

# Agent conversationnel et chatbot

Utilisation de Large Language Models pour publier une  
base de connaissances  
Partie 1

Nicolas Debeissat  
[nicolas.debeissat@mail-formateur.net](mailto:nicolas.debeissat@mail-formateur.net)

# Plan du module

## Pourquoi faire un agent conversationnel ?

Comment défendre votre projet en entreprise ?

## Principes de base des LLMs

Ce qui va faire qu'il va donner des bons résultats, ou non

## Entraînement - Fine Tuning

Comment adapter ses réponses

## Retrieval Augmented Generation

Répondre à partir de vos données

## Intégration dans une application

Créer une API à partir de votre modèle

## Evaluation

Entraîner votre IA à faire votre évaluation

# Pourquoi faire un agent conversationnel ?

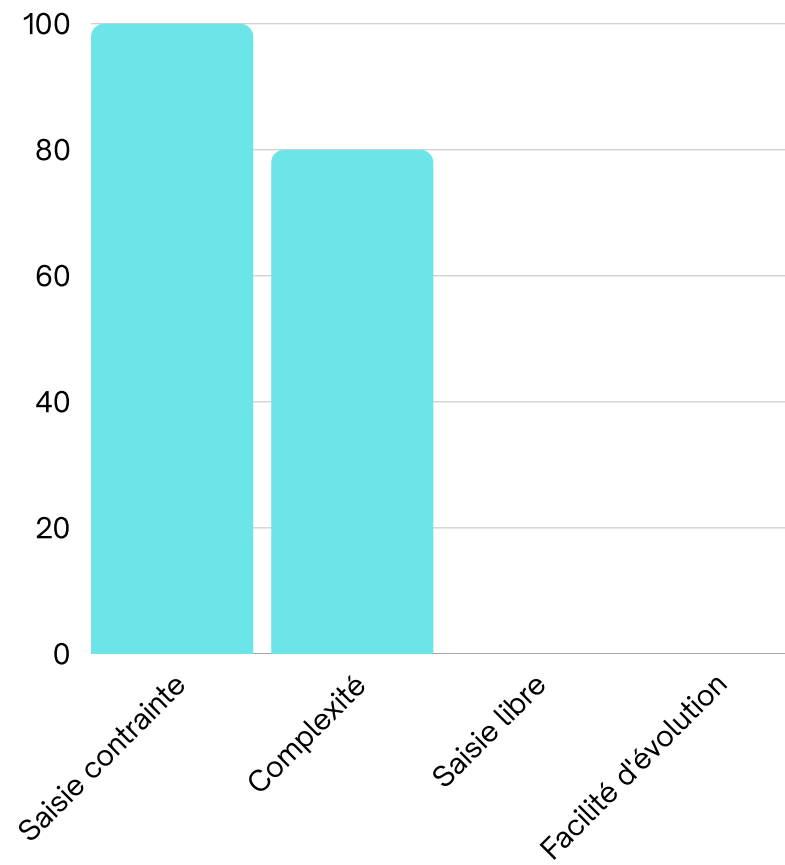
Comment défendre votre projet en entreprise ?

# Techniques traditionnelles d'Interfaces Hommes-Machines

# Techniques traditionnelles d'Interfaces Hommes-Machines



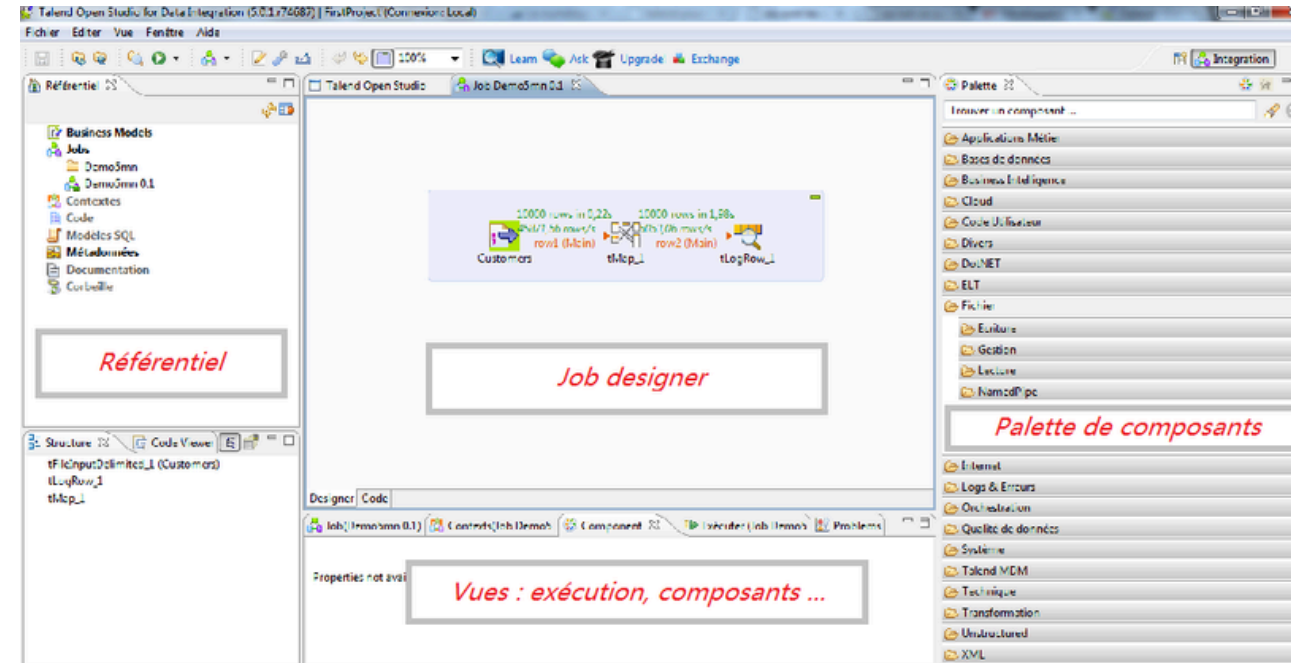
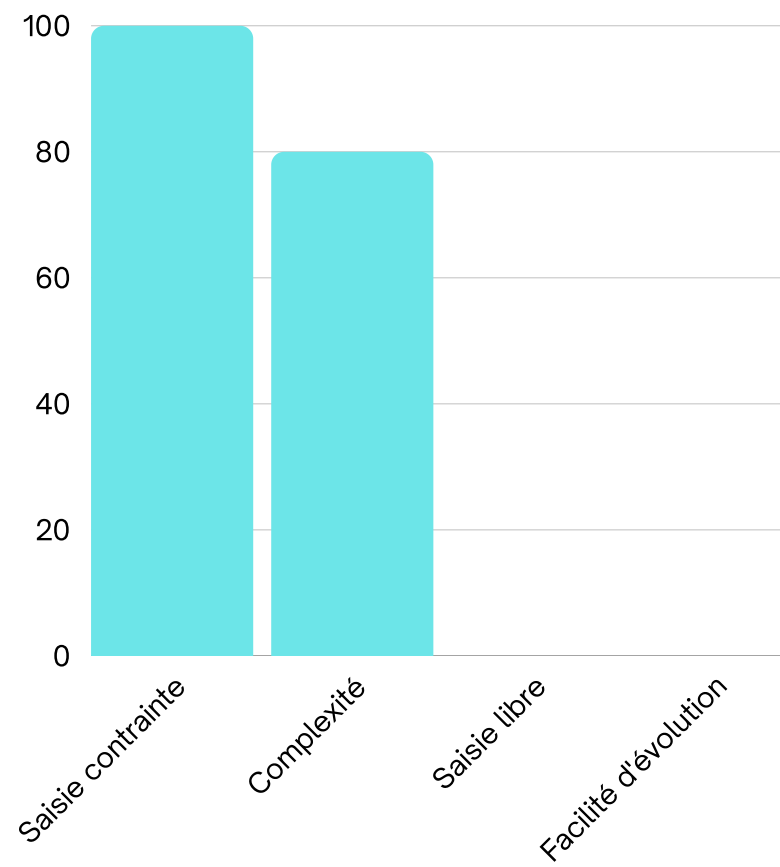
## Click-bouton



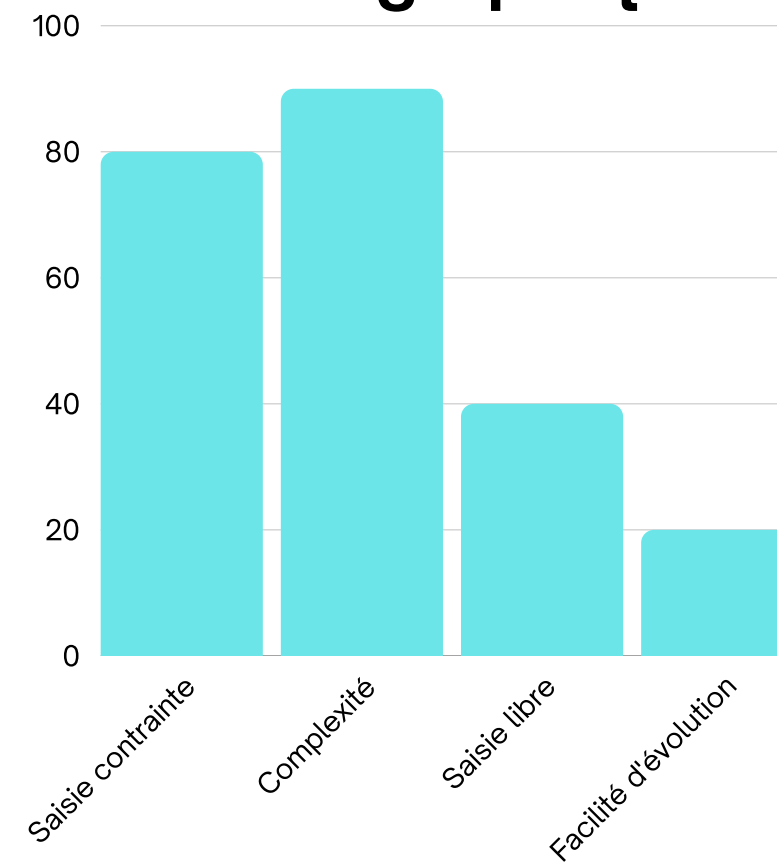
# Techniques traditionnelles d'Interfaces Hommes-Machines



Click-bouton



Palette graphique

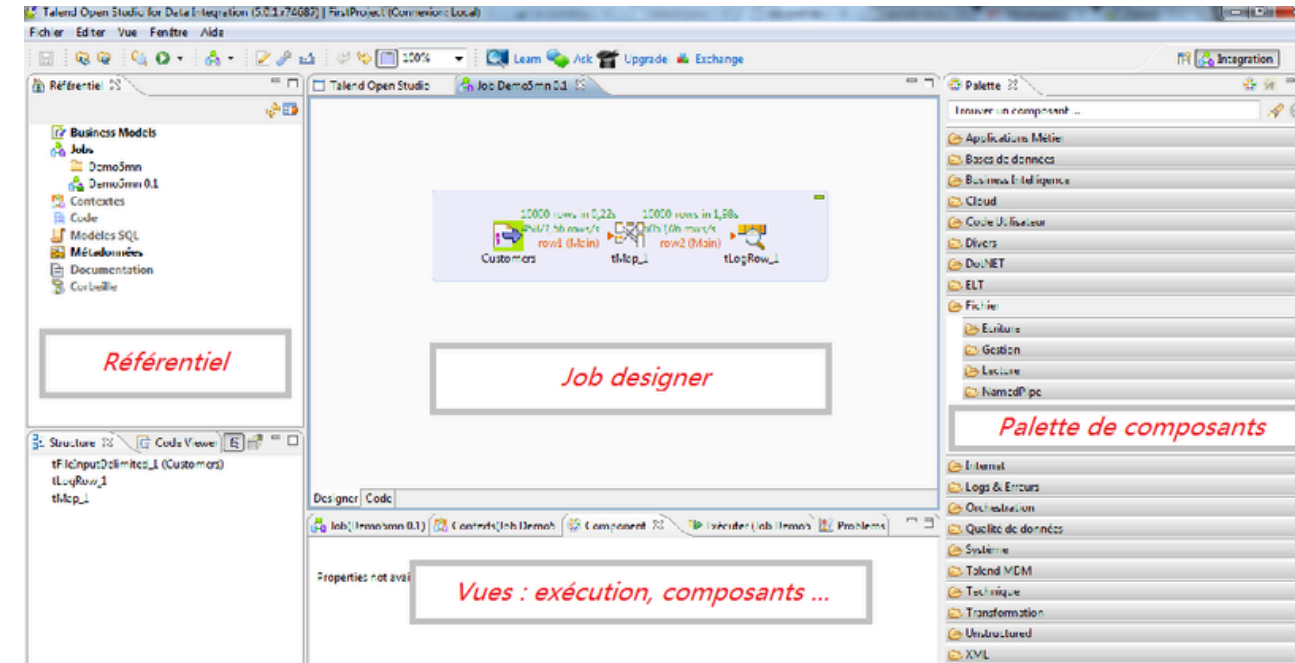
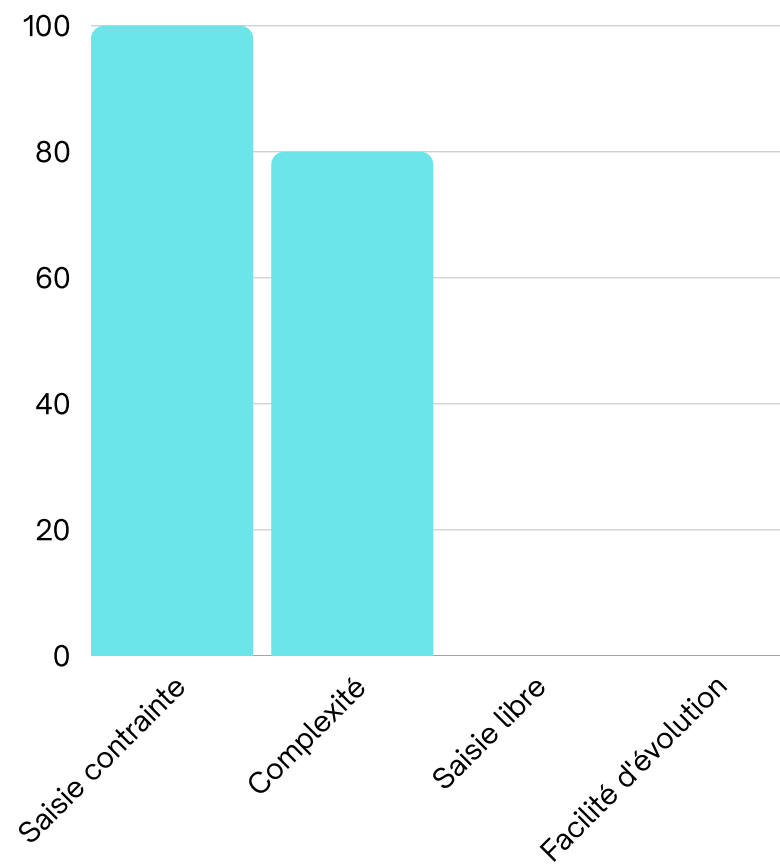




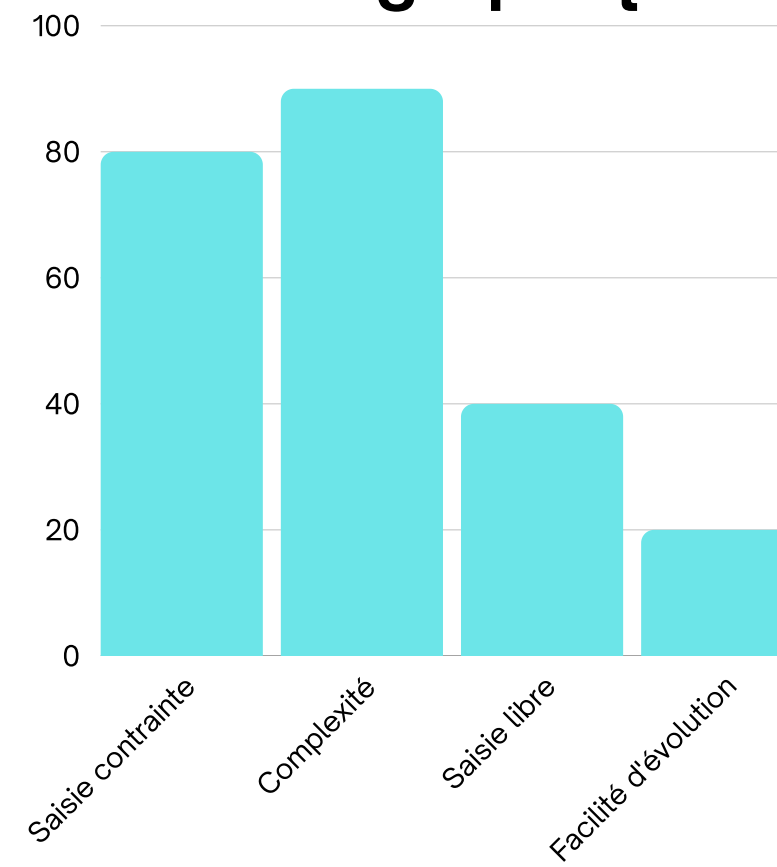
# Techniques traditionnelles d'Interfaces Hommes-Machines



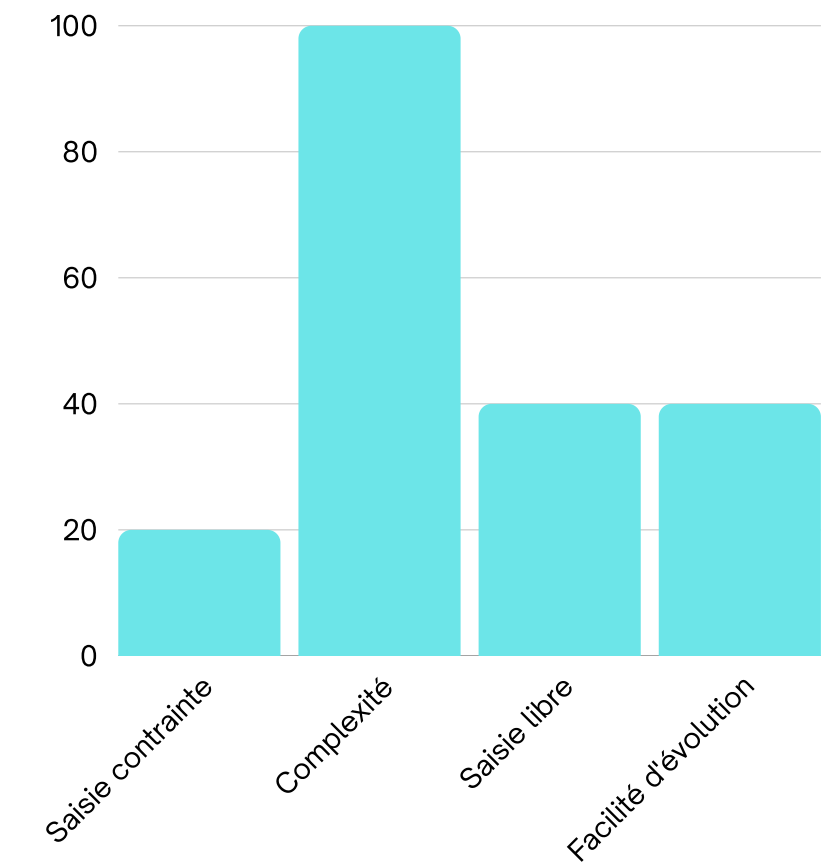
Click-bouton



Palette graphique



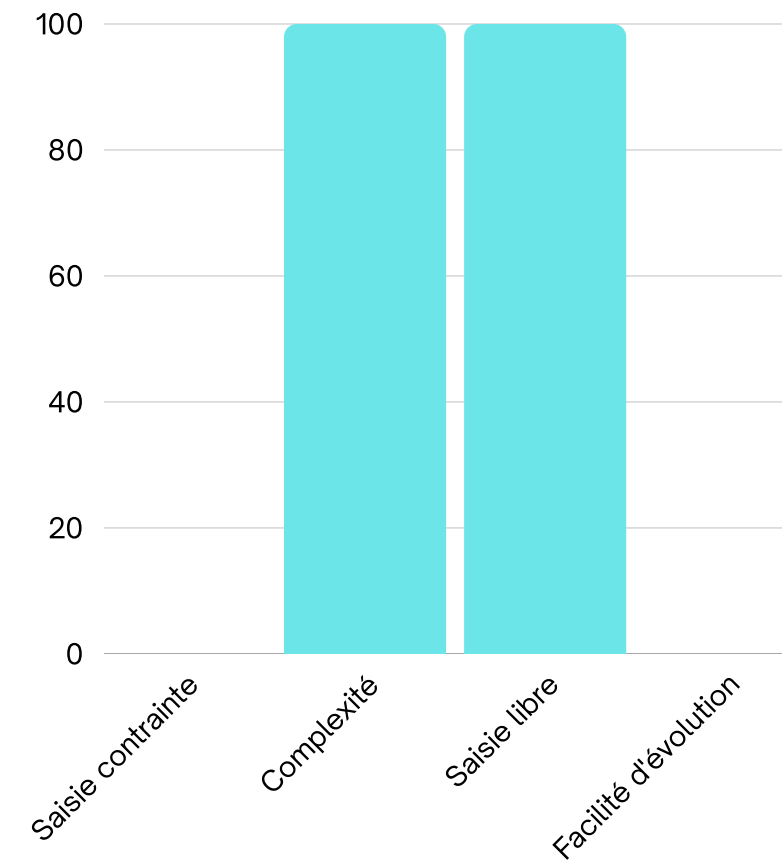
Recherches avancées



# Techniques traditionnelles d'Interfaces Hommes-Machines

```
34  if (FlagNewNum) {
35      FKeyPad.ReadOut.value = Num;
36      FlagNewNum = false;
37  }
38  else {
39      if (FKeyPad.ReadOut.value == "0")
40          FKeyPad.ReadOut.value = Num;
41      else
42          FKeyPad.ReadOut.value += Num;
43  }
44  }
45
46  function Operation (Op) {
47      var Readout = FKeyPad.ReadOut.value;
48      if (FlagNewNum && PendingOp != "=");
49      else
50      {
51          FlagNewNum = true;
52          if ( '+' == PendingOp )
53              FKeyPad.ReadOut.value = parseFloat(FKeyPad.ReadOut.value);
54          else if ( '-' == PendingOp )
55              Accumulate -= parseFloat(Readout);
56          else if ( '/' == PendingOp )
57              FKeyPad.ReadOut.value = parseFloat(FKeyPad.ReadOut.value);
58          else if ( '*' == PendingOp )
59              Accumulate *= parseFloat(Readout);
60          else
61              Accumulate = parseFloat(Readout);
62              FKeyPad.ReadOut.value = Accumulate;
63              PendingOp = Op;
64      }
```

Code

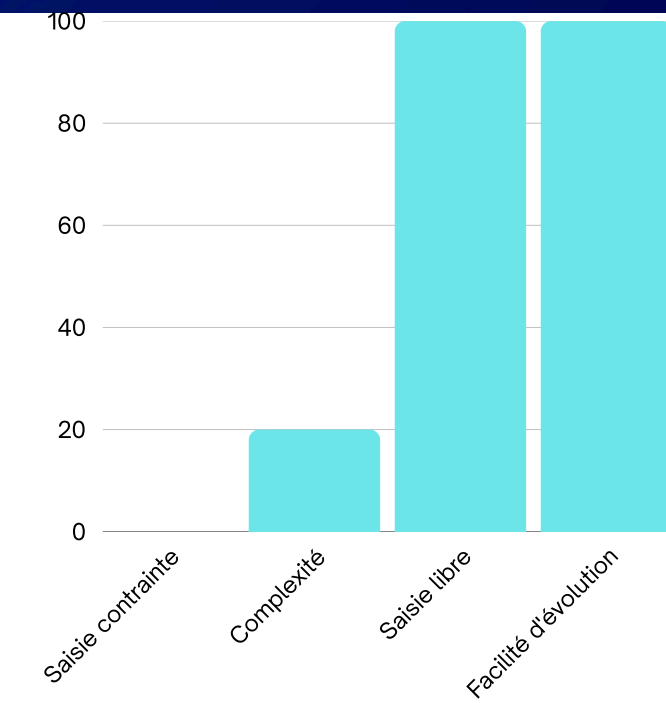
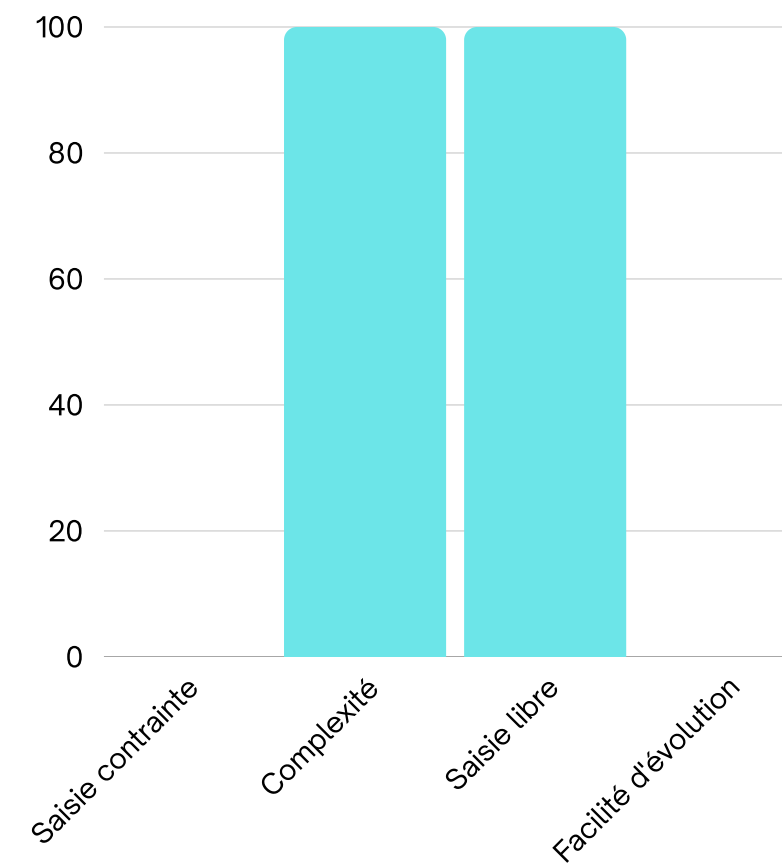




# Techniques traditionnelles d'Interfaces Hommes-Machines

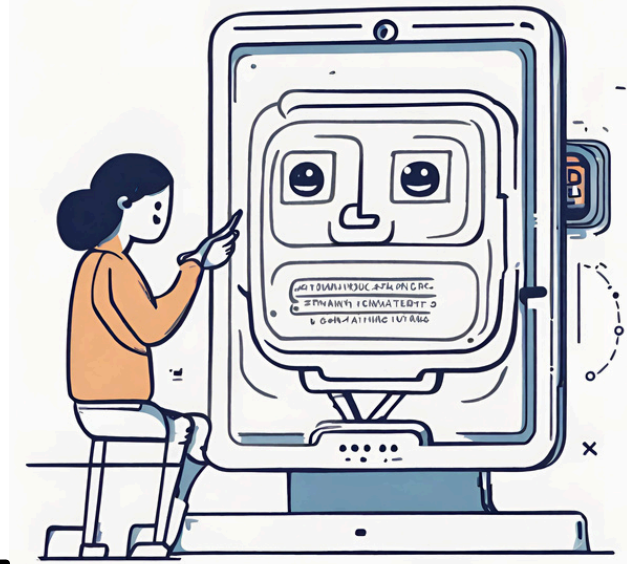
```
34  if (FlagNewNum) {
35      FKeyPad.ReadOut.value = Num;
36      FlagNewNum = false;
37  }
38  else {
39      if (FKeyPad.ReadOut.value == "0")
40          FKeyPad.ReadOut.value = Num;
41      else
42          FKeyPad.ReadOut.value += Num;
43  }
44  }
45
46  function Operation (Op) {
47      var Readout = FKeyPad.ReadOut.value;
48      if (FlagNewNum && PendingOp != "");
49      else
50      {
51          FlagNewNum = true;
52          if ( '+' == PendingOp )
53              FKeyPad.ReadOut.value = parseFloat(FKeyPad.ReadOut.value);
54          else if ( '-' == PendingOp )
55              Accumulate -= parseFloat(Readout);
56          else if ( '/' == PendingOp )
57              FKeyPad.ReadOut.value = parseFloat(FKeyPad.ReadOut.value);
58          else if ( '*' == PendingOp )
59              Accumulate *= parseFloat(Readout);
60          else
61              Accumulate = parseFloat(Readout);
62              FKeyPad.ReadOut.value = Accumulate;
63              PendingOp = Op;
64      }
65  }
```

Code



# Comment faire un chatbot

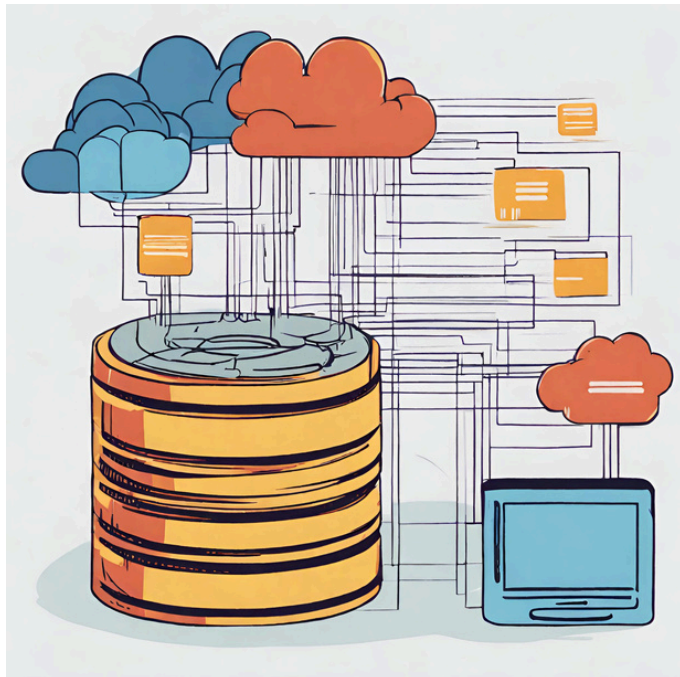
LLM généraliste



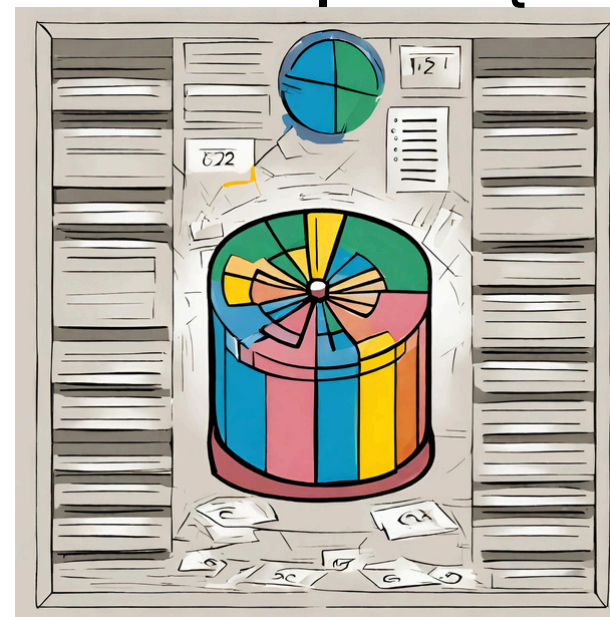
pré-entraîné à  
compléter  
une phrase

entraîné à  
répondre à  
des questions

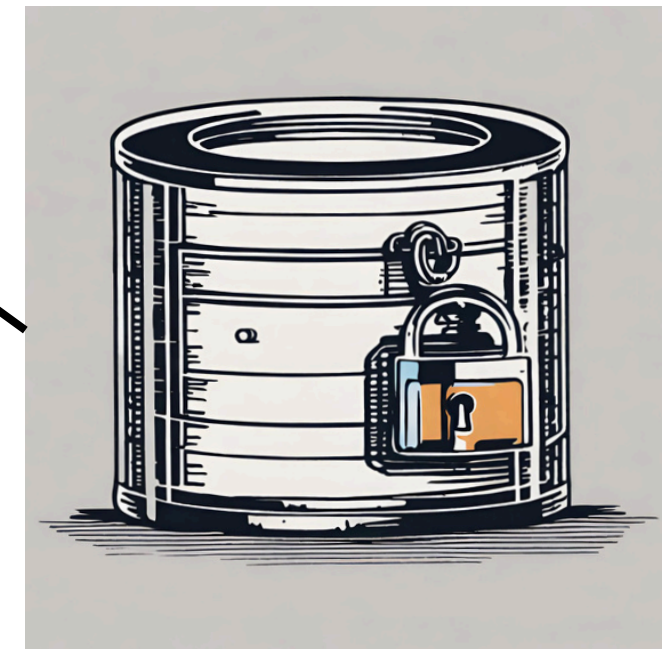
Fine-tuné sur  
les données  
propriétaires



Données publiques  
massives



Corpus de question-réponses  
<https://huggingface.co/datasets/bigscience/P3>



Données privées  
de l'entreprise

# Parmi les tâches qu'un LLM sait nativement faire

- répondre à des QCMs de culture générale (Multiple-Choice QA)

*Wall St. Bears Claw Back Into the Black (Reuters) Reuters - Short-sellers, Wall Street's dwindling\band of ultra-cynics, are seeing green again. Is this a piece of news regarding world politics, sports, business, or science and technology?*

- extraire des informations d'un texte donné en contexte (Extractive QA)

*Given the following passage "A generation later,*

*...*

*Berkeley's approach to empiricism would later come to be called subjective idealism.", answer the following question. Note that the answer is present within the text. Question: what concept is mentioned last?*

- répondre à des questions de culture générale (Close-book QA)

*Give me the correct facts to answer this: what is the name of justin bieber brother?*

# Parmi les tâches qu'un LLM sait nativement faire

- comprendre les concepts évoqués dans un texte, les extraire sous la forme de quelques mots, ou générer une phrase avec (Structure-to-text)

*Read the bio below and try to give details on walter extra 's: - birth date - name - nationality - occupation Bio: ...*

- analyser la satisfaction ou l'insatisfaction exprimée dans un texte, sous la forme d'une note (Sentiment)

*Is this product review positive? Title: Stuning even for the non-gamer Review: This sound track was beautiful!*

*...*

*Answer:*

- résumer un texte (Summarization)

*Generate a summary for this dialogue: Tim: Hi, what's up?*

*...*

*Kim: thanks, maybe I'll do that Tim: I also like using post-its in kaban style*



# Parmi les tâches qu'un LLM sait nativement faire

- classer un texte parmi des thèmes (Topic Classification)

*Carlyle Looks Toward Commercial Aerospace (Reuters) Reuters - Private investment firm Carlyle Group, \which has a reputation for making well-timed and occasionally \controversial plays in the defense industry, has quietly placed \its bets on another part of the market. What label best describes this news article?*

- reformuler une phrase (Paraphrase Identification)

*Paraphrase the sentence: The world 's first laser was developed in 1960 by the American scientists Nikolay Basov and Alexander Prokhorov and the Russian scientist Charles H. Townes .*

- à partir d'un texte, répondre à une question (Natural Language Inference)

*Alexandra Lendon Bastedo (9 March 1946 – 12 January 2014) was a British actress, best known for her role as secret agent Sharron Macready in the 1968 British espionage/science fiction adventure series "The Champions". She has been cited as a sex symbol of the 1960s and 1970s. Bastedo was a vegetarian and animal welfare advocate. Question: Bastedo didn't keep any pets because of her views on animal rights. True, False, or Neither?*

# Parmi les tâches qu'un LLM sait nativement faire

- compléter une phrase à partir des informations contenues dans une autre phrase (Coreference Resolution)

*John moved the couch from the garage to the backyard to create space. The \_ is small. Replace the \_ in the above sentence with the correct option: - garage - backyard*

- comprendre dans un texte à quoi fait référence un pronom (Word Sense disambiguation)

*Sentence A: Do you want to come over to my place later? Sentence B: A political system with no place for the less prominent groups. "place" has a similar meaning in sentences A and B. True or False?*

- trouver une relation de cause à effet entre 2 phrases (Sentence Completion)

*Complete the sentence: Then, the man writes over the snow covering the window of a car, and a woman wearing winter clothes smiles. then*



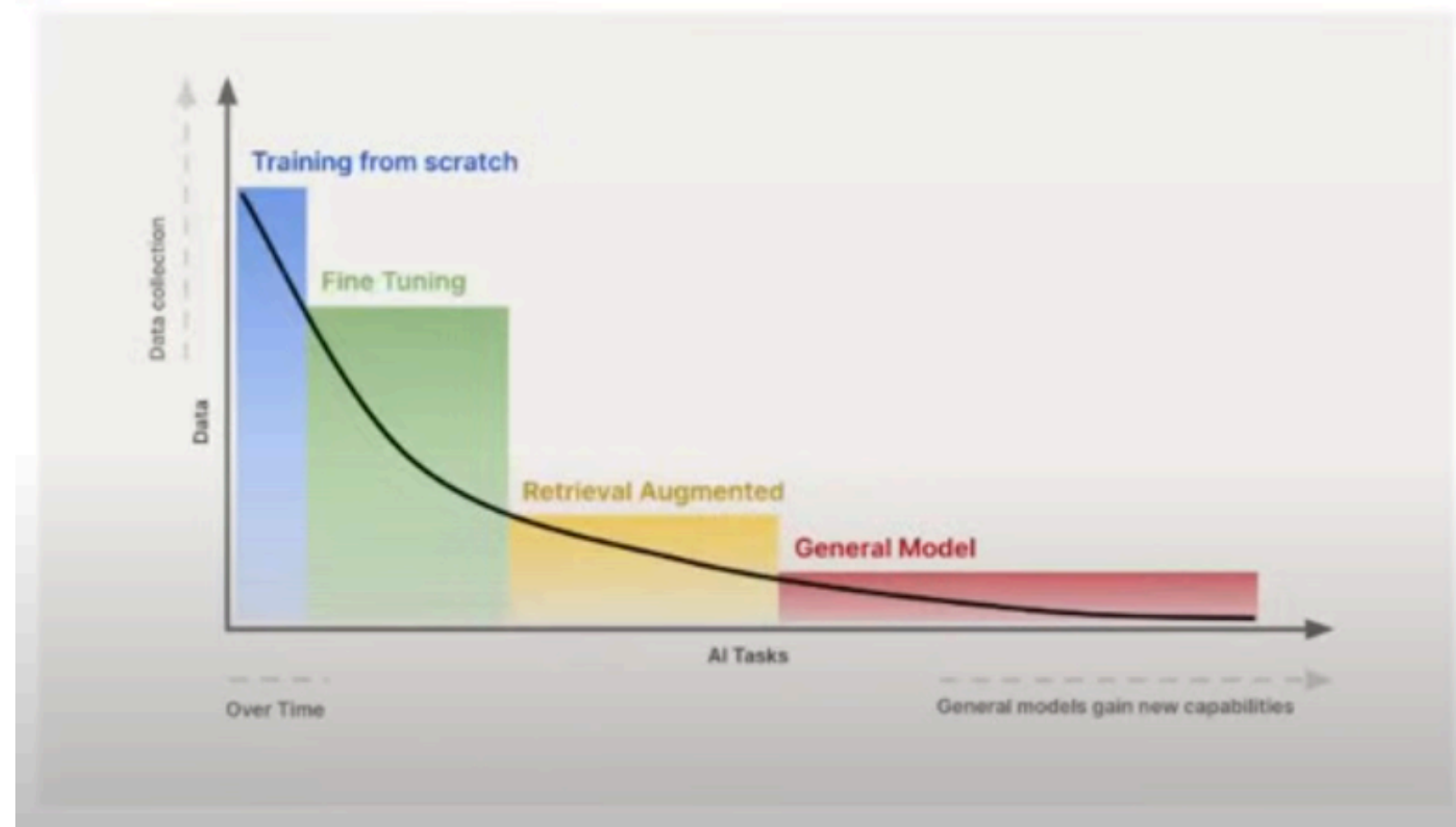
# Combien de données nécessaires

Des recherches sont en cours pour le déterminer et le comprendre.

Intuitivement : LLM  $\sim$  7M paramètres  $\Rightarrow$  7M d'inconnues

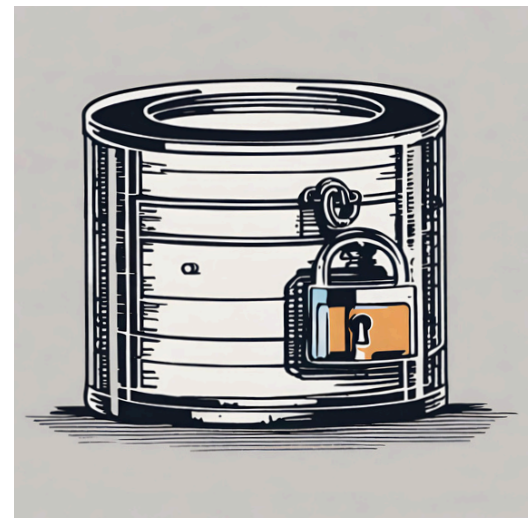
Pour 1 réseau de 1 neurone, 2 inconnues :  $\text{gelu}(w * x + b) = y$

Il faut sans doute significativement plus d'exemples que d'inconnues



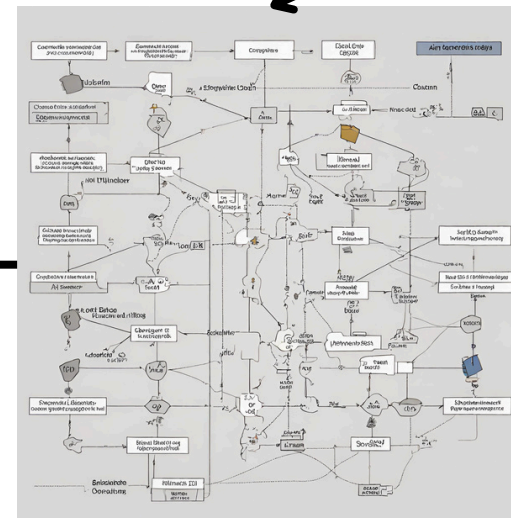
# Génération des données d'exemple

Questions  
utilisateurs



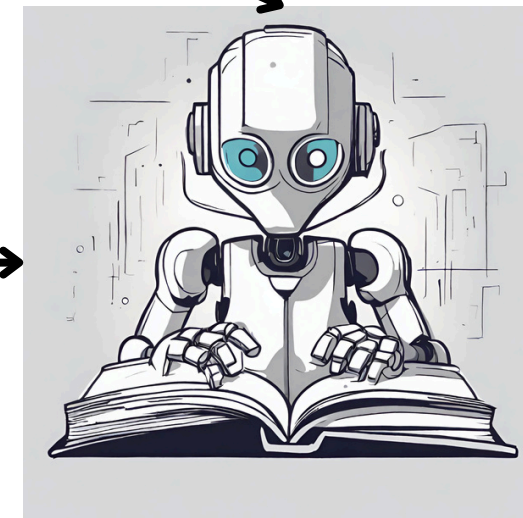
Données privées

analyse



Algo Non-IA  
ou RAG

alimente le  
fine-tuning



Modèle IA

# Pourquoi faire de l'IA ?

Exemple : extraction d'un profil à partir de données

Algo non IA	Algo IA

# Pourquoi faire de l'IA ?

Exemple : extraction d'un profil à partir de données

Algo non IA	Algo IA
Limité aux données connues	Généralisation à tout le langage
Bugs, nécessité de maintenir de gros volumes de code	Si détection d'un problème, enrichissement des données d'exemple et ré-entraînement
Parfois plus lent !	Performances consistantes