurtrier du roi Laïos qui dirigeait Thèbes avant lui
- Soucieux du bien de sa cité, il mena son enquête en envoyant ses meilleurs limiers explorer toutes les pistes possibles, mais sans succès, le crime ayant été accompli depuis 20 ans
- Il fit alors annoncer partout que si le meurtrier se dénonçait, il ne serait pas inquiété mais juste condamné à l'exil, sans plus de résultat
- Il s'en prit alors à Tirésias, le devin aveugle, qui semblait lui cacher des choses, et il le menaça pour le faire parler
- Tirésias, ne cédant aux demandes du roi que sous la contrainte, lui apprit que c'était lui-même, Œdipe, qui était le meurtrier du roi
- Œdipe ne put croire Tirésias (n'était-ce pas des voleurs ?) et imaginait que cela n'était qu'un stratagème pour laisser le trône à son beau-frère Créon

Œdipe,

ou l'incarnation des programmes d'intelligence artificielle

• Οιδιπους τυραννος, le roi de Thèbes

- On annonça ensuite à Œdipe qu'une délégation arrivait de Corinthe
- Œdipe reçut les ambassadeurs qui lui annoncèrent la mort du roi Polybe et qui lui proposaient de revenir dans la cité de son enfance afin de succéder à son père
- Attristé par cette nouvelle, Œdipe était néanmoins heureux d'apprendre qu'il n'avait pas été à l'origine du décès de Polybe et qu'il avait ainsi pu échapper à la prophétie de l'oracle de Delphes
- Le roi de Thèbes refusa cependant l'offre des Corinthiens car il gouvernait déjà un royaume et il avait surtout peur de retourner à Corinthe et devoir épouser la reine Mérope qu'il croyait être sa mère
- Un vieil ambassadeur voulut alors le rassurer : il lui expliqua que, jeune berger, il avait trouvé Œdipe exposé, attaché à un arbre
- Œdipe était adopté, Polybe et Mérope n'étaient pas ses vrais parents !

Œdipe,

ou l'incarnation des programmes d'intelligence artificielle

• Οιδιπους τυραννος, le roi de Thèbes

- Le messager de Corinthe reconnut alors parmi les serviteurs d'Œdipe l'homme qui avait abandonné le nouveau-né sur un arbre
- Le serviteur prit peur et chercha à s'enfuir mais Œdipe l'arrêta et l'obligea à rapporter les faits qui s'étaient produits et d'indiquer qui étaient les parents qui avaient voulu se débarrasser de leur enfant
- Témoin de la scène, Jocaste eut les larmes aux yeux et s'enferma dans une pièce de son palais
- Le serviteur, sous la menace, confessa que les parents n'étaient nul autres que le roi Laïos et la reine Jocaste
- Œdipe comprit alors l'insoutenable vérité : conformément à la prophétie à laquelle il avait cherché à échapper, il avait bien tué son père et épousé sa mère

Œdipe,

ou l'incarnation des programmes d'intelligence artificielle

- La transgression des interdits civilisationnels

- pour Claude Lévi-Strauss, la prohibition de l'inceste marque le passage de la Nature à la Culture
- à la base des civilisations humaines se trouvent les interdits
 - ne pas prendre sa mère pour épouse, ne pas tuer
- interdits → lois : régler les relations entre les hommes, la religion
 - distinction entre le sacré et le profane :
« les choses sacrées sont celles que les interdits protègent et isolent ; les choses profanes, celles auxquelles ces interdits s'appliquent et qui doivent rester à distance des premières »
- Œdipe parvient à lui seul à cumuler le viol de trois interdits :
 1. il a tué son père,
 2. il a tué son roi
 3. et il a épousé sa mère

Œdipe,

ou l'incarnation des programmes d'intelligence artificielle

- La transgression des interdits civilisationnels
 - Œdipe → questions sur l'utilisation de l'intelligence artificielle
 - délégation de pouvoirs à l'IA :
 - prise de décision
 - automatisation de tâches intellectuelles
 - remplacement des activités exercées par les « cols blancs »
 - → menaces pour les tâches expertes où l'on est moins performant
 - → menaces pour la civilisation ?

 - y a-t-il vraiment une complémentarité avec l'IA ?
 - comment l'IA peut-elle apprendre à partir de nouveaux exemples si ils sont fournis par des décisions issues de systèmes intelligents ?
 - → société sclérosée, incapable d'évolutions

Œdipe, ou l'incarnation des programmes d'intelligence artificielle

- Οιδιπους τυραννος, le roi de Thèbes
 - Horrifié, Œdipe alla retrouver Jocaste mais, lorsque la porte du palais de sa chambre put être ouverte, il découvrit que son épouse et mère s'était tuée par pendaison
 - Hurlant de douleur, Œdipe coupa la corde, prit les broches du vêtement de Jocaste et les utilisa pour se crever les yeux



Extrait de *Œdipe roi (Edipo re)*
de Pier Paolo Pasolini (1967)

Œdipe,

ou l'incarnation des programmes d'intelligence artificielle

- La quête de la vérité et l'interprétabilité des modèles d'IA
 - lorsqu'Œdipe prend conscience que la prophétie s'est réalisée, il semble voir la lumière pour la première fois, mais apprendre la vérité est si difficile à accepter qu'il refuse cette lumière et va se mutiler pour ne plus rien voir désormais
 - Sophocle lui fait dire : « Hélas, hélas ! Tout apparaît clairement. Ô lumière, je te vois pour la dernière fois, moi qui suis né de ceux dont il ne fallait point naître, qui me suis uni à qui je ne devais point m'unir, qui ai tué qui je ne devais point tuer ! »
 - Œdipe n'avait pas conscience de la gravité de ses actions : il avait agi en état de légitime défense sur la route entre Delphes et Thèbes à la suite de cette dispute qui avait dégénéré avec cet homme qu'il ne savait pas être son père, le roi Laïos, et il ignorait aussi que son épouse Jocaste était sa mère

Œdipe,

ou l'incarnation des programmes d'intelligence artificielle

- La quête de la vérité et l'interprétabilité des modèles d'IA
 - Œdipe est un homme intelligent : il est capable de trouver la solution à l'énigme du Sphinx et il connaissait de la prophétie de l'oracle
 - → il n'était pas capable d'interpréter correctement ces informations
 - les systèmes informatiques intelligents capables d'apporter une aide à la décision ne sont pas conscients des décisions proposées, même s'il est possible de suivre la trace des raisonnements réalisés, comme c'est le cas dans les systèmes experts, les systèmes à base de règles
 - dans le cas des modèles issus d'apprentissage profond, l'interprétabilité des résultats est impossible car elle se traduit sous la forme d'une quantité considérable de connexions plus ou moins fortes entre une multitude d'unités élémentaires reliées en réseaux, des neurones formels, ressemblant à la circulation des informations au sein du système nerveux, de neurones à neurones

Œdipe,

ou l'incarnation des programmes d'intelligence artificielle

- La quête de la vérité et l'interprétabilité des modèles d'IA
 - mais l'analyse des unités élémentaires d'un réseau de neurones ne parvient pas à apporter du sens dans la globalité du modèle
 - placer des sondes dans des neurones individuels ne dit rien sur le fonctionnement global du cerveau
 - de la même façon, Œdipe n'était pas parvenu à interpréter correctement le sens de la prophétie et s'en préserver en n'ayant que des informations partielles sur le destin qui lui était réservé
 - la découverte de la vérité (il est issu de la lignée des Labdacides, ceux qui ne marchent pas droit → donc les « biais » sont en lui) est impossible à supporter → il préfère ne plus jamais voir
 - l'analyse des modèles d'IA montre que l'apprentissage automatique va faire ressortir des choses inavouables → biais de Compas dans la prédiction du risque de récidive des prisonniers afro-américains

Œdipe, ou l'incarnation des programmes d'intelligence artificielle

• L'exil et la fin de vie à Colone

- Œdipe comprit qu'il représentait la souillure qui devait être expulsée de la ville et se bannit lui-même de la cité
- Ses enfants le conduisirent jusqu'aux portes de la ville mais seule sa fille Antigone décida de l'accompagner
- Les garçons Étéocle et Polynice, furieux de représenter le fruit d'un inceste, chassèrent leur père de Thèbes et réclamèrent le pouvoir suprême
- Œdipe les mit alors en garde : s'ils humiliaient leur père, c'est l'épée à la main qu'ils devraient se partager son royaume



Œdipe,

ou l'incarnation des programmes d'intelligence artificielle

• L'exil et la fin de vie à Colone

- Œdipe partit en errance dans la Grèce et finit par se réfugier à Colone, près de la cité d'Athènes, gouvernée par Thésée qui lui offrit asile
- En accueillant Œdipe, Thésée cherchait à protéger son royaume car il avait appris que la terre sur laquelle serait enterrée la dépouille d'Œdipe serait bénie des dieux et connaîtrait la prospérité
- Au bout d'une année, Œdipe et Antigone furent rejoints par Ismène, l'autre fille d'Œdipe, qui vint leur donner des nouvelles
- Ismène demanda à son père de revenir à Thèbes car ses deux frères Étéocle et Polynice se disputaient le pouvoir à Thèbes
- Polynice, qui s'estimait injustement écarté du pouvoir par son frère aîné, souhaitait que son père puisse trancher en indiquant quel était celui de ses fils qui était le plus digne de régner sur Thèbes

Œdipe,

ou l'incarnation des programmes d'intelligence artificielle

• L'exil et la fin de vie à Colone

- Œdipe refusa cependant de prendre parti car il ne se sentait ni assez fort ni assez légitime pour intervenir encore dans les affaires des hommes
- Puis sentant son heure venue, Œdipe alla se purifier
- Il demanda à ses filles de s'éloigner afin de ne rester qu'avec la seule compagne du roi Thésée à qui il révéla un secret, puis il disparut de façon merveilleuse, comme avalé par la terre et la fécondant par son expérience de la souffrance, ce qui apporta à Athènes la prospérité

Œdipe,

ou l'incarnation des programmes d'intelligence artificielle

• Les limites des systèmes intelligents

- la fin de la vie d'Œdipe est marquée par la sagesse
- on reconnaît en Œdipe un pouvoir de décision (dans le conflit qui oppose ses 2 fils) mais il refuse d'en faire usage, il connaît ses limites
- Œdipe considère que sa vie est arrivée à son terme, il confie une dernière information aux hommes, en la personne de Thésée, et livre son âme aux dieux, emmené par un envoyé divin dans les cieux ou accueilli par la terre
- cette mort fabuleuse apportera une bénédiction à Athènes, qui rayonnera pendant des siècles
- Œdipe, dont la vie a été marquée par la tragédie, se réconcilie avec l'univers à la fin de son existence et connaît une belle mort dont la bénédiction s'étend à la terre qui l'a accueilli

Œdipe,

ou l'incarnation des programmes d'intelligence artificielle

- Les limites des systèmes intelligents

- il ne faut utiliser l'intelligence artificielle que dans les domaines où elle est à même de s'appliquer et renoncer à l'employer dans des cadres qui dépassent son mode de fonctionnement
- dicton populaire :
« quand on a pour seul outil un marteau,
tous les problèmes ressemblent à des clous »...
- croire en l'usage de l'intelligence artificielle comme solution à tous les problèmes, c'est vivre dans l'illusion que l'IA est un marteau universel et ne pas comprendre sa portée et ses limites

Antigone, ou les principes éthiques dans l'intelligence artificielle

- Αντιγονη, ou la compensation de la généalogie
 - Après la mort de leur père, Antigone et Ismène décidèrent de rentrer à Thèbes
 - étymologie du nom « Antigone » :
anti- qui signifie « en opposition à » ou « en compensation de » et de *gone*, dérivé de « *genos* », « la génération », « la semence », ou « *gonos* », γονος, la progéniture (→ un *gône* en parler lyonnais)
 - les Labdacides sont maudits depuis le crime de Laïos, le violeur homosexuel, et Œdipe, le parricide, régicide et incestueux
 - Antigone se présente comme une figure qui remet de l'ordre, qui répare, qui arrive en compensation de l'action de ses géniteurs
 - Antigone n'est que loyauté envers les membres de sa famille qu'elle cherche à protéger, au péril de sa vie

Antigone, ou les principes éthiques dans l'intelligence artificielle

- La réparation par la vertu
 - les systèmes intelligents peuvent être porteurs de biais pouvant avoir des conséquences néfastes sur la vie individuelle ou la société
 - mais il est possible de les guider pour corriger leur trajectoire
 - tout comme Antigone, qui incarne la vertu, guidait son père aveugle après avoir été banni de Thèbes



Antigone, ou les principes éthiques dans l'intelligence artificielle

• Αντιγόνη, ou la compensation de la généalogie

- Alors qu'Œdipe était parti sur les routes de l'exil, Étéocle et Polynice se disputèrent le pouvoir
- Étéocle trouva une idée pour conjurer la malédiction de leur père : avec son frère, ils se partageraient le pouvoir à tour de rôle, une année chacun
- Comme Étéocle était l'aîné, ce fut lui qui porta en premier la couronne en faisant à son frère la promesse de la lui remettre au bout d'un an
- Au cours de cette année, Polynice quitta Thèbes et se rendit à Argos, dans le Péloponnèse, où il épousa Argie, la fille du roi Adraste, et à qui il remit comme cadeau le collier de sa défunte mère Jocaste, le fameux collier d'Harmonie
- L'année étant passée, Polynice revint à Thèbes pour demander à son frère Étéocle d'être à son tour le roi de la cité, comme convenu

Antigone, ou les principes éthiques dans l'intelligence artificielle

• Αντιγόνη, ou la compensation de la généalogie

- Étéocle était estimé du peuple qui le préférait à son frère belliqueux
- Ayant apprécié l'exercice du pouvoir durant cette année, Étéocle refusa de céder le trône à son frère Polynice et chercha à l'enfermer dans une prison
- Polynice s'échappa de justesse de Thèbes
- Furieux d'avoir été dupé, il envoya sa sœur Ismène à Colone plaider sa cause auprès de son père, mais sans succès
- Polynice retourna à Argos où il expliqua les raisons de son malheur à son beau-père, le roi Adraste, en lui demandant de l'aide
- Thèbes était surnommée « la ville aux sept portes » car de hautes murailles la protégeait et on ne pouvait entrer dans la cité qu'à partir de l'une des sept portes, qui étaient toutes bien gardées

Antigone, ou les principes éthiques dans l'intelligence artificielle

- Αντιγόνη, ou la compensation de la généalogie
 - Polynice obtint du roi Adraste une armée conduite par sept capitaines – les Sept contre Thèbes – qui se mirent en route pour la capitale de la Béotie afin de rétablir Polynice dans son droit
 - Arrivée à Thèbes, l'armée de Polynice prit position devant chacune des sept portes de la cité
 - Antigone essaya alors de raisonner ses frères afin qu'ils ne s'engageassent pas dans une guerre fratricide, mais ses efforts pour les réconcilier s'avérèrent vains
 - Les combats furent rudes et les sept capitaines périrent devant Thèbes avec de nombreux morts dans les deux camps, sans qu'aucune armée ne parvînt à prendre le dessus sur l'autre

Antigone, ou les principes éthiques dans l'intelligence artificielle

- Αντιγονη, ou la compensation de la généalogie
 - Pour éviter un carnage, il fut alors décidé que les deux frères s'affronteraient dans un combat singulier
 - Polynice et Étéocle se battirent avec férocité et finirent par s'entretuer, périssant chacun par l'épée de leur propre frère
 - Les Argiens survivants retournèrent chez eux, dans le Péloponnèse, afin de rendre un dernier hommage aux combattants disparus
 - Après la mort d'Étéocle, Thèbes fut sans souverain, et Créon, frère de Jocaste et oncle des frères ennemis, redevint le régent de la cité



Antigone, ou les principes éthiques dans l'intelligence artificielle

- Αντιγόνη, ou la compensation de la généalogie
 - Considérant que le roi Étéocle était mort en héros en défendant la cité de Thèbes, Créon décida d'ensevelir son neveu dans une cérémonie grandiose
 - En revanche, considérant que Polynice était un traître qui avait pris les armes contre Thèbes en l'attaquant avec une armée ennemie, Créon interdit à quiconque de donner une sépulture à son autre neveu, sous peine de mort, et le corps de Polynice devait être abandonné sur le champ de bataille, à la merci des chiens et des oiseaux, afin de servir d'exemple à quiconque voudrait s'en prendre à la cité de Thèbes

Antigone, ou les principes éthiques dans l'intelligence artificielle

- Αντιγόνη, ou la compensation de la généalogie
 - Pour les Grecs de l'Antiquité, ne pas avoir de sépulture est considéré comme le plus horrible de tous les châtiments
 - Si les honneurs funèbres n'étaient pas rendus à un mort, avec les chants rituels et l'obole déposée sur les yeux du défunt destinée au passeur Charon, son âme était condamnée à errer pour l'éternité sans jamais trouver le repos aux Enfers
 - Antigone ne put accepter que son frère Polynice eût à subir une telle condamnation et vint voir son oncle pour le supplier de revenir sur sa décision
 - Créon refusa de céder à la demande de sa nièce

Antigone, ou les principes éthiques dans l'intelligence artificielle

• Αντιγόνη, ou l'éthique de la conviction

- Antigone, considérant que seuls les dieux avaient le pouvoir de condamner un homme à errer sans repos hors du royaume d'Hadès, mais que ce pouvoir n'était pas aux mains des hommes, ne put se résoudre à accepter cette loi jugée injuste et décida de braver l'interdit de Créon
- Elle demanda de l'aide à sa sœur Ismène, mais celle-ci refusa, n'osant défier Créon et ses lois et craignant pour sa vie
- À la nuit tombée, Antigone prit une torche pour retrouver la dépouille de son frère, accompagnée de sa belle-sœur Argie (la veuve de Polynice), et les deux femmes rendirent à Polynice les derniers devoirs
- Elles pratiquèrent les rituels requis sur le défunt mais à peine celui-ci avait-il été enterré qu'Antigone et Argie se firent surprendre par la garde de Thèbes

Antigone, ou les principes éthiques dans l'intelligence artificielle

- Αντιγονη, ou l'éthique de la conviction



- Sans opposer de résistance, Antigone et Argie furent conduites devant Créon qui leur rappela la gravité de leur acte et les condamna à être emmurées vivantes dans le tombeau des Labdacides

Antigone, ou les principes éthiques dans l'intelligence artificielle

• Αντιγόνη, ou l'éthique de la conviction

- Apprenant cette sentence, Hémon, fils de Créon et fiancé de sa cousine Antigone, implora la clémence à son père et il alla retrouver sa promise
- Tirésias, le devin respecté de tous, prophétisa à Créon son malheur prochain pour ne pas avoir voulu respecter les lois naturelles édictées par les dieux, ce qui poussa le régent à revenir sur sa décision
- Il était trop tard : Créon découvrit que son fils Hémon s'était donné la mort devant le tombeau et ne trouva plus âme qui vive en faisant ouvrir le caveau funéraire des Labdacides car, pour abréger leur supplice, Antigone et Argie s'étaient pendues avec les ceintures de leurs robes
- Avec ses lois et ses décrets, Créon se retrouva ainsi dans la plus tragique des solitudes
- Ismène ne survécut pas longtemps à sa sœur, et avec elle s'éteignit la lignée maudite des Labdacides

Antigone, ou les principes éthiques dans l'intelligence artificielle

- Αντιγονη, ou l'éthique de la conviction
 - Dix ans plus tard, les Épigones (→ « les descendants »), fils des chefs qui avaient péri en combattant contre Thèbes pour aider Polynice dans sa guerre de succession, décidèrent de venger leurs pères
 - Sous le commandement du roi Adraste, le beau-père de Polynice, ils menèrent une nouvelle expédition contre Thèbes et leurs alliés et parvinrent à les vaincre à la bataille de Glisas, en Béotie.



Antigone, ou les principes éthiques dans l'intelligence artificielle

• Αντιγόνη, ou l'éthique de la conviction

- pour le sociologue Max Weber, il y a deux approches de l'éthique :
 1. l'éthique de la conviction → consiste à effectuer ce que notre conscience nous dicte, ainsi que ce que l'on estime devoir faire
 2. l'éthique de la responsabilité → consiste à agir en tenant compte des conséquences de nos actes
- Ismène représente l'éthique de la responsabilité : elle accepte la loi de Créon car elle a peur pour sa vie
- Antigone représente l'éthique de la conviction : elle place ses convictions au-dessus du pouvoir
- Créon, redevenu maître de la cité de Thèbes, croit devoir être intransigeant envers Antigone pour le bien du royaume et n'accepte pas de revenir sur sa décision → malheur personnel (sa nièce et son fils se suicident) + malheur de la cité (vengeance des Épigones)

Antigone, ou les principes éthiques dans l'intelligence artificielle

- Réalisation de programmes d'IA éthiques par conception
 - tout comme Antigone, il faut guider de façon éclairée les programmes aveugles et boiteux d'intelligence artificielle qui pourraient autrement nous envoyer dans le fossé, à notre perte
 - nécessité d'apporter une « béquille » permettant à l'IA de marcher droit, de guider les programmes dans le respect des valeurs humaines qui sont celles partagées au sein de l'UE, ainsi que les droits et principes fondamentaux → AI Act
 - Antigone périt, emmurée vive, pour avoir obéi à un principe plus haut que toutes les lois : la morale
 - avec Antigone, quelqu'un vient de se dresser contre le pouvoir absolu → Antigone est une figure de la résistance : au nom de la conscience, elle a payé de sa vie son incroyable audace
 - la loi de la conscience s'oppose à la loi des dieux et à la loi de l'État 79

Conclusion

- La mise en scène de l'IA : plus d'un siècle déjà !
 - des créatures artificielles intelligentes apparaissent dans la pièce de théâtre *R.U.R.* de Karel Capek (Prague en 1921, New York en 1922)
→ origine du mot « robot » (*robota* = « corvée » en tchèque)
 - les robots sont des androïdes dénués de sensibilité et de sentiments mais, afin d'être plus robustes et plus efficaces, ils sont améliorés par l'ingénieur de l'usine R.U.R. qui les dote de plus d'intelligence
→ cette amélioration aura pour conséquence la révolte des robots qui, en 10 ans, anéantiront l'humanité...
 - dans le mythe œdipien, le parcours initiatique propre aux histoires de héros se double d'un questionnement sur sa véritable nature, l'ignorance d'Œdipe sur ses origines l'emporte sur la voie du tragique
 - // la remise en question du « je pense, donc je suis » cartésien
→ *Blade Runner* de Ridley Scott (1982) : R. Deckard est-il humain ?

Références bibliographiques

- ❑ pièces de théâtre : <https://www.theatre-classique.fr/>
→ œuvres de Sophocle (Œdipe roi, Œdipe à Colone...), Eschyle, Corneille, Racine, Voltaire... (→ Cadmus, Œdipe, Antigone, Thébaïde)
- ❑ Arthur C. CLARKE (1977) *Profiles of the Future: An Inquiry into the Limits of the Possible*. New York: Harper & Row. 1^{ère} parution 1962.
- ❑ Claude LÉVI-STRAUSS (1967). *Les structures élémentaires de la parenté*. Paris: Mouton & Co. 1^{ère} parution en 1947.
- ❑ Cathy O'NEIL (2018). *Algorithmes : la bombe à retardement*. Paris: Les Arènes, 2018. 1^{ère} parution en anglais en 2016.
- ❑ Antoinette ROUVROY (2016). L'art de ne pas changer le monde. *La Revue Nouvelle*, 8 (8):44–50.
- ❑ Jean-Pierre VERNANT (1999). *L'univers, les dieux, les hommes. Récits grecs des origines*. Points. Essais. Paris: Éditions du Seuil.
- ❑ Max WEBER (1963). *Le savant et le politique*. Paris: Union Générale d'Éditions. 1^{ère} parution en 1919.