

Méthodes neuronales pour une génération de texte conditionnelle, multilingue et informée

Claire Gardent

CNRS / LORIA, Nancy



Génération de texte conditionnelle

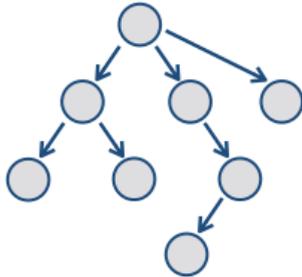
génère du texte à partir de



Bases de données, données tabulaires



Textes



Graphe

... d'images ou de vidéos

A person riding a motorcycle on a dirt road.



Two dogs play in the grass.



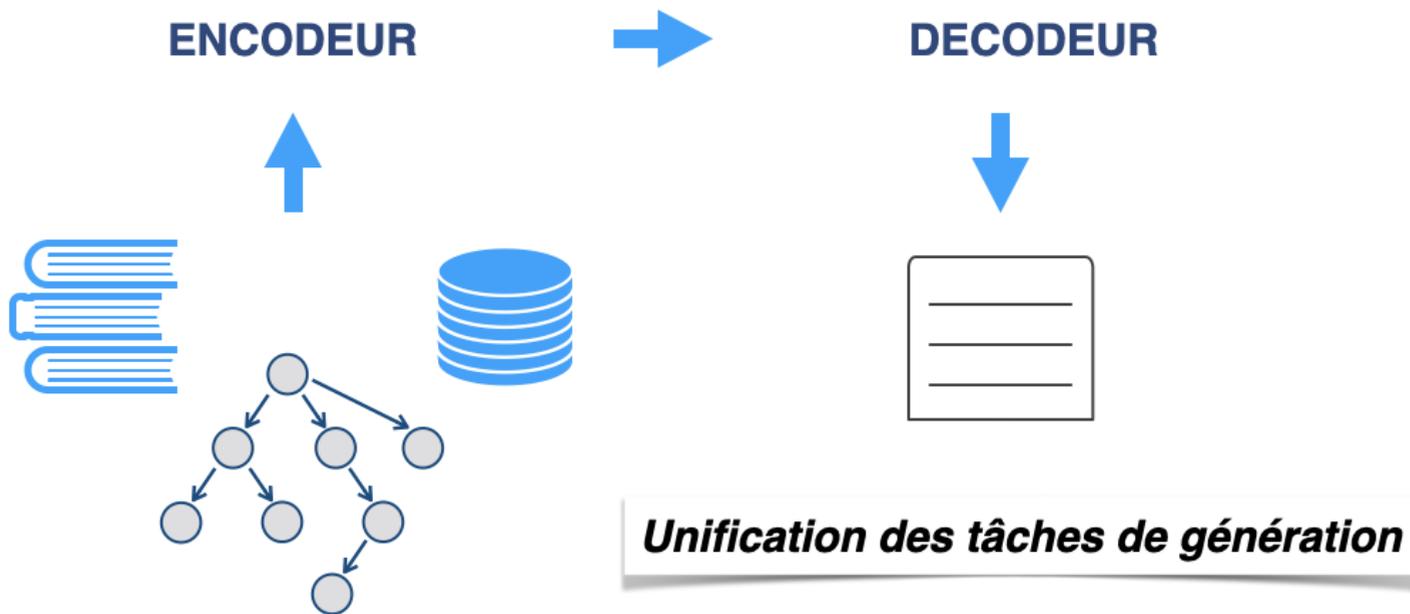
A group of young people playing a game of frisbee.

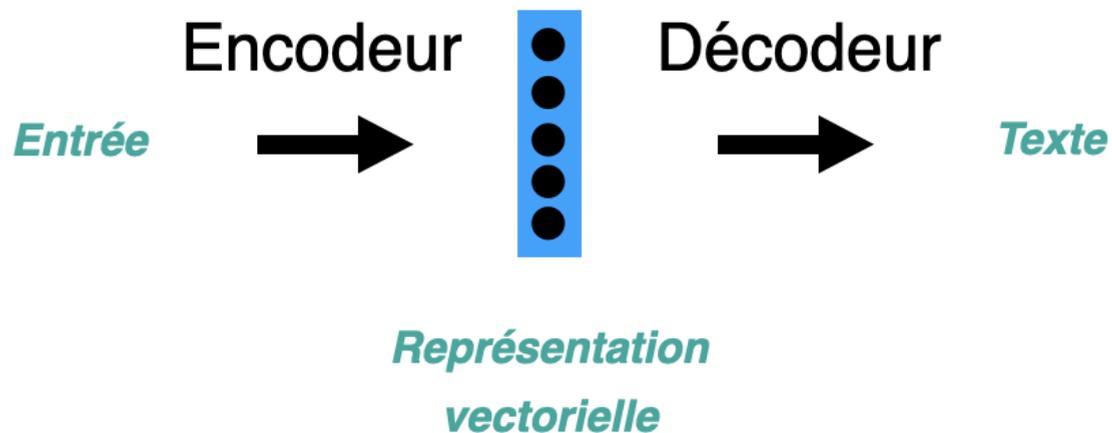


Two hockey players are fighting over the puck.



L'encodeur-décodeur, un modèle dominant





Deux réseaux de neurones

- L'encodeur *crée une représentation vectorielle* de l'entrée
- Le décodeur est un *modèle de langue* qui génère du texte à partir de cette représentation vectorielle

Quelques méthodes neuronales pour une génération de texte conditionnelle, multilingue et informée

Générer en plusieurs langues (Modèles multilingues)

- Plongements et encodeur-décodeur multilingue
- *Example* : Verbaliser des représentations sémantiques abstraites (AMRs) dans les 21 langues de l'UE

Optimiser l'apprentissage (Pré-apprentissage et Affinage)

- Pré-apprentissage
- Affinage (Fine tuning) et affinage efficace (Parameter Efficient Fine Tuning)
- *Examples* : Verbaliser des graphes (RDF, AMR) dans des langues peu dotées

Interfacer génération et recherche d'information (Modèles informés)

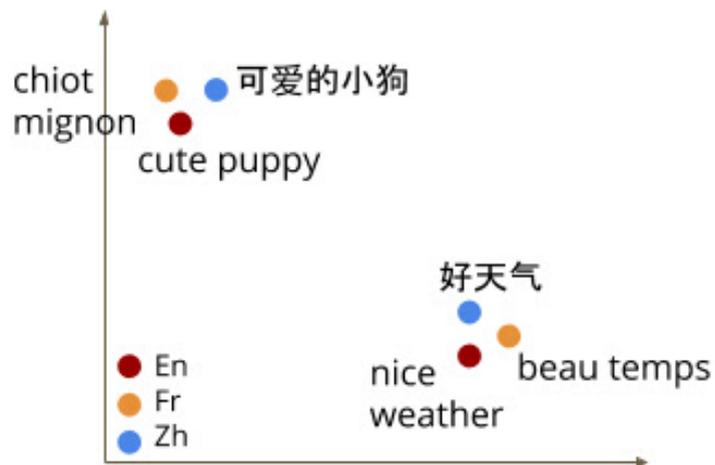
- Retrieval Augmented Generation (RAG)
- *Examples* : Dialogue Humain-Machine, Génération de biographies Wikipedia

Multilinguisme

Générer en plusieurs langues

Génération conditionnelle multilingue

Plongements multilingues

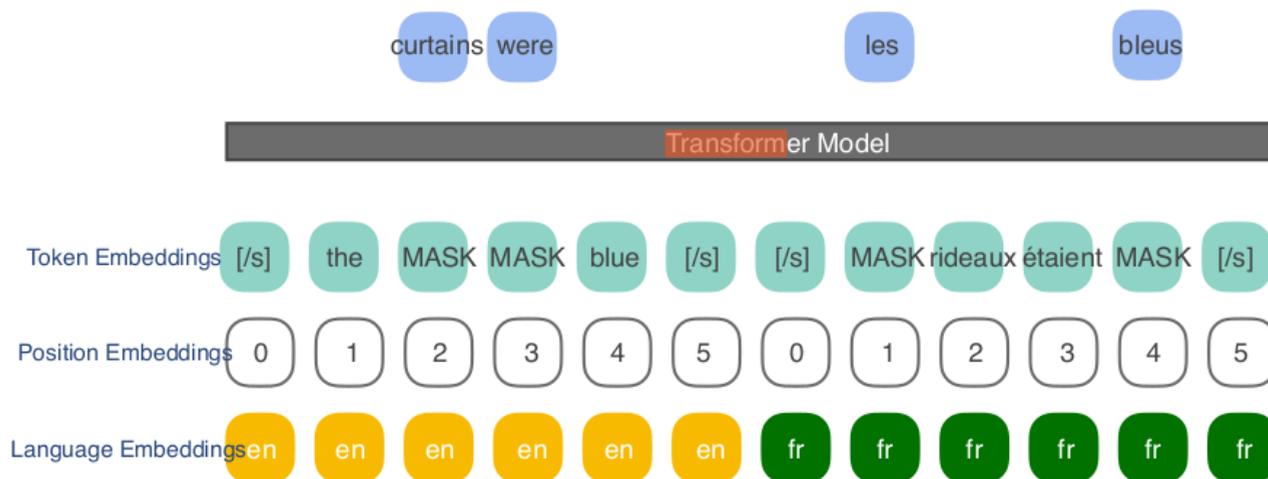


Les mots de différentes langues qui ont le même sens ont des plongements proches dans l'espace sémantique

Plongements multilingues

Représentations à la BERT pré-entraînées sur des données parallèles

Objectif MLM : *Deviner un mot sachant son contexte gauche et droit*



L'auto-attention aide à apprendre l'alignement entre mots similaires

Encodeur-Décodeur Multilingue

Un seul modèle

- Entraîné sur un corpus parallèle multilingue e.g., fr/en, de/fr, es/it ...

Contrôle de la langue cible

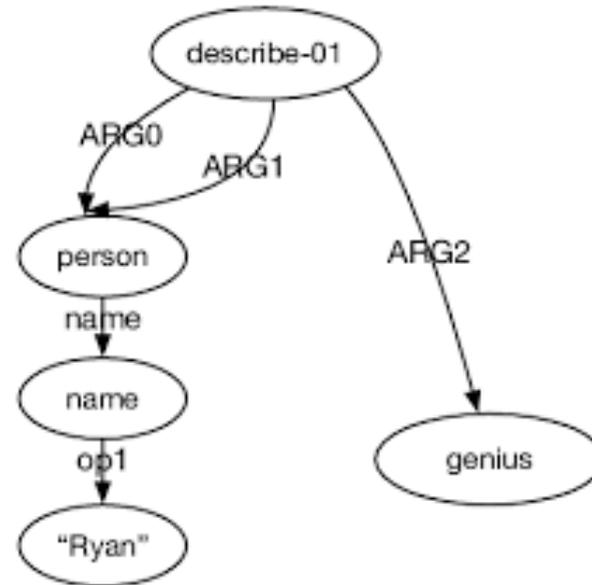
- Chaque instance est préfixée par un *token de contrôle* indiquant la langue cible

AMR --> 21 Languages

Fan and Gardent, EMNLP 2020

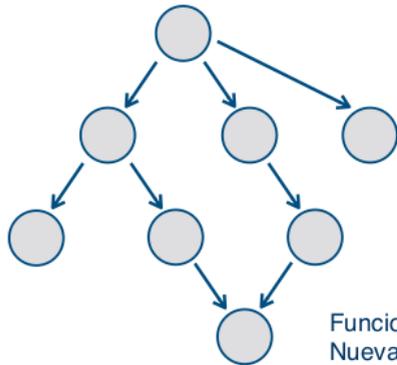
Représentations Sémantiques Abstraites (AMR)

Ryan describes himself as a genius



AMR --> 21 Languages

Amerikanska tjänstemän höll ett expertgruppsmöte i januari 2002 i New York.



Americký predstavitelia usporiadali stretnutie expertnej skupiny v januári 2002 v New Yorku.

US officials held an expert group meeting in January 2002 in New York.

Des responsables américains ont tenu une réunion d'un groupe d'experts en janvier 2002 à New York.

Funcionarios estadounidenses celebraron una reunión de un grupo de expertos en enero de 2002 en Nueva York.

Romance, Germanic, Slavic, Uralic

Encodeur

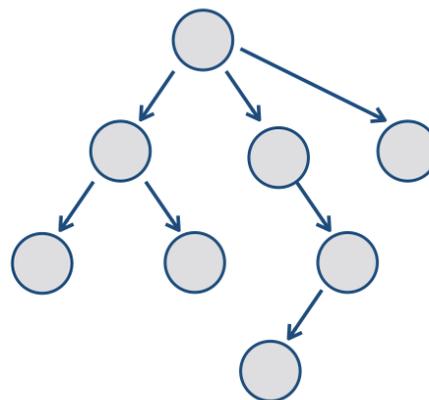
Objectif

Apprendre un encodeur qui crée de bonnes représentations pour les données d'entrée i.e., les graphes AMR

Méthode

- Encodeur de type Transformer
- Linéarisation des AMRs
- Ajout de plongements structurels
- Pré-apprentissage (MLM) sur 30M AMRs argent

Linéarisation



hold

:ARG0 person : ARG0-of have-org-role :ARG1 :op1 **United** :op2

States :ARG2 **official**

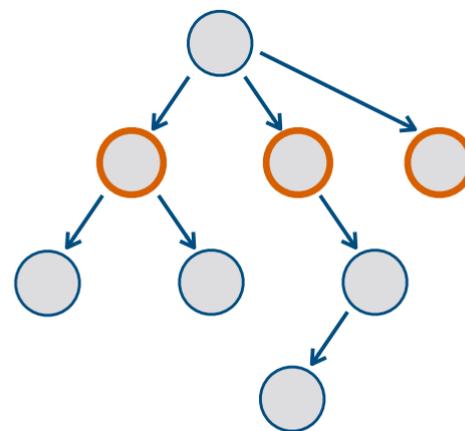
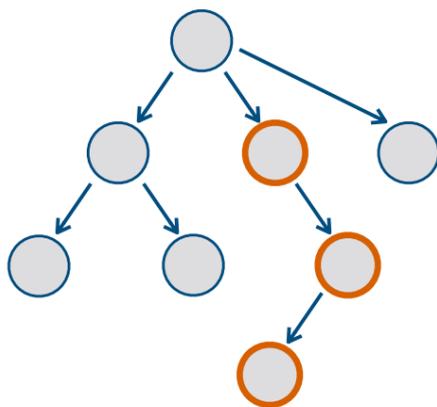
:ARG1 **meet** :ARG0 person :ARG1-of **expert** :ARG2-of **group**

:time date-entity :year **2002** :month **1**

:location city :op1 **New** :op2 **York**

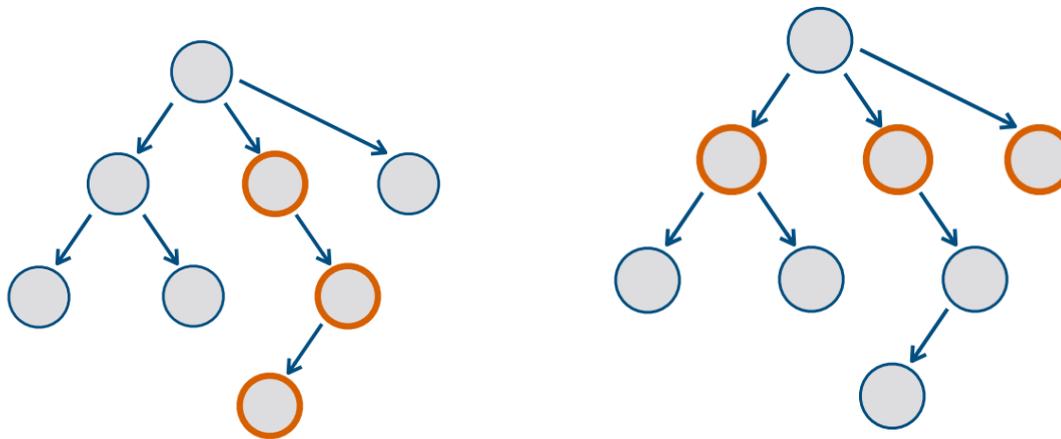
Encoder la structure

Ajout de plongements "branche" et "noeuds frères"



Encoder la structure

Ajout de plongements "branche" et "noeuds frères"



Input	[CLS]	my	dog	is	cute	[SEP]	he	likes	play	#ing	[SEP]
Token Embeddings	$E_{[CLS]}$	E_{my}	E_{dog}	E_{is}	E_{cute}	$E_{[SEP]}$	E_{he}	E_{likes}	E_{play}	$E_{\#ing}$	$E_{[SEP]}$
	+	+	+	+	+	+	+	+	+	+	+
Segment Embeddings	E_A	E_A	E_A	E_A	E_A	E_A	E_B	E_B	E_B	E_B	E_B
	+	+	+	+	+	+	+	+	+	+	+
Position Embeddings	E_0	E_1	E_2	E_3	E_4	E_5	E_6	E_7	E_8	E_9	E_{10}

Encodeur-Décodeur Multilingue

Decoding into Slovak



hold

:ARG0 person : ARG0-of have-org-role :ARG1 :op1
United :op2 States :ARG2 official
:ARG1 meet :ARG0 person :ARG1-of expert :ARG2-
of group
:time date-entity :year 2002 :month 1
:location city :op1 New :op2 York



Amerikanska tjänstemän höll ett
expertgruppsmöte i januari 2002 i New York.

Decoding into French



hold

:ARG0 person : ARG0-of have-org-role :ARG1 :op1
United :op2 States :ARG2 official
:ARG1 meet :ARG0 person :ARG1-of expert :ARG2-
of group
:time date-entity :year 2002 :month 1
:location city :op1 New :op2 York



Des responsables américains ont tenu une
réunion d'un groupe d'experts en janvier 2002 à
New York.

Un seul modèle pour les 21 langues

Données d'apprentissage

hold

:ARG0 person : ARG0-of have-org-role :ARG1 :op1 **United** :op2

States :ARG2 **official**

:ARG1 **meet** :ARG0 person :ARG1-of **expert** :ARG2-of **group**

:time date-entity :year **2002** :month **1**

:location city :op1 **New** :op2 **York**

Des responsables américains ont tenu une réunion d'un groupe d'experts en janvier 2002 à New York.

French

Funcionarios estadounidenses celebraron una reunión de un grupo de expertos en enero de 2002 en Nueva York.

Spanish

Americkí predstavitelia usporiadali stretnutie expertnej skupiny v januári 2002 v New Yorku.

Slovak

Американските служители проведоха среща на експертна група през януари 2002 г. в Ню Йорк.

Bulgarian

Amerikanska tjänstemän höll ett expertgruppsmöte i januari 2002 i New York.

Swedish

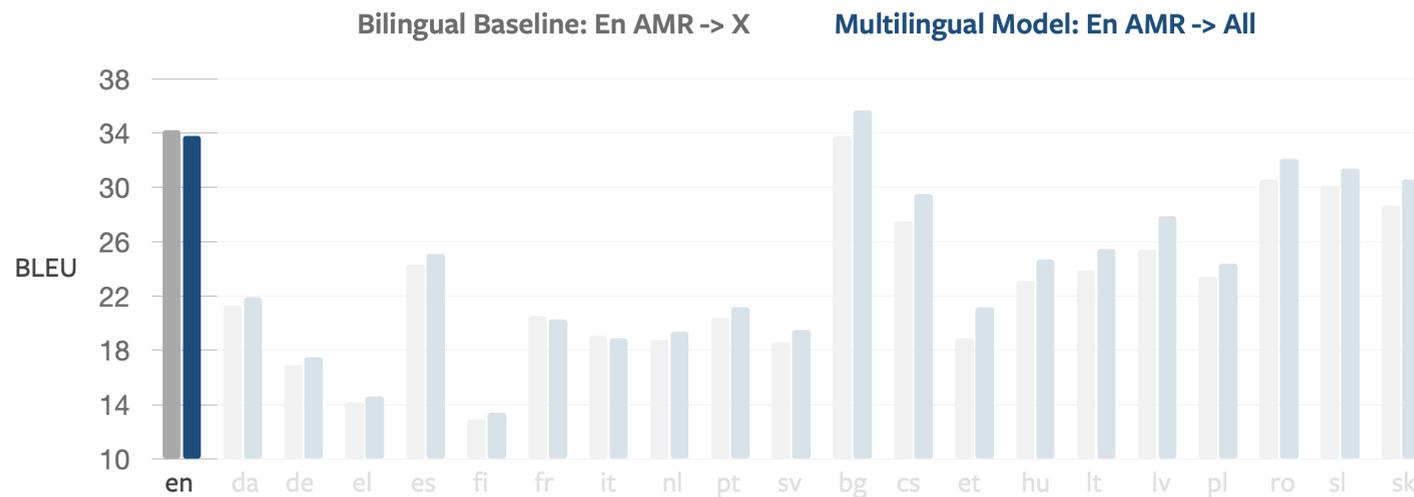
- Europarl: 21 Langues
- AMR: dérivée de l'anglais (analyseur sémantique JAMR)

Données de test

- AMR argent : 21 langues, Europarl
- AMR or : 4 langues

Comparaison: Bilingue vs Multilingue

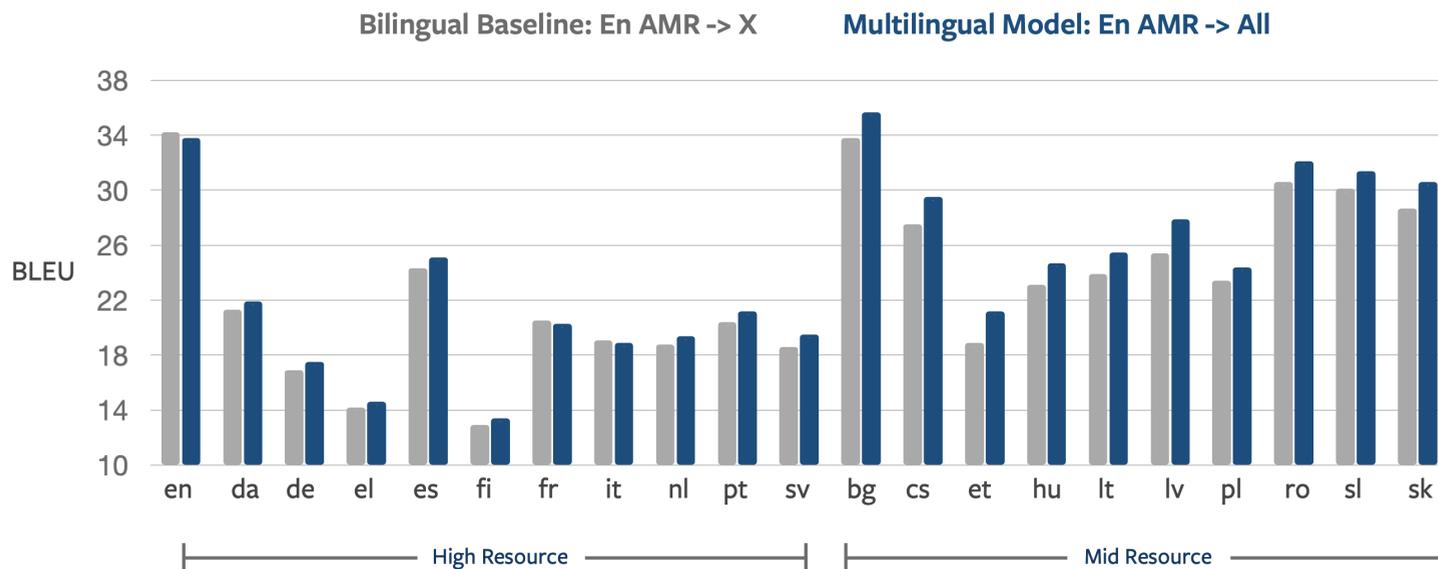
Résultats: Europarl



Sur l'anglais, le modèle bilingue AMR/Anglais est un peu meilleur que le modèle multilingue

Comparaison: Bilingue vs Multilingue

Résultats: Europarl

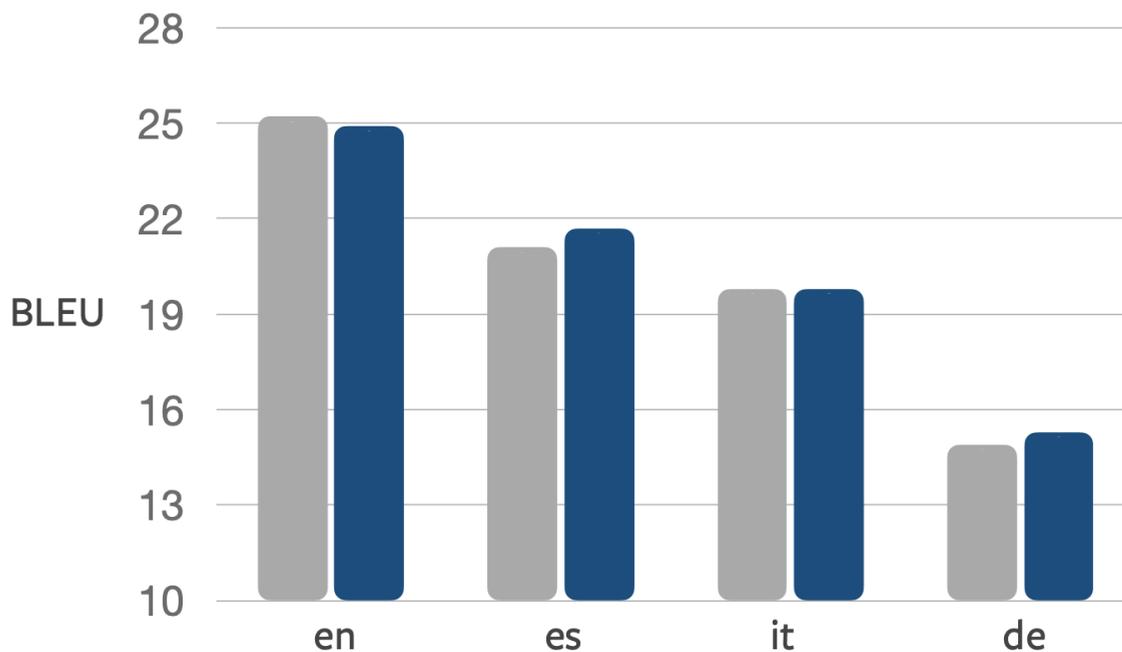


Le modèle multilingue donne de meilleurs résultats pour les langues moins dotées (effet de transfert, plus de données d'apprentissage)

Résultats: AMR Or

Bilingual Baseline: En AMR -> X

Multilingual Model: En AMR -> All



Le modèle multilingue donne des résultats comparables et parfois meilleurs pour les langues bien dotées et les représentations "OR"

Comparaison: Hybride vs Multilingue

Hybrid Translation Model

hold

:ARG0 person : ARG0-of have-org-role :ARG1 :op1
United :op2 States :ARG2 official
:ARG1 meet :ARG0 person :ARG1-of expert :ARG2-of group
:time date-entity :year 2002 :month 1
:location city :op1 New :op2 York



AMR to English

US officials held an expert group
meeting in January 2002 in New York.

Multilingual Model



hold

:ARG0 person : ARG0-of have-org-role :ARG1 :op1
United :op2 States :ARG2 official
:ARG1 meet :ARG0 person :ARG1-of expert :ARG2-of group
:time date-entity :year 2002 :month 1
:location city :op1 New :op2 York



Des responsables américains ont tenu une réunion
d'un groupe d'experts en janvier 2002 à New York.

Comparaison: Hybride vs Multilingue

Hybrid Translation Model

hold
:ARG0 person : ARG0-of have-org-role :ARG1 :op1
United :op2 States :ARG2 official
:ARG1 meet :ARG0 person :ARG1-of expert :ARG2-of group
:time date-entity :year 2002 :month 1
:location city :op1 New :op2 York



AMR to English

US officials held an expert group
meeting in January 2002 in New York.



Translation Model

Des responsables américains ont tenu une réunion
d'un groupe d'experts en janvier 2002 à New York.

Multilingual Model



hold
:ARG0 person : ARG0-of have-org-role :ARG1 :op1
United :op2 States :ARG2 official
:ARG1 meet :ARG0 person :ARG1-of expert :ARG2-of group
:time date-entity :year 2002 :month 1
:location city :op1 New :op2 York



Des responsables américains ont tenu une réunion
d'un groupe d'experts en janvier 2002 à New York.

Comparaison: Hybride vs Multilingue

Hybrid Translation Model

hold
:ARG0 person : ARG0-of have-org-role :ARG1 :op1
United :op2 States :ARG2 official
:ARG1 meet :ARG0 person :ARG1-of expert :ARG2-of group
:time date-entity :year 2002 :month 1
:location city :op1 New :op2 York



AMR to English

US officials held an expert group
meeting in January 2002 in New York.



Translation Model

Des responsables américains ont tenu une réunion
d'un groupe d'experts en janvier 2002 à New York.

Multilingual Model



hold
:ARG0 person : ARG0-of have-org-role :ARG1 :op1
United :op2 States :ARG2 official
:ARG1 meet :ARG0 person :ARG1-of expert :ARG2-of group
:time date-entity :year 2002 :month 1
:location city :op1 New :op2 York

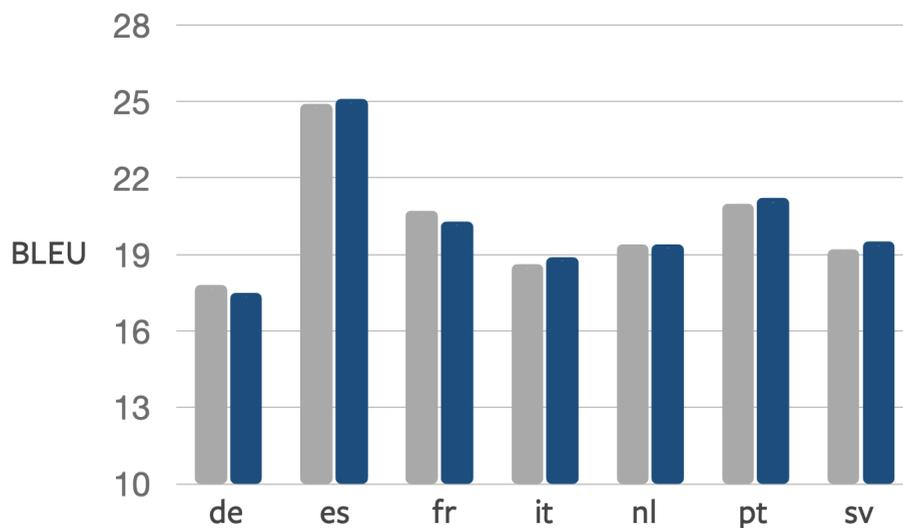


Des responsables américains ont tenu une réunion
d'un groupe d'experts en janvier 2002 à New York.

Comparaison: Hybride vs Multilingue

Hybrid Translation: En AMR -> En -> Translate to X

Multilingual Model: En AMR -> All



Le modèle multilingue donne des résultats comparables et parfois meilleurs aux modèles en cascade

Evaluation par l'humain



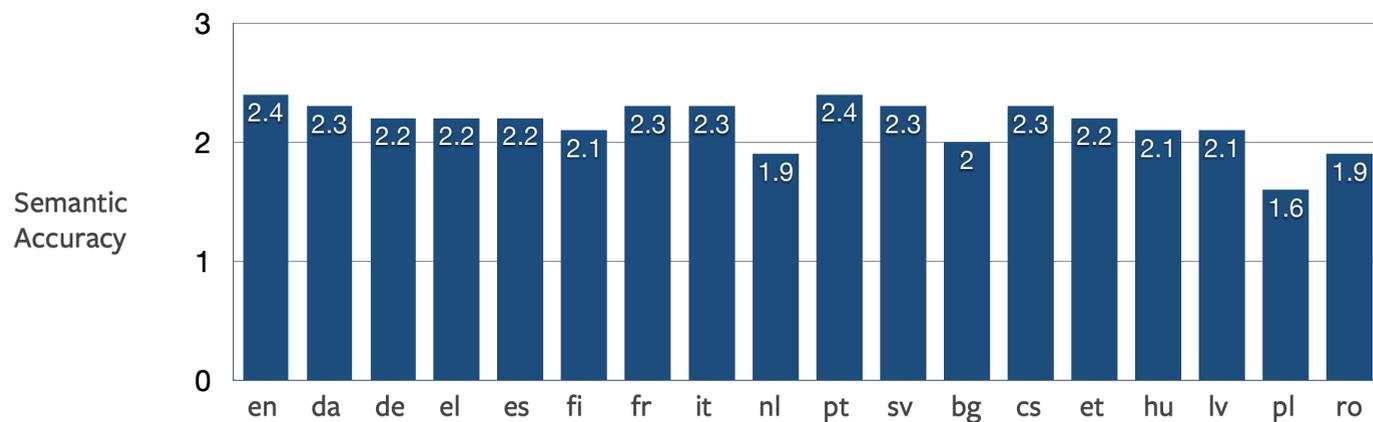
- Evaluateurs: collègues, listes TAL/NLP
- 50 phrases par langue
 - 1/2 BLEU bas
 - 1/2 BLEU élevé

Evaluation par l'humain



- Sémantique :
Le texte généré a-t'il le même sens que la référence ?
- Morphologie:
La morphologie est-elle correcte? Les contraintes d'accord sont-elles respectées e.g., accord verbe/sujet, nom/adjectif ?
- Ordre des mots: L'ordre des mots est il naturel ?

Evaluation par l'humain : Sémantique



Le modèle multilingue préserve la sémantique de l'entrée (2 = différence mineure) pour la plupart des langues

Résumé

L'approche proposée permet de générer à partir d'une AMR anglo-centrique des textes dans 21 langues différentes

Résumé

L'approche proposée permet de générer à partir d'une AMR anglo-centrique des textes dans 21 langues différentes

L'évaluation par l'humain indique que les techniques multilingues permettent de généraliser à 21 langues

Résumé

L'approche proposée permet de générer à partir d'une AMR anglo-centrique des textes dans 21 langues différentes

L'évaluation par l'humain indique que les techniques multilingues permettent de généraliser à 21 langues

Le modèle multilingue bénéficie d'une plus grande quantité de données et donne de meilleurs résultats que les modèles monolingues

Optimiser l'apprentissage

Pré-apprentissage, affinage, prompting

Pré-apprentissage et affinage

Pre-train once, fine-tune many times

Deux idées clé

- *Pré-apprentissage auto-supervisé* sur de grandes quantité de données (modèle générique)
- Affinage pour une tâche/langue donnée : *Apprentissage supervisé* sur des données annotées (modèle ciblant une tâche ou une langue donnée)

Exemples

- BERT: pré-apprentissage de plongements lexicaux
- GPT: pré-apprentissage de modèles de langues
- (m)T5, (m)BART: pré-apprentissage de modèles de type encodeur-décodeur

GPT-2

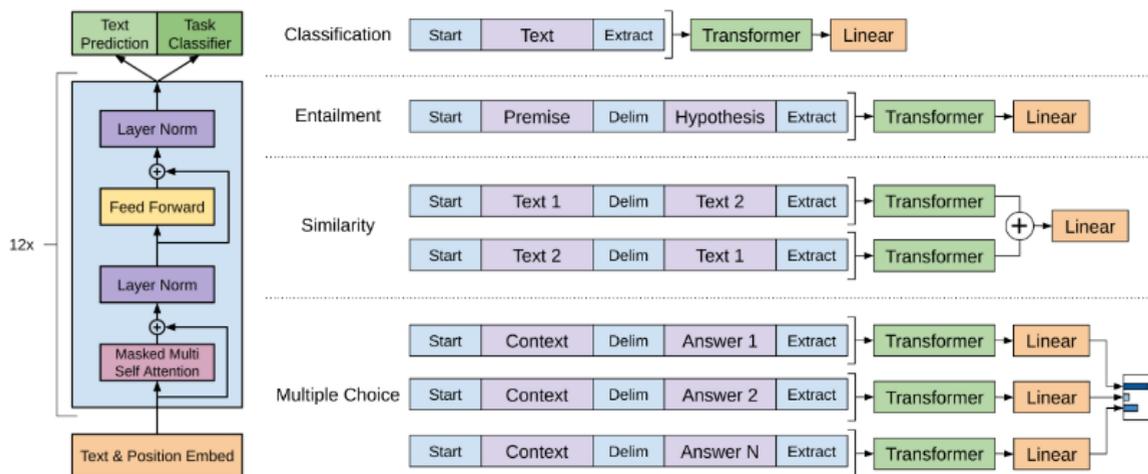
Generative Pre-Trained Transformer

- *Modèle de langue*
- De type Transformer, 12 couches, 117M paramètres
- Appris sur BookCorpus (7K livres)
- Pré-apprentissage génératif + affinage discriminatif

GPT (Affinage)

- GPT peut être affiné sur des tâches d'analyse (classification, détection d'implications, question réponse)
- L'affinage modifie les paramètres du modèle

GPT (Affinage)



Pendant l'affinage, le modèle a deux têtes :

- la tête standard *LM head* pour la prédiction du prochain mot
- une *tête spécifique à la tâche*

Impact

Améliore l'état de l'art pour 9 tâches d'analyse (sur 12)

Détection d'implications

Method	MNLI-m	MNLI-mm	SNLI	SciTail	QNLI	RTE
ESIM + ELMo [44] (5x)	-	-	<u>89.3</u>	-	-	-
CAFE [58] (5x)	80.2	79.0	<u>89.3</u>	-	-	-
Stochastic Answer Network [35] (3x)	<u>80.6</u>	<u>80.1</u>	-	-	-	-
CAFE [58]	78.7	77.9	88.5	<u>83.3</u>		
GenSen [64]	71.4	71.3	-	-	<u>82.3</u>	59.2
Multi-task BiLSTM + Attn [64]	72.2	72.1	-	-	82.1	61.7
Finetuned Transformer LM (ours)	82.1	81.4	89.9	88.3	88.1	56.0

Meilleure génération de texte

GPT-2 (1.5B)

Génération d'histoires

Context (human-written): In a shocking finding, scientist discovered a herd of unicorns living in a remote, previously unexplored valley, in the Andes Mountains. Even more surprising to the researchers was the fact that the unicorns spoke perfect English.

GPT-2: The scientist named the population, after their distinctive horn, Ovid's Unicorn. These four-horned, silver-white unicorns were previously unknown to science.

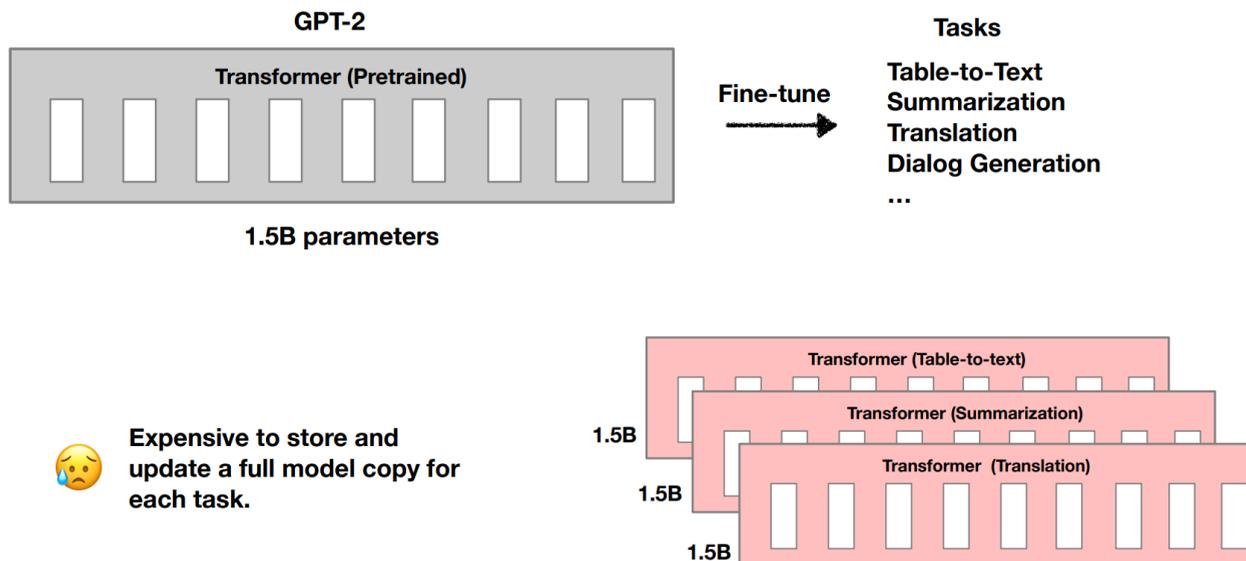
Now, after almost two centuries, the mystery of what sparked this odd phenomenon is finally solved.

Dr. Jorge Pérez, an evolutionary biologist from the University of La Paz, and several companions, were exploring the Andes Mountains when they found a small valley, with no other animals or humans. Pérez noticed that the valley had what appeared to be a natural fountain, surrounded by two peaks of rock and silver snow.

Affinage complet vs. Affinage léger

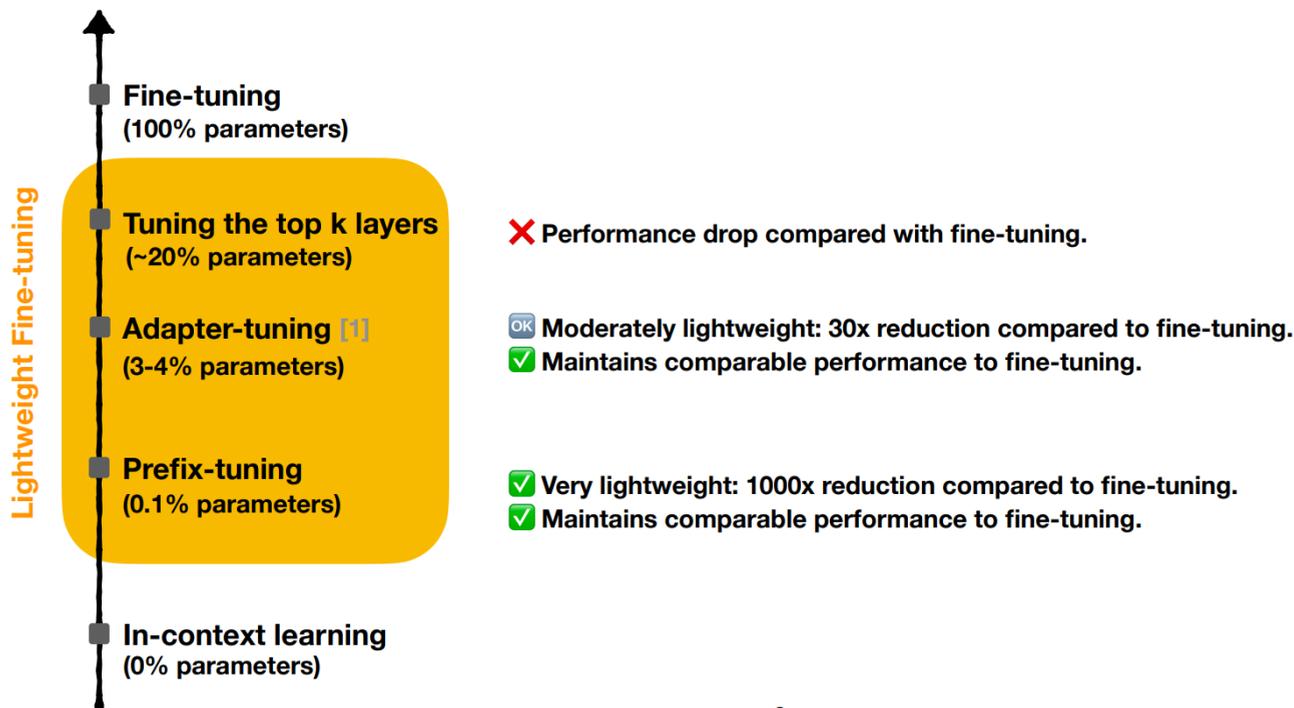
L'affinage complet (Full Fine-Tuning) est couteux en temps/espace car il

- met à jour *tous les paramètres* du modèle
- opère sur des modèles pré-appris de très grande taille
- Ex : GPT-2, 774M paramètres. GPT-3, 175B paramètres
- exige de stocker une copie par affinage



Src: Li and Liang 2021

Affinage léger



Ne modifie pas les paramètres du modèle pré-entraîné, seulement un petit ensemble de paramètres additionnels

- **Adapter** : ajout de paramètres dans le modèle (couches additionnelles)
- **Prefix tuning** : affine un vecteur de paramètres
- **Lora (Low Rank Matrices)** affine et intègre de petites matrices
- **In context learning or prompting** : pas d'affinage

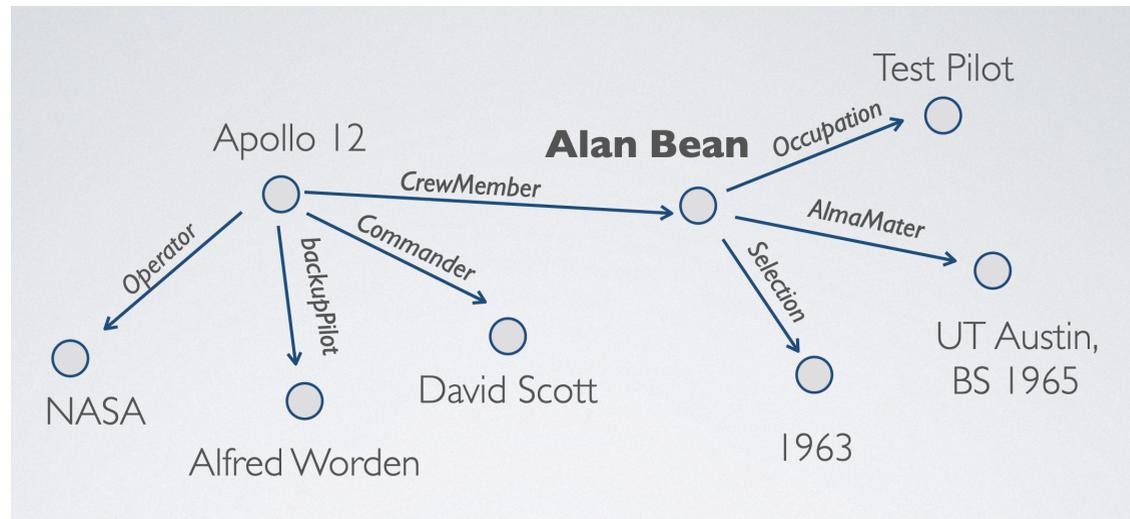
Graphe RDF --> Langue peu dotée

Soto-Martinez et al. ACL-IJCNLP 2023

La tâche partagée WebNLG



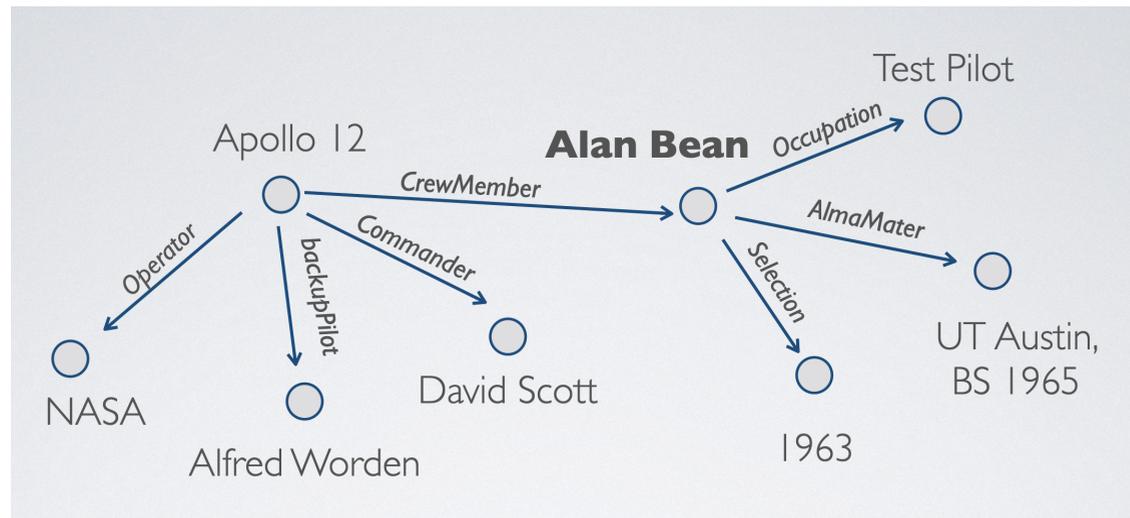
RDF Graph



La tâche partagée WebNLG



Graphe RDF



Texte anglais

Alan Bean graduated from UT Austin in 1955 with a Bachelor of Science degree. He was hired by NASA in 1963 and served as a test pilot. Apollo 12's backup pilot was Alfred Worden and was commanded by David Scott

Graphe RDF --> Breton

- NLG+MT pipeline



Graphe RDF --> Breton

- NLG+MT pipeline



- Full-fine tuning (BLEU: 0.10)



Graphe RDF --> Breton

- NLG+MT pipeline



- Full-fine tuning (BLEU: 0.10)

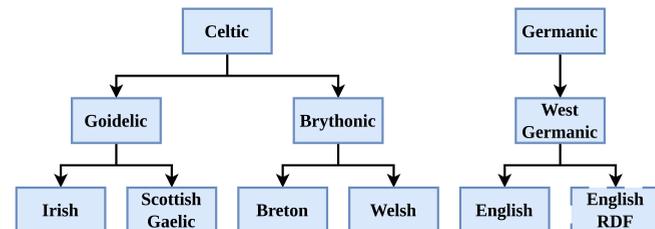


- Affinage de préfixe



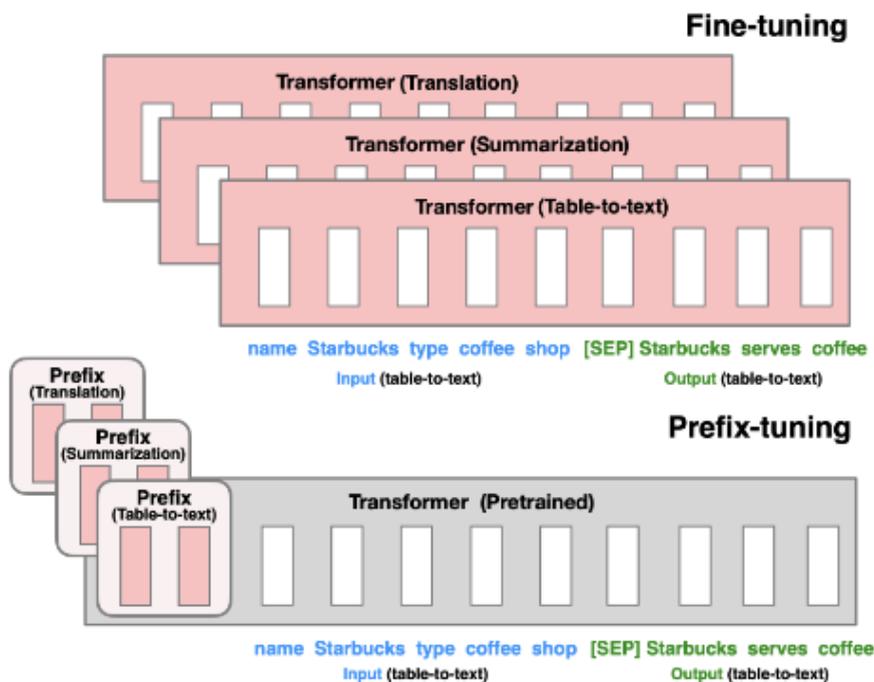
Graphe RDF --> Breton

- NLG+MT pipeline
- Full-fine tuning (BLEU: 0.10)
- Affinage de préfixe
- Préfixe structuré pour capturer les relations phylogénétiques



Affinage de préfixe continu

- Gèle tous les paramètres
- Apprend un préfixe de paramètres

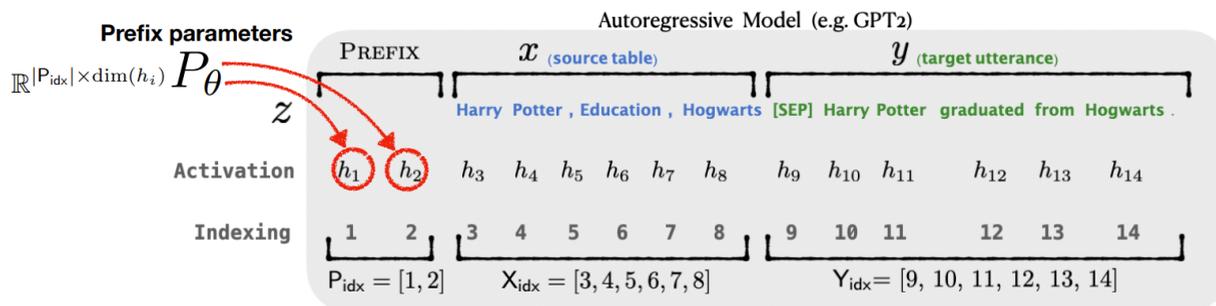


Affinage de préfixe continu

$$h_i = \begin{cases} P_\theta[i, :], & \text{if } i \in P_{\text{idx}}, \\ \text{LM}_\phi(z_i, h_{<i}), & \text{otherwise.} \end{cases}$$

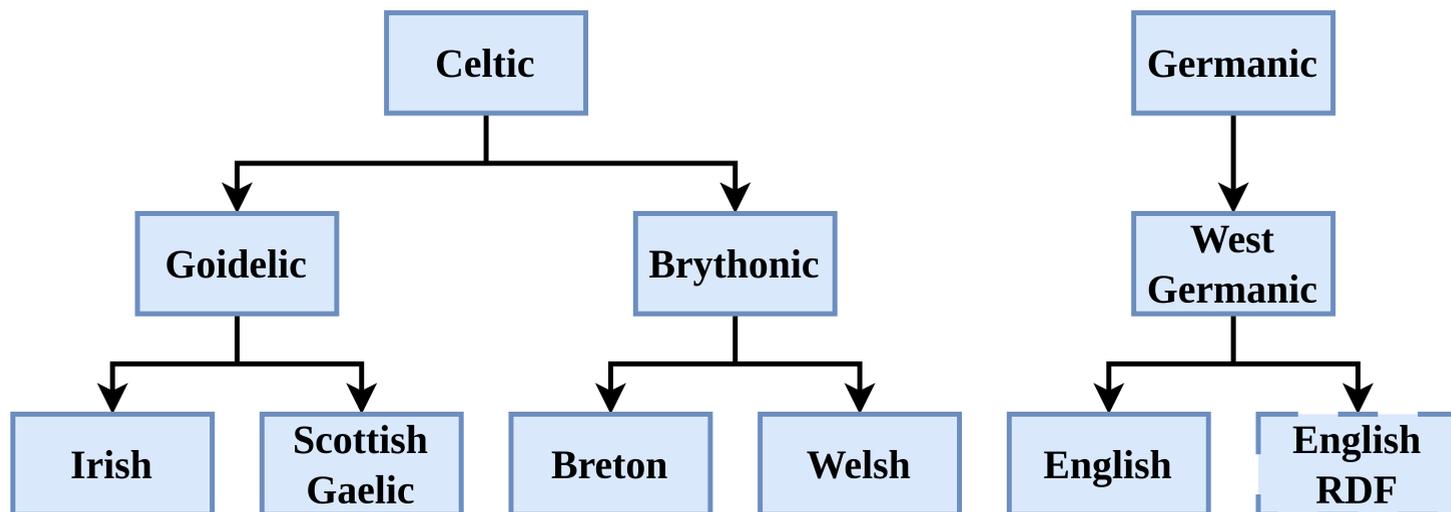
$$\max_{\theta} \log p_{\phi, \theta}(y | x) = \sum_{i \in Y_{\text{idx}}} \log p_{\phi, \theta}(z_i | h_{<i})$$

freeze LM parameters ϕ
update prefix parameters θ



Les paramètres du modèle initial sont gelés. Seuls les paramètres du préfixe sont appris.

Arbre phylogénétique

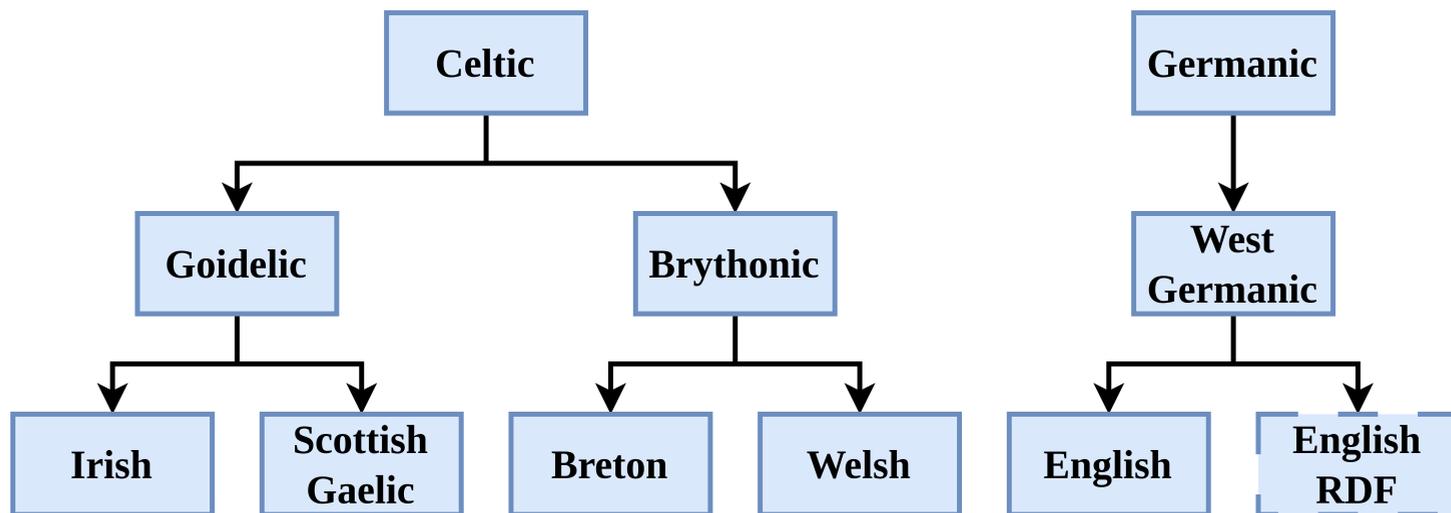


Préfixe

50 Tokens Task	15 Tokens Source Family	15 Tokens Source Genus	15 Tokens Source Language	15 Tokens Target Family	15 Tokens Target Genus	15 Tokens Target Language	n Tokens Input Sequence
-------------------	-------------------------------	------------------------------	---------------------------------	-------------------------------	------------------------------	---------------------------------	-------------------------------

Les parties "Language" du préfixe sont spécifiques à chaque langue. Les parties "Family" et "Genus" sont partagés par plusieurs langues et sont mis à jour plus fréquemment.

Arbre phylogénétique



Préfixe

50 Tokens Task	15 Tokens Source Family	15 Tokens Source Genus	15 Tokens Source Language	15 Tokens Target Family	15 Tokens Target Genus	15 Tokens Target Language	n Tokens Input Sequence
-------------------	-------------------------------	------------------------------	---------------------------------	-------------------------------	------------------------------	---------------------------------	-------------------------------

E.g., Breton et Welsh (Gallois) mettent à jour les sous-prompts appris pour "Brythonic" et "Celtic".

Apprentissage

Etape 1: Apprentissage auto-supervisé (Modèle de langue)

Monolingue

	Task	Source			Target			Original Input Sequences					
		Family	Genus	Lang.	Family	Genus	Lang.						
Input Batch	Masked LM	Germanic	West Germanic	RDF	Germanic	West Germanic	RDF	<S>	Einstein	<P>	<mask>	<P>	Poland
	Prefix LM	Germanic	West Germanic	English	Germanic	West Germanic	English	Thank	you	for	<mask>	<pad>	<pad>
	Suffix LM	Celtic	Britonic	Welsh	Celtic	Britonic	Welsh	<mask>	honno	?	<pad>	<pad>	<pad>
	Deshuffling	Celc	Britonic	Breton	Celtic	Britonic	Breton	skuizh	?	out	Ha	<pad>	<pad>
	Generate	Celc	Goidelic	Irish	Celtic	Goidelic	Irish	Seo	<mask>	<pad>	<pad>	<pad>	<pad>

Apprentissage

Etape 1: Apprentissage auto-supervisé (Modèle de langue)

Monolingue

	Task	Source			Target			Original Input Sequences					
		Family	Genus	Lang.	Family	Genus	Lang.						
Input Batch	Masked LM	Germanic	West Germanic	RDF	Germanic	West Germanic	RDF	<S>	Einstein	<P>	<mask>	<P>	Poland
	Prefix LM	Germanic	West Germanic	English	Germanic	West Germanic	English	Thank	you	for	<mask>	<pad>	<pad>
	Suffix LM	Celtic	Britonic	Welsh	Celtic	Britonic	Welsh	<mask>	honno	?	<pad>	<pad>	<pad>
	Deshuffling	Celc	Britonic	Breton	Celtic	Britonic	Breton	skuizh	?	out	Ha	<pad>	<pad>
	Generate	Celc	Goidelic	Irish	Celtic	Goidelic	Irish	Seo	<mask>	<pad>	<pad>	<pad>	<pad>

Etape 2: Affinage sur les données WebNLG (Verbalisation de graphe)

Graphe --> Texte

Apprentissage

Etape 1: Apprentissage auto-supervisé (Modèle de langue)

Monolingue

	Task	Source			Target			Original Input Sequences					
		Family	Genus	Lang.	Family	Genus	Lang.						
Input Batch	Masked LM	Germanic	West Germanic	RDF	Germanic	West Germanic	RDF	<S>	Einstein	<P>	<mask>	<P>	Poland
	Prefix LM	Germanic	West Germanic	English	Germanic	West Germanic	English	Thank	you	for	<mask>	<pad>	<pad>
	Suffix LM	Celtic	Britonic	Welsh	Celtic	Britonic	Welsh	<mask>	honno	?	<pad>	<pad>	<pad>
	Deshuffling	Celc	Britonic	Breton	Celtic	Britonic	Breton	skuizh	?	out	Ha	<pad>	<pad>
	Generate	Celc	Goidelic	Irish	Celtic	Goidelic	Irish	Seo	<mask>	<pad>	<pad>	<pad>	<pad>

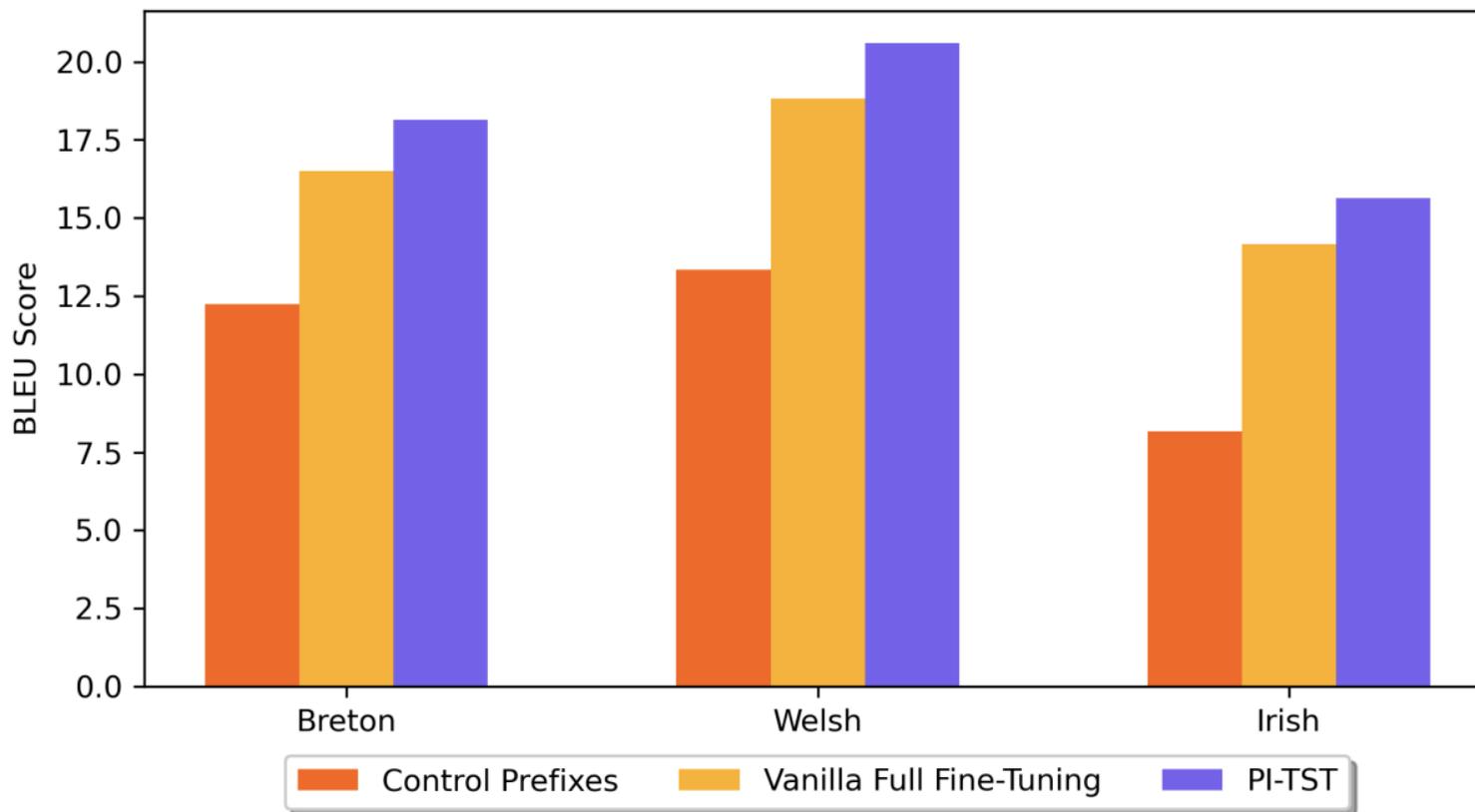
Etape 2: Affinage sur les données WebNLG (Verbalisation de graphe)

Graphe --> Texte

Inférence

Le préfixe est celui de la langue cible.

Résultats



L'affinage de préfixe améliore les résultats : BLEU (NLG+MT) : 10 --> BLEU (Préfixe) : 18.15

Graphe AMR --> Langues bien et peu dotées

Soto-Martinez et al. 2024, In Submission

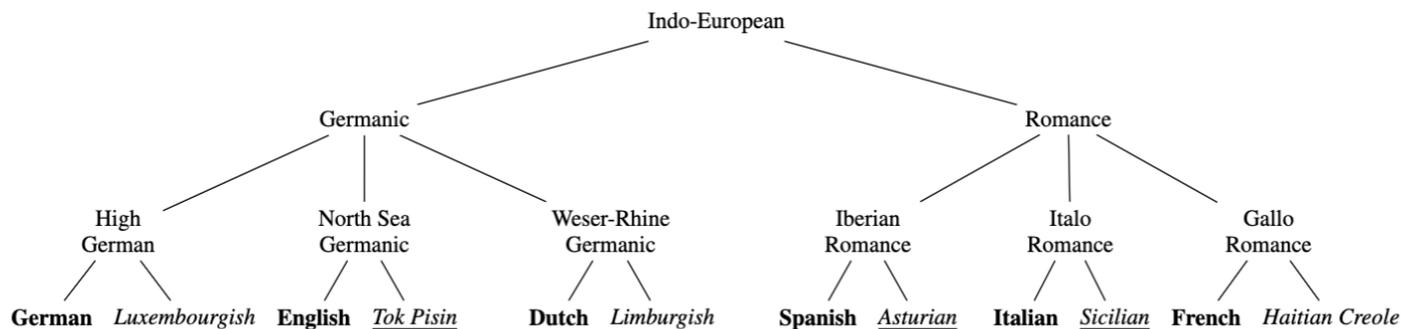
Affinage hiérarchique phylogénétique

Affinage hiérarchique

- Affinage itératif d'un modèle multilingue (12 langues) en 12 modèles multilingues

Connaissances phylogénétiques

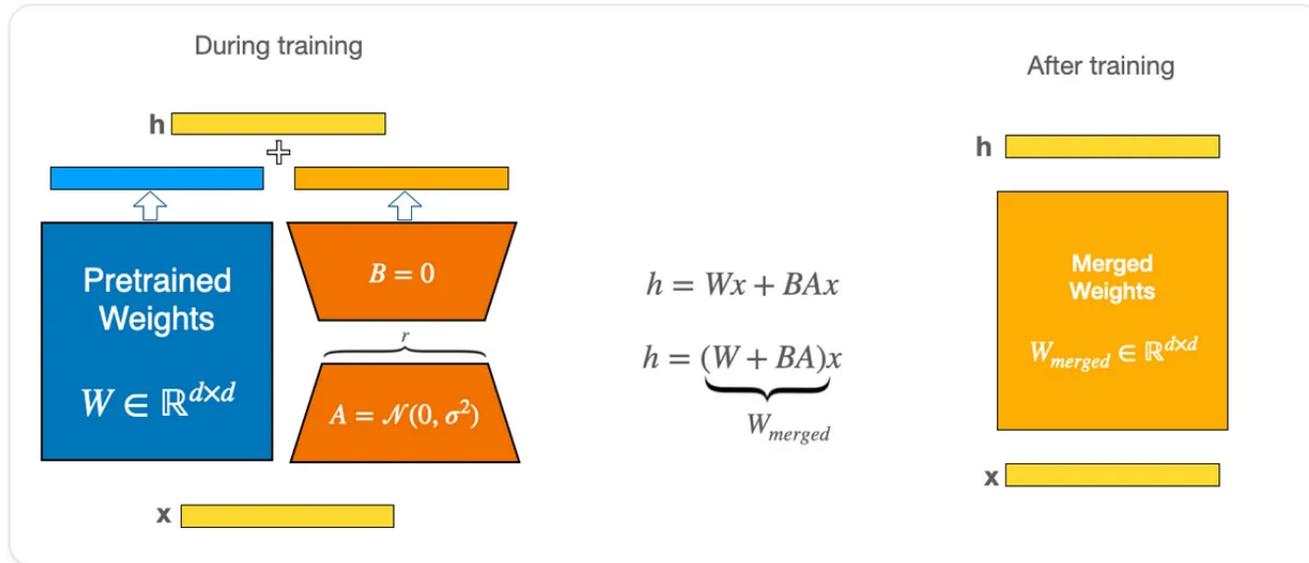
- A chaque itération, les langues considérées pour l'apprentissage sont déterminées par des connaissances phylogénétiques



(b) Phylogenetic Tree Hierarchy (PTL)

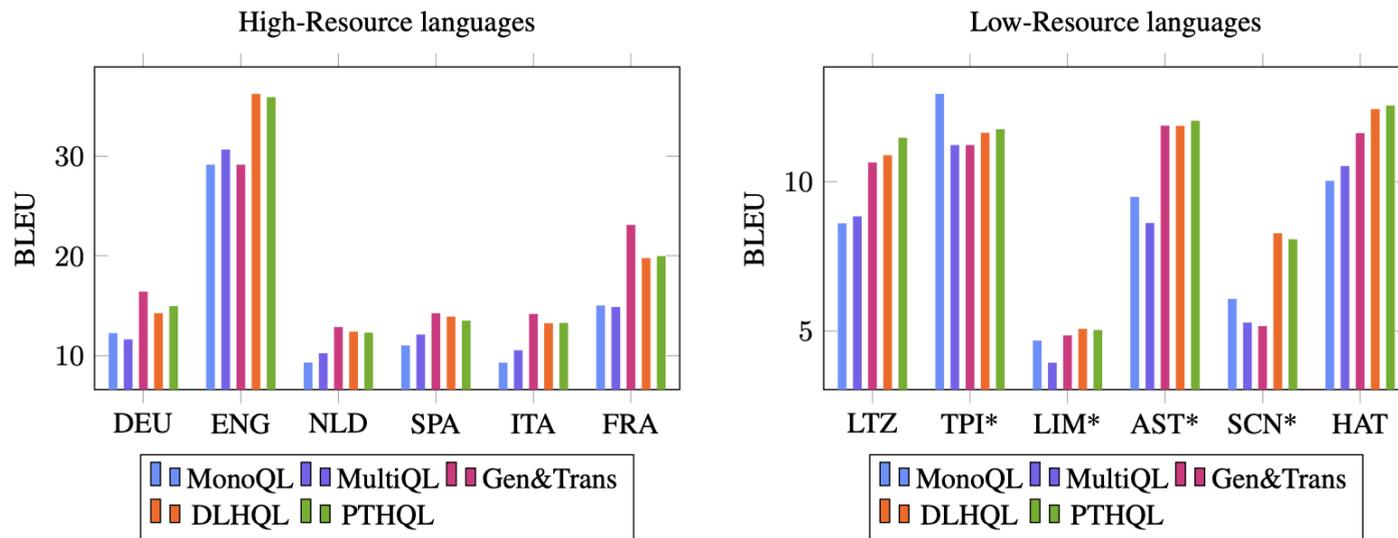
Affinage Hiérarchique avec LoRA

Affinage LoRA (Low Rank Matrices)



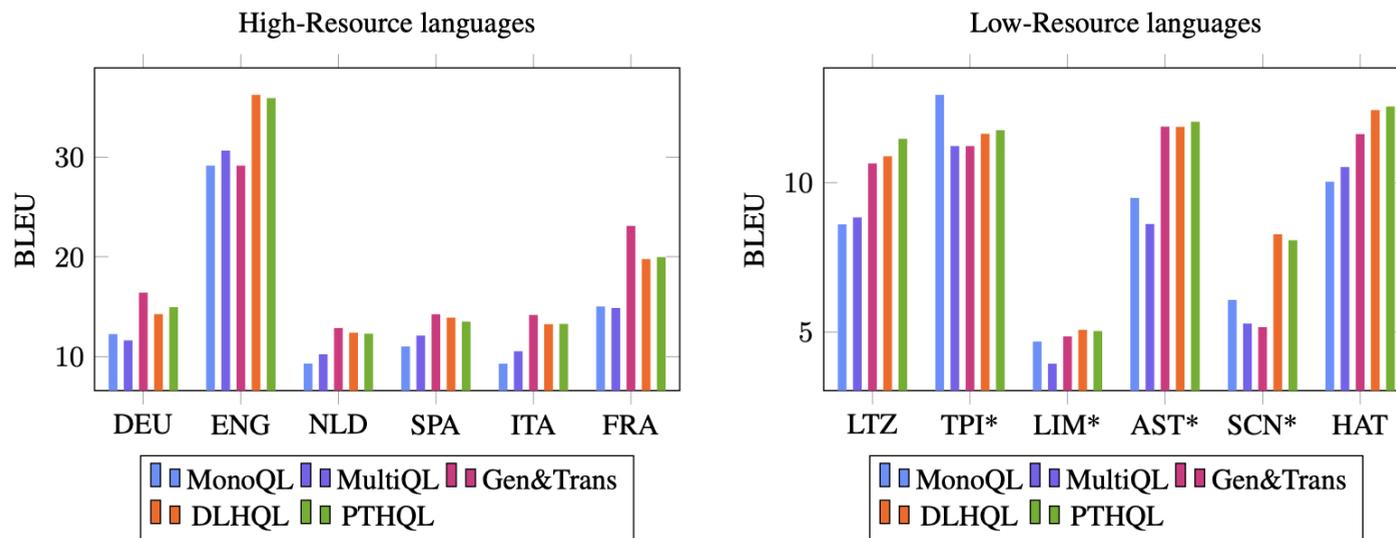
Un nouveau modèle est créé. Pas d'overhead pendant l'inférence

Comparaison avec les lignes de base



HQL surpasse ou est comparable aux approches multi- et monolingues.

Comparaison avec les lignes de base



HQL surpasse l'approche Gen&Trans pour toutes les langues peu dotées

Comparaison avec les lignes de base

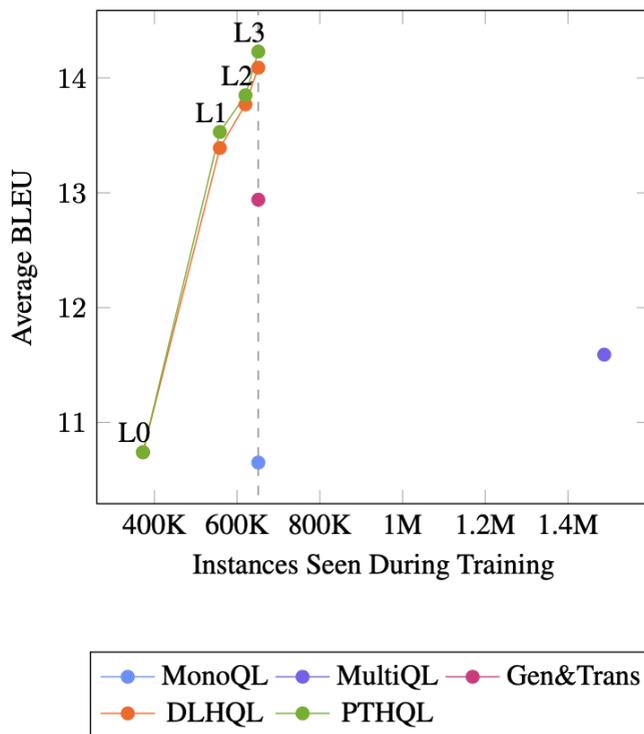


Figure 4: Average BLEU score of all languages against total number of seen training instances. HQL models include results on the intermediary levels of the hierarchy.

HQL s'optimise plus vite que les 3 baselines models et en moyenne les surpasse.

Comparaison avec les approches existantes

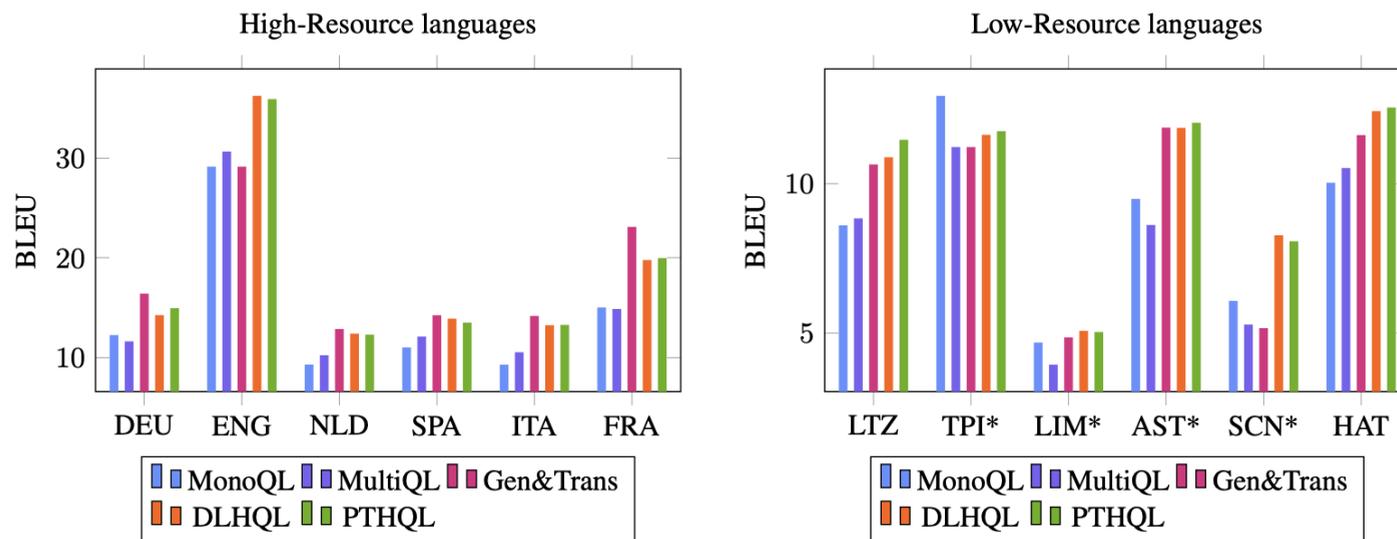
Model	DEU	ENG	SPA	ITA
F&G	15.3	24.9	21.7	19.8
Ribeiro	20.6	—	30.7	26.4
Xu	25.7	—	31.4	28.4
Martinez	23.2	44.8	34.6	29.0
MonoQL	18.2	49.2	38.6	22.7
MultiQL	19.8	42.9	34.1	27.2
Gen&Trans*	28.0	49.2	39.6	33.8
DLHQL	21.2	44.2	37.4	29.2
PTHQL	22.8	43.4	37.2	29.7

Table 4: BLEU score on AMR3.0 test data.

English Gen&Trans is simply the result of MonoQL.

Sur les langues bien dotées, HQL a des résultats similaires aux approches existantes.

Comparaison avec les approches existantes



HQL a des résultats similaires aux approches existantes bien qu'entraîné sur moins de données.

Intégrer des connaissances

Interfacer Génération et Extraction d'Information

Retrieval Augmented Generation (RAG)

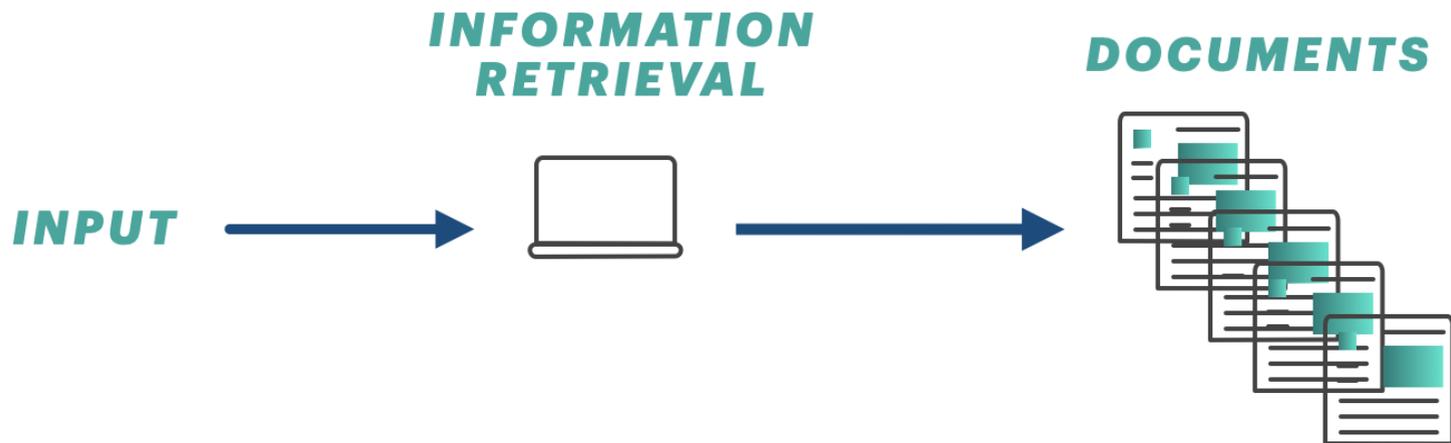
Recherche d'information

- Sélectionne des données externes

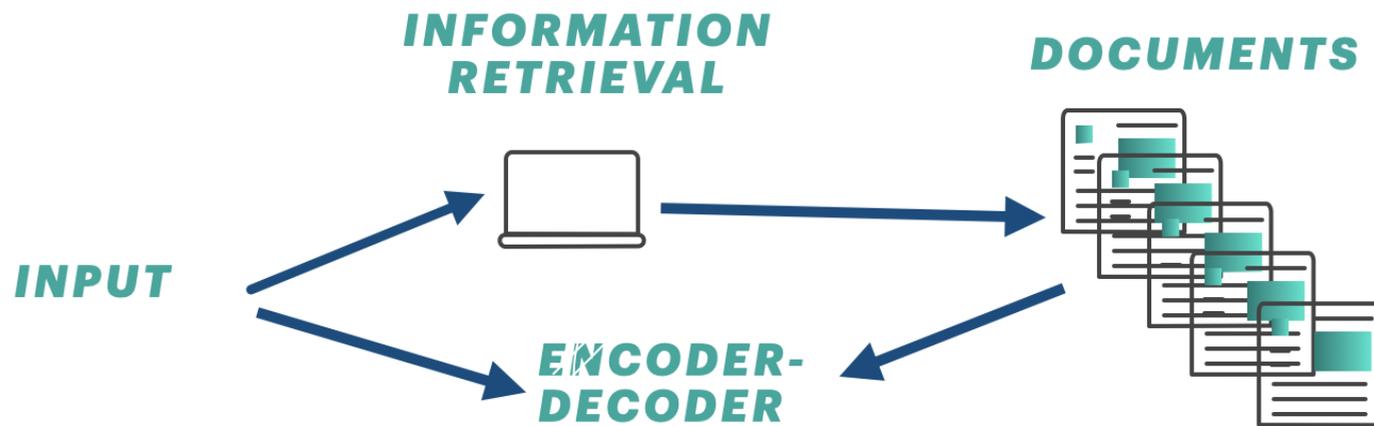
Génération

- Génère à partir de ces données et des données d'entrée

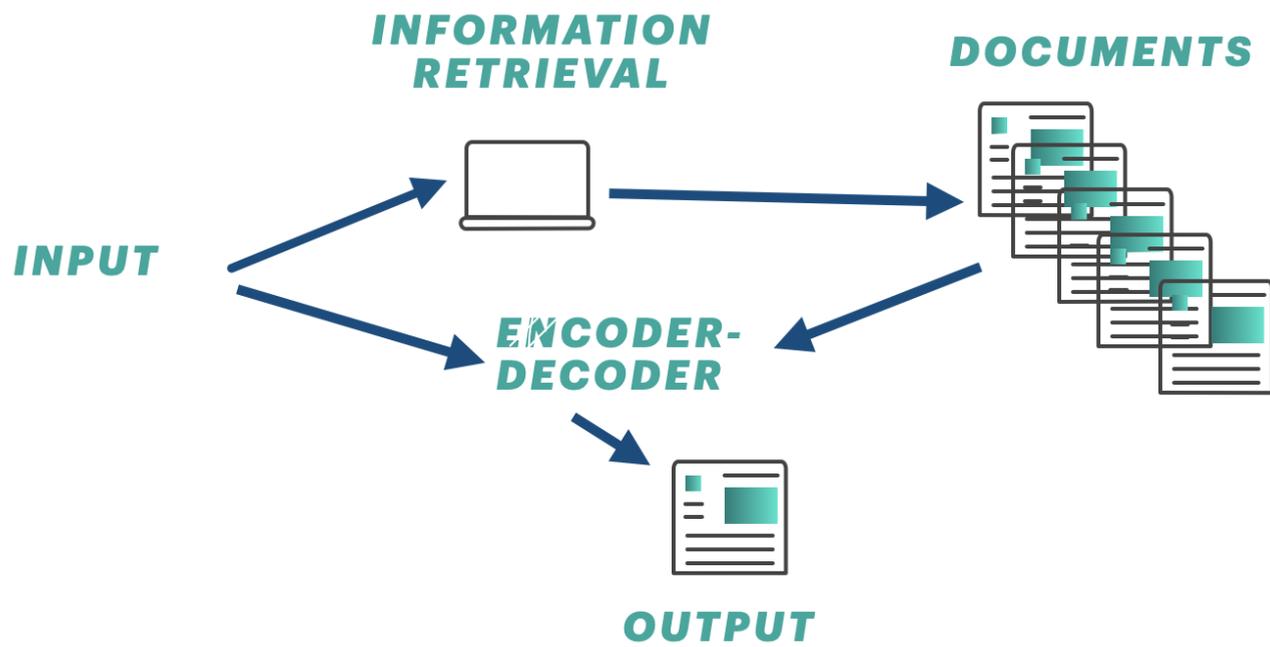
Sélectionner des données externes



Encoder les données



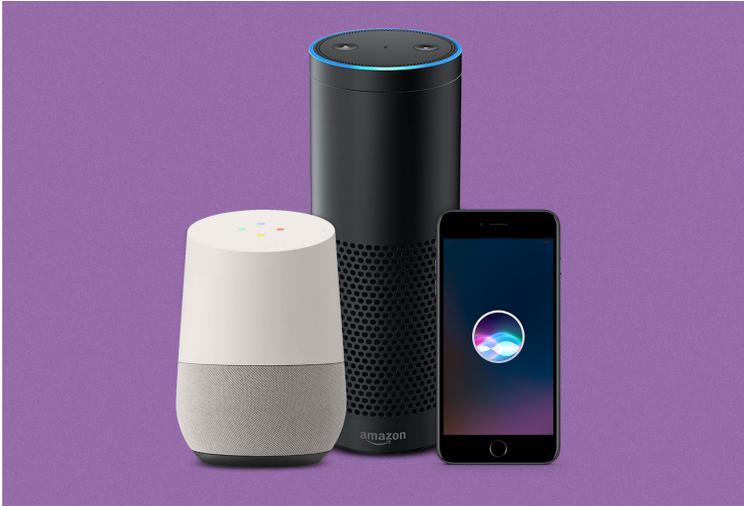
Générer



Intégrer des connaissances dans les modèles de dialogue

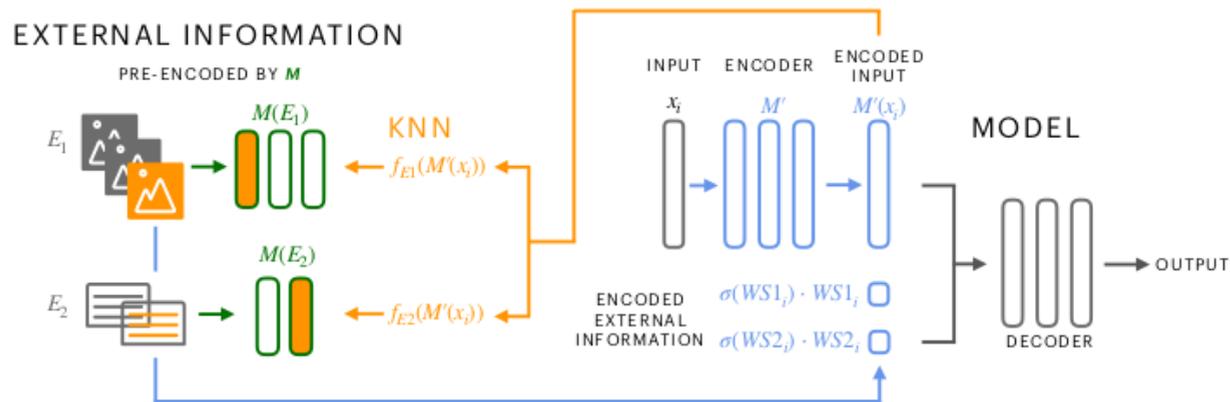
Fan et al. TACL 2021

Modèles de dialogue standard



Génère ou sélectionne une réponse à partir du contexte (les tours de dialogue précédents)

Modèles de dialogue avec des connaissances



- Interface génération et recherche d'information
- Génère des tours de dialogue à partir du contexte et de connaissances externes
- Recherche des K plus proches voisins dans des corpus de dialogue

Deux sources de connaissances externes

Human: What's your favorite Disney movie ?

Model: I love the incredibles, they are my favorite Disney movie.

Fetches Knowledge: Disney announced intentions to develop additional superhero films after the success by the incredibles

Fetches Template: I love kiteboarding, it is one of my favorite activities on the water.

- Connaissances sur le monde (contenu)
- Tours de dialogues (forme)

Données

Wizard of Wikipedia (Dinan et al. 2018)

- Dialogues centrés sur différents topics
- Connaissances : Phrases pertinentes extraites de Wikipedia
- E.g., parle de Star Wars quand le topic est la science fiction
- Données : 34 phrases, 170K tours de dialogue

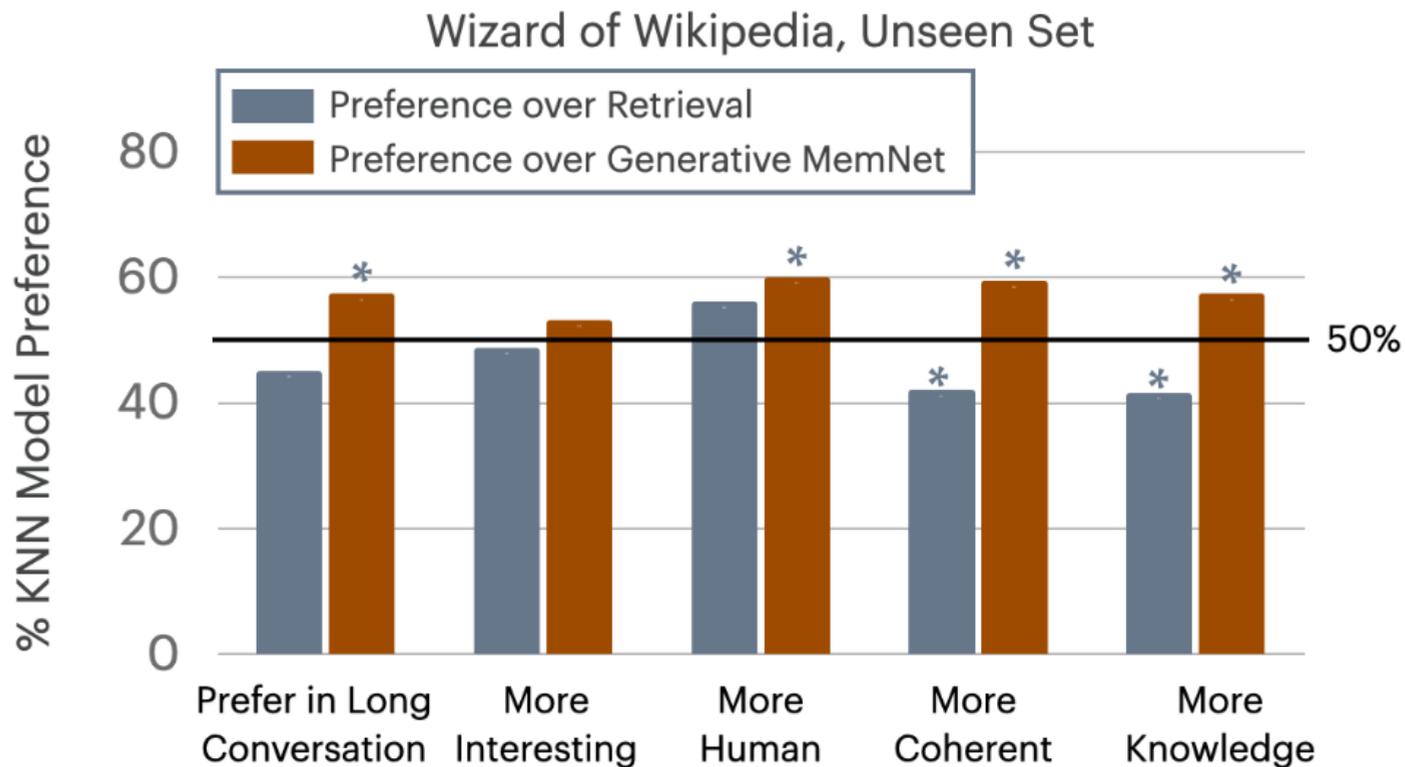
Image Chat (Shuster et al, 2020)

- Dialogues centrés sur une image
- Connaissances : images + dialogue
- Extraction d'information: phrases qui sont typiquement produites pour des images similaires (e.g., image avec chien, phrase sur la race).
- Données : 189K images, 350K tours de dialogue

Exemples

Knowledge	Training Utterance	Generation
buzz lightyear 's name is in honor of astronaut edwin 'buzz' aldrin	my favorite character in that book series is hermione granger	cool! my favorite character in that movie is buzz lightyear
mr potato head is based on the real-life mr. potato head toy	my favorite character in that book series is hermione granger	my favorite character in that movie is real-life mr potato head
slinky dog is a toy dachschund with a metal slinky for a body	my favorite character in that book series is hermione granger	cool! my favorite character is the slinky dog
slinky dog is a toy dachschund with a metal slinky for a body	i really like the character hermione granger	cool! i really like slinky dog
slinky dog is a toy dachschund with a metal slinky for a body	my favorite character of all time has to be hermione granger	i love that movie, my favorite character has to be slinky dog the dachshund
slinky dog is a toy dachschund with a metal slinky for a body	i agree with you! that's my favorite character as well	i think so too! my favorite is slinky

Evaluation par l'humain



Générer des biographies Wikipedia

Fan et Gardent, ACL 2022

Objectif

PERSON NAME



**Retrieval Augmented Generation (RAG)
Generation + Information Retrieval**

WIKIPEDIA

Joan Paton

Joan Burton Paton AM née Cleland (1916–April 2000) was an Australian teacher, naturalist, environmentalist and ornithologist. One of the first women to become a member of the exclusive Adelaide Ornithologists Club, of which she was elected President 1991–1993, she also served as president of the South Australian Ornithological Association (1979-1982). Her father was Professor Sir John Burton Cleland, a notable microbiologist and pathologist who strongly encouraged her early interest in natural history.

Contents

- [Early life and education](#)
- [Career](#)
- [Legacy and honours](#)
- [References](#)
- [External References](#)

Early life and education

Joan Burton Paton was born in Sydney, New South Wales, the daughter of [John Burton Cleland](#) (1878–1971) and his wife, [Dora Isabel Paton](#) (1880–1955).^[1] She had three sisters, [Dr Margaret Burton Cleland](#), [Elizabeth Robson Cleland](#) and [Barbara Burton Cleland](#); and a brother, [William Paton 'Bill' Cleland](#), who became a surgeon. The father encouraged his children's interest in science. Joan Paton was educated at the [University of Adelaide](#), where she majored in [organic chemistry](#) and [biochemistry](#). In 1951 she married [Erskine Norman Paton](#) (1922–1985), son of [Adolph Ernest Paton](#) and [Ida Marie Poynton](#). Their son is [Prof David Cleland Paton](#).^[2]

Career

In 1967 Paton became a lecturer on ornithology in South Australia's [Workers' Educational Association](#).^{[3][4]} Among those she inspired to work in ornithology and environmental conservation was [Margaret Cameron](#), who became the President of the [Royal Australasian Ornithologists Union](#) (RAOU).^[5]

Paton was active in the RAOU, as well as in the [South Australian Ornithological Association](#) (SAOA), of which she was elected Vice-President 1974–1979, and President 1979–1982. She was one of the first women to become a member of the exclusive [Adelaide Ornithologists Club](#), of which she was elected president (1991–1993).^[6]

Legacy and honours

- 1990, she was made an Honorary Member of the SAOA.
- 1996, she was made an Honorary Member of the Adelaide Ornithologists Club.

Défis

- Collecter des informations pertinentes
- Générer un texte structuré (sections, titres etc.)
- Factualité

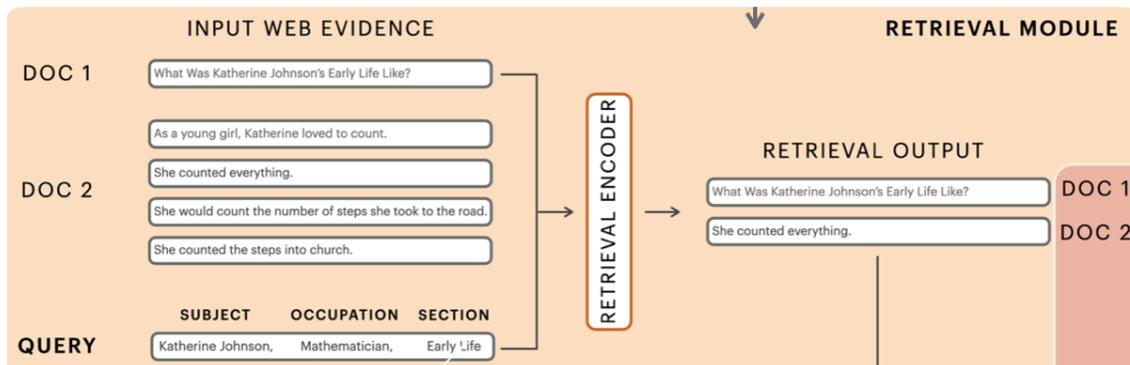
Extraction



QUERY

Katherine Johnson
Mathematician
Early Life

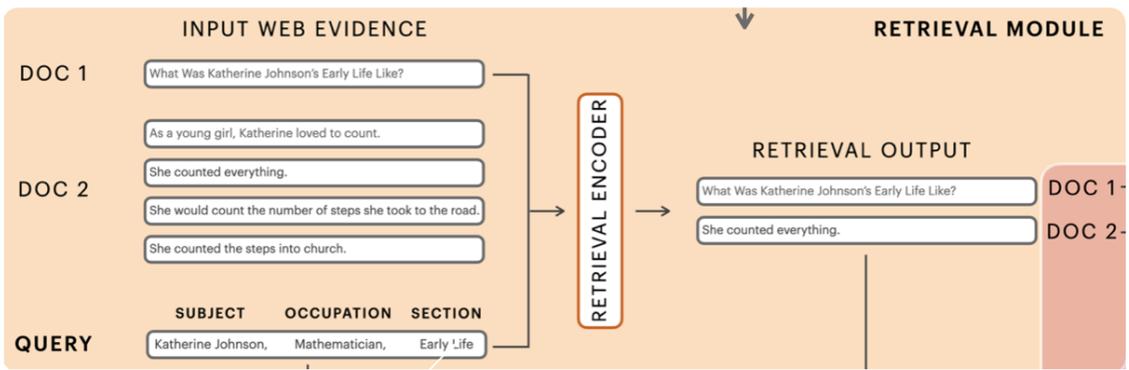
Requête web



Extraction



QUERY
Katherine Johnson
Mathematician
Early Life



RESULT
20 first documents
segmented in
sentences

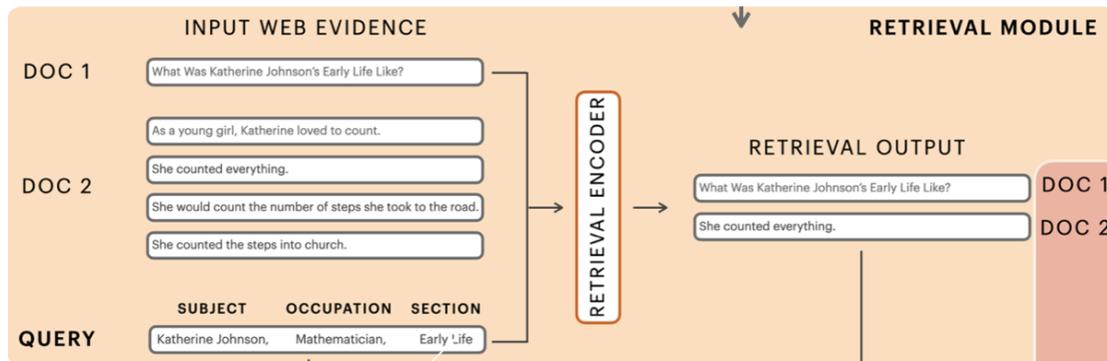
Requête web --> 20 documents

Extraction



QUERY

Katherine Johnson
Mathematician
Early Life



Web Search

20 first documents
segmented in
sentences

Semantic Filtering

40 most similar
sentences (1,000
words)

Filtrage sémantique --> 40 phrases

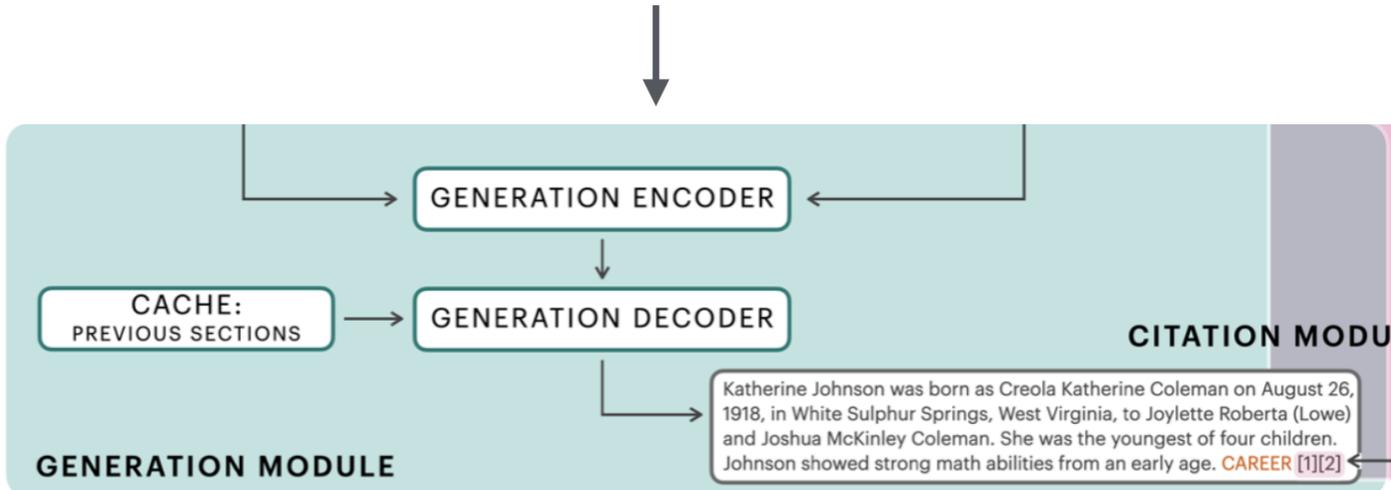
Generation

QUERY

Katherine Johnson
Mathematician
Early Life

EXTRACTION

1,000 words



Résultats

Model	ROUGE-L	Entailment	Named Entity Coverage
BART Pretraining + Finetuning	17.4	15.8	21.9
+ Retrieval Module	18.8	17.2	23.1
+ Caching Mechanism	19.3	17.9	23.4

IR and the cache mechanism allow for statistically significant improvement

Données biaisées

WikiSum Evaluation Dataset	
Average Number of Sections	7.2
Average Length of a Section	151.0
Average Length of Total Article	892.3
Avg overlap of Web Hits and Biography	39.8%
Our Evaluation Dataset	
Average Number of Sections	5.8
Average Length of a Section	132.3
Average Length of Total Article	765.9
Avg Number of Web Hits (max 20)	18.1
Avg overlap of Web Hits and Biography	24.9%

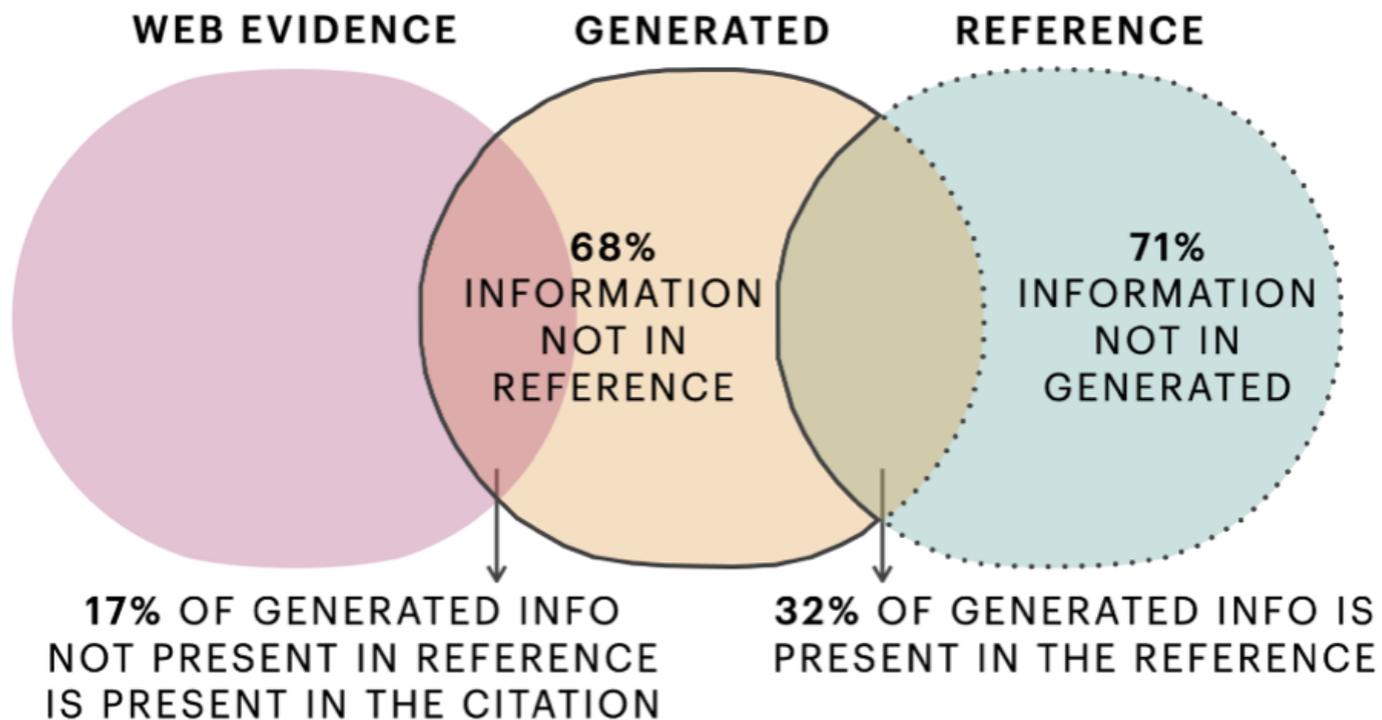
Plus d'information sur le web pour les hommes (Wikisum) que pour les femmes (Our dataset)

Impact sur les résultats

Model	WikiSum Test	Women	Scientists	Women in Asia	Women in Africa
BART Pretraining	19.0	17.4	18.2	16.7	16.4
+ Retrieval	21.4	18.8	19.3	17.9	17.1
+ Caching	21.8	19.3	19.7	18.4	17.3

La qualité des textes générés (ROUGE-L) est moindre pour les biographies de femmes.

Evaluation par l'humain



Environ 13% du texte généré n'est pas validé (ni dans le texte de référence ni dans les données web extraites).

Questions ?

