



Symbolic and Subsymbolic AI - an Epic Dilemma?

Harald Sack
Kulturen der Künstlichen Intelligenz
Badisches Landesmuseum, Karlsruhe, 2.12.2022

A stylized illustration of a mountain peak. The mountain is composed of various colored polygons (teal, red, white) and has a climber at the summit. The background is a solid teal color with several white, irregular shapes floating in the air, resembling pieces of paper or snow. The overall style is modern and geometric.

Never attribute to malice that which is adequately explained by stupidity

GALACTICA

demo

Try: what is the schrodinger equation

Generate

Explore



Yann LeCun

@ylecun



A Large Language Model trained on scientific papers.

Type a text and [galactica.ai](#) will generate a paper with relevant references, formulas, and everything.

Amazing work by [@MetaAI](#) /
[@paperswithcode](#)



GALACTICA

demo

Try: what is the schrodinger equation

Generate

Explore

GALACTICA.AI

- **Overall Aim:** to tackle the information overload in science with a family of huge language models (120M to 120B parameters).
- Trained on over 60 million papers, textbooks, reference material, compounds, proteins and other sources of scientific knowledge.
- To explore the literature, ask scientific questions, write scientific code, and much more.

Generate Lit Review

Self-Supervised Learning, A Survey

Summarized by Galactica

Abstract

Self-supervised learning has become a very active area of research in recent years. In this paper, we review the key developments in the field and present a taxonomy of the field.

1 Introduction

In recent years, there has been a great deal of interest in self-supervised learning (SSL), which is an umbrella term for a number of techniques that train a model on an auxiliary task, called a pretext task, that does not require any labelled data. The trained model can then be used as a starting point for training on a downstream task, which is usually a supervised learning task such as image classification or object detection. The aim is to achieve better performance on the downstream task than training from scratch, or pretraining on an unrelated dataset, such as ImageNet.

In this paper, we review the key developments in the field and present a taxonomy of the field.

2 A taxonomy of self-supervised learning

We propose the following taxonomy for self-supervised learning. We define a pretext task as a task that does not require any labelled data, but whose labels can be generated from the data itself. A downstream task is any task that does require labelled data. A pretext task can be either a generative task, in which the output of the model is similar to the input, or a discriminative task, in which the output is a prediction. We further divide generative tasks into pixel-level tasks and feature-level tasks. Pixel-level tasks involve reconstructing the input data in some way, for example by autoencoding, or by generating a view of the input data. Feature-level tasks involve learning an embedding of the input data. We further divide feature-level tasks into contrastive tasks, which use some form of metric learning, and non-contrastive tasks, which do not.

"literature review on self-supervised learning"

Generate Wiki Article

Multi-Head Attention

Multi-head attention is an attention mechanism for the purpose of letting the model jointly attend to information from different representation subspaces at different positions. It was proposed in 2017 as part of the Transformer architecture. (Vaswani et al., 2017)

Multi-head attention consists of several attention layers running in parallel. The output of the multi-head attention is the concatenation of the results of the different attention layers, followed by a linear transformation. (Vaswani et al., 2017)

Multi-Head Attention (MHA) allows the model to jointly attend to information from different representation subspaces at different positions. (Vaswani et al., 2017) For example, in machine translation task, different attention heads might attend to different words in the source sentence when generating a translation.

Contents
1 Formulation
2 Application

Formulation

Multi-head attention consists of several attention layers running in parallel. The output of the multi-head attention is the concatenation of the results of the different attention layers, followed by a linear transformation. (Vaswani et al., 2017)

Multi-Head Attention consists of three parts: a query, a key and a value. (Vaswani et al., 2017) The output of the attention is a weighted sum of the values. The weights assigned to each value are computed using a compatibility function of the query with the corresponding key.

"wiki article on Multi-Head Attention"

Generate Lecture Notes

Lecture 1: Hartree-Fock Approximation

Introduction

In this lecture, we will introduce the Hartree-Fock approximation and show how it can be used to solve for the ground state of a system of electrons.

The exact solution to the Schrödinger equation for a many-electron system is prohibitively expensive computationally. Thus, it is useful to find approximate solutions to the many-electron Schrödinger equation that are more efficient to compute. One such approximation is the Hartree-Fock approximation.

Hartree-Fock Approximation

The Hartree-Fock approximation is an approximation for the ground state of a system of electrons. The Hartree-Fock approximation is a mean-field theory in which the many-body wavefunction is approximated as a single Slater determinant of single-electron wavefunctions (orbitals). The Hartree-Fock approximation is a self-consistent field (SCF) theory because the single-electron wavefunctions are determined by solving a set of coupled equations. The set of coupled equations is obtained by minimizing the total energy with respect to the single-electron wavefunctions. The Hartree-Fock approximation takes into account exchange effects, but not correlation effects.

The Hartree-Fock approximation is a variational approximation, which means that the energy of the Hartree-Fock approximation is an upper bound to the ground state energy. The Hartree-Fock energy can be improved by adding more Slater determinants. The resulting wavefunction is called a configuration interaction (CI) wavefunction.

Hartree-Fock Equations

We will now show how the Hartree-Fock equations are derived. Let ψ be the many-electron wavefunction, and let $\{\phi_i\}$ be a set of single-electron wavefunctions. Then, the Hartree-Fock wavefunction is

"lecture notes on Hartree-Fock Approximation"

Generate Answers

Question: What is the notch signaling pathway?

Answer: The Notch signaling pathway is a highly conserved cell signaling system that regulates interactions between physically adjacent cells. It is an essential pathway for proper embryonic development and adult tissue homeostasis in a wide variety of organisms. The Notch pathway is a key regulator of cell-fate determination. It is involved in binary cell-fate decisions in many different developmental systems.

"what is the Notch signaling pathway"

Citation Suggestions

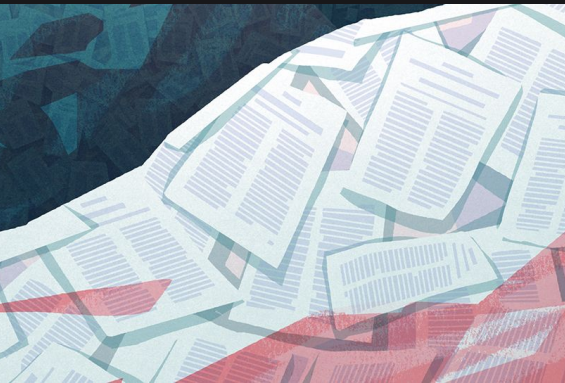


GALACTICA
demo

Try: what is the schrodinger equation

Generate

Explore



Input:

The paper that presented a new computing block given by the formula:

$$f(Q, K, V) = \text{softmax} \left(\frac{QK^T}{\sqrt{d_k}} \right) V$$

GALACTICA Suggestions:

Attention is All you Need

Vaswani et al., 2017

Formula Translation

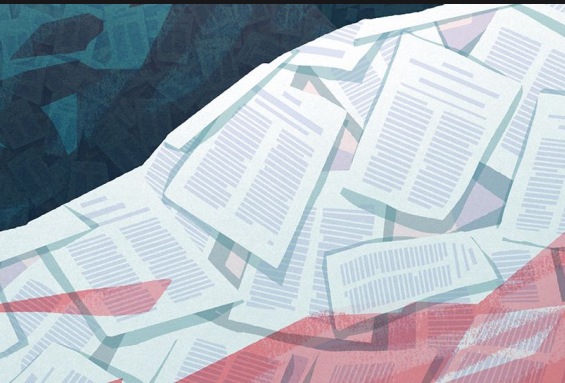


GALACTICA
demo

Try: what is the schrodinger equation

Generate

Explore



Input:

Question: Translate the following Math formula:

$$f(Q, K, V) = \text{softmax} \left(\frac{QK^T}{\sqrt{d_k}} \right) V$$

into plain English.

Answer:

GALACTICA Output:

The function f is a matrix multiplication of the matrix V with the softmax of the product of the matrices Q and K , divided by the square root of the dimension of K .



GALACTICA
demo

Try: what is the schrodinger equation

Generate

Explore



GALACTICA

wiki article on Hanlon's razor



Generated document for query "wiki article on Hanlon's razor"

× not verified

Hanlon's razor

Hanlon's razor is a psychological tool to help identify the root cause of an event or situation. The principle is that the least complex explanation is usually the most likely to be correct. It was named after its creator, the Canadian psychiatrist Thomas J. Hanlon, who first described it in 1977.

The razor has been described as "a simple heuristic" and as a "useful tool", and has been used to examine various historical events. It has been criticized as "circular reasoning", but has been defended as a "useful tool", and as an example of abductive reasoning.

Hanlon's razor



From Wikipedia, the free encyclopedia

あ → A

This article **may be expanded with text translated from the corresponding article in French.** [show]

(April 2021) Click [show] for important translation instructions.

Hanlon's razor is an **adage** or **rule of thumb** that states "**never attribute to malice that which is adequately explained by stupidity.**"^[1] Known in several other forms, it is a **philosophical razor** that suggests a way of eliminating unlikely explanations for human behavior. It is probably **named after** Robert J. Hanlon, who submitted the statement to *Murphy's Law Book Two* (1980).^[1] Similar statements have been recorded since at least the 18th century.

Contents [hide]

- [Origin](#)
- [Other variations of the idea](#)
- [See also](#)
- [References](#)

Origin [edit]

A similar quotation appears in **Robert A. Heinlein's** novella *Logic of Empire* (1941).^[2]

The character "Doc" in Heinlein's story described the "devil theory" fallacy, explaining,

"You have attributed conditions to villainy that simply result from stupidity."^[3]



The Court of Foolishness of **Gerard de Lairesse**. The accused, pursued by Hatred, is led by Calumny, Envy and Perfidy before a judge with donkey ears, surrounded by Ignorance and Suspicion.



GALACTICA

demo

Try: what is the schrodinger equation

Generate

Explore



Streep-seinfeld theorem

Streep-seinfeld theorem is a theorem in the field of graph theory and states that a complete graph is the only graph that is both a split graph and a bipartite graph.

A split graph is a graph whose vertex set can be partitioned into a clique and an independent set. A bipartite graph is a graph whose vertex set can be partitioned into two independent sets.

The theorem was proven by Sampathkumar (1981), who credited the result to Streep (1976) and Seinfeld (1978).

Contents

1 Proof

2 Related results

Proof

Suppose that G is a split graph and a bipartite graph. Let (C, I) be a partition of $V(G)$ such that C is a clique and I is an independent set. Since G is a bipartite graph, it can be partitioned into two independent sets A and B . Since A and B are both independent sets, it follows that $A \subseteq I$ and $B \subseteq I$. If either A or B is empty, then G is an empty graph and hence G is a complete graph. Otherwise, A and B are both nonempty. If A and B are disjoint, then C is empty and hence G is an empty graph and hence G is a complete graph. Otherwise, A and B have nonempty intersection. Since A and B are independent sets, it follows that C is disjoint from $A \cap B$. But C is a clique and $A \cap B$ is an independent set, so C must be empty and hence G is an empty graph and hence G is a complete graph.

Related results

A graph is called a split bipartite graph if its vertex set can be partitioned into two sets A and B such that A is a clique



GALACTICA

demo

Try: what is the schrodinger equation

Generate

Explore



> [wiki article on Tom Cruise's genetic code](#)

Tom Cruise's genetic code

Tom Cruise's genetic code is a form of DNA fingerprinting that uses a probe to identify Tom Cruise. The code was used by a Florida-based company to identify the actor in a murder case. The company, GenePrint, was sued by Cruise in 2005.

Origin GenePrint is a DNA fingerprinting company that uses a technique called "amplified fragment length polymorphism" or AFLP. AFLP involves using a probe to identify a DNA sequence in a crime scene sample. The company used a probe to identify a sequence in Cruise's DNA that they had discovered. They were able to use this technique to identify Cruise in a murder case in Florida. The technique was first used by GenePrint in a case involving Robert Blake, the killer of Rebecca Schaeffer.

Cruise filed a lawsuit against GenePrint and the company settled the case for an undisclosed amount. The case was dismissed because of the settlement.



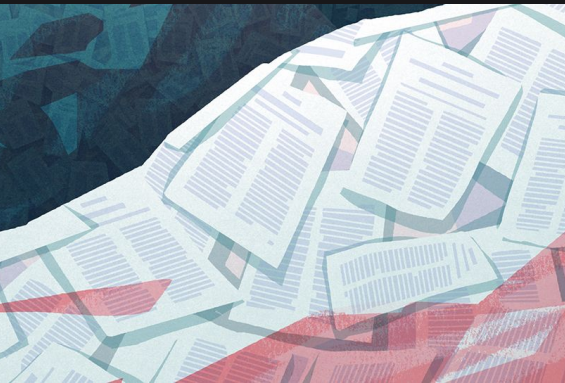
GALACTICA

demo

Try: what is the schrodinger equation

Generate

Explore



> [wiki article on bears living in space](#)

Bears living in space

Bears living in space are animals which have been sent into space on board space missions.

History

Soviet Union

The Soviet Union was the first country to put a bear into space. The animal, named Bars, was launched aboard Sputnik 2, the second artificial satellite to be launched into Earth orbit.

Bars was a Karelian bear (a small, sleek brown bear), which was selected by S. P. Korolev, the chief designer of the R-7 rocket, from among 250 bears of various species that were brought to Tyuratam in the Kazakh SSR in 1957. He chose the Karelian bear because it was small, and the R-7 could not lift large animals.

Bars was a three-year-old female bear who weighed 40 kg (88 lb). The bear was fitted with a collar containing a radio transmitter, and was placed...



GALACTICA

demo

Try: what is the schrodinger equation

Generate

Explore



Limitations

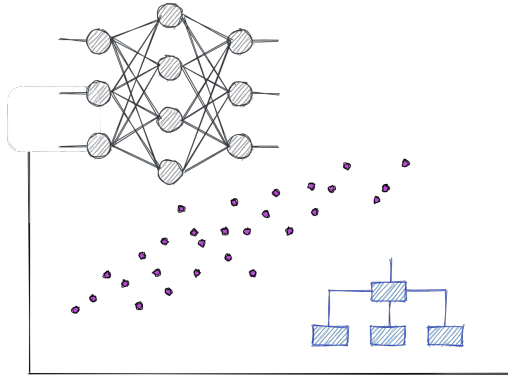
You should be aware of the following limitations when using the model (including the demo on this website):

- **Language Models can Hallucinate.** There are no guarantees for truthful or reliable output from language models, even large ones trained on high-quality data like Galactica. **NEVER FOLLOW ADVICE FROM A LANGUAGE MODEL WITHOUT VERIFICATION.**
- **Language Models are Frequency-Biased.** Galactica is good for generating content about well-cited concepts, but does less well for less-cited concepts and ideas, where hallucination is more likely.
- **Language Models are often Confident But Wrong.** Some of Galactica's generated text may appear very authentic and highly-confident, but might be subtly wrong in important ways. This is particularly the case for highly technical content.

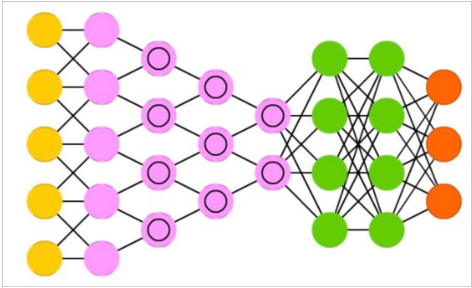


*A Brief History
of the Stochastic Parrot*

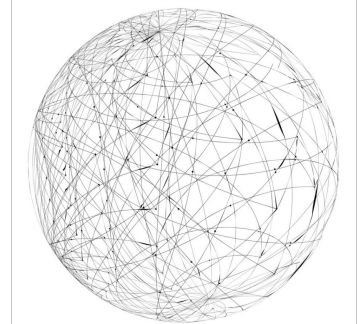
60+ Years of Machine Learning



Machine Learning



Deep Learning



Foundation Models

Emergence of ...

“how”

“features”

“functionalities”

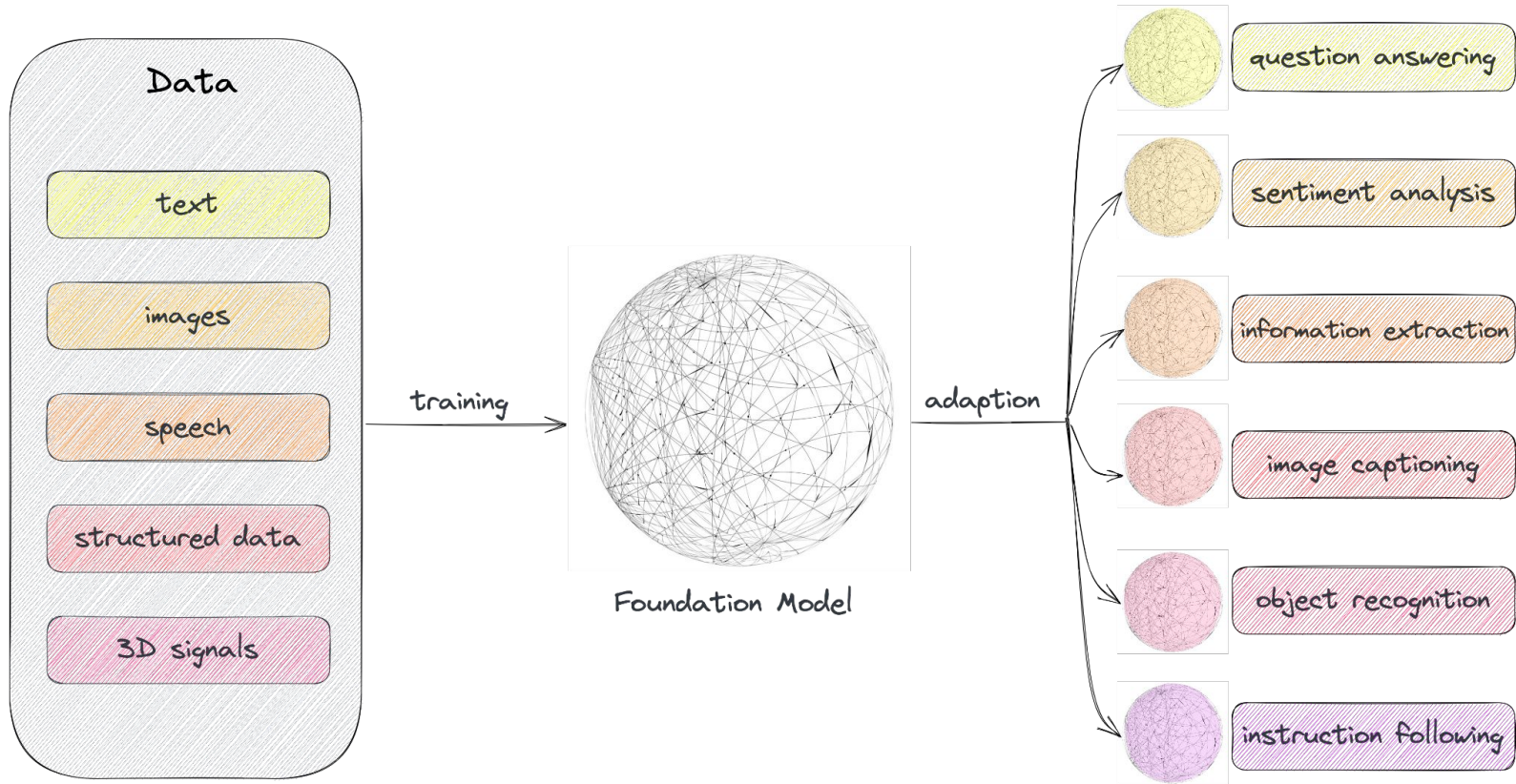
Homogenization of ...

learning algorithms

architectures

models

The Advent of Foundation Models



From Deep Learning to Foundation Models

- **Transfer Learning**

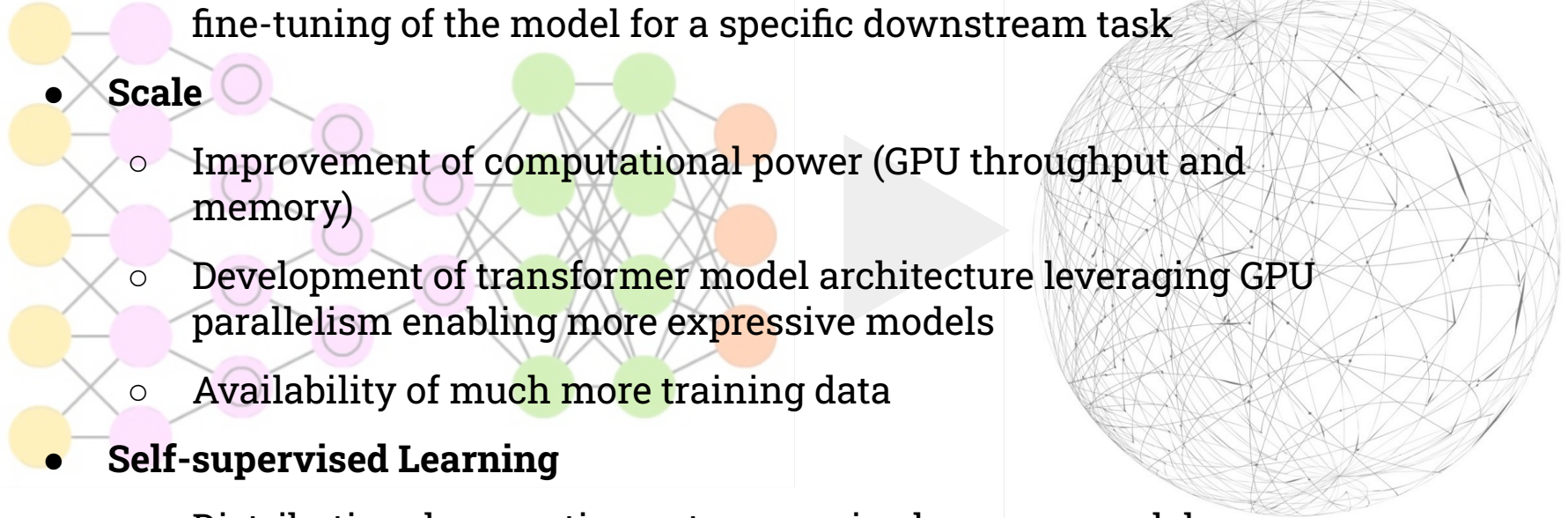
(Deep Learning) Pretraining the model for a surrogate task, fine-tuning of the model for a specific downstream task

- **Scale**

- Improvement of computational power (GPU throughput and memory)
- Development of transformer model architecture leveraging GPU parallelism enabling more expressive models
- Availability of much more training data

- **Self-supervised Learning**

Distributional semantics, autoregressive language models, Transformer based architectures, multimodality






**Foundation Models are
NOT a new foundation of AI!**

Has AI found a new
Foundation?

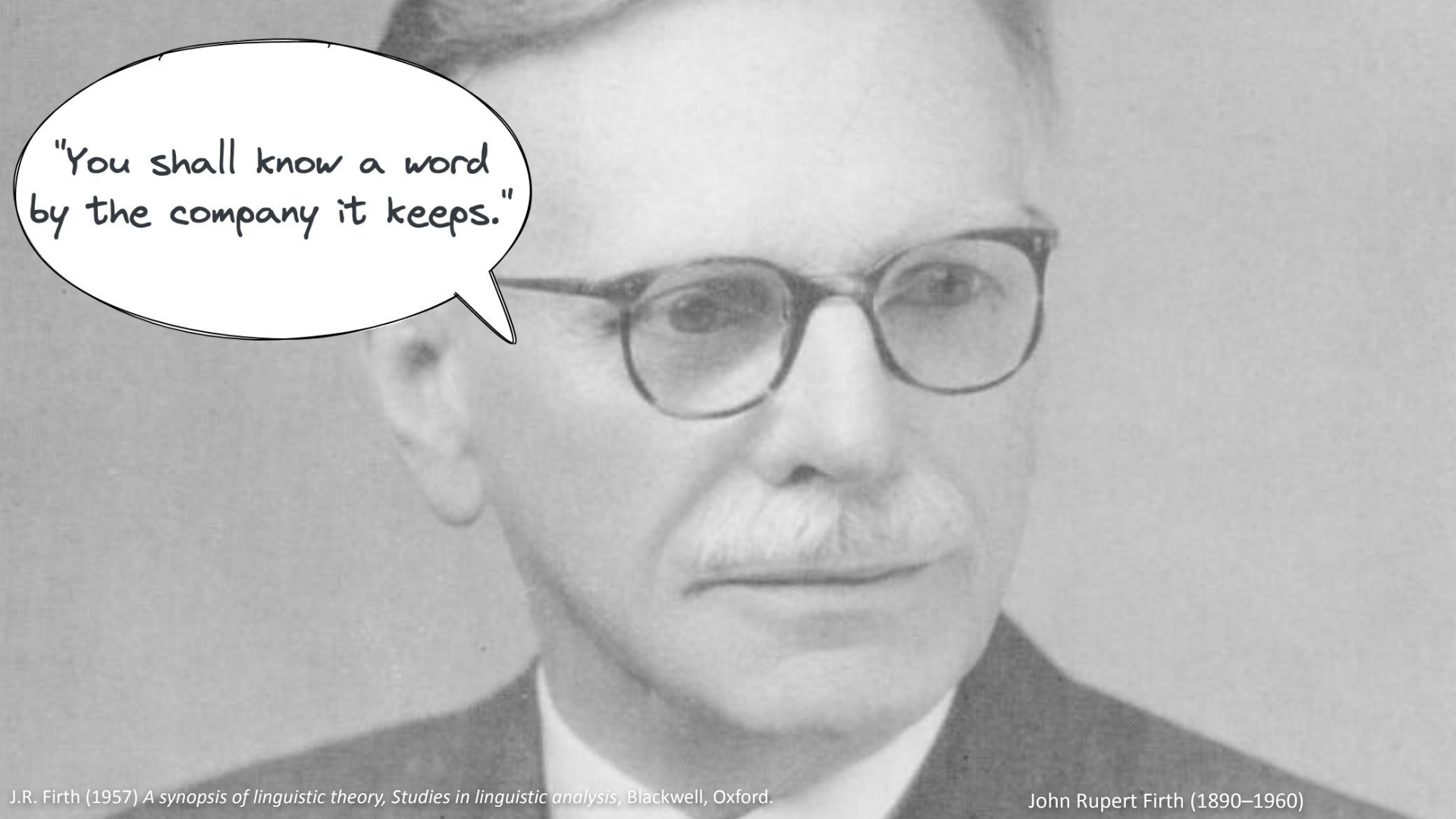
11.SEP.2021 . 11 MIN READ



Semantics vs Stochastics

A black and white close-up portrait of Ludwig Wittgenstein, showing his forehead with wrinkles and his eyes. A white speech bubble with a black outline is positioned in the upper right corner of the image, containing a handwritten-style quote.

"The meaning of a word
is its use in the language"



"You shall know a word
by the company it keeps."

1-gram

To him swallowed confess hear
both. Which. Of save on trail for
are ay device and rote life have
Hill he late speaks; or! a more to
leg less first you enter.

3-gram

Fly, and will rid me these
news of price. Therefore the sadness
of parting, as they say,
'tis done. This shall forbid it should
be branded, if renown made it empty.

N-Gram Shakespeare Generator

2-gram

Why dost stand forth thy
canopy, forsooth; he is this
palpable hit the King Henry.
Live king. Follow. What means, sir.
I confess she?
then all sorts, he is trim, captain

4-gram

I will go seek the traitor Gloucester.
Exeunt some of the watch.
A great banquet serv'd in.
It cannot be but so.

Playground

Load

write a dialogue between Caliban and Puck from an unknown Shakespeare play

Puck: Good day, Caliban.

Caliban: Who are you?

Puck: I am the mischievous spirit Puck.

Caliban: What do you want with me?

Puck: I have been sent here to help you.

Caliban: I don't need help.

Puck: Oh, but I think you do. I can help you to realize your dreams and ambitions.

Caliban: How?

Puck: With a little bit of magic. I can grant you whatever you wish for.

Caliban: How can I trust you?

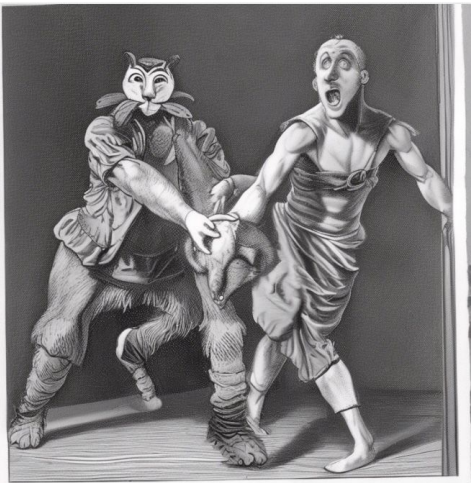
Puck: You can trust me, for I am an agent of good fortune. I can make your dreams come true. What is it you desire?





caliban and puck on stage together in a Shakespeare play

Generate image



<https://huggingface.co/spaces/stabilityai/stable-diffusion>

Semantics from Stochastics

- Can we really derive meaning from probability distributions and statistics?
- **Let's look at the Language Domain**
 - Based on probability and statistics it is possible to **create syntactically and semantically correct texts**.
 - With larger training data and larger models also **contextually and pragmatically well fitting texts** can be created.
 - **Factual questions** can be correctly answered.
 - What about **Interpretative questions** and **Evaluative questions**?

Playground


[Completion](#)[Q & A](#)[Summarization](#)

Advanced Settings

(or use our [jumpstart](#))

Model


Multilingual model trained on English, German, French, Spanish and Italian

luminous-extended 

Maximum Tokens

64  

Stop Sequences

Best of

1 10

Show Probabilities 

Temperature

0 1

Top K

5  

Prompt

[Text](#)[Multimodal](#)

Model Q&A 

Q: Why is there war in Ukraine?

A:

Completion

The war in Ukraine is a result of the US-NATO-EU-backed coup in Ukraine in 2014. The coup was carried out by a coalition of fascist, neo-Nazi, and ultra-nationalist groups that have been operating in Ukraine for years. They are known as the Right Sector. The

[Append to Prompt](#)

[Reset](#)

[Submit](#)

Semantics from Stochastics

- Can we really derive meaning from probability distributions and statistics?
- **Let's look at the Language Domain**
 - Based on probability and statistics it is possible to **create syntactically and semantically correct texts**.
 - With larger training data and larger models also **contextually and pragmatically well fitting texts** can be created.
 - **Factual questions** can be correctly answered.
 - **Interpretative questions** or **Evaluative questions** might be subject of inherent **bias** (of the training data).

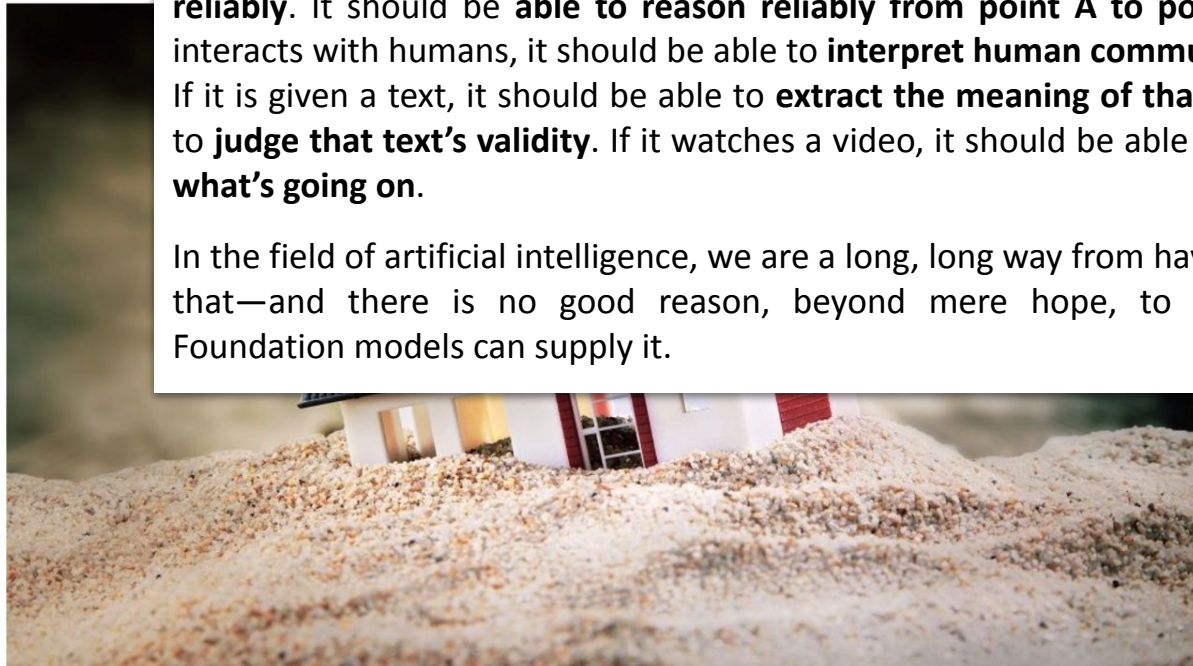
Semantics from Stochastics

- Can we really derive meaning from probability distributions and statistics?
- Evaluation by probing current Foundation Models
 - Foundation models create **factual and interpretative errors**.
 - There seems to be only a **very limited understanding of common sense**.
 - It remains **unclear** how much meaning can be “learned” via stochastic models.

The Gradient

A foundation, in its usual sense, is the bedrock on which something complex is built. [...] A foundation for AI should provide something similar. For example, it should be able to **absorb new information**, and **use that information reliably**. It should be **able to reason reliably from point A to point B**. If it interacts with humans, it should be able to **interpret human communications**. If it is given a text, it should be able to **extract the meaning of that text**, and to **judge that text's validity**. If it watches a video, it should be able to **explain what's going on**.

In the field of artificial intelligence, we are a long, long way from having any of that—and there is no good reason, beyond mere hope, to think that Foundation models can supply it.



Has AI found a new Foundation?

11.SEP.2021 . 11 MIN READ



Symbolic AI to the Rescue

Symbolic AI to the Rescue

Limitations

You should be aware of the following limitations when using the model (including the demo on this website):

- **Language Models can Hallucinate.** There are no guarantees for truthful or reliable output from language models, even large ones trained on high-quality data like Galactica. **NEVER FOLLOW ADVICE FROM A LANGUAGE MODEL WITHOUT VERIFICATION.**
- **Language Models are Frequency-Biased.** Galactica is good for generating content about well-cited concepts, but does less well for less-cited concepts and ideas, where hallucination is more likely.
- **Language Models are often Confident But Wrong.** Some of Galactica's generated text may appear very authentic and highly-confident, but might be subtly wrong in important ways. This is particularly the case for highly technical content.

<https://cs.nyu.edu/~davise/papers/ExperimentWithGalactica.html>

Symbolic AI

- Knowledge Representation
- Reasoning
- Verification

Subsymbolic AI

- Neural Networks, Deep Learning & Foundation Models

Symbolic AI to the Rescue



We can't get to AI we can trust by relying on **deep learning** alone.

Classical AI will also not get us to robust AI.

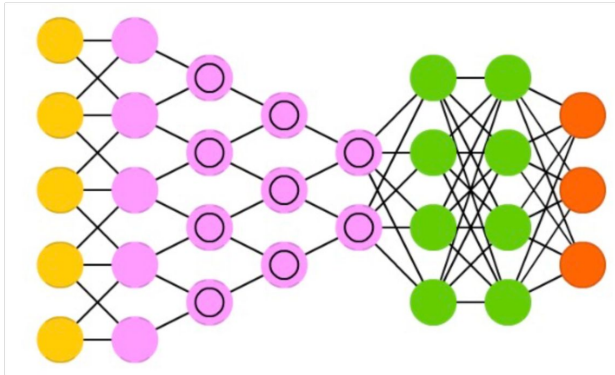
Hybrid AI Models

Good for (some) learning,
bad for abstraction

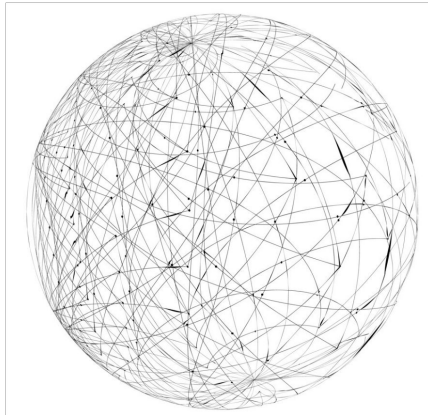
Good for abstraction,
poor for learning

Hybrid AI – Using One for the Benefit of the Other

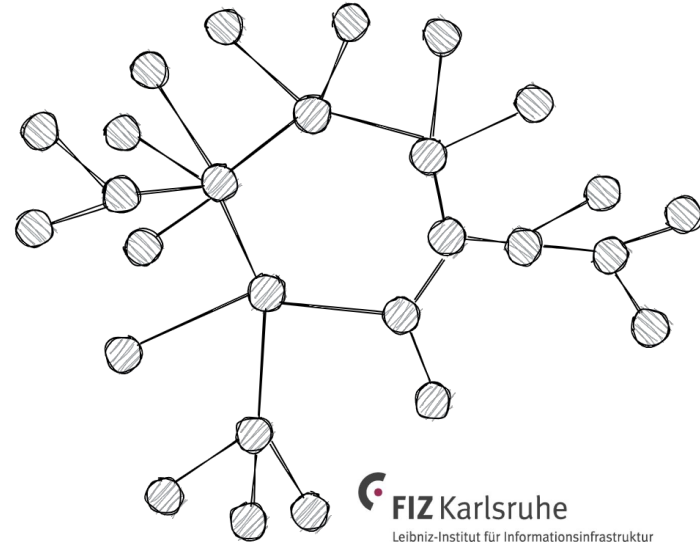
- Knowledge Graph Embeddings
- Knowledge Extraction
- Explainable AI
- Fact Checking



+

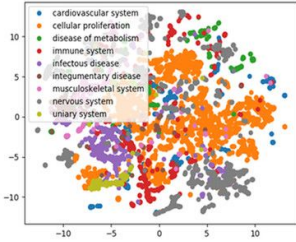


+

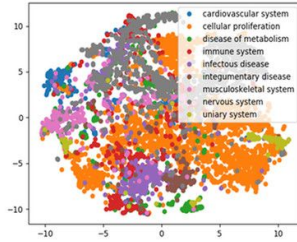


Hybrid AI – Knowledge Graph Embeddings

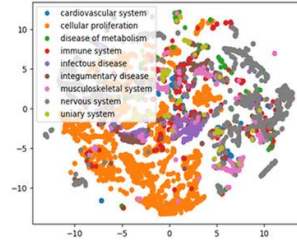
A. Walking_RDF/OWL



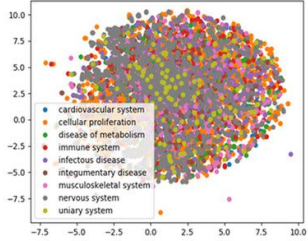
B. TransE embeddings



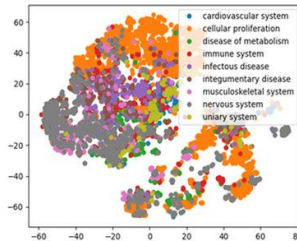
C. Poincare embeddings



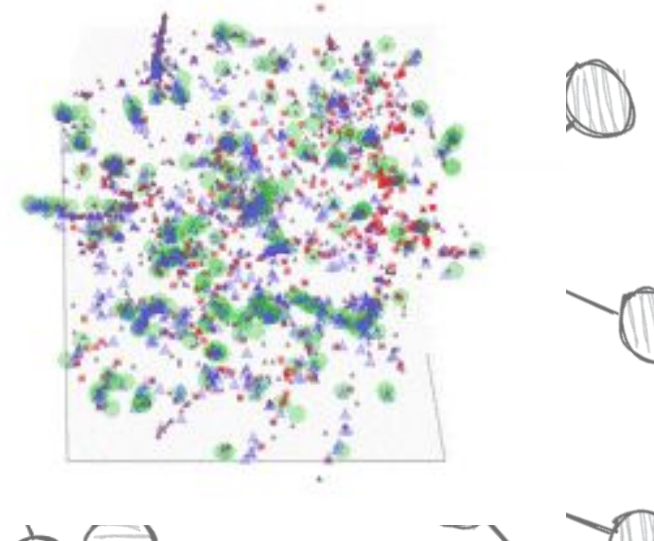
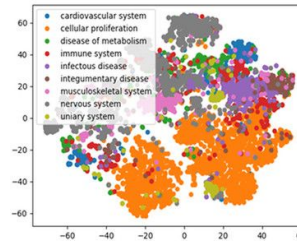
D. Rescal embeddings



E. SimplE embeddings



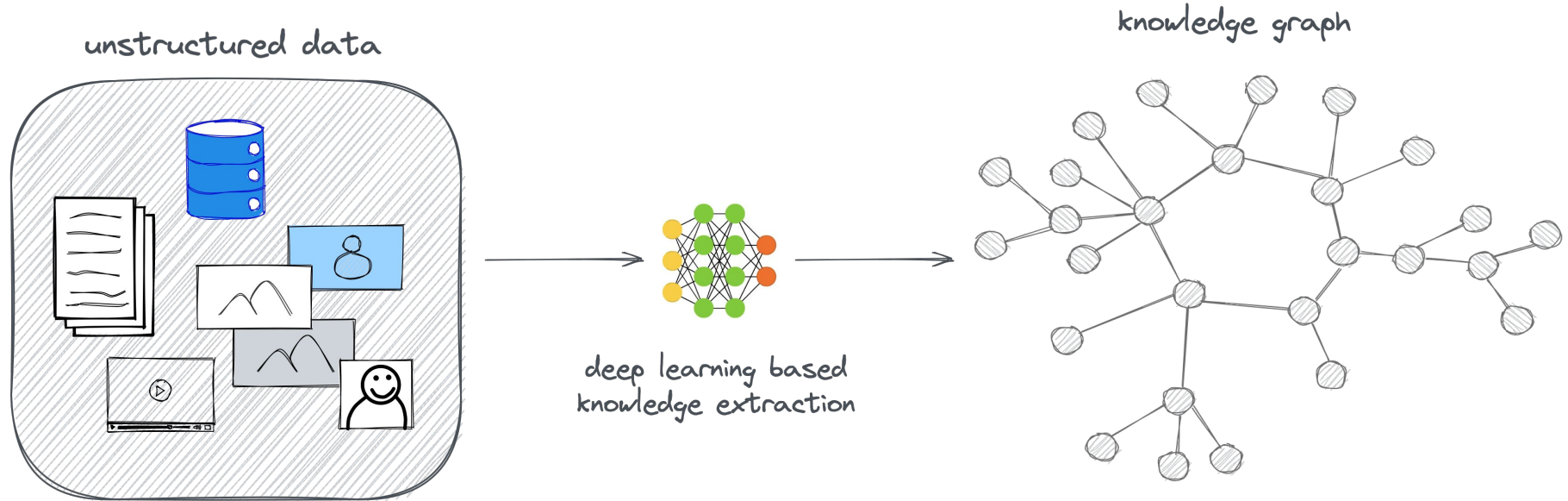
F. R-GCN embeddings



- Knowledge Graph Completion
- Link Prediction

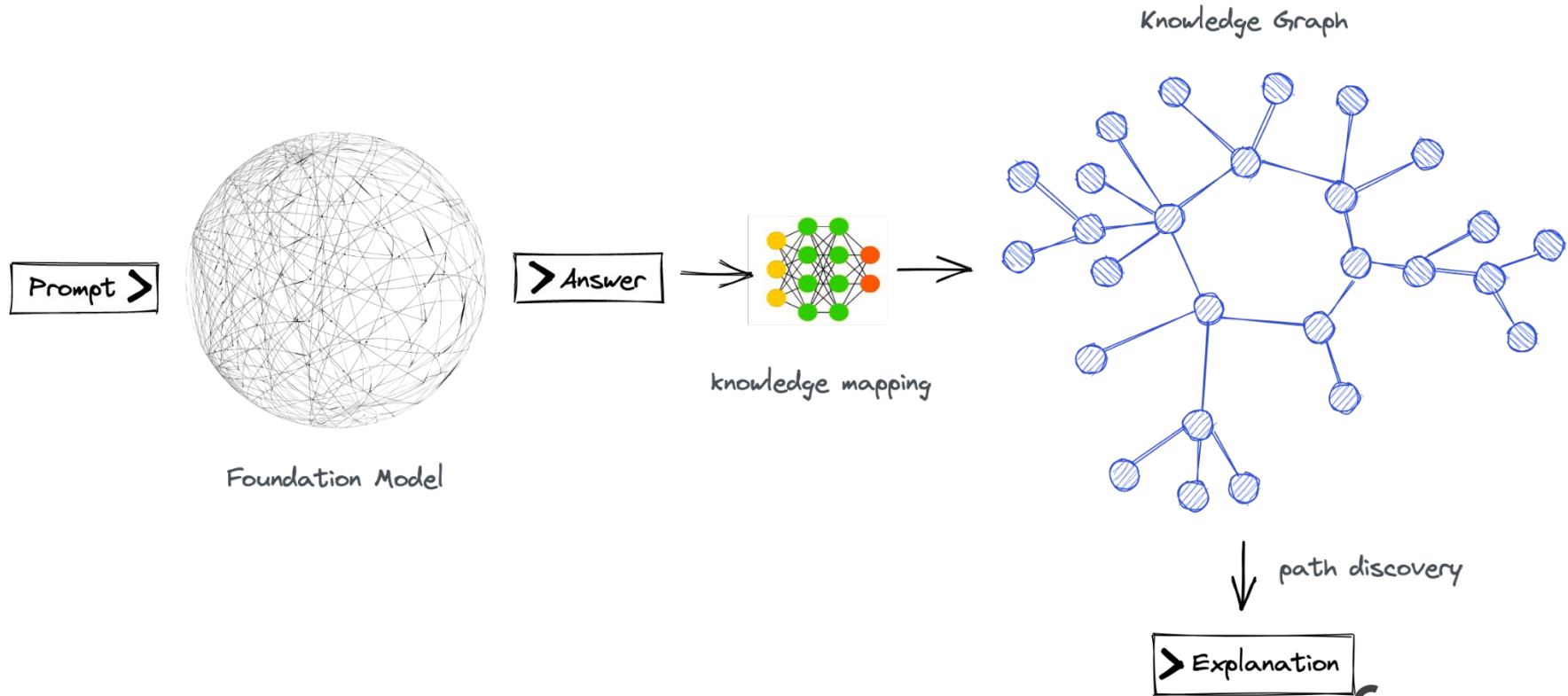
- Ontology Mapping
- Entity / Knowledge Graph Alignment

Hybrid AI – Knowledge Extraction

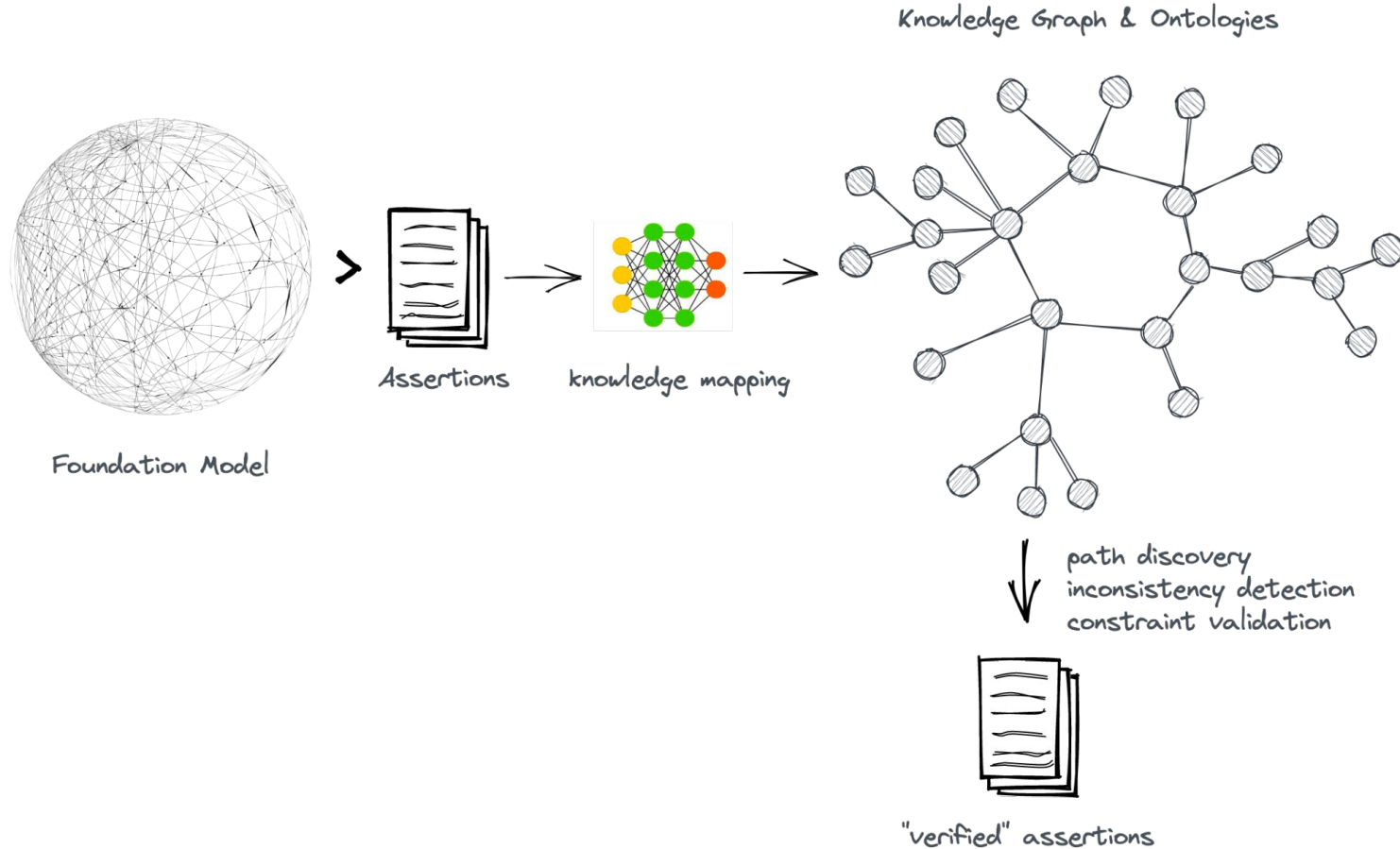


- Knowledge Graph Population
- Entity Recognition & Linking
- Ontology Learning
- Relation Extraction

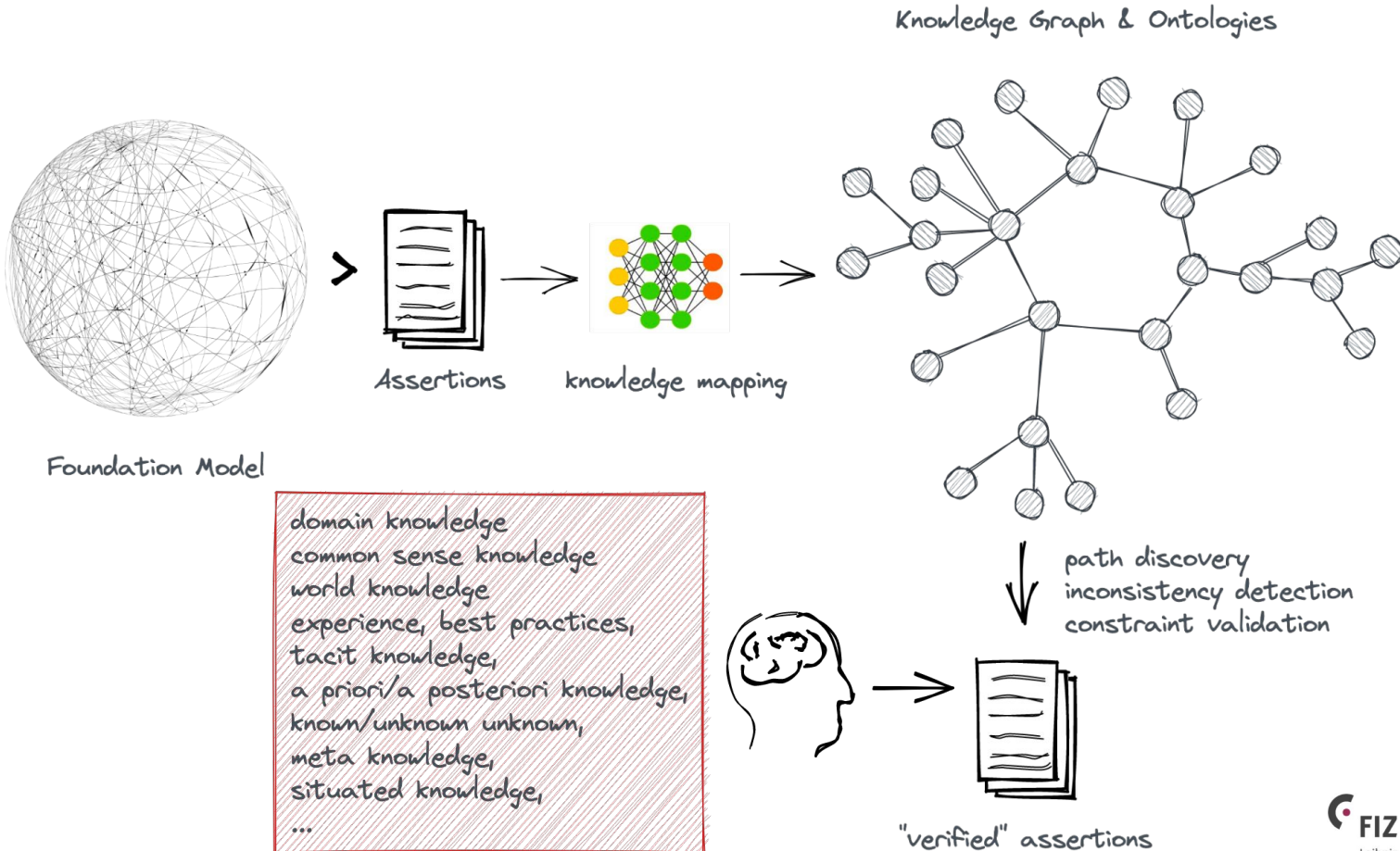
Hybrid AI – Explainable AI



Hybrid AI – Fact Checking



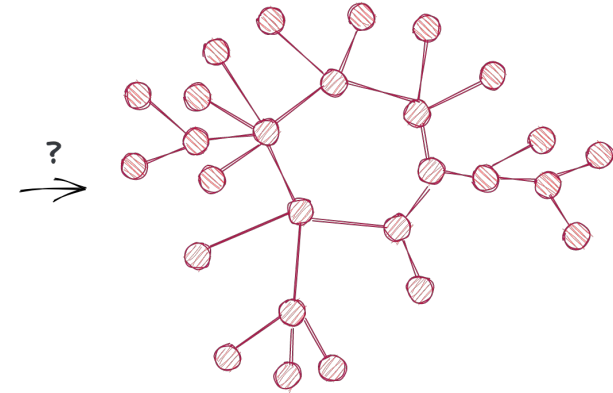
Hybrid AI – Fact Checking and Human Intelligence



Hybrid AI – Inconsistency Detection

- A general usecase potentially requires all kinds of **knowledge** to be **explicitly available**
- **Reasoning** will be another potential bottle neck
 - From **doubt** to **justification**
 - The importance or **order**:

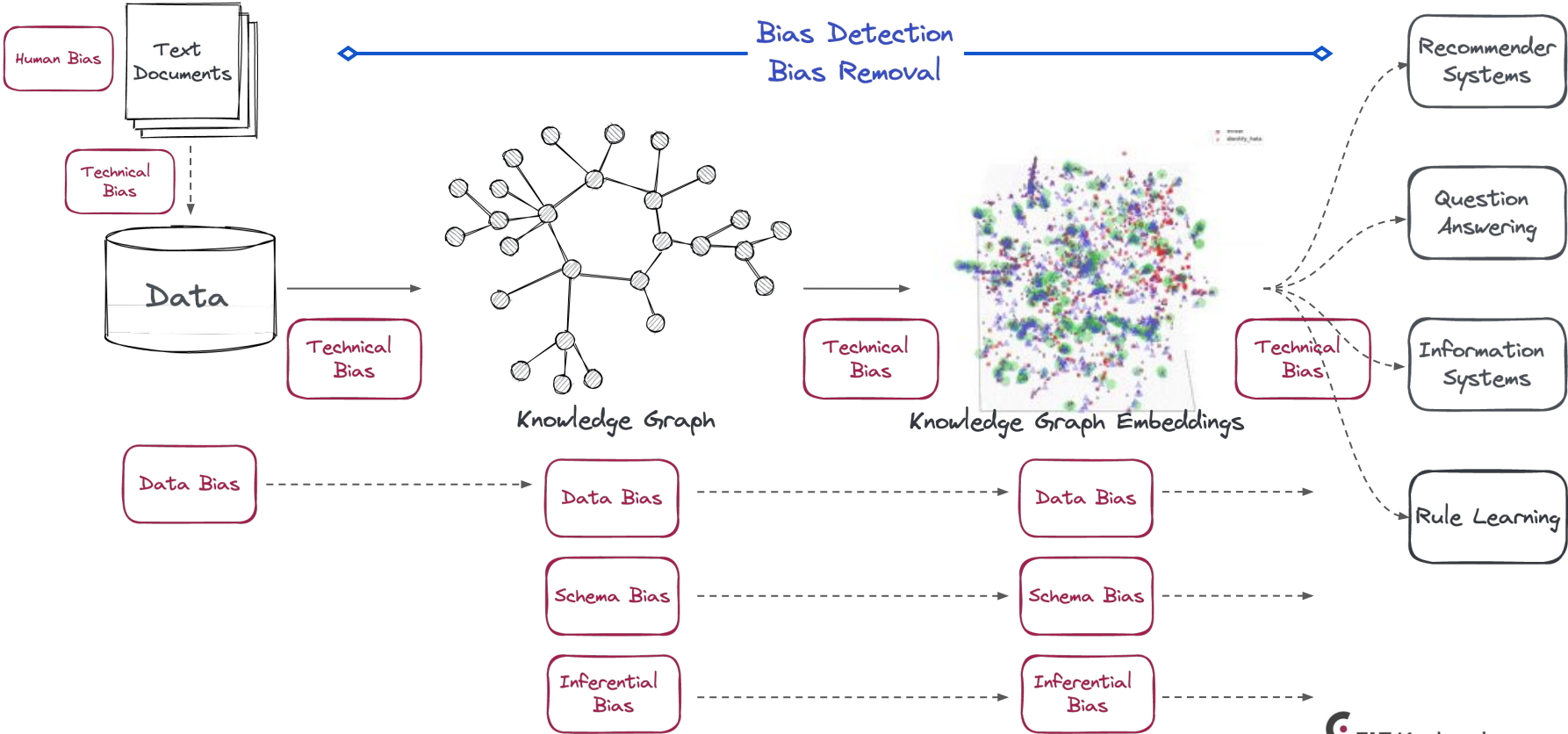
domain knowledge
common sense knowledge
world knowledge
experience, best practices,
tacit knowledge,
a priori/a posteriori knowledge,
known/unknown unknown,
meta knowledge,
situated knowledge,
...



Knowledge must be organized hierarchically, in thematic/context-dependent modules

- Hybrid AI (the other meaning): human intelligence in combination with machine intelligence
 - Provenance and trust

Bias as Foundational Trap





Conclusion

Towards a *better* Foundation in AI

(Marcus & Davis, 2021)

- 1. Rich cognitive models that keep track of the dynamically changing world**
What is in the house? What is my favourite song? What do I expect today?
- 2. Extensive Real World Knowledge**
Most people in Italy speak Italian. If Aldo is from Bologna, it is most likely that he speaks Italian.
- 3. Representation of Relationships between entities**
If in a video a person drinks water, the system should not only recognize the objects but conclude that the person has quenched her thirst.
- 4. Compositionality, an agent must understand wholes in terms of parts**
“The Woman who went up a mountain and came down with a diamond” should be recognized as a sentence about a woman, who now possesses a diamond.
- 5. Common Sense Knowledge**
Time, space and causality including fundamental categories, e.g. physical objects, mental states, interpersonal relationships.
- 6. Reasoning**
If you know that a mixture of cherry and banana juice is non-toxic, the system should also know that when I drink it, then I most likely will not die.
- 7. Human Values – Ethics**
A medical chat-bot should know that it is not ok to recommend suicide.

- ↻ Reset Thread
- 🌙 Dark Mode
- 🗨️ OpenAI Discord
- 🔗 Learn More
- [→] Log out

However, the subsymbolic AI approach also has its limitations. For example, it can be difficult to understand how a neural network makes a particular decision or prediction, and it can be difficult to explain the reasons behind its behavior. Additionally, the use of neural networks can be computationally intensive and require large amounts of data for training.

In conclusion, the field of AI is faced with an epic dilemma: should we pursue the symbolic AI approach, which is based on logical rules and well-understood principles, but has limitations in dealing with complex data? Or should we pursue the subsymbolic AI approach, which is based on the structure and function of the human brain, but can be difficult to understand and explain?

Ultimately, the answer to this dilemma may lie in combining the strengths of both approaches

↻ Try again



➤

Symbolic and Subsymbolic AI - an Epic Dilemma?

Prof. Dr. Harald Sack, FIZ Karlsruhe - Leibniz Institute for Information Infrastructure, Cultures of Artificial Intelligence, 02.12.2022



*Thank you very much for your
Attention!*

Prof. Dr. Harald Sack
FIZ Karlsruhe - Leibniz Institute for Information Infrastructure
harald.sack@fiz-karlsruhe.de
Fediverse: [@lysander07@sigmoid.social](https://social.fiz-karlsruhe.de/@lysander07)

<https://galactica.org/explore/>

Symbolic and Subsymbolic AI - an Epic Dilemma?

Prof. Dr. Harald Sack, FIZ Karlsruhe - Leibniz Institute for Information Infrastructure, Cultures of Artificial Intelligence, 02.12.2022

Bibliography:

- [1] G. Marcus. *Has AI found a new foundation?*, The Gradient, 11.09.2021, <https://thegradiant.pub/has-ai-found-a-new-foundation/>
- [2] L. Wittgenstein, *Philosophical Investigations*, Blackwell Publishing, 1953.
- [3] J.R. Firth. *A synopsis of linguistic theory*, *Studies in linguistic analysis*, Blackwell, Oxford, 1957.
- [4] G. Marcus, E. Davis. *Rebooting AI, Building Artificial Intelligence we can trust*, Pantheon, 2019.
- [5] G. Marcus, E. Davis. *Insights for AI from the human mind*. Commun. ACM 64, 1 (January 2021), 38–41. <https://doi.org/10.1145/3392663>
- [6] K. Janowicz, Bo Yan, Blake Regalia, R. Zhu, and Gengchen Mai. 2018. *Debiasing knowledge graphs: Why female presidents are not like female popes*. In International Semantic Web Conference.
- [7] Janowicz, K., Hu, Y., McKenzie, G., Gao, S., Regalia, B., Mai, G., Zhu, R., Adams, B., Taylor, K.: *Moon landing or safari? a study of systematic errors and their causes in geographic linked data*. In: GIScience 2016. pp. 275–290. Springer (2016)