

# Commonsense Knowledge and Controllable Techniques for an Effective and Efficient Approach to Text Generation

Iván Martínez-Murillo

*Dept. of Software and Computing Systems, University of Alicante, Apdo. de Correos 99, E-03080, Alicante, Spain*

## Abstract

The Natural Language Generation (NLG) field has advanced at a breakneck speed, favoured by the development of Large Language Models (LLMs). Notwithstanding, these models also have some drawbacks. On the one hand, these models can introduce some risks such as hallucination or bias which can be used in an unethical way to potentially generate dis- and mis-information. On the other hand, the expense of time and cost of training these models is too high. In account of this, the purpose of this paper is to propose a new research line for my PhD thesis. During the research, I will propose an efficient architecture, that could generate quality text in a controllable way, while integrating external commonsense knowledge. The objective is that this proposed architecture could achieve similar performance to state-of-the-art models while being more efficient.

## Keywords

Natural Language Generation, Controllable techniques, Hallucination, Efficient architectures, Task-agnostic, Commonsense Knowledge

## 1. Justification of the research

The rapid development of generative Artificial Intelligence (AI) has caused an augment of interest in society in AI tools. These tools can produce a positive impact in lots of areas, saving the time and effort of solving some tasks [1, 2, 3].

In particular, state-of-the-art Natural Language Generation (NLG) tools can produce text that, in some cases, can be indistinguishable from human-generated texts. This could have lots of benefits in some sectors such as academia, tourism or marketing [4]. Nonetheless, these tools also have some drawbacks. First of all, text generated by these tools may contain hallucinations, which is the phenomena that occur when a text is nonsensical or unfaithful to the provided source [5]. Secondly, AI-generated text could be biased in some cases, which is the misrepresentation or attribution errors that result in favouring certain groups or ideas [6]. Finally, these tools also lack of logical reasoning, a fact that it is essential to human intelligence [7]. In the wake of these limitations, these tools can be used in a bad and unethical way to potentially generate dis- and mis-information.

Moreover, the core of these tools are Large Language Models (LLMs). The expense of time and cost needed to train these models is extremely high, being only within the reach of large

---

*Doctoral Symposium on Natural Language Processing from the Proyecto ILENIA, 28 September 2023, Jaén, Spain.*

✉ [ivan.martinezmurillo@ua.es](mailto:ivan.martinezmurillo@ua.es) (I. Martínez-Murillo)



© 2023 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

 CEUR Workshop Proceedings (CEUR-WS.org)

companies.

Therefore, the motivation for the present research arises from the need in the academia to find efficient architectures that could produce text in a controlled manner, achieving a similar performance to state-of-the-art models, but solving the hallucination issue.

The remainder of this article is organised as follows: Section 2 presents an overview of the relevant literature concerning NLG; Section 3 shows the main hypotheses and objectives planned for this research; Finally, Section 4, and Section 5 detail the methodology this PhD will follow and some relevant research topics for discussion.

## 2. Background and Related Work

Before introducing my proposal, this section aims to contextualise this study within the state of the art of the NLG.

NLG is the subfield in the Natural Language Processing (NLP) area that aims to produce meaningful sentences to meet a communicative goal [8]. Depending on several aspects of the generation, NLG can be classified according to two criteria:

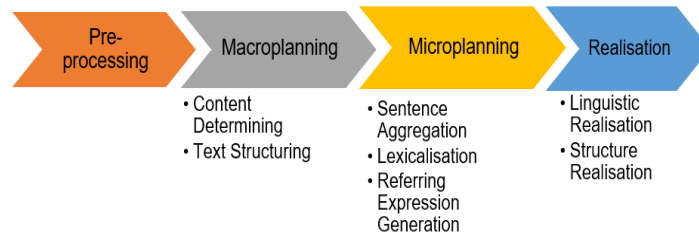
- **Type of input:** Depending on the type of input, NLG can be catalogued as (1) text-to-text generation (T2T) and (2) data-to-text generation (D2T) [9]. In D2T, input data can assume different types such as binary data, images, voice, database, ontologies, etc. Recently, another concept of NLG has emerged, (3) none-to-text generation (N2T) [10], which corresponds to the generation to which no input is received.
- **Task typology:** Based on the communicative goal, NLG can be grouped into (1) text abbreviation; (2) text expansion; and (3) text rewriting and reasoning. *Text abbreviation* tasks consist in detecting the most important information in a text and fusing that information into a short text, e.g. text summarisation. *Text expansion* tasks aim to generate complete sentences from some meaningful words, e.g. topic-to-essay. Finally, *Text rewriting and reasoning* tasks try to rewrite a text into another style or apply reasoning methods, e.g. text simplification.

To achieve the communicative goal of these tasks, the NLG area has been studied for a long time. First researches date by the end of 1970 [11]. Notwithstanding, it has not been until recent years that the NLG field has achieved an exponential improvement, producing text in a very similar way to humans. But, how did we get to this?

In a first stage, the NLG task was seen as a sequential scheme of four different stages (pre-processing, macroplanning, microplanning and realisation). **Modular architectures** followed this scheme, making a clear distinction between the distinct sub-tasks of each stage. The most famous modular architecture was proposed by Reiter [12]. Figure 1 shows the sub-task division in this architecture.

Other works within this architecture can be found in [13, 14, 15, 16].

Later, that clear distinction between the distinct sub-tasks became more flexible originating what is known as **planning perspectives**. This scheme was similar to the employed in modular architectures, but it allows to combine and implement two or more different sub-tasks in one



**Figure 1:** Sub-task division in the modular architecture for the stages proposed by Reiter [8]

sub-task, e.g. to combine text structuring and sentence aggregation sub-tasks. Some examples of this approach are present in [17, 18, 19, 20, 21, 22, 23, 24].

Finally, the sub-task division started to disappear, originating **global approaches**. This type of architecture does not make a distinction among sub-tasks, performing every task as a whole, and relying on statistical learning and neural networks. Some proposed architectures within global approaches are: Graph Neural Networks [25], Generative Adversarial Nets [26], Recurrent Neural Networks [27], Pre-trained Models [28], Memory Networks [29], Transformers [30] and Copy and Pointing Mechanism [31]. This group of approaches have made the major development in the NLG area. The most important proposal in this group was the Transformers architecture and its concept of attention. Models based on this architecture achieve a high performance at NLG tasks. The best-performing models based on Transformers are LLMs such as GPT4 [32] or LLaMa [33], which have neural networks with billions of parameters. Nowadays, most of the research in the industry is focused on developing bigger LLMs, as it is thought that a bigger LLM would achieve better performance. The cost and time of training these models are unassumable for the academia. On account of that issue, there is a need in the academia to find more efficient architectures that could perform similarly to LLMs.

Consequently, my line of work will focus on exploring efficient architectures that could generate text with similar results to state-of-the-art models. Moreover, controllable generation methods, techniques to integrate external commonsense knowledge and task-agnostic architectures will be studied in order to reduce the phenomena known as hallucination.

### 3. Main Hypothesis and Objectives

This PhD thesis is based on the hypothesis that integrating external commonsense knowledge along with controllable text generation techniques in an efficient architecture will help to reduce the hallucination issue, and besides performing similarly to state-of-the-art models. Thus, the main objective of this research is the proposal of an efficient architecture that could achieve a good performance in different NLG tasks, e.g. text summarisation, and text simplification, and could reduce hallucination as much as possible. In order to complete this main objective, several sub-objectives have been proposed:

- **A1.** To explore optimal controllable text generation techniques.

- **A2.** To examine hallucination mitigation techniques.
- **A3.** To study how to integrate external commonsense knowledge.
- **A4.** To analyse and test different task-agnostic architectures incorporating the previously studied techniques.
- **B1.** To compare the performance of open-source state-of-the-art architectures using a common benchmark.
- **B2.** To propose a cost-effective architecture that can generate text in a controllable way and evaluate it.
- **C1.** To adapt the proposed architecture to perform in some NLG tasks, e.g., summarisation or text simplification.

The planned schedule of these sub-objectives can be seen in Figure 2, starting from February 2023. **Group A** corresponds with the study and test of state-of-the-art techniques. After an initial study, during **Group B**, an efficient architecture will be proposed, tested and compared with other open-source architectures using a common benchmark. Finally, in **Group C** the proposed architecture will be adapted to perform in different NLG tasks.

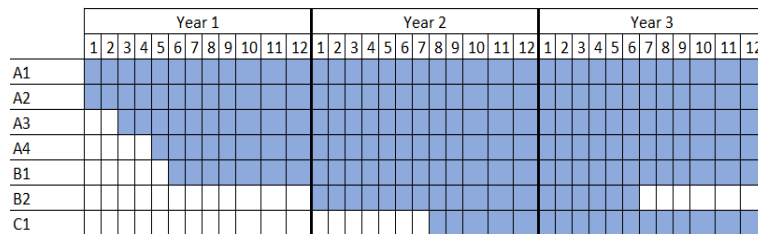


Figure 2: PhD project schedule

## 4. Methodology and proposed experiment

The proposed methodology to carry out this research is based on a complete and comprehensive training in all areas of NLG, including general training on NLP. After having the basic notions of NLG, the research focuses on an exhaustive analysis of the state of the art of NLG, especially on deep learning techniques that allow controlled language generation and integrate commonsense knowledge. Subsequently, the experimentation also starts, testing different open-source architectures, along with the most relevant studied techniques. After having tested several architectures, an efficient base model will be proposed, integrating commonsense knowledge and controllable generation techniques into it. Then, it will be evaluated against other architectures using a common benchmark. Finally, the proposed architecture will be adapted to perform different tasks.

At present, I am experimenting with the CommonGen dataset [34]. The CommonGen dataset consists of a set of common concepts and some reference sentences using those concepts and the main idea is to test machines for the ability of generative commonsense reasoning. I am testing with different types of approaches such as SimpleNLG, Factorised Language Models, or

Neural Models over this dataset. With the proposed experiment, the main idea is to combine the best-obtained architecture with controllable generation techniques in order to obtain a base model.

## 5. Research issues to discuss

In order to advance towards an effective and efficient approach for controllable text generation, several research issues are suggested and briefly discussed.

**What does controllable text generation mean, and what are the most efficient methods to incorporate it?** Controllable text generation is the task of producing text in a way that its attributes can be controlled [35]. These attributes can adopt a wide variety of ranges, such as stylistics, to include specific information in the content, based on the demographic attributes of the interlocutor, etc. As seen in [36], there are three ways to approach controllable text generation.

**1. Via hyperparameters:** Training data in LLMs can be unbalanced due to the fact that it is difficult to balance that huge amount of data. Modifying hyperparameters may generalise the knowledge better and consequently raise obtained results.

**2. Via additional input:** To fine-tune a pre-trained model with more information than just the text could enhance its performance.

**3. Via conditional training:** Using internal control variables could enrich the generation with specific capabilities.

**What is hallucination and what are the ways to reduce its occurrence?** Hallucination in NLG occurs when a text generated by an AI lacks of coherence or deviates from the intended sense of the source input [5]. It can be classified into two categories: *intrinsic hallucinations*, which appear when the generated text contradicts the source input, and *extrinsic hallucinations* which arise when the source input cannot substantiate the generated text.

There exist different types of approaches to minimise the occurrence of hallucinations. Firstly, constructing a reliable dataset, which does not contain any type of contradiction in the data. Secondly, modifying the encoder/decoder architecture can enhance the ability to better understand and represent the knowledge. Thirdly, proposing an optimal training strategy such as controllable text generation could benefit the model. Finally, one important approach is to integrate external commonsense knowledge into the models.

**How to integrate external commonsense knowledge?** Commonsense knowledge is an important factor in human communication, as it facilitates inference without the explicit mention of context [37]. Although current state-of-the-art models exhibit some common sense abilities, it is not complete yet. Traditionally, commonsense has been injected into NLG systems in the form of rules and ontologies. Nowadays, the approaches have focused on injecting commonsense into neural NLG models through pre-trained models and using commonsense graphs. But there is still much work to do in this field in order to reach a complete commonsense knowledge.

**Can a smaller architecture obtain similar performance than LLMs?** There are some structures such as Plug and Play models or Variational Autoencoders that are more efficient than LLMs. Integrating commonsense knowledge and controllable generation techniques into these models could help to perform like LLMs while being smaller and more efficient models.

## Acknowledgements

This research work 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”

## References

- [1] W. P. Walters, M. Murcko, Assessing the impact of generative ai on medicinal chemistry, *Nature biotechnology* 38 (2020) 143–145.
- [2] S. Mayahi, M. Vidrih, The impact of generative ai on the future of visual content marketing, *arXiv preprint arXiv:2211.12660* (2022).
- [3] G. Cooper, Examining science education in chatgpt: An exploratory study of generative artificial intelligence, *Journal of Science Education and Technology* (2023) 1–9.
- [4] Y. K. Dwivedi, N. Kshetri, L. Hughes, E. L. Slade, A. Jeyaraj, A. K. Kar, A. M. Baabdullah, A. Koochang, V. Raghavan, M. Ahuja, et al., “so what if chatgpt wrote it?” multidisciplinary perspectives on opportunities, challenges and implications of generative conversational ai for research, practice and policy, *International Journal of Information Management* 71 (2023) 102642.
- [5] Z. Ji, N. Lee, R. Frieske, T. Yu, D. Su, Y. Xu, E. Ishii, Y. J. Bang, A. Madotto, P. Fung, Survey of hallucination in natural language generation, *ACM Comput. Surv.* 55 (2023). URL: <https://doi.org/10.1145/3571730>. doi:10.1145/3571730.
- [6] E. Ferrara, Should chatgpt be biased? challenges and risks of bias in large language models, *arXiv preprint arXiv:2304.03738* (2023).
- [7] H. Liu, R. Ning, Z. Teng, J. Liu, Q. Zhou, Y. Zhang, Evaluating the logical reasoning ability of chatgpt and gpt-4, *arXiv preprint arXiv:2304.03439* (2023).
- [8] E. Reiter, R. Dale, Building applied natural language generation systems, *Natural Language Engineering* 3 (1997) 57–87. doi:10.1017/S1351324997001502.
- [9] M. Vicente, C. Barros, F. S. Peregrino, F. Agulló, E. Lloret, La generación de lenguaje natural: análisis del estado actual, *Computación y Sistemas* 19 (2015) 721–756.
- [10] K. R. Chandu, A. W. Black, Positioning yourself in the maze of neural text generation: A task-agnostic survey, 2020. URL: <https://arxiv.org/abs/2010.07279>. doi:10.48550/ARXIV.2010.07279.
- [11] D. D. McDonald, Natural language generation., *Handbook of natural language processing* 2 (2010) 121–144.
- [12] E. Reiter, Has a consensus nl generation architecture appeared, and is it psycholinguistically plausible?, 1994. *arXiv:cmp-lg/9411032*.
- [13] W. C. Mann, J. A. Moore, Computer generation of multiparagraph english text, *American Journal of Computational Linguistics* 7 (1981) 17–29.
- [14] E. Hovy, Generating natural language under pragmatic constraints, *Journal of Pragmatics* 11 (1987) 689–719.
- [15] W. Levelt, *Speaking: From intention to articulation* mit press, Cambridge, MA (1989).

- [16] S. Nirenburg, V. R. Lesser, E. Nyberg, Controlling a language generation planner., in: IJCAI, 1989, pp. 1524–1530.
- [17] R. E. Fikes, N. J. Nilsson, Strips: A new approach to the application of theorem proving to problem solving, *Artificial intelligence* 2 (1971) 189–208.
- [18] D. Appelt, *Planning english sentences*. cambridge university press, 1985.
- [19] E. H. Hovy, *Approaches to the planning of coherent text*, Springer, 1991.
- [20] J. A. Bateman, Enabling technology for multilingual natural language generation: the kpml development environment, *Natural Language Engineering* 3 (1997) 15–55.
- [21] A. Koller, M. Stone, Sentence generation as a planning problem, in: *Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics*, Association for Computational Linguistics, Prague, Czech Republic, 2007, pp. 336–343. URL: <https://aclanthology.org/P07-1043>.
- [22] V. Rieser, O. Lemon, Natural language generation as planning under uncertainty for spoken dialogue systems, *Empirical Methods in Natural Language Generation: Data-oriented Methods and Empirical Evaluation* (2009) 105–120.
- [23] C. Nakatsu, M. White, Generating with discourse combinatory categorial grammar, *Linguistic Issues in Language Technology* 4 (2010).
- [24] O. Lemon, Learning what to say and how to say it: Joint optimisation of spoken dialogue management and natural language generation, *Computer Speech & Language* 25 (2011) 210–221.
- [25] F. Scarselli, M. Gori, A. C. Tsoi, M. Hagenbuchner, G. Monfardini, The graph neural network model, *IEEE transactions on neural networks* 20 (2008) 61–80.
- [26] M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, I. J. Goodfellow, J. Pouget-Abadie, Generative adversarial nets, *Advances in neural information processing systems* 27 (2014) 2672–2680.
- [27] I. Sutskever, O. Vinyals, Q. V. Le, Sequence to sequence learning with neural networks, *Advances in neural information processing systems* 27 (2014).
- [28] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, J. Dean, Distributed representations of words and phrases and their compositionality, *Advances in neural information processing systems* 26 (2013).
- [29] S. Sukhbaatar, J. Weston, R. Fergus, et al., End-to-end memory networks, *Advances in neural information processing systems* 28 (2015).
- [30] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, I. Polosukhin, Attention is all you need, 2017. [arXiv:1706.03762](https://arxiv.org/abs/1706.03762).
- [31] A. See, P. J. Liu, C. D. Manning, Get to the point: Summarization with pointer-generator networks, *arXiv preprint arXiv:1704.04368* (2017).
- [32] OpenAI, Gpt-4 technical report, 2023. [arXiv:2303.08774](https://arxiv.org/abs/2303.08774).
- [33] H. Touvron, T. Lavril, G. Izacard, X. Martinet, M.-A. Lachaux, T. Lacroix, B. Rozière, N. Goyal, E. Hambro, F. Azhar, A. Rodriguez, A. Joulin, E. Grave, G. Lample, Llama: Open and efficient foundation language models, 2023. [arXiv:2302.13971](https://arxiv.org/abs/2302.13971).
- [34] B. Y. Lin, W. Zhou, M. Shen, P. Zhou, C. Bhagavatula, Y. Choi, X. Ren, CommonGen: A constrained text generation challenge for generative commonsense reasoning, in: *Findings of the Association for Computational Linguistics: EMNLP 2020*, Association for Computational Linguistics, Online, 2020, pp. 1823–1840. URL: <https://www.aclweb.org/>



anthology/2020.findings-emnlp.165.

- [35] S. Prabhunoye, A. W. Black, R. Salakhutdinov, Exploring controllable text generation techniques, in: Proceedings of the 28th International Conference on Computational Linguistics, International Committee on Computational Linguistics, Barcelona, Spain (Online), 2020, pp. 1–14. URL: <https://aclanthology.org/2020.coling-main.1>. doi:10.18653/v1/2020.coling-main.1.
- [36] E. Erdem, M. Kuyu, S. Yagcioglu, A. Frank, L. Parcalabescu, B. Plank, A. Babii, O. Turuta, A. Erdem, I. Calixto, et al., Neural natural language generation: A survey on multilinguality, multimodality, controllability and learning, *Journal of Artificial Intelligence Research* 73 (2022) 1131–1207.
- [37] S. Mahamood, M. Clinciu, D. Gkatzia, It’s common sense, isn’t it? demystifying human evaluations in commonsense-enhanced nlg systems (2021).