

TOWARDS AN EFFICIENT APPROACH FOR CONTROLLABLE TEXT GENERATION

- Name: Iván Martínez Murillo
- @: ivan.martinezmurillo@ua.es
- Supervisors: Paloma Moreda, Elena Lloret
- Research Group: Dept. of Software and Computing Systems, University of Alicante, Spain

INDEX



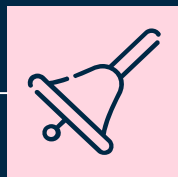
01

INTRODUCTION



02

BACKGROUND



03

LIMITATIONS OF
SOTA MODELS



04

PROPOSAL



05

OPEN RESEARCH
QUESTIONS



06

WHAT COMES
NEXT?

INTRODUCTION

01

01

INTRODUCTION

02

BACKGROUND

03

LIMITATIONS OF
SOTA MODELS

04

PROPOSAL

05

OPEN RESEARCH
QUESTIONS

06

WHAT COMES
NEXT?

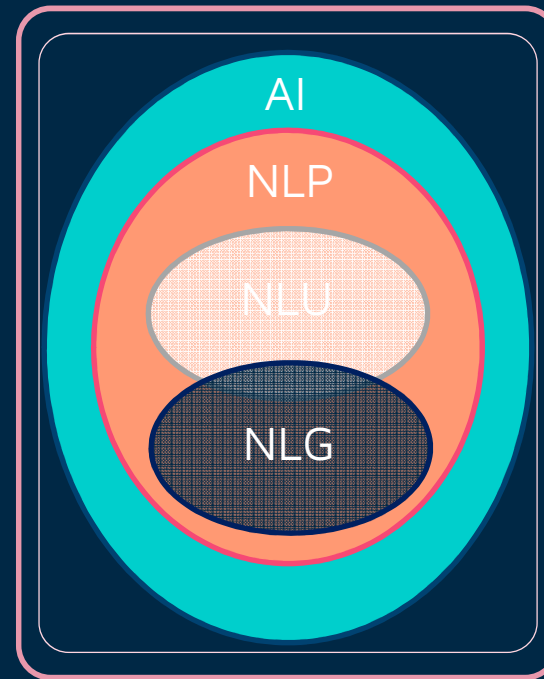
1. INTRODUCTION

NLG: DEFINITION

The process of producing meaningful sentences to meet a communicative goal

1. INTRODUCTION

- AI: Artificial Intelligence
- NLP: Natural Language Processing
- NLU: Natural Language Understanding
- NLG: Natural Language Generation



1. INTRODUCTION

NLG: CLASSIFICATION

T2T
Text-to-Text



D2T
Data-to-Text

N2T
None-to-Text

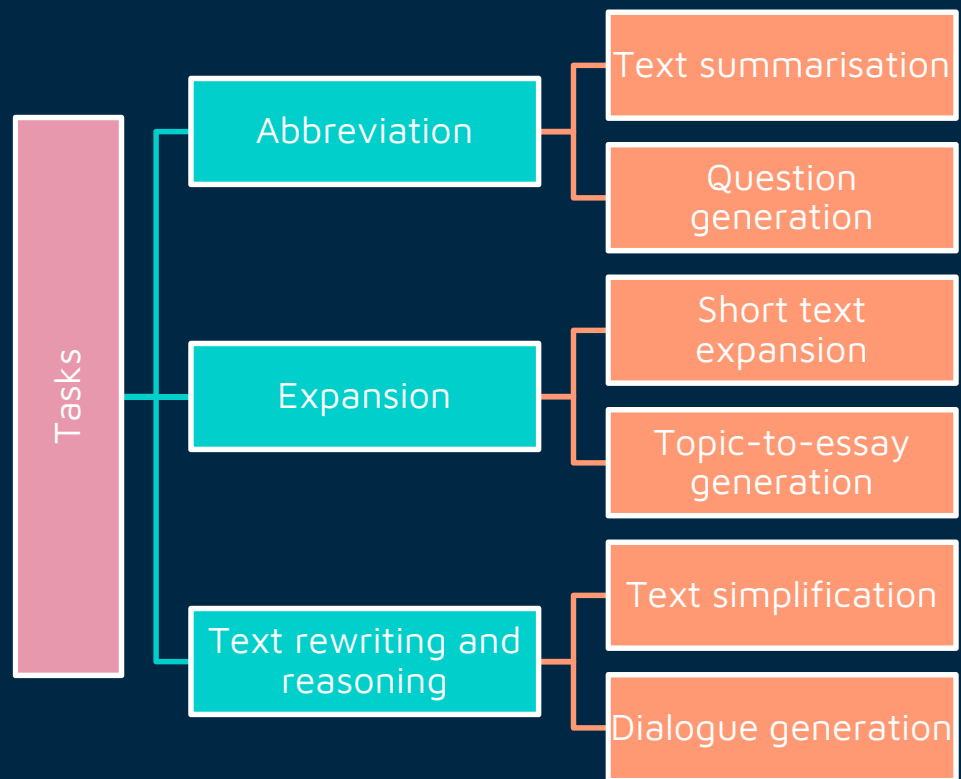


1010
1010



1. INTRODUCTION

NLG: TASKS



BACKGROUND

02

01

INTRODUCTION

02

BACKGROUND

03

LIMITATIONS OF
SOTA MODELS

04

PROPOSAL

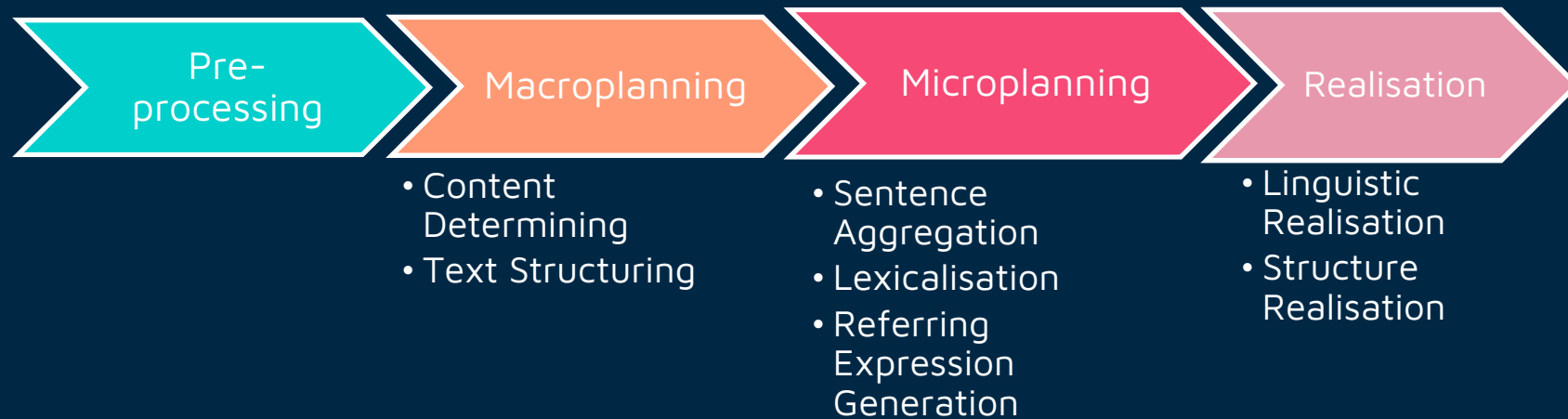
05

OPEN RESEARCH
QUESTIONS

06

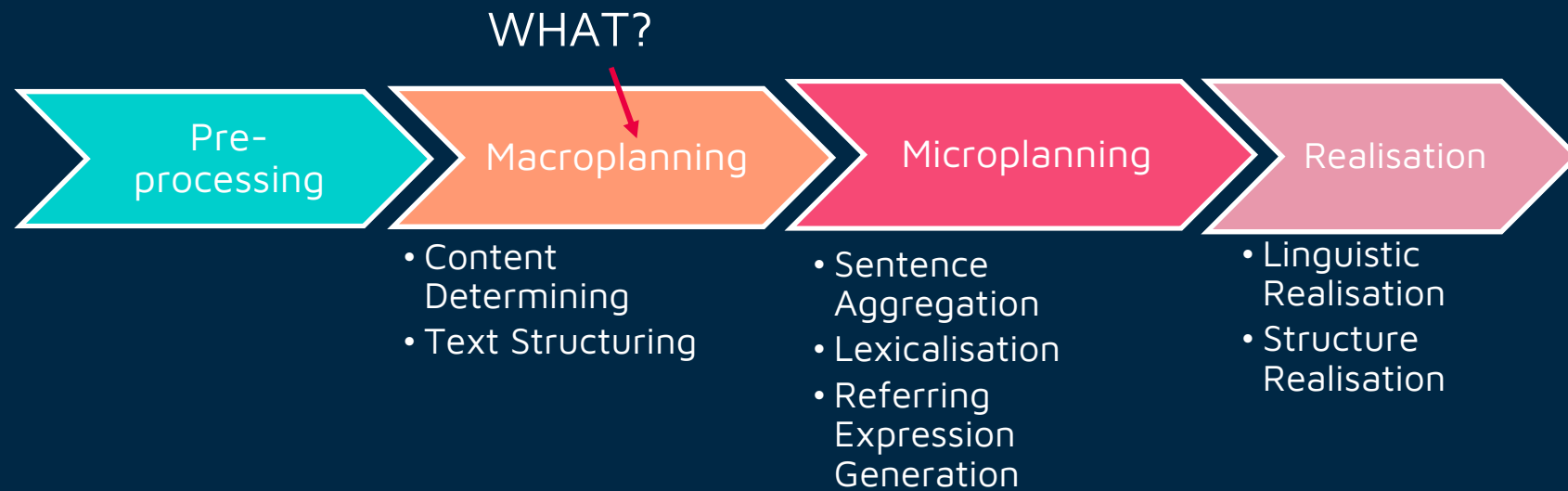
WHAT COMES
NEXT?

2. BACKGROUND



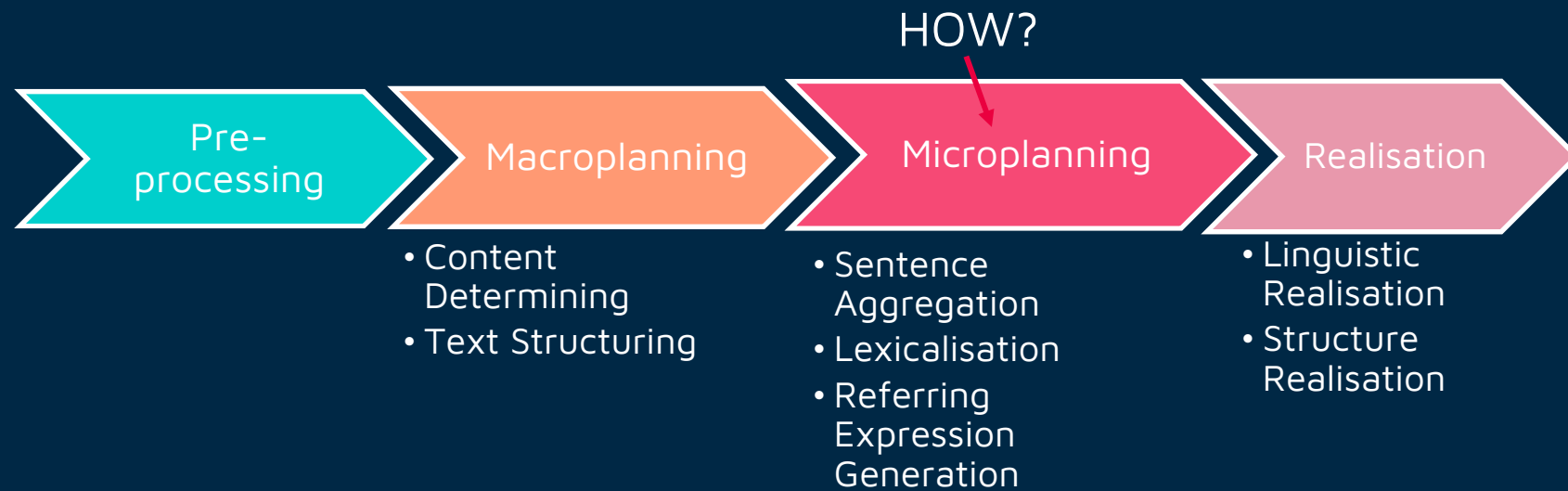
Reiter, E. (1994). Has a consensus NL generation architecture appeared, and is it psycholinguistically plausible?. *arXiv preprint cmp-lg/9411032*.

2. BACKGROUND



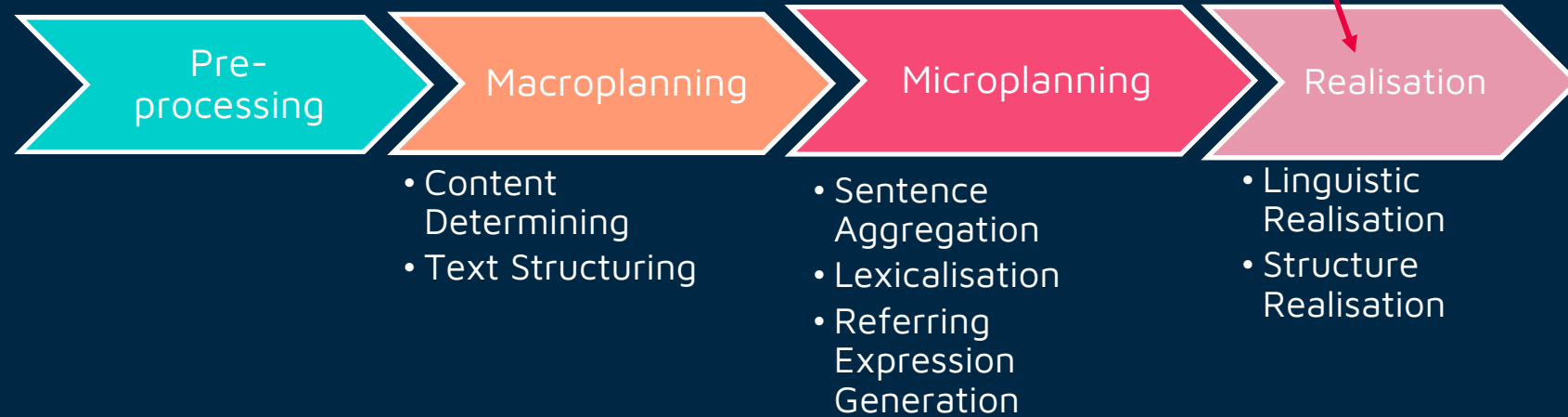
Reiter, E. (1994). Has a consensus NL generation architecture appeared, and is it psycholinguistically plausible?. *arXiv preprint cmp-lg/9411032*.

2. BACKGROUND



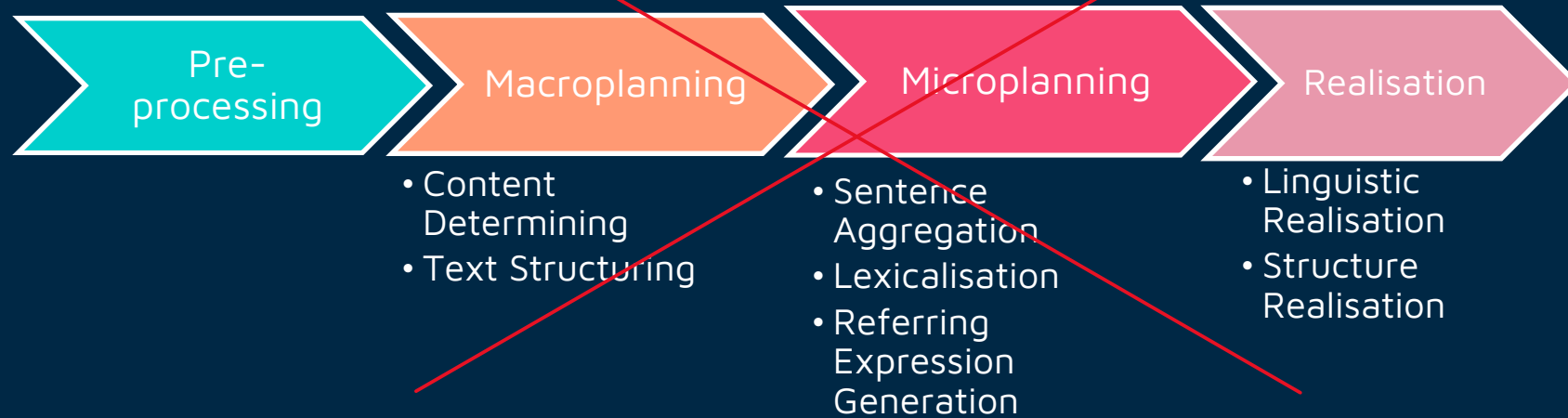
Reiter, E. (1994). Has a consensus NL generation architecture appeared, and is it psycholinguistically plausible?. *arXiv preprint cmp-lg/9411032*.

2. BACKGROUND



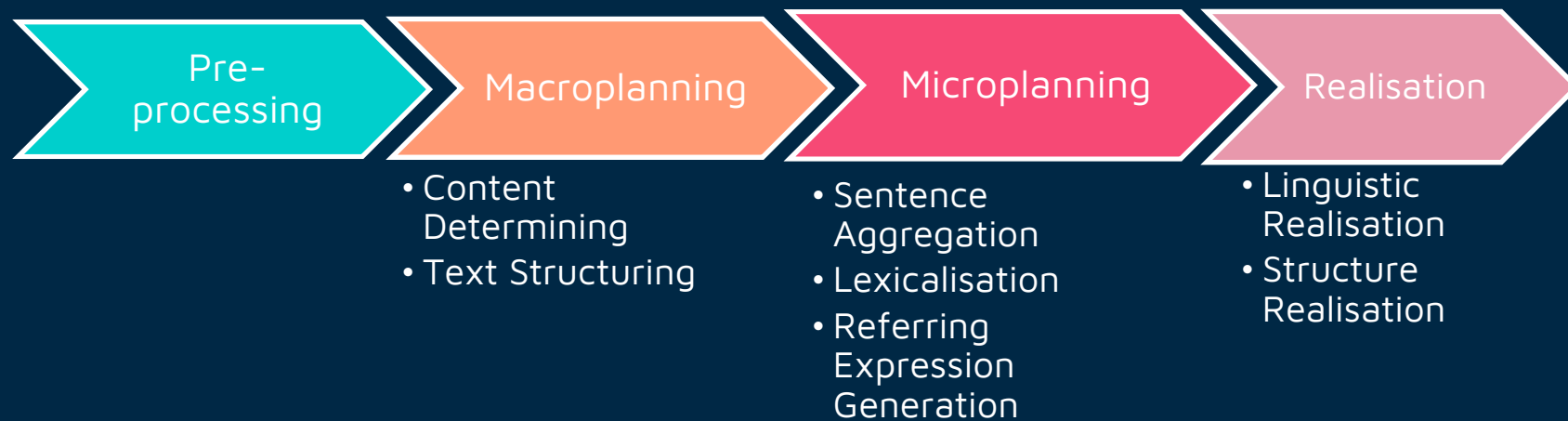
Reiter, E. (1994). Has a consensus NL generation architecture appeared, and is it psycholinguistically plausible?. *arXiv preprint cmp-lg/9411032*.

2. BACKGROUND



Reiter, E. (1994). Has a consensus NL generation architecture appeared, and is it psycholinguistically plausible?. *arXiv preprint cmp-lg/9411032*.

2. BACKGROUND



Reiter, E. (1994). Has a consensus NL generation architecture appeared, and is it psycholinguistically plausible?. *arXiv preprint cmp-lg/9411032*.

2. BACKGROUND

Modular

- Clear division between distinct subtasks, and usually sequentially

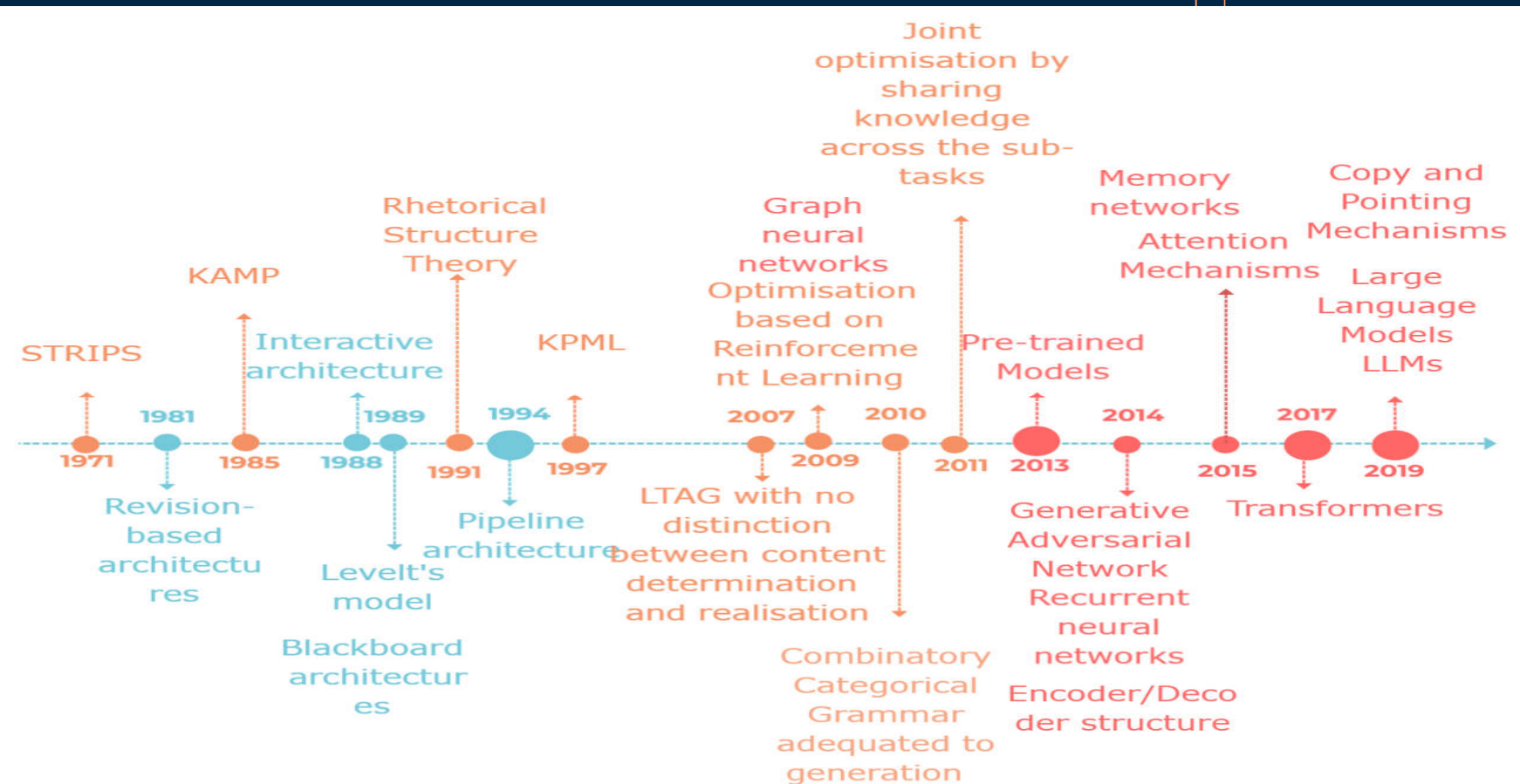
Planning perspectives

- Similar subtask division than modular architectures.
- They allows to perform two or more subtask into the same task.

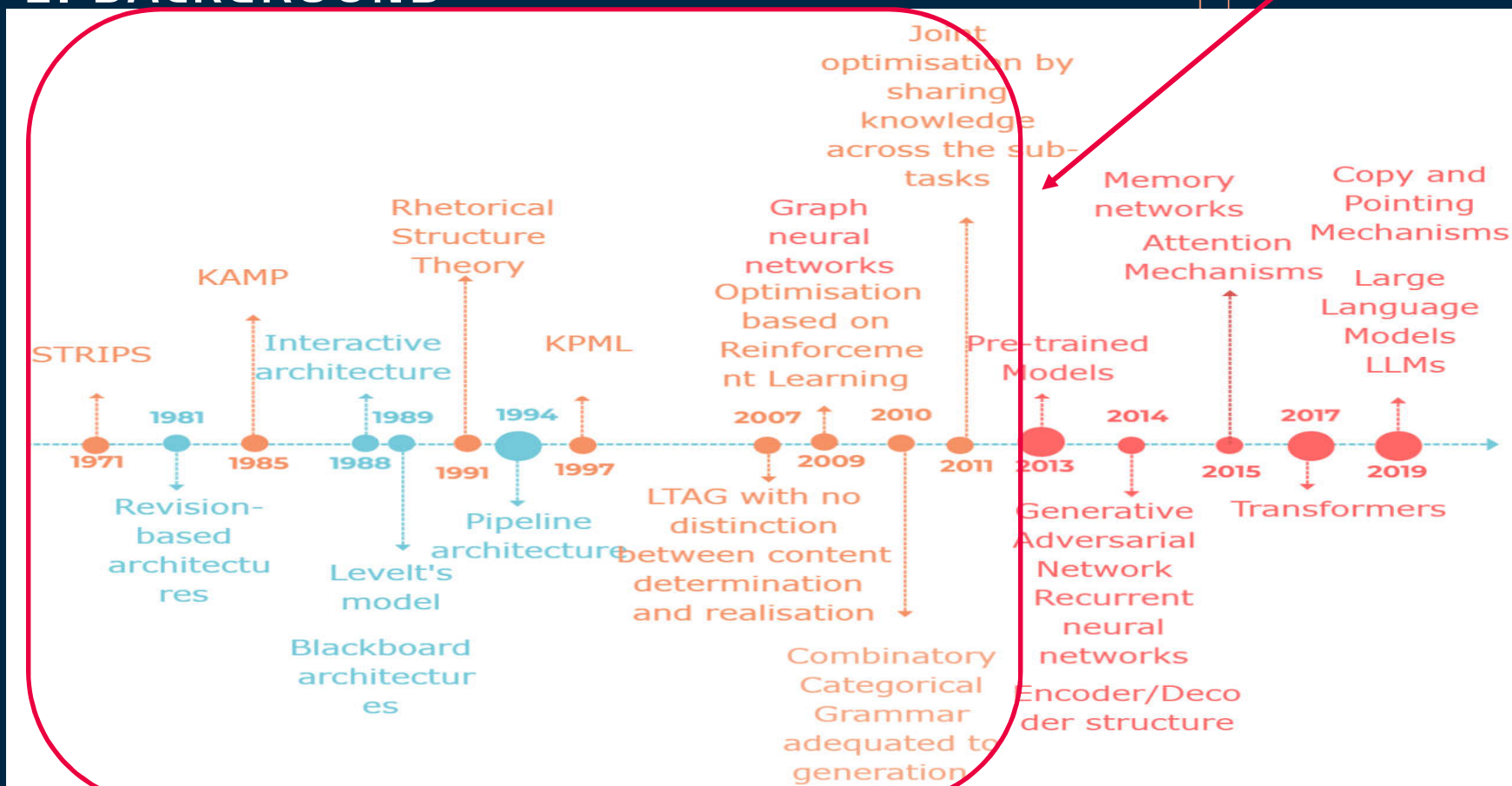
Global approaches

- They hardly establish divisions between the different subtasks.
- Strong reliance in statistical learning.

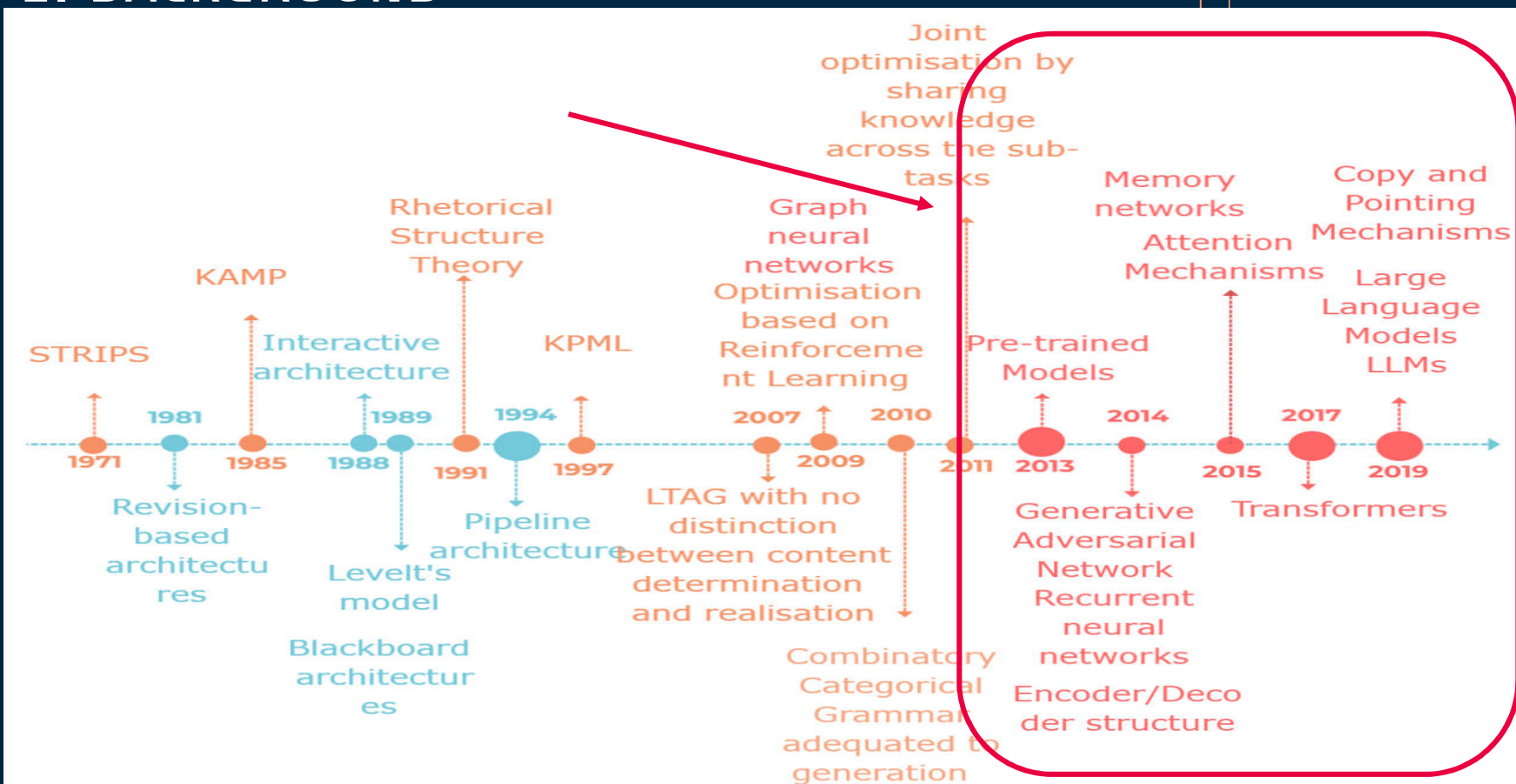
2. BACKGROUND



2. BACKGROUND

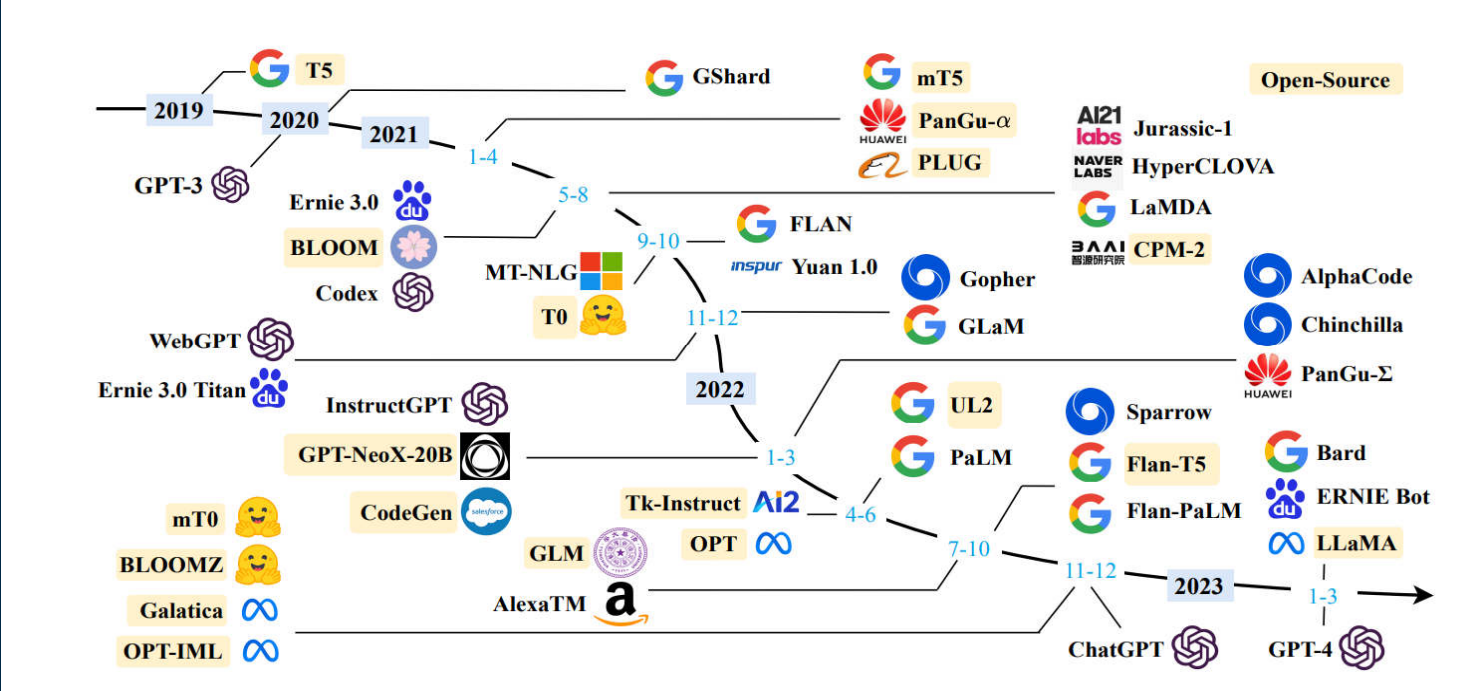


2. BACKGROUND



2. BACKGROUND

Many of the last efforts in the industry have focused on developing LLMs



Zhao, W. X., Zhou, K., Li, J., Tang, T., Wang, X., Hou, Y., Min, Y., Zhang, B., Zhang, J., Dong, Z., Du, Y., Yang, C., Chen, Y., Chen, Z., Jiang, J., Ren, R., Li, Y., Tang, X., Liu, Z., ... Wen, J.-R. (2023). *A Survey of Large Language Models*. <https://arxiv.org/abs/2303.18223v1>

LIMITATIONS OF SOTA MODELS

03

01

INTRODUCTION

02

BACKGROUND

03

LIMITATIONS OF SOTA MODELS

04

PROPOSAL

05

OPEN RESEARCH QUESTIONS

06

WHAT COMES NEXT?

3. LIMITATIONS OF SOTA MODELS

The temporal and economic expense of training these models is extremely high for Academia

01

02

Sometimes these models hallucinate

Generated text can be biased in some cases

03

04

Lack of commonsense and human insight

PROPOSAL

04

01

INTRODUCTION

02

BACKGROUND

03

LIMITATIONS OF
SOTA MODELS

04

PROPOSAL

05

OPEN RESEARCH
QUESTIONS

06

WHAT COMES
NEXT?

4. PROPOSAL

To explore more efficient architectures than SOTA models that could generate text in a

- controllable way and could perform similar to LLMs

- The main hypothesis is that implementing **controllable** generation techniques along with external **commonsense** knowledge can help to **reduce the hallucination** problem

OPEN RESEARCH QUESTIONS

05

01

INTRODUCTION

02

BACKGROUND

03

LIMITATIONS OF
SOTA MODELS

04

PROPOSAL

05

OPEN RESEARCH
QUESTIONS

06

WHAT COMES
NEXT?

5. OPEN RESEARCH QUESTIONS

What is controllable text generation, and what are the most common techniques to address it?

Generating texts whose **attributes** (stylistics, based on **demographic** attributes, or based on the **content**) can be **controlled**.

Three strategies:

1. Via hyperparameters
2. Via additional input
3. Via conditional training

5. OPEN RESEARCH QUESTIONS

What is hallucination, what causes hallucination, and which are the best ways to mitigate it?

A text generated that is nonsensical or unfaithful to the source input. Can be intrinsic and extrinsic.

5. OPEN RESEARCH QUESTIONS

What is hallucination, **what causes hallucination**, and which are the best ways to mitigate it?

During the construction of **datasets**, and during the **training** of the models.

5. OPEN RESEARCH QUESTIONS

What is hallucination, what causes hallucination, and which are the best ways to mitigate it?

Create a faithful dataset, alter the encoder/decoder structure, propose an optimal training strategy and include external commonsense knowledge.

5. OPEN RESEARCH QUESTIONS

Is it possible to obtain an architecture that performs equally to LLMs without being as computationally demanding as them?

The time and computational **expense** needed to train LLMs are **inaccessible** to academia.

Others architectures such as **Plug and Play** models, or **Variational Autoencoders** are much more **efficient** than LLMs. But **can they perform** equally to them?

5. OPEN RESEARCH QUESTIONS

Is there a task-agnostic architecture able to perform well for different tasks?

Most researches in the NLG area are focused on a specific task. They perform correctly in one task but they underperform in the others.

This study will analyse task-agnostic architectures to propose a model that can achieve a high performance at every task.

WHAT COMES NEXT?

06

01

INTRODUCTION

02

BACKGROUND

03

LIMITATIONS OF SOTA MODELS

04

PROPOSAL

05

OPEN RESEARCH QUESTIONS

06

WHAT COMES NEXT?

BIBLIOGRAPHY

- Zhao, W. X., Zhou, K., Li, J., Tang, T., Wang, X., Hou, Y., Min, Y., Zhang, B., Zhang, J., Dong, Z., Du, Y., Yang, C., Chen, Y., Chen, Z., Jiang, J., Ren, R., Li, Y., Tang, X., Liu, Z., ... Wen, J.-R. (2023). *A Survey of Large Language Models*. <https://arxiv.org/abs/2303.18223v1>
- Gatt, A., & Krahmer, E. (2018). Survey of the State of the Art in Natural Language Generation: Core tasks, applications and evaluation. In *Journal of Artificial Intelligence Research* (Vol. 61). <https://www.narrativescience.com>
- Yu, W., Zhu, C., Li, Z., Hu, Z., Wang, Q., Ji, H., & Jiang, M. (2022). A Survey of Knowledge-enhanced Text Generation. *ACM Computing Surveys*, 54(11s), 1–38. <https://doi.org/10.1145/3512467>
- Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., Ishii, E., Bang, Y., Dai, W., Madotto, A., & Fung, P. (2022). *Survey of Hallucination in Natural Language Generation*. <https://doi.org/10.1145/3571730>
- Tang, C., Guerin, F., Li, Y., & Lin, C. (2022). *Recent Advances in Neural Text Generation: A Task-Agnostic Survey*. <http://arxiv.org/abs/2203.03047>
- Chandu, K. R., & Black, A. W. (2020). *Positioning yourself in the maze of Neural Text Generation: A Task-Agnostic Survey*. <http://arxiv.org/abs/2010.07279>



Do you have any questions?

Ivan.martinezmurillo@ua.es

<https://gplsi.dlsi.ua.es/>

THANK YOU FOR YOUR ATTENTION



This work has been funded by the European Commission ICT COST Action "Multi-task, Multilingual, Multi-modal Language Generation" (CA18231).



The research work conducted is part of the R&D projects "CORTEX: Conscious Text Generation" (PID2021-123956OB-I00), funded by MCIN/ AEI/10.13039/501100011033/ and by "ERDF A way of making Europe"; 34