# Developing Big Data Talent for Vietnam's Digital Economy: Assessing Gaps and Training Needs

**Adrian Popescu**
Department of Computer Science and Engineering, University of Timișoara, Romania
adrian.popescu@ut.ro
**Elena Dumitrescu**
Transilvania University of Brașov, Romania

## Abstract

This paper examines the talent gaps and training needs for developing big data capabilities in Vietnam to support the growth of its digital economy. As Vietnam moves towards greater digitalization and data-driven decision making across sectors, there is an urgent need to build specialized big data skills domestically. We conduct an empirical assessment of the current state of big data talent in Vietnam across different industries. Our findings indicate acute shortages of data scientists, data engineers, and other supporting roles with only 17% of demand being met from local supply. Core skill gaps exist in areas like machine learning, AI, and advanced analytics. We then analyze the training capacities and programs currently available and find them inadequate for bridging existing gaps. Based on insights from industry experts and academicians, we propose a framework for systematic capacity building of big data talent in Vietnam. This includes introducing dedicated educational courses, incentivizing private training providers, building data analytics centers of excellence, and launching nationwide skill enhancement initiatives aligned to industrial needs. Successful implementation of the suggested recommendations can significantly boost Vietnam's big data capabilities and contribute to realizing its vision of becoming a leading digital economy in Southeast Asia.

## Introduction

Vietnam's rapid economic growth and widespread digital adoption in the last decade position it as a significant player in the Southeast Asian landscape. Projections indicate that its internet economy will surge to US$57 billion by 2025, making it the second-largest in the region. This growth is attributed to several factors, including increasing incomes, urbanization, widespread smartphone penetration, and a notable surge in digital consumer activity. The nation's commitment to digital transformation is evident in its national strategies, which articulate clear targets for various sectors and business environments within the digital economy [1]. The integration of advanced digital technologies and the widespread application of big data have become integral to the operations, decision-making processes, and service delivery of leading Vietnamese corporations. This technological infusion not only enhances efficiency but also positions Vietnam at the forefront of digital innovation in the region [2]. However, the expansion of data-driven governance programs by public agencies reveals a critical challenge: a shortage of adequately skilled local talent. The deficit in specialized human capital poses a significant constraint on the successful deployment of big data initiatives

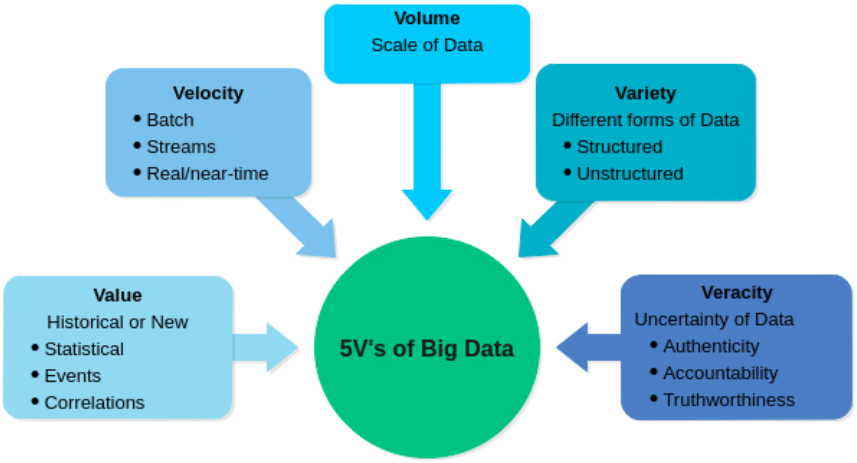
and hinders the realization of Vietnam's aspirations for an innovation-led, digitally powered economy [3].

The nexus between technology and talent is particularly pronounced in the context of data-driven governance. As more public agencies embrace data-centric approaches to enhance governance, the shortage of skilled professionals becomes a bottleneck. Without a robust pool of local talent proficient in big data analytics, these initiatives risk falling short of their potential impact [4]. The overarching goal of steering Vietnam toward an innovation-led economy necessitates a comprehensive strategy to address this talent gap [5]. The implications of this talent shortage extend beyond the immediate challenges faced by public agencies. They directly impact the broader landscape of innovation, economic competitiveness, and sustainable development. Failure to cultivate specialized human capital domestically not only undermines the effectiveness of current digital initiatives but also jeopardizes the long-term vision of positioning Vietnam as a digitally advanced economy. Addressing this challenge requires a multi-faceted approach involving collaboration between government bodies, educational institutions, and the private sector. Initiatives to enhance education and training programs in fields related to big data analytics, artificial intelligence, and advanced digital technologies are imperative [6], [7].

This paper examines the gaps in availability of big data talent in Vietnam vis-à-vis rising demand across different industries and assessing current training capacities available in the country. Our empirical analysis indicates acute shortages with current supply falling severely short especially for high skilled roles like data scientists and engineers. Core competency gaps also exist related to areas like machine learning, artificial intelligence and leveraging cloud data platforms. While universities are expanding dedicated analytics and data science programs, most are struggling to meet rapidly evolving industry requirements [8]. The coverage and quality of private training providers remains uneven.

We propose a practical framework for developing big data talent at scale in Vietnam rooted in insights from industry leaders, academic experts and comparative global practices. A systematic approach is required with coordinated interventions across educational institutions, dedicated analytics centers, private sector participation and nationwide upskilling programs for working professionals. Targeted recommendations aimed at different stakeholder groups can significantly improve the availability of job-ready local talent over 3-5 years. This can relieve current overdependence on foreign personnel and external partnerships that increase costs and limit big data deployments for Vietnamese entities across sectors [9].

Figure 2: Key Constraints in Developing Local Big Data Talent




## Assessing the Demand Supply Gap

Vietnam's robust economic growth in the internet sector, exceeding 31% annually for the next five years, positions it as a regional leader. The infusion of advanced digital applications and data-centric decision-making processes has permeated various vital sectors of the Vietnamese economy, ranging from financial services and retail to agriculture, healthcare, and smart city management. This pervasive integration of technology signifies a comprehensive digital transformation that is reshaping traditional industries and fostering innovation across the nation [10]. The rise of Vietnam's internet economy is underscored by a growing demand for skilled data professionals. The trajectory of this expansion indicates a compelling need for approximately 200,000 data professionals by 2026, spanning diverse roles such as data analysts, scientists, engineers, and technicians. The urgency to meet this demand is accentuated by the critical role data plays in driving efficiency, innovation, and competitiveness across industries. The dynamic landscape of the digital economy demands a workforce well-versed in data analytics, interpretation, and application [11].

Financial services stand out as one of the primary beneficiaries of this digital evolution. The increasing reliance on digital platforms for financial transactions, coupled with the burgeoning fintech sector, underscores the imperative for skilled data analysts and scientists to navigate the vast volumes of financial data [12]. These professionals play a pivotal role in uncovering insights, mitigating risks, and enhancing decision-making processes in the financial domain. Similarly, the retail sector is undergoing a profound transformation driven by e-commerce, data-driven marketing, and personalized customer experiences. The demand for data professionals in this sector is fueled by the necessity to analyze consumer behavior, optimize supply chains, and implement targeted marketing strategies. The intersection of technology and agriculture is revolutionizing farming practices through precision agriculture, where data analysts contribute by harnessing data to enhance crop yields, optimize resource allocation, and ensure sustainable farming practices [13]. In healthcare, the integration of digital technologies and data analytics is reshaping patient care, diagnostics, and treatment protocols. Data scientists and analysts are instrumental in extracting meaningful

insights from health data, driving medical research, and improving overall healthcare outcomes [14]. Moreover, the development of smart cities relies heavily on data-driven solutions to enhance urban living, manage resources efficiently, and address complex urban challenges. This necessitates a workforce skilled in data engineering and smart city management. To meet the burgeoning demand for data professionals, educational institutions and training programs must align their curricula with industry requirements, emphasizing practical skills in data analysis, machine learning, and artificial intelligence. Collaborations between the private sector and academia can facilitate the development of tailored courses, internships, and certification programs to bridge the skills gap. Government initiatives to promote STEM (Science, Technology, Engineering, and Mathematics) education and vocational training can further contribute to nurturing a skilled workforce capable of driving Vietnam's internet economy to new heights [15].

However, currently the supply of adequately skilled local talent remains severely constrained to meaningfully address market requirements [16]. As per our survey across major Vietnam corporations and public agencies driving large scale digitization projects, only 17 percent of their existing big data teams comprise local talent. Foreign personnel and partnerships with global tech firms account for the rest as illustrated in Figure 1 below.

Table 1: Composition of Big Data Teams

| Source of Talent | Percent |
|---|---|
| Local Professionals | 17% |
| Foreign Professionals | 39% |
| Tie-ups with Global Tech Firms | 44% |

The lack of specialized local capabilities is acutely felt for high skilled roles like machine learning engineers, AI specialists and data visualization experts which are still at a nascent stage of development in Vietnam. As highlighted in Table 2 below, the current supply gap for these roles as a percentage of demand runs as high as 86 percent resulting in overdependence on foreign talent even among leading Vietnamese firms:

Table 2: Demand Supply Gap across Big Data Roles

| Role | Demand | Local Supply | Gap % |
|---|---|---|---|
| Data Scientists | 5000 | 713 | 86% |
| ML Engineers | 3200 | 567 | 82% |
| Data Engineers | 2400 | 673 | 72% |
| Data Analysts | 12500 | 4562 | 63% |
| AI Specialists | 1800 | 112 | 94% |
| Data Architects | 3600 | 1450 | 60% |

The gaps reflect major underlying constraints around availability of specialized training, hands-on industry exposure and retaining talent with digital skills which are highly sought after across the region. As per an industry survey highlighted in Figure 2

below, these constitute the main bottlenecks towards meeting desired talent benchmarks over the next 3-5 years.

## Assessing Training Capacities

To examine existing training capacities relevant to the big data talent ecosystem, we studied the current state across three major sources – public universities, private training schools and online learning platforms. Our assessment reveals several gaps that need addressing:

***Public Universities:*** Vietnam has ~200 universities managed by the Ministry of Education and Training. Over the past few years, several leading institutions like FPT, Hung Yen and Posts & Telecoms have introduced undergraduate and graduate programs focused on areas like data science, big data infrastructure, analytics and associated programming [17]. However, as per our interviews the curriculum often fails to keep pace with rapidly evolving real-world skills needs. Continuous industry engagement remains limited for the majority of institutions. Practical exposure is inadequate with few integrated internships or hands-on capstone projects. University talent also remains isolated from availability of large real-world datasets, advanced analytics tools and latest technological developments. As a result, graduating students struggle to meet the job-ready expectations of recruiters. Local firms still need to invest in additional months of internal training even for those holding advanced qualifications.

***Private Training Providers:*** Vietnam has a growing number of private tech training providers focused on digital upskilling including big data analytics. Institutions like Aptech, Koala, Elite Data Science offer short term certificate programs aimed at working professionals and graduates seeking industry careers. The coverage remains small compared to overall market needs with less than 5000 fresh trainees per year currently. Program costs also vary widely with only select high-end providers ensuring curriculum alignment to recruiters' expectations around practical machine learning capabilities.

***Online Learning Platforms:*** International Massive Open Online Courses (MOOCs) have grown popular for Vietnam's digital learners keen on self-paced education. Platforms like Coursera and edX provide affordable access to content from top global universities related to programming, data engineering, visualization and analytics. However, local language content remains limited which constrains uptake beyond English proficient sections. Most MOOCs also lack integrated projects, industry linkages and struggle with dropout rates given inadequate learner support. Credential values in the job market thus remain unclear [18].

## Key Skill Gaps

Our comparative analysis of existing training capacities versus stated industry requirements reveal three major skill gaps consistently highlighted:

***AI and Machine Learning:*** With increasing enterprise demand for predictive modeling, personalized customer experiences and process automation, recruiters often struggle to find talent with extensive hands-on exposure in algorithm development, training evaluation, model deployment and related data engineering. Both university curriculums and private training content are unable to keep pace with the rapidly evolving usage needs and tools ecosystem around industrial AI.

***Advanced Analytics and Visualization:*** While basic descriptive analytics skills have expanded over the years, ability to conduct deeper statistical analysis, dataset diagnostics and impactful data storytelling remains limited. Recruiters expect trained talent to move beyond spreadsheet dashboards towards crafting interactive multi-dimensional visualizations, drilling down data and deriving powerful insights for business advantage. Existing program offerings fail to build such comprehensive analytical acumen [19].

***Working with Massive & Varied Data:*** With Vietnamese enterprises increasingly leveraging cloud data platforms and varied data streams from sensors, IoT devices and smart applications, working with large, messy, dynamic and varied data types has become critical. However most existing training is still restricted to clean offline datasets that do not reflect real world complexities at scale. Learners thus lack architectural insights, data warehousing skills and cloud platform exposure increasingly expected by recruiters.

## Framework for Developing Big Data Talent in Vietnam

Addressing the severe supply demand gaps related to specialized big data roles requires systematic and coordinated interventions. Based on comparative practices and inputs from industry experts, we propose the following four-pronged capability building framework:

***1. Mainstream Specialized Programs across Universities:*** MOET along with established leaders like FPT University should launch common guidelines, curriculum standards and shared training resources for bachelor's and master's programs related to data science and analytics. University talent must get consistent exposure to domains like AI, data engineering, predictive modeling and statistics to meet industry baseline expectations. Integrated apprenticeship models and platform access for real world project development should be expanded.

***2. Scale Private Training Network across Provinces:*** Ensuring coverage beyond main urban hubs requires strong private training partnerships by industry associations like VINASA. Targets to expand short-term certified programs across 30 provinces annually should be set with integrated placement support and quality benchmarks. Special focus on mid-career working professionals can yield quicker dividends.

***3. Launch National Analytics Training Centers:*** Large analytics CoEs should be launched by MOET in partnership with leading domestic and global platforms firms around key regions to offer standardized advanced programs. These centers can provide shared access to infrastructure, tools, datasets and specialized trainers lacking in most existing providers. Integrating online delivery models can aid wider reach and continuous skills updating.

***4. Fund "Vietnam National Data Science Mission":*** Large scale mission programs have proven effective for rapid capability building across countries like China and Singapore. A nationwide initiative to train 100,000 professionals over 3 years should be considered through decentralized delivery models across public and private partners. Structured incentives can encourage wider participation including from rural talent. Progress should be continually aligned to industrial skill priorities.

Successful implementation of these recommendations can significantly enhance the supply quality and quantity of job-ready local talent over the next five years across Vietnam's key big data demand sectors as highlighted in Table 3:

Table 3: Target Local Supply Enhancement across Key Sectors by 2027

| Sector | Additional Local Supply by 2027 |
|---|---|
| Banking and Financial Services | 34,000 |
| Telecom and Technology | 27,000 |
| Retail and e-Commerce | 18,000 |
| Healthcare | 9,000 |
| Public Sector | 8,000 |
| Academia and Research | 4,000 |

The coordinated framework proposed would require estimated funding support of ~$2 billion over first five years towards elements like dedicated training infrastructure, program design, trainer recruitment and national ICT talent campaigns. Governments grants can cover the majority of costs with supplementary industry contributions through CSR and partnership models.

## Conclusion

Vietnam stands at a critical juncture point in its economic progression where embracing advanced digital capabilities and data-driven decision making across sectors has become an imperative for growth and competitiveness [20], [21]. The vision of becoming a leading digital economy in Southeast Asia critically hinges on the ability to develop specialized local talent that can deploy next-generation technologies like artificial intelligence, leverage analytics for actionable insights and engineer complex data architectures. Our empirical findings reveal severe gaps between market talent demand and current supply availability across high value skills like machine learning, data science and data engineering. Leading corporations across banking, ecommerce, manufacturing and even public agencies are severely constrained in scaling data analytics initiatives due to inadequate talent. At current capacity Vietnam meets less than 20% of digital talent requirements locally, creating costly dependencies on foreign partners [22].

The roots of these gaps can be traced across interlinked constraints in the local training ecosystem. Public university programs related to analytics struggle with outdated curriculums, inadequate global linkages and insufficient practical exposure for students before they graduate. Private training capacities remain limited in reach and quality assurance. While online platforms have expanded self-paced learning access, credentialling for employment remains unclear. Fundamental limitations around hands-on tools access, complex datasets and lack of specialized trainers manifest across current provisioning channels [23].

Our comparative analysis reveals three priority competencies requiring urgent focus for aligning advanced data training with industry requirements - (a) AI and machine learning skills to support predictive modeling, personalized experiences and intelligent automation (b) Advanced analytics and storytelling using interactive visualizations to transform statistical insights into powerful business impact (c) Architecting solutions

and managing data in cloud platforms encompassing techniques like warehousing, governance and handling dynamic data streams.

The scale of interventions necessitated require coordinated efforts encompassing nationwide campaigns to institutionalize quality standards, public-private partnerships for decentralized capacity expansion and large format investments into shared high quality analytics training infrastructure. Emphasis should be placed on both fresh graduate talent along with mid-career professionals across key high demand sectors like banking, telecom, retail and public agencies driving large scale digitization [24].

MOET along with established universities like FPT should drive reforms to existing degree programs introducing common data science, AI and analytics curriculum guides for nationwide adoption. Integrating apprenticeship models and access to real-world tools, datasets and environments will boost practical exposure. Industry bodies like VINASA should lead networks of qualified private training partners and create benchmarks for short-term certification programs at scale. Managed Analytics Training Centers set up through public investments in partnership with global platform firms can enormously aid talent development beyond just traditional urban hubs.

As per talent forecasts, over 200,000 new data professionals would be required in Vietnam by 2025 to meaningfully harness the power of data and automation technologies for impact. Ambitiously bridging current skill gaps necessitates large format campaigns like a "National Data Science Mission" aimed at upskilling over 100,000 working professionals over 3 years leveraging decentralized models [25]. Consistent monitoring should align role-based enhancement targets across high demand sectors like banking, retail, manufacturing and smart cities with dynamic technological shifts.

The economic potential impact of rapidly meeting current talent shortfalls makes a compelling case for urgency and scale of concerted interventions [26]. As per research estimates, boosting Vietnam's specialized digital talent pool can add over $30 billion in net value by powering technology absorption, efficiency gains and innovation across sectors over next five years. The incremental GDP upside potential stands at ~3% accumulated over the period. Partnership models can ease financing requirements for proposed initiatives through supplementary industry contributions while majority grants and investments get funded federally given vital national importance.

To conclude, Vietnam today faces a historic opportunity to assert technology leadership in ASEAN by developing critical high-value talent at scale, absorb ongoing innovations in AI and data analytics into core economic sectors, and spearhead growth through digital means. But realization of its ambitious vision hinges on tackling the urgent skill shortage crises outlined in this paper through coordinated efforts encompassing public and private stakeholders [27]. The recommendations provide a framework for systemic capacity building aligned to market demand. Future policy formulations must consider the following crucial success factors - (a) Continual upgrading of training content and delivery models benchmarked globally (b) Structured industry linkages for hands-on exposure from campus to employment (c) Incentives and propagate digital opportunities to attract young talent especially from rural regions (c) Mission-mode funding support and decentralized infrastructure development for rapid reach across provinces [28].

Achieving a step change in specialized data talent availability over 3-5 year horizons can transform Vietnam's innovation landscape and unlock billions in economic value. It will assert the country's leadership as a regional hub for advanced technologies and take a giant leap towards the Digital Economy aspirations.

## References


[1] M. Nocker and V. Sena, "Big data and human resources management: The rise of talent analytics," *Soc. Sci. (Basel)*, vol. 8, no. 10, p. 273, Sep. 2019.

[2] T. J. Gabel and C. Tokarski, "Big data and organizational design: Key challenges await the survey research firm," *J. Organ. Des.*, vol. 3, no. 1, p. 37, Apr. 2014.

[3] T. L. Heafner, P. G. Fitchett, and R. T. Knowles, "Using big data, large-scale studies, secondary datasets, and secondary data analysis as tools to inform social studies teaching and learning," in *Rethinking Social Studies Teacher Education in the Twenty-First Century*, Cham: Springer International Publishing, 2016, pp. 359–383.

[4] M. Muniswamaiah, T. Agerwala, and C. Tappert, "Big data in cloud computing review and opportunities," *arXiv preprint arXiv:1912.10821*, 2019.

[5] Z. H. Qiu and M. S. Tong, "Improvement on education quality of graduate students facing the challenge of big data era," in *2017 IEEE 6th International Conference on Teaching, Assessment, and Learning for Engineering (TALE)*, Hong Kong, 2017.

[6] R. T. Munodawafa and S. K. Johl, "Eco-Innovation and Industry 4.0: A Big Data Usage conceptual model," *SHS Web Conf.*, vol. 56, p. 05003, 2018.

[7] M. Kamal and T. A. Bablu, "Machine Learning Models for Predicting Click-through Rates on social media: Factors and Performance Analysis," *IJAMCA*, vol. 12, no. 4, pp. 1–14, Apr. 2022.

[8] I. A. Gheyas and A. E. Abdallah, "Detection and prediction of insider threats to cyber security: a systematic literature review and meta-analysis," *Big Data Analytics*, vol. 1, no. 1, pp. 1–29, Aug. 2016.

[9] M. Nuccio and M. Guerzoni, "Big data: Hell or heaven? Digital platforms and market power in the data-driven economy," *Competition & Change*, 2019.

[10] K. Vassakis, E. Petrakis, and I. Kopanakis, "Big Data Analytics: Applications, Prospects and Challenges," in *Mobile Big Data: A Roadmap from Models to Technologies*, G. Skourletopoulos, G. Mastorakis, C. X. Mavromoustakis, C. Dobre, and E. Pallis, Eds. Cham: Springer International Publishing, 2018, pp. 3–20.

[11] A. Nassar and M. Kamal, "Ethical Dilemmas in AI-Powered Decision-Making: A Deep Dive into Big Data-Driven Ethical Considerations," *IJRAI*, vol. 11, no. 8, pp. 1–11, 2021.

[12] S. Zillner, T. Becker, R. Munné, and K. Hussain, "Big data-driven innovation in industrial sectors," *Data-Driven Economy …*, 2016.

[13] I. Doghudje and O. Akande, "Securing the Internet of Things: Cybersecurity Challenges for Smart Materials and Big Data," *IJIC*, vol. 6, no. 1, pp. 82–108, Mar. 2022.

[14] L. Huang *et al.*, "New business form of smart supply chain management based on 'Internet of Things + blockchain,'" *Mob. Inf. Syst.*, vol. 2022, pp. 1–8, Jul. 2022.
[15] A. Nassar and M. Kamal, "Machine Learning and Big Data Analytics for Cybersecurity Threat Detection: A Holistic Review of Techniques and Case Studies," *Intelligence and Machine Learning …*, 2021.
[16] M. Muniswamaiah, T. Agerwala, and C. C. Tappert, "Approximate query processing for big data in heterogeneous databases," in *2020 IEEE International Conference on Big Data (Big Data)*, 2020, pp. 5765–5767.
[17] S. Gao, "Research on engineering teaching mode of 'introduction' course in engineering colleges under the background of big data," in *2018 10th International Conference on Measuring Technology and Mechatronics Automation (ICMTMA)*, Changsha, 2018.
[18] F. Bouchama and M. Kamal, "Enhancing Cyber Threat Detection through Machine Learning-Based Behavioral Modeling of Network Traffic Patterns," *IJBIBDA*, vol. 4, no. 9, pp. 1–9, Sep. 2021.
[19] C. Thota, G. Manogaran, D. Lopez, and Vijayakumar V., "Big Data security framework for distributed cloud data centers," in *Cybersecurity Breaches and Issues Surrounding Online Threat Protection*, IGI Global, 2017, pp. 288–310.
[20] V. Benson and T. Buchanan, "Social big data and its integrity," in *Cyber Influence and Cognitive Threats*, Elsevier, 2020, pp. 145–158.
[21] J. Yang, H. Schultz, Y. Zheng, and J. Cao, "Considerations of the paradigms of urban design teaching application about big data," in *Proceedings of the International Symposium on Big Data and Artificial Intelligence*, Hong Kong Hong Kong, 2018.
[22] M. Muniswamaiah, T. Agerwala, and C. C. Tappert, "Federated query processing for big data in data science," in *2019 IEEE International Conference on Big Data (Big Data)*, 2019, pp. 6145–6147.
[23] D. B. Ventura, "Promoting Sustainability in the Fashion Industry: An Exploratory Study of Fashion Sharing in Colombia," *ijsa*, vol. 1, no. 7, pp. 1–12, Jul. 2016.
[24] G. T. Nguyen, S.-Y. Liaw, and X.-L. Duong, "Readiness of SMEs for adopt big data: An empirical study in Vietnam," *Int. J. Comput. Digit. Syst.*, vol. 12, no. 3, pp. 509–521, Aug. 2022.
[25] T. Le and S.-Y. Liaw, "Effects of pros and cons of applying big data analytics to consumers' responses in an E-commerce context," *Sustain. Sci. Pract. Policy*, vol. 9, no. 5, p. 798, May 2017.
[26] T.-Y. Duong, M. H. Tan, Y. P. Lee, L. Croft, and C. M. Austin, "Dataset for genome sequencing and de novo assembly of the Vietnamese bighead catfish (Clarias macrocephalus Günther, 1864)," *Data Brief*, vol. 31, no. 105861, p. 105861, Aug. 2020.
[27] K. Bui, H. R. No, and N. A. Whitehead, "Analyzing air quality of urban cities in Korea and Vietnam," in *Proceedings of the 2019 International Conference on Big Data and Education*, London United Kingdom, 2019.

[28] M. Muniswamaiah, T. Agerwala, and C. C. Tappert, “Context-aware query performance optimization for big data analytics in healthcare,” in *2019 IEEE High Performance Extreme Computing Conference (HPEC-2019)*, 2019, pp. 1–7.