From Product Sheet to Text and Video: A NLG Pipeline to Transform Structured Data into Comprehensive Descriptions*

Andrea Avignone[1,*,†], Alessandro Fiori[1,†], Silvia Chiusano[1,†] and Giuseppe Rizzo[2,*,†]

[1]Dept. of Control and Computer Engineering, Politecnico di Torino, Corso Duca degli Abruzzi, 24, Torino, 10129, Italy
[2]LINKS Foundation, Torino, Italy

Abstract

The recent improvement of powerful Large Language Models is the key to automatically produce satisfactory written and spoken language, in contrast to the constraints of conventional template-based solutions. However, the most advanced models can be costly and complex to integrate into practical applications, especially in business contexts where the output quality significantly matters. This study presents a tailored pipeline for data-to-text and text-to-speech generation, primarily harnessing the availability of open source pre-trained language models and leveraging established Natural Language Processing tasks. As a use case, we worked on the automatic generation of both textual and video product descriptions from the structured information about the product features. The pipeline involves all the required steps, providing the final trained and customized model. The obtained descriptions showed the capability of replicating the overall semantic, lexical and linguistic style of the corresponding human counterpart, despite being based on a cost-effective model.

Keywords

Natural Language Processing, Large Language Model, Text Generation, Video Generation

1. Introduction

As the number of applications for Artificial Intelligence (AI) has grown, several technologies have gained particular interest, especially with regard to automation techniques. This aspect increased the importance of Natural Language Processing (NLP) and Natural Language Generation (NLG), allowing computers to understand human language and respond accordingly, showing significant potential and a wide array of applications based on solid deep learning solutions [1, 2]. NLG tasks aim to produce natural language outputs that closely resemble human language, effectively conveying the intended meaning while embodying a distinct style and lexical structure.

Data-to-text architectures serve as the foundation for enhancing services by transforming structured data into useful information. They enable the generation of text to compensate for

SEBD 2024: 32nd Symposium on Advanced Database Systems, June 23-26, 2024, Villasimius, Sardinia, Italy
*Corresponding author.
†These authors contributed equally.
andrea.avignone@polito.it (A. Avignone); alessandro.fiori@polito.it (A. Fiori); silvia.chiusano@polito.it (S. Chiusano); giuseppe.rizzo@linksfoundation.com (G. Rizzo)
0000-0001-7003-129X (A. Avignone); 0000-0003-3819-2696 (A. Fiori); 0000-0002-5740-5004 (S. Chiusano); 0000-0003-0083-813X (G. Rizzo)

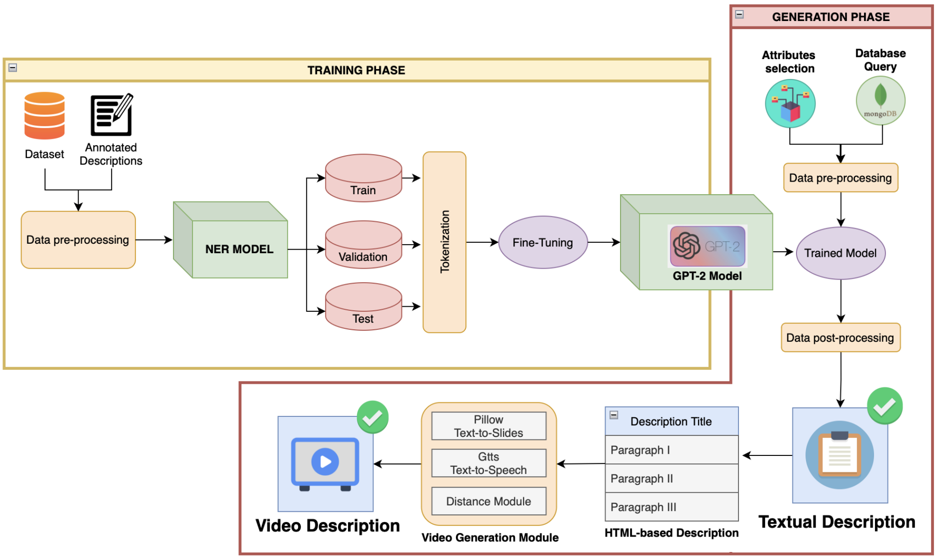


Figure 1: Proposed pipeline with the two main phases: a training phase and a generation phase.

the shortage of human resources. Over the years, different approaches have been proposed for different domains [3, 4, 5]. However, the success of the Transformer-based solutions [6] highlighted the effectiveness of transferring the general knowledge of more sophisticated models to specific sub-domains using fine-tuning [7]. In particular, popular models like BERT [8] and GPT [9] have provided opportunities for engaging implementations [10, 11].

This paper represent an extended abstract of a recent proposal [12]. Using structured data as input, our work exploits NLP techniques to design an integrated pipeline for generating complete textual and audio-video descriptions. As a real-world use case scenario, we designed a system for automatically generating technological product descriptions from the specifications of the product itself. The frequent launch of new devices makes it difficult to find detailed and complete descriptions for each product. While structured information is highly beneficial for comparisons, individuals typically favor a narrative approach to enhance comprehension, as noted by [13].

Our main contributions include: (1) the development and evaluation of a Named Entity Recognition (NER) model to detect product features and product title; (2) an integrated pipeline using the open GPT-2 model; (3) the generation of a complete textual product description and the corresponding video based on the list of product features.

2. Methods

Fig. 1 shows the general overview of the proposed pipeline for producing complete descriptions starting from the list of attributes. We worked on two main steps: (i) the training phase, for fine-tuning the pre-trained model (Section 2.1); and (ii) the generation phase, for creating a new textual description and the corresponding video according to the given input (Section 2.2).

2.1. Training phase

The core element of our pipeline is the usage of a Neural Network (NN)-based model. In the current instance of the proposed pipeline, we selected GPT-2 as reference solution for the text generation model. In fact, GPT-2 is open source, free of charge and gives access to the generated model. The overall process is designed to enhance the ability of the language model to produce coherent and relevant product descriptions. Therefore, it involves specific data pre-processing steps. In particular, we developed a customized Named Entity Recognition (NER) model to transform the given dataset and map the descriptions with the corresponding list of features.

2.1.1. Data pre-processing

We collected several product descriptions from different sources (see Section 3). Irrelevant features have been eliminated and the stored information has been reduced to *product title* and *product description* only. Since short descriptions (less than 100 characters) are unhelpful for training, they have been removed, whereas long descriptions (i.e. more than 750 words for GPT-2 small) have been split multiple times to comply with the length-related constraints of the model.

2.1.2. NER model

A primary objective was to automatically establish a link between the structured data and the corresponding textual description. For detecting product features and titles within the textual descriptions, we proposed a customised NER model. It was based on the introduction of two specific tags to denote the product title (*<prod>*) and each feature/attribute (*<attr>*). Then, the final dataset for training was built, obtaining the desired semi-structured data linking the list of features (extracted using NER) and the corresponding text (i.e., the original description).

2.1.3. Fine-tuning the LLM model

For each element of the training dataset, we merged the input data extracted using the NER model and the corresponding description into one single object to feed the GPT-2 trainer. As shown in Table 1, specific tokens were included to enrich the tokenizer of GPT-2: (1) *product name*, (2) *product features list*, (3) *product description*. The sequence tokens (*<OVERV_START>* and *<OVERV_END>*) replaced the original GTP-2 ones (i.e. GPT-2 *bos* and GPT-2 *eos*, respectively).

Table 1
Definition of the special tokens used for fine-tuning the GPT-2 model.

Role	Token
Indicate beginning and end of sequence	<OVERV_START><OVERV_END>
Delimit start and end of product title	<NAME_START><NAME_END>
Indicate beginning and end of features list	<FEAT_START><FEAT_END>
Separate one feature from another	<NEXT_FEAT>
Indicate beginning and end of the description	<DESCR_START><DESCR_END>

2.2. Generation phase

Once obtained the new trained model, it was possible to proceed with the automatic generation of the product description. However, post-processing was still required to improve the overall quality of the results. It returns a complete HTML-based textual description which is used to generate the video presentation accordingly.

2.2.1. Input manipulation

Considering our real-world use case, the proposed approach automatically retrieved the required information from a MongoDB instance, manipulating the collected data to be a suitable input for our model (i.e., including the introduced special tokens). Filtering is supported by the pipeline to define the desired list of features (and categories of features) employed for the description generation and/or a *block-list* to remove unessential information (e.g. packaging, suppliers). Each product feature is defined by the union of the corresponding key-value pair (or the key alone for boolean values) to better support GPT-2 in understanding the received input.

2.2.2. Text description

For the actual text generation, there are different decoding solutions [14]. *Top-k sampling* [15] is a simple schema that, once the parameter k is set, creates a list of the k most probable items to choose from. Since this method is strictly dependent on the choice of k, *Top-p* (or *nucleus sampling*) [16] tries to enhance this aspect by choosing a set of words (with flexible set size) and checking if the corresponding cumulative probability exceeds the probability p. We found effective to implement a combined solution based on both *Top-k* and *Top-p* sampling.

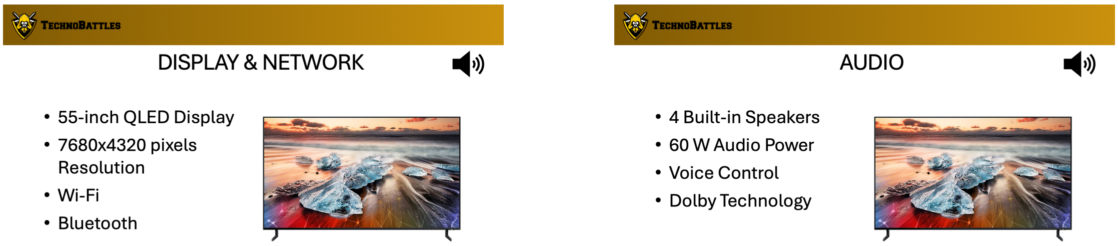
Finally, the *temperature* is a crucial parameter to be set carefully. In fact, it defines the creativity of the model in generating text, thus higher values may lead to incorrect, misleading or out of topic text.

Minor post-processing is then required to ensure better quality and usability. Based on the special tokens *<DESCR_START>* and *<DESCR_END>*, the description is automatically retrieved and processed. Common mistakes are removed, as well as redundant sentences according to a similarity score. Lastly, subtitles and paragraphs are highlighted to create the final comprehensive and organized textual description.

2.2.3. Video generation

The *Video Generation module* returns the video description given the obtained text. We divided this block in two main actors: a *distance module* and a *video and audio generation module*.

Distance module. The video is designed as a slideshow about the given product, with a text-to-speech model for reading the description. The synchronization of audio and video ensures that whenever a sentence featuring a particular feature is spoken, it promptly displays on the screen. It is based on sentence embedding, a technique mapping sentences to real-valued vectors, considering as input both the textual description and the list of features. As model, we used *all-MiniLM-L6-v2*. For improving the accuracy of the sentence embedding, we created simple sentences from the original features (e.g., "this product has a 4K resolution").



(a) Display and Network section.

(b) Audio section.

Figure 2: Example of the automatically generated slides describing a Smart TV.

Audio and video generation. Slides consider the title and the image of the product for providing suitable graphics. Each paragraph triggers the creation of a new slide, with the paragraph title and features from the sentences displayed in sync with the audio narration. The Google Text-to-Speech (gTTS) library orchestrates the synchronized audio generation. To elevate consumer engagement, supplementary elements like background music and visually appealing slide templates are introduced. Fig. 2 illustrates two sample slides for reference.

3. Results

In this section, we report the obtained results, both for the NER model and the generated text evaluation. For training and testing the pipeline, we merged different free datasets about technological products of popular e-commerce [17, 18, 19] and the data from icecat catalogue[1]. The final dataset was composed by 23,677 entries (90% for training, 7% for validation, and 3% for testing). We start reporting the NER experiments and then we describe the results for the generated descriptions using our pipeline. All the experiments have been conducted using the Kaggle platform with NVIDIA TESLA P100 GPUs resources. The most popular Python libraries have been used for NLP tasks (e.g., NLTK [20] and spaCy [21]).

3.1. Assessing the NER model

The *10-fold Cross Validation* method has been used over 58 product descriptions and considering two distinct tags (i.e., *ATTR* and *PROD* tags). For describing the overall performance of the NER model, we defined some specific metrics. The *Perfect match rate*, which is the ratio of exactly retrieved attributes and product names with respect to the whole set of detected elements; *Partial match rate*, indicating the detected entities with partial overlapping, thus not the complete original element; *False alarm*, counting the portion of incorrectly detected entities with no counterpart in the reference set (complementary of the *Total match rate*); *Miss rate*, providing the rate of undetected entities, neither partial; *Misclassification rate*, the ratio of both perfect and partial matches assigned to the wrong tag. The complete assessment is detailed in Table 2.

[1]https://icecat.com

The *Total match rate*, comprising *Perfect match* and *Partial match*, serves as a key indicator of system performance. Notably, it stands at 82.55% for attributes, while 96.90% for product names, indicating highly satisfactory result. Misinterpreting a feature as a product name occurs less frequently (6.69% only), whereas it is more prevalent for product attributes (14.68%). This discrepancy arises from the limited number of product names relative to the diverse array of attributes within textual descriptions.

3.2. Generated descriptions

Fig. 3 reports an example for a gaming keyboard (provided product features: *wired keyboard, illuminated keys, USB port, USB plug, aluminium, Windows key lock, anti-ghost 19 keys, 3 LED color light, Windows OS*), comparing our generated solution (Fig. 3a) with the reference description (Fig. 3b) . The generated text successfully organized the majority of the given features with a coherent structure, including subtitles, despite some minor inaccuracies (e.g. "Windows and Windows operating system"). The overall writing style aligns with the reference text, with our model attempting to inject creativity where possible. In comparison, the reference description relies on standard and redundant expressions towards the end.

3.2.1. Statistical assessment of the generated text

The statistical assessments was performed by using the common metrics BLEU-4 and GLEU (Google-BLEU) [22]. These metrics are usually in the range $0.35 - 0.65$ for consistent results.

A grid search has been performed to identify the model response to different configurations, according to the values of $temperature$ (i.e., 0.5, 0.8, 1.2), k (i.e., 35, 50) and p (i.e., 0.9, 0.95). By averaging over 3 runs, we obtained better results for $temperature = 1.2$, $k = 50$, $p = 0.9$. For our test dataset, we had 0.450 ± 0.007 for BLEU and 0.510 ± 0.005 for GLEU. This indicates that the generated text generally exhibits strong correlation with the reference text.

3.2.2. Human survey

We have launched an online survey to gather public opinions on the obtained product descriptions to have a human evaluation. In particular, we proposed two text comparisons including both the original and the generated description for a specific product. Users were asked to select the most fitting description, according to their individual perception.

	ATTR		PROD	
	Mean	**Std**	**Mean**	**Std**
Perfect match	56.35%	8.70%	68.66%	28.17%
Partial match	26.19%	4.85%	28.25%	27.86%
Total match	82.55%	6.15%	96.90 %	6.55%
False alarm	17.45%	6.15%	3.10%	6.55%
Miss	25.78%	8.04%	21.33%	17.94%
Misclassification	6.69%	4.07%	14.68%	13.34%

Table 2
NER evaluation for different metrics, $k = 10$.

Redgear Blaze LED Wired USB Gaming Keyboard (MT01, Black)

Get the powerful gaming experience with Redgear MT-01 Blaze 3 Color LED Wired Gaming Keyboard. The minimalist yet stylish black keyboard is crafted from the finest material to provide you with a comfortable typing experience. With its compact, ergonomic design, you can easily carry this gaming keyboard wherever you go. It uses the latest USB port technology, so you can connect your compatible mouse or pen drive to the USB plug and play..

Convenient Usage
Redgear MT-01 Blaze 3 Color LED Wired Gaming Keyboard is compatible with Windows and Windows operating systems. With the Redgear MT-01 Blaze 3 Color LED Wired Gaming Keyboard, you can easily toggle between the mouse and pen drive. It is also easy to set up and operate, as it comes with 3 LED Color Light. Equipped with aluminum housing, you can use this gaming keyboard as your gaming mouse, plug and play, or turn it on and off as per your needs. It is equipped with anti-ghost 19 keys.

(a) Generated description

Redgear Blaze LED Wired USB Gaming Keyboard (MT01, Black)

A wired keyboard means there is a wire connecting your keyboard to your computer. At the end of the wire is a USB plug that goes into a USB port on your computer. Presenting you the Redgear MT-01 Blaze 3 Color LED Wired Gaming Keyboard which provides with perfect gaming experience to the user.

Perfect Usage
The Redgear MT-01 Blaze 3 Color LED Wired Gaming Keyboard is Windows Operating System supported. It comes with Illuminated Keys. It is equipped with 3 LED Color Light. Comes with Floating Switch: 45g trigger pressure, anti-ghost 19 keys.

Awesome Built
The Redgear MT-01 Blaze 3 Color LED Wired Gaming Keyboard give you one of the strongest body ever for keyboard. It comes with Wired USB interface. It comes with Windows key lock, Aluminum structure. It has floating keycaps for mechanical feel.

(b) Reference description

Figure 3: Description comparison between (a) the generated text; (b) the reference text.

We collected feedback from 87 volunteers with diverse ages and backgrounds. The survey revealed that in 56.3% of cases, our generated text was rated as either better or equal to the original text. Therefore, the product descriptions produced by our system closely resemble the original ones, making it evident that the proposed model possesses the capability to replicate the linguistic elements of the training dataset, including text structure and paragraph delineation, relying on the provided list of product features as structured input only.

3.3. Real world use case

Finally, we used structured information collected from a MongoDB database of a real website[2] to generate product descriptions, assessing the effectiveness of our pipeline in a practical scenario. For online e-commerce and websites, the length of the description is a key aspect to provide a useful service: therefore, longer descriptions are required. However, the implemented version of GPT-2 (768 tokens) is a strong restriction. Then, our solution allows features selection by list of categories and the final text is obtained by iteration, organizing the description in the corresponding paragraphs. As a reference, Fig. 4 shows a complete example about a Smart TV: the text is consistent with common online product descriptions, describing all the different aspects of the product in each paragraph.

[2]https://technobattles.com

PHILIPS 65PUS7600 SMART TV

GENERAL OVERVIEW

Enjoy a high-quality viewing experience with the brilliant PHILIPS 65 Inch TV that allows you to enjoy watching content with an ease. The television comes loaded with various connectivity features to ensure you enjoy watching content according to your convenience at all times. Apart from also giving you amazing picture quality and a viewing experience, the television gives you a very ergonomic and comfortable design which facilitates long hours of usage at all times. The amazing performance coupled with a sleek design makes the television a perfect option for you.

Brilliant Visuals

Enjoy an immersive viewing experience as the PHILIPS 65 Inch TV comes with various connectivity features that ensure you enjoy watching content according to your convenience. The Wi-Fi connectivity allows you to connect your television to a wide range of Wi-Fi devices so that you can enjoy your favorite movies, shows and songs on a bigger screen.
The television comes with a 4K resolution that ensures you enjoy visuals that are up to 3840 x 2160 (4K) pixels.

KEY FEATURES & AUDIO

Enjoy the power of Dolby Vision on your television with the amazing PHILIPS 65PUS7600 TV which is sure to give you an amazing audio experience. The Dts studio sound Dolby Technology gives you cinematic sound like never before to give you distortion-free audio and the Dts audio recorder gives you a hands-free TV experience so that you can enjoy your television viewing experience anywhere and anytime. The amazing features coupled with a user-friendly design makes it the perfect option for you.

Amazing Audio Output

Experience enhanced audio quality with the PHILIPS 65PUS7600 TV which is sure to mesmerize one and all with its cutting-edge technology. The Dolby audio technology allows you to enjoy distortion-free audio so that you can enjoy any dialogue, every scene, from the highest settings to the darkest recesses of your mind. The Dts studio sound Dolby Technology gives you Dolby sound quality like never before and allows you to enjoy distortion-free audio so that you can enjoy your television viewing experience anywhere and anytime.

SPECIAL FEATURES, NETWORK & DISPLAY

The video player is responsible for providing the viewer with high resolution television content that is suitable for the viewing pleasure of the buyer. In order to fulfil the buyer's expectation of viewing this television you need to have access to a high quality video player which offers you with perfect usage. The PHILIPS 65PUS7600 TV offers you with perfect usage. It comes with Miracast technology. It has 1.400 Hz Motion Interpolation Frequency. It has Ethernet Port. It has Wi-Fi, RJ-45 Support. It supports Miracast (IPv4).

Perfect Usage

The PHILIPS 65PUS7600 TV offers with flat screen resolution. It comes with HDMI 2.1, 3 USB 2.0 Ports. It enhances the viewing experience of the buyer. It supports Pmr (Perfect motion Rate) 1400 hz Motion Interpolation Technology. It also comes with Ethernet Port.

Awesome Design

It comes with Flat screen technology. It comes with 3840 x 2160 (4K) pixels resolution.

Figure 4: Complete product description generated from the corresponding structured list of features.

4. Conclusions

The advent of significantly powerful NLG models has made possible the integration of systems capable of automating time-consuming tasks. However, the resources required for more advanced models are not negligible. In our work, we studied the feasibility and effectiveness of automatically transforming structured data into engaging textual and visual narratives, by leveraging on the capabilities of the open-source GPT-2 model. Our proposed pipeline deals with all the necessary steps to fully integrate an NLG solution in a real-world scenario, starting from data manipulation towards the generation of both text and video descriptions. Using structured data as entry point (i.e., list of product features) implies more constraints, as the model must remain consistent with the given input to avoid reporting incorrect information. By comparing the generated text with the reference one, our analysis showed that the system was actually able to generate a coherent text and replicate the linguistic style of the given context, transforming raw data into more compelling and narrative information.

Currently, our solution is confined to the small version of GPT-2 which provided interesting results and human-like structures and contents. However, moving to the medium or large versions would be optimal for fully harnessing the model's potential, offering better performance. Our future work will focus on analyzing different models and configurations, supported by the acquisition of a more suitable training dataset, both in terms of quantity and linguistic quality.

References

[1] P. Johri, S. K. Khatri, A. T. Al-Taani, M. Sabharwal, S. Suvanov, A. Kumar, Natural language processing: History, evolution, application, and future work, in: A. Abraham, O. Castillo, D. Virmani (Eds.), Proceedings of 3rd International Conference on Computing Informatics and Networks, Springer Singapore, Singapore, 2021, pp. 365–375.

[2] I. Lauriola, A. Lavelli, F. Aiolli, An introduction to deep learning in natural language processing: Models, techniques, and tools, Neurocomputing 470 (2022) 443–456. doi:https://doi.org/10.1016/j.neucom.2021.05.103.

[3] R. Lebret, D. Grangier, M. Auli, Neural text generation from structured data with application to the biography domain, in: J. Su, K. Duh, X. Carreras (Eds.), Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics, Austin, Texas, 2016, pp. 1203–1213. doi:10.18653/v1/D16-1128.

[4] A. Chisholm, W. Radford, B. Hachey, Learning to generate one-sentence biographies from Wikidata, in: Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 1, Long Papers, Association for Computational Linguistics, Valencia, Spain, 2017, pp. 633–642.

[5] J. Wang, Y. Hou, J. Liu, Y. Cao, C.-Y. Lin, A statistical framework for product description generation, in: Proceedings of the Eighth International Joint Conference on Natural Language Processing (Volume 2: Short Papers), Asian Federation of Natural Language Processing, Taipei, Taiwan, 2017, pp. 187–192.

[6] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. u. Kaiser, I. Polosukhin, Attention is all you need, in: I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, R. Garnett (Eds.), Advances in Neural Information Processing Systems, volume 30, Curran Associates, Inc., 2017. URL: https://proceedings.neurips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf.

[7] X. Han, Z. Zhang, N. Ding, Y. Gu, X. Liu, Y. Huo, J. Qiu, Y. Yao, A. Zhang, L. Zhang, W. Han, M. Huang, Q. Jin, Y. Lan, Y. Liu, Z. Liu, Z. Lu, X. Qiu, R. Song, J. Tang, J.-R. Wen, J. Yuan, W. X. Zhao, J. Zhu, Pre-trained models: Past, present and future, AI Open 2 (2021) 225–250. doi:https://doi.org/10.1016/j.aiopen.2021.08.002.

[8] J. Devlin, M.-W. Chang, K. Lee, K. Toutanova, BERT: Pre-training of deep bidirectional transformers for language understanding, in: J. Burstein, C. Doran, T. Solorio (Eds.), Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), Association for Computational Linguistics, Minneapolis, Minnesota, 2019, pp. 4171–4186. URL: https://aclanthology.org/N19-1423. doi:10.18653/v1/N19-1423.

[9] A. Radford, J. Wu, R. Child, D. Luan, D. Amodei, I. Sutskever, Language models are unsupervised multitask learners (2019).

[10] J.-S. Lee, J. Hsiang, Patent claim generation by fine-tuning openai gpt-2, World Patent Information 62 (2020) 101983. URL: https://www.sciencedirect.com/science/article/pii/S0172219019300766. doi:https://doi.org/10.1016/j.wpi.2020.101983.

[11] M. Bień, M. Gilski, M. Maciejewska, W. Taisner, D. Wisniewski, A. Lawrynowicz, RecipeNLG: A cooking recipes dataset for semi-structured text generation, in: Proceedings of the 13th International Conference on Natural Language Generation, Association for

Computational Linguistics, Dublin, Ireland, 2020, pp. 22–28.
[12] A. Avignone, A. Fiori, S. Chiusano, G. Rizzo, Generation of textual/video descriptions for technological products based on structured data, in: 2023 IEEE 17th International Conference on Application of Information and Communication Technologies (AICT), 2023, pp. 1–7. doi:10.1109/AICT59525.2023.10313177.
[13] N. D. Gershon, W. Page, What storytelling can do for information visualization, Communications of the ACM 44 (2001) 31 – 37.
[14] P. von Platen, How to generate text: using different decoding methods for language generation with transformers, 2020. https://huggingface.co/blog/how-to-generate, Last access on 2 November 2022.
[15] A. Fan, M. Lewis, Y. Dauphin, Hierarchical neural story generation, in: I. Gurevych, Y. Miyao (Eds.), Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Association for Computational Linguistics, Melbourne, Australia, 2018, pp. 889–898. URL: https://aclanthology.org/P18-1082. doi:10.18653/v1/P18-1082.
[16] A. Holtzman, J. Buys, L. Du, M. Forbes, Y. Choi, The curious case of neural text degeneration, arXiv preprint arXiv:1904.09751 (2019).
[17] H. Mehta, Croma electronic products dataset, 2021. https://www.kaggle.com/hetulmehta/croma-electronic-products-dataset.
[18] AK47, Amazon-phone-dataset, 2019. https://www.kaggle.com/ak47bluestack/amazonphonedataset.
[19] S. Tripath, Amazon mobile dataset, 2021. https://www.kaggle.com/daishinkan002/amazon-mobile-dataset.
[20] S. Bird, E. Klein, E. Loper, Natural language processing with Python: analyzing text with the natural language toolkit, " O'Reilly Media, Inc.", 2009.
[21] M. Honnibal, I. Montani, spaCy 2: Natural language understanding with Bloom embeddings, convolutional neural networks and incremental parsing, 2017. To appear.
[22] Y. Wu, M. Schuster, Z. Chen, Q. V. Le, M. Norouzi, W. Macherey, M. Krikun, Y. Cao, Q. Gao, K. Macherey, J. Klingner, A. Shah, M. Johnson, X. Liu, Łukasz Kaiser, S. Gouws, Y. Kato, T. Kudo, H. Kazawa, K. Stevens, G. Kurian, N. Patil, W. Wang, C. Young, J. Smith, J. Riesa, A. Rudnick, O. Vinyals, G. Corrado, M. Hughes, J. Dean, Google's neural machine translation system: Bridging the gap between human and machine translation, 2016. arXiv:1609.08144.