

Challenges of LLMs in Geographic Information Management and Extraction

Prof. **Christophe Claramunt**
Naval Academy Research
Institute

Atelier Humanités
numériques spatialisées





Main principles behind **LLMs** & fundamentals

We address three complementary issues:

- (1) the ability of **LLMs** to extract and structure **spatial and temporal knowledge** from **texts**,
- (2) their **limits** in the face of the nature and **imperfections** of **geographic data**,
- (3) the tendencies of **LLMs** to **hallucinate**, that is, to generate erroneous but convincingly formulated information.



Early contribution of **AI to GIScience** in the 80s and 90s has been favored by the decline of **quantitative geography**, **lack of solid theories** to model GI phenomena on the one hand

On the other hand, this major trend was considered as particularly appropriate regarding the **complexity of geographical data**, the emergence of new **GIS computing capabilities** and the need for novel approaches

Expectations were high on **expert systems**, artificial-based, supervised and unsupervised **neurocomputing approaches**, and also encouraged by the development of **formal reasoning languages**



Basically, a **LLM** is a type of machine-learning model that can understand and generate human language via **deep neural networks**.

The basic task of a **LLM** is to calculate the probability of a word following a given input in a sentence: for example, “The sky is ____” with the most likely answer being “blue”. This is done using a large text dataset (or corpus) that constitutes the **pre-trained language model**.

What makes a language model “**large**” is the size of its architecture.

A **LLM** is based on **neural networks** that “learn” from and process information

Furthermore, **LLMs** comprise a large number of **parameters** (for example, GPT has 100+ billion) trained on vast quantities of unlabeled text data via self-supervised or semi-supervised learning.



Attention: a statistical evaluation of the impact of each token fed through an LLM.

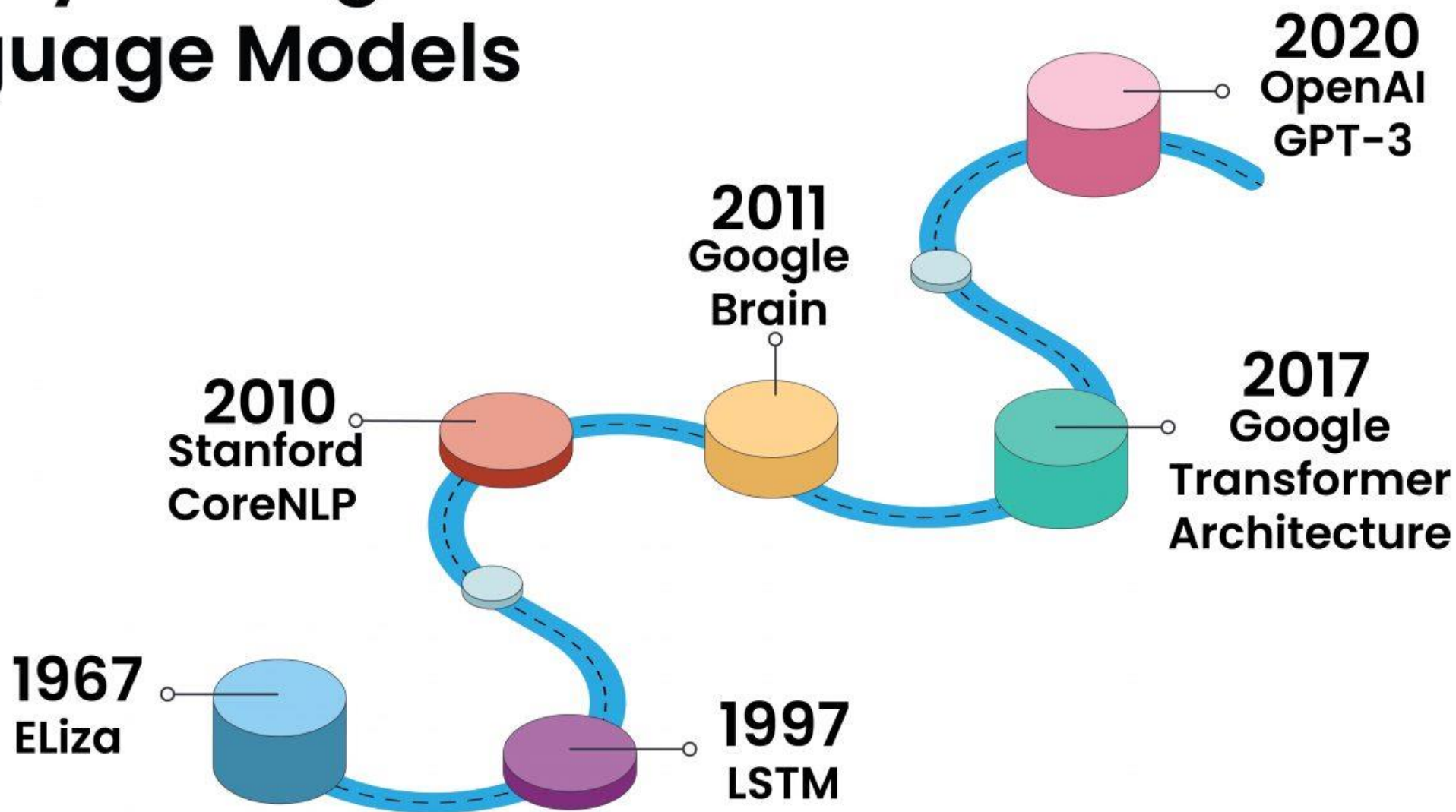
Embeddings: numerical representations of words that capture their meanings and relationships in terms of context.

Transformers: neural network architecture that forms the basis of most LLMs.

Prompt: the input a user provides to an LLM to illicit a response or carry out a task.

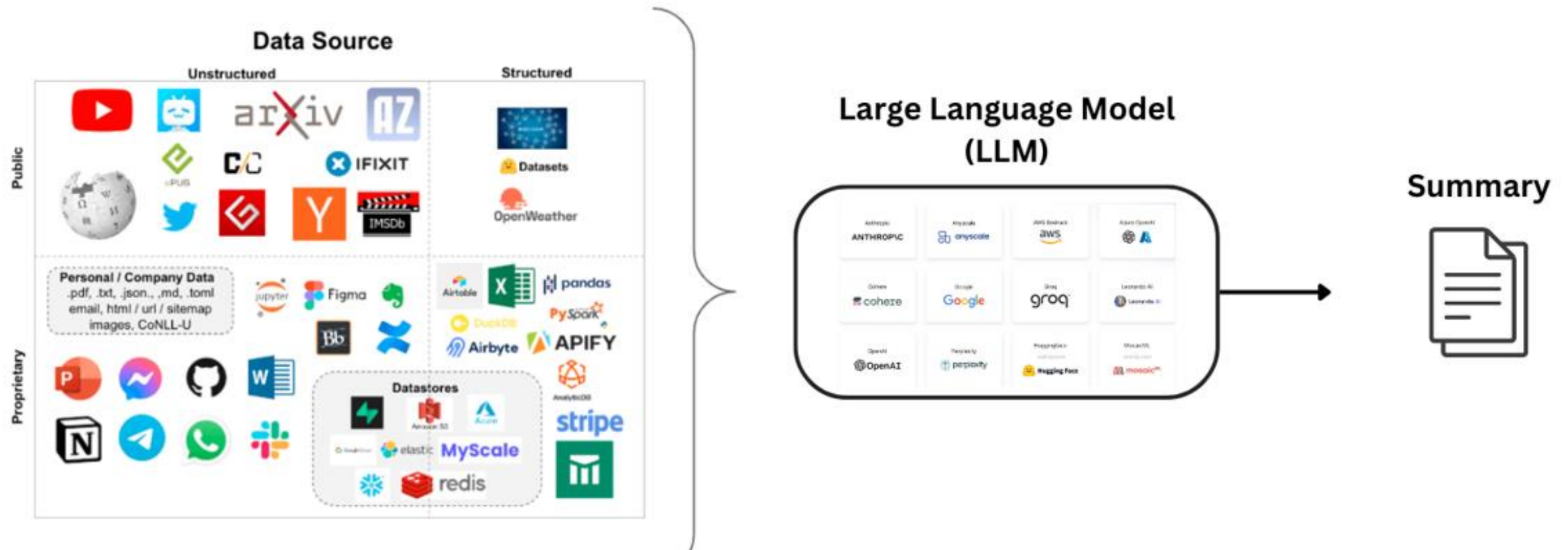
Instruction tuning: training a language model to answer different prompts to learn how to answer new ones.

History of Large Language Models



LLMs can be used to summarize information in a wide range of unstructured content, such as detecting spatial references in text and images to answer questions

Pre-training models (e.g. GPT-3/GPT-3.5 or 4.1) are trained on vast amounts of data, allowing them to learn a wide range of language patterns and structures.



LLMs use algorithms and statistical models to analyze and infer patterns from vast amounts of data.

The acronym **GPT** refers to **ChatGPT**'s architecture; it (1) **generates** the next word in a sequence based on (2) extensive **pretraining** based on the collective written works of humanity, then (3) **transforms** multiple words simultaneously to generate meaning and context.

Generative

Predicts the next
token in a sequence



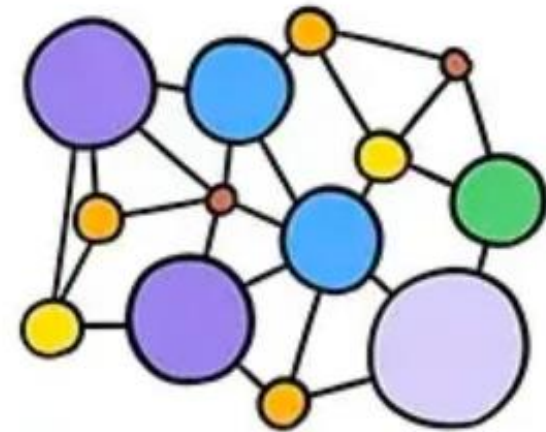
Pretrained

Primed with
massive datasets



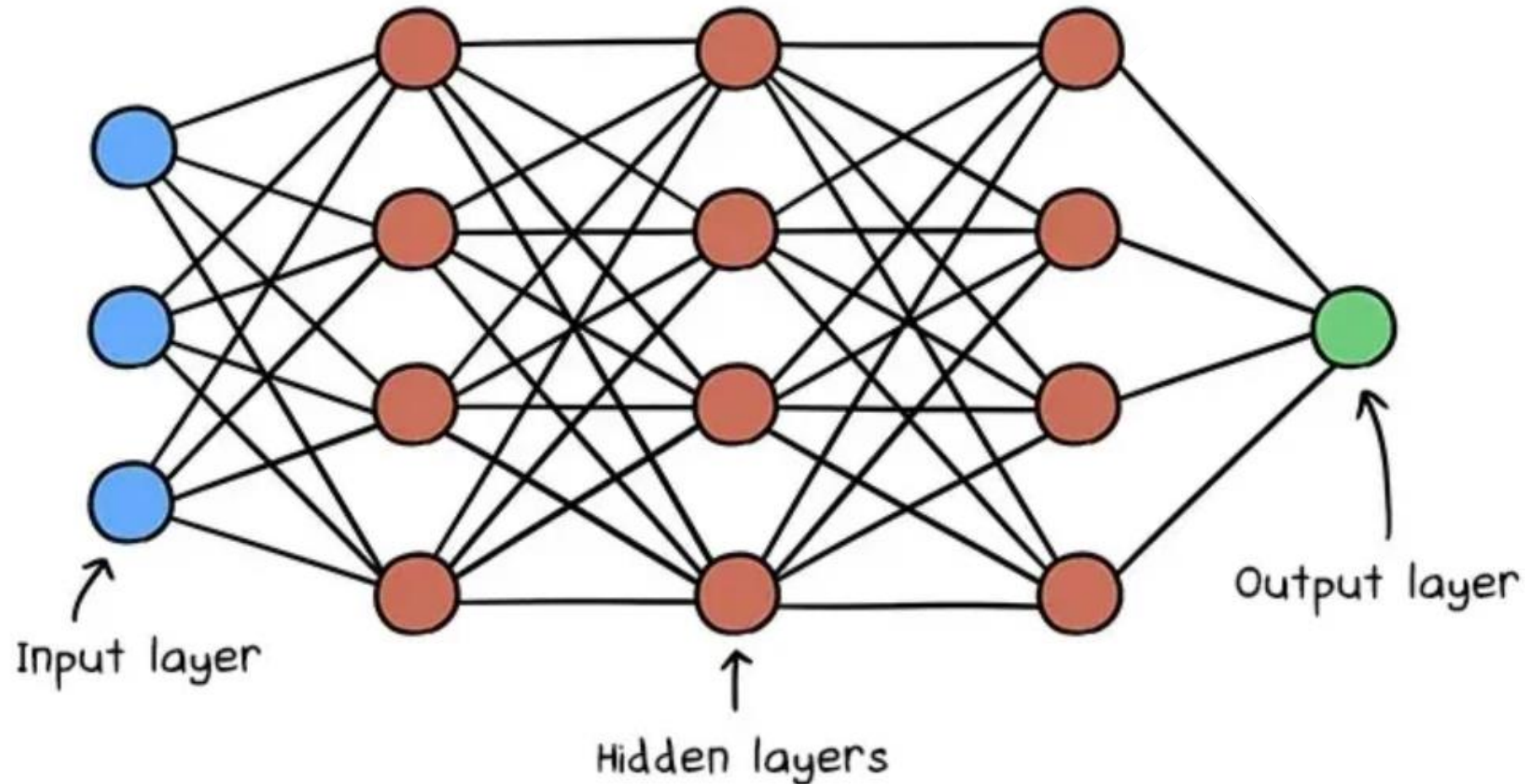
Transformer

Values input tokens
simultaneously



ChatGPT's transformer model uses an (artificial) neural network, much like the human brain.

This structure allows both systems to generate a vast array of outputs, or thoughts, thanks to the immense number of unique connections between nodes



**Large Language Models
(LLMs) + GIScience:
what should we
expect?**

So what?





Textual data



Web data



Sensor data



GIS data

LLMs

- NLP
- Textual interpretation
- Semantic analysis
- Spatial narrative analysis
- Language generation
- Contextual processing
- Location-based narratives

Synergy

- Spatial queries & interpretation
- Spatial analysis generation
- Map language description
- Spatial data accessibility
- Textual/spatial data generation
- Code generation



GIScience

- Spatial data collection
- Location-based services
- Spatial analysis
- Spatial visualisation
- AI GIS
- Volunteer GIS
- Digital twins

**Large Language Models
(LLMs) + GIScience:
what should we
expect?**

**But some
challenges
ahead!**





ChatGPT



Est-ce-qu'il y a un problème sur cette image ?



Non, il n'y a pas de problème apparent sur cette image. Il s'agit d'une main humaine levée avec les doigts écartés. Si vous avez une question spécifique ou pensez qu'il y a un détail inhabituel, pourriez-vous préciser ce que vous voulez examiner ?



Geographical/geopolitical Bias in LLMs

Does Ceuta belong to Spain?

ChatGPT -> Response if the question is asked in

Arabic (translated):

Ceuta did not become a natural part of Spain in a geographical sense, because it is located in North Africa. However, it is considered an integral part of Spain politically and administratively. Ceuta has been under Spanish control since the 15th century and became an autonomous city within the Spanish political system. Therefore, it is considered legally and politically part of Spain, although it is not geographically connected to the Spanish mainland.

Spanish (translated):

Yes, Ceuta belongs to Spain. It is an autonomous city located in North Africa, on the coast of the Mediterranean Sea, right on the Strait of Gibraltar. Although it is geographically on the African continent, it is an integral part of Spain and, therefore, of the European Union. Ceuta, along with Melilla, another Spanish autonomous city in Africa, has a special status within the Spanish territorial administration



Cardinal direction bias in LLMs (+granularity)

Which direction to take from Paris to Luxembourg

LLAMA-2

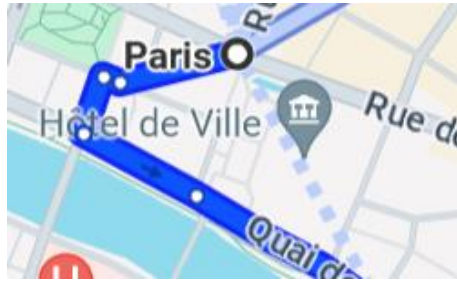
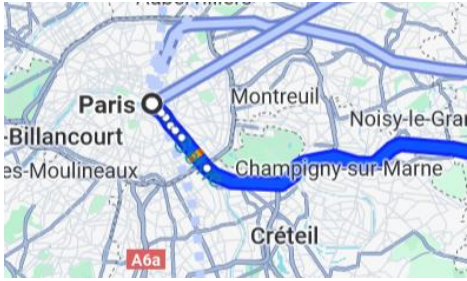
The direction from Paris, France to Luxembourg City, Luxembourg is **southwest**.

To get to Luxembourg City from Paris, you can take a train or drive on the A4 motorway. The journey takes around 3-4 hours by train and around 5-6 hours by car, depending on traffic and weather conditions.

Alternatively, you can also take a flight from Paris to Luxembourg Airport, which is located about 8 km west of Luxembourg City. The flight duration is around 1 hour.

ChatGPT

To travel from Paris to Luxembourg, you would **generally** head **northeast**.



Cardinal direction bias in LLMs (+granularity)

Which direction to take from Paris to Luxembourg?

LLAMA-2

Any reason for this mistake?

Yes, there are several reasons :

Outdated knowledge: I am a machine learning model, and sometimes my training data may be incomplete or outdated.

Typos or errors: As a machine learning model, I am not perfect and may make typos when generating text. I mistakenly stated that Luxembourg is southwest of Paris instead of north.

Misunderstanding the context: I can misunderstand the context of your question or misinterpret the location you were referring to. For example, "Luxembourg," could refer to either Luxembourg City or Luxembourg Province, both of which are located in the north of France.

LLAMA-2

The direction from Paris, France to Luxembourg City, Luxembourg is **southwest**.

.....

Really Southwest, I thought Luxembourg was North of Paris?

My apologies, you are correct! Luxembourg is actually north of Paris, not southwest. Thank you for pointing that out. Here are the updated directions:



Cardinal direction bias in LLMs (+granularity/hierarchy)

Which direction to take from Paris to Luxembourg?
Same question to LLAMA-2 after a few minutes...

LLAMA-2

Since you are coming from Paris, the best direction to travel to Luxembourg would be south. So, you will need to head towards the southern part of France before crossing into Luxembourg.

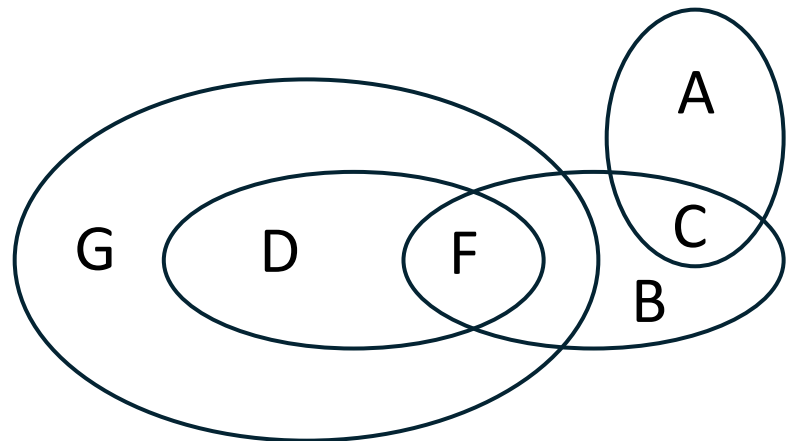
Here are some approximate directions:

From Paris, take the A6 motorway towards Marseille.

Continue on the A6 for about 200 miles (320 km) until you reach the French border.

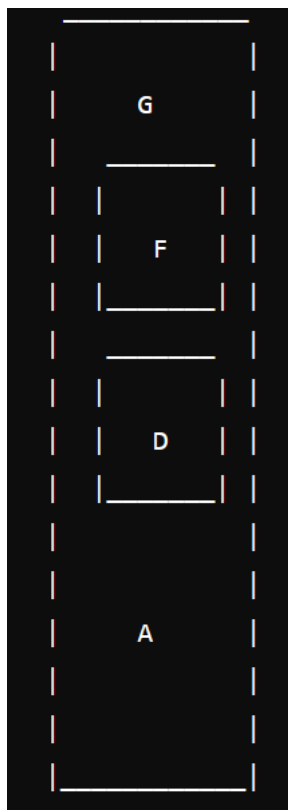
Once in Luxembourg, follow the signs towards the city center or your final destination.

Note: These directions assume you are starting from the city center of Paris. If you are staying elsewhere in Paris, you may want to factor in additional time for your journey. Additionally, traffic and road conditions can affect travel times, so it's always a good idea to check for updates before embarking on your trip.

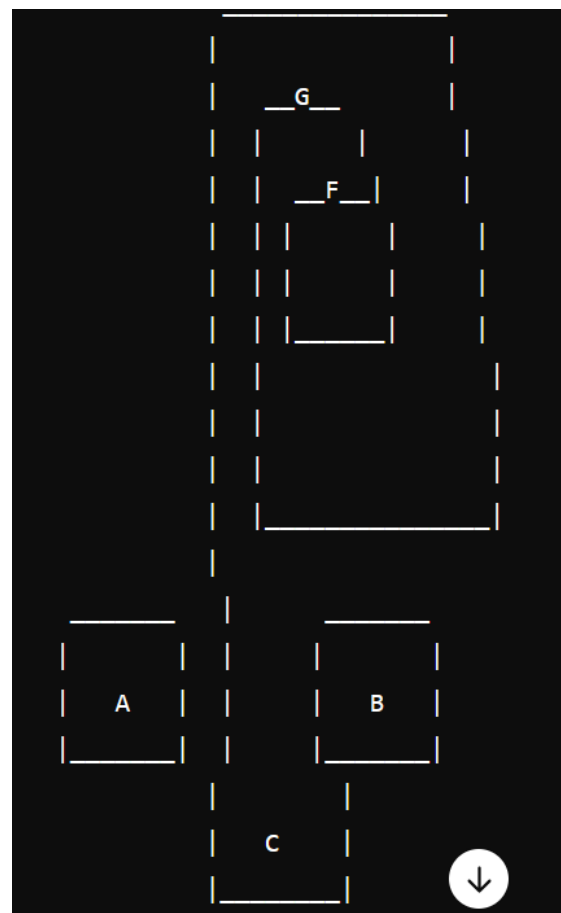


Topological bias in LLMs

To ChatGPT: In a Venn diagram, G encompasses F, A and B overlap to make area C, B and D overlap to form F. Finally, A does not intersect G. Could you draw the Venn diagram for me?



First and 2nd attempt...



3rd attempt...

Sample Prompt:

“Generate a detailed map of the D-Day landings in Normandy, highlighting key landing zones such as Utah Beach and Omaha Beach, along with other significant landmarks.”



Sample Prompt:“ Make me a sketch of the Crozon Peninsula: to the South is Morgat, to the Northwest is Camaret, to the north is Lanvéoc. ”





Challenges facing LLMs (when applied to GIS)

Prompt understanding (queries, spatial relations)

Data privacy and security

Training, bias & debiasing (still no auto-training with generated data)

Data diversity, regional diversity

Scalability and efficiency (auto-validation not available)

Interpretability and explanation

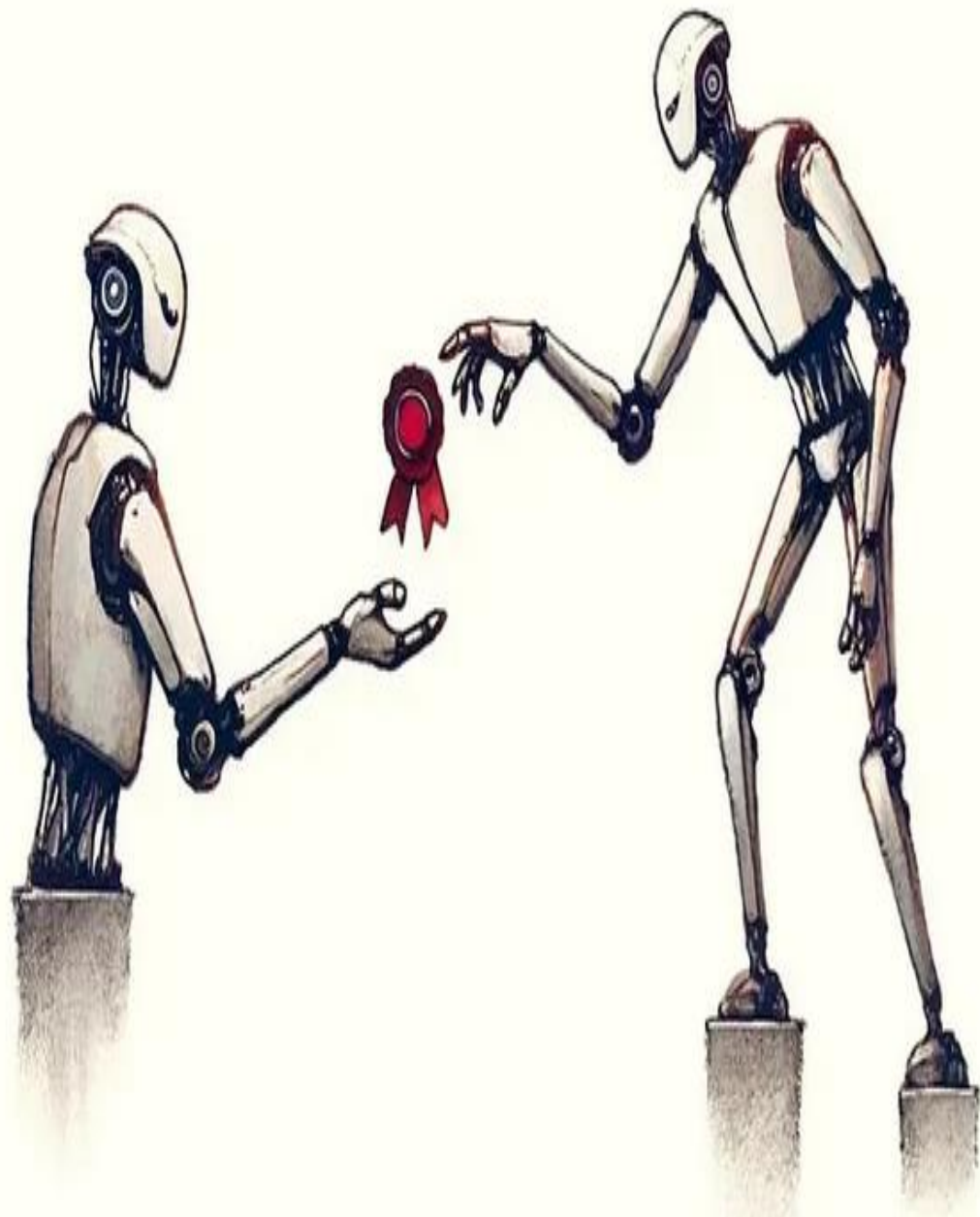
Context understanding (reasoning capabilities are limited)

Lack of disruptions

Ethics and diversity

Standards and industry-driven, “lack” of open systems

Sustainability & education



Bias in LLMs

LLM models are firmly entrenched in the data used to train them

This can lead to outputs that reflect gender, race, or other forms of bias

This is not only ethically problematic but can also impact the utility and acceptance of LLMs

Addressing bias is about ensuring fairness, promoting inclusivity, and building trust in AI.

It's about creating AI systems that respect and understand the diversity of human experiences rather than reinforcing harmful stereotypes



Metrics in evaluating bias in LLMs

Quantitative semantic and probabilistic fairness metrics, to quantify differences in different model's behaviour across a range of demographic groups

But by a far integration of spatio-temporal properties, especially for geographical prompts: landmarks, image classification and segmentation, detection of change

Integration of human/expert knowledge

This implies active learning processes iterative interactions with domain experts, and crowdsourcing collaborative platforms to gather feedback and validations (e.g. using reinforcement learning methods).



Mitigating bias in LLMs

Leveraging pre-trained models on general data and fine-tuning them with domain-specific and preselected geographical data to improve performance

Regular updates to the data to avoid further propagation of bias and unexpected emergent bias (temporal bias)

Extending the distribution of geographical data to new examples and appropriate diversity

Removing inappropriate geographical data and generating new data

Identifying hidden data and integrating them

Prompt 'engineering' & fairness training mitigation (words, attention), relevant metrics required

Combine the outputs of multiple neural networks trained on different data to develop robust and reliable models that accurately represent the complexity of geographical data

Prompt 'engineering'

Use clear and specific instructions

Be mindful of your own biases in your assumptions and opinions, avoid using harmful stereotypes

Use a variety of prompts

Use inclusive language

Test the LLM's responses (humans in the loop)

Use multiple LLMs





Brain vs LLMs similarities

Both are based on neural nets (inferences on the data)

Both are knowledge-based, and have training processes

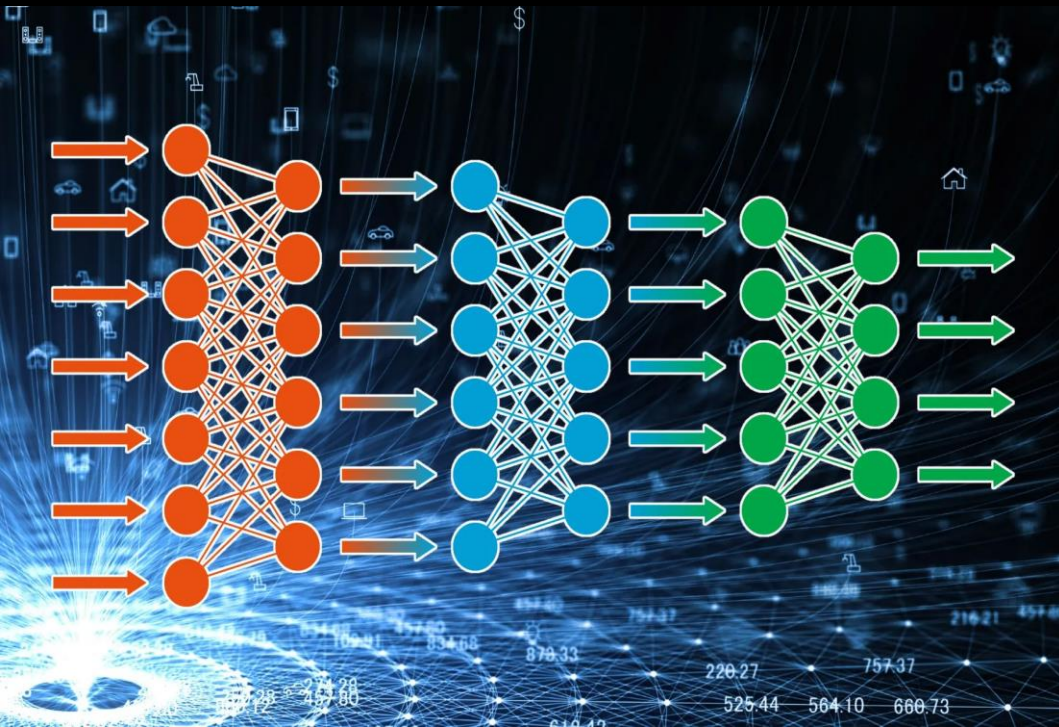
Brain vs LLMs differences

The human brain operates through complex biochemical processes and has consciousness, emotions, and experiences

LLMs operate purely through mathematical computations and algorithms without consciousness or subjective experiences

Human memory is associative and contextual, influenced by emotions, sensory experiences, and personal history. AI "knowledge" is based on the data it was trained on

Humans often use intuition, personal judgment, and deep contextual understanding while LLMs inferences are based on statistical patterns and algorithms



The
Future
of Large
Language Models



The **Future** of Large Language Models

Autonomous and sensor-based LLMs

Robotics-based LLMs

Multi-LLMs

Biological-inspired LLMs

Learning capabilities

Reasoning capabilities

Agent-based capabilities



Opportunities and challenges of integrating Geographic Information Science and LLMs, 2025, N. van de Weghe, L. de Sloover, A. G Cohn, H. Huang, S. Scheider, R. E. Sieber, S. Timpf, C. Claramunt, JOSIS, 25, 1, to appear .

Shaping Future Sustainable Cities with AI-powered Urban Informatics: Toward Human-AI Symbiosis, 2025, Y Yue; G. Yan; T. Lan; R. Cao; Q. Gao; W. Gao; B. Huang; G. Huang; Z. Huang; Z. Kan; Xi. Li; D. Liu; Xi. Liu; D. Ma; L. Wang; J. Xia; X. Yang; M. Zhou; A. G. Yeh; R. Guo; C. Claramunt, Computational Urban Science, to appear.

Rethinking GIScience education in an age of disruptions, 2025, A. E. Frazier, T. Nelson, P. Kedron, E. Shook, S. Dodge, A. Murray, M. Goodchild, S. Battersby, J I. Blanford, C. Claramunt, J. Holler, C. Koylu, A. Lee, S. Manson, S. Şalap-Ayça, J. P. Wilson, B. Zhao, L. Bennett, C. Cabrera, R. Franklin, G. McKenzie, H. Miller, T. Oshan, S. Rey, F. Rowe, S. Spielman, W. Xu, Transactions in GIS, 29(2).

A research agenda for GIScience in a time of disruptions, 2024, T. Nelson, A. E. Frazier, P. Kedron, S. Dodge, B. Zhao, M. Goodchild, A. Murray, S. Battersby, L. Bennett, J. I. Blanford, C. Cabrera-Arnau, C. Claramunt, R. Franklin, J. Holler, C., A. Lee, S. Manson, G. McKenzie, H. Miller, T. Oshan, S. Rey, F. Rowe, S. Ş. Ayça, E. Shook, S.man, W. Xu & J. Wilson, 2024, International Journal of Geographical Information Science.

Collaboration between artificial intelligence and Earth science communities for mutual benefit, 2024, M. Chen, Z. Qian, N. Boers, F. Creutzig, G. Camps-Valls, K. Hubacek, C. Claramunt, J P. Wilson, S. Nativi, A. J. Jakeman, R. D. Müller, M. Batty, C. Zhou, F. Chen, Q. Wang, F. Zhang, C. M. Barton, J. Strobl, M. Meadows, C. Ratti, P. Hess, Q. Xu, Z. Zhang, Q. Gu, A-X. Zhu, H. Lin, L. Yuan & G. Lü, Nature Geoscience, 17, 947.

Artificial intelligence and visual analytics in geographical space and cyberspace: Research opportunities and challenges, Chen, M., Claramunt, C., Çöltekin, A., Liu, X., Peng, P., Robinson, A. C., Wang, D., Strobl, J., Wilson, J. P., Batty, M., Kwan, M.-P., Lotfian, M., Golay, F., Joost, S., Ingensand, J., Senousi, A. M., Cheng, T., Bandrova, T., Konecny, M., Torrens, P M., Klippel, A., Li, S., Zhang, F., He, L., Wang, J., Ratti, C., Kolditz, O., Lin, H., Lü, G., 2023,, Earth-Science Reviews, 241, 104438.

<https://sites.google.com/site/christopheclaramunt/>

<https://chatgpt.com/>

<https://www.perplexity.ai/>

<https://copilot.microsoft.com/>

<https://www.llama2.space/>

<https://mistral.ai>

<https://gemini.google.com/>

<https://openrouter.ai/>

<https://www.deepseek.com/>

