IA Générative, LLMs et Traitement Automatique des Données Médicales

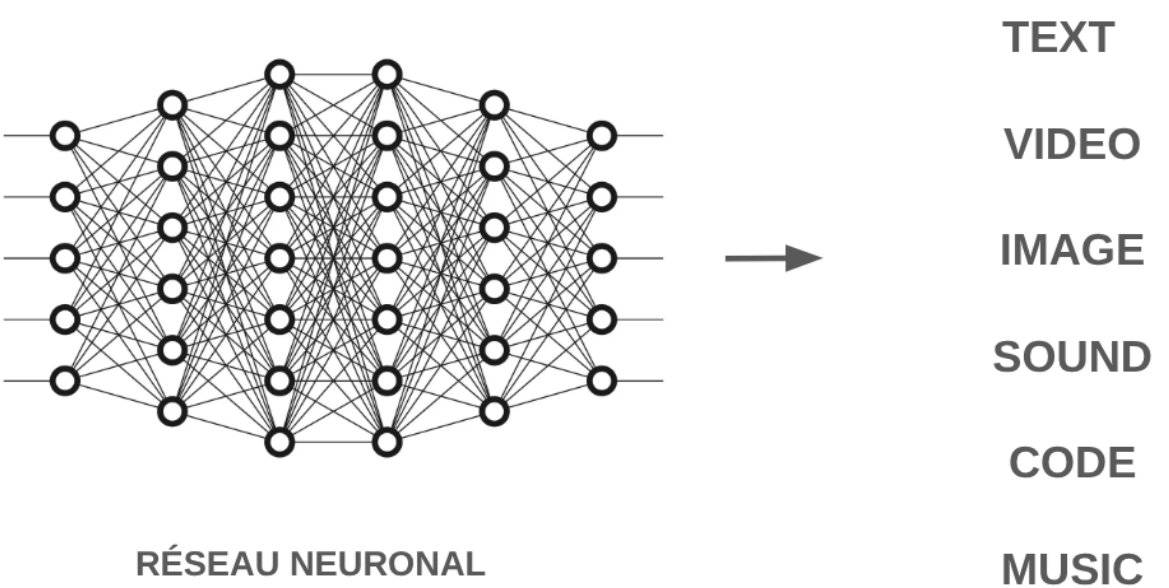
Claire Gardent

CNRS / LORIA, Nancy

IA Générative, Réseaux Neuronaux et LLMs

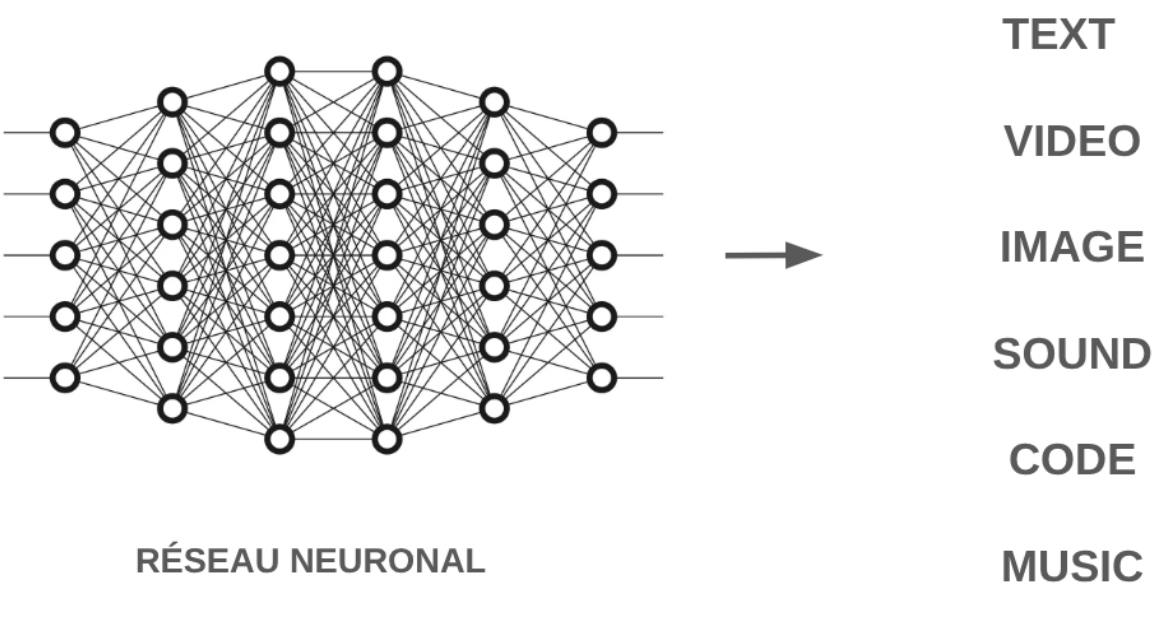
Qu'est ce que l'IA generative ?

Une branche de l'IA qui ***génère*** du nouvelles données à l'aide de réseaux neuronaux

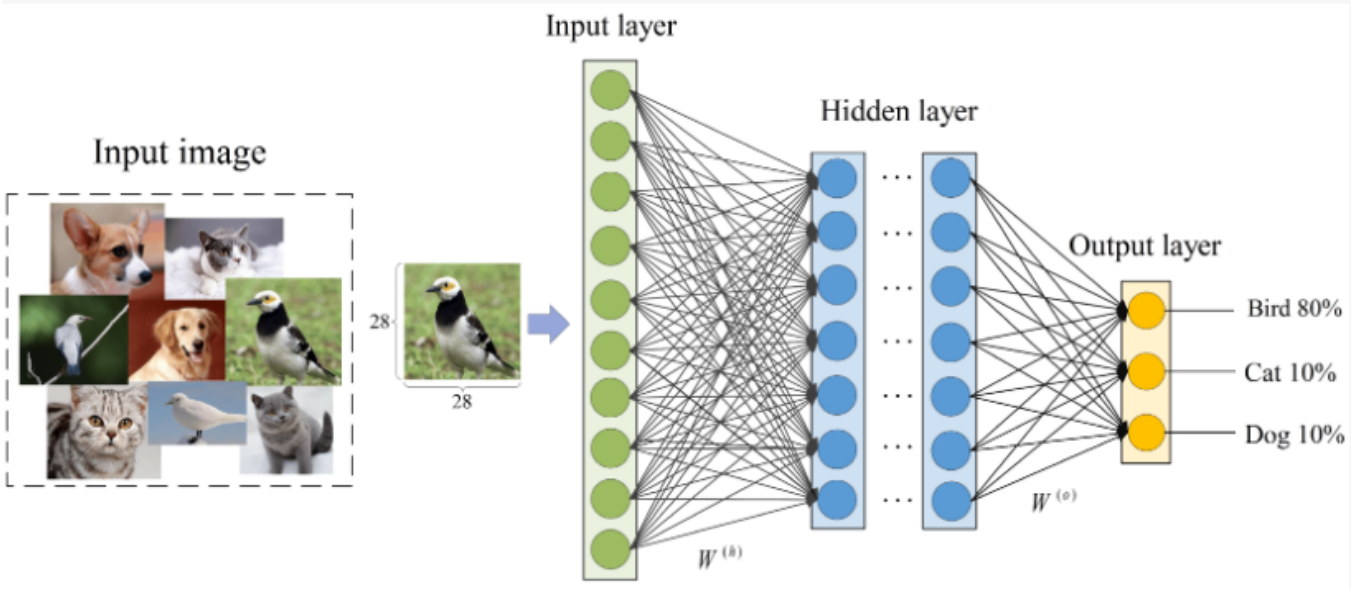


Des données réalistes

L'IA générative désigne un ensemble de modèles capables d'***apprendre la distribution d'une donnée*** afin de générer de nouveaux échantillons plausibles.

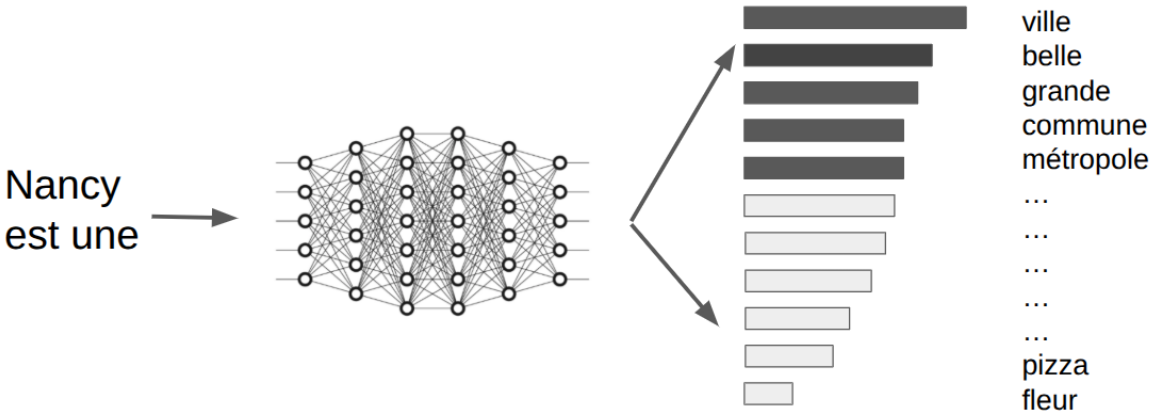


Comment un réseau neuronal produit il une distribution de probabilité ?



Source: Deng et al. 2022

Comment générer du texte ?



Un ***modèle de langue neuronal*** génére une phrase mot par mot

- Il produit une **distribution de probabilité sur l'ensemble du vocabulaire**
- et sélectionne **un mot ayant une probabilité élevée**.
- Il prédit le mot suivant à partir des mots précédents (**modèle auto-régressif**).

Des NLMs aux LLMs

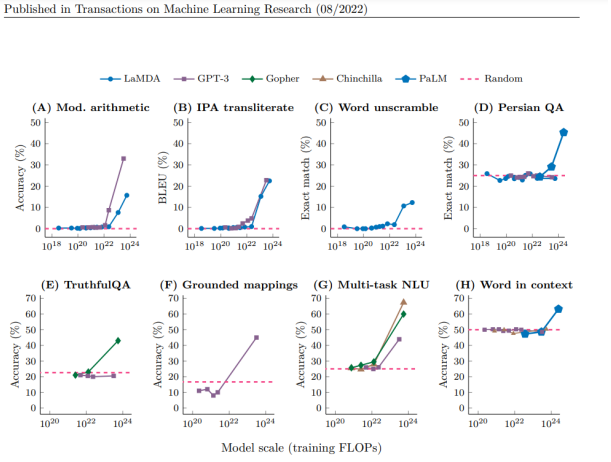
Montée en échelle

Sutzkever (2014)

- 340 millions de paramètres
- Traduction automatique par des méthodes neuronales état de l'art

GPT3 (2020) -- Apparition des LLMs

- **175 milliards** de paramétres
- Augmentation nette des performances
- Propriétés émergentes: n'émerge qu'à partir d'une certaine taille (e.g., calcul arithmétique)

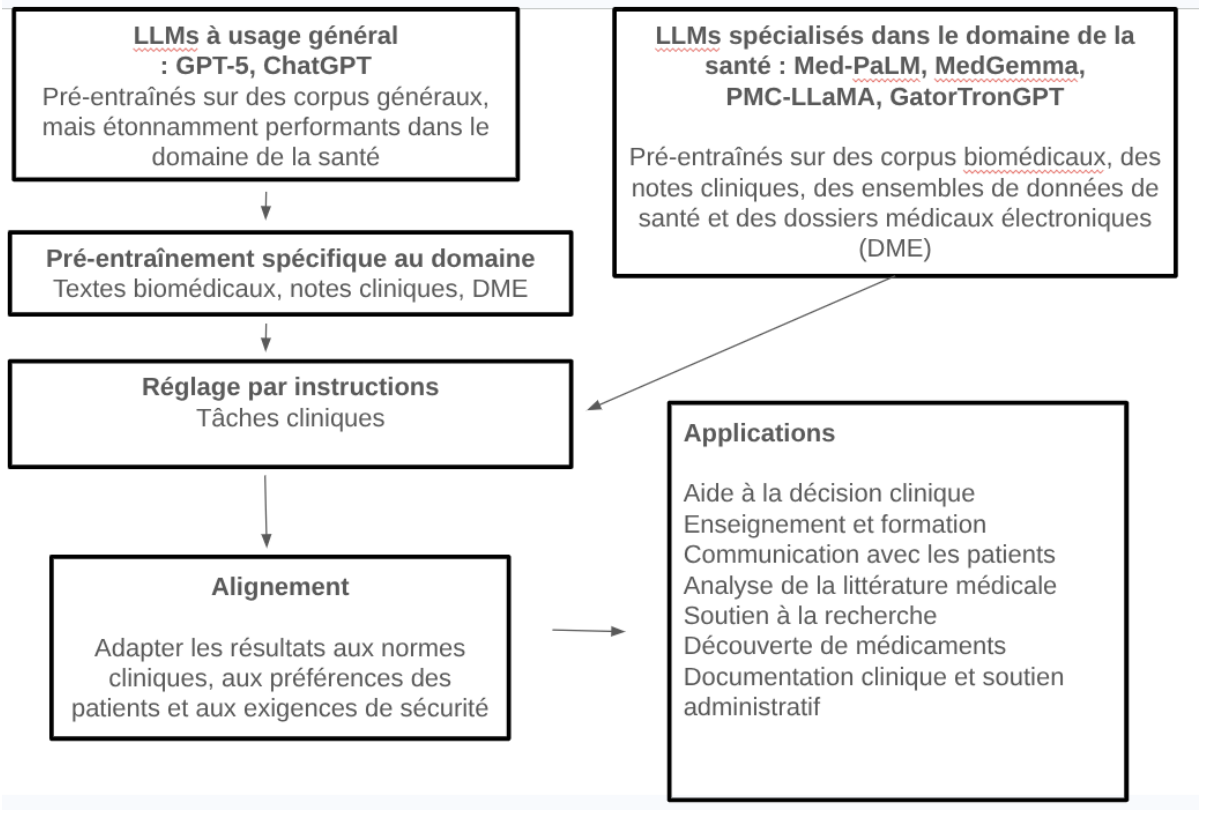


Post-Processing

- Instruction-Tuning
  Modèles de langage génériques $\rightarrow$ Modèles capables de réaliser une grande variété de tâches.

- Alignement
  Permet d'adapter les modèles à suivre plus fidèlement les instructions humaines.

LLMs pour le médical



Applications dans le domaine de la santé

Aide à la décision clinique

- **GatoTron** (NIH) analyse les dossiers médicaux électroniques afin de détecter d'éventuelles interactions médicamenteuses

Enseignement et formation

- Génération de contenu pédagogique personnalisé avec retour d'information immédiat
- La simulation de consultations patients donne aux étudiants l'occasion de développer leurs capacités de raisonnement clinique dans un environnement sûr et contrôlé

Communication avec les patients

- Les assistants médicaux virtuels peuvent faciliter le triage des patients, l'évaluation des symptômes et l'orientation vers les soins appropriés (chatbot **Florence** du NHS et le **chatbot de Babylon Health**)

Applications dans le domaine de la santé

Analyse et synthèse de la littérature médicale

- aident les professionnels de santé à se tenir informés des dernières avancées et des pratiques

Soutien à la recherche

- Accélèrent les découvertes en analysant de vastes ensembles de données issues de dossiers médicaux, d'essais cliniques et de la littérature scientifique.
- Permettent d'identifier de **nouveaux traitements potentiels, de mettre au point des thérapies efficaces et de comprendre les mécanismes des maladies**

Documentation clinique et assistance aux tâches administratives.

- Rédaction de rapports, organisation et synthèse des informations relatives aux patients

Défis et limites des LLMs dans le domaine de la santé

Techniques

- Hallucinations : génération d’informations plausibles mais factuellement erronées
- Perte d'information
- Compréhension contextuelle limitée et lacunes dans les connaissances
- Exigences informatiques et consommation d’énergie
- Latence des réponses du modèle, problématique dans les cas urgents

Biais et équité en matière de santé

- Risque de perpétuer les biais issus des données médicales historiques
- Exacerbation des inégalités en matière de santé

Défis et limites des LLMs dans le domaine de la santé

Considérations éthiques et gouvernance

- Préoccupations relatives à la vie privée des patients et respect des réglementations
- Consentement éclairé et transparence
- Responsabilité

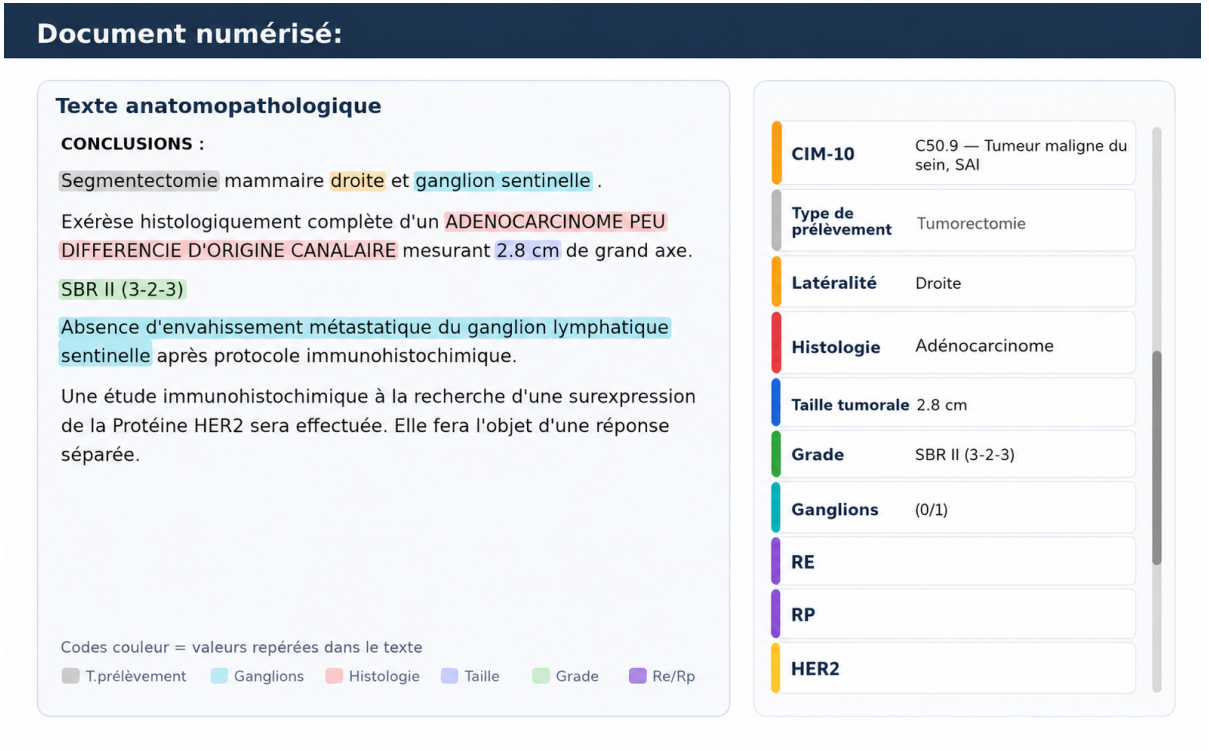
Intégration dans le flux de travail clinique

- Perturbation des routines et des environnements cliniques existants
- Interprétabilité avec les systèmes de dossiers médicaux électroniques

LLMs, confidentialité des données des patients, extraction d'informations et génération de données synthétiques

A. Lorenzo, A. Coulet et C. Gardent. Privacy-Preserving Generation of Synthetic Pathology Reports for Information Extraction. AIME 2026.

Extraction d'Information



Rapports d'anatomopathologie $\rightarrow$ 13 clinical variables

3,880 dossiers de patientes atteintes d'un cancer du sein

Données réelles

Rapports

- Comptes rendus d'anatomopathologie rédigés par des médecins des quatre cliniques concernant des patientes atteintes d'un cancer du sein.

Données

- Treize variables cliniques incluant notamment le *diagnostic, la taille et la localisation de la tumeur*, ainsi que le *statut des récepteurs aux œstrogènes*

Attribute	Type	Description / Possible values
Test used for diagnosis	Enumerate	lumpectomy; biopsy; mastectomy; surgical excision; lymph node dissection
Morphology	Enumerate	Morphology code according to the International Classification of Diseases for Oncology (ICD-O), French version
SBR grade	Enumerate	Scarff-Bloom-Richardson (SBR) grade: SBR 1; SBR 2; SBR 3
Tumor size	Numeric	Tumor size (usually expressed in millimeters or centimeters)
Clear margins	Enumerate	yes; no. Indicates whether healthy tissue surrounds the excised tumor
Number of lymph nodes taken	Numeric	Total number of lymph nodes removed during surgery
Number of lymph nodes affected	Numeric	Number of lymph nodes showing tumor involvement
Presence of lymphovascular invasion	Enumerate	yes; no. Indicates tumor invasion of lymphatic or blood vessels
Presence of extracapsular tumor spread	Enumerate	yes; no. Indicates tumor extension beyond the lymph node capsule
Estrogen receptor status	Enumerate	positive; negative
Progesterone receptor status	Enumerate	positive; negative
HER2 status	Enumerate	Human epidermal growth factor receptor 2: positive; negative
Ki-67 percentage	Numeric	Percentage of Ki-67 nuclear protein expression

Traitement des données réelles conforme au RGPD

- Pseudonymisation des données
- Information des patients
- Dépot d'un projet conforme à la méthodologie de référence MR-004 de la CNIL sur les site web du Health Data Hub
- Traitement des données effectué localement au sein de l’infrastructure du groupe U2R.

Générer des paires « rapport-données »

Étape 1 : génération de **données synthétiques**

- Utilisation d'un synthétiseur de données tabulaires pour générer des données synthétiques (ensembles de 13 paires variable-valeur)
- $\rightarrow$ Cela préserve la **confidentialité des données patients**

Étape 2 : génération des rapports à partir des données synthétiques

- Distillation de connaissances basée sur les données à partir de 4 modèles de langage (LLM)
- $\rightarrow$ Cela favorise la **diversité des rapports**

Générer des données synthetiques

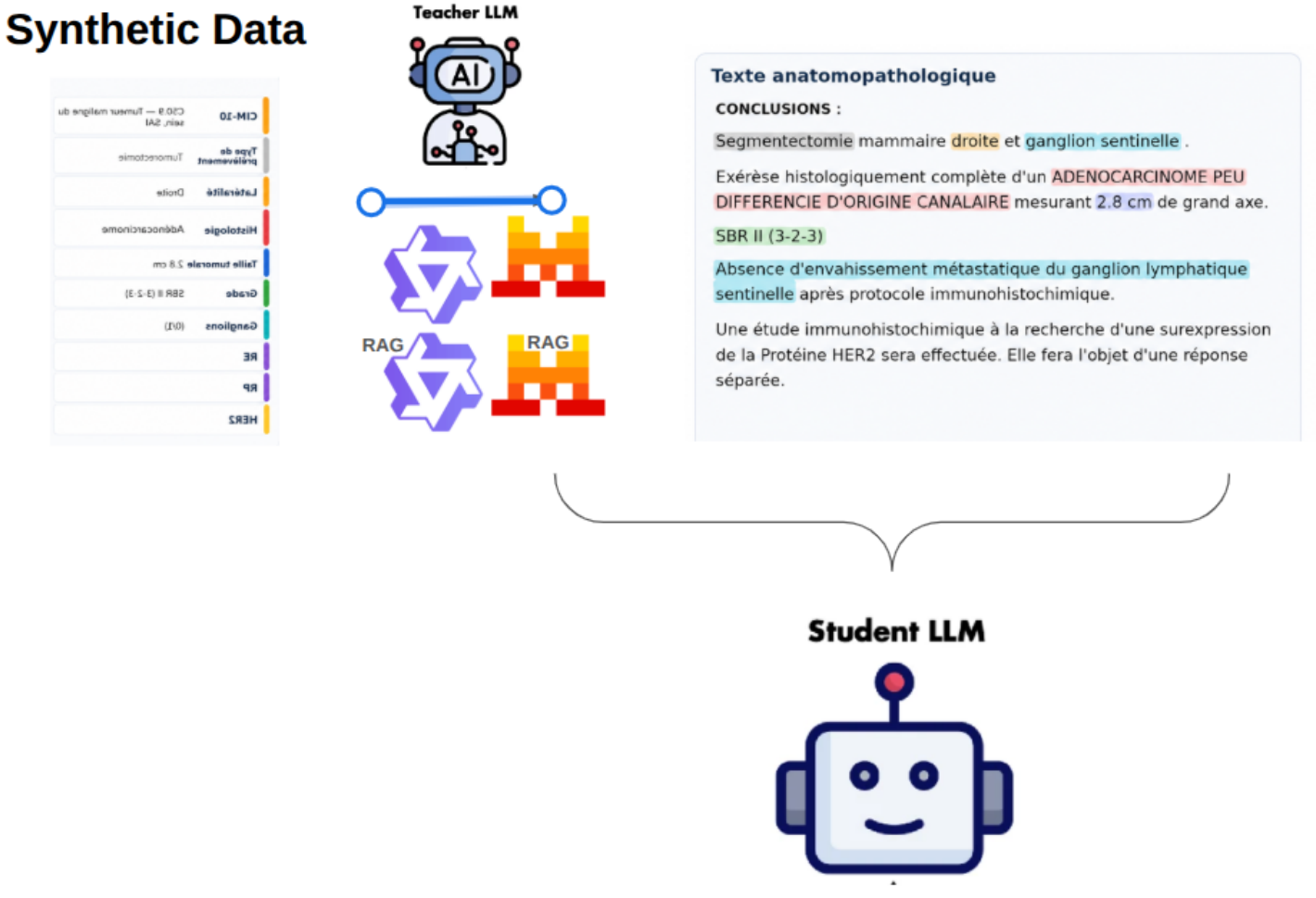
Comparaison de 4 synthésiseurs de données tabulaires

Metric \ Synthesizer	Gaussian Copula	CTGAN	PATE-GAN	MST
Fidelity				
Overall Accuracy	86.59%	74.57%	82.56%	**92.22%**
Accuracy - Univariate	96.57%	86.68%	93.46%	**98.36%**
Accuracy - Bivariate	86.46%	74.17%	82.34%	**92.22%**
Accuracy - Trivariate	76.72%	62.86%	71.87%	**86.07%**
Similarity				
Cosine Similarity	0.984	0.886	0.973	**0.999**
Novelty				
↓ % Identical Matches	5.85%	**1.76%**	3.68%	17.62%
↑ NNDR	0.961	0.560	**1.080**	1.189
↓ DCR – Share	55.10%	56.47%	54.90%	**52.36%**
DCR Overfitting Score	**1.0**	**1.0**	**1.0**	**1.0**

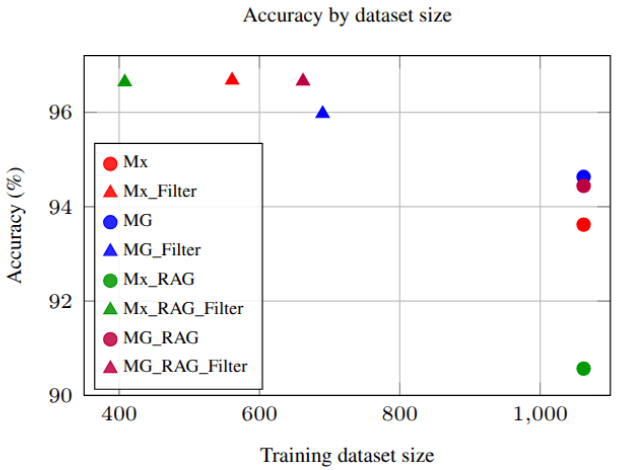
Faithful and Privacy Preserving

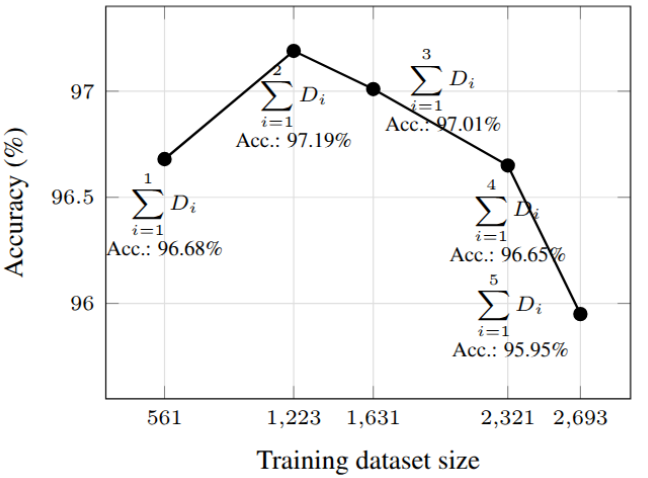
Distillation de connaissances guidées par les données

Data Driven Knowledge Distillation (DDKD)



La qualité l'emporte sur la quantité





Les modèles entraînés sur des données filtrées obtiennent de meilleurs résultats avec moins de données.

Extraction d'information

	Base Model				Fine-Tuned Models					
	Mi0s		Mi3s		MiS12		MiS123		MiS12_T12	
attribute	acc.	f1	acc.	f1	acc.	f1	acc.	f1	acc.	f1
diagnostic	0.22	0.28	0.16	0.23	0.36	0.52	0.37	0.52	**0.41**	**0.54**
echantillon	0.52	0.68	0.66	0.79	0.68	0.81	0.71	0.83	0.71	0.83
emboles	0.68	0.68	0.73	0.76	0.94	0.94	0.93	0.93	**0.95**	**0.96**
grade	0.55	0.59	0.58	0.66	0.87	0.89	0.87	0.90	**0.88**	0.90
her2	0.49	0.59	0.77	0.81	0.82	0.86	0.92	0.94	0.91	0.93
ki67	0.84	0.84	0.95	0.95	0.96	0.96	0.96	0.96	**0.97**	**0.97**
marges_saines	0.54	0.51	0.66	0.61	0.75	0.39	0.74	0.36	**0.81**	**0.64**
numero_gang_att	0.69	0.43	0.84	0.76	0.95	0.93	0.96	0.95	0.95	0.94
numero_gang_prel	0.89	0.84	0.94	0.93	0.95	0.94	0.97	0.96	0.97	0.96
re	0.59	0.65	0.80	0.84	0.93	0.95	0.93	0.93	**0.94**	0.95
rp	0.56	0.60	0.74	0.77	0.92	0.93	0.85	0.85	**0.95**	**0.95**
rupture_capsulaire	0.54	0.22	0.88	0.52	0.93	0.36	0.96	0.72	0.95	0.62
taille	0.36	0.32	0.34	0.41	0.81	0.75	0.80	0.75	**0.84**	**0.77**
Total	0.57	0.57	0.69	0.71	0.84	0.84	0.84	0.84	**0.86**	**0.86**

	Base Model				Fine-Tuned Models					
	Mg0s		Mg3s		MgS12		MgS123		MgS12_T12	
attribute	acc.	f1	acc.	f1	acc.	f1	acc.	f1	acc.	f1
diagnostic	0.23	0.30	0.45	0.60	0.42	0.55	**0.56**	**0.69**	0.46	0.59
echantillon	**0.89**	**0.93**	0.81	0.88	0.86	0.92	0.85	0.92	0.80	0.89
emboles	0.95	0.95	0.94	0.94	0.95	0.95	0.95	0.96	0.94	0.95
grade	0.88	0.91	0.80	0.85	**0.95**	**0.96**	0.83	0.88	0.90	0.93
her2	0.94	0.95	0.88	0.91	0.94	0.95	0.93	0.95	**0.95**	**0.96**
ki67	0.98	0.98	0.86	0.88	0.98	0.98	0.98	0.98	0.98	0.98
marges_saines	0.86	0.80	0.86	0.80	0.84	0.72	**0.86**	0.79	0.86	0.78
numero_gang_att	0.98	0.97	0.88	0.87	0.98	0.97	0.98	0.97	0.95	0.93
numero_gang_prel	0.98	0.97	0.89	0.87	0.98	0.98	0.96	0.95	0.95	0.94
re	0.93	0.95	0.87	0.90	0.93	0.94	0.91	0.93	**0.97**	**0.98**
rp	0.94	0.95	0.86	0.89	0.93	0.94	0.91	0.93	**0.97**	**0.98**
rupture_capsulaire	0.98	0.87	0.95	0.69	0.98	0.90	0.93	0.36	0.98	0.90
taille	0.58	0.64	0.67	0.68	0.87	0.82	0.85	0.83	0.86	0.81
Total	0.85	0.86	0.82	0.84	**0.89**	**0.90**	0.88	0.89	**0.89**	0.89

L’entraînement du modèle sur des données synthétiques

- améliore les performances.
- permet d'utiliser un modèle plus petit (Mistral-7B vs MedGemma-27B)

Résumé

La ***génération de données synthétiques*** permet de préserver la confidentialité des données.

L'utilisation de ***LLMs, de la DDKD, du filtrage et de la combinaison de données*** pour générer les rapports correspondants permet de créer des paires (données, rapports) de haute qualité, qui peuvent être utilisées efficacement pour entraîner un modèle d'extraction d'information.

Questions ?