International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 10 Issue: 01 | Jan - 2026 SJIF Rating: 8.586 ISSN: 2582-3930

# Smart Crawler Based Lead Generator for Freelancers

**Kunal Harad[1], Prabhat Jha[2], Wasim Patel[3], Dr. Arun Saxena***

[1]*Information Technology , St. John College of Engineering and Management*
[2]*Information Technology , St. John College of Engineering and Management*
[3]*Information Technology , St. John College of Engineering and Management*
[4] Associate *Professor, Department of Information Technology, St. John College of Engineering and Management*

---

## Abstract

This project addresses the major issues that independent contractors confront, especially the dearth of chances and leads. For independent contractors, this procedure is among the most laborious and ineffective. We propose a system to overcome these problems by using a combination of technologies such as artificial intelligence, agentic AI, and web crawling. The system is developed to automate the entire first phase of lead discovery and opportunity identification by leveraging autonomous web crawling and advanced reasoning. The system scans web pages to generate leads and identify potential opportunities.

It also uses maps to take advantage of geolocation data and social media platforms to find leads, which speeds up the process of searching and finding opportunities and increases the overall efficiency of freelancers. The core of the system uses AI agents to scan web pages, find leads, identify sources, and validate their authenticity. It evaluates each webpage based on predefined parameters such as inactive websites, poor branding, offline websites, and potential demand, while eliminating inactive websites. The system architecture ensures efficiency and authenticity and also considers the overall SEO of websites. This process allows freelancers to bypass inefficient search methods and connect more effectively while expanding their opportunities.

## Introduction

The growing pace at which the digital economy is growing has greatly augmented the chances of freelancers in different fields including web development, digital marketing, graphic designing, and content production. Nonetheless, client acquisition is one of the most challenging issues even though the number of freelance services demanded continues to increase.

A lot of freelancer's waste much of their time in search of prospective clients at the expense of their main areas of professional competence. This inequity makes the business less productive and restricts business growth. Conventional methods of lead generation including cold emailing, hand-searching of directories and using job portals are unproductive and unreliable. These techniques are very laborious and use human judgment and this may lead to time wastage and lost opportunities.

Moreover, freelancers tend to lack the analysis tools that would allow them to conclude whether a business is indeed in need of their services, which makes the process vagrant and disappointing. The development of Artificial intelligence, machine learning, and cloud computing has opened new opportunities in the process of automating repetitive and date-heavy tasks. Nowadays, smart web crawlers and AI-based analytics systems have the ability to work with vast amounts of online information and extract patterns and produce actionable insights with high precision.

Moreover, freelancers tend to lack the analysis tools that would allow them to conclude whether a business is indeed in need of their services, which makes the process vagrant and disappointing. The development of Artificial intelligence, machine learning, and cloud computing has opened new opportunities in the process of automating repetitive and date-heavy tasks. Nowadays, smart web crawlers and AI-based analytics systems have the ability to work with vast amounts of online information and extract patterns and produce actionable insights with high precision.

The capabilities allow it to be possible to create systems that can independently find and analyze possible business leads. The Smart Crawler-Based Lead Generator project will aim at sealing the gap between freelancers and businesses that require a digital transformation.

The system actively searches the digital space to identify businesses that lack online presence or have outdated digital infrastructure instead of relying on passive job listing or competitive marketplaces. This proactive strategy would make sure that the freelancers get relevant and timely lead.

The proposed solution, which is a combination of automation, data analytics, and real-time communication technologies, will provide a holistic solution in terms of intelligent lead generation and outreach. Not only does it enhance efficiency and accuracy, but also enhances sustainability in business practices by matching competent professionals with businesses that actually need their services. The net effect of this change is that it helps both the freelancers and small businesses in an ever-digitized market.

## Methodology

The initial phase of the approach is aimed at automated lead discovery by means of web crawling and data extraction via API. The system will make use of Google Maps API and other publicly available online directories to collect data on local and international business. It tracks down businesses that are missing websites, have low ratings, do not have social media accounts or do not have complete business profiles. This is automated in terms of data discovery to provide continuous and scalable inflow of raw business data without human intervention.

After collecting data, the AI analysis module analyzes the data of each business through machine learning and Natural Language Processing. Clustering algorithms can be used to group similar business as per industry and the level of digital maturity, sentiment analysis can be used to evaluate the perception of the people in terms of reviews and comments. The stage aids in identifying the potential that a business may be in need of services like SEO, redesigning websites, or online marketing that helps in transforming raw data into qualified leads.

The system then automatically creates customized outreach content based on text generation models powered by AI after qualification. These studies bring out certain shortcomings in the online presence of a business and present personalized solution. Automating this step will guarantee consistency, professionalism, and speediness at the same time keeping a personalized tone that is likely to boost engagement.

The proposals that are generated are transmitted via several communication methods such as email (through SMTP) and voice synthesis platforms through voice synthesis programs that run on AI. This channeling method enhances visibility and response rates. The WebSocket's-powered real-time dashboard enables freelancers to track outreach campaigns, see the statistics of engagement, and immediately get notified when the lead responds to the calls or follow-ups.

The full data is stored safely in Google BigQuery where all of the gathered and processed data can be easily queried and analyzed. The backend is coded in FastAPI and is put in a containerized version of Docker to ensure a consistent and portable system. Servers on Google cloud run setup offer serverless and auto-scaling platform that delivers reliability and performance even in the face of changing load. This infrastructure ensures scalability, security, and cost-effectiveness of the system.

## Results and Discussion

Use of Smart Crawler Based Lead Generator presented encouraging and quantifiable outcomes in automating the lead discovery and qualification process. It was able to utilize Google Maps and online directories as online sources through which the system was able to extract business information with high precision and minimum human interference. It showed the team the capability to generate new leads on an ongoing basis in accordance with the set filters such as missing websites, low digital activities, and unfinished profiles. Such automation greatly saved time that used to be dedicated to manual search and verification by the freelancers.

The artificial intelligence-based analysis unit was useful in the process of assessing and classifying the leads based on their potential worth. The clustering and sentiment analysis helped the system to differentiate between high-priority and low-priority prospects so that freelancers could strategically concentrate on them. Individualized generation of proposals and automated outreach via email and voice communication led to better response rates as opposed to generic outreach techniques. The real time dashboard also proved to be reliable as it offered immediate updates and analytics which assisted users monitor engagement and campaign performance effectively.

The system was found to be scalable and stable when it was used on a cloud-based infrastructure and during testing and pilot usage. The combination of FastAPI, Docker, and Google Cloud Run provided the expected seamless functioning even in the context of working with large data. With the help of BigQuery, it was possible to store data and retrieve it, which made the work with the histories of leads and campaign analytics effective. These technical results validated that the architecture can be used in real-life application scenarios without affecting its performance.

According to the observed findings, the project may be regarded as successful in reaching its key goal of automating the process of generating leads and reaching out to freelancers. As the discussion shows, intelligent automation, in addition to making the tasks more efficient, makes decision-making more agile by delivering data-driven insights. Although additional upgrades, including more personalized functionality, more diverse data, and enhanced ethical protection measures can be used to better the system, the existing implementation proves that it is practically feasible and has a greater likelihood of commercial implementation. The conclusion is that it is a useful tool that can be applied by the contemporary freelance workers because this system is technically feasible and operationally advantageous.

Results:

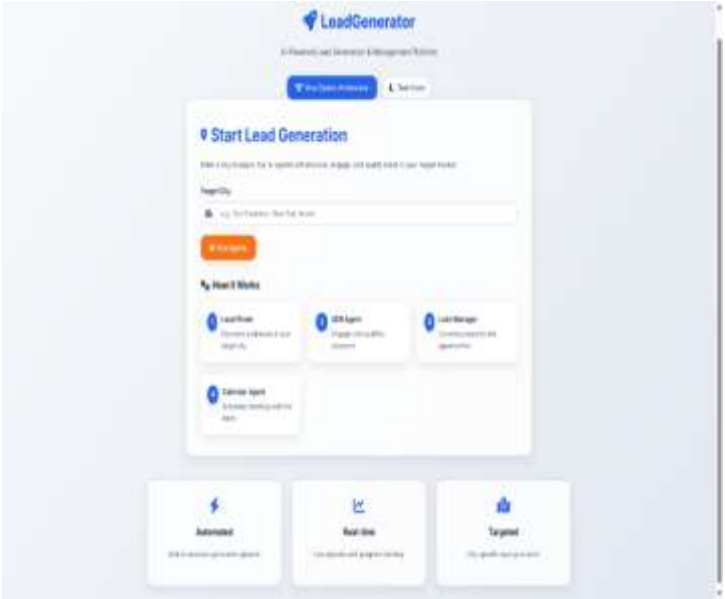

**Fig. 1** This screen shows the main dashboard of the AI-Powered Lead Generator platform where users can start the lead discovery process. The user enters a target city and clicks Run Agents to automatically find and qualify business leads. The interface also explains the workflow through different AI agents such as Lead Finder, SDR Agent, and Lead Manager. At the bottom, feature cards highlight key benefits like automation, real-time updates, and targeted lead generation, making the system simple and efficient to use.

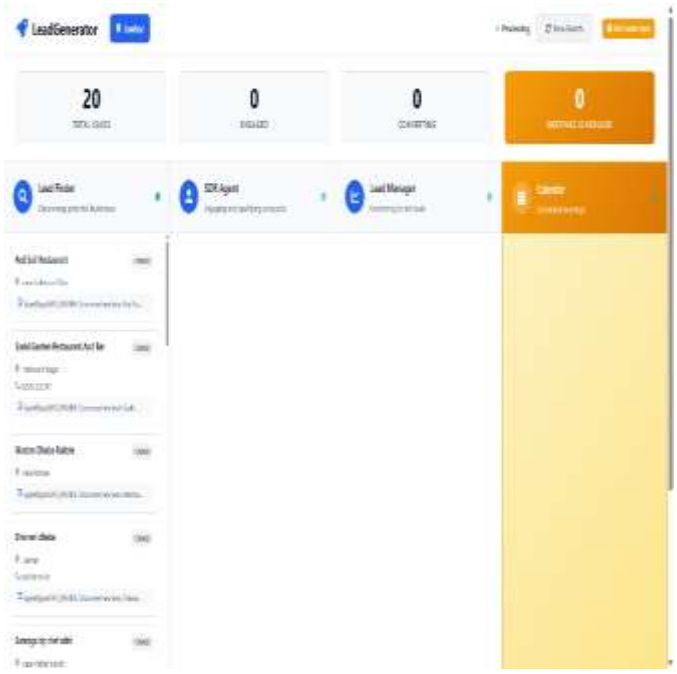


**Fig. 2** shows the Lead Management Dashboard of the Smart Lead Generator system. It provides a real-time overview of total leads, engaged prospects, conversions, and scheduled meetings at the top. The dashboard is divided into modules such as Lead Finder, SDR Agent, Lead Manager, and Calendar, each handling a specific stage of the lead pipeline. On the left, discovered businesses are listed, while the right panel displays scheduled meetings and progress status. This interface helps freelancers efficiently track, manage, and convert leads in one centralized view.

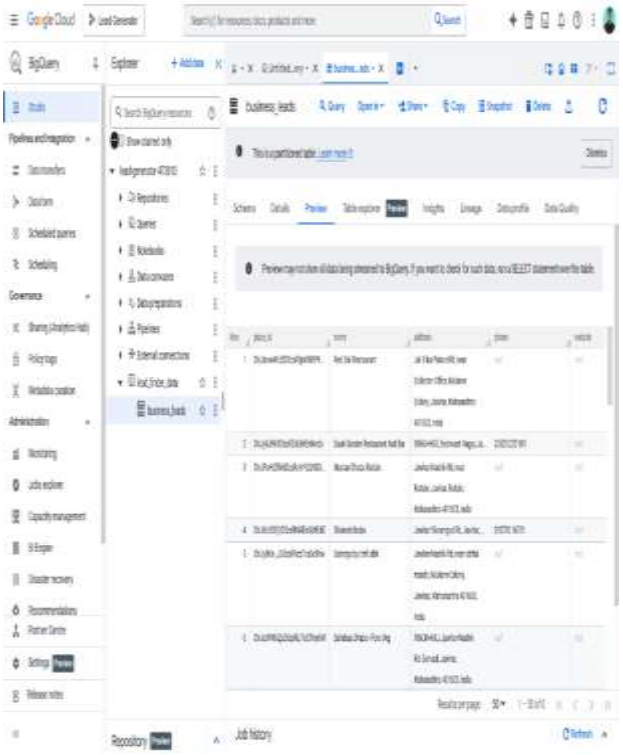


**Fig. 3** This figure shows the Google BigQuery Database used in the Smart Lead Generator system. It displays the stored business lead data in a structured table format including fields such as name, address, phone number, and website. The left panel shows the project dataset and tables, while the main panel previews the collected lead records. This database enables secure storage, fast querying, and efficient analysis of large volumes of lead information for real-time system performance.

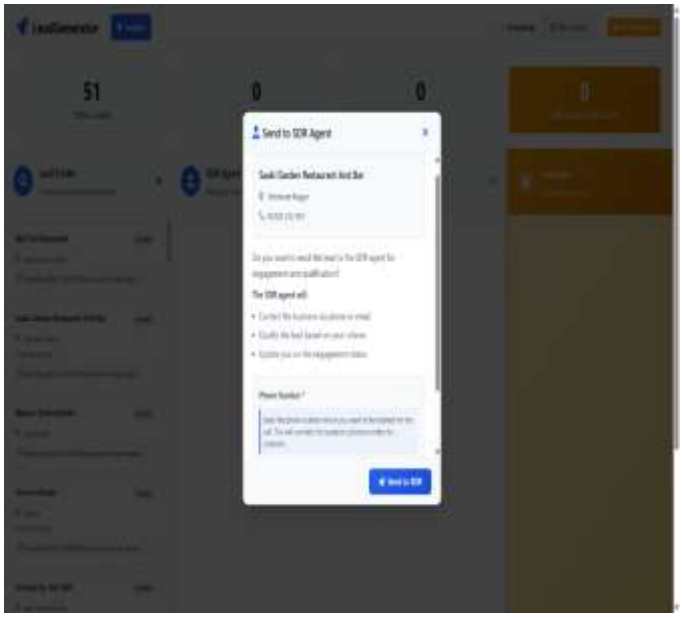


**Fig. 4** shows the Actual Result of the Smart Lead Generator system during live operation. The dashboard displays the total number of leads discovered along with engagement and meeting statistics. A pop-up window titled “Send to SDR Agent” is visible, demonstrating how a selected business lead is forwarded

for qualification and communication. This result confirms that the system successfully discovers businesses, allows user interaction, and initiates automated outreach processes, proving the practical functionality of the platform.

## Conclusion

The Smart Crawler-Based Lead Generator for Freelancers is an effective way of illustrating how artificial intelligence and automation can bring sanity to the process of acquiring clients, traditionally a manual one. With the help of web crawling, data analytics, and smart lead qualification, the system can find businesses with a weak online presence and transform untamed online data into valuable and actionable opportunities. This saves time and effort that the freelancers would have used in prospecting and they are better positioned to concentrate on providing quality services.

This is augmented with automated proposal creation, multi-channel outreach and a management dashboard that will be effective in real-time. Fast API, Docker, Google Cloud Run, and Big Query technologies used make the system technically sound and capable of working in the real world as it offers scalability, reliability, and secure data management. The outcomes suggest that there is enhanced efficiency, increased targeting of the lead, and an increased potential of engagement as opposed to the traditional methods.

On the whole, the project reminds about the increased significance of AI-based solutions in contemporary digital business settings. It is an efficient and affordable tool that helps freelancers and also helps the small business to reinforce their online presence. As the system may further develop with improvements like greater personalization and a broader range of sources of data, there is a high potential of the system becoming a full-fledged intelligent helper in the autonomous development of the freelance business.

## References

[1]. Zhao, L., Seton, O., Reddivari, H. R., Jena, S., Zhao, S., Kumar, R., & Wei, C. (2025). Causal predictive optimization and generation for business AI. arXiv.

[2]. T. Dhar, S. Mazumder, S. Dhar, S. Karak, and D. Chatterjee, "An Approach to Design and Implement Parallel Web Crawler," in Proc. 2021 Int. Conf. Smart Generation Computing, Communication and Networking (SMART GENCON), 2021, pp. 1–4, doi: 10.1109/SMARTGENCON51891.2021.9645918.

[3]. W. Zhu, J. Qin, R. Kong, H. Lin, and Z. He, "A System Framework for Efficiently Recognizing Web Crawlers," in Proc. 2018 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI), Oct. 2018, pp. 1130–1133, doi: 10.1109/SmartWorld.2018.00196.

[4]. Rane, N., Choudhary, S., & Rane, J. (2024, July). Artificial intelligence and machine learning in business-to-business (B2B) sales and marketing: A review. International Journal of Data Science and Big Data Analytics, 4(1), 17–33.

[5]. Ye, T., Zheng, J., Jin, J., Qiu, J., Ai, W., & Mei, Q. (2024). Using artificial intelligence to unlock crowdfunding success for small businesses. arXiv.

[6]. Schwartz, S., Yaeli, A., & Shlomov, S. (2023). Enhancing trust in LLM-based AI automation agents: New considerations and future challenges. arXiv.

[7]. Dumas, M., Fournier, F., Limonad, L., Marrella, A., Montali, M., Rehse, J.-R., Accorsi, R., Calvanese, D., De Giacomo, G., Fahland, D., Gal, A., La Rosa, M., Völzer, H., & Weber, I. (2023). AI-augmented business process management systems: A research manifesto. ACM Transactions on Management Information Systems, 14(1), 1–19.

[8]. Kaur, S., Singh, A., Geetha, G. et al. IHWC: intelligent hidden web crawler for harvesting data in urban domains. Complex Intell. Syst. **9**, 3635–3653 (2023).]

[9]. Le Dinh, T., Vu, M.-C., & Tran, G. T. C. (2025). Artificial Intelligence in SMEs: Enhancing Business Functions Through Technologies and Applications. Information, 16(5), 415.

[10]. Enholm, I.M., Papagiannidis, E., Mikalef, P. et al. Artificial Intelligence and Business Value: a Literature Review. Inf Syst Front **24**, 1709–1734 (2022).

[11]. Wikipedia contributors. (2025, July 26). Artificial intelligence marketing. In Wikipedia, The Free Encyclopedia.Retrieved05:34,July27,2025,from

[12]. Wei, R., & Pardo, C. (2022, November). Artificial intelligence and SMEs: How can B2B SMEs leverage AI platforms to integrate AI technologies? Industrial Marketing Management, 107, 466–483.

[13]. Sharma, M., & Tribhuvan, P. P. (2021, December). Smart Crawler. International Journal of Computer Science and Information Technology,

[14]. Khare, A., Dalvi, A., & Kazi, F. (2020). Smart crawler for harvesting deep web with multi-classification. In 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT) (pp. 1–5). IEEE.

[15]. Nemeslaki, A., & Pocsarovszky, K. (2011, September 18–21). Web crawler research methodology. Paper presented at the 22nd European Regional Conference of the International Telecommunications Society (ITS2011): Innovative ICT Applications – Emerging Regulatory, Economic and Policy Issues, Budapest, Hungary

[16]. Elhajjar, S., Yacoub, L., & Yaacoub, H. (2023). Automation in business research: Systematic literature review. Information Systems and e-Business Management, 21, 1–24.

[17]. Sabir Hossain, M., Arefin, M.S. (2020). An Intelligent System to Generate Possible Job List for Freelancers. In: Sharma, H., Govindan, K., Poonia, R., Kumar, S., El-Medany, W. (eds) Advances in Computing and Intelligent Systems. Algorithms for Intelligent Systems. Springer, Singapore.

[18]. Sahnoun, I., & Elhadjamor, E. A. (2024). Enhanced freelance matching: Integrated data analysis and machine learning techniques. Journal of Computing Theories and Applications, 1(4), 507–517.