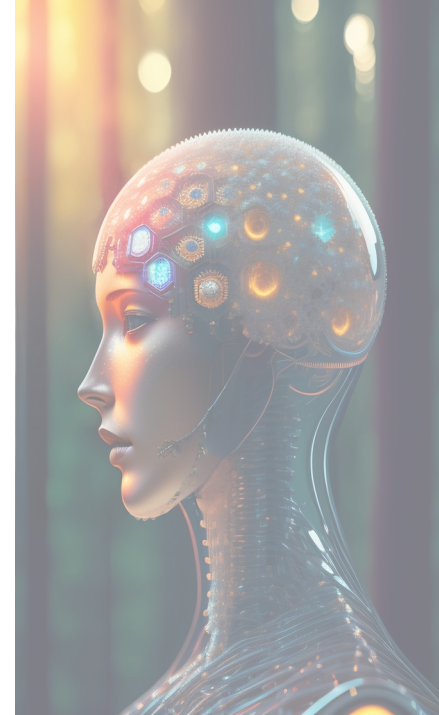


ChatGPT comprend-il
quelque chose à ce qu'il
raconte ?

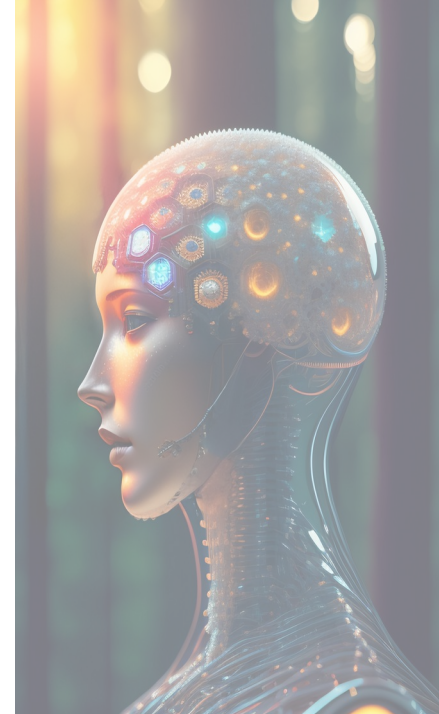
Le point de vue (pas du tout)
modeste d'un linguiste...



La linguistique et les LLMs

- Peu de liens entre la linguistique théorique et le deep learning en général
- Pourtant, tradition de forte interdisciplinarité en linguistique
- Des travaux qui mettent à l'épreuve les théories linguistiques
- Un formidable outil pour la recherche (potentiellement)

E. Schang (LLL)



Baroni (2022)

On the proper role of linguistically-oriented deep
net analysis in linguistic theorizing

Marco Baroni^{1,2}

¹Universitat Pompeu Fabra

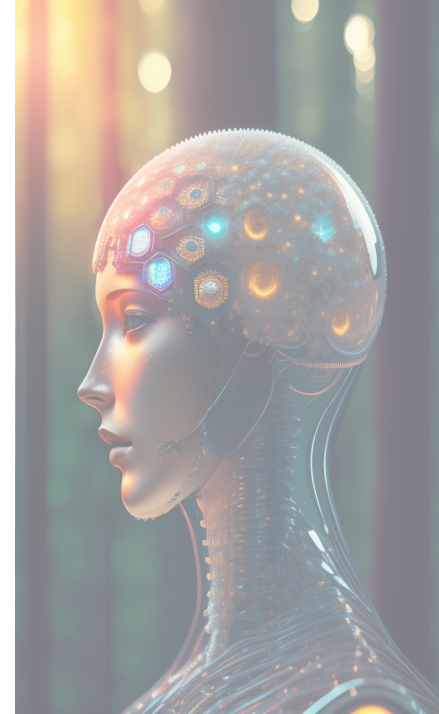
²Catalan Institution for Research and Advanced Studies (ICREA)

March 25, 2022

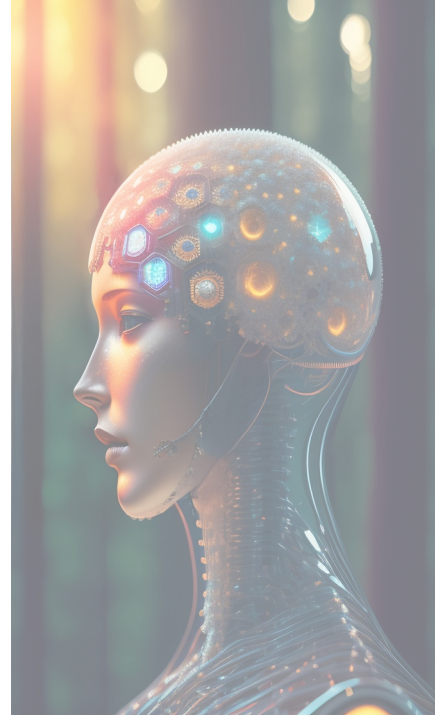
Abstract

A lively research field has recently emerged that uses experimental methods to probe the linguistic behavior of modern deep networks. While work in this tradition often reports intriguing results about the grammatical skills of deep nets, it is not clear what their implications for linguistic theorizing should be. As a consequence, linguistically-oriented deep net analysis has had very little impact on linguistics at large. In this chapter, I suggest that deep networks should be treated as theories making explicit predictions about the acceptability of linguistic utterances. I argue that, if we overcome some obstacles standing in the way of seriously pursuing this idea, we will gain a powerful new theoretical tool, complementary to mainstream algebraic approaches.

E. Schang (LLL)



LLMs et linguistique formelle



Linguistique formelle : fondée sur l'idée de computations syntaxiques qui seules peuvent rendre compte de l'infinité de phrases possibles dans une langue

« The problem of language acquisition has often been summarized in terms of the problem of the poverty of the stimulus. Our linguistic capacity, for instance our knowledge of English, goes beyond the evidence we have been exposed to in our childhood. The linguist wants to account for the fact that the linguistic competence is attained in spite of important inadequacies in the stimulus, the linguistic experience. » *Haegemann 1994*

LLMs et linguistique formelle

- Gold 1967 montre que les preuves positives ne suffisent pas pour que les apprenants identifient la classe de la langue cible de leurs parents.

INFORMATION AND CONTROL **10**, 447–474 (1967)

Language Identification in the Limit

E MARK GOLD*

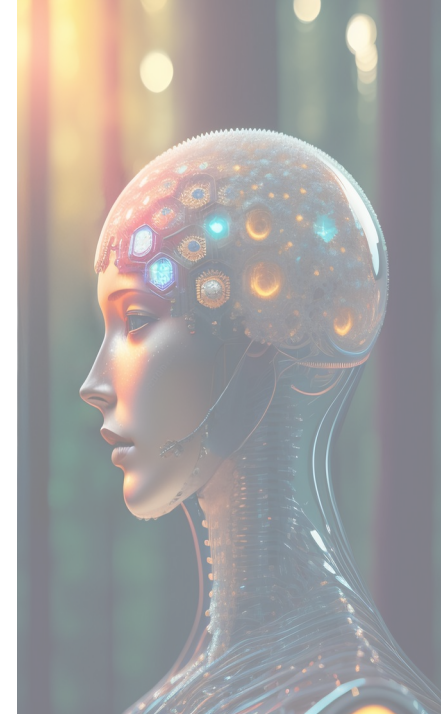
The RAND Corporation

Language learnability has been investigated. This refers to the following situation: A class of possible languages is specified, together with a method of presenting information to the learner about an unknown language, which is to be chosen from the class. The question is now asked, "Is the information sufficient to determine which of the possible languages is the unknown language?" Many definitions of learnability are possible, but only the following is considered here: Time is quantized and has a finite starting time. At each time the learner receives a unit of information and is to make a guess as to the identity of the unknown language on the basis of the information received so far. This process continues forever. The class of languages will be considered *learnable* with respect to the specified method of information presentation if there is an algorithm that the learner can use to make his guesses, the algorithm having the following property: Given any language of the class, there is some finite time after which the guesses will all be the same and they will be correct.

In this preliminary investigation, a *language* is taken to be a set of strings on some finite alphabet. The alphabet is the same for all languages of the class. Several variations of each of the following two basic methods of information presentation are investigated: A *text* for a language generates the strings of the language in any order such that every string of the language occurs at least once. An *informant* for a language tells whether a string is in the language, and chooses the strings in some order such that every string occurs at least once.

It was found that the class of context-sensitive languages is learnable from an informant, but that not even the class of regular languages is learnable from a text.

E. Schang (LLL)

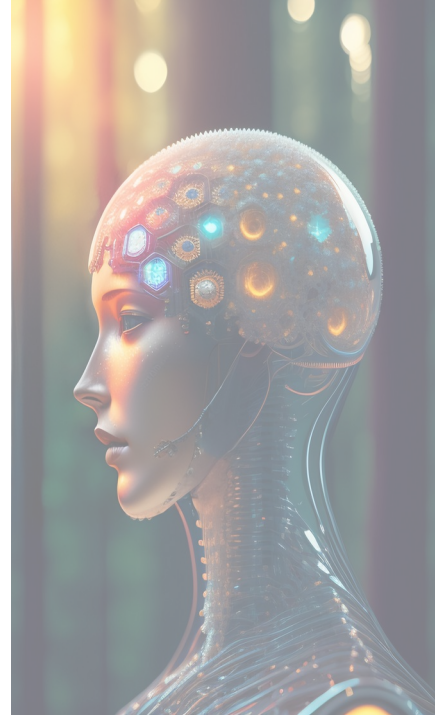


LLMs et linguistique formelle

Carnie 2002 :13

« Many parts of Language are built in , or **innate**. Much of Language is an ability hard-wired into our brains by our genes.

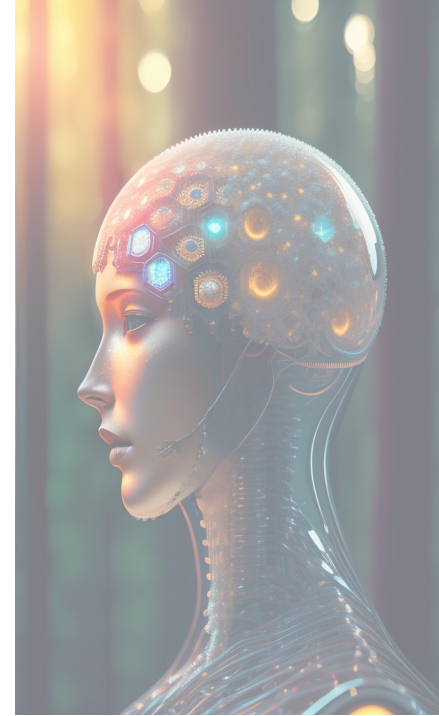
(...) Infinite systems are both unlearnable and unacquirable. Since we all have such an infinite system in our heads, and we can't have learned it, it must be the case that it is built in. »



Défaite de la syntaxe formelle



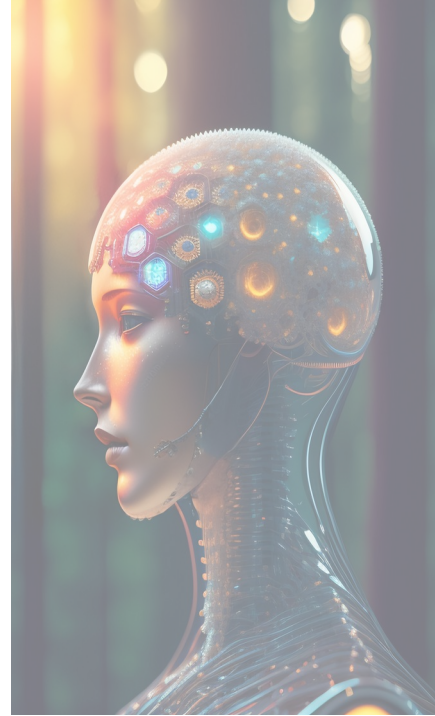
E. Schong (LLL)



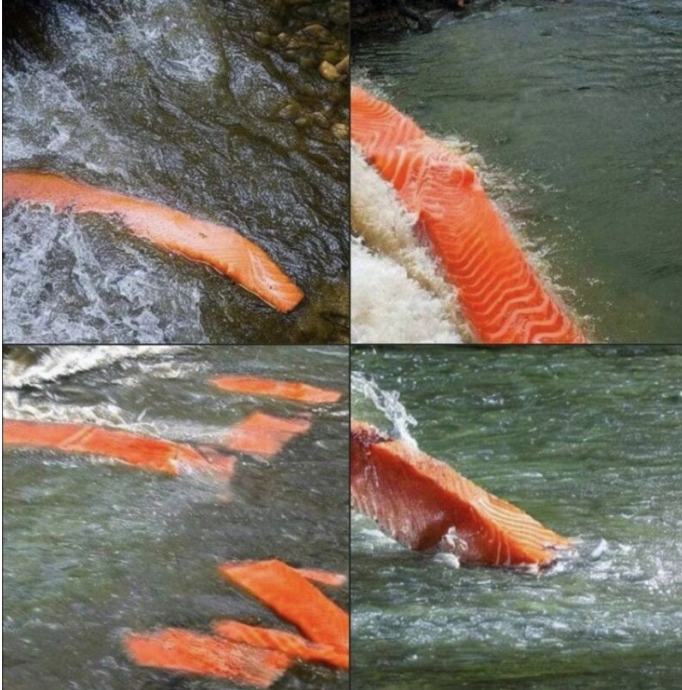
Steven T. Piantadosi (UC Berkeley) @ILFC

- Abstract: A major target of linguistics and cognitive science is to understand what class of learning systems can acquire the key structures of natural language. Until recently, the computational requirements of language have been used to argue that learning is impossible without a highly constrained hypothesis space. Here, we **describe a learning system** that is maximally unconstrained, operating over the space of all computations, and is **able to acquire several of the key structures** present **natural language** from **positive evidence alone**. We demonstrate this by providing the same learning model with data from 70 distinct formal languages which have been argued to capture key features of language, have been studied in experimental work, or come from an interesting complexity class. The model is able to successfully induce the latent system generating the observed strings from positive evidence in almost all cases, including regular, context-free, and context-sensitive formal languages, as well as languages studied in artificial language learning experiments. These results show that **relatively small amounts of positive evidence** can support learning of rich classes of generative computations over structures.

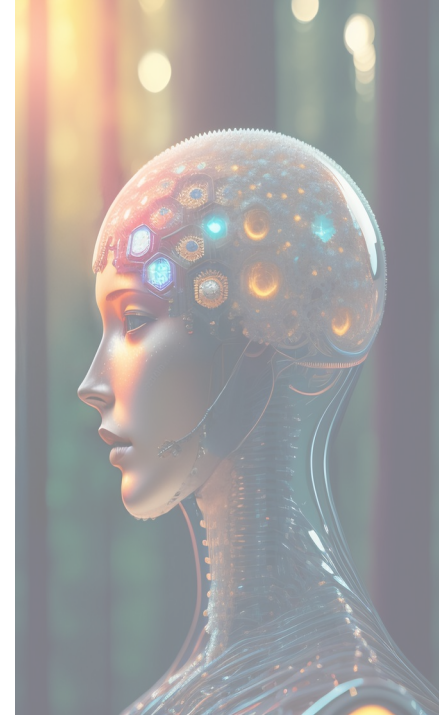
E. Schang (LLL)



OK, mais on se détend...



Tiré d'un post d'Andrea
Bensaid sur LinkedIn



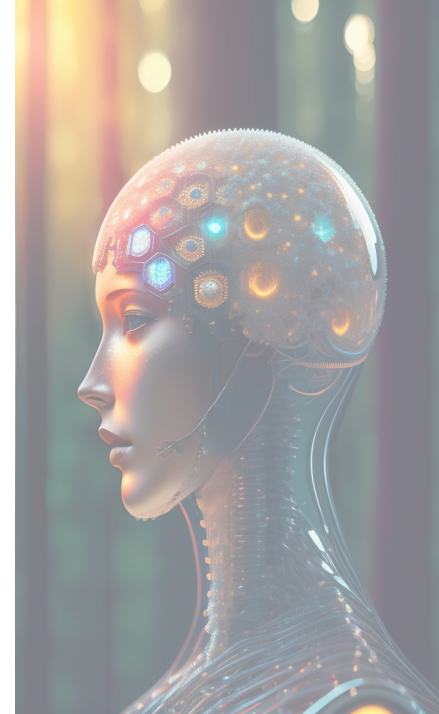
NLP is not NLU

Manipuler correctement des séquences de mots ne présuppose pas la compréhension de ces séquences.

- On trouve chez Devlin&al.2018 l'affirmation suivante :

« In order to train a model that **understands** sentence relationships, we pre-train for a binarized next sentence prediction task. »

- Que dit GPT-3 d'Hamlet ?



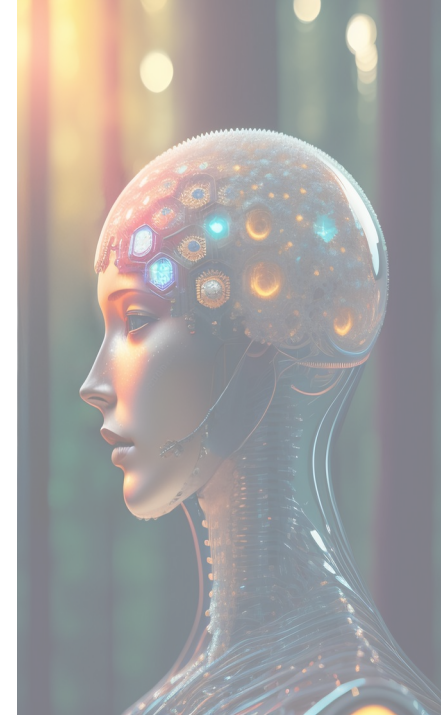
L'article 'Hamlet' sur Wikipédia revu par GPT-3

Prompt : *The protagonist of Hamlet is Prince Hamlet of Denmark, son of the recently deceased King Hamlet, and nephew of King Claudius, his father's brother and successor.*

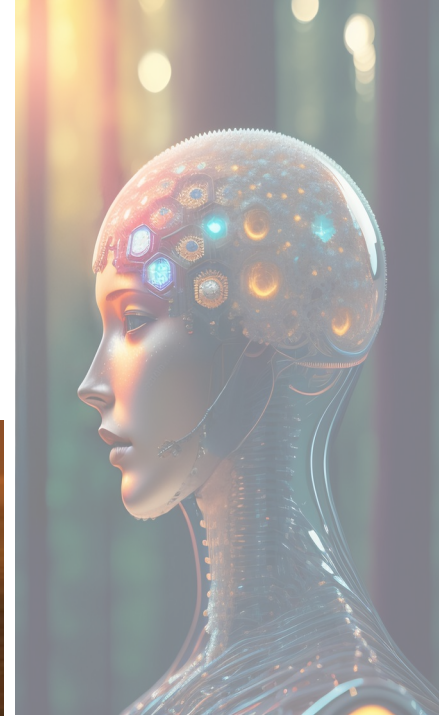
« Hamlet, a 19th-century writer, is best known for his short stories, such as *The Decline and Fall of the Roman Empire* and *The Rise and Fall of the Roman Empire*. He also wrote books such as *The Diary of a Young Woman* and *The Life and Times of a Young Man*.

Hamlet was born in the St. Petersburg suburb of Youngstown, Ohio, on October 13, 1844. He was educated at the University of Pennsylvania, where he was a student until his death in 1896.

Hamlet's father, a former soldier, was a leading figure in the European War of Independence, the struggle for independence in Germany during World War I, and the war between England and France during World War II. Hamlet was... » [GPT-3, <https://demo.allennlp.org/>]



Deep Octopus et Perroquet Stochastique



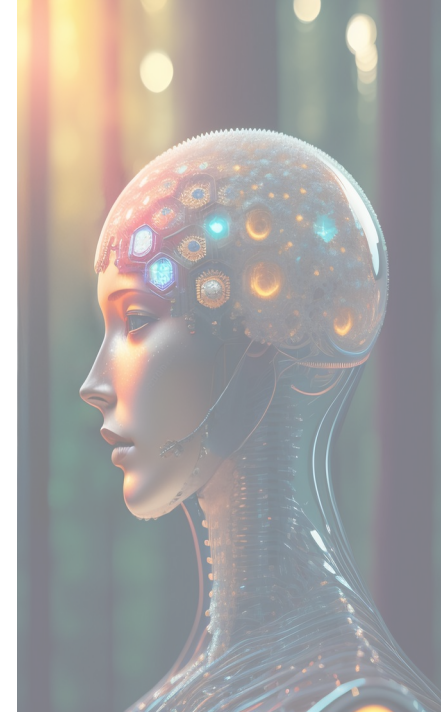
Stochastic Parrots

Emily M. Bender*
ebender@uw.edu
University of Washington
Seattle, WA, USA

Angelina McMillan-Major
aymm@uw.edu
University of Washington
Seattle, WA, USA

Timnit Gebru*
timnit@blackinai.org
Black in AI
Palo Alto, CA, USA

Shmargaret Shmitchell
shmargaret.shmitchell@gmail.com
The Aether

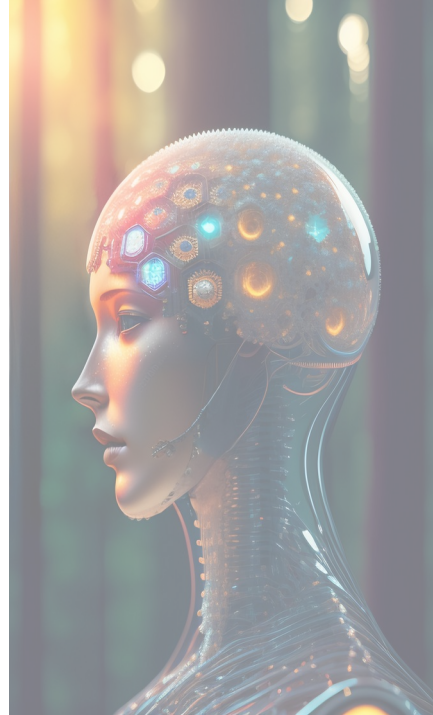


Stochastic Parrots

ABSTRACT

The past 3 years of work in NLP have been characterized by the development and deployment of ever larger language models, especially for English. BERT, its variants, GPT-2/3, and others, most recently Switch-C, have pushed the boundaries of the possible both through architectural innovations and through sheer size. Using these pretrained models and the methodology of fine-tuning them for specific tasks, researchers have extended the state of the art on a wide array of tasks as measured by leaderboards on specific benchmarks for English. In this paper, we take a step back and ask: How big is too big? What are the possible risks associated with this technology and what paths are available for mitigating those risks? We provide recommendations including weighing the environmental and financial costs first, investing resources into curating and carefully documenting datasets rather than ingesting everything on the web, carrying out pre-development exercises evaluating how the planned approach fits into research and development goals and supports stakeholder values, and encouraging research directions beyond ever larger language models.

E. Schang (LLL)



LeCun et l'AGI

A Path Towards Autonomous Machine Intelligence

Version 0.9.2, 2022-06-27

Yann LeCun

Courant Institute of Mathematical Sciences, New York University yann@cs.nyu.edu

Meta - Fundamental AI Research yann@fb.com

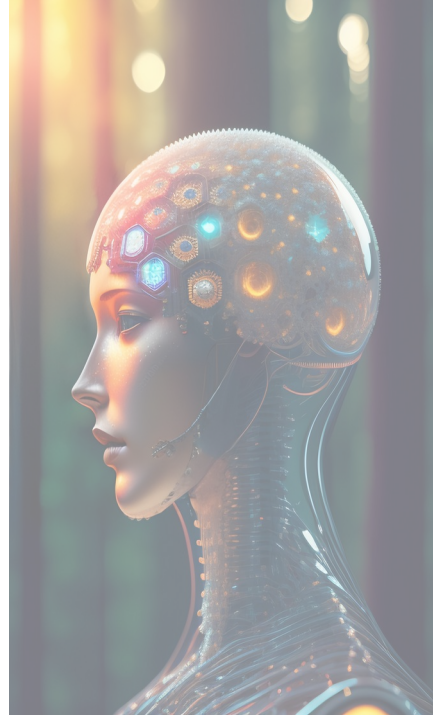
June 27, 2022

Abstract

How could machines learn as efficiently as humans and animals? How could machines learn to reason and plan? How could machines learn representations of percepts and action plans at multiple levels of abstraction, enabling them to reason, predict, and plan at multiple time horizons? This position paper proposes an architecture and training paradigms with which to construct autonomous intelligent agents. It combines concepts such as configurable predictive world model, behavior driven through intrinsic motivation, and hierarchical joint embedding architectures trained with self-supervised learning.

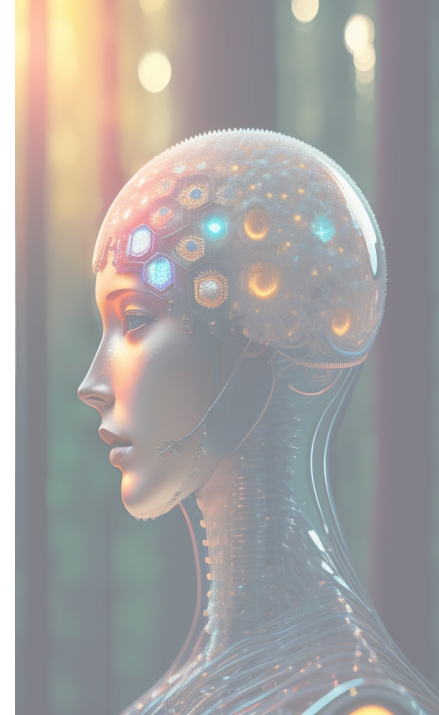
Keywords: Artificial Intelligence, Machine Common Sense, Cognitive Architecture, Deep Learning, Self-Supervised Learning, Energy-Based Model, World Models, Joint Embedding Architecture, Intrinsic Motivation.

E. Schang (LLL)



General AI

- Les LLM orientent la recherche en IA du côté de l'accumulation de données. Plus de données et plus de calcul.
- Prémisse implicite :
« Knowing is more important than thinking »

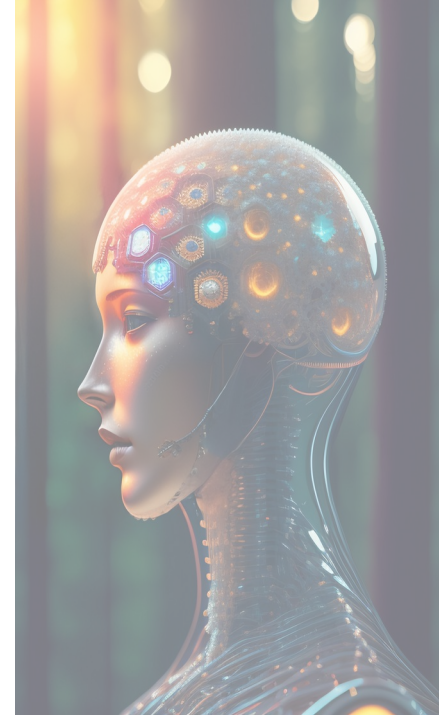


General AI

- Mais d'autres approches sont possibles et incorporent des connaissances issues du fonctionnement de l'esprit humain.

Qu'est-ce qu'on veut ?

- Une IA utile ou une IA humaine ?

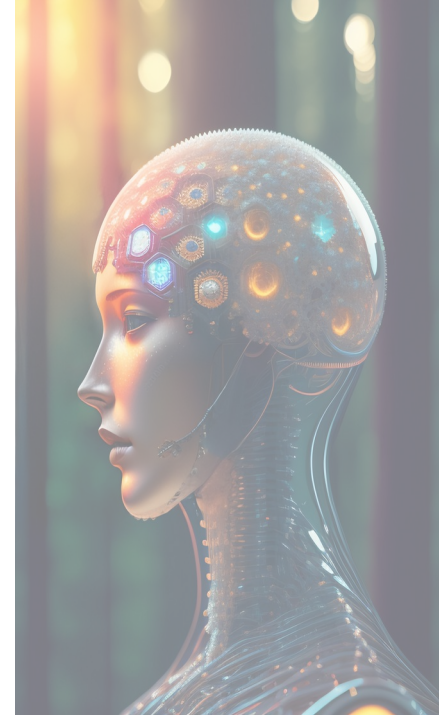


General AI

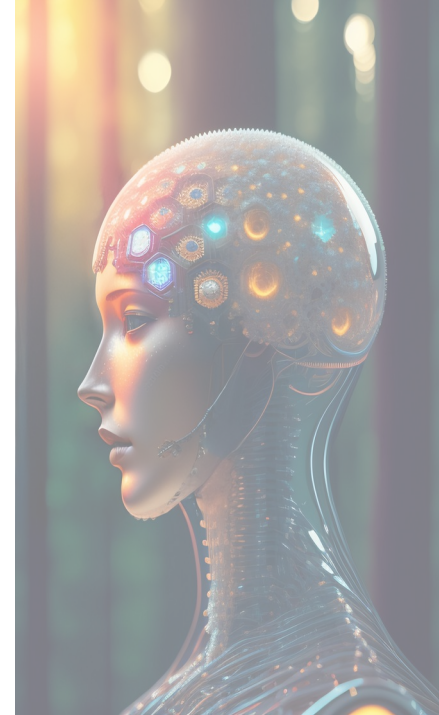
- Plein de questions pour les linguistes :
 - Quelle est la place de la langue dans l'intelligence ?
 - Existe-t-il une seule intelligence ?
 - Les humains sont-ils uniques (intelligence animale et mécanique) ?



E. Schang (LLL)



On a de quoi discuter, non ?



E. Schang (LLL)