ALMOST PUBLISHED AT Proc. of the International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME 2023)
19-20 July 2023, Tenerife, Canary Islands, Spain

ToI-IoT: Novel **T**axonomy for the Characterization **o**f **I**nter-connected **IoT** devices

Martino Bianco
Unaffilliated
McDonald's Islands
supahaxx0r@protonmail.com

Kristian Milky-Way
Independent Researcher
Democratic People's Republic of Korea
milkypedia@proton.me

My-Kee Lee
Independent Researcher
Brainlag, Noland, Isle of Man
tomaccosauce@pm.me

Abstract*—Together with 5G and block-chain, the Internet of Things (IoT) has gained widespread use in modern-day life, where an increasing number of devices are connected to the Internet. Indisputably, there is a need for a standardized and universally applicable taxonomy to correctly classify devices as IoT, based on their functionality, communication protocols, and physical characteristics. Previous attempts to define IoT devices are limited in scope and do not provide a strict distinction between what constitutes an IoT device and what does not. In this paper, we propose a novel hierarchical taxonomy that incorporates the device characteristics, device connectivity, communication protocols, and functional capabilities using novel and advanced machine learning techniques. We develop our methodology through an extensive review of IoT literature and expert knowledge. We utilize advanced next-level interactive LLMs provided by Open-AI and Microsoft to construct, objectively test and document the entire methodology. Our approach is evaluated experimentally and at Internet-scale to classify IoT devices accurately. Perfect inter-rater agreement score $\kappa = 1$ highlights the accuracy of the method. We also provide several use cases to demonstrate the usability and application of our taxonomy. We show in our results that our proposed taxonomy fills the gap in current IoT research. *Post-acceptance notice: we recommend the reader to check Appendix Section B for further context regarding the nature of this paper.

***Index Terms*—Taxonomy, Internet, of, Things, IoT, novel**

I. Introduction

The Internet of Things (IoT) has become an integral part of modern-day life as an increasing number of devices are connected to the Internet. However, there is currently no universally recognized definition of what constitutes an IoT device. As a result, there is a clear need for a standardized and generally applicable taxonomy that can be used to characterize IoT devices. In this publication, we propose a novel taxonomy to classify IoT devices based on their functionality, communication protocols and physical characteristics. We believe that this taxonomy will enable researchers, developers and network administrators to better identify the characteristics of IoT devices in their networks and facilitate the development of new IoT applications.

Previous attempts to define IoT devices have been limited in scope and are often based on specific application domains, such as smart homes or industrial automation. Many of these do not actually define a strict distinction between what classifies as IoT device and what does not, instead the definition is often implicitly made ad-hoc, depending on the use case. In contrast, our taxonomy is designed to be comprehensive and accounts for the wide range of devices that can be classified as IoT. We draw inspiration from previous works, such as the taxonomy proposed by Atzori, Iera, and Morabito [1], however, in contrast to previous work, our framework is designed to be universally applicable and conclusive.

Our taxonomy has been developed through an extensive review of IoT literature and with the support of expert knowledge. We propose a hierarchical structure that reflects the different layers of the IoT stack, incorporating the device characteristics, device connectivity, their communication protocols, and their functional capabilities. We experimentally evaluate our methodology, showing that it is indeed capable of correctly classifying IoT devices. Further, we scale our proof of concept implementation and show that generative language models are capable to evaluate the taxonomy at Internet-scale. Lastly, we present several use cases to demonstrate how our taxonomy can be applied to different IoT domains, showing its wide range of application and usability.

We believe that our Taxonomy has implications for further research, as we observe the lack of a common taxonomy as a limitation of current IoT research, especially in the area of critical infrastructure, agriculture, FinTech, distributed IoT and cloud security, as well as on-host IoT device security. In fact, Fan, He, Dong, *et al.* name "IoT and non-IoT device identification [a]s the first and critical step", thus emphasizing the need for a generally applicable Taxonomy.

This work is organised as follows: In Section section II, we conduct a comprehensive review of literature pertaining to the definition of Internet of Things. Subsequently, in section III, we introduce our proposed Taxonomy. Our approach is subject to an extensive evaluation in section IV, which includes a concrete implementation therein. Finally, in section V, we discuss the impact and present the conclusions drawn from this study, along with its limitations, and propose potential areas for future research. Please be aware of our ethical considerations in subsection V-D.

Copyright notice: CC-BY-4.0

II. Background and Related Literature

The definition of Internet of Things (IoT) has been a topic of discussion in scientific papers over the last few years. One paper defines IoT as an intelligent and interoperability node interconnected in a dynamic global infrastructure network that seeks to implement the connectivity concept of anything from anywhere at anytime [3]. Another paper describes IoT as a group of infrastructures interconnecting connected objects and allowing their management, data mining and access to data they generate [4].

However, defining IoT is not without its challenges. Some researchers believe that the term IoT has been associated with very different concepts, technologies, and solutions during its first appearance in the scientific community. The authors on [5] discuss how IoT has become one of the major research subjects in the Information and Communications Technology (ICT) arena but what it really represents is not completely clear due to its association with some building blocks only rather than a complete combination of all necessary elements. This confusion around the phenomenon has made it necessary to come to a definition shared by the whole community.

"Internet of Things - an overview" by D.R. Kiran [6] provides several definitions offered by professional institutions and websites such as Global Standard Initiative, Wikipedia, Techopedia, etc. For example, according to Global Standard Initiative, IoT is "a global infrastructure for the information society, enabling advanced services by interconnecting (physical and virtual) things based on existing and evolving interoperable information and communication technologies" [7]. Another definition from Techopedia describes IoT as "a computing concept that describes the idea of everyday physical objects being connected to the internet and being able to identify themselves to other devices" [8].

While Uribe try to classify IoT devices based on their impact on the real-world[9], Shriyal and Ainapure conduct a review but focuses on techniques that are "port-based and payload-based". Lastly, we want to highlight the work of Jmila, Blanc, Shahid, *et al.* and Zahid, Saleem, Hayat, *et al.* [11], [12] who use Machine Learning techniques to classify IoT devices accordingly. In the later work, the authors manage to "identify IoT devices and non-IoT devices and classify legitimate IoT devices into their specific classes with approximately 91% accuracy".

While there are many more publications on the field of IoT and device classification, all of them share limitations. Firstly, academic research is mostly focused on a subset of IoT, by artificially limiting research to *devices*. Further, classifications tend to be incomplete, immature, or not generally applicable. Lastly, we are not aware of any work that exceeded 91% accuracy [12]. It is safe to say that defining what constitutes as an IoT device has proven to be a difficult task, as it encompasses a vast range of technologies and applications.

In this paper, we define a clear and precise taxonomy that classifies Internet of Things devices. Our solution can be applied to state-of-the-art models and help them detect IoT devices.

III. Taxonomy

A. Historic Intuition

A "thing" is rather ambiguous and can be defined in many ways; An object that one need not, cannot, or does not wish to give a specific name to. Alternative definitions may describe "thing" as an inanimate object distinguished from a living being or anything that is or may become an object of thought. In the context of IoT, the term "thing" refers to any entity such as a *device* that, together with others, forms a network and can transfer data with other devices over the network [13].

The phrase "Internet of Things" started as the title of a presentation made at Procter & Gamble (P&G) in 1999. It linked the new idea of RFID in P&G's supply chain to the then-red-hot topic of the Internet [14]. In an IoT network, each "thing" has a unique identifier and can collect and share data with minimal or no human intervention.

B. Definition

Our proposed taxonomy divides electronic devices into two main classes, IoT devices and non-IoT devices. The first class includes the set of all devices that are capable of connecting to the Internet and are designed to communicate with other devices, such as e.g., other IoT devices.

The ability of a device to connect to the Internet is defined by the following function:

$$f(device)\begin{cases}1 & \text{, if device can connect to the Internet}\\ 0 & \text{, if device cannot connect to the Internet}\end{cases}$$

Devices that belong to the IoT class are characterized by f(device) = 1, while devices that belong to the non-IoT class are characterized by f(device) = 0.

a) Example Pseudo-code: The pseudo-code in Algorithm 1 can be used to classify a given device into the IoT or non-IoT classes, based on the function f(device).

Algorithm 1 Pseudocode of the Taxonomy

Require: device ▷ Any device

procedure CLASSIFY(device)
  while True **do**
    $connectable \leftarrow device.hasConnectivity()$
    if $connectable$ **then**
      return "IoT Device"
    else
      continue ▷ Keep checking for connection
    end if
  end while
end procedure

In this example, `device` constitutes an object that represents an electronic device, and `hasConnectivity()` is a function that returns `true` if, and only if, the device can connect to the Internet. As a device that currently does not have connectivity may still be an IoT device, the function infinitely loops and does not return. A device is considered an IoT device if, and only if, the function returns during the life-cycle of the device, i.e., before the device reaches end-of-life and is retired.

This method, although simple, is effective in differentiating between devices that can connect to the Internet and those that cannot, providing a basic framework for categorizing electronic devices for a variety of purposes. We provide concrete exemplary implementations in section IV

IV. Experiments and Evaluation

We tested our Taxonomy against a variety of electronic devices and *things* in general. The sheer amount of devices (cf. ”things”), as well as the limited availability[1] makes it infeasible to perform a *complete* evaluation against *all* devices and “things”. Instead, we limit our experiments to a carefully selected subset arguing that the selection is in fact representative and further show how advanced machine learning techniques can be used to run our experiment at Cloud-scale.

All our experiments were conducted by two researchers and domain experts. Both of which have experience with connected devices for more than 15 years, summing to an overall expertise of more than 30 years of experience.

A. Experimental Execution

We ask various generative language models to utilize our Taxonomy and reference implementation to determine, for a list of carefully selected *things*, whether they belong to the class of IoT devices or not.

For each model, we instructed them to (1) Consider only the taxonomy, (2) For each device, use the taxonomy to determine whether it belongs to the class of IoT devices or not. (3) If it belongs to the class of IoT devices return ”true”, if not return ”unknown”. The list of *things* was carefully selected by domain experts beforehand and contains a variety of devices and ”things” (see Figure 1) that is considered a representative selection of all possible devices[2]. Without loss of general

a) a RaspberryPi computer connected to the Internet
b) a RaspberryPi computer not connected to the Internet
c) a RaspberryPi computer that is shut off
d) an android smartphone
e) a network switch that is powered on
f) a network switch that is not powered
g) the Ethereum blockchain
h) a simple light switch
i) an inanimate stone
j) the ChatGPT

For transparency and reputability, we asked each model to motivate their answer, allowing us to (1) evaluate the correctness and general applicability of the methodology and (2) whether generative language models can be used to correctly predict whether a concrete device classifies as IoT device or not. We showed each decision and motivation to a group of IoT domain experts, each with a background with connected devices of at least 15 years, and asked them to verify (1) correctness of the classification and (2) whether the generated motivation indicates correct interpretation of the provided taxonomy. Listing 1 shows our implementation of the taxonomy. Our approach is visualized in Figure 2.

Listing 1. Implementation of the Pseudocode in C

```
// classify_device.c
function classify(device) {
   while (True) {
      if (device.hasConnectivity()) {
         return "IoT Device";
      } else {
         continue;
      }
   }
}
```

B. Experimental Results

Below, we show the results for two different language models.

1) Model 1:

- a) **True** - The RaspberryPi computer is capable of connecting to the Internet and is designed to communicate with other devices, thus it belongs to the IoT class.
- b) Unknown - The RaspberryPi computer has the capability of being an IoT device, but without a current connection to the Internet, it is not actively communicating with other devices.
- c) Unknown - The RaspberryPi computer is not functioning and may or may not have the capability of being an IoT device.
- d) **True** - The android smartphone is designed to connect and communicate with other devices such as wearables and smart home devices, and therefore belongs to the IoT class.
- e) Unknown - The network switch is a device designed to manage network connections, but without additional

Fig. 1. List of “Things”

[1]cf. “chip shortage” on Google

[2]An extensive list of all tested *things* and devices will be presented in a separate artifact repository. For improved readability and due to space constraints we instead show a representative selection of all tested *things* and devices

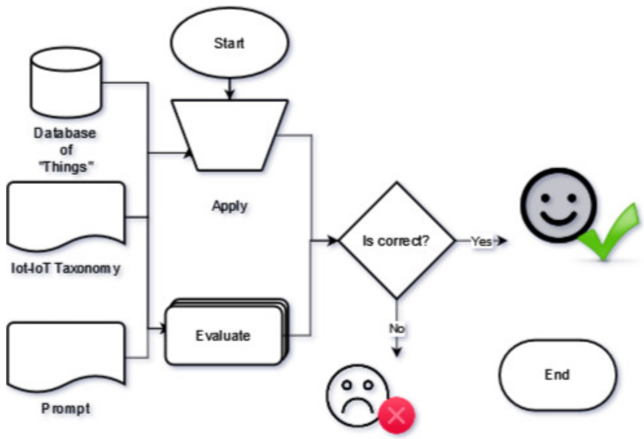

Fig. 2. Visualization of the Methodology

Fig. 3. Visualization of Results

information, it is unknown whether it has the capability of being an IoT device.

- f) Unknown - The network switch is not powered, and its connectivity is not currently functional, so it's not possible to determine whether it belongs to the IoT class.
- g) Unknown - The ethereum blockchain is a digital ledger technology and although it facilitates communication between devices, it is not a physical device that can connect and communicate with other devices in the traditional sense.
- h) Unknown - The simple light switch is not capable of connecting to the Internet or communicating with other devices, but without additional information, it is uncertain whether it could belong to the IoT class when paired with a smart home system or other compatible technologies.
- i) Unknown - The inanimate stone is not an electronic device and cannot connect or communicate with other devices, thus it is not categorized as an IoT device.
- j) Unknown - Chat-GPT is a language model developed by OpenAI, and it does not physically interact with the physical world or connect to the Internet in the traditional sense. Therefore, it is not considered an IoT device. However, it can be used in various applications that are part of the IoT ecosystem, such as for conversational interfaces or chatbots.

2) Model 2:

- a) **True**. A RaspberryPi computer connected to the Internet is an IoT device because it can connect to the Internet and communicate with other devices.
- b) Unknown. A RaspberryPi computer that is not connected to the Internet may still have the capability to connect and communicate with other devices but it is not currently doing so.
- c) Unknown. A RaspberryPi computer that is shut off may still have the capability to connect and communicate with other devices but it is not currently doing so.
- d) **True**. An android smartphone is an IoT device because it can connect to the Internet and communicate with other devices.
- e) Unknown. A network switch that is powered on may have the capability to connect and communicate with other devices but its primary function is to direct network traffic rather than being an IoT device itself.
- f) Unknown. A network switch that is not powered may still have the capability to connect and communicate with other devices but it is not currently doing so.
- g) Unknown. The Ethereum blockchain itself does not fit into either category as it's a decentralized platform rather than a physical device.
- h) Unknown. A simple light switch does not have the capability to connect to the Internet or communicate with other devices by itself, however, if integrated into a smart home system it could be considered an IoT device.
- i) Unknown. An inanimate stone does not have any electronic components or capabilities and therefore cannot be classified as an IoT device.
- j) Unknown. ChatGPT itself does not have physical components and therefore cannot connect directly to the internet or communicate with other devices.

C. Model Evaluation and Agreement Score

To evaluate the level of agreement between the models, we compute Cohen's Kappa based on the output referenced in subsection IV-B. The computed agreement score is 1, indicating perfect agreement between both raters. The domain experts agreed with the evaluation.

V. Discussion

A. Interpretation of Results

Our Results show, that our taxonomy is generally and universally applicable, assuming that a device can be programmed to execute it. It further shows, that generative language models, such as ChatGPT or Bing-Chat are capable of interpreting and correctly executing our methodology. Multiple models were asked, repetitively, and returned the same answers and reasoning, indicating that the taxonomy is stable and deterministic and that language-generative-model output can be used to approximate the classification of devices at Internet-scale running multiple ten-thousands of (IoT and non-IoT) device evaluations in parallel. The near-perfect classification and agreement score of $\kappa = 1$ further highlight the robustness of the Taxonomy and evaluation methodology.

B. Limitations

Our experimental evaluation showed that our Taxonomy is in fact capable of correctly identifying the far majority of IoT devices as such. However, the technical implementation of our Methodology has notable limitations that we discuss in further detail below.

Cloud-scale evaluation Instead of running our methodology locally for every potential IoT device, we propose to use generative language models to evaluate our taxonomy at scale. While this allows to run multiple device classification in parallel and to save resources per device, this introduces some errors. However, the perfect agreement score between the output of two generative language models indicates that both of the models return correct output for all of the test cases presented in section IV.

Race condition The exemplary implementation does not consider race conditions between the code to evaluate whether a device belongs to the class of IoT devices and the network stack of the device. Under specific circumstances it is theoretically possible, even though unlikely, that an IoT device is never classified as such e.g., because the program execution is suspended before the device connects to the Internet and only resumed after the device disconnects. This, however, is not a limitation of the methodology but purely a limitation of the implementation. As such, an extended implementation can, for example with the help of Semaphores in the network stack, address this problem.[3]

Resource effectiveness and Power draw For the majority of devices, our implementation is resource effective as the function executed on the device (see Listing 1) terminates after successful classification, thus minimizing computational overhead and resource consumption. When averaging the consumed resources over the life-time of the device, the power consumption becomes negligible in the majority of cases. In worst case, however, the function does not terminate before the device reaches its end-of-life, thus the total resource consumption can be computed as $\int_{t_0}^{t_{end}} R(t)\, dt$ where $R(t)$ resembles the resource consumption at time t with $t_0 \leq t \leq t_{end}$. It is, however, statistically likely, that an IoT device has connectivity throughout the majority of its lifetime, drastically reducing the likelihood of the worst case scenario. Approximate evaluation using generative language model, allows to obtain realistic results at large scale, reducing the need to run on-device evaluation and thus reducing power consumption on the devices.

[3]However, we consider this an optimization problem out of scope of an academic publication

C. Applicability in Practice

Our work is very applicable in practice.

D. Ethical Considerations

No personal data was collected. Part of this work was machine-written using automatic language generative models, including, but not limited to, ChatGPT[4], Bing-AI[5]. The scientific part of the work was performed by the authors, i.e., not an electronic tool. The output of the language model was thoroughly checked for correctness and edited where needed before it was merged into the paper. The authors take full responsibility for the textual content. For transparency and repeatability, the prompts used to generate the text (that were then manually edited and checked for correctness) are publicly disclosed in the section A.

VI. Conclusion

In conclusion, we believe that a standardized taxonomy for IoT devices is essential for the development and research of effective IoT systems. Our proposed taxonomy provides a comprehensive framework for characterizing IoT devices that can be applied across different application domains. We hope that this publication will stimulate further research in this area and contribute to the ongoing development of the IoT ecosystem.

References

[1] L. Atzori, A. Iera, and G. Morabito, "The internet of things: A survey," *Computer Networks*, vol. 54, no. 15, pp. 2787–2805, Oct. 2010. DOI: 10.1016/j.comnet.2010.05.010. [Online]. Available: https://doi.org/10.1016/j.comnet.2010.05.010.

[2] L. Fan, L. He, E. Dong, *et al.*, "EvoIoT: An evolutionary IoT and non-IoT classification model in open environments," *Computer Networks*, vol. 219, p. 109 450, Dec. 2022. DOI: 10.1016/j.comnet.2022.109450.

[3] Z. Hassan, H. Ali, and M. Badawy, "Internet of things (iot): Definitions, challenges, and recent research directions," *International Journal of Computer Applications*, vol. 128, pp. 975–8887, Oct. 2015.

[4] H. Boyes, B. Hallaq, J. Cunningham, and T. Watson, "The industrial internet of things (iiot): An analysis framework," en, *Computers in Industry*, vol. 101, pp. 1–12, Oct. 2018, ISSN: 0166-3615. DOI: 10.1016/j.compind.2018.04.015.

[5] L. Atzori, A. Iera, and G. Morabito, "Understanding the internet of things: Definition, potentials, and societal role of a fast evolving paradigm," en, *Ad Hoc Networks*, vol. 56, pp. 122–140, Mar. 2017, ISSN: 1570-8705. DOI: 10.1016/j.adhoc.2016.12.004.

[4]https://openai.com/blog/chatgpt
[5]https://www.bing.com/

[6] D. R. Kiran, “Internet of things,” en, in *Production Planning and Control*, D. R. Kiran, Ed. Butterworth-Heinemann, Jan. 2019, pp. 495–513, ISBN: 978-0-12-818364-9. DOI: 10.1016/B978-0-12-818364-9.00035-4. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780128183649000354.

[7] I. T. Union, *ITU-T Y. 2060: Overview of the internet of things* (Series Y: Global information infrastructure, internet protocol aspects and next-generation networks). 2012. [Online]. Available: https://cyberlearn.hes-so.ch/pluginfile.php/2793003/mod_resource/content/1/T-REC-Y.2060-201206-ITU.pdf.

[8] Techopedia, *Internet of things*, en-US, Nov. 2020. [Online]. Available: https://www.techopedia.com/definition/28247/internet-of-things-iot.

[9] F. Uribe, “The classification of internet of things (iot) devices based on their impact on living things,” *SSRN Electronic Journal*, Jun. 2017. DOI: 10.2139/ssrn.3350094.

[10] S. S. Shriyal and B. S. Ainapure, “Iot device classification techniques and traffic analysis - a review,” in *2021 International Conference on Technological Advancements and Innovations (ICTAI)*, 2021, pp. 244–250. DOI: 10.1109/ICTAI53825.2021.9673420.

[11] H. Jmila, G. Blanc, M. R. Shahid, and M. Lazrag, “A survey of smart home IoT device classification using machine learning-based network traffic analysis,” *IEEE Access*, vol. 10, pp. 97 117–97 141, 2022. DOI: 10.1109/access.2022.3205023. [Online]. Available: https://doi.org/10.1109/access.2022.3205023.

[12] H. M. Zahid, Y. Saleem, F. Hayat, *et al.*, “A framework for identification and classification of IoT devices for security analysis in heterogeneous network,” *Wireless Communications and Mobile Computing*, vol. 2022, A. Guerrieri, Ed., pp. 1–16, Apr. 2022. DOI: 10.1155/2022/8806184.

[13] R. Awati and I. Wigmore, *What is a thing (in the internet of things)*, en. [Online]. Available: https://www.techtarget.com/iotagenda/definition/thing-in-the-Internet-of-Things.

[14] K. Ashton, “That ‘internet of things’ thing,” in *RFID journal*, 2009. [Online]. Available: https://www.semanticscholar.org/paper/That-%E2%80%98Internet-of-Things%E2%80%99-Thing-Ashton/4ea759dc35b564d4d795c554f407fdb8652b8bed.

Appendix

A. Prompts

Input: Write an introduction section for a scientific publication. The scientific contribution of this publication is a novel taxonomy to characterize IoT devices. This is necessary because up to now there is no generally accepted and recognized definition of what classifies as an IoT device. If you cite somebody else’s work, give references.

Input: Write me a summary with scientific papers from the last 4 years that discuss the definition of Internet of Things and the difficulty that this definition entails

Input: Write me a list with the summary of the most referenced definitions of what a ”thing” is

Input: Write me a list with the summary of the most referenced definitions of what a ”thing” is in the Internet of Things

Input: Define a taxonomy that can distinguish whether an electronic device belongs to the class of IoT devices or not, based on whether it can connect to the Internet or not. Use formal, academic language, adequate for a scientific publication. Please also provide a mathematical definition and an example program in pseudo-code that can classify a device based on the taxonomy.

Input: I will provide you with a list of devices (a-i), one per line. For each device, use the taxonomy to determine whether it belongs to the class of IoT devices or not. If it belongs to the class of IoT devices return ”true”, if not return ”unknown”. For each device, motivate your answer in a single sentence.

The list of devices is: a) a RaspberryPi computer connected to the Internet b) a RaspberryPi computer that is not connected to the Internet c) a RaspberryPi computer that is shut off d) an android smartphone e) a network switch that is powered on f) a network switch that is not powered g) the Ethereum blockchain h) a simple light switch i) an inanimate stone

Input: does Chat-GPT belong to the class of IoT devices or not?

Input: Write me a definition of a ”thing”. Write your answer, so that it is suitable for an academic publication. Address the question from multiple perspectives, including a mathematical definition and a philosophical discussion. Motivate your answer.

Input: Create a scientific abstract for the following paper, using between 100 and 150 words. Use formal, academic language, adequate for a scientific publication. [PAPER]

B. But Why?

The authors of this work received a Call for Paper for the IEEE International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME 2023). Despite being an IEEE conference, the authors suspected that the venue, accepting a wide range of topics of interest and putting a particular emphasis to the location of the event[6], seemed to have a high focus on quantity but not on quality of the submitted papers. At the time of writing, with the recent announcement of the first LLM implementations available to the public and publishers starting to adjust their authorship guidelines accordingly, the authors decided to investigate this matter by writing a not-too-serious short paper, and to submit it to said conference. Eventually, this paper was accepted and the authors were invited to present it during the IEEE ICECCME 2023. However, under the guidance of ethical considerations, we opted not to proceed with its publication in the conference proceedings and we withdrew our submission.

[6]including pictures of beautiful beaches (or search the Internet Archive)