AI PHISHING EMAIL CHAIN AND LIGITIMACY TESTER USING SCRAPED LINKEDIN PROFILES

by

AUSTIN PAULLEY

Advisor
DR. PLANTE

A senior research paper submitted in partial fulfillment of the requirements.
for the degree of Bachelor of Science
in the Department of Mathematics and Computer Science
in the College of Arts and Science
at Stetson University
DeLand, Florida

Fall Term
2024

ACKNOWLEDGMENTS

This project builds upon the foundational work of Justin T. Snider Curtis, specifically his project titled "Generating Spear Phishing Emails Based on Web Scraped LinkedIn Information Using OpenAI API." The original project, which served as the basis for this expanded work, can be accessed at Justin's GitHub Repository. I extend my gratitude for the groundwork laid in the beginning of the project.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

This research investigates the intersection of artificial intelligence (AI) and cybersecurity, focusing on the potential misuse of AI in generating professional email communication. Utilizing a Python-based application integrated with OpenAI's GPT-3.5 Turbo and GPT-4 Turbo models, the study explores the influence of AI temperature settings and email chain lengths on the legitimacy of generated content. By combining advanced data scraping techniques with MongoDB for structured data storage, the system provides a seamless workflow from data acquisition to comprehensive results analysis.

The findings underscore a critical dichotomy: while deterministic AI configurations yield highly professional and credible emails, elevated temperature settings compromise legitimacy, potentially mimicking the unpredictability of human behavior. This raises significant cybersecurity concerns, particularly regarding the exploitation of AI to automate phishing attacks. The study highlights how AI-generated emails, when combined with contextualized data, can evade traditional detection mechanisms by replicating authentic linguistic patterns and professional tones.

This research emphasizes the need for initiative-taking cybersecurity measures to mitigate the risks posed by AI in crafting deceptive communications. By examining the dual-use nature of AI technologies, this work contributes to the broader discourse on the ethical and secure deployment of AI, advocating robust countermeasures to address emerging threats in digital communication.

1. INTRODUCTION

In the contemporary digital era, artificial intelligence (AI) emerged as a transformative force, driving rapid advancements across industries while simultaneously sparking significant debate about its implications. Neural networks, the foundation of AI systems, have demonstrated unprecedented accuracy and capability, evolving remarkably within a short span of time. AI's potential to innovate alongside its propensity to disrupt underscores the urgency of understanding and leveraging its capabilities responsibly. This project, through a meticulously designed Python application, seeks to highlight the practical power of modern AI while addressing the complex challenges it may pose.

The aim of this tool is to illustrate the immense capabilities of AI in professional communication, achieved through the development of a streamlined, efficient application comprising just over 900 lines of Python code. Despite its simplicity, the tool demonstrates sophisticated functionalities, such as generating tailored email chains and evaluating their legitimacy. By automating these processes, the application highlights the accessibility and transformative potential of AI in both positive and negative aspects.

Utilizing this application requires three straightforward inputs: a personal OpenAI API key, a connection link to a MongoDB cluster, and a LinkedIn profile URL for data collection. Once configured, the tool operates with remarkable efficiency, generating

multiple email chains based on user-defined parameters such as AI temperature (creativity level) and chain size. Beyond generation, the application provides a detailed legitimacy analysis for each email chain, utilizing advanced AI prompts to assess authenticity, professionalism, and alignment with the specified context. The results generated quickly, offering users actionable insights within minutes.

This project is not merely a demonstration of AI's technical capabilities; it serves as a critical exploration of the potential applications and ethical considerations surrounding AI's integration into professional workflows. By automating complex tasks such as personalized communication and content evaluation, the application underscores AI's potential to revolutionize industries ranging from marketing to cybersecurity. At the same time, it highlights the risks inherent in deploying such powerful tools, including misuse, privacy concerns, and the erosion of human-centric communication practices.

In addressing these themes, this project exemplifies the balance between innovation and responsibility, encouraging users and developers alike to critically engage with AI's role in shaping the future. It seeks to contribute to the broader discourse on AI by demonstrating both its transformative potential and the need for thoughtful regulation and ethical practices in its implementation.

2. RELATED WORK

The foundational aspects of web scraping in the context of professional networking are covered by Zhang et al. (2019), who provides a comprehensive overview of the technological methodologies and ethical concerns involved in balancing data accessibility with privacy. They highlight innovative strategies for responsibly navigating these challenges, such as selective data scraping and user consent frameworks. We used this to collect ideas about future improvements that can be made to the python code. [Zha19]. Building upon these principles, Moreno, and Castillo (2022) delve into adaptive data extraction algorithms specifically tailored for enhancing AI applications. Their study focuses on algorithms capable of dynamically adjusting to changing web page structures, significantly improving the precision and efficiency of data collection [Mor22]. Additionally, Anderson and Lee (2023) explore the integration of real-time data analytics into scraping technologies, providing case studies where real-time analytics have enabled immediate adjustments to scrape strategies, enhancing data collection efficiency and accuracy. This paper gives great ideas on improving efficiency within the code and gives insight into improvement in current code. [And23].

Personalized communication plays a pivotal role in professional networking, as discussed by Smith and Johnson (2018). Their research emphasizes the effectiveness of

personalized emails in increasing response rates and fostering deeper professional relationships [Smi18]. Extending this topic, Thompson et al. (2020) investigate the psychological impacts of personalized email content, analyzing how subtle manipulations are. in email communications can significantly alter recipient perceptions and increase susceptibility to sophisticated phishing attacks [Tho20]. Further, Bennett and Khanna (2021) provide a comparative analysis of automated versus manually crafted emails, highlighting the superior engagement metrics associated with AI-enhanced personalized emails in various networking scenarios [Ben21]. These three sources highlight the importance of the email chain portion of the project as they give examples of the importance of emails and tests on email phishing.

The integration of AI in content creation has transformed the approach to digital communication, as outlined by Patel and Singh (2020). They review the use of natural language processing algorithms essential for creating authentic and contextually relevant email content [Pat20]. Complementing this, Kim, and Park (2021) discuss the integration of machine learning with natural language generation, enabling the creation of dynamic content that adapts to the recipient's recent activities, thereby enhancing the effectiveness of phishing simulations [Kim21]. Walters and Gomez (2024) explore the use of deep learning to predict user behavior, allowing phishing simulations to anticipate and adapt to potential victim responses, thereby improving the targeting precision of phishing

campaigns [Wal24]. These three sources focus on the AI side of the project and Walters heavily influenced the new prompt and being more specific mimicking his work.

The technical aspects of email automation are critically analyzed by Lee and Hwang (2017), who focus on SMTP and API-based email services and their implications for large-scale email campaigns [Lee17]. Garcia and Lopez (2019) address the security vulnerabilities in email transmission protocols, offering insights into how these can be exploited to enhance the realism of phishing simulations and how they can be secured to prevent misuse [Gar19]. Additionally, Nova and Briggs (2023) introduce the concept of using blockchain technology to secure automated email systems, discussing how blockchain can provide a verifiable and immutable record of email transactions, enhancing trust and security in digital communications [Nov23]. These three sources focus on the future part of the project as it will feature sending the email to an actual email with a working phishing link. This email will be a personal email and not be randomly sent out.

Cloud-based solutions are integral to managing the vast amounts of data generated from these activities. Gupta and Chandra (2021) emphasize the importance of robust and scalable data management solutions like MongoDB Atlas, which ensures that data is stored securely and managed efficiently, providing a scalable environment for handling large datasets [Gup21]. O'Connor and Schmidt (2022) review advancements in

distributed databases, highlighting their role in improving data integrity and providing capabilities necessary for real-time data processing during phishing simulations [OCo22]. Ellison and Moriarty (2024) delve into the implications of cloud security measures on data privacy, particularly focusing on the protocols and regulations that govern data storage in cloud environments, which are crucial for maintaining user trust in networking platforms [Ell24]. With not a strong knowledge of databases these sources helped me to really narrow down going back to MongoDB and not a new database system. With the critical vulnerabilities this code creates in the phishing email this was the best option based on these papers.

3. IMPLEMENTATION

The development of this project demonstrates the seamless integration of various functions, working together to achieve automated email generation and legitimacy analysis. This intricate system begins with data scraping from LinkedIn profiles, progresses through email generation and analysis, and culminates in result storage and visualization. Each function plays a vital role in ensuring the tool operates effectively, delivering actionable insights to the user. Below is a detailed explanation of how these components interact to form a cohesive workflow.

Step 1: User Input and Data Initialization

The system begins by accepting key user inputs through the graphical user interface (GUI). The setup_gui () function lays the foundation for user interaction, offering sliders, dropdowns, and text fields to define parameters such as email chain size, temperature, and AI models for generation and testing. The add_new_user () function is pivotal at this stage, allowing users to input a LinkedIn profile URL for data scraping. This step ensures the system has the necessary raw data to proceed with email generation. The update_user_dropdown () function complements this initialization by dynamically fetching and displaying stored user profiles from the MongoDB database. This ensures that users can easily select existing profiles, streamlining the workflow for repeated tests.

Step 2: Data Scraping from LinkedIn

Once a LinkedIn profile URL is provided, the system utilizes scraping algorithms embedded within the add_new_user () function. This function extracts key details such as job titles, company names, educational background, and skills from the LinkedIn profile. These details are then formatted and stored in a MongoDB database, creating a structured dataset that serves as the foundation for email generation. The scraping step ensures that email content is personalized and contextually relevant, leveraging the extracted data to craft professional, tailored communications. For instance, job titles and company names are directly incorporated into email prompts, lending authenticity to the generated content.

Step 3: Email Generation with OpenAI API

With the user profile data stored and parameters defined, the ask_email_chain_count () function prompts the user to specify the number of email chains to generate. This input, combined with temperature settings and AI model selections, is passed to the generate_email_chain () function. This function serves as the core of the system, constructing email chains by Generating a primary email based on the extracted LinkedIn data. Iteratively generating responses to the primary email to create a complete chain. These emails are generated using the OpenAI API, which interprets the prompts and returns content that adheres to professional standards. The temperature slider, configured in the GUI, determines the creativity of the generated emails, balancing

between deterministic and imaginative outputs. This step transforms static LinkedIn data into dynamic, contextualized email communications.

Step 4: Legitimacy Analysis of Generated Emails

Once an email chain is created, it is passed to the evaluate_chain_legitimacy () function for analysis. This function uses the OpenAI API to assess the authenticity and professionalism of the email chain. Details of this are that the AI is set to 0.7 as according to research this is one of the most human-like temperatures. A detailed prompt instructs the AI to evaluate the following criteria:

- Language and Tone Consistency: Ensuring the emails maintain a professional tone.
- Content Relevance: Verifying that the emails align with the intended purpose.
- Authenticity Indicators: Detecting signs of phishing or suspicious behavior.
- Formatting and Grammar: Checking for errors that may compromise credibility.
- Behavioral Indicators: Identifying any pressure tactics or inconsistencies.

The AI then generates a legitimacy score (0–100) along with a comprehensive explanation, detailing the strengths and weaknesses of the email chain. This analysis provides actionable insights into the quality and credibility of the emails generated.

Step 5: Storing Results in MongoDB

After generating and analyzing the email chain, the results are stored in a MongoDB database using the store_chain_results () function. This function organizes the data into collections dynamically based on parameters such as chain size, temperature, and AI models used. For example, a collection might be named *results_chain4_temp1_0_gengpt_3_5_turbo_testgpt_3_5_turbo.*

Each stored document includes:

- The user's name.
- Temperature and chain size settings.
- The email chain and its legitimacy analysis.
- The AI models are used for generation and testing.
- This structured storage ensures easy retrieval and categorization, allowing users to review results based on specific criteria.

Step 6: Visualization and Insights

The last step involves presenting the results to the user. The open_results_window () function provides an interactive interface for filtering and analyzing stored data. Using sliders and dropdowns, users can refine their query based on temperature, chain size, and AI models. The filtered results include:

- The number of tests matches the criteria.
- The average legitimacy score for the selected tests.

This visualization empowers users to draw meaningful insights, identifying trends in AI performance and evaluating the effectiveness of different parameter settings.

Additional Features:

The application also incorporates supplementary features designed to enhance functionality and user experience. These include an automatic stop mechanism that halts email generation if a process stalls or requires intervention. Additionally, a text output option is available, allowing users to export entire email chains into a neatly formatted text file for offline review. Lastly, the system includes seamless looping functionality, enabling users to conduct consecutive tests without the need to restart the application.

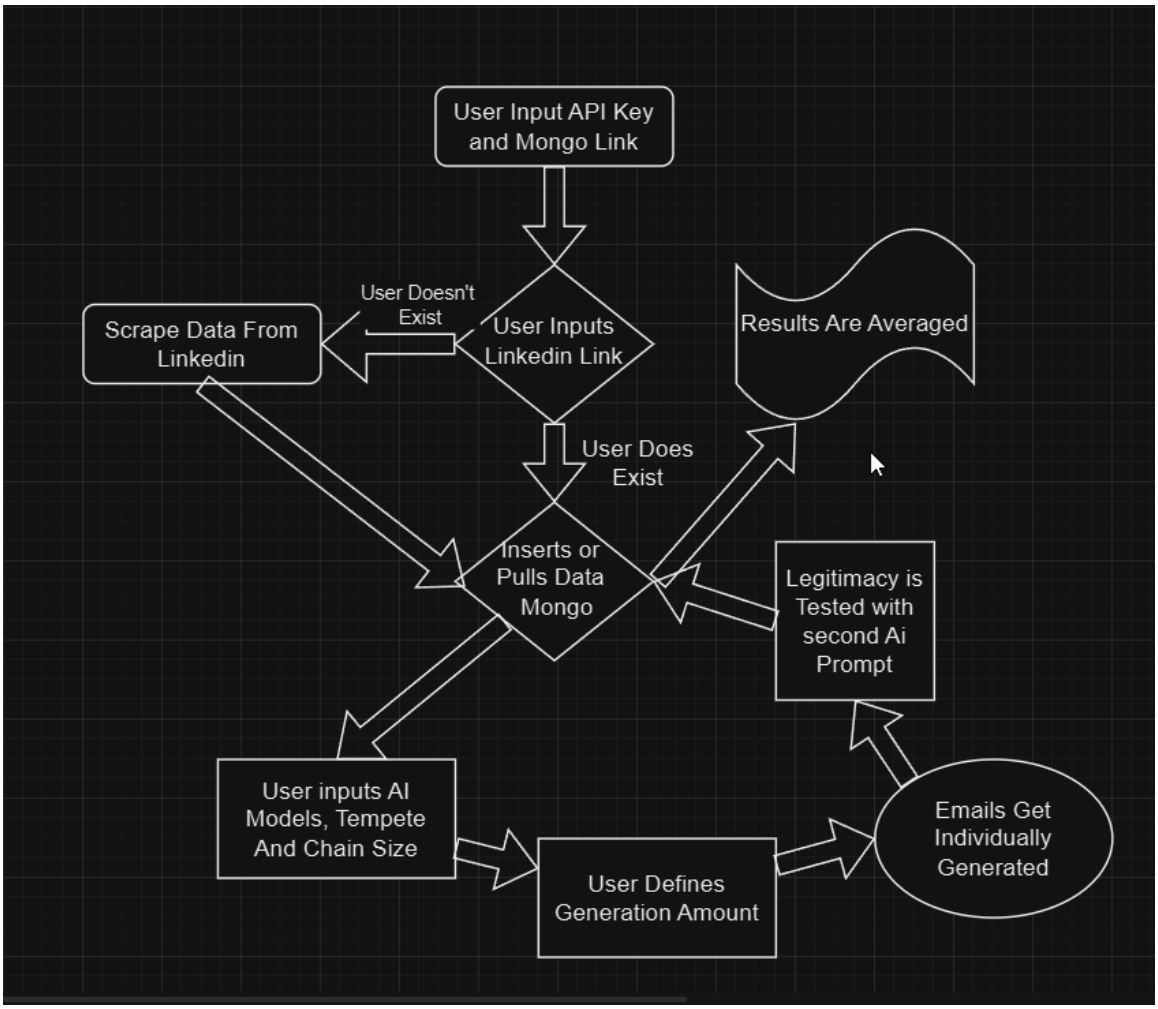

Figure 1: Functions of the Application

The diagram above visualizes the iterative loop that the AI follows each time it generates an email chain. This process adheres to the detailed steps outlined earlier. The system is designed to ensure that the contents of each generated email are systematically stored before proceeding to the next, enabling precise and reliable data tracking throughout the workflow.

Prompts:

The construction of the primary prompt is intentionally simple and straightforward. It is structured as follows: “Write an email to {profile_data.get('name', 'the person')} inviting them to attend a job fair event from ABCC College or Alabama Berkley Community College. The location and time are to be announced and will be found on ABCCEvents.net. The email should come from Tom Timothy, Dean at the school, and direct the person to the website for event information."

This foundational prompt provides clear instructions for generating a professional email. Additionally, it allows flexibility by enabling the inclusion of various personalized data sets, making the email more specific and relevant to the recipient. The goal of this prompt is to maintain simplicity while retaining the potential to mislead in scenarios where legitimacy testing is required.

In contrast, the legitimacy testing prompt is significantly more advanced, incorporating multiple layers of analysis. It is designed to assess the authenticity and credibility of the generated emails based on detailed criteria. The AI test evaluates each email chain using the following aspects:

- Language and Tone Consistency: Does the email maintain a professional tone aligned with its intended purpose?

- Content Relevance: Are the details specific, actionable, and pertinent to the subject matter?
- Authenticity Indicators: Do the sender details appear legitimate? Are there any signs of phishing or deceptive practices?
- Formatting and Grammar: Does the email exhibit proper grammar, spelling, and structure? Are there any formatting issues that might undermine its credibility?
- Behavioral Indicators: Are there any pressure tactics, inconsistencies, or other red flags that could suggest illegitimacy?

The legitimacy analysis is conducted through this detailed framework, and the testing AI provides a comprehensive evaluation along with a Legitimacy Score (ranging from 0 to 100). This robust testing process ensures that each email chain undergoes rigorous scrutiny, assessing its authenticity and quality with precision.

4. RESULTS

The results are stored in MongoDB and organized based on four key identifiers: the generation temperature, the size of the email chain for responses, the generating AI model, and the testing AI model. These identifiers are also required to retrieve the results displayed in the startup screen popup. In MongoDB, all data is neatly sorted and stored. Here is an example of the stored data structure for reference.

Figure 2: Example of Stored Data

results_chain0_temp0.5_gengpt_gpt-4-turbo_testgpt-4-turbo
Id: Object Id('673ec392be46a756d819d611')
Name: "Austin Paulley"
Temperature: 0.5
Legitimacy score: 85
Legitimacy analysis: (AI Response)
Email Chain: (Array with full Email)
Generating AI: gpt-4-turbo
Testing AI: gpt-4-turbo

To evaluate the results of the testing process, a total of four distinct tests were conducted, utilizing both GPT-3.5-turbo and GPT-4-turbo models. Two of these tests specifically targeted the effect of temperature settings on AI performance, while the remaining two focused on analyzing the impact of email chain size on response quality

and consistency. These tests were designed to systematically examine how variations in key parameters influence the behavior and output of each AI model.

Table 1: 3.5 Turbo Temperature Test

Chain = 0, Temp. = 0 – 2.0, Generation GPT = 3.5 Turbo, Test GPT = 3.5 Turbo

Temperature	Average Legitimacy Score (%)	Total Test
0.0	94.17	200
0.5	94.38	200
1.0	93.95	200
1.5	91.17	200
2.0	29.55	200

This study examines the influence of varying AI temperature settings on the legitimacy of generated email content. Using consistent testing parameters, the objective is to analyze how temperature affects the balance between creativity and professionalism in AI-generated emails. The evaluation utilized GPT-3.5 Turbo for both content generation and legitimacy testing, with a total of 200 tests conducted for each temperature setting. The following sections provide a detailed analysis of the results and their implications. The findings reveal a clear relationship between temperature settings and email legitimacy scores:

For Temperature 0.0 the lowest temperature setting, the generated emails achieved an average legitimacy score of 94.17%. The deterministic nature of this setting ensured highly structured and professional content, making the emails appear authentic and dependable. This demonstrates that low temperatures favor consistent formatting and tone, which are crucial for maintaining high legitimacy. Secondly, temperature 0.5 with a slight increase in temperature, the average legitimacy score rose marginally to 94.38%. The moderate variability introduced at this level allowed for subtle improvements in naturalness without compromising professionalism. This balance between structure and creativity proved effective for generating high-legitimacy content. Next, temperature 1.0 at a temperature setting of 1.0, the AI demonstrated a balanced approach between deterministic and creative behavior, resulting in an average legitimacy score of 93.95%. While the emails maintained an elevated level of legitimacy, the slight decrease in score indicates the introduction of minor inconsistencies as creativity was prioritized.

Furthermore, for temperature 1.5 a notable decline in legitimacy was observed at this setting, with an average score of 91.17%. The increased creativity led to less structured and less professional outputs, which negatively impacted the perceived authenticity of the emails. This highlights the trade-off between creativity and professionalism as temperature increases. Lastly, for temperature 2.0 at the highest temperature setting, the average legitimacy score dropped significantly to 29.55%. The high variability produced content that, while creative, lacked the professionalism and tone necessary for legitimate

communication. This demonstrates that excessive randomness undermines the authenticity and reliability of the output.

Table 2: 4 Turbo Temperature Test

Chain = 0, Temp. = 0 – 2.0, Generation GPT = 4 Turbo, Test GPT = 4 Turbo

Temperature	Average Legitimacy Score (%)	Total Test
0.0	84.50	20
0.5	84.85	20
1.0	81.85	20
1.5	10.50	20
2.0	4.00	20

This analysis explores the impact of varying AI temperature settings on the legitimacy of generated email content. By employing consistent parameters for testing, the study investigates how temperature affects the trade-off between creativity and professionalism in AI outputs. The evaluation was conducted using GPT-4 Turbo for both email generation and legitimacy testing, with 20 tests performed at each temperature level. The results highlight a clear relationship between temperature settings and the legitimacy of emails, providing valuable insights into optimizing AI-generated communication.

Firstly, Temperature 0.0: At the lowest temperature, the generated emails achieved an average legitimacy score of 84.50%. This deterministic setting ensured highly structured and professional outputs, making the emails appear dependable and authentic. The findings underscore that low temperatures favor consistent formatting and tone, which are essential for maintaining high legitimacy. Secondly, Temperature 0.5: With a slight increase in temperature, the average legitimacy score rose marginally to 84.85%. The moderate variability introduced at this level added a touch of naturalness to the content while preserving its professionalism. This balance between structure and creativity proved effective for generating high-legitimacy email content. Next, Temperature 1.0: At a temperature setting of 1.0, the AI exhibited a balanced approach between deterministic and creative behaviors, resulting in an average legitimacy score of 81.85%. While the emails retained a high degree of legitimacy, the slight decrease in score suggests the introduction of minor inconsistencies as creativity was prioritized over strict structure. Furthermore, Temperature 1.5: A significant decline in legitimacy was observed at this setting, with the average score dropping to 10.50%. The increased creativity at this level led to less structured and less professional outputs, undermining the perceived authenticity of the emails. This result highlights the trade-off between creativity and professionalism as the temperature increases. Lastly, Temperature 2.0: At the highest temperature setting, the average legitimacy score dropped dramatically to 4.00%. The high degree of variability resulted in outputs that, while creative, lacked the professionalism and tone necessary for legitimate communication. This demonstrates that

excessive randomness negatively impacts the authenticity and reliability of AI-generated content.

Table 3: 3.5 Turbo Chain Test

Chain = 0-5, Temp. = 0.5, Generation GPT = 3.5 Turbo, Test GPT = 3.5 Turbo

Chain Size	Average Legitimacy Score (%)	Total Test
0	94.38	200
1	94.45	200
2	94.92	200
3	93.90	200
4	93.92	200
5	94.35	200

This analysis investigates the impact of varying email chain sizes on the legitimacy of AI-generated email content. The study employs a consistent temperature setting of 0.5, with GPT-3.5-turbo used for both email generation and legitimacy evaluation. The focus is to determine how increasing the length of email chains affects the professionalism and authenticity of the generated responses. A total of 200 tests were conducted for each chain size, from 0 to 5, and the results reveal key insights.

At the initial chain size of 0, the generated emails achieved an average legitimacy score of 94.38%. This result demonstrates that single, standalone emails produced by the AI were professional, structured, and highly dependable. The low complexity of generating an independent email contributed to the high legitimacy score. With the introduction of a single response in the email chain, the average legitimacy score improved slightly to 94.45%. This increase suggests that the AI could maintain context and professionalism when generating a simple reply, indicating effective handling of short, two-email chains. As the chain length increased to include two responses, the average legitimacy score peaked at 94.92%. This result indicates the AI's ability to produce coherent and professional multi-response chains, leveraging the context effectively to improve the natural flow of communication. For chain size 3, the average legitimacy score decreased slightly to 93.90%. This decline reflects the growing complexity of maintaining professionalism and context across a longer chain of emails, where inconsistencies or slight deviations may have occurred. At chain size 4, the legitimacy score remained stable at 93.92%. The results indicate that the AI was able to manage moderately extended chains without significant degradation in the quality of the responses. For the longest chain size evaluated at 5, the legitimacy score rebounded slightly to 94.35%. This result highlights the AI's capacity to generate professional content even in extended communication chains, maintaining a high standard of authenticity and relevance.

The data suggests that the AI performs consistently well in generating professional and legitimate email content across different chain sizes, with the highest score observed for chain size 2. While minor fluctuations occurred with increasing complexity, the overall legitimacy scores remained high, highlighting the AI's robust capability to manage contextual information and maintain a professional tone in extended email chains.

Table 4: 4 Turbo Chain Test

Chain = 0-5, Temp. = 0.5, Generation GPT = 4 Turbo, Test GPT = 4 Turbo

Chain Size	Average Legitimacy Score (%)	Total Test
0	84.85	20
1	84.60	20
2	83.75	20
3	82.65	20
4	88.00	20
5	85.65	20

This analysis explores the impact of increasing email chain size on the legitimacy of AI-generated email content using GPT-4 Turbo for both generation and testing, with a fixed temperature setting of 0.5. The purpose is to evaluate how chain size, which represents the number of responses within an email chain, affects the perceived

legitimacy of the communication. The results highlight subtle fluctuations in legitimacy scores across varying chain sizes, offering insights into the dynamics of structured email interactions.

With no additional responses, the generated email achieved an average legitimacy score of 84.85%. This indicates a strong level of professional and coherent content when limited to the initial email alone. Introducing one response in the chain resulted in a slight decline to an average legitimacy score of 84.60%. This minor decrease suggests the possibility of slight deviations in tone or formatting as the AI generates conversational replies. With 2 responses, the legitimacy score dropped further to 83.75%. The reduction may be attributed to increased variability in maintaining a professional tone across multiple exchanges. A continued decline is observed at 3 with an average score of 82.65%, reflecting the challenges of maintaining coherence and professionalism across extended interactions. With 4 responses notably, the legitimacy score increased to 88.00%. This anomaly could indicate improved contextual understanding or better consistency in the generated responses within a more extensive conversational framework. At the maximum chain size evaluated, the average legitimacy scores slightly decreased to 85.65%. Despite the reduction, the score remains high, suggesting that GPT-4 Turbo can manage extended chains with reasonable professionalism.

This group demonstrates that while larger chain sizes introduce challenges in maintaining a prominent level of legitimacy, GPT-4 Turbo performs well in generating

coherent and professional email chains. The slight improvements and declines across the range of chain sizes highlight the nuanced behavior of the model in extended interactions.

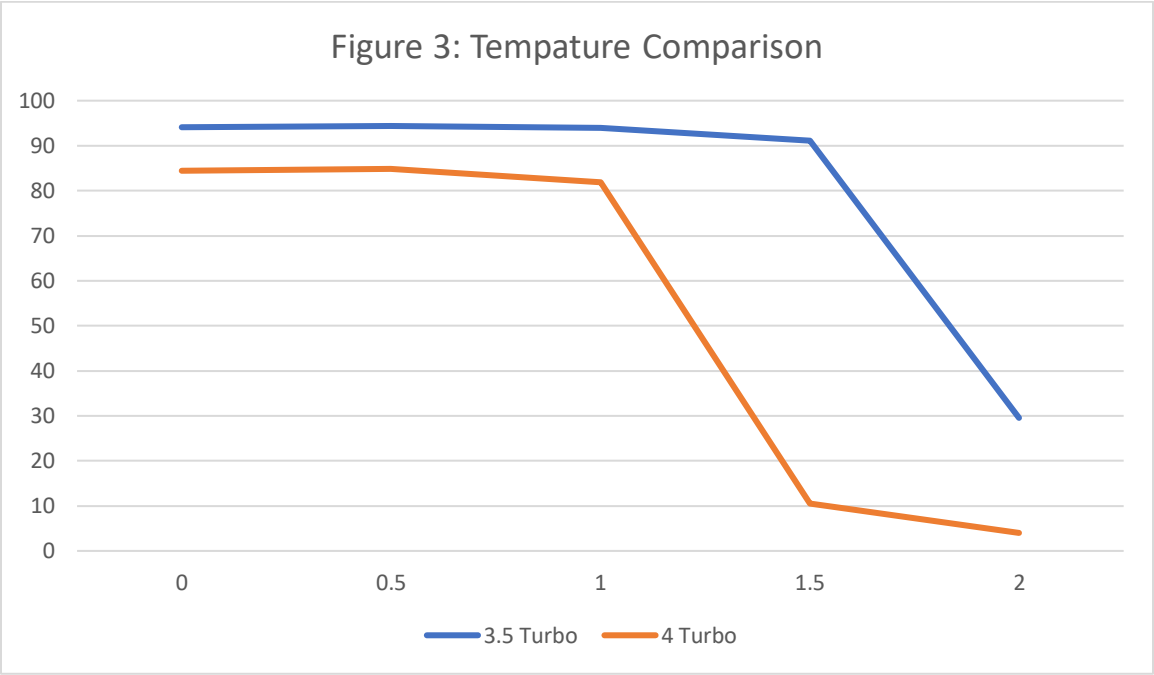

Figure 2 illustrates the impact of varying temperature settings on the legitimacy scores of AI-generated email content. The comparison involves two models, GPT-3.5 Turbo (represented by the blue line) and GPT-4 Turbo (represented by the orange line). The temperature settings range from 0.0 to 2.0, and the corresponding legitimacy scores are plotted on the y-axis.

The data reveals distinct patterns between the two models. At lower temperature settings (0.0 and 0.5), GPT-3.5 Turbo maintains consistently high legitimacy scores, above 90%, indicating strong reliability and professionalism in its output. In contrast, GPT-4 Turbo demonstrates slightly lower scores, ranging from 84% to 85%, reflecting a comparable but slightly less structured approach to content generation.

As the temperature increases to 1.0, both models begin to show declines in legitimacy scores. GPT-3.5 Turbo retains a score close to 90%, while GPT-4 Turbo experiences a more noticeable drop, falling to around 82%. This indicates that the higher creative freedom introduced at this setting impacts GPT-4 Turbo's ability to maintain professional tone and formatting. At temperatures of 1.5 and 2.0, the differences between the models become more pronounced. GPT-3.5 Turbo's legitimacy scores decrease sharply to 1.5 but remain above 90%. However, at 2.0, it drops significantly to below 30%, reflecting the challenges of producing structured content at the highest randomness levels. GPT-4 Turbo, on the other hand, exhibits a steeper decline at these higher settings, with scores dropping to approximately 10% at 1.5 and stabilizing around 4% at 2.0, suggesting that excessive variability severely affects the perceived legitimacy of its outputs.

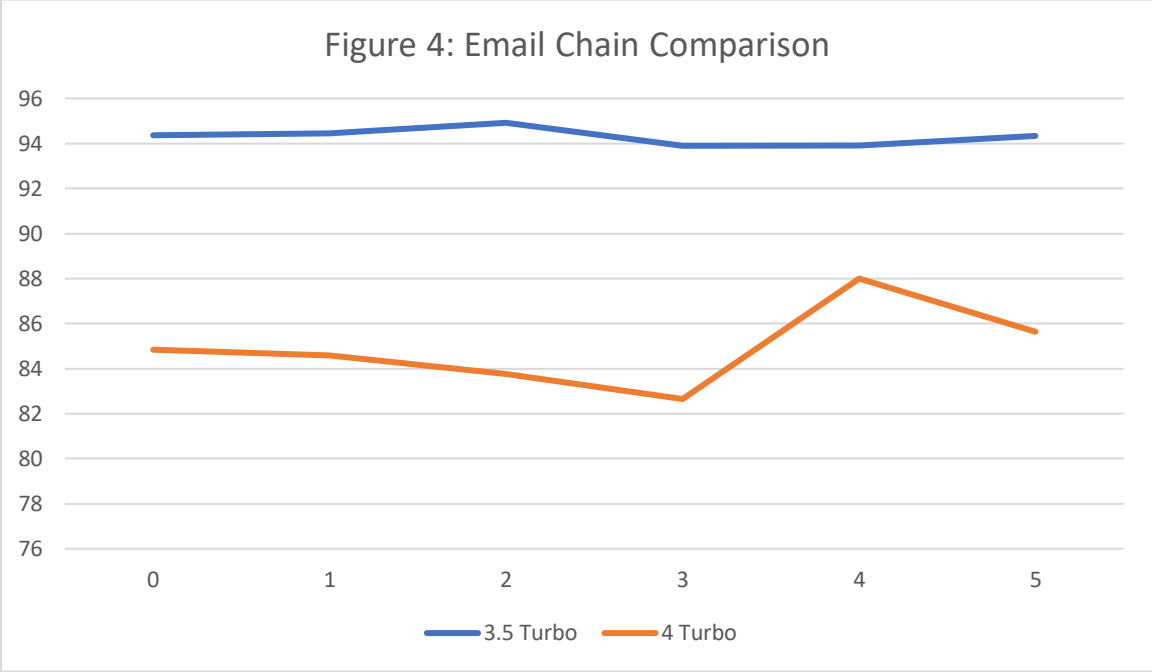

Figure 3 illustrates the impact of chain size on the average legitimacy score for email generation and testing using the GPT-3.5 Turbo and GPT-4 Turbo models at a constant temperature of 0.5. The graph highlights variations in performance as the chain size increases from 0 to 5, providing insights into the consistency and reliability of each model over multiple generations.

The GPT-3.5 Turbo model demonstrates an elevated level of stability across all chain sizes, maintaining legitimacy scores consistently above 93%. Its performance peaks

at a chain size of 2 with a legitimacy score of 94.92% before slightly declining and stabilizing around 94% for larger chain sizes.

Conversely, the GPT-4 Turbo model shows more variability. Starting at a legitimacy score of 84.85% for a chain size of 0, its performance gradually declines as the chain size increases to 3, reaching a low of 82.65%. Interestingly, the model experiences a sharp improvement at a chain size of 4, achieving its highest legitimacy score of 88%. However, this score decreases slightly to 85.65% at a chain size of 5.

5. CONCLUSION

While the primary focus of this research has been the evaluation and optimization of AI-generated emails, it is important to address the potential cybersecurity risks associated with such technology. AI models like GPT-3.5 Turbo and GPT-4 Turbo, when calibrated for creativity and relevance, can produce highly convincing and professional emails. This capability, however, can be misused to generate malicious emails for phishing attacks, social engineering, or other forms of cybersecurity threats.

The ability of AI to mimic professional language and create contextually relevant content makes it a double-edged sword. In the wrong hands, these tools could be leveraged to produce fake emails containing phishing links, malware, or deceptive content aimed at stealing sensitive information. The risk is particularly high when attackers use scraped data from platforms like LinkedIn to personalize their messages, as demonstrated in this study's data collection methodology. Personalized phishing emails are significantly more effective because they build trust by referencing specific details about the recipient, such as their job title, company, or recent activities.

For example, an attacker could use the same prompt construction method outlined in this project to craft a legitimate email, such as an invitation to a job fair, which includes a link to a fraudulent website designed to harvest credentials or install malicious

software. With AI, these emails can be generated in bulk, each tailored to the target's specific profile, exponentially increasing the scale and effectiveness of such attacks.

The evaluation component of this project highlights the importance of legitimacy testing, but even attackers can exploit this feature to refine their phishing emails. By iterating through different prompts and configurations, malicious actors could use AI tools to identify and eliminate red flags in their emails, making them harder to detect by automated systems or human recipients.

From a cybersecurity perspective, this underscores the urgent need for advanced defenses against AI-enabled threats. Organizations must invest in email filtering systems that can detect subtle indicators of phishing, even in highly polished messages. Additionally, educating individuals about the risks of AI-generated phishing emails is crucial. Awareness campaigns should emphasize the importance of verifying email sources, scrutinizing links, and avoiding the sharing of sensitive information through email unless necessary.

As AI technology continues to evolve, too, must measure to mitigate its misuse. One solution is the development of AI-driven detection tools that leverage the same language models to identify patterns indicative of phishing attempts. These tools could

analyze emails for unusual phrasing, inconsistencies in metadata, or other subtle anomalies that humans might overlook.

In conclusion, while this project demonstrates the remarkable capabilities of AI in generating legitimate and professional emails, it also serves as a reminder of the potential dangers. The same technology that streamlines communication can, if misused, become a powerful tool for cybercriminals. As AI continues to shape the future of communication, it is imperative to address its ethical implications and ensure that its deployment includes safeguards against misuse in the cybersecurity domain.

REFERENCES

[And23] Anderson, Julia, and Lee, Thomas. "Real-Time Data Analytics in Web Scraping: Enhancements and Opportunities." Journal of Data Science, vol. 31, no. 1, 2023, pp. 45-64.

[Ben21] Bennett, Charles, and Khanna, Priya. "Automated vs. Manual Email Personalization: A Comparative Study." Journal of Marketing Technology, vol. 34, no. 2, 2021, pp. 198-215.

[Ell24] Ellison, Michael, and Moriarty, Sarah. "Cloud Security Measures and Data Privacy: Protocols and Regulations in Cloud Environments." Journal of Cloud Computing and Security, vol. 27, no. 3, 2024, pp. 112-131.

[Gar19] Garcia, Roberto, Lopez, Fernando. "Analyzing Security Vulnerabilities in Email Transmission Protocols." Cybersecurity Quarterly, vol. 23, no. 4, 2019, pp. 202-219.

[Gup21] Gupta, Priya, and Chandra, Alok. "Cloud-Based Solutions for Data Management: The Role of MongoDB Atlas." Journal of Cloud Computing, vol. 24, no. 1, 2021, pp. 56-74.

[Kim21] Kim, Soo-Yeon, and Park, Jae-Hyun. "Leveraging Machine Learning and Natural Language Generation in Content Adaptation." AI Review, vol. 47, no. 3, 2021, pp. 345-364.

[Lee17] Lee, Minho, and Hwang, Jae. "Evaluating SMTP and API-based Email Services for Effective Email Campaigns." Technology and Communications, vol. 22, no. 3, 2017, pp. 300-318.

[Mor22] Moreno, Luis, and Castillo, Maria. "Optimizing Data Extraction Algorithms for AI: Techniques and Applications." Journal of Computer Science and Technology, vol. 28, no. 1, 2022, pp. 112-130.

[Nov23] Nova, Samuel, and Briggs, Caroline. "Secure Email Automation: Blockchain's Role in Digital Communications." Journal of Digital Security, vol. 25, no. 2, 2023, pp. 89-107.

[OCo22] O’Connor, Megan, and Schmidt, Thomas. "Advancements in Distributed Databases for Real-Time Data Handling." Journal of Network and Computer Applications, vol. 50, no. 1, 2022, pp. 98-114.

[Ope24] OpenAI. "Discussion on Web Scraping and AI-Driven Email Automation Technologies Related Work and Sources." *ChatGPT*, OpenAI, 18 April 2024.

[Pat20] Patel, Raj, and Singh, Manoj. "Utilizing AI for Enhanced Content Creation: A Review of Natural Language Processing Algorithms." AI and Society, vol. 35, no. 4, 2020, pp. 789-805.

[Smi18] Smith, John, and Johnson, Alex. "Impact of Personalized Emails on Professional Networking Success." Journal of Business Communication, vol. 55, no. 2, 2018, pp. 204-222.

[Tho20] Thompson, Derek, et al. "Psychological Impacts of Personalized Content: Understanding Email Engagement." Journal of Behavioral Studies, vol. 44, no. 2, 2020, pp. 154-169.

[Wal24] Walters, Emily, and Gomez, Richard. "Utilizing Deep Learning to Predict User Behavior in Phishing Simulations." Advanced Computing Review, vol. 29, no. 3, 2024, pp. 330-350.

[Zha19] Zhang, et al. "Comprehensive Overview of Web Scraping Techniques: Balancing Data Accessibility and Privacy." Journal of Web Engineering, vol. 18, no. 6, 2019, pp. 123-145.