

Reading between the lines: Adapting GenAI for **Implicit Information Retrieval**



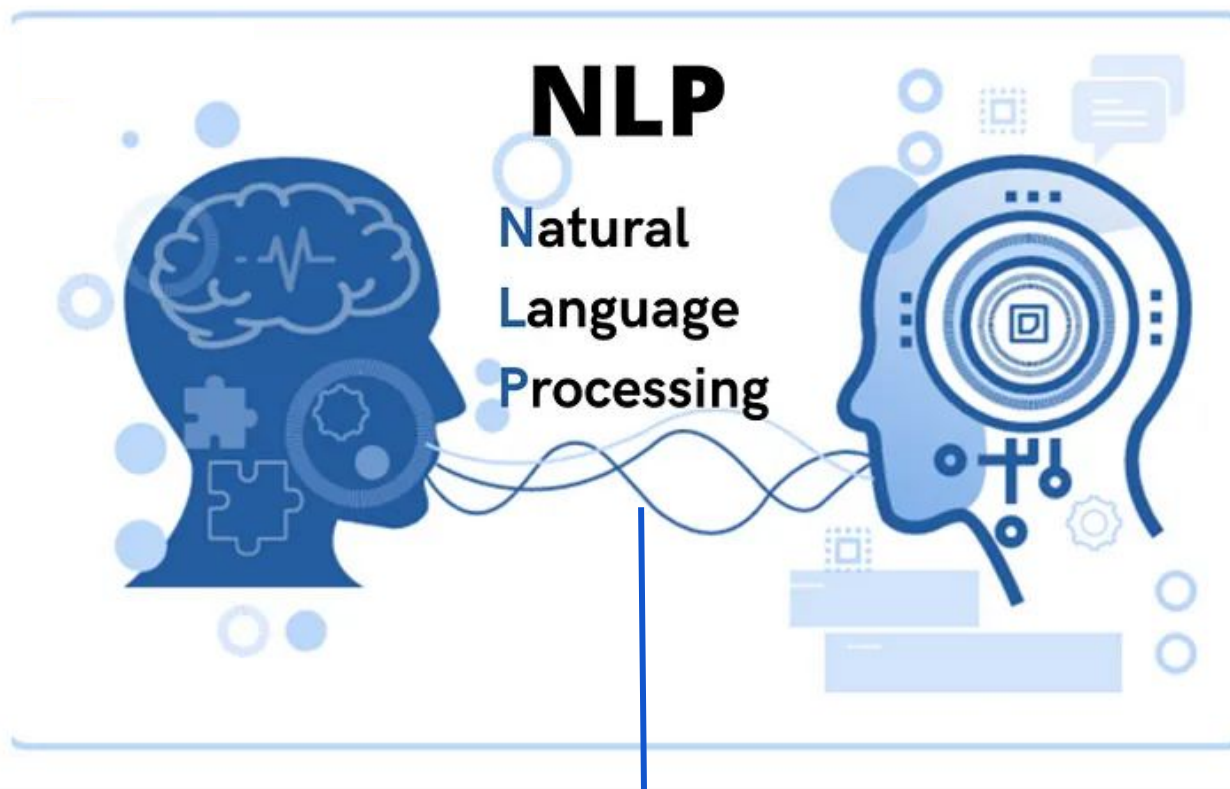
Nicolay Rusnachenko

[nicolay-r.github.io](https://github.com/nicolay-r)

*Research Fellow in
Marking Medical Images with NLP
Bournemouth University, CfACTs+*



What is Information Retrieval?



When we wish to extract something out of this communication we are dealing with **Information Retrieval**

Implicit Information Retrieval

Implicit IR

Is related to

Author / Person / Character

Behind the text

Implicit IR [Example]

Hi, how are you?

Observations:

- Open-end question
- No aspects of interest
- Asking for attention

Implicit IR [Example]

Hi, how are you?

Observations:

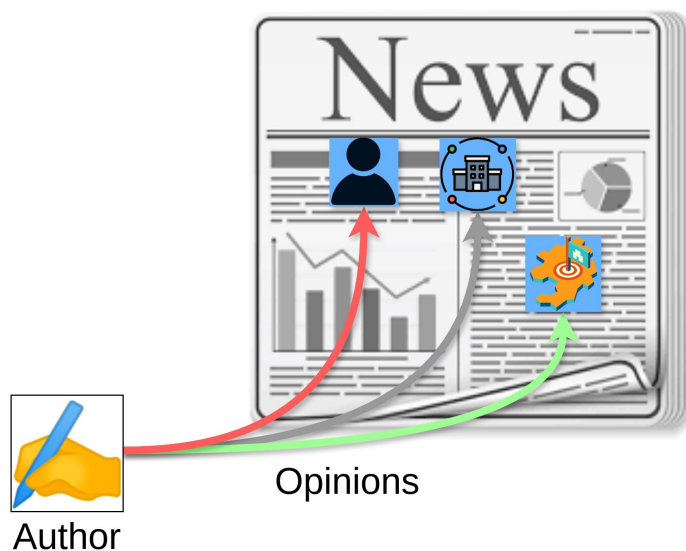
- Open-ended question
- No aspects of interest
- Asking for attention

Hey, who is that nice person?

Observations:

- Close-ended question
- Interest in person
- Emotion State
- Emotion Cause (Joy)

Implicit IR: Sentiment Analysis



Example:

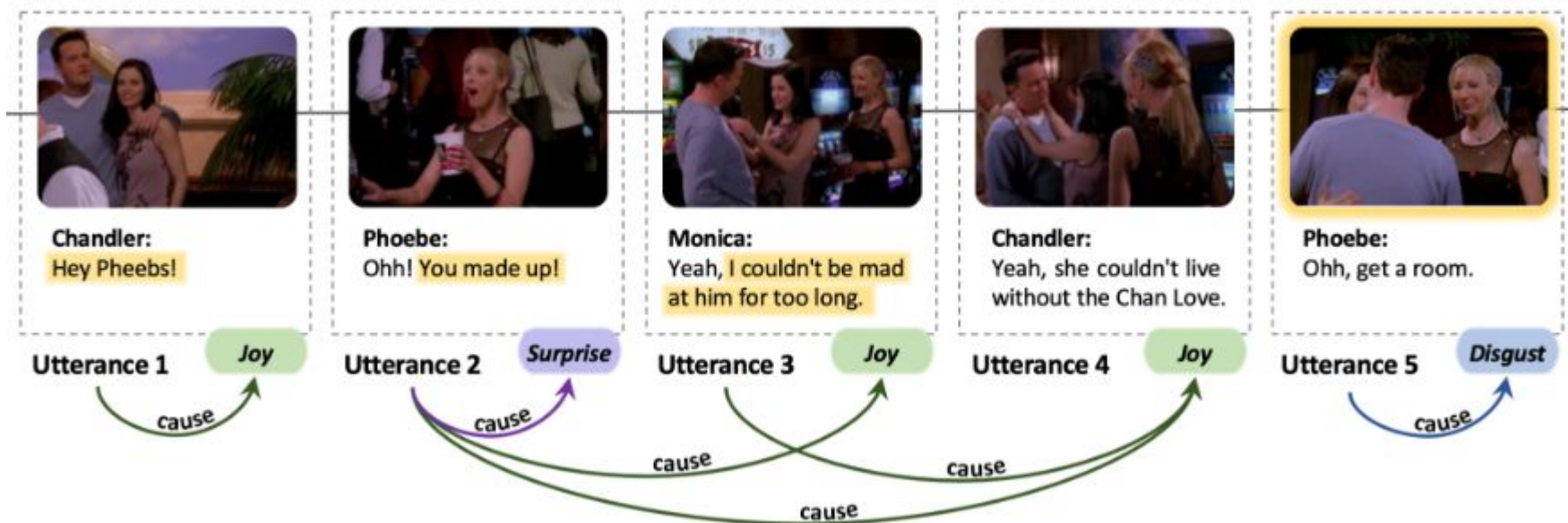


<https://arxiv.org/pdf/2305.17679>

Implicit IR: Emotion Causes

Example: ECAC-2024 task

Conversations for the F.R.I.E.N.D.S. TV show



https://nustm.github.io/SemEval-2024_ECAC

Key takeaways from the talk

- ✓ Key advances in AI that lead us to **GenAI**
- ✓ How to master your GenAI to perform implicit IR
- ✓ How to quick apply IIR for your CSV/JSON-lines data

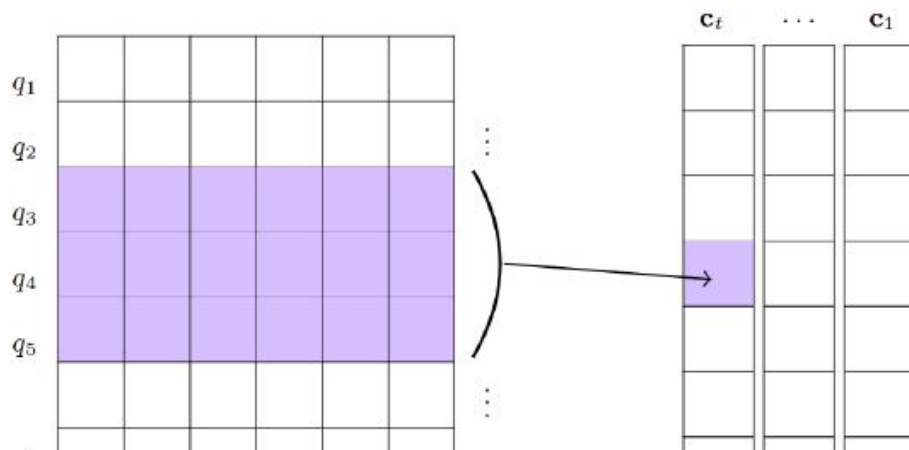
Towards Generative AI

Outline

Directions of advances:

- 1. Architectural**
- 2. Data-related**

Architectural [Conventional Networks]

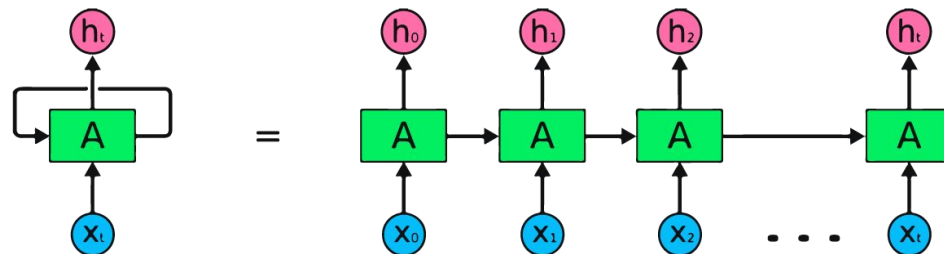


Convolutional neural networks

Sliding window across the embedded words

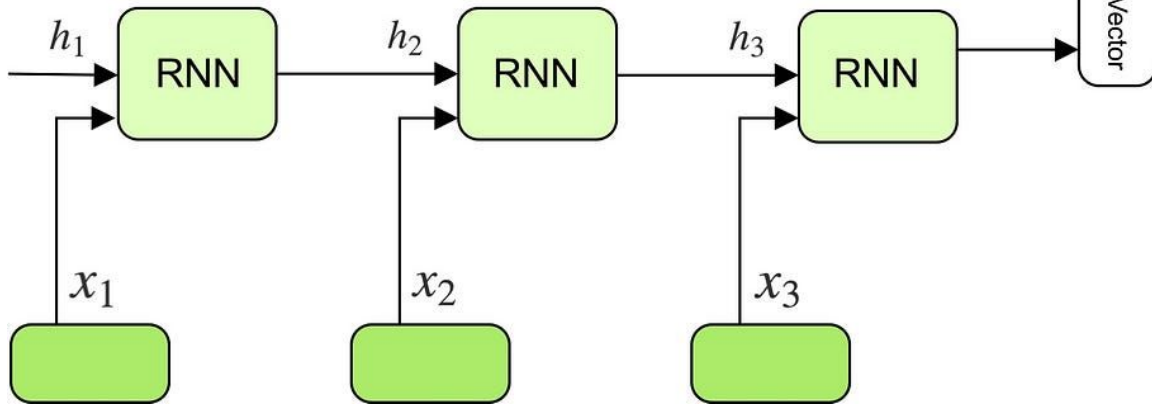
Recurrent Neural Networks (RNN)

Using one state that

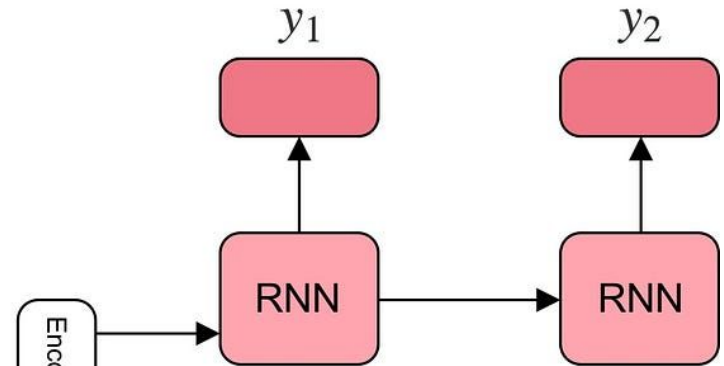


Architectural [Encoder-Decoder]

Encoder



Decoder



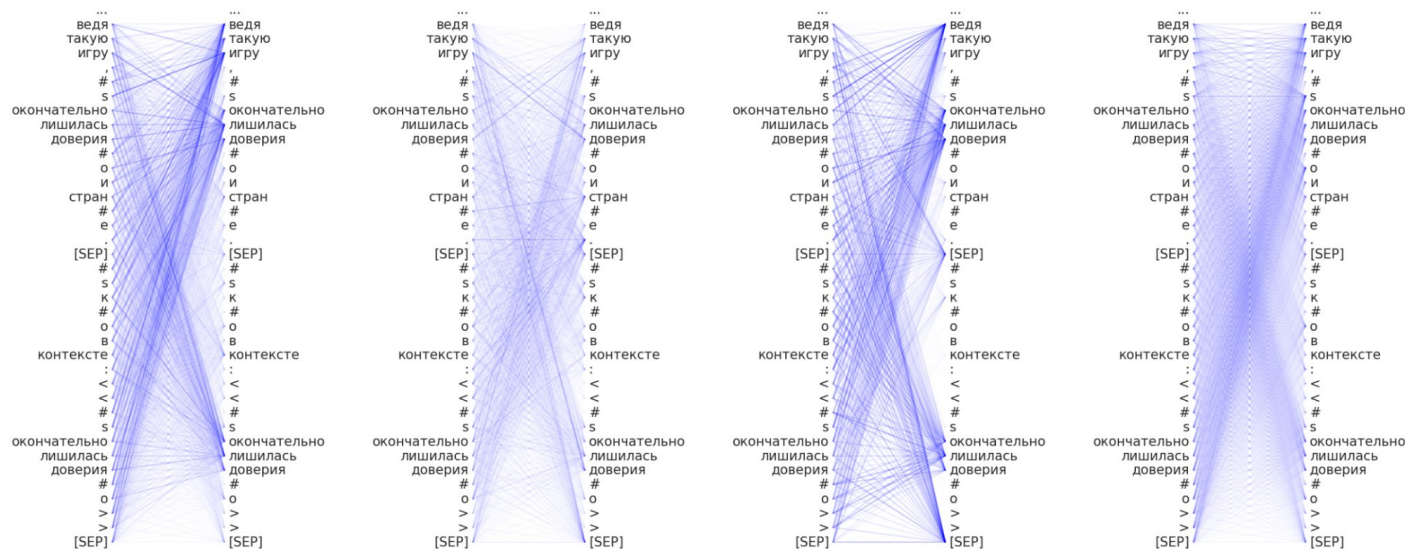
Architectural Transformers

The concept of Encoder-Decoder architectures + *Attention Mechanism* + *Concept of Heads*

<https://arxiv.org/abs/1706.03762>

What does BERT look at? An Analysis of BERT's Attention

<https://aclanthology.org/W19-4828.pdf>



<https://nicolay-r.github.io/#ruattitudes-2021>

Data-related advances

1. Supervised Learning

2. Unsupervised Learning

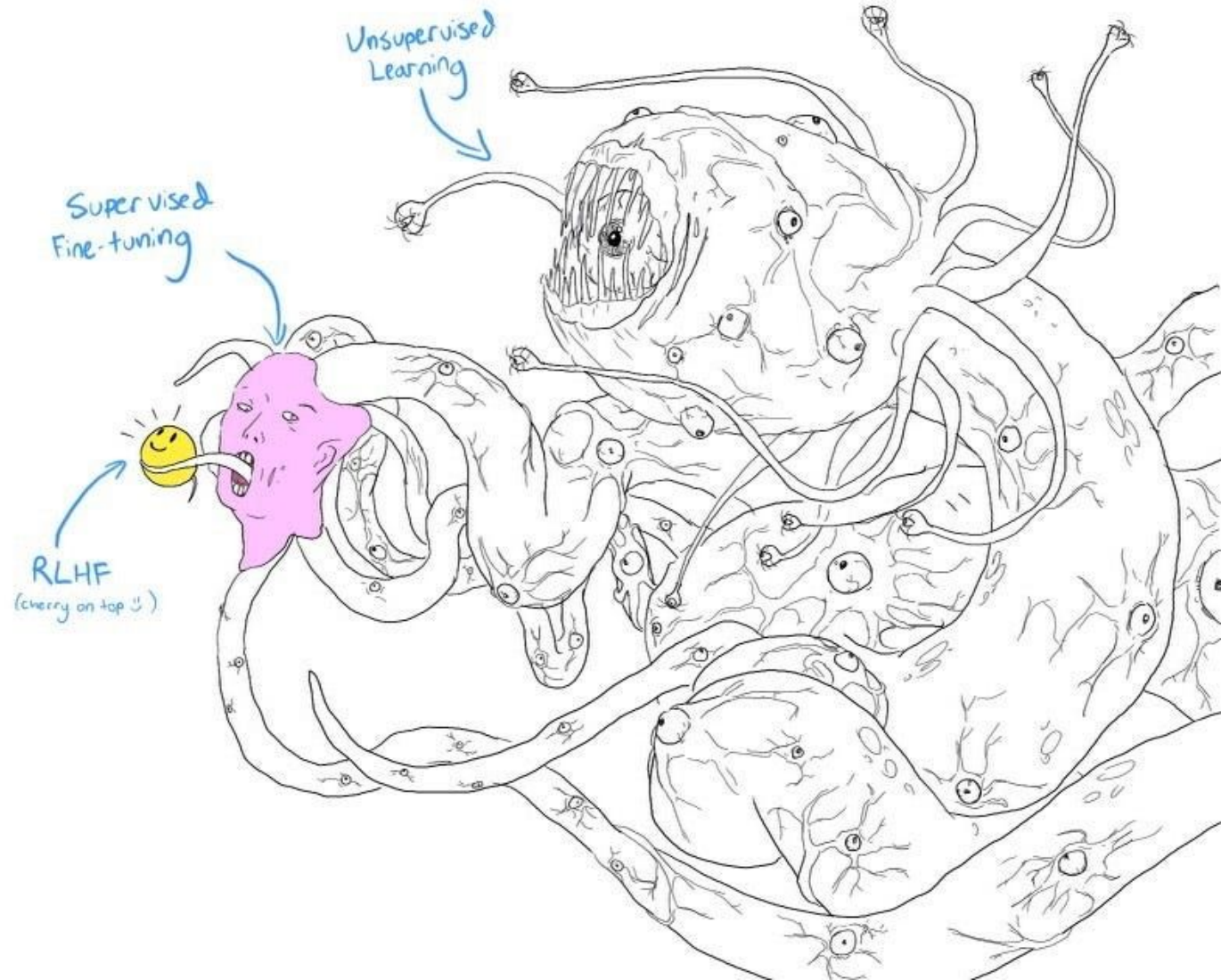
a. + Distant Supervision (optional)

3. Supervised Learning

a. Instructions

4. Reinforcement Learning With Human Feedback (*RLHF* Framework)

LLM: Transformers + Instruction tuning



Mastering IIR In Generative AI Era

Experiential Data

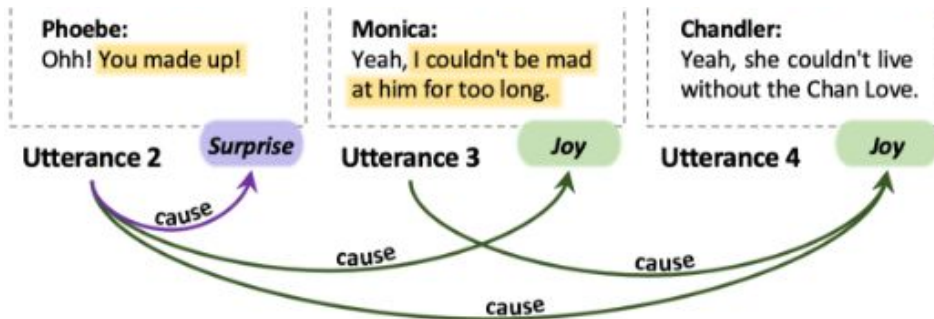
Task: **RuSentNE-2023**

Data: 10K sentences [Training]
2K sentences [Test]

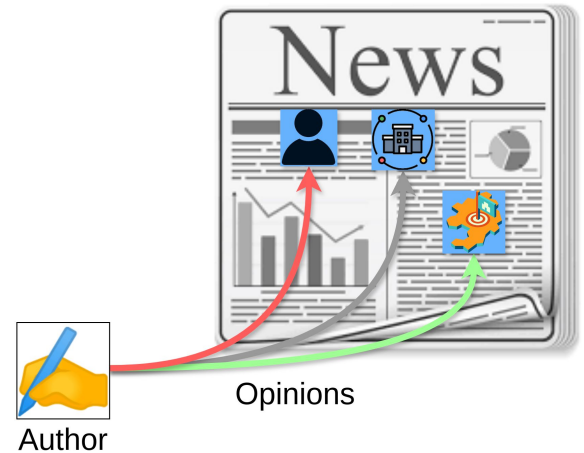
Language: Russian

Classes: positive, negative, neutral

Emotion States and Causes Extraction



Sentiment Analysis



Task: **ECAC-2024**

Data: 300 dialogues [Train]
100 dialogues [Test]

Language: English

Classes: {SURPRISE,
SADNESS, JOY, DISGUST,
FEAR, ANGER}

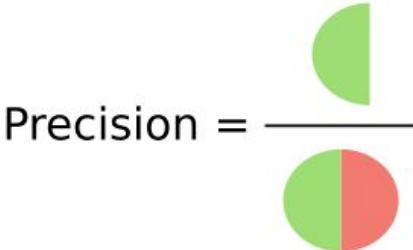
Evaluation

We wish to assess our methods, but how?

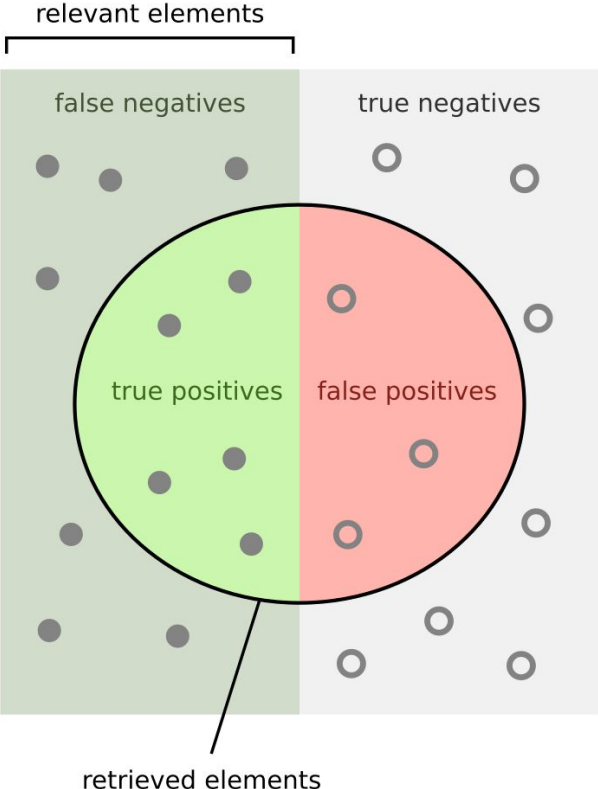
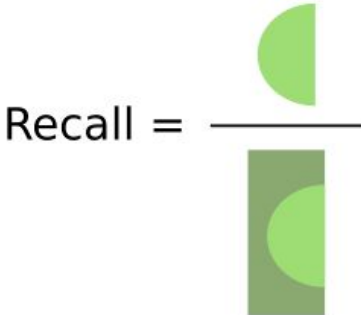
Solution:

F1-measure

How many retrieved items are relevant?



How many relevant items are retrieved?



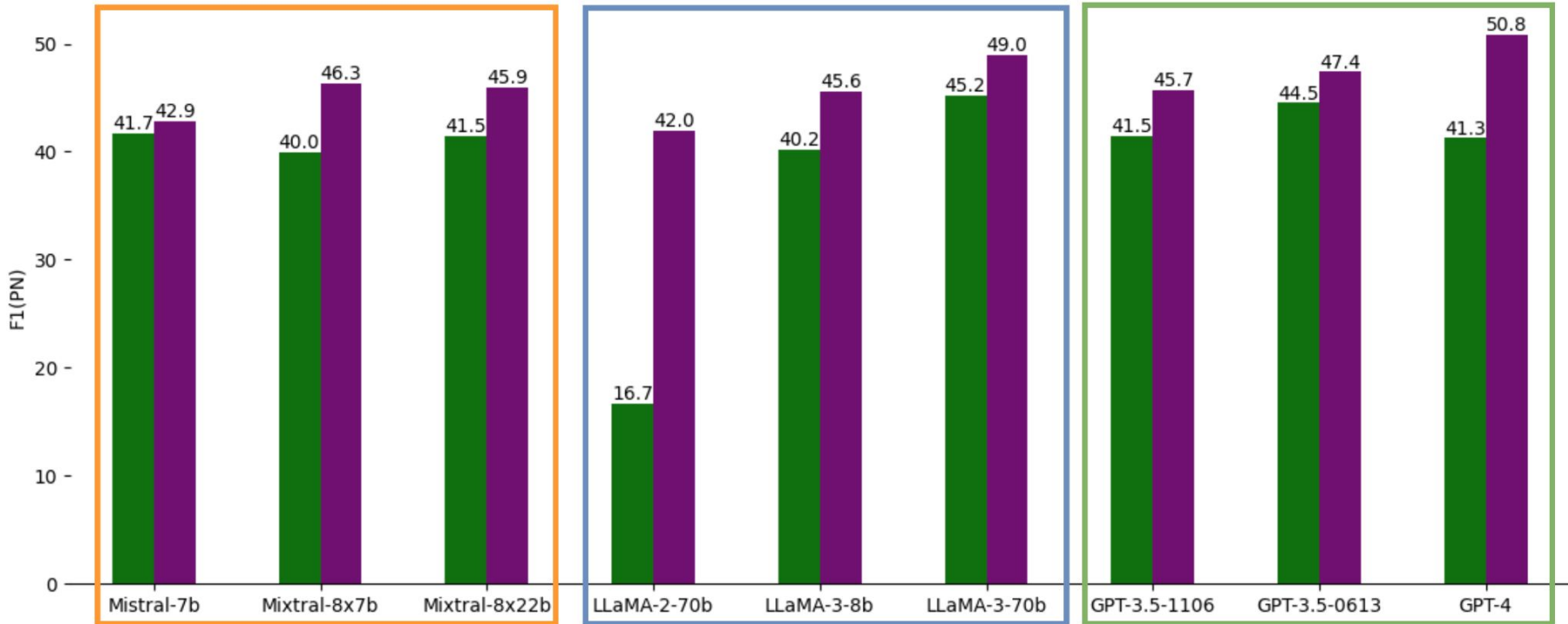
Initial Experiment





Zero-shot Key Takeaways

[Translate Texts into English]



<https://github.com/nicolay-r/RuSentNE-LLM-Benchmark>



Zero-shot Key Takeaways


[Masking Entities]



Concept

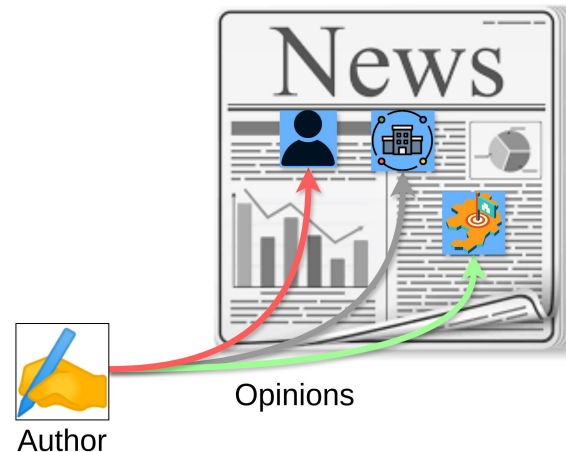
Removing bias in between *entities* and *text labels* in dataset.

Empirical

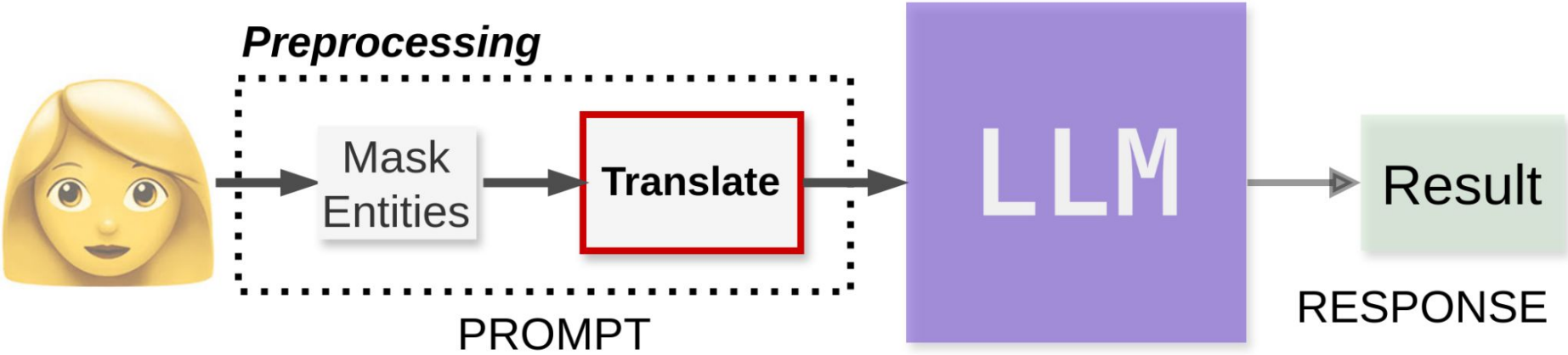
 *HALf-MAsked Model for Named Entity Sentiment analysis*

<https://www.dialog-21.ru/media/5923/podberezkoplusetal112.pdf>

Showcase the performance benefit in **1-4%**



Do text Translation



Mastering IIR
In Generative AI Era
Reasoning

Common Sense Reasoning

Key: Common Sense Reasoning problems could be decomposed into intermediate reasoning steps that lead to answer.

(c) Zero-shot

Q: A juggler can juggle 16 balls. Half of the balls are golf balls, and half of the golf balls are blue. How many blue golf balls are there?

A: The answer (arabic numerals) is

(Output) 8 X



(d) Zero-shot-CoT (Ours)

Q: A juggler can juggle 16 balls. Half of the balls are golf balls, and half of the golf balls are blue. How many blue golf balls are there?

A: **Let's think step by step.**

(Output) *There are 16 balls in total. Half of the balls are golf balls. That means that there are 8 golf balls. Half of the golf balls are blue. That means that there are 4 blue golf balls. ✓*

**Large Language Models are
Zero-Shot Reasoners**

<https://arxiv.org/pdf/2205.11916>

Reasoning Concept in IIR

Chain:  Aspect   Opinion   Retrieve

*Reasoning Implicit Sentiment with
Chain-of-Thought Prompting*

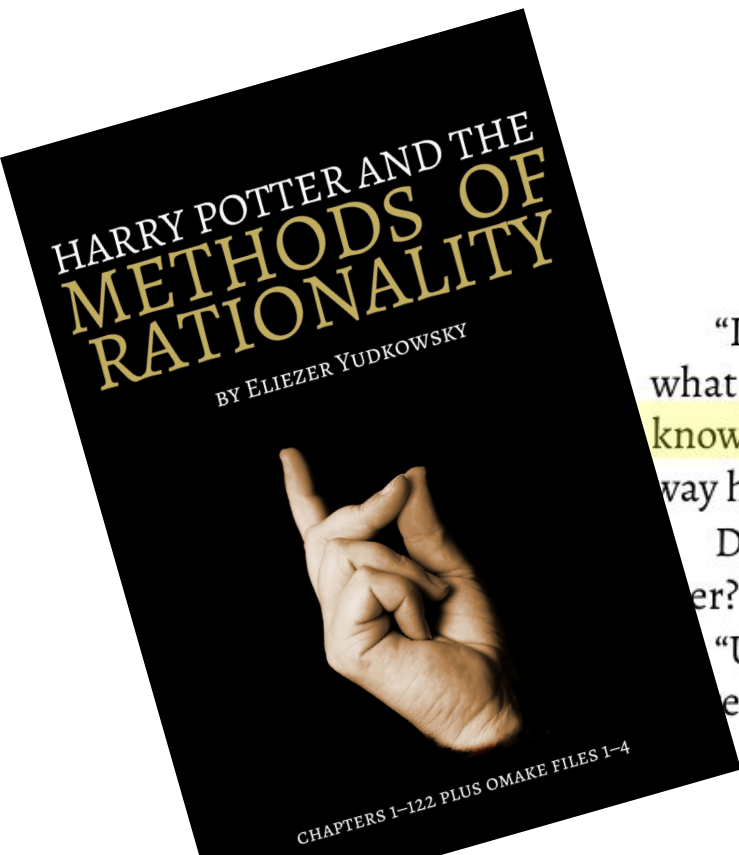
<https://arxiv.org/abs/2305.11255>

🧠 Reasoning Concept in IIR

Chain: 💡 Aspect → 🤔 Opinion → 🧠 Retrieve

Reasoning Implicit Sentiment with Chain-of-Thought Prompting

<https://arxiv.org/abs/2305.11255>



“I ask the fundamental question of rationality: why do you believe what you believe? What do you think you know and how do you think you know it? What makes you think Lucius wouldn’t sacrifice you the same way he’d sacrifice anything else for power?”

Draco shot Harry another odd look. “Just what do *you* know about Father?”

“Um...seat on the Wizengamot, seat on Hogwarts’ Board of Governors, incredibly wealthy, has the ear of Minister Fudge, has the confidence of



Three-hop Reasoning Concept

1

Given the sentence [SENTENCE].
Which specific aspect of [ENTITY] is
possibly mentioned?





Three-hop Reasoning Concept

1

Given the sentence [SENTENCE].
Which specific aspect of [ENTITY] is
possibly mentioned?



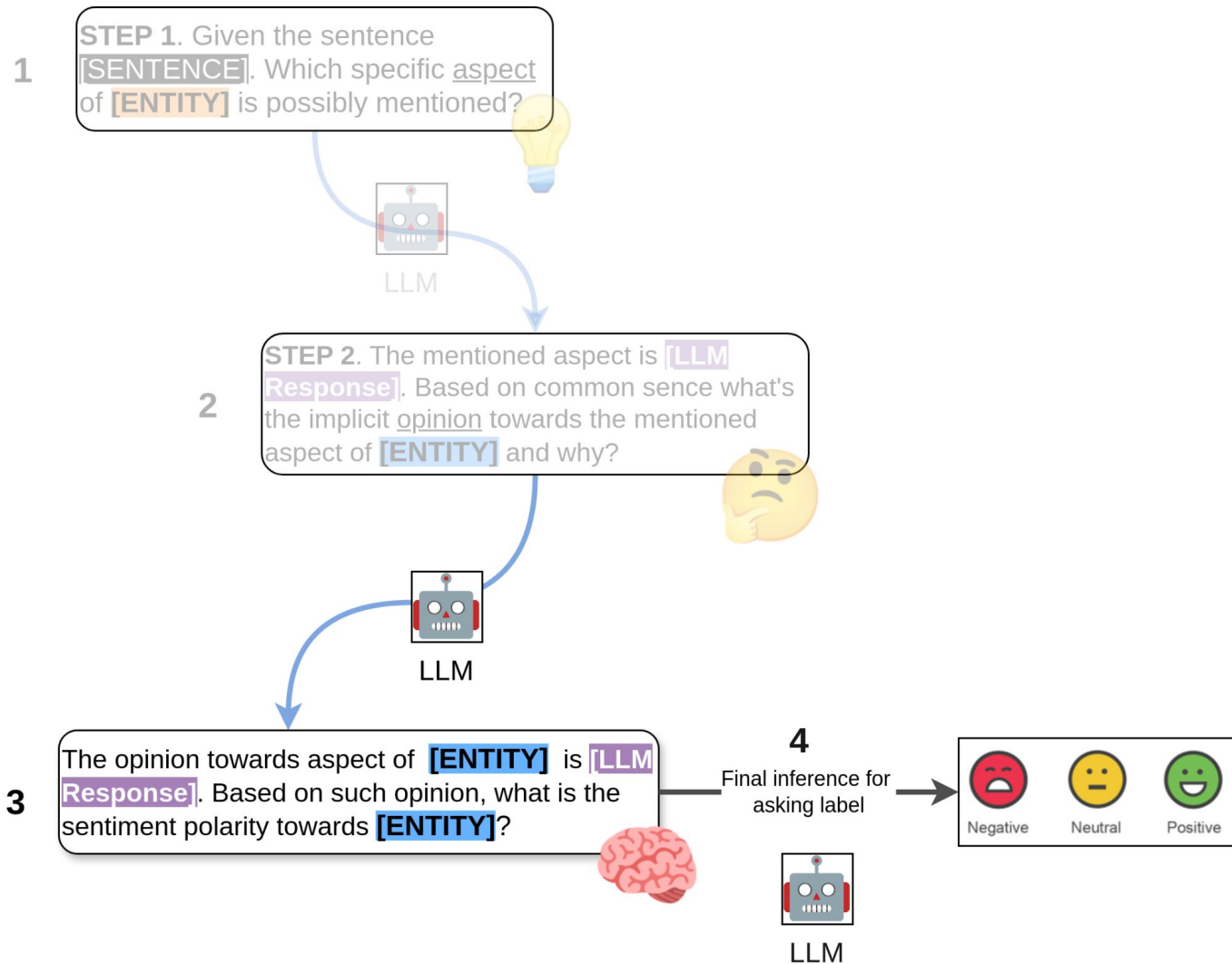
LLM

2

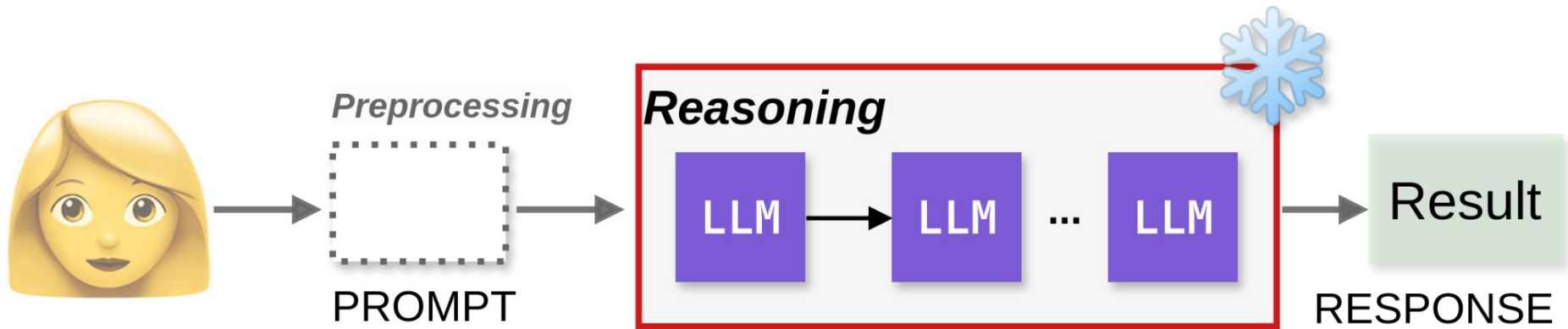
The mentioned aspect is [LLM Response].
Based on common sense what's the implicit
opinion towards the mentioned aspect
of [ENTITY] and why?



Three-hop Reasoning Concept



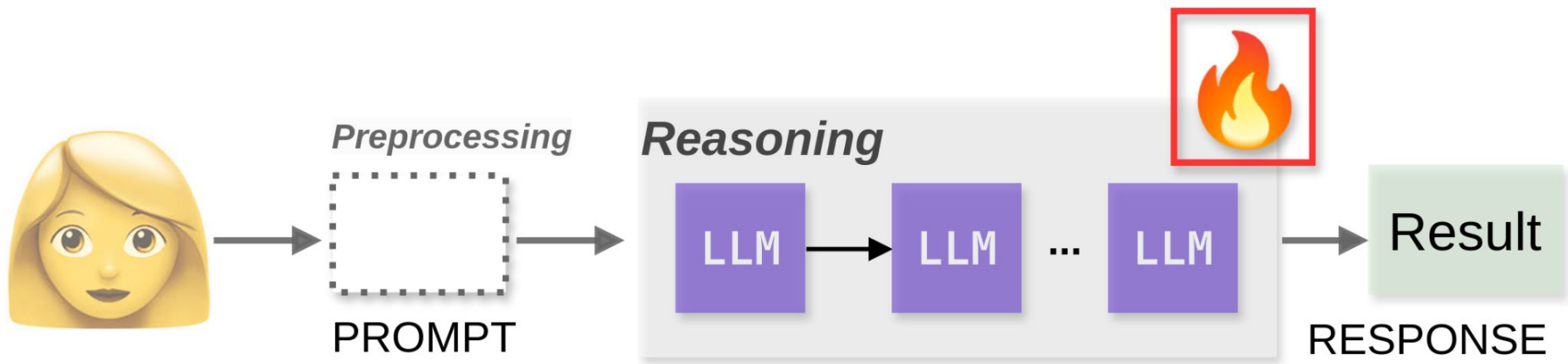
Pipeline with Reasoning



** LLM models are the same*

Chain: 💡 Aspect + 🤔 Opinion + 🧠 Think

🔥 Pipeline with Reasoning + **Tuning**



** LLM models are the same
but **fine-tuned** 🔥*

Chain: 💡 Aspect + 🤔 Opinion + 🧠 Think



Experiments

Emotion extraction

[https://nicolay-r.github.io/
#semeval2024-nicolay](https://nicolay-r.github.io/#semeval2024-nicolay)

Method	F1 (C) weighted
Zero-shot (Mistral-7B-v1)	2.54
Fine-tuned (FlanT5-BASE)	22.27
Fine-tuned-CoT (FlanT5-BASE)	24.28

Sentiment Analysis

[https://nicolay-r.github.io/
#ljom2024](https://nicolay-r.github.io/#ljom2024)

Method	F1 (P,N)
Zero-shot* (Mistral-7B-v1)	49.46
Fine-tuned* (FlanT5-BASE)	57.01
Fine-tuned-CoT* (FlanT5-BASE)	59.75

* texts were translated into English

Quick Start

Quick Start with **Objects Extraction**

BULKNER

A no-strings Named Entity Recognition (NER) framework of wrapped **LM models** for quick processing data in **CSV**



Input data



NER model



Annotated entities

Quick Start with Chain-of-Thought

Schema
of prompts



BULKchain

A tiny Chain-of-Thought
framework for processing data in
CSV tables via any **LLM**.

Supports applying series of
prompts formed into **schema**



Input: Very
long data




ChatGPT / any
your own



OUTPUT: With
columns And
answers

Models for Reasoning to Use

- **Curated list** of models for reasoning:
 - <https://github.com/Hannibal046/Awesome-LLM>
 - **Local:**
 - <https://huggingface.co/models>
 - **Third-party:**
 - <https://openrouter.ai/>
 - <https://replicate.com/>
- **Language oriented:** most recent list of reasoning models for 
Chinese language:
 - <https://x.com/AdinaYakup/status/1861908631807017007>

Conclusion

- ✓ importance of implicit IR from texts
- ✓ how to master your GenAI to perform implicit IR
Mask Entities, Translate Texts, adopt CoT +
Fine-tuning.
- ✓ how to quick start

Thank you for your attention!



Nicolay Rusnachenko

[nicolay-r.github.io](https://github.com/nicolay-r)

*Research Fellow in
Marking Medical Images with NLP
Bournemouth University, CfACTs+*

