Philosophy and Software Engineering: A Transdisciplinary Perspective on LLMs and the Patterns of Human Language Use

Christoph Durt (Technical University of Munich)

Tobias Hey (Karlsruhe Institute of Technology)

Which patterns do LLMs uncover, and how do they transform and recombine them?

What role do these patterns play in human language use?

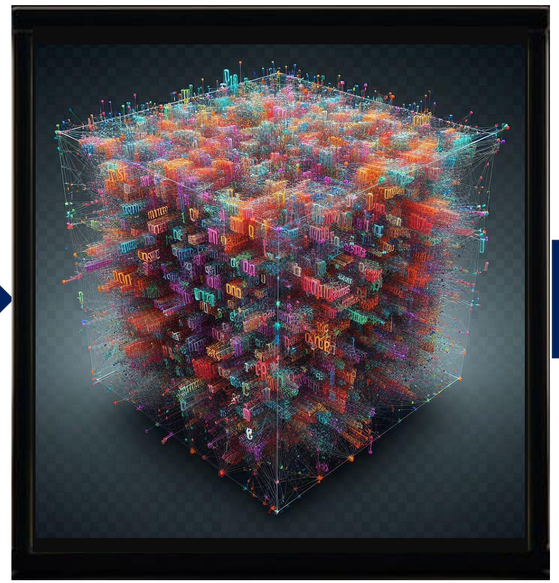

Text Input
LLM
Text Output

Text Input

Text Output

Training Data

Digital Input



Digital Output

Phil

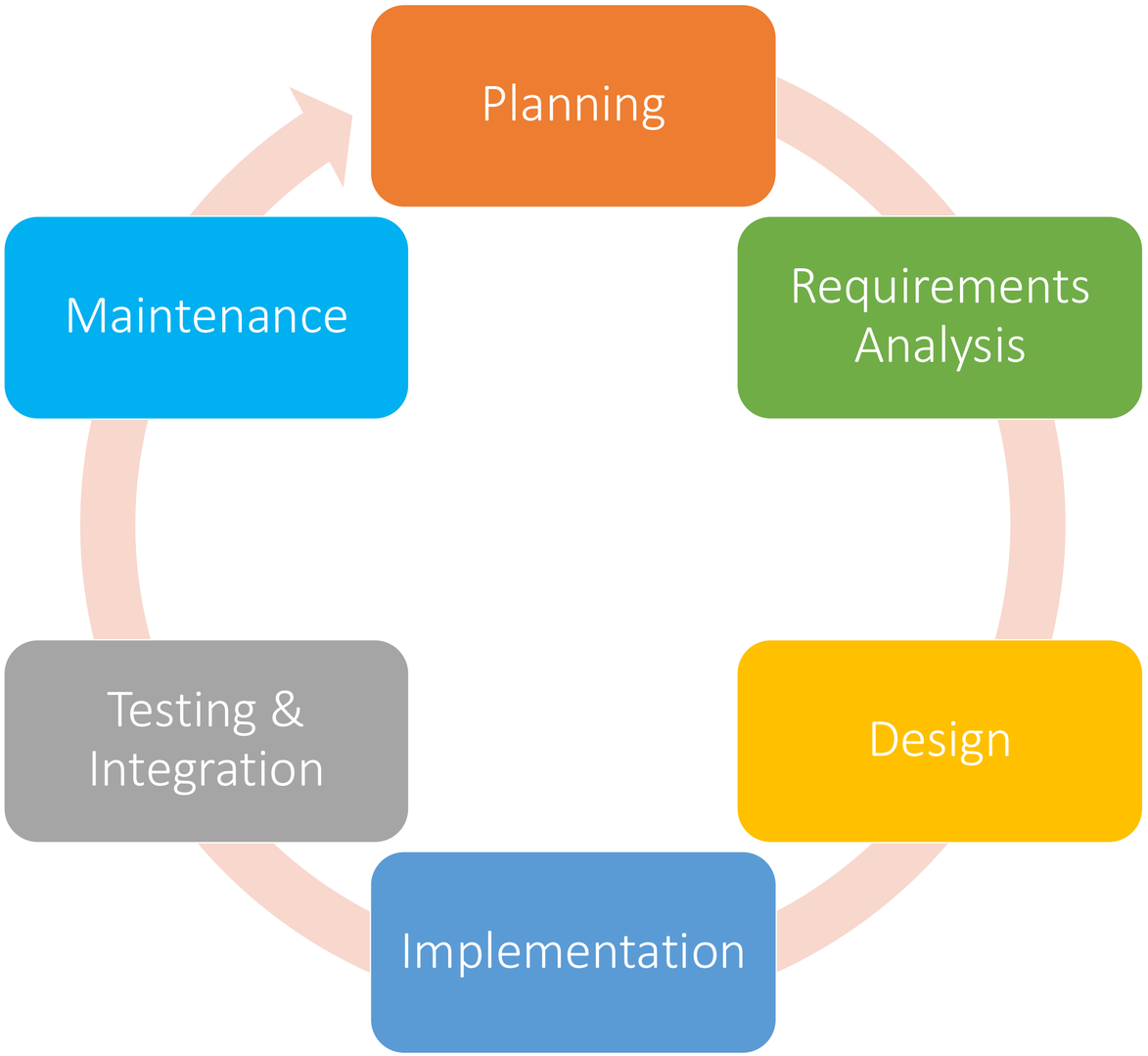
The conditions and nature of meaningful language use and its relationship to computation.

SE

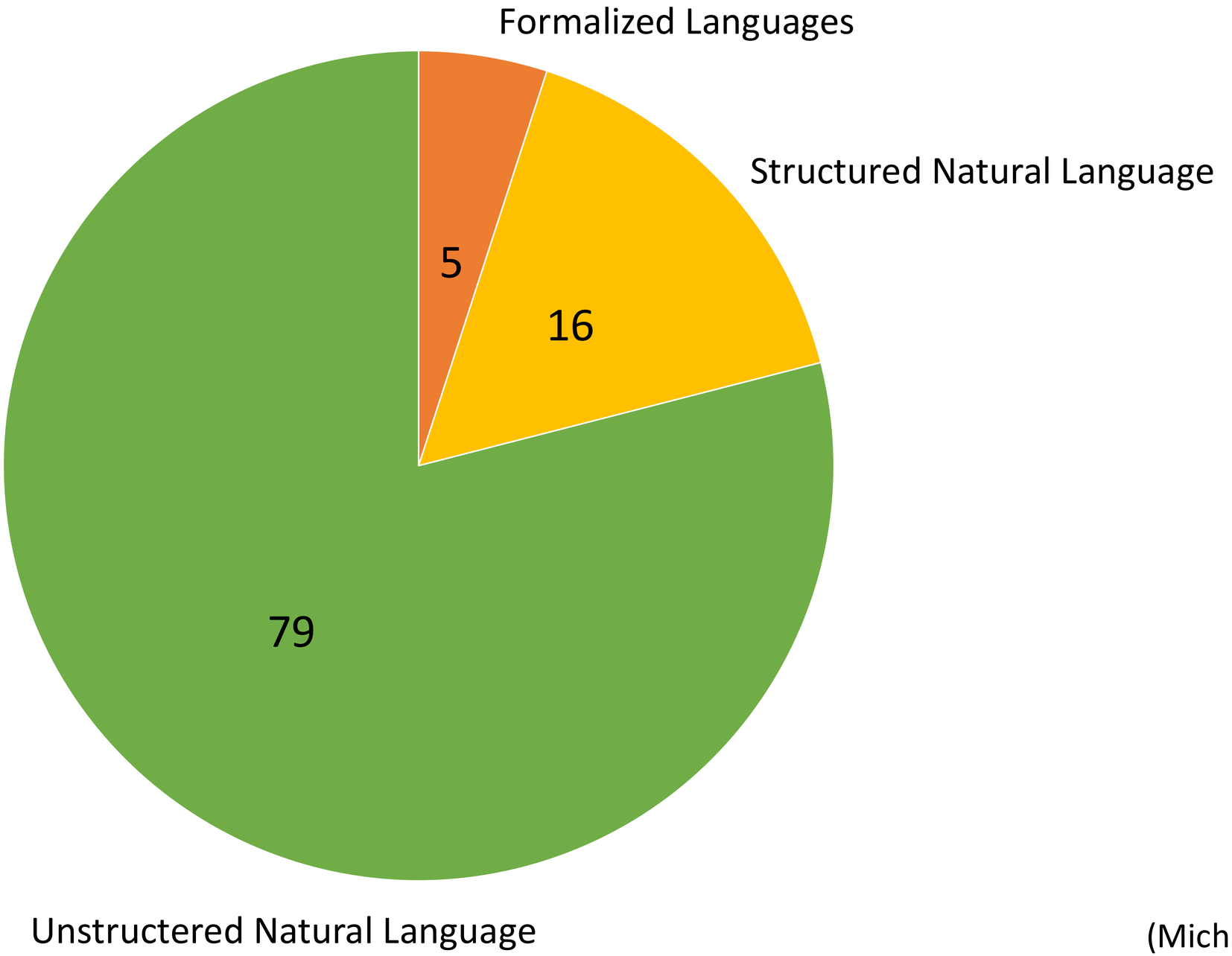
Objectification of requirements into highly objectified code to be performed by a machine.

In how far do the forms and patterns of natural language use matter in Software Engineering?

Software Development Lifecycle

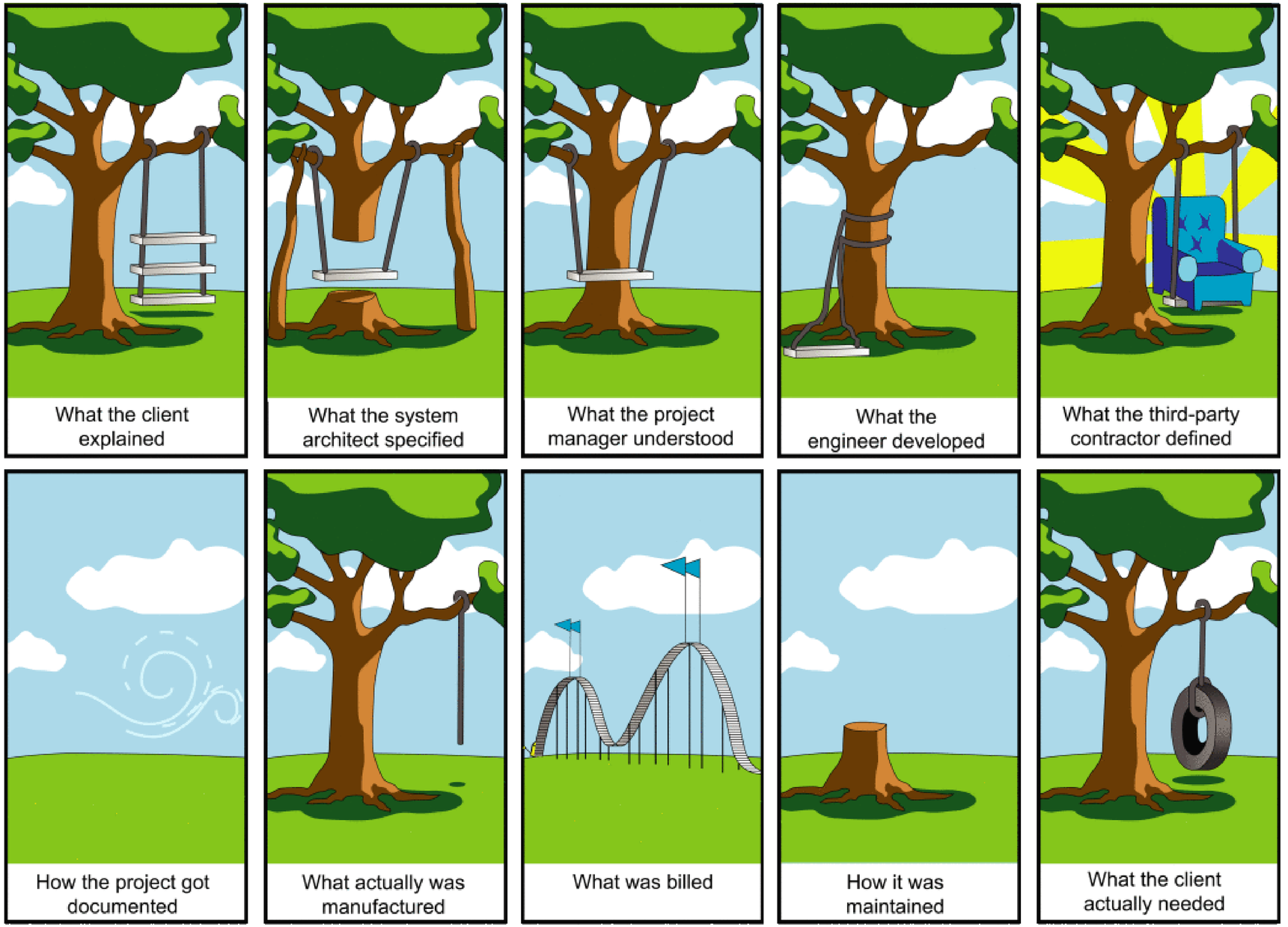


Software Requirements



(Mich2004)

Software Requirements



Explicit Shared Understanding (ESU):
stakeholders all interpret “explicit specifications, such as requirements, design documents, and manuals, in the same way”

Implicit Shared Understanding (ISU):
“the common understanding of non-specified knowledge, assumptions, opinions, and values”
(Glinz & Fricker 2015)

Phil

Explicit Shared Understanding (ESU):
shared interpretation of spoken or written language

- core topics of philosophy: interpretation, meaning, reference, intention, intentionality, world, language ...

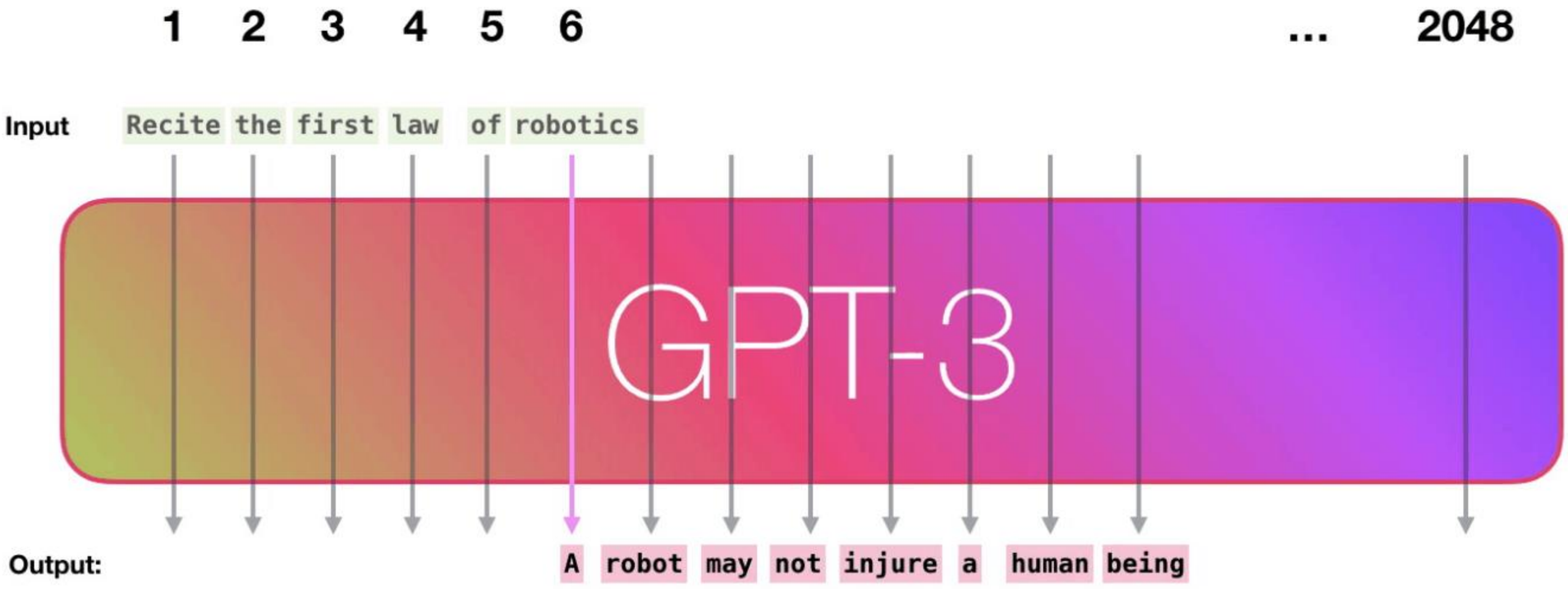
Implicit Shared Understanding (ISU):
the unspoken background of language use, tacit knowledge, common sense

- symbolic language use is only one tip of an iceberg in a vast sea of factors involved in understanding

When we open the black box,
is there understanding anywhere in the LLM?

Are LLMs doing anything beyond
converting one set of numerical patterns into another?

Next Token Prediction



Source: jalammar.github.io

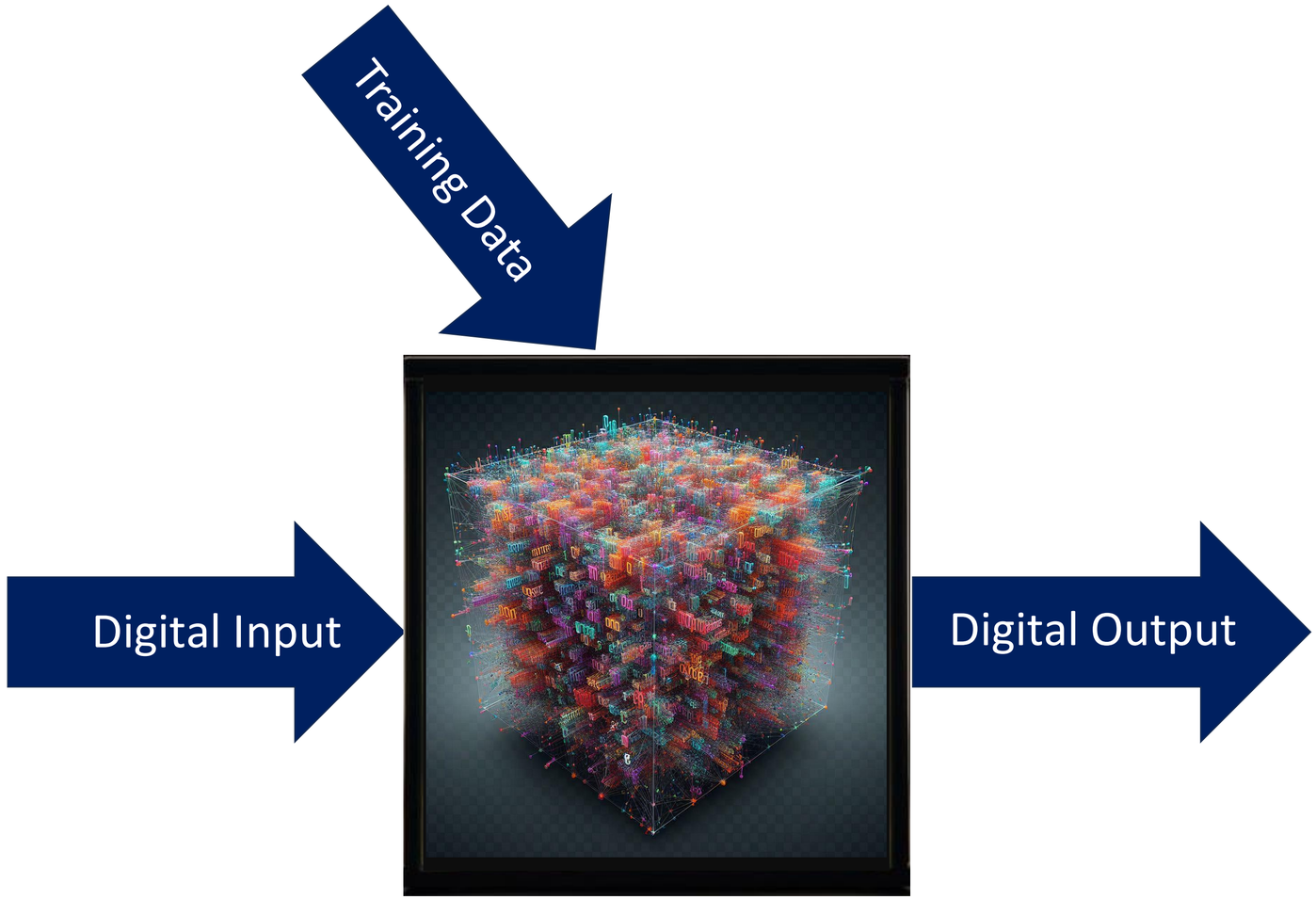
Training Data (GPT-3)

Dataset	Tokens	Weight
Common Crawl (filtered)	410 billion	60%
WebText2	19 billion	22%
Books1	12 billion	8%
Books2	55 billion	8%
Wikipedia	3 billion	3%

GPT-3:
 # parameters = 175 billion
 # layers = 96

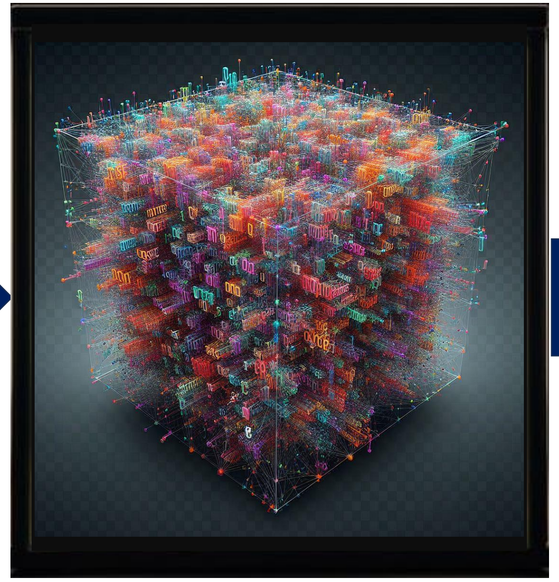
GPT-4:
~= 1.8 trillion
= 120

What is the role of text in human language use?


Training Data
Digital Input
Digital Output

Training Data

Digital Input

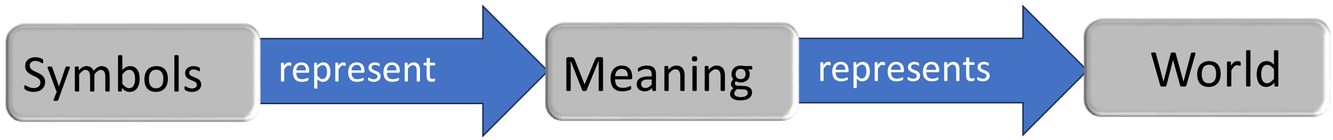


Digital Output

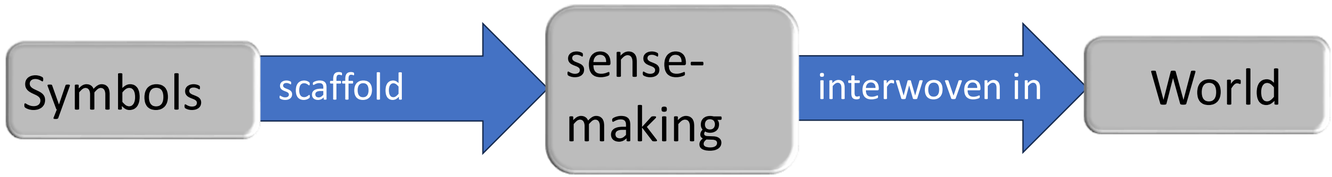
In human language use, writing scaffolds:

- interaction
- the replication of living discourse
- sense-making, understanding
- thought, emotion, and perception

Representational Theories of Meaning

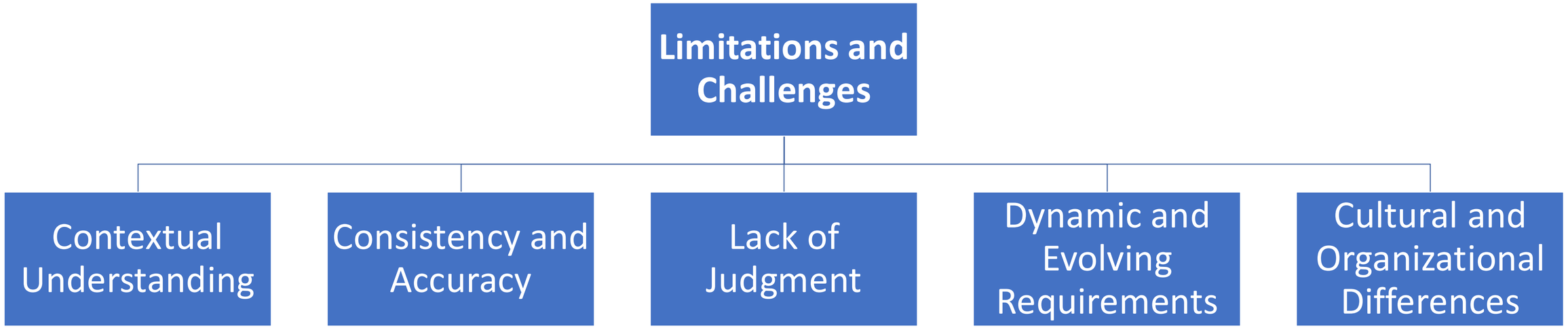


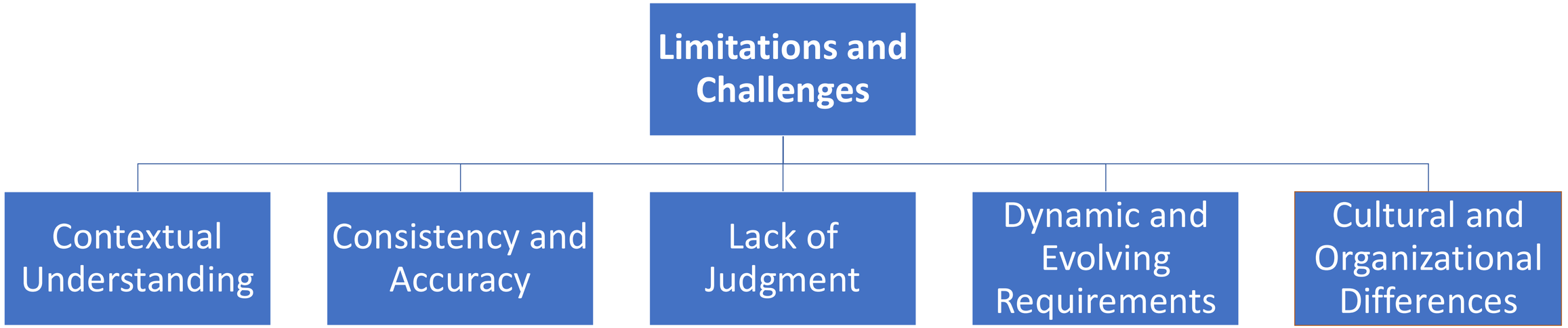
Use Theory of Meaning



What are the current limitations and challenges of LLMs regarding shared understanding?

GPT‘s answers




Limitations and Challenges
Contextual Understanding
Consistency and Accuracy
Lack of Judgment
Dynamic and Evolving Requirements
Cultural and Organizational Differences

Is that all?

Sources

(Glinz & Fricker 2015) Glinz, M., & Fricker, S. A. (2015). On shared understanding in software engineering: an essay. *Computer Science-Research and Development*, *30*, 363-376.

(Mich2004) Mich, L.; Franch, M. & Inverardi, P. N. : Market research for requirements analysis using linguistic tools. Requirements Engineering, Springer, 2004, 9, 40-56. 10.1007/s00766-003-0179-8