

# La recherche et l'enseignement en IA et SHS à l'ère de ChatGPT

Ludovic Tanguy<sup>1</sup>, Phillipe Muller<sup>2</sup>, Cassia Trojahn<sup>3</sup>

<sup>1</sup>UT2J/CLLE, <sup>2</sup>UPS/IRIT/ANITI, <sup>3</sup>UT2/IRIT

Cycle de Séminaires IA&SHS  
18 Septembre 2023



EFELIA X

ANITI

# Agenda

1. Introduction
2. IA et langage, le TAL et les modèles de langue
3. Impact sur la recherche (notamment en SHS)
4. Impacts sur l'enseignement

# Agenda

1. Introduction
2. IA et langage, le TAL et les modèles de langue
3. Impact sur la recherche (notamment en SHS)
4. Impacts sur l'enseignement

# L'IA nous a envahis !

L'IA ne date pas d'il y a 10 ans ...

- 1943, le premier modèle mathématique du neurone biologique "Le Perceptron de McCulloch et Pitts"
- 1950, le Test de Turing (capacité d'une machine à imiter la conversation humaine)
- 1956, apparition du terme 'Intelligence Artificielle'
- 1996, Garry Kasparov versus Deep Blue (IA symbolique)

Depuis 2011 environ, l'IA (les réseaux de neurones, en particulier) a connu un essor important (2017, Ke Jie versus AlphaGo).

De la reconnaissance vocale, la traduction automatisée et la génération de textes, en passant par la prédiction de maladies, les techniques d'IA ont atteint des performances spectaculaires ...  
Et récemment, l'IA générative !

# L'IA nous a envahis !

Révolution, buzz, catastrophe, tendance, sous-estimée, sur-estimée, dangereuse, hype, fantastique, imprévisible, inéluctable, magique ...

(Le Nguyen Hoang. Le fabuleux chantier - Rendre l'IA robustement bénéfique, 2019).

Plusieurs définitions, souvent partagées, controversées concernant les notions d'**intelligence**, de **conscience** ... percevoir, intégrer, raisonner, décider (ou puissance de calcul combinée au traitement massive – statistique – de données ?)

... mais surtout concernant leur implication et impact dans la société.

# Une définition très généraliste d'intelligence artificielle (IA)

**Intelligence** : faculté de connaître, de comprendre ; qualité de l'esprit qui comprend et s'adapte facilement.

**Intelligence artificielle** : ensemble des théories et des techniques développant des programmes informatiques complexes capables de simuler certains traits de l'intelligence humaine (raisonnement, apprentissage...).

(Dictionnaire Le Robert)

# Définition d'IA par la Commission Européenne

## Article 3, point 1

... un *système d'intelligence artificielle*, est un logiciel développé au moyen d'approches énumérées à l'annexe I et qui peut ... générer des résultats tels que des contenus, des prédictions, des recommandations ou des décisions influençant les environnements avec lesquels il interagit.

## Annexe 1

- (a) Approches d'apprentissage automatique, y compris d'apprentissage supervisé, non supervisé et par renforcement, utilisant une grande variété de méthodes, y compris l'apprentissage profond.
- (b) Approches fondées sur la logique et les connaissances, y compris la représentation des connaissances, la programmation inductive (logique), les bases de connaissances, les moteurs d'inférence et de déduction, le raisonnement (symbolique) et les systèmes experts.
- (c) Approches statistiques, estimation bayésienne, méthodes de recherche et d'optimisation.

(Inspiré de L. Risser "Définir l'Intelligence Artificielle", IA et enseignement supérieur : des opportunités pour une pédagogie innovante, 13/06/2023)

# Différentes approches : IA symbolique

S'appuie sur la représentation des faits, une base de règles et un moteur d'inférence.

Woman	≡	Person $\sqcap$ Female
Man	≡	Person $\sqcap$ $\neg$ Woman
Mother	≡	Woman $\sqcap$ $\exists$ hasChild.Person
Father	≡	Man $\sqcap$ $\exists$ hasChild.Person
Parent	≡	Father $\sqcup$ Mother
Wife	≡	Woman $\sqcap$ $\exists$ hasHusband.Man

# Différentes approches : IA connexionniste

Optimisation des paramètres d'un modèle mathématique à partir d'exemples.

- Simulation d'un neurone biologique
- Fonction d'activation
- Apprentissage (rétro-propagation)

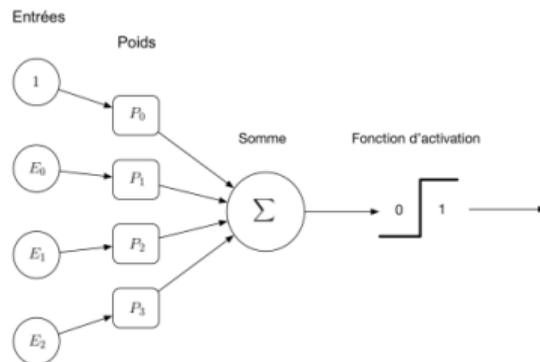
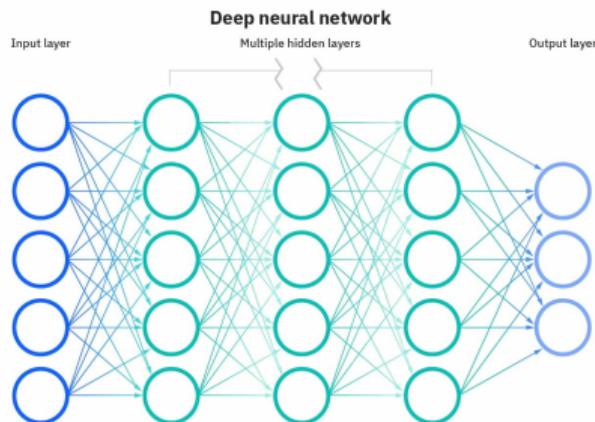


Image: D. Cardon, J. Cointet, A. Mazières. La revanche des neurones : L'invention des machines inductives et la controverse de l'intelligence artificielle. Réseaux : communication, technologie, société, 2018, 5 (211), pp.173-220.

## ChatGPT a plus de 2 milliards de paramètres !



‘GPT-4 is a Transformer-style model pre-trained to predict the next token in a document, using both publicly available data (such as internet data) ... this report contains no further details about the architecture (including model size), hardware, training compute, dataset construction, training method, or similar’ (*OpenAI (2023). GPT-4 Technical Report. ArXiv, abs/2303.08774*).

# Agenda

1. Introduction
2. IA et langage, le TAL et les modèles de langue
3. Impact sur la recherche (notamment en SHS)
4. Impacts sur l'enseignement

# Intelligence Artificielle et Langage

Manipuler le langage : la compétence principale envisagée dès le début de l'IA

Nombreuses tâches et applications : corriger, transcrire, traduire, résumer, donner des instructions, dialoguer...

La jauge ultime de l'IA : le test de Turing

Une discipline spécialisée de l'IA : le Traitement Automatique des Langues (TAL)

Différentes époques :

- Systèmes symboliques : règles établies par des linguistes
- Apprentissage probabiliste : statistiques issues de l'observation guidée de données
- Apprentissage neuronal : systèmes complexes et opaques, peu ou pas de supervision

Décrire et formaliser le langage : la linguistique et le TAL

# Traitement Automatique des langues

Décrire et formaliser le langage : la linguistique et le TAL

Dialogue idéal entre les deux disciplines :

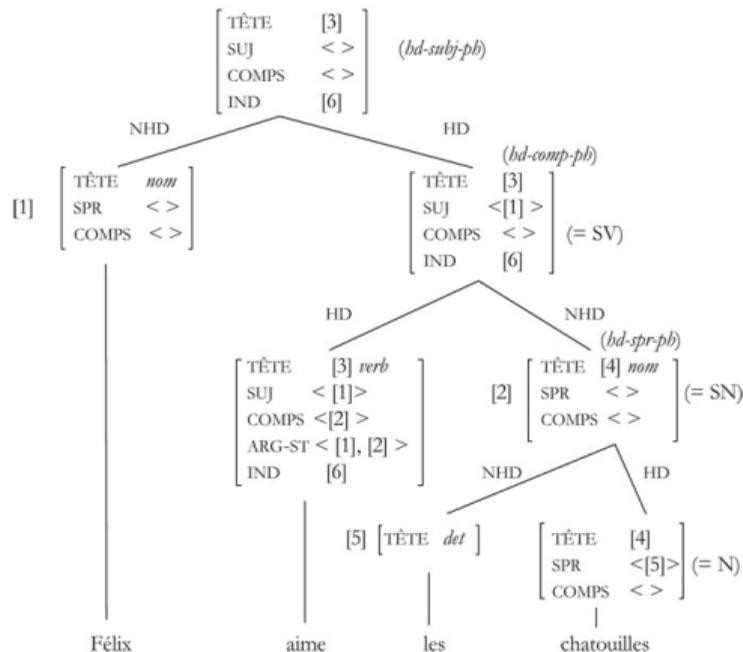
- Les linguistes produisent des connaissances sur le langage qui sont utilisées par les systèmes du TAL
- Les systèmes de TAL aident l'observation et la construction de modèles

# Traitement Automatique des langues

Décrire et formaliser le langage : la linguistique et le TAL

Dialogue idéal entre les deux disciplines :

- Les linguistes produisent des connaissances sur le langage qui sont utilisées par les systèmes du TAL
- Les systèmes de TAL aident l'observation et la construction de modèles

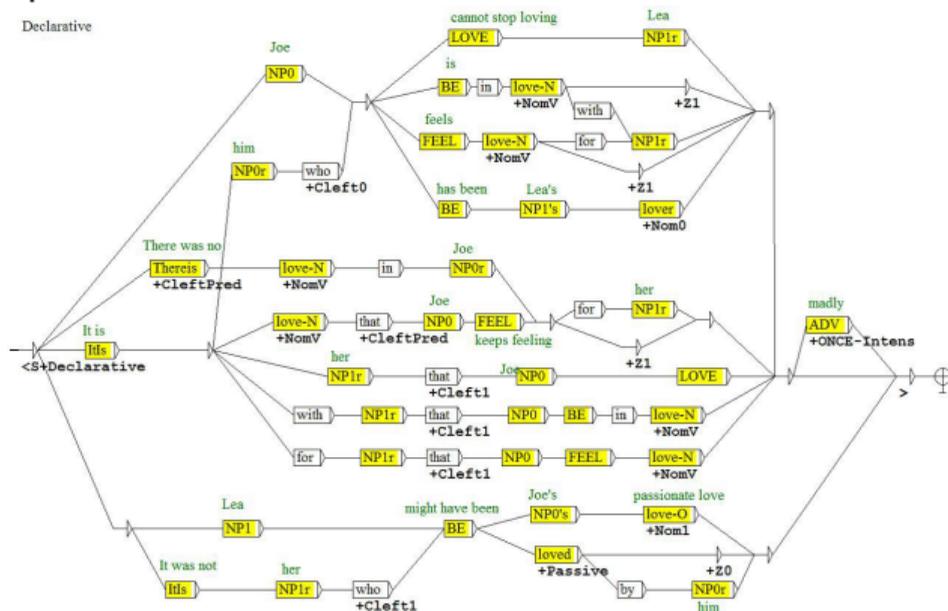


# Traitement Automatique des langues

## Décrire et formaliser le langage : la linguistique et le TAL

Dialogue idéal entre les deux disciplines :

- Les linguistes produisent des connaissances sur le langage qui sont utilisées par les systèmes du TAL
- Les systèmes de TAL aident l'observation et la construction de modèles



# Traitement Automatique des langues

## Décrire et formaliser le langage : la linguistique et le TAL

### Dialogue idéal entre les deux disciplines :

- Les linguistes produisent des connaissances sur le langage qui sont utilisées par les systèmes du TAL
- Les systèmes de TAL aident l'observation et la construction de modèles

```
inépui sée 100 adj [pred="inépui sée 1<Suj:(sn),Objde:(de-scompl|de-sinf|de-sn),Objâ:(â-scompl|â-sinf|â-sn)>",&pers,c at=adj,&]
inépui sées 100 adj [pred="inépui sée 1<Suj:(sn),Objde:(de-scompl|de-sinf|de-sn),Objâ:(â-scompl|â-sinf|â-sn)>",&pers,c at=adj,&]
inépui sés 100 adj [pred="inépui sée 1<Suj:(sn),Objde:(de-scompl|de-sinf|de-sn),Objâ:(â-scompl|â-sinf|â-sn)>",&pers,c at=adj,&]
inéquation 100 nc [pred="inéquation 1<Objde:(de-sinf|de-sn),Objâ:(â-sinf)>",&cat=nc,&fs] inéquation 1 Default fs
inéquations 100 nc [pred="inéquation 1<Objde:(de-sinf|de-sn),Objâ:(â-sinf)>",&cat=nc,&fp] inéquation 1 Default fp
inéquitable 100 adj [pred="inéquitable 1<Suj:(de-sinf|scompl|sn)>",&impers,c at=adj,&fs] inéquitable 1 Default s &adj_per
inéquitable 100 adj [pred="inéquitable 1<Suj:(scompl|sinf|sn)>",&pers,c at=adj,&fs] inéquitable 1 Default s &adj_per
inéquitablement 100 advp [pred="inéquitablement 1",&clives=+,&cat=advp] inéquitablement 1 Default &default
inéquitablement 100 adv [pred="inéquitablement 1<Suj:(de-sinf|scompl|sn)>",&impers,c at=adj,&fp] inéquitablement 1 Default p
inéquitablement 100 adj [pred="inéquitablement 1<Suj:(scompl|sinf|sn)>",&pers,c at=adj,&fp] inéquitablement 1 Default p &adj
inéquité 100 nc [pred="inéquité 1<Objde:(de-sinf|de-sn),Objâ:(â-sinf)>",&cat=nc,&fs] inéquité 1 Default fs &default
inéquités 100 nc [pred="inéquité 1<Objde:(de-sinf|de-sn),Objâ:(â-sinf)>",&cat=nc,&fp] inéquité 1 Default fp &default
inétendu 100 adj [pred="inétendu 1<Suj:(sn),Objde:(de-scompl|de-sinf|de-sn),Objâ:(â-scompl|â-sinf|â-sn)>",&pers,c at=adj,&]
inétendus 100 adj [pred="inétendu 1<Suj:(sn),Objde:(de-scompl|de-sinf|de-sn),Objâ:(â-scompl|â-sinf|â-sn)>",&pers,c at=adj,&]
inétendus 100 adj [pred="inétendu 1<Suj:(sn),Objde:(de-scompl|de-sinf|de-sn),Objâ:(â-scompl|â-sinf|â-sn)>",&pers,c at=adj,&]
inétendus 100 adj [pred="inétendu 1<Suj:(sn),Objde:(de-scompl|de-sinf|de-sn),Objâ:(â-scompl|â-sinf|â-sn)>",&pers,c at=adj,&]
inévitable 100 nc [pred="inévitable 1<Objde:(de-sinf|de-sn),Objâ:(â-sinf)>",&cat=nc,&fs] inévitable 1 Defa
inévitable 100 nc [pred="inévitable 1<Objde:(de-sinf|de-sn),Objâ:(â-sinf)>",&cat=nc,&fp] inévitable 1 Defa
inévitable 100 adj [pred="inévitable 1<Suj:(sn),Objde:(de-scompl|de-sinf|de-sn),Objâ:(â-scompl|â-sinf|â-sn)>",&pers,c at=adj
inévitablement 100 advp [pred="inévitablement 1",&detach = +,&modnc = +,&cat=advp] inévitablement 1
inévitablement 100 adj [pred="inévitablement 1<Suj:(sn),Objde:(de-scompl|de-sinf|de-sn),Objâ:(â-scompl|â-sinf|â-sn)>",&pers,c at=adj
ioda 100 v [pred="ioder 1<Suj:cln|scompl|sinf|sn,Obj:(cl|a|sn)>",&pers,c at=v,&J3s] ioder 1 ThirdSing J3s &actif
iodai 100 v [pred="ioder 1<Suj:cln|scompl|sinf|sn,Obj:(cl|a|sn)>",&pers,c at=v,&J1s] ioder 1 Default J1s &actif
iodaient 100 v [pred="ioder 1<Suj:cln|scompl|sinf|sn,Obj:(cl|a|sn)>",&pers,c at=v,&I3p] ioder 1 Default I3p &actif
iodais 100 v [pred="ioder 1<Suj:cln|scompl|sinf|sn,Obj:(cl|a|sn)>",&pers,c at=v,&I12s] ioder 1 Default I12s &actif
iodait 100 v [pred="ioder 1<Suj:cln|scompl|sinf|sn,Obj:(cl|a|sn)>",&pers,c at=v,&I13s] ioder 1 ThirdSing I3s &actif
iodant 100 v [pred="ioder 1<Suj:cln|scompl|sinf|sn,Obj:(cl|a|sn)>",&pers,c at=v,&G] ioder 1 Default G &actif
iodas 100 v [pred="ioder 1<Suj:cln|scompl|sinf|sn,Obj:(cl|a|sn)>",&pers,c at=v,&J2e] ioder 1 Default J2s &actif
iodasse 100 v [pred="ioder 1<Suj:cln|scompl|sinf|sn,Obj:(cl|a|sn)>",&pers,c at=v,&T1s] ioder 1 Default T1s &actif
iodassent 100 v [pred="ioder 1<Suj:cln|scompl|sinf|sn,Obj:(cl|a|sn)>",&pers,c at=v,&T3p] ioder 1 Default T3p &actif
iodasses 100 v [pred="ioder 1<Suj:cln|scompl|sinf|sn,Obj:(cl|a|sn)>",&pers,c at=v,&T2s] ioder 1 Default T2s &actif
iodassiez 100 v [pred="ioder 1<Suj:cln|scompl|sinf|sn,Obj:(cl|a|sn)>",&pers,c at=v,&T2p] ioder 1 Default T2p &actif
iodassions 100 v [pred="ioder 1<Suj:cln|scompl|sinf|sn,Obj:(cl|a|sn)>",&pers,c at=v,&T1p] ioder 1 Default T1p &actif
iodate 100 nc [pred="iodate 1<Objde:(de-sinf|de-sn),Objâ:(â-sinf)>",&cat=nc,&ns] iodate 1 Default ns &default
iodates 100 nc [pred="iodate 1<Objde:(de-sinf|de-sn),Objâ:(â-sinf)>",&cat=nc,&fp] iodate 1 Default fp &default
iode 100 nc [pred="iode 1<Objde:(de-sinf|de-sn),Objâ:(â-sinf)>",&cat=nc,&ns] iode 1 Default ns &default
iode 100 v [pred="ioder 1<Suj:cln|scompl|sinf|sn,Obj:(cl|a|sn)>",&imperative,&pers,c at=v,&I2s] ioder 1 Imperative
iode 100 v [pred="ioder 1<Suj:cln|scompl|sinf|sn,Obj:(cl|a|sn)>",&pers,c at=v,&P13s] ioder 1 ThirdSing P13s
iodent 100 v [pred="ioder 1<Suj:cln|scompl|sinf|sn,Obj:(cl|a|sn)>",&pers,c at=v,&P3p] ioder 1 Default P3p &actif
```

# Les modèles de langue

Le degré zéro de la modélisation du langage.

Modèle statistique des séquences d'unités (mots, sons...) = n-grammes

# Les modèles de langue

Le degré zéro de la modélisation du langage.

Modèle statistique des séquences d'unités (mots, sons...) = n-grammes

le petit

# Les modèles de langue

Le degré zéro de la modélisation du langage.

Modèle statistique des séquences d'unités (mots, sons...) = n-grammes

le petit matin 0.06165

# Les modèles de langue

Le degré zéro de la modélisation du langage.

Modèle statistique des séquences d'unités (mots, sons...) = n-grammes

le	petit	matin	0.06165
		séminaire	0.05956

# Les modèles de langue

Le degré zéro de la modélisation du langage.

Modèle statistique des séquences d'unités (mots, sons...) = n-grammes

le	petit	matin	0.06165
		séminaire	0.05956
		ami	0.03030

# Les modèles de langue

Le degré zéro de la modélisation du langage.

Modèle statistique des séquences d'unités (mots, sons...) = n-grammes

le	petit	matin	0.06165
		séminaire	0.05956
		ami	0.03030
		village	0.05120

# Les modèles de langue

Le degré zéro de la modélisation du langage.

Modèle statistique des séquences d'unités (mots, sons...) = n-grammes

le	petit	matin	0.06165
		séminaire	0.05956
		ami	0.03030
		village	0.05120
		monde	0.00208
		...	...

## Génération aléatoire de texte (3-grammes de la Wikipédia)

Le petit

# Les modèles de langue

Le degré zéro de la modélisation du langage.

Modèle statistique des séquences d'unités (mots, sons...) = n-grammes

le	petit	matin	0.06165
		séminaire	0.05956
		ami	0.03030
		village	0.05120
		monde	0.00208
		...	...

## Génération aléatoire de texte (3-grammes de la Wikipédia)

Le petit monde

# Les modèles de langue

Le degré zéro de la modélisation du langage.

Modèle statistique des séquences d'unités (mots, sons...) = n-grammes

le	petit	matin	0.06165
		séminaire	0.05956
		ami	0.03030
		village	0.05120
		monde	0.00208
		...	...
petit	monde	de	0.26462
		.	0.13125
		vit	0.13003
		va	0.06648
		...	...

Génération aléatoire de texte  
(3-grammes de la Wikipédia)

Le petit monde

# Les modèles de langue

Le degré zéro de la modélisation du langage.

Modèle statistique des séquences d'unités (mots, sons...) = n-grammes

le	petit	matin	0.06165
		séminaire	0.05956
		ami	0.03030
		village	0.05120
		monde	0.00208
		...	...
petit	monde	de	0.26462
		.	0.13125
		vit	0.13003
		va	0.06648
		...	...

## Génération aléatoire de texte (3-grammes de la Wikipédia)

Le petit monde de

# Les modèles de langue

Le degré zéro de la modélisation du langage.

Modèle statistique des séquences d'unités (mots, sons...) = n-grammes

le	petit	matin	0.06165
		séminaire	0.05956
		ami	0.03030
		village	0.05120
		monde	0.00208
		...	...
petit	monde	de	0.26462
		.	0.13125
		vit	0.13003
		va	0.06648
		...	...

Génération aléatoire de texte  
(3-grammes de la Wikipédia)

Le petit monde de BMX

# Les modèles de langue

Le degré zéro de la modélisation du langage.

Modèle statistique des séquences d'unités (mots, sons...) = n-grammes

le	petit	matin	0.06165
		séminaire	0.05956
		ami	0.03030
		village	0.05120
		monde	0.00208
		...	...
petit	monde	de	0.26462
		.	0.13125
		vit	0.13003
		va	0.06648
		...	...

Génération aléatoire de texte  
(3-grammes de la Wikipédia)

Le petit monde de BMX en

# Les modèles de langue

Le degré zéro de la modélisation du langage.

Modèle statistique des séquences d'unités (mots, sons...) = n-grammes

le	petit	matin	0.06165
		séminaire	0.05956
		ami	0.03030
		village	0.05120
		monde	0.00208
		...	...
petit	monde	de	0.26462
		.	0.13125
		vit	0.13003
		va	0.06648
		...	...

Génération aléatoire de texte  
(3-grammes de la Wikipédia)

Le petit monde de BMX en 1988

# Les modèles de langue

Le degré zéro de la modélisation du langage.

Modèle statistique des séquences d'unités (mots, sons...) = n-grammes

le	petit	matin	0.06165
		séminaire	0.05956
		ami	0.03030
		village	0.05120
		monde	0.00208
		...	...
petit	monde	de	0.26462
		.	0.13125
		vit	0.13003
		va	0.06648
		...	...

## Génération aléatoire de texte (3-grammes de la Wikipédia)

Le petit monde de BMX en 1988 pour

# Les modèles de langue

Le degré zéro de la modélisation du langage.

Modèle statistique des séquences d'unités (mots, sons...) = n-grammes

le	petit	matin	0.06165
		séminaire	0.05956
		ami	0.03030
		village	0.05120
		monde	0.00208
		...	...
petit	monde	de	0.26462
		.	0.13125
		vit	0.13003
		va	0.06648
		...	...

## Génération aléatoire de texte (3-grammes de la Wikipédia)

Le petit monde de BMX en 1988 pour le

# Les modèles de langue

Le degré zéro de la modélisation du langage.

Modèle statistique des séquences d'unités (mots, sons...) = n-grammes

le	petit	matin	0.06165
		séminaire	0.05956
		ami	0.03030
		village	0.05120
		monde	0.00208
		...	...
petit	monde	de	0.26462
		.	0.13125
		vit	0.13003
		va	0.06648
		...	...

## Génération aléatoire de texte (3-grammes de la Wikipédia)

Le petit monde de BMX en 1988 pour le jardinage

# Les modèles de langue

Le degré zéro de la modélisation du langage.

Modèle statistique des séquences d'unités (mots, sons...) = n-grammes

le	petit	matin	0.06165
		séminaire	0.05956
		ami	0.03030
		village	0.05120
		monde	0.00208
		...	...
petit	monde	de	0.26462
		.	0.13125
		vit	0.13003
		va	0.06648
		...	...

## Génération aléatoire de texte (3-grammes de la Wikipédia)

Le petit monde de BMX en 1988 pour le jardinage  
et

# Les modèles de langue

Le degré zéro de la modélisation du langage.

Modèle statistique des séquences d'unités (mots, sons...) = n-grammes

le	petit	matin	0.06165
		séminaire	0.05956
		ami	0.03030
		village	0.05120
		monde	0.00208
		...	...
petit	monde	de	0.26462
		.	0.13125
		vit	0.13003
		va	0.06648
		...	...

## Génération aléatoire de texte (3-grammes de la Wikipédia)

Le petit monde de BMX en 1988 pour le jardinage  
et s'

# Les modèles de langue

Le degré zéro de la modélisation du langage.

Modèle statistique des séquences d'unités (mots, sons...) = n-grammes

le	petit	matin	0.06165
		séminaire	0.05956
		ami	0.03030
		village	0.05120
		monde	0.00208
		...	...
petit	monde	de	0.26462
		.	0.13125
		vit	0.13003
		va	0.06648
		...	...

## Génération aléatoire de texte (3-grammes de la Wikipédia)

Le petit monde de BMX en 1988 pour le jardinage  
et s'attache

# Les modèles de langue

Le degré zéro de la modélisation du langage.

Modèle statistique des séquences d'unités (mots, sons...) = n-grammes

le	petit	matin	0.06165
		séminaire	0.05956
		ami	0.03030
		village	0.05120
		monde	0.00208
		...	...
petit	monde	de	0.26462
		.	0.13125
		vit	0.13003
		va	0.06648
		...	...

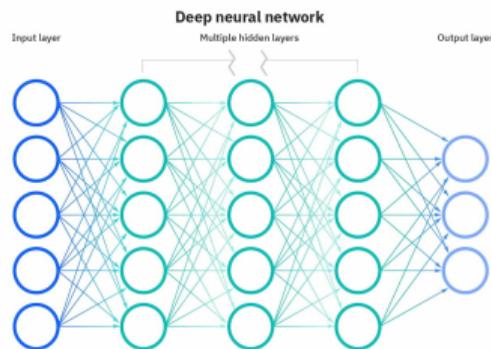
## Génération aléatoire de texte (3-grammes de la Wikipédia)

Le petit monde de BMX en 1988 pour le jardinage et s'attache à démontrer que le barrage Gouin peut entraîner des variations significatives du XIe siècle, beaucoup de lui et sa tour d'angle en bois polychrome et représente l'Espagne, dans la vallée du peuple votent la suspension d'Éric Cantona, Cristiano Ronaldo, Élodie est invitée à dîner par le reggae dub, et son rival.

Exemple d'utilisation : guider les choix. Écriture prédictive, reconnaissance de la parole : préférer "baguette de pain" à "baguette de bain"

# IA génératives : de quoi s'agit-il ?

Même principe que les modèles n-grammes, mais avec bien plus de moyens  
Réseau de neurones profond (deep learning), entraîné à prédire la suite d'une séquence de mots en entrée (modèle autorégressif).



- Pas de connaissances symboliques, uniquement des séquences de mots observées.
- Très grosse capacité de représentation internes : des connexions (poids) établies lors de l'observation (apprentissage)

Différence avec les modèles ngrammes :

- représentations internes des mots et séquences de mots
- capacité à s'abstraire des mots spécifiques
- prise en compte de séquences plus longues (plusieurs centaines/milliers de mots)

# Comment on construit un modèle

- On construit un réseau de neurones vierge : taille, architecture (dizaines de milliards de connexions)
- On lui montre des textes (des centaines de milliards de mots), autrement dit des mots qui se suivent avec une certaine logique
- Les algorithmes d'apprentissage modifient les connexions du réseau pour qu'il soit le plus efficace à prédire le texte qui suit (mobilise des milliers de processeurs pendant plusieurs mois)
- On obtient un modèle générique (*foundation model*) : GPT, Bloom, LLaMa, LaMDA...
- On peut ensuite l'entraîner à faire des choses plus spécifiques : dialoguer (ChatGPT), écrire des programmes (Copilot), analyser des textes scientifiques (Consensus), rechercher de l'information (Bard, Bing)...

Qui est "on" ? Des entreprises dédiées (OpenAI), des branches des GAFAM et bien plus rarement des laboratoires publics (souvent en partenariat)

# Comment on s'en sert ?

Entrée d'un texte (prompt) et récupération du retour (complétion) du modèle

## Exemple de texte continué (ChatGPT)

**Le petit monde** dans lequel nous évoluons peut parfois sembler étroit, mais il est rempli de richesses insoupçonnées. Chaque jour, nous croisons des visages familiers, des amis loyaux, et des lieux chargés d'histoires personnelles. Ce petit monde est le théâtre de nos joies, de nos peines, et de nos aventures les plus mémorables...

- Le texte ainsi produit est une suite plausible du prompt car il respecte les schémas des séquences de mots observées lors de l'apprentissage
- Il est issu d'un processus aléatoire (stochastique).
- La plupart de ces grands modèles ont été entraînés sur des textes écrits dans différentes langues, y compris des langages informatiques. Ils peuvent donc aussi les mélanger.
- Il est possible d'utiliser ces modèles pour effectuer des tâches plus complexes en se basant sur le principe de la complétion de texte

# L'art du prompt

Avec un modèle génératif de base, il est simple de concevoir un texte dont la continuation donnera la réponse à une tâche/question.

## Phrase à compléter

*"Noël a lieu durant le mois de "*  
→ décembre

## Format standard

*"Question : Le crapaud est-il un mammifère ? Réponse :"*  
→ non

## Série : on donne un quelques exemples pour une tâche nouvelle (*few-shot learning*)

*"Animal : chat, groupieux : oui. Animal : chien, groupieux : non. Animal : souris, groupieux : non. Animal : oiseau, groupieux : "*  
→ non

# Les grandes questions

- Le sens : à aucun moment le programme n'a "compris" les textes qu'il a observés : aucun lien n'a été fait avec le contenu ni le contexte  $\Rightarrow$  Ces modèles sont des *perroquets stochastiques* (Bender et al. 2021)
- La fiabilité : les *hallucinations* (une sortie jugée correcte n'est qu'un cas particulier d'hallucination)
- Le fonctionnement : opacité inhérente aux méthodes
- Le coût (machines, temps, énergie) : se compte en dizaines de millions de dollars ; quel accès avec des moyens limités ?
- Les données : d'où viennent-elles, quelle représentation des langues, des cultures ?
- Les biais : ces systèmes reproduisent les schémas qu'ils ont observés, avec leurs tendances majoritaires
- Le contrôle : un système non ouvert (la plupart de ceux qui sont mis à disposition du public) peut censurer/corriger ses réponses

# Agenda

1. Introduction
2. IA et langage, le TAL et les modèles de langue
3. Impact sur la recherche (notamment en SHS)
4. Impacts sur l'enseignement

Usages possibles/constatés pour les disciplines ayant des données textuelles :

- annotation automatique de données
- génération de données textuelles
- transformation de données textuelles

Pour toutes les disciplines :

- (aide à) la rédaction d'article
- aide à la bibliographie

# Annotation automatique de données

**Choose the sentiment of the given text from Positive and Negative.**

**Text:** a feast for the eyes

**Sentiment:** Positive

...

**Text:** boring and obvious

**Sentiment:** Negative

**Text:** [Unlabeled Data]

**Sentiment:** [Label]

Approach	Num. of Samples	Cost (USD)	Time (Mins)	Results
PGDA	3000	11.31	14†	87.75
	6000	22.63	27†	<b>89.29</b>
PGDG	3000	0.91	4†	73.81
	6000	1.83	8†	76.55
DADG	3000	7.18	23†	68.04
	6000	14.37	46†	71.51
Human Labeled	3000	221 - 300	1000	88.47
	67349	4800 - 6700	22740	93.52
PGI	1821	7.33	12	95.77

**Is GPT-3 a Good Data Annotator?** (Ding et al., ACL 2023)

<https://aclanthology.org/2023.acl-long.626/>

cf aussi **Want To Reduce Labeling Cost? GPT-3 Can Help** (Wang et al., EMNLP 21)

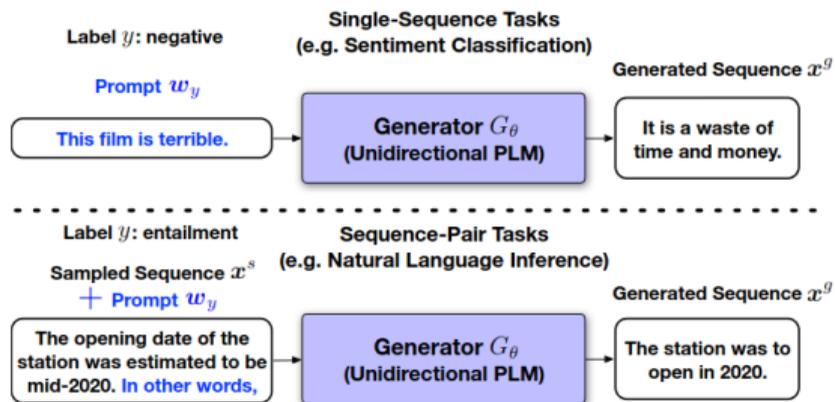
<https://aclanthology.org/2021.findings-emnlp.354/>

# Questions

- faiblesses sur des domaines spécifiques (médical, juridique,...)  
peut évoluer, car il y a beaucoup de travaux dédiés
- robustesse : erreurs + biais systématique (très étudié ... mais pas facile à résoudre)
- traçabilité: modèles fermés / explicabilité  
auto-explicable ? mais problème circulaire de validation
- modèles évolutifs → reproductibilité douteuse
- ok s'il y a une adjudication / un contrôle qualité → intérêt/cout total ?
- ≈ok si c'est utilisé pour entraîner d'autres modèles.

# Génération automatique de données

Dans le cas où la présence de "bruit" est assumée (apprentissage automatique)



**Generating Training Data with Language Models: Towards Zero-Shot Language Understanding.** (Meng et al., NeurIPS 2022) [https://openreview.net/pdf?id=4G1Sfp\\_1sz7](https://openreview.net/pdf?id=4G1Sfp_1sz7)

# Génération automatique de données

Avec les modèles "conversationnels" (ChatGPT, Bard, ...) : il suffit de demander ?

> Generate a review about a terrible movie in a small paragraph:

> "Simply put, this movie was a colossal disappointment. The plot was a confusing mess, the characters were one-dimensional, and the special effects looked like they were from the last century. It's hard to find anything redeeming about this cinematic disaster. Save your money and avoid this one at all costs."

> Generate a review about a mediocre movie in a small paragraph:

> "This movie falls squarely in the realm of mediocrity. It's not particularly bad, but it's far from exceptional. The plot is predictable, and the characters lack depth, making it difficult to fully engage with their journeys. The cinematography is serviceable, but there's nothing visually striking to remember. While it won't leave you feeling outraged, it won't leave a lasting impression either. It's the kind of film you might watch on a lazy Sunday afternoon when you're looking for something to pass the time, but it won't leave you raving about it afterward."

force: *fluency* impressionnante sur petits ou moyens contextes; modification de style

problèmes :

- trouver la bonne amorce (*prompt*)
- contrôle de la distribution/ des biais
- cohérence sur contextes longs
- limites sur composition ("nouveauté" des constructions générées)
- pas clair sur langues moyennement/peu dotées
  - pas forcément un problème si bruit assumé / évalué (approches apprentissage semi-supervisées)
  - pas le cas d'études linguistiques
- ok sur volumétrie faible / facile à évaluer, corriger

# Transformation de textes

ce qui marche assez bien:

- paraphrase
- reformulation/style
- résumé

toujours les mêmes problèmes:

- hallucinations/altération du contenu
- contrôle difficile
- langues peu dotées

# Tout du sol au plafond ?

Exemple dans le monde du sondage

**Employing large language models in survey research** (Jansen et al., Natural language processing Journal, 2023)



Extraits

- (...) developing the survey questions, response options, item construct
- (...) LLMs could suggest appropriate samples and techniques for recruiting participants. As part of Sampling, LLMs can perform intelligent interviewing through conversational AI instead of the conventional survey where text is read and responded to by the respondents.
- (...) LLMs could, perhaps, detect inconsistent and uniform selections, resulting in low-quality entries by analyzing close-ended responses and identifying gibberish and spelling mistakes in open-ended responses.
- Data analysis uses statistical and qualitative methods to analyze the survey data (...)
- Reporting and dissemination (...) LLMs that can implement data science code could help facilitate this process.

# Aide à la recherche ?

Ou bien ... juste de la fraude ?

**Artificial Intelligence Can Generate Fraudulent but Authentic-Looking Scientific Medical Articles: Pandora's Box Has Been Opened**, (Majovsky et al., J Med Internet Res, 2023)

- générer des bibliographies/état de l'art = résumé automatique
- problèmes déjà mentionnés: cohérence, factualité, ...
- du bon côté: peut aider la communication scientifique des non-anglophones (*fluency*)

# Agenda

1. Introduction
2. IA et langage, le TAL et les modèles de langue
3. Impact sur la recherche (notamment en SHS)
4. Impacts sur l'enseignement

# Ignorer, bannir ou s'adapter ?

Les modèles d'IA génératives sont devenus les nouveaux moteurs de recherche (non déterministe ?).

Trois façon d'aborder leur usages :

- Ignorer (approche naïve)
- Bannir (approche naïve)
- S'adapter

# S'adapter

## *Don't Ban ChatGPT in Schools. Teach With It.*

OpenAI's new chatbot is raising fears of cheating on homework, but its potential as an educational tool outweighs its risks.



### ARTIFICIAL INTELLIGENCE

## ChatGPT is going to change education, not destroy it

The narrative around cheating students doesn't tell the whole story. Meet the teachers who think generative AI could actually make learning better.

By Will Douglas Heaven

April 6, 2023

## How one elite university is approaching ChatGPT this school year

Why Yale never considered banning the technology.

By Tate Ryan-Mosley

September 4, 2023

## ChatGPT in Higher Education

Artificial Intelligence and its Pedagogical Value

Rob Rose

ChatGPT in Higher Education explores the positive integration of AI in academia and the ways that it might enhance pedagogy, research, and professional development.

Creative Commons Attribution

CHATGPT IN  
HIGHER  
EDUCATION

Artificial Intelligence and its Pedagogical Value

Rob Rose

Sources : <https://www.nytimes.com/2023/01/12/technology/chatgpt-schools-teachers.html>

Tanguy, Muller and Trojahn

# Comment les utiliser ?

## Des exemples d'utilisation de ChatGPT

Voici quelques exemples d'utilisation de ChatGPT en contexte d'enseignement.

Notez que les réponses générées par ChatGPT peuvent comporter des inexactitudes ou même parfois des informations inventées de toutes pièces. Cette génération fortuite de fausses informations, néanmoins présentées sous une forme plausible, est surnommée [hallucination](#).

-  [Générer des réponses à divers types de questions](#)
-  [Résumer ou synthétiser un article](#)
-  [Écrire une étude de cas sur un sujet précis](#)
-  [Effectuer une revue de littérature](#) (Exemple d'hallucination : toutes les références citée)
-  [Écrire des lignes de codes](#)
-  [Produire une banque de questions sur un sujet](#)
-  Générer des calculs statistiques : [image 1](#) - [image 2](#) - [image 3](#)
-  [Synthétiser et critiquer le contenu d'un guide](#)
-  [Résumer le chapitre d'un livre](#)

etc.

## Évaluation des apprentissages et intégrité académique

Cet outil soulève évidemment certains enjeux et des préoccupations en matière d'évaluation des apprentissages. Certaines méthodes d'évaluation ou certains contextes d'évaluation en ligne sont davantage propices à des possibilités de plagiat avec ChatGPT:

- Examens et quiz en ligne;
- Productions écrites, communément appelées «travaux longs» qui évaluent la compréhension d'un contenu disciplinaire, la compétence linguistique ou la compréhension d'une langue seconde;
- Problèmes mathématiques;
- Programmation en rédigeant du code informatique;
- Etc.

Source : <https://www.enseigner.ulaval.ca/ressources-pedagogiques/chatgpt>

# Comment les utiliser ?

Role <sup>6</sup>	Description	Example of implementation
<b>Possibility engine</b>	AI generates alternative ways of expressing an idea	Students write queries in ChatGPT and use the Regenerate response function to examine alternative responses.
<b>Socratic opponent</b>	AI acts as an opponent to develop and argument	Students enter prompts into ChatGPT following the structure of a conversation or debate. Teachers can ask students to use ChatGPT to prepare for discussions.
<b>Collaboration coach</b>	AI helps groups to research and solve problems together	Working in groups, students use ChatGPT to find out information to complete tasks and assignments.
<b>Guide on the side</b>	AI acts as a guide to navigate physical and conceptual spaces	Teachers use ChatGPT to generate content for classes/courses (e.g., discussion questions) and advice on how to support students in learning specific concepts.
<b>Personal tutor</b>	AI tutors each student and gives immediate feedback on progress	ChatGPT provides personalized feedback to students based on information provided by students or teachers (e.g., test scores).
<b>Co-designer</b>	AI assists throughout the design process	Teachers ask ChatGPT for ideas about designing or updating a curriculum (e.g., rubrics for assessment) and/or focus on specific goals (e.g., how to make the curriculum more accessible).
<b>Exploratorium</b>	AI provides tools to play with, explore and interpret data	Teachers provide basic information to students who write different queries in ChatGPT to find out more. ChatGPT can be used to support language learning.
<b>Study buddy</b>	AI helps the student reflect on learning material	Students explain their current level of understanding to ChatGPT and ask for ways to help them study the material. ChatGPT could also be used to help students prepare for other tasks (e.g., job interviews).
<b>Motivator</b>	AI offers games and challenges to extend learning	Teachers or students ask ChatGPT for ideas about how to extend students' learning after providing a summary of the current level of knowledge (e.g., quizzes, exercises).
<b>Dynamic assessor</b>	AI provides educators with a profile of each student's current knowledge	Students interact with ChatGPT in a tutorial-type dialogue and then ask ChatGPT to produce a summary of their current state of knowledge to share with their teacher/for assessment.

## ChatGPT and Artificial Intelligence in higher education. UNESCO.

<https://www.iesalc.unesco.org/wp-content/uploads/2023/04/>

ChatGPT-and-Artificial-Intelligence-in-higher-education-Quick-Start-guide\_EN\_FINAL.pdf (15/09/2023)

# Les nombreux enjeux d'IA dans l'enseignement

- Qualité, pertinence et validité des données
- Nouvelles compétences / nouvelles pédagogies
- Modalités d'évaluation
- Usage des outils d'IA dans l'enseignement (données d'apprentissage ?) – (ce qui va bien au delà des usages des IA génératives)
- Formation des enseignants aux nouveaux outils / temps de travail - plan de formation
- Évolution du métier d'enseignant / rôle de l'enseignant
- Rôle de l'étudiant / engagement / posture
- Aspects liés au droit d'auteur / plagiat
- Inclusion, inégalités, biais des données
- Manque de régulation
- Impact climatique de ces nouveaux usages
- ...

# Plusieurs débats

## Intelligence artificielle et enseignement supérieur : des opportunités pour une pédagogie innovante - Mardi 13 juin 2023

mer 14/06/2023 - 11:21 | Lisa MARCHETTI

En collaboration avec Toulouse Tech et l'Ecole Française de l'Intelligence Artificielle portée par ANITI, le SiUP a organisé le mardi 13 juin 2023 de 13h30 à 18h un événement autour des liens qui se tissent entre intelligence artificielle et enseignement supérieur.

*Comment fonctionne l'intelligence artificielle ? Quels impacts peut-elle avoir sur l'enseignement ? Comment l'utiliser pour mieux faire apprendre ?*

L'objectif de cette demi-journée était de sensibiliser aux enjeux de l'IA dans l'enseignement supérieur, d'en définir les contours et de présenter plusieurs applications pédagogiques d'IA en cours.

Définition de l'IA par Laurent Risser, Ingénieur de recherche en Sciences de données (CNRS)



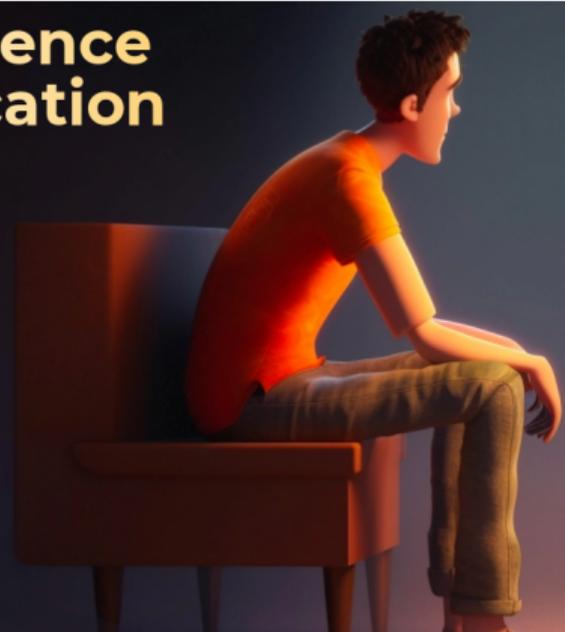
Sources : <https://sia.univ-toulouse.fr/ressources/intelligence-artificielle-et-enseignement-superieur-des-opportunités-pour-une-pédagogie-innovante/>  
<https://iaconference.education/>

## La conférence IA & éducation

8 & 9 juin 2023

au Dock B · Paris Pantin

Rencontres & débats  
Présentations & tables rondes  
Présentations flash  
Prompt party



# À l'UT2J : détection de contenus IA ?

  **ENT**

 **OUTILS NUMÉRIQUES** ↓  **ALADIN**  **MINGUS**    **CASSIA** ↓

**RESSOURCES TECHNIQUES & LOGISTIQUES // IMAGE ET NUMÉRIQUE // NUMÉRIQUE AU SERVICE DE LA FORMATION**

 <b>MON ESPACE</b>	▼
 <b>MES OUTILS</b>	▼
 <b>MES RESSOURCES</b>	▼
 <b>BIBLIOTHÈQUES</b>	▼
 <b>MON INSCRIPTION</b>	▼
 <b>UNIVERSITÉ</b>	▼
 <b>FORMATION &amp; VIE ÉTUDIANTE</b>	▼
 <b>RECHERCHE</b>	▼
 <b>SERVICE COMMUN DE LA DOCUMENTATION</b>	▼
 <b>PERSONNELS</b>	▼

## Compilatio Détection de contenus IA

Publié le 31 mars 2023 – Mis à jour le 4 avril 2023

*La société Compilatio, dont nous utilisons le logiciel de contrôle de similitudes (sensibilisation au plagiat), est en train de développer un logiciel similaire pour la détection de contenus produits par des intelligences artificielles (IA).*

La dernière version performante de ChatGPT a suscité un certain émoi chez les enseignants. Pour autant, aussi performantes soient-elles, ces IA présentent encore des points de faiblesse qui peuvent être repérés par... des IA !

Compilatio propose de tester gratuitement son logiciel en développement. Nous vous invitons à réaliser ces tests et à nous faire un retour d'expérience. L'enjeu ? A terme Compilatio nous proposera d'inclure le logiciel dans son offre actuelle en réévaluant le prix de l'abonnement. Le bilan de vos tests nous permettra de se positionner sur cette offre.

**Les liens utiles :**  
Les infos sur le logiciel Compilatio  
[Le démonstrateur Compilatio](#)

Les retours à adresser à [eric.ferrante@univ-tlse2.fr](mailto:eric.ferrante@univ-tlse2.fr)

