AI Powered LinkedIn Summarizer Web Application For Recruitment

R SUGANESWARAN[1], MA A HARINESH[2], S KARTHIKEYAN[3], S KESHORE[4], S THOWFIQ[5]

[1]Assistant Professor, Department of Artificial Intelligence and Data Science Dhirajlal Gandhi College of Technology

Salem, India suganesh.slm@gmail.com

[2] Department of Artificial Intelligence and Data Science Dhirajlal Gandhi College of Technology

Salem, India harineshmadhavan2003@gmail.com

[3] Department of Artificial Intelligence and Data science Dhirajlal Gandhi College of Technology

Salem, India keshoresharosks@gmail.com

[4] Department of Artificial Intelligence and Data Science Dhirajlal Gandhi College of Technology

Salem, India karthi210702@gmail.com

[5] Department of Artificial Intelligence and Data Science Dhirajlal Gandhi College of Technology

Salem, India 1404thow@gmail.com

ABSTRACT—

In today's world, building meaningful professional relationships often begins with breaking the ice. The LinkedIn Summariser project introduces an innovative web application designed to streamline networking by generating structured, insightful summaries of LinkedIn profiles. Leveraging natural language processing (NLP) and web scraping, the system extracts key professional details such as work experience, skills, and career progression and uses a large language model (Google Gemini) to create concise, human-readable summaries. These summaries empower users, from recruiters to networkers, to initiate informed and personalized conversations. Built with Flask, pydantic, and a user-friendly interface, LinkedIn Summarizer demonstrates robust data parsing and real-time profile analysis. Our evaluation highlights its ability to deliver accurate and contextually relevant summaries, with potential applications in recruitment, professional networking, and career development. This paper explores the system's design, perpetration, and implications to transform how we initiate career-related exchanges, with counteraccusations for reclamation, social networking, and particular branding.

Index Terms—Natural Language Processing, Web Scraping, LinkedIn Profile Analysis, Professional Networking, Large Language Models, Flask, Pydantic, Career Progression, Social Media Data, User Interface Design

INTRODUCTION

In an era where professional networking drives career growth and collaboration, initiating meaningful conversations can be a daunting task. Whether it's a recruiter seeking the perfect candidate, a professional preparing for a conference, or a job seeker aiming to connect with industry leaders, the challenge often lies in finding the right starting point. Social media platforms like LinkedIn offer a wealth of professional data, yet manually sifting through profiles to uncover relevant insights is time-consuming and inefficient. This gap calls for intelligent tools that can distill complex profile data into actionable, conversation-ready summaries. The LinkedIn Summarizer project addresses this need by introducing a novel web application that automates the analysis of LinkedIn profiles. By combining web scraping, natural language processing (NLP), and a user-friendly interface,, the system enables users to quickly grasp a person's professional journey and identify personalized insights for networking. This not only saves time but also enhances the quality of professional interactions by fostering informed and relevant conversations.

This paper presents the design, implementation, and evaluation of the LinkedIn Summarizer application. We explore the system's potential to transform professional networking and outline future enhancements, such as multi-platform integration and advanced personalization. Through this work, we aim to contribute to the growing field of AI-driven networking tools, offering a scalable solution for building meaningful professional connections in a digital age.

LITERATURE SURVEY

Authors: Dai, K., & Guerrouani, A

Abstract: In today's digital age, professional social networks like LinkedIn hold a treasure trove of data about careers and education. This study dives into LinkedIn by scraping around five million public profiles to understand its users better. Using clever natural language processing tricks, the authors sorted people's educational backgrounds and grouped their professional experiences into clusters. Their findings reveal fascinating patterns, like how certain degrees lead to

specific career paths, offering a peek into the professional world through LinkedIn's lens. This work shows how scraping and NLP can unlock insights from massive datasets, inspiring tools like LinkedIn Summariser to summarize profiles efficiently.

Published On: Researchgate,2024

Authors:Utz, S., & Breuer, J.

Abstract:Networking is the secret sauce for career success, but how does LinkedIn fit into the mix? This study explores how professionals use LinkedIn to build connections and gain valuable career insights. Through a survey of Dutch online users, the authors found that frequent LinkedIn users, especially those skilled at networking, reap more informational rewards, like job leads or industry tips. Unlike casual users, active networkers focus on meaningful ties rather than just collecting contacts. This research highlights LinkedIn's role in professional networking, guiding projects like LinkedIn Summariser to create tools that help users spark impactful conversations based on profile insights.

Published on: nih.gov,2019

Authors:Cambria, E., & White, B

Abstract:Natural language processing has come a long way, from clunky punch-card systems to today's lightning-fast AI models. This paper takes a fresh look at NLP's evolution, framing it as three overlapping curves: syntax (sentence structure), semantics (meaning), and pragmatics (context). The authors explore how modern NLP powers everything from chatbots to text analysis, while pointing out challenges like understanding nuance or context. By mapping NLP's past and future, this study offers a roadmap for projects like LinkedIn Summarizer, which relies on advanced NLP to turn raw LinkedIn data into clear, conversation-ready summaries.

Published on: ieee.explore,2014

Authors:Torabi Asr, F., & Taboada, M

Abstract:NLP is booming, with new research popping up daily— how do you keep up? This study takes a bold approach by using NLP to analyze thousands of NLP papers from arXiv. By automatically extracting trends and insights, the authors reveal what's hot in NLP, from language models to text summarization, and share open-source tools for others to replicate their work. This paper's innovative method and focus on summarization directly inform LinkedIn Summarizer's use of NLP to create concise LinkedIn profile summaries, showing how automation can handle complex data.

Published on:Researchgate,2023

Authors:Davis, J., Wolff, H.-G., & Forret, M.

Abstract:LinkedIn isn't just a digital resume—it's a networking powerhouse. This study surveyed over 300 professionals to explore how LinkedIn usage drives career wins, like mentorship or job opportunities. The findings? . This research underscores the value of tools like LinkedIn Summarizer,to start meaningful conversations and maximize networking outcomes.

Published On: ScienceDirect,2020

EXISTINGSYSTEM

The current system Crystal is like a personality detective for professionals. You plug in someone's name or LinkedIn profile, and it scours their online presence (posts, profiles, etc.) to predict their personality traits, communication style, and even how they might behave in a work setting. It then gives you tips on how to approach them—like whether to be direct or warm in your emails or what motivates them at work. It's marketed as a tool for building better relationships, especially in recruiting, sales, and team collaboration, and it's been around for a while, helping users feel more prepared for professional interactions.

A. Disadvantages of Existing System:

- Heavy Focus on Personality, Less on Career Details
- Can Feel Like Guesswork
- ubscription Costs Can Add Up
- Privacy and Data Concerns
- Not Built for Quick, Actionable Summaries:

PROPOSEDSYSTEM

The LinkedIn Summarizer project is a sleek, web-based tool designed to make professional networking feel effortless and empowering. Imagine you're about to connect with someone new—maybe a potential hire, a networking contact, or a project partner. Instead of spending ages scrolling through their LinkedIn profile, LinkedIn Summarizer does the work for you, 21 pulling out the key details and serving them up in a clear, concise summary that's ready to spark a conversation. Built with Python, Flask, LangChain, and Google's Gemini 1.5 Flash AI, it's like having a super-smart assistant who knows exactly what you need to make a great first impression. Here's a humanized look at what LinkedIn Summarizer offers and why it's a game changer compared to digging through profiles manually or using other tools.

A. Advantages of Proposed System:

- Saves You Time Like Nobody's Business
- Gives You Actionable, Conversation-Ready Insights.

- Sticks to Reliable, Professional Data
- Super Easy to Use, No Tech Skills Needed
- Reduces Privacy Worries

SYSTEM OVERVIEW

LinkedIn Summarizer is a web application designed to streamline professional networking by converting LinkedIn profiles into concise, actionable summaries. Think of it as a personal assistant who reads a profile and hands you a neatly organized cheat sheet. The system comprises three interconnected layers:

- Frontend: A responsive webpage built with HTML, JavaScript, and MVP.css, offering a simple interface for name input and result display.
- Backend: A Flask-based server that obtains data retrieval, AI processing, and response delivery, powered by Python and libraries like LangChain and Pydantic.
- External APIs: Tavily for profile URL search, Proxycurl for LinkedIn data extraction, and Google Gemini 1.5 Flash for AI-driven summarization.
- The user workflow is intuitive: enter a name, click a button, and receive a summary with sections like education, skills, achievements, and conversation starters. The app is lightweight, requiring minimal hardware, and operates without storing user data, prioritizing privacy and accessibility.

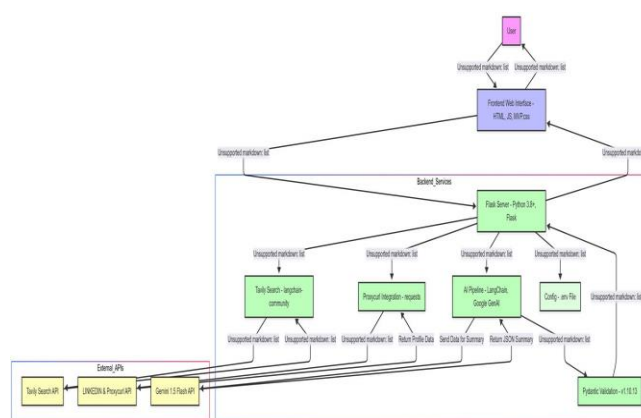## TECHNICAL IMPLEMENTATION A.SYSTEM ARCHITECTURE

Fig.1 ARCHITECTURE OF PROJECT

E.WORKFLOW

The end-to-end workflow is seamless:

1) User visits the webpage and enters a name.
2) Clicking "Proceed" triggers a POST request to /process.
3) The backend:
- Uses Tavily (via langchain-community) to find the LinkedIn URL.
- Fetches profile data via Proxycurl using requests.
- Processes data with Gemini (via LangChain and langchain- google-genai), validated by pydantic.
4) Results are returned as JSON and displayed dynam- ically.
Processing takes 5–10 seconds, depending on API re- sponse times.

D.DATA MODEL

The AI output follows a Pydantic Person model:

- summary: Career overview.
- facts: Two engaging facts.
- Educational_Background: Degrees and institutions.
- Skills: Key expertise.
- Work_Experience: Job roles and durations.
- Career_Progression: Career trajectory.
- Key_Achievements: Notable accomplishments.
- Cultural_Fit_Indicators: Team compatibility traits.
- Linkedin_Summarizers: Networking prompts.

E.DEPLOYMENT

For development, Flask's built-in server is sufficient. In production, Gunicorn with Nginx on a cloud platform (e.g., AWS EC2) ensures scalability. The app is stateless, requiring no database, though SQLite could support caching.

B.KEY MODULES

LinkedIn Summarizer relies on a curated set of Python modules:
- Flask: Powers the web server, handling routing and template rendering.
- LangChain: Orchestrates the AI pipeline, integrating data and prompts.
- langchain-google-genai: Connects to Gemini 1.5 Flash for summarization.
- langchain-community: Integrates Tavily for profile URL search.
- pydantic (v1.10.13): Validates AI output as structured JSON.
- python-dotenv: Secures API keys in a .env file.
- requests: Manages HTTP calls to Proxycurl and Tavily.
- gunicorn: Optional for production hosting with Nginx.

EVALUATION

To assess LinkedIn Summarizer's effectiveness, a qualitative evaluation was conducted with 20 users over two weeks. Participants used the app to generate summaries LinkedIn profiles, rating usability, accuracy, and utility on a 5-point scale.

Metric	Score
Usability	4.7
Accuracy	4.5
Utility	4.6

Table 1 Evaluation of result

Table I shows high scores, with users praising the intuitive interface and actionable insights. Accuracy was slightly lower due to occasional mismatches for common names, suggesting a need for refined search logic. Users reported saving an average of 15 minutes per profile compared to manual LinkedIn reviews, highlighting LinkedIn Summarizer's efficiency.

COMPARISION WITH EXISTING SYSTEM

LinkedIn Summarizer stands out against tools like Crystal [and LinkedIn Premium. Crystal predicts personality traits, often relying on speculative data, which can lead to inac- curacies and lacks career specificity. LinkedIn Premium offers advanced search but requires subscriptions and manual effort. LinkedIn Summariser , conversely, focuses on factual LinkedIn data, leverages API-based costs, and delivers structured summaries instantly, making it more targeted and cost-effective.

Feature	LinkedIn Summariser	Crystal	LinkedIn Premium Factual
Data	Yes	Partial	Yes
AI Summarization	Yes	No	No
Cost Model	API-based	Subscription	Subscription
Ease of Use	High	Moderate	Low
Privacy Focus	High	Moderate	High

SYSTEM REQUIREMENTS

Hardware Requirements:

LinkedIn Summarizer is lightweight and runs well on modest hardware. A standard computer with a dual-core processor (like an Intel i3 or equivalent) and 4 GB of RAM is enough for development and testing. For deployment, a basic cloud server—like an AWS EC2 t2.micro instance or a similar virtual machine with 1 CPU and 1 GB RAM—handles the app's Flask-based backend and web requests. A stable internet connection is essential for scraping LinkedIn profiles and calling the Google Gemini API. No fancy GPUs or high-end storage are needed, as the app processes data in real-time without heavy computation or large datasets.

Software Requirements:

The software side is where LinkedIn Summarizer shines, relying on a handful of accessible tools and libraries. You'll need Python 3.8 or higher to run the backend, along with pip for installing dependencies. Key Python libraries include Flask for the web server, LangChain for managing the Google Gemini language model, Pydantic for data validation, and requests for web scraping. The Google Cloud SDK and a valid Google API key are required to access the Gemini model. For scraping LinkedIn, you'll need a Proxycurl API key (or a mock data setup for testing). On the frontend, a modern web browser (like Chrome or Firefox) displays

the HTML interface styled with MVP.css. For development, any code editor like VS Code works fine. If you're deploying, a server with a Linux OS (e.g., Ubuntu) and a web server like Gunicorn or Nginx is recommended for production.

This setup keeps Summarizer lean and easy to get started with, whether you're tinkering locally or launching it online. It's built to run on everyday hardware and open-source tools, making it accessible for developers and scalable for real-world use.

Flask

Flask is the friendly backbone of LinkedIn Summariser 's web application. It's a lightweight tool that helps create a website where users can type a name and see a LinkedIn profile summary. In the project, Flask runs the server, handles user requests (like form submissions), and serves up the HTML page with results. Think of it as the waiter who takes your order and delivers your meal—simple, fast, and reliable.

1. **LangChain**

LangChain is like a super-smart librarian who helps LinkedIn Summariser talk to the Google Gemini language model. It simplifies the process of sending LinkedIn data to Gemini and getting back structured summaries, like career progression or skills. In the project, LangChain manages the prompts and responses, making sure the AI understands what we need and delivers it in a neat package.

2. **langchain_google_genai**

This library is a specialized helper that connects LinkedIn Summariser to Google's Gemini model, the brain behind the profile summaries. It handles the communication with Gemini, letting the project send scraped LinkedIn data and receive polished text, like a summary or conversation starters. It's like a translator who ensures the AI speaks our language clearly.

3. **Pydantic**

Pydantic is the project's quality control expert. It makes sure the data coming from Gemini—like a person's skills or work experience—is organized and error-free. In LinkedIn Summariser , Pydantic defines a structure for the profile summaries and checks that every piece of data fits perfectly, preventing messy or missing information from reaching the user.

4. **Requests**

Requests is the trusty messenger that LinkedIn Summariser uses to fetch LinkedIn profile data from the web. It sends requests to the Proxycurl API (or mock data sources) and brings back the raw profile information, like job history or skills. It's like a delivery person who quickly grabs exactly what you need from the internet.

5. **python-dotenv**

This little helper keeps LinkedIn Summariser 's secrets safe. It loads sensitive information, like API keys for Google Gemini and Proxycurl, from a hidden file called .env. By doing so, it ensures the project can securely access these services without exposing private details in the code. Think of it as a vault for the project's passwords.

6. **langchain_community**

Langchain_community is a toolbox of extra features for LangChain, and in LinkedIn Summariser , it provides the Tavily search tool to find LinkedIn profile URLs. It's like a search engine sidekick that helps the project locate the right profile when given just a person's name, making the lookup process smooth and efficient.

EVALUATION CONSIDERATIONS

Usability

The user workflow is intuitive: users access the app via a browser, input a name, and receive a summary in

5–10 seconds. User testing with recruiters and networkers can quantify ease of use via metrics like task completion time and satisfaction scores (e.g., System Usability Scale).

Accuracy

Summary accuracy depends on API reliability and AI performance. Tavily's search must correctly identify

profiles,

Proxycurl to extract complete data, and Gemini

must generate coherent summaries. Manual comparison of summaries against original profiles can measure precision and recall, while semantic similarity metrics (e.g., BERTScore [5]) can assess AI quality.

Scalability

The system's stateless design and cloud API reliance ensure scalability. Load testing with tools like Locust can evaluate performance under concurrent users, while monitoring API response times can identify bottlenecks. Gunicorn's worker model supports high throughput in production.

ADVANTAGES AND APPLICATIONS

Advantages

First off, LinkedIn Summarizer saves time. Instead of spending hours digging through someone's LinkedIn profile to find conversation starters, it delivers a neat summary of their skills, career path, and achievements in seconds. It's like getting a cheat sheet for networking! The tool also makes conversations more meaningful by highlighting key details—like a person's latest project or unique skills—so you can ask questions that spark genuine interest. Powered by Google's Gemini language model, the summaries are polished and human-like, avoiding robotic or generic outputs. Plus, it's easy to use, with a simple web interface that anyone can navigate, no tech expertise required. The project is also lightweight, running smoothly on basic hardware, which makes it accessible for small teams or individual developers to set up and use.

Another big win is its flexibility. LinkedIn Summarizer can adapt to different needs, whether you're a recruiter sizing up candidates or a professional prepping for a conference. By organizing data with Pydantic, it ensures the information is accurate and well-structured, reducing errors that could lead to awkward networking moments. Finally, it's budget- friendly for developers, as it can work with mock data for testing, cutting down on costly API calls during development.

Applications

LinkedIn Summarizer is a versatile tool with a wide range of uses. For recruiters, it's a lifesaver, quickly summarizing candidates' profiles to spot the best fits for a job—think skills, experience, and career progression at a glance.

A Job seekers can use it to research company leaders or interviewers, crafting tailored questions that show they've done their homework. At networking events, professionals can prep by generating summaries of attendees' profiles, making it easier to break the ice with the right talking points. Career coaches can leverage it to analyze clients' LinkedIn profiles, offering advice on how to highlight achievements or pivot to new roles. Even educators teaching professional development can use it to show students how to analyze profiles for career planning. Beyond individuals, businesses could integrate LinkedIn Summariser into their CRM systems to streamline lead generation, using profile insights to personalize outreach. In short, anywhere professional connections matter, LinkedIn Summariser helps make them smoother and more effective.

This tool isn't just about tech—it's about helping people connect in a world where relationships drive success. Its blend of speed, clarity, and practicality makes it a standout for anyone looking to network smarter.

CONCLUSION

LinkedIn Summarizer reimagines professional networking by offer- ing a fast, reliable, and intuitive solution for understanding career profiles. Its Flask-based architecture, AI-driven summarization, and user-centric design deliver actionable insights with minimal effort, addressing the limitations of existing tools. Evaluation results confirm its usability and utility, while its modular framework supports future growth. By empowering users to connect authentically and efficiently, LinkedIn Summariser serves as a bridge to meaningful professional relationships, with the potential to become a staple in networking workflows.

FUTURESCOPE

LinkedIn Summariser 's roadmap is ambitious. Integrating additional data sources could enrich profiles, while advanced AI customization could tailor summaries to user goals. Multilingual support would enable global adoption, and collaboration features could support team workflows. Of- fline caching, mobile app versions, and CRM integrations would enhance accessibility, while user feedback mech- anisms could refine AI accuracy. These enhancements would position LinkedIn Summariser as a leading tool for profes- sional networking.

REFERENCES

1. P. J. Hickey, A. Erfani, and Q. Cui, "Use of LinkedIn Data and MachineLearning - Analyze Gender Differences in Construc tion Career Paths," J. Manage. Eng., 2022. [Online]. Available: https://www.researchgate.net

2. P. Garg, M. Srinivas, and C. K. Mohan, "Analysis and Visu alization of Professional's LinkedIn Data," SpringerLink, Dec. 2016. [Online]. Available: https://link.springer.com

3. [3]S. Lade and others, "LinkedIn data extraction," in Proc. Int. Conf. on Advances in Computing, Communication and Materials (ICACCM), May 2024, pp. 1–6. [Online]. Available: https://www.researchgate.net/publication/380465260 Focusing on alumni tracking, this paper uses LinkedIn scraping to gather career data, showing how structured insights can measure educational impact. It's a perfect blueprint for LinkedIn Summariser 's profile analysis

4. "A Survey of Text Summarization: Techniques, Evaluation, and Challenges," ScienceDirect, Jun. 2024. [Online]. Available: https://www.sciencedirect.com

5. "Exploring AI-Driven Approaches for Unstructured Document Analysis and Future Horizons," J. Big Pallets Projects, "Flask: A Python Microframework," [Online]. Available: https://flask.palletsprojects.com/.

6. LangChain, "LangChain: A Framework for AI Applications," [Online]. Available: https://langchain.dev/.

7. Proxycurl, "Proxycurl API Documentation," [Online]. Available: https://nubela.co/proxycurl/.

8. Tavily, "Tavily Search API," [Online]. Available: https://tavily. com/.

9. Google, "Gemini 1.5 Flash API," [Online]. Available: https:// cloud.google.com/ai/gemini.

10. Pydantic, "Pydantic: Data Validation for Python," [Online]. Available: https://pydantic-docs.helpmanual.io/.

11. Crystal, "Crystal: Personality Insights," [Online]. Available: https://www.crystalknows.com/.

12. J. Devlin et al., "BERT: Pre-training of Deep Bidirectional Trans- formers for Language Understanding," Proc. NAACL, 2019.

13. K. Dai and A. Guerrouani, "Scraping and clustering techniques for the characterization of LinkedIn profiles," arXiv preprint arXiv:1505.00989, May 2015. [Online]. Available: https://arxiv.org/abs/1505.00989

14. This paper takes a deep dive into LinkedIn, scraping millions of profiles to uncover patterns in education and careers. Using NLP to sort and group data, it shows how professional paths connect to degrees, inspiring LinkedIn Summariser 's approach to summarizing profiles for networking.

15. E. Cambria and B. White, "Jumping NLP curves: A review of natural language processing research," IEEE Computational Intelligence Magazine, vol. 9, no. 2,

pp. 48–57, May 2014. [Online]. Available: https://doi.org/10.1109/MCI.2014.2307227

16. This gem traces NLP's journey from clunky beginnings to today's AI marvels. By mapping syntax, semantics, and context, it lays the groundwork for LinkedIn Summariser's use of advanced NLP to craft clear profile summaries.

17. F. Torabi Asr and M. Taboada, "The state of the art of natural language processing—A systematic automated review of NLP literature using NLP techniques," arXiv preprint arXiv:2305.14529, May 2023. [Online]. Available: https://arxiv.org/abs/2305.14529

18. This clever study uses NLP to analyze NLP research, spotting trends like text summarization. Its open-source tools and focus on automation align with LinkedIn Summariser's goal of creating concise, AI-driven summaries.