SCOPE
DATABASE
INDEXED

OPEN ACCESS

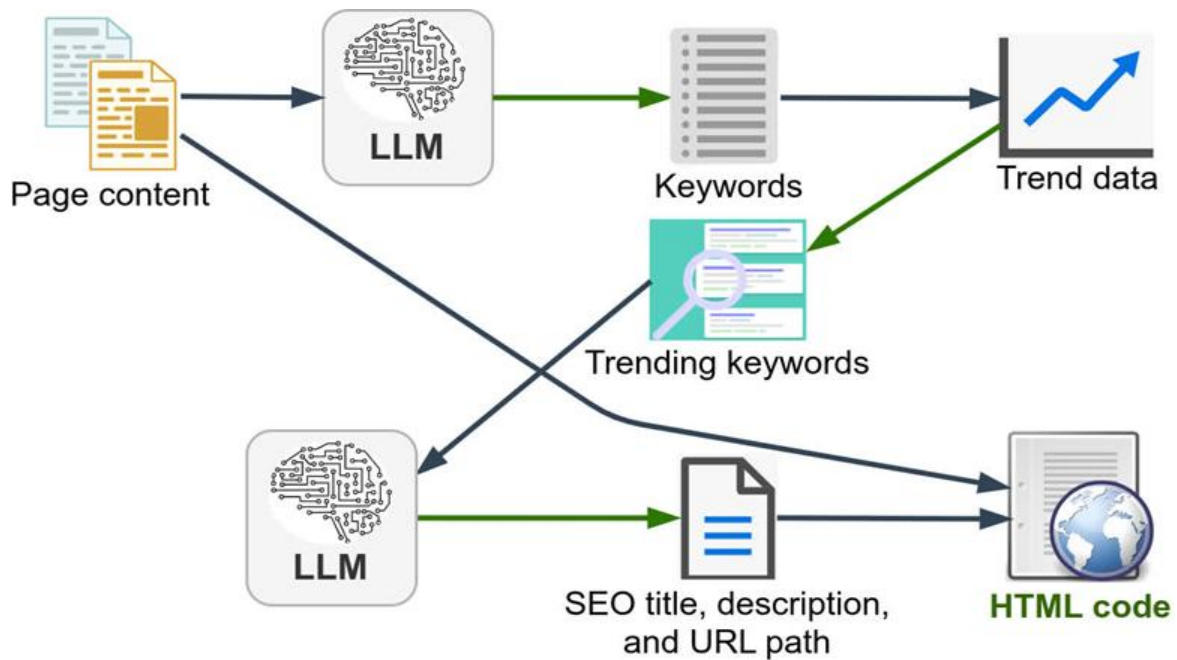# AUTOTRENDYKEYWORDS: REAL-TIME AI-DRIVEN TREND-BASED SEO USING LLMS

Praneeth Vadlapati

University of Arizona, USA.
ORCID: 0009-0006-2592-2564

ABSTRACT

Considering the evolution of Search Engine Optimization (SEO) methods, it is crucial to generate and update relevant key-words continuously based on trends in a dynamic method to maintain the relevance and visibility of a web page, which en-hances user engagement. To automate the process of continuous SEO, this paper proposes a method called Auto-TrendyKeywords to leverage Large Language Models (LLMs) to automate the constant generation of SEO keywords and the selection of trending keywords based on changing data. Unlike the manual selection method, the new system automates the selection using real-time trend data. This unsupervised approach ensures up-to-date relevance of the page and can ensure satisfactory traffic for pages in unpopular languages when the web is dominated by content in popular languages. The meth-od streamlines the keyword update process for digital content creators and marketers by eliminating the need for manual intervention, thereby constantly ensuring visibility across search results. The system has wide-ranging societal benefits, such as facilitating access to desired information and promoting diverse languages. The system successfully generated keywords using an LLM, selected the most trending keywords, and integrated the keywords and the description into HTML code. The system generated short URL paths that contain the trending keywords.

1. Introduction

Search Engine Optimization (SEO) is important to ensure that online content reaches the intended audience who expect to find the desired content for a web page [1]. Keywords are essential in SEO to connect user queries to relevant content across the web, improving content discovery [1]. Numerous keywords can be derived from a web page.

1.1. *Limitations of current fixed-keyword approaches*

Traditional SEO practices involve the manual selection of keywords and embedding in the content at the time of writing page content [1]. Content creators select the keywords before content creation, commonly based on topics of the content, trends, and historical data. The popularity and relevance of each query shift over time based on user interests [2], often resulting in the keywords becoming outdated. As a result, maintaining the visibility and relevance of a web page remains a challenge without continuous manual adjustments.

1.2. Disadvantages of manual-only approaches

Considering a high number of pages to maintain, reliance on manual SEO practices introduces challenges related to scalability, timeliness, cost, and effectiveness. The practices require dedicated SEO teams [3]. Manual processes can introduce challenges for smaller businesses and non-profit organizations that lack the resources to obtain and maintain dedicated SEO teams. Additionally, reliance on manual approaches fails to persist with trends during non-working days, when SEO teams are unavailable to react to shifts in user behavior. Such reliance can result in missed opportunities for traffic during critical moments, such as viral trends of a query related to a web page the organization owns. Furthermore, manually chosen keywords might reflect human biases.

1.3. Proposed system and its benefits

SEO keywords can be generated using LLMs and embedded into the HTML code of a page [4-6]. AutoTrendyKeywords is proposed as a system that continuously generates keywords using Large Language Models (LLMs) and shortlists the most trending keywords, sustaining the effectiveness of keywords in SEO. The system uses unsupervised usage of LLMs, which allows it to autonomously generate and update SEO keywords without requiring manual input or predefined datasets for training. This continuous process ensures that the system can adapt in real-time to changing trends. The system is transparent as it displays the trend data before selecting the keywords and selects keywords based on a fixed procedure. This transparency enhances trust in the system and enables further improvements. Page content changes rapidly for the majority of web pages [7]. The system provides significant advantages for such dynamic content by continuously updating keywords, thereby ensuring new traffic from relevant audiences. Short and readable URL paths that include keywords are helpful for users in a decision to open a page, as well as for search engines to know what the page is about [8]. Hence, the process involves generating a new relevant path using the LLM based on trending keywords when creating a web page. Constantly adding new paths that redirect to the page might lead to a consequence of retaining too many paths.

1.4. Societal impact of the system

AutoTrendyKeywords benefits organizations and also has a broader positive impact on users and society. It allows a broader number of users to find the desired content using the desired query. The system can generate keywords in the native language of the content, ensuring a presence within search engine results for queries in that language. Such presence broadens access to information in unpopular languages, thus promoting inclusivity and linguistic diversity.

1.5. Related work

Chodak et al. [4] explored the integration of LLMs into SEO activities for online retailers, in addition to content creation and keyword generation. Automating the generation of HTML code and content optimization for search engines is a notable advancement in using LLMs for SEO. However, a research gap exists on how keywords can be automatically changed based on changes in trends. Ziakis et al. [9] have emphasized the efficiency gained through AI for automating SEO tasks such as keyword research, content creation, backlink analysis, and technical aspects of SEO. However, existing literature predominantly addresses the short-term benefits of AI automation, with limited exploration of its efficacy in adapting to long-term fluctuations in search trends. LLMs have gained considerable attention in SEO processes, particularly in e-commerce. Previous research has highlighted several advancements in using LLMs in SEO for content generation and product visibility. However, there are significant research gaps regarding maintaining reasonable visibility. The gaps allow further exploration into the development of unique AI applications for SEO to preserve visibility in search results in the long run.

2. Materials and Methods

2.1 Selecting an LLM and trend data source and loading page content

The first step involves selecting an appropriate LLM that will be capable of generating keywords based on the topic of the web page and its content. Llama 3.1 (70B) [10-11] is selected due to its proven capabilities in natural language generation and handling text. Trend data is sourced from Google Trends [12]. The language selected is English. The page content is written for a blog article and saved in a file. The LLM, trend data source and page content are loaded into the system for the experiment.

2.2 Generating keywords

The LLM is used to generate two sets of keywords. The first set of 10 keywords is generated using the LLM based on the page title and content. The keywords are ensured to be directly related to the content, ensuring these keywords are closely aligned with the core message of the content. These keywords are directly associated with the page's theme and are considered essential SEO keywords. Another set of 5 keywords generated are broader terms related to the title of the content. For example, if the content is about the benefits of apples, the first set of keywords includes keywords such as "apples and health" and "health benefits of

apples," while the second set includes "benefits of eating fruits" and "healthy snacks for digestive health." The keywords from both sets are combined, and any duplicates are removed. The keywords are converted to lower-case to maintain uniformity with search queries since most queries in search engines are in lower-case. The keywords generated in this step are short-tail keywords and have up to 4 words.

2.3 Fetching trend data

Trend data is fetched for each generated keyword for a defined time period of one week. The attributes used from the data are the demand for the keyword at the start and the end of the time period. The trend is calculated as the percentage of change in the demand of the query from the start to the end. The keywords with zero current demand are eliminated and not considered for the next steps, ensuring only active keywords are retained for further utilization. The remaining keywords are then sorted primarily using the growth, based on the assumption that future demand might be the same or higher than the current demand for the most trending queries. If the growth is the same for two queries, they are sorted based on the current demand. This process ensures that only the keywords with a high future potential are prioritized and considered for the next steps. The time period selected is not more than one week since the system can be used to change keywords once a week.

2.4 Generating long-tail keywords from most trending keywords and fetching their trend data

Users frequently use more specific queries called long-tail keywords when querying a search engine. Hence, long-tail keywords are necessary for a wide reach of the page. Long-tail keywords are generated in this step using the LLM based on the top 10 most popular short-tail keywords. The trend data of the long-tail keywords is fetched and processed using a method similar to processing short-tail keywords in the previous step. Each long-tail keyword generated in this step typically contains up to 7 words.

2.5 Generating a description for metadata

A meta description is crucial for SEO as it helps search engines find the summary of the page, which helps find the relevance of the page for a query. A value for metadata description is generated for SEO using the LLM based on the top 5 most demanded keywords from both sets of keywords. The value is ensured to contain all the selected keywords using an auto-retry mechanism. The value is also ensured to exclude any words that are not in the selected keywords. It is crucial to include all the trending keywords in the meta description.

2.6 Generating tags based on keywords, page title, and content

Tags play an essential role in organizing and categorizing the content for both search engines and users. Tags are supported by multiple blogging platforms. Tags play a vital role if the page is published on such platforms, as tags attract more potential readers to the page. The LLM is used to generate a set of 10 initial tags based on the page content. Trend data for each tag is fetched and evaluated using a method similar to fetching trend data for keywords. The top 5 most trending tags are retained, as most blogging sites allow only five tags for each article.

2.7 Generating a title for SEO

A title is generated for SEO using the LLM based on the page content, keywords, and the page title. A distinct title for SEO facilitates the SEO process without altering the page title decided by the author. The SEO title allows better visibility of the page across the search engine results.

2.8 Usage of the keywords in the tags of the HTML page

All the tags are added to the description value to ensure a broader coverage in search engines. HTML code is created for the article. The generated keywords and the description are included in the HTML code using metadata tags to ensure the effectiveness of the utilization of the generated keywords. The first 155 characters of the SEO description are considered as per the limitations of search engines. There is a possibility of adding keywords in HTML hidden heading tags at the start of the page to allow better visibility across search engines.

2.9 Generating relevant paths

URL paths are also critical for SEO, as they contribute to the overall visibility and navigability of the website. Two paths are generated using the LLM based on the page title, content, and keywords. The second path generated would be helpful if the first path is already used. Since the paths are keyword-rich, the search engines can detect the keywords in the URL path, and the users searching using a trending query can easily find the page. Hence, paths make the content more reachable to the intended audience.

3. Results

3.1 Generated keywords

The blog was written based on a research paper on AI. The LLM generated 15 unique keywords after processing the page title and content. The keywords generated are "machine learning," "language models," "explainable predictions," "data-augmented prediction,"

"transparent ai," "interpretable machine learning models," "lml-dap," "large language models," "dataset analysis," "data prediction methods," "explainable ai," "natural language processing applications," "transparent machine learning," "interpretable predictive models," and "data-driven decision making."

These keywords reflect the core focus of the blog and include terms related to cutting-edge trends in machine learning and AI research. The keywords span a variety of important themes. Additionally, these terms align with emerging interests in AI. This keyword selection contains both general concepts and specialized topics within the field of AI, thus broadening the potential audience base.

3.2 Keywords sorted and filtered based on trends

After generating the initial set of keywords, trend data was successfully fetched to assess their popularity and relevance. The results shown in Table I below reveal key insights into the trend data. The term "explainable ai" exhibited the highest growth, with a dramatic increase from 1 to 89, representing an 8800% surge in popularity. The results acknowledge that the content of the blog aligns with trending topics. By focusing on keywords with a demonstrated growth in demand, the SEO strategy is likely to improve the blog's visibility in search results, capitalizing on these upward trends.

Table 1. Keywords sorted and filtered based on trends

Keyword	Oldest Value	Latest Value	Growth
explainable ai	1	89	**8800%**
Large Language Models	37	70	89%
transparent ai	46	84	83%
language models	43	67	56%
machine learning	53	71	34%
dataset analysis	42	51	21%

3.3 Generated long-tail keywords and their trend data

To create more specific search queries, the system generated five long-tail keywords based on the trending short-tail keywords. The keywords generated are "Language models for data prediction," "Language models for machine learning applications," "Training language

models on small datasets," "Language models for data augmentation techniques," and "Using language models for predictive analytics."

The trend data for the generated long-tail keywords is shown in Table II below. Trends revealed that none of them had current demand, with each showing a negative growth trend. The lack of current demand suggests that the keywords relevant earlier are no longer relevant. Hence, they were filtered out before further consideration. The shifts imply that these keywords would have been beneficial if the system had been run in the past, and the keywords trending at a time might eventually lose the trend over time. The changes in the trend emphasize the importance of using the system to continuously monitor trends and adapt to shifting trends.

Table 2. Long-tail keywords sorted and filtered based on trends

Keyword	Oldest Value	Latest Value	Growth
language models for data prediction	1	0	-100%
training language models on small datasets	1	0	-100%
language models for data augmentation techniques	1	0	-100%
using language models for predictive analytics	1	0	-100%

3.4 *Generated metadata description*

To create a concise, SEO-optimized meta description, five of the most trending keywords were considered step. The keywords considered are "language models," "transparent ai," "Large Language Models," "explainable ai," and "machine learning." The description is successfully generated by combining all the keywords and contains all the keywords. The description generated is "Large Language Models transparent ai explainable ai machine learning." By focusing on keywords with high demand, the meta description is successfully optimized to increase the likelihood of the page appearing in relevant search results.

3.5 *Generated tags*

The ten tags generated by the LLM are "Transparent AI," "Explainable AI," "Machine Learning," "Large Language Models," "Natural Language Processing," "AI for Healthcare," "Data Augmented Prediction," "Interpretable AI," "Artificial Intelligence," and "Language Model Learning." The top 5 most demanded tags are considered for the next steps. The trend data of all the generated tags are mentioned below. The top 5 most trending tags are selected for the next steps, which are "AI Model Explainability," "Machine Learning," "Language

Model Learning," "Large Language Models," and "Explainable AI." The keywords mentioned earlier are in lower-case, and the trend data is different for the same word in lower-case. Tags require the words in regular casing instead of lower-case, and the trend data of tags is different from that of keywords.

Table 3. Tags sorted and filtered based on trends

Keyword	Oldest Value	Latest Value	Growth
AI Model Explainability	1	100	**9900%**
Machine Learning	51	72	41%
Language Model Learning	46	65	41%
Large Language Models	44	58	32%
Explainable AI	40	50	25%
AI Predictions	42	48	14%
Transparent AI	57	64	12%

3.6 *Sample HTML file after embedding the values*

The generated keywords, SEO title, description, and keywords were incorporated into the HTML structure of the page to optimize its SEO performance. The HTML code includes trending keywords in relevant tags to ensure effective SEO utilization.

```
<!DOCTYPE html>
<html>
<head>
        <title> (SEO title here) </title>
        <meta charset="UTF-8">
        <meta property="og:title" content="(SEO title here)">
        <meta name="description" content="Content on Large Language Models
transparent ai explainable ai machine learning AI Model Explainability Machine
Learning Language Model Learning Large Lang">
        <meta property="og:description" content="Large Language Models
transparent ai explainable ai machine learning AI Model Explainability Machine
Learning Language Model Learning Large Language Models">
        <meta name="keywords" content="language models, transparent ai,
Large Language Models, explainable ai, machine learning">
        <meta name="viewport" content="width=device-width, initial-
scale=1.0">
</head>
<body>
<h1 style="display: none;"> Large Language Models transparent ai explainable ai
machine learning AI Model Explainability Machine Learning Language Model
```

```
Learning Large Language Models </h1>

<h1> (page title here) </h1>

(page content here)

Keywords: Large Language Models transparent ai explainable ai machine
learning AI Model Explainability Machine Learning Language Model Learning
Large Language Models Explainable AI
</body>
</html>
```

Figure 1. Sample HTML code of the page, excluding the title and content

3.7 Generated relevant URL paths

The paths generated by the LLM are "/explainable-ai-predictions" and "/lml-dap-alternative-ml." The paths are short and successfully include keywords relevant to the page, users, and search engines.

4. Discussion

AutoTrendyKeywords has proven the potential to change conventional practices by implementing a real-time automated system that adapts to emerging trends by updating the keywords based on trends. Using keywords generated by the system disallowed manual selection of keywords and allowed mitigation of human bias to create a more inclusive and fair online ecosystem. A challenge is the cost of running the system. However, the growing affordability of the usage of LLMs ensures that the system is affordable for individual creators and non-profit organizations. The system remains further useful if the page content changes on a regular basis. The system can be modified to generate keywords for content in endangered languages to preserve the visibility of the content by allowing visibility of the content in the languages.

The method does not include analysis of competitors to manipulate rankings to over-rank competitors for profit, as it remains only academic. The aim is to keep a web page discoverable to relevant audiences who intend to find the content and not to subdue the online marketplace for financial gain. AutoTrendyKeywords is not designed for profit-maximizing strategies or LLM-based content optimization, unlike existing SEO tools on the market. The experiment uses only one source of trend data, and more sources of real-time trend data could be used.

5. Conclusions

AutoTrendyKeywords offers a solution to the challenges of SEO by successfully using the power of LLMs to continuously generate keywords and select the trending keywords based on real-time trends to update the keywords. In the experiment, the system generated 15 relevant keywords, with some showing a highly growing trend over the selected time period of one week. The system demonstrated a successful generation and selection of keywords, along with the generation of the title and description of the page for SEO. The system generated URL paths that are short and contain the trending keywords. Filtering keywords based on trend data resulted in the selection of the most impactful terms to maintain SEO performance. The focus on sustaining web page relevance and discoverability for users in the long term remained the core focus and strength of the system.

The approach is set to become essential for SEO professionals and organizations to maintain the relevance of a web page for the long term, stimulating a future where page content continues to remain accessible. As LLMs continue to become more affordable and faster, the system can be run frequently. The system can be tested for videos since keywords can be generated based on descriptions and subtitles. With future research, the system can be modified to predict future trends based on the historical data of scheduled regular events and festivals to change keywords on a schedule. The system can be modified to generate the keywords for the images using multimodal LLMs to enhance sustained visibility across the results of image searches.

Appendix A

1. Prompt templates used to process using LLMs

```
Page content:
{page_content}
---
Page title: {page_title}
```

Generate 10 SEO keywords for the above page based on the page content.

All keywords must be in the same language as the page title and content.
Respond with a list of keywords separated by commas inside tags like this:
Sample response:
<keywords>
        keyword1, keyword2, keyword3
</keywords>

Figure A1. Prompt template to generate keywords based on the page content

Page content:
{page_content}

Page title: {page_title}
Find five related topics to the above page.

Respond with a list of SEO keywords separated by commas inside tags like this:
Sample response:
<keywords>
        keyword1, keyword2, keyword3
</keywords>

All keywords must be in the same language as the page title and content. Keywords
should not be directly related to the content but should be related to the topic.

Figure A2. Prompt template to generate keywords related to the topic of the page

Top trending short-tail SEO keywords:
`{top_short_keywords}`

Topic: `{page_title}`

Generate five long-tail SEO keywords for each of the top trending short-tail
keywords. Generate those that might have more demand.
Respond with a list of long-tail SEO keywords separated by commas inside tags
like this:
Sample response:
<keywords>
        keyword1, keyword2, keyword3
</keywords>

Figure A3. Prompt template to generate long-tail keywords based on trending short-tail

keywords

Act as an SEO expert.
Combine multiple keywords without altering their content.
For example, "A" + "A B" + "B" + "B C D" -> "A B C D".
For example, input: ['uses', 'uses of car',
        'cars', 'cars in the world', 'drive a car']
output: <keywords>uses of cars in the world drive a car</keywords>

Use the following SEO keywords as input:
<keywords>{top_keywords}</keywords>

Important Rules:
1. Combine only the keywords provided. **Do not add new words**.
2. **All input keywords** must be present in the output exactly as provided. Do not change, modify, or substitute them.
3. Handle duplicates intelligently by removing redundancy (e.g., "cars" + "cars in the world" -> "cars in the world").
4. The order of keywords can be flexible but should maintain logical flow and readability.
5. The final output should not exceed the total number of characters of the original keywords combined.
All the input keywords that must exist in the output keyword:
<keywords>{top_keywords}</keywords>

Respond with a single combined keyword inside tags like this:
Sample response:
<keywords>(keyword here)</keywords>

Figure A4. Prompt template to generate metadata description for SEO

Page content:
{page_content}

Page title: {page_title}

Act as an SEO expert. Create a list of the ten most popular tags for the blog post.
Generate in regular case. Tags need not be in lower-case.
Respond with a list of tags separated by commas inside tags like this:
Sample response:
<tags>
        tag1, tag2, tag3
</tags>

Figure A5. Prompt template to generate tags

Page content:
{page_content}

SEO keywords:
`{keywords}`
Page title: {page_title}

Generate an SEO title for the blog post based on the keywords.
Respond with the title inside tags like this:
Sample response:
<title>(title here)</title>

Figure A6. Prompt template to generate a title for SEO

```
Page content:
{page_content}
---
SEO keywords: `{keywords}`
Page title: {page_title}
Example path: /benefits-of-apples/

Generate meaningful paths for the blog post.
A path must have the least possible number of words (max 3).

Return a list of 2 paths line-by-line, inside tags like this:
Sample response:
<paths>
        /alias1 \n /alias2
</paths>
```

Figure A7. Prompt template to generate paths

References

[1]   M. Nagpal and J. A. Petersen, Keyword Selection Strategies in Search Engine Optimization: How Relevant is Relevance?, *J. Retailing* **2021**, 97, doi: 10.1016/j.jretai.2020.12.002.

[2]   S.-P. Jun, H. S. Yoo, and S. Choi, Ten years of research change using Google Trends: From the perspective of big data utilizations and applications, *Technological Forecasting and Social Change* **2018**, 130, doi: 10.1016/j.techfore.2017.11.009.

[3]   Here's When – and How – You Should Hire an SEO Expert for Your Business, Entrepreneur. Available online: https://www.entrepreneur.com/growing-a-business/hiring-an-seo-expert-will-transform-your-business-heres-7/435990 (accessed on 12 Oct. 2024).

[4]   G. Chodak and K. Błazyczek, Large Language Models for Search Engine Optimization in E-commerce. In Proceedings of the Advanced Computing Conference, Cham, Switzerland, Mar. 2024: Springer Nature, doi: 10.1007/978-3-031-56700-1_27.

[5]   G. Matošević, J. Dobša, and D. Mladenić, Using Machine Learning for Web Page Classification in Search Engine Optimization, *Future Internet* **2021**, 13, doi: 10.3390/fi13010009.

[6]   R. Maragheh et al., LLM-TAKE: Theme-Aware Keyword Extraction Using Large Language Models. In Proceedings of the 2023 IEEE International Conference on Big Data (BigData), Los Alamitos, CA, USA, 15-18 December 2023: IEEE Computer Society, doi: 10.1109/BigData59044.2023.10386476.

[7]   V. Mallawaarachchi, L. Meegahapola, R. Madhushanka, E. Heshan, D. Meedeniya, and S. Jayarathna, Change Detection and Notification of Web Pages: A Survey, *ACM Comput. Surv*. **2020**, 53, doi: 10.1145/3369876.

[8]   SEO Friendly URLs. Available online: https://backlinko.com/hub/seo/urls (accessed on 12 Oct. 2024).

[9]   C. Ziakis and M. Vlachopoulou, Artificial Intelligence's Revolutionary Role in Search Engine Optimization, In Proceedings of the The International Conference on Strategic Innovative Marketing and Tourism, Cham, Switzerland, Jun. 2024: Springer Nature, doi: 10.1007/978-3-031-51038-0_43.

[10]   Llama 3.1 [Language model]. Available online: https://github.com/meta-llama/llama3/blob/main/MODEL_CARD.md (accessed on 12 Oct. 2024).

[11]   A. Dubey et al., The Llama 3 Herd of Models, 2024, arXiv:2407.21783. Available online: https://arxiv.org/abs/2407.21783 (accessed on 12 Oct. 2024).

[12]   Google Trends [Data source]. Available online: https://trends.google.com/trends/explore?date=now%207-d (accessed on 12 Oct. 2024).

✉ **editor@iaeme.com**