VIPUL H. HARIHAR

LEAD SOFTWARE ENGINEER | Enterprise AI • Wealth Technology • Cloud Infrastructure • Product Architecture

New York, NY • vhh2105@columbia.edu • LinkedIn • GitHub • Portfolio • Credly • Link Hub

PROFESSIONAL PROFILE

Lead Software Engineer and end-to-end platform architect with 6+ years of experience across enterprise AI, financial technology, cloud infrastructure, full-stack systems, machine learning, healthcare engineering, and government research. At GenTrust, independently transformed a legacy modernization assignment into AIRIA and AURA: a secure, AI-powered proprietary operating environment that connects investment, trading, benchmark, risk, research, reporting, advisor, CRM, document, compliance, and operational workflows for a registered investment adviser managing more than $5 billion in assets.

Scope extends beyond individual contribution into technical governance, product ownership, and production accountability. Responsibilities include architecture decisions, cloud and production operations, cybersecurity controls, code and design review, supervision of contractors and collaborating software or infrastructure engineers, release acceptance, performance, incident resolution, user experience, roadmap translation, and direct collaboration with executive, investment, compliance, trading, reporting, and operations stakeholders. Built the flagship AI platform across front end, back end, APIs, mobile integration, dashboards, data pipelines, deployment, and visual design within approximately four months while retaining the official title of Lead Software Engineer.

LEADERSHIP SCOPE & CORE EXPERTISE

- Enterprise AI and LLM systems: agent frameworks, retrieval and context engines, document intelligence, prompt and workflow design, bounded model behavior, multilingual interaction, clarification flows, and high-stakes output controls.
- Financial technology: portfolio oversight, trading workflows, benchmark management, risk exposure, tax-loss harvesting, research automation, advisor productivity, CRM integration, reporting, and investment-data operations.
- Cloud and production architecture: AWS and Azure, APIs, authentication, role-based access, data pipelines, distributed services, monitoring, logging, deployment, scalability, reliability, performance optimization, and incident response.
- Engineering leadership: architecture review, code review, technical acceptance criteria, contractor direction, cross-stack integration, documentation, testing strategy, release governance, mentoring, quality standards, and production accountability.
- Product ownership: 0-to-1 design, requirements discovery, stakeholder translation, UX and visual design, roadmap prioritization, adoption, user support, executive demonstrations, and commercialization-readiness.
- Contractor and engineering supervision: assign application-development and cloud-infrastructure work, review pull requests and infrastructure changes, validate security and performance, require rework where needed, approve or reject deliverables, and remain accountable for production deployment and integration.
- Cost and operating leverage: design reusable in-house software, replace fragmented external applications and manual workflows, reduce model-token consumption, avoid duplicate vendor spend, and convert recurring operational work into scalable platform capabilities.

PROFESSIONAL EXPERIENCE

GenTrust | Lead Software Engineer

New York, NY | Full-time, on-site | Feb 2023 - Present | AI platforms, cloud infrastructure, engineering governance, investment systems, product ownership

- Own the architecture, engineering, deployment, security, reliability, performance, and continued evolution of AIRIA and AURA, the firm's proprietary investment-technology and enterprise-AI platforms.
- Designed and built an AI-powered proprietary operating environment in approximately four months; the platform launched in July 2025 and became a core component of the firm's trading-technology environment.
- Built the system end to end as its primary architect and developer across front-end interfaces, back-end services, APIs, mobile integration, dashboards, data pipelines, databases, cloud deployment, and product visual design, initially operating on an API budget of approximately $50 per month.
- Serve as final technical reviewer for code, architecture, cloud-infrastructure changes, deployment plans, security controls, and production deliverables created by contractors and collaborating software or infrastructure engineers.
- Review pull requests and implementation work for correctness, maintainability, performance, security, compatibility with the AIRIA/AURA architecture, and readiness for production; identify defects, define rework, and approve technical acceptance.
- Set engineering direction, break initiatives into executable workstreams, establish technical standards and acceptance criteria, coordinate external contributors, mentor developers, and retain accountability for the integrated production system.
- Direct testing, user-support, adoption, documentation, and smaller engineering work assigned to contractors or supporting contributors while protecting the architectural consistency of the core platform.

- Supervise contractors and external contributors across both software development and cloud infrastructure. Define scope and priorities, review implementation plans, inspect code and configuration changes, validate access and security design, test integrated behavior, direct corrections, and approve production readiness.
- Retain final accountability for contractor output even when work is delegated, including architectural compatibility, data protection, cost control, observability, rollback planning, deployment quality, documentation, and offboarding of access when an engagement ends.
- Architect and operate AWS/Azure infrastructure, authentication, APIs, databases, secrets, monitoring, logging, testing, debugging, deployment, backups, and performance optimization for business-critical applications.
- Resolved production memory and concurrency constraints as platform adoption expanded to 20+ simultaneous users; optimized service behavior, resource management, and responsiveness for firm-wide use.
- Designed privacy and AI-governance controls including Anthropic Zero Data Retention access, PII masking and redaction, restricted task scopes, role-based access, audit trails, data separation, and clarification when context is insufficient.
- Built a custom knowledge and context system that connects thousands of internal documents, workflows, procedures, compliance guidance, portfolio data, and business rules to bounded AI experiences without exposing firm data for model training.
- Developed AURA's application and agent framework for advisor, trader, researcher, portfolio, quantitative, planning, legal, fixed-income, accounting, Salesforce, Addepar, reporting, and document-analysis workflows.
- Built AURA itself as an in-house enterprise platform, not only as a thin connection to an external model. Designed the user experience, application builder, agent routing, context assembly, memory and knowledge behavior, guardrails, permissions, model integration, auditability, document handling, and connections to AIRIA and firm systems.
- Designed reusable AURA applications to replace or consolidate standalone vendor tools and repetitive internal processes across research, portfolio analysis, reporting, document review, meeting summaries, CRM notes, advisor support, knowledge search, and operational requests.
- Reduced model and token expense through scoped prompts, retrieval-first context, selective context loading, reusable system instructions, structured calculations outside the model, response-size controls, model selection, caching and precomputation where appropriate, and routing only the necessary work to higher-cost models.
- Created role-aware and multilingual interaction patterns that adjust tone, detail, terminology, and explanation depth so outputs are useful to traders, advisors, researchers, operations teams, executives, and non-technical staff.
- Built portfolio and investment capabilities spanning benchmark history and overrides, current performance, underweight/overweight analysis, risk exposure, betas, tax-loss harvesting, trading support, portfolio automation, client intake, and firm-wide reporting.
- Integrated Salesforce advisor logs, Addepar reports, document repositories, internal databases, secure LLM APIs, PDF/Excel analysis, benchmark data, and reporting templates into a unified natural-language operating layer.
- Developed document-intelligence workflows for research reports, financial statements, planning and tax materials, Excel workbooks, and password-protected PDFs, producing structured summaries, extracted data, recommendations, and role-specific outputs.
- Built meeting and CRM intelligence capabilities for speech-to-text, summaries, action items, advisor notes, and structured follow-through, reducing repetitive administrative work while preserving business context.
- Designed reporting automation for presentations, portfolio summaries, client-ready bullets, structured data tables, and internal support workflows, including automated routing of user requests to the appropriate support channels.
- Partner directly with the CIO, CFO, CCO, COO, president, investment leadership, trading, compliance, reporting, advisory, operations, and technology teams to convert ambiguous business requirements into secure production products.
- Lead architecture reviews, design reviews, code reviews, documentation, testing, demonstrations, usability improvements, guardrails, redacted sample environments, user training, support, and release readiness.
- Coordinate cross-stack integration between Python-based AI services and broader enterprise application components, preserving interoperability without requiring a full rewrite of proven systems.
- Documented more than 80 hours of firm-wide time savings in a 30-day usage period. Public portfolio and internal program estimates attribute 25+ automated data pipelines and projected operating savings above $2 million to the platform program, with further avoided cost from in-house application replacement, reduced external development dependence, lower software licensing requirements, and disciplined LLM token usage.
- Support a firm with approximately 45-48 employees, 28 advisors, roughly 250 clients, and more than $5 billion in assets under management through platforms used across investment and operating teams.
- Contribute to platform commercialization, external demonstrations, intellectual-property discussions, and product strategy while maintaining strict separation between internal production data and approved external examples.

AI AGENT GOVERNANCE, RISK AND COMPLIANCE FRAMEWORK

- Access delta review: for each proposed agent, identify every new system, file type, database, document repository, data source, API, tool, credential, and permission it would access compared with existing agents. Document the business purpose, minimum necessary access, approved users, and whether the access creates a new security or compliance boundary.

- Action and identity review: determine whether the agent can act on behalf of a user, use a user credential or service account, impersonate a user, send communications, change records, initiate workflows, execute trades, create reports, modify permissions, or take any action in firm systems beyond the authority of existing agents. Require explicit controls, approvals, logging, and human confirmation for higher-risk actions.
- Data-use and third-party review: evaluate whether the agent changes how information is collected, transmitted, retained, cached, logged, used for model training, shared with vendors, or made available to subprocessors. Confirm Zero Data Retention or equivalent contractual controls where required and prohibit use of firm or client information for external model training.
- Restricted-data review: determine whether the agent can encounter client data, personally identifiable information, material nonpublic information, portfolio holdings, transactions, performance, tax information, authentication data, legal or compliance materials, employee information, or confidential firm intellectual property. Apply masking, redaction, segmentation, role-based access, data minimization, and environment separation based on the sensitivity of the information.
- Control and lifecycle review: define additional access controls, monitoring, audit logs, usage limits, human review, cybersecurity testing, vendor due diligence, change management, incident escalation, compliance approval, business-owner accountability, periodic recertification, and offboarding steps before the agent is released or materially changed.
- Operational approval process: require pre-production testing with redacted or synthetic data, document expected and prohibited behaviors, validate failure and rollback paths, test prompt injection and unauthorized-access scenarios, confirm logging and alerting, and obtain the appropriate technology, cybersecurity, compliance, and business approvals before broader deployment.

GenTrust | Software Engineer Intern

New York, NY | Aug 2022 - Jan 2023

- Supported software engineering, data science, testing, documentation, and internal application development before promotion to Lead Software Engineer.
- Contributed to Java-based systems, data workflows, troubleshooting, platform modernization, and early architecture work supporting the firm's investment-technology environment.

NewYork-Presbyterian / Columbia University Medical Center | Research Assistant & Engineer

New York, NY | Volunteer research | May 2022 - Dec 2022

- Engineered an encrypted, offline-capable medical passport application supporting a research cohort of 79 pediatric cardiac patients while minimizing server-side exposure of sensitive information.
- Led two application-development initiatives, mentored three research assistants in coding and debugging, and built two data pipelines to improve client-side analysis speed and accuracy.
- Applied privacy-conscious application architecture, local encryption, secure data handling, and user-centered product design in a sensitive clinical-research environment.

NYC Department of Design and Construction | Data Science Research Fellow - Wind Energy Capstone

New York, NY | Sep 2022 - Dec 2022

- Led a five-member Columbia team analyzing approximately 15 GB of data to model the carbon impact and logistics of repurposing about 70,800 decommissioned wind-turbine blades.
- Developed a geographic optimization approach to minimize greenhouse-gas emissions across transportation and reuse scenarios and delivered a cloud-hosted sustainability-planning application.
- Combined machine learning, geospatial analysis, data engineering, visualization, cloud deployment, and stakeholder presentation for a New York City infrastructure use case.

Columbia University | Enhanced Teaching Assistant - Analysis of Algorithms

New York, NY | Jan 2022 - Dec 2022

- Supported CSOR W4231 for Professor Eleni Drinea through grading, exam proctoring, office hours, student correspondence, course administration, and explanation of advanced algorithmic concepts.
- Received the Best Data Science Course Assistant Award for 2022 and 2023.

IBM | Application Developer

Bengaluru, India | Feb 2021 - Aug 2021

- Developed and supported full-stack enterprise applications using Java, SQL, JavaScript, web technologies, and relational databases; diagnosed production defects and implemented fixes across application and data layers.
- Modernized financial application infrastructure by contributing to migration of Java modules toward Red Hat/OpenShift and microservice-oriented deployment patterns.
- Automated SQL and reporting workflows, supported MySQL/DB2-related data operations, and improved release and CI/CD processes for production applications.
- Public portfolio materials report a 33% reduction in hosting cost associated with modernization work and streamlined global production-release processes.

IBM | Associate System Engineer

Bengaluru, India | Jan 2021 - Feb 2021

- Supported enterprise application operations, technical troubleshooting, testing, SQL and data workflows, and the transition into the Application Developer position.

IBM | Java Full-Stack Development Engineer Trainee

Dehradun, India | Feb 2020 - Dec 2020

- Completed full-stack engineering training across Java, SQL, web technologies, application architecture, debugging, testing, source control, and the software-development lifecycle.
- Completed extensive IBM technical learning and credential tracks across artificial intelligence, Python, Cognos, business intelligence, predictive analytics, data science, big data, Spark, Kafka, Hadoop, and visualization.
- Received IBM Golden Employee recognition in Feb 2021 for distinguished learning effort, skill development, and contribution to project value.

Indian Road Safety Campaign | Software Engineer Intern / Policy Intern

India | Jul 2020 - Dec 2020

- Built a full-stack accident-risk and road-defect mapping platform using Angular and Spring technologies, supporting road-safety awareness for more than 16,000 commuters.
- Developed real-time map-based risk identification and alerting for 17 accident-prone areas, enabling users to identify nearby hazards and submit structured reports to authorities.
- Combined software engineering, location-aware data, public-policy objectives, analytics, and accessible product design for a national safety initiative.

Oil and Natural Gas Corporation (ONGC) | Big Data Engineer Intern

Dehradun, India | May 2019 - Jun 2019

- Supported big-data analytics using Hadoop, Google BigQuery, Python-based data collection, and open-source datasets.
- Built processing and reporting workflows to extract, clean, analyze, and present large datasets for operational and research use cases.

Oil and Natural Gas Corporation (ONGC) | Summer Intern, Application Developer

Dehradun, India | May 2018 - Jun 2018

- Supported application-development and data-visualization work using JavaScript and operational datasets.
- Contributed to reporting interfaces and technical components used to organize and present processed information.

GENTRUST PLATFORMS & PRODUCT PORTFOLIO

- AIRIA - Investment Operating Platform: Cloud-hosted environment for portfolio oversight, benchmark management, trading support, risk analysis, tax workflows, reporting, application access, and operating automation.
- AURA - Enterprise AI Assistant and Application Builder: Secure agent platform integrated with AIRIA, approved firm knowledge, Salesforce, Addepar, documents, portfolio data, model APIs, and role-specific workflows.
- Benchmark and Portfolio Intelligence: Historical benchmarks, current performance, overrides, asset-class mapping, exposure analysis, underweight/overweight review, betas, risk tracking, and portfolio comparison.
- Trading and Tax Automation: Trading support, portfolio actions, tax-loss harvesting, investment workflow coordination, and decision-support outputs with guardrails against unauthorized live or algorithmic execution.
- Advisor and Salesforce Intelligence: Recent intake, advisor logs, meeting context, CRM notes, client preparation, follow-up, structured summaries, and searchable business history.
- Addepar and Reporting Automation: Template-driven reports, portfolio and client summaries, presentation content, structured tables, automated support routing, and reporting-data mapping.
- OTO - Meeting Intelligence: Client-side speech-to-text, secure meeting summaries, action items, and CRM integration designed for rapid post-meeting completion.
- ACE - Concierge Analytics: Forecasting, dashboarding, and advisor-facing insights designed for rapid interpretation and operational follow-through.
- Quantitative and Backtesting Tools: Flask-based research and backtesting workflows, high-frequency data refresh, quantitative analysis, and micro-hedging experimentation.
- Research and Document Intelligence: PDF, Excel, research, password-protected document, financial statement, and planning-document extraction, reasoning, summarization, and reporting.

PUBLIC WEBSITES, APPLICATIONS & CODE PORTFOLIO

- Personal Portfolio Website - self-coded professional site hosted on GitHub Pages; presents experience, projects, recognition, credentials, education, contact information, and the Columbia Data Science Institute feature.

- GitHub - public engineering portfolio with 47 public repositories, covering machine learning, neural networks, algorithms, statistics, NLP, web applications, APIs, data visualization, Java/Spring services, cloud projects, and full-stack prototypes.
- VeeCreate Audio Books - browser application that converts user-provided text into spoken audio, supports background music and preloaded stories, and combines storytelling, text-to-speech, and interactive media design.
- VeeMusic DJ - JavaScript/YouTube Data API application with advanced music-video search, random playback, rapid-sequence DJ mode, and live-camera visual backgrounds.
- VeePalm-Oracle - responsive camera-enabled palmistry entertainment application using getUserMedia, secure HTTPS capture, randomized readings, and front-end interaction design.
- VeeTasks - dynamic task-management web application with text and image entries, drag-and-drop ordering, local-storage persistence, responsive layout, and clear-all controls.
- DJ VeePULL - immersive online DJ proof of concept with virtual mixing controls, pitch experimentation, track recording, background-music integration, and mobile compatibility.
- VeeLens - real-time browser object identification using TensorFlow.js, COCO-SSD, webcam input, detection across 90 object classes, YouTube media integration, and interactive overlays.
- Deep Learning DJ - Columbia applied-deep-learning project using camera-based emotion recognition and music recommendation to support mental-health-oriented interaction and positive user experience.
- Wind Turbines Repurposing - AI Carbon Emission Calculator - geospatial and machine-learning application for evaluating recovery and direct reuse of decommissioned wind-turbine blades while minimizing greenhouse-gas emissions.
- Medical Passport - encrypted, offline-first healthcare application designed for pediatric cardiac research workflows and privacy-conscious local data access.
- Omnifood - Spring Boot food-service and analytics platform designed during COVID-19 to support restaurant delivery, sales insights, customer discounts, food donations, NGOs, and reduction of food waste.
- Machine Learning-Based Imaging System for Road Defect Inspection - location-aware public-safety platform for potholes, damaged infrastructure, congestion, accident reports, escalating map severity, and nearby-user alerts.
- Real-Time Data Processing Using Big Data Analytics - Python scraping, cleaning, keyword bucketing, regression, trend counting, and visualization pipeline using open-source news data from multiple websites.
- TechHub - public collection of code examples, projects, and algorithms across Python, Java, HTML, CSS, JavaScript, data science, machine learning, NLP, networking, and cloud computing.
- Quiz App Spring Boot Backend - Java API and database-backed service for retrieving quizzes and associated questions.
- EDAV Community Contribution: Kickstart R - public Columbia learning resource covering exploratory data analysis, R fundamentals, data structures, functions, missing values, and visualization.
- Selected public ML repositories: DecisionTrees--RandomForests-Hypertuning-Caliberation; Nueral-Networks-and-CNN-and-RNN; RandomSampler_SMOTE_Classification_methods; Statistics_Inferencial_Modelling.
- Professional Link Hub - consolidated public access to LinkedIn, the self-coded portfolio, Credly, GitHub, Quora, YouTube, Instagram, and other public profiles.

TECHNICAL SKILLS

Languages: Python, Java, JavaScript, SQL, Bash, R, C/C++, HTML, CSS; cross-stack integration with C# enterprise components

AI and Machine Learning: Anthropic Claude APIs, OpenAI GPT APIs, enterprise LLM integration, agent workflows, retrieval and context systems, NLP, TensorFlow, Keras, PyTorch, scikit-learn, XGBoost, classification, regression, forecasting, deep learning, computer vision

Web and Application Frameworks: Flask, Spring Boot, React, Angular, Node.js, jQuery, REST APIs, responsive web applications, server-side services, mobile integration

Cloud and DevOps: AWS EC2, S3, RDS, CloudFormation, Microsoft Azure, GCP/BigQuery, Docker, Git, CI/CD, Red Hat OpenShift, cloud monitoring, deployment automation, production operations

Data and Analytics: ETL and data pipelines, OLAP, data warehousing, MySQL, NoSQL, Oracle, DB2, pandas, Spark, Kafka, Hadoop, IBM Cognos, Power BI, Tableau, Matplotlib

Architecture and Security: Distributed systems, microservices, full-stack architecture, API design, role-based access, PII masking, Zero Data Retention, audit logging, secure document processing, data isolation, secrets management, performance optimization

Engineering and Product: Architecture review, code review, contractor oversight, technical acceptance, testing strategy, incident troubleshooting, documentation, mentoring, stakeholder management, product design, UX, prototyping, Agile/SDLC, release governance

EDUCATION

Columbia University | Master of Science in Data Science

New York, NY | GPA: 3.6/4.0 | Aug 2021 - Dec 2022

- Coursework: Applied Machine Learning, Natural Language Processing, Deep Learning, Big Data Technologies, Time Series Analysis, Algorithms for Data Science, Computer Systems for Data Science, Database Management and SQL Optimization, Neural Networks, Probabilistic Graphical Models, Data Warehousing, Cloud Computing, Distributed Systems, Bayesian Statistics, and Inferential Modeling.
- Leadership and service: Data Science Department Representative on the Engineering Graduate Student Council, Professional Development and Leadership Fellow, Enhanced Teaching Assistant, and Data Science Day support lead.
- Languages: English and Hindi - full professional proficiency; IELTS 7.0 overall (Nov 2019): Listening 7.0, Reading 7.0, Speaking 7.0, Writing 7.5.

DIT University | Bachelor of Technology in Computer Science and Engineering - Big Data Analytics

Dehradun, India | CGPA: 9.62/10.0 | Aug 2016 - May 2020

- Graduated first in the B.Tech Computer Science and Engineering program; University Program Topper, Gold Medal recipient, Roll of Honor student, and Academic Excellence Scholar.
- Coursework included Artificial Intelligence, Database Management Systems, Distributed Systems, Software Engineering, Computer Networks, Object-Oriented Programming, Big Data Engineering, and full-stack application development.

PROFESSIONAL CERTIFICATIONS & VERIFIED DIGITAL CREDENTIALS

- Google Data Analytics Professional Certificate - Google/Coursera | Jul 2022. Multi-course professional program covering data preparation, cleaning, analysis, visualization, case-study development, spreadsheets, SQL, R, and communication of analytical findings.
- Microsoft Certified: Azure Fundamentals - Microsoft | May 2021. Foundational certification covering cloud concepts, core Azure services, security, privacy, compliance, governance, pricing, service-level agreements, and cloud support models.
- Enterprise Design Thinking Practitioner - IBM. Credential recognizing practical use of IBM's human-centered design framework, including user outcomes, stakeholder alignment, iterative prototyping, multidisciplinary collaboration, and continuous learning through feedback.
- Artificial Intelligence Analyst - Mastery Award 2019 - IBM | issued Apr 2019. Advanced IBM learning recognition covering artificial intelligence concepts, analytical problem solving, applied machine learning, and interpretation of AI-supported business outcomes.
- Artificial Intelligence Analyst - Explorer Award 2019 - IBM | issued Apr 2019. Foundational IBM credential demonstrating structured learning in artificial intelligence terminology, use cases, data-driven analysis, and responsible application of AI methods.
- Applied Data Science with Python - IBM. Applied credential covering Python-based data analysis, manipulation, visualization, model development, and practical use of common data-science libraries.
- Data Science Foundations - Level 1 - IBM. Foundational credential covering the data-science lifecycle, analytical roles, problem framing, data preparation, model concepts, and communication of findings.
- Data Analysis Using Python - IBM. Credential focused on importing, cleaning, exploring, transforming, and analyzing datasets with Python and common analytical libraries.
- Data Science Methodologies - IBM. Credential covering structured approaches to business understanding, analytical problem definition, data requirements, preparation, modeling, evaluation, deployment, and feedback.
- Data Visualization Using Python - IBM. Credential covering creation and interpretation of analytical charts, dashboards, and visual narratives using Python visualization libraries.
- Predictive Analytics Modeler - Explorer Award - IBM. Learning recognition in predictive-model development, variable selection, evaluation, interpretation, and use of analytical models for decision support.
- Business Intelligence Analyst - Mastery Award for Students - IBM. Advanced student credential covering business intelligence concepts, reporting, dashboarding, data interpretation, and translation of information into business decisions.
- Business Intelligence Analyst - Explorer Award for Students - IBM. Foundational student credential in reporting, data exploration, business metrics, visualization, and analytical decision support.
- Big Data Foundations - Level 1 - IBM. Introductory credential covering big-data concepts, distributed storage, processing ecosystems, analytical use cases, and the role of technologies such as Hadoop.
- Big Data Foundations - Level 2 - IBM. Continued credential covering more advanced big-data architecture, distributed processing, data pipelines, and practical analytics in large-scale environments.
- Big Data Engineer - Explorer Award - IBM. Credential recognizing learning in large-scale data ingestion, storage, distributed processing, engineering workflows, and big-data platform concepts.
- IBM Certified Python and Cognos Developer - credential grouping listed on the public professional portfolio.
- IBM technical learning portfolio across Python, Cognos, business intelligence, predictive analytics, big data, data science, Spark, Kafka, Hadoop, analytics, and visualization.

- Aspiring Minds Business Analyst - issued Apr 2019 | Credential ID 11459070-94.
- Aspiring Minds Data Processing Specialist - issued Apr 2019 | Credential ID 11459070-211.
- Aspiring Minds English Comprehension - issued Apr 2019 | Credential ID 11459070-283.
- Aspiring Minds Proficiency in English - issued Apr 2019 | Credential ID 11459070-284.
- Software Development Trainee - Aspiring Minds certification listed on LinkedIn; covered open-source data collection, Python scraping, cleaning, keyword bucketing, regression, analysis, visualization, and report generation.
- Complete public credential wallet and verification links: credly.com/users/vipul-harihar/badges.

HONORS, LEADERSHIP & RECOGNITION

- Columbia Data Science Institute Alumni Feature | Sep 16, 2025. Selected for the Columbia profile “A Data Science Alum Was Hired to Fix Code. He Built an Operating System Instead.” The feature documents the progression from a legacy modernization assignment to the independent end-to-end design of a secure AI-powered operating platform for GenTrust, including front end, back end, mobile integration, dashboards, pipelines, security, and product design.
- Forbes 30 Under 30 2026 Nominee. Nomination listed on the public professional portfolio in recognition of work across enterprise AI, financial technology, and product engineering. This is accurately presented as a nomination and not as selection to the final Forbes list.
- Best Data Science Course Assistant Award | Columbia University | May 2023. The LinkedIn-listed award title references 2022 and 2023 and recognizes exceptional course-assistant support, technical explanation, grading quality, responsiveness, and contribution to the student learning experience.
- Gold Graduate Marshal | Columbia University | May 2023. Ceremonial and leadership recognition associated with commencement service and representation of the graduating class, reflecting academic engagement and trusted participation in the university community.
- Professional Development and Leadership Fellow | Columbia Engineering | Feb 2022. Faculty-nominated fellowship focused on professional growth, leadership judgment, communication, collaboration, self-awareness, and the ability to lead effectively across technical and multidisciplinary environments.
- Data Science Department Representative | Columbia Engineering Graduate Student Council | Sep 2021 - Dec 2022. Served as a formal liaison between data-science students, faculty, and the graduate student council. Gathered concerns, communicated updates, supported events, represented departmental needs, and helped strengthen cross-program engagement.
- IBM Golden Employee | Feb 2021. Internal recognition for distinguished learning effort, rapid skill development, contribution to project value, and consistent engagement with IBM technical and professional-development programs while supporting enterprise application work.
- University Program Topper, Gold Medal, and Roll of Honor | DIT University | Dec 2020. Graduated first in the Bachelor of Technology Computer Science and Engineering program with a 9.62/10.0 CGPA. The combined recognition reflects the highest program-level academic standing, sustained performance, and formal placement on the university honor roll.
- University Bronze Medal / Academic Honor | DIT University | Dec 2020. Separate university-level academic recognition recorded on the public professional profile, reflecting high overall academic performance and achievement during the undergraduate program.
- Merit-Based Scholarship and Academic Brilliance Award | DIT University. Awarded for strong and sustained academic performance, providing formal recognition of scholastic excellence during the undergraduate degree.

Recognition presentation note: Awards, nominations, fellowships, academic honors, and credentials are labeled precisely so that a reader can distinguish selection, nomination, certification, service, and academic standing.

PUBLIC PROFILES, MEDIA & PROFESSIONAL LINKS

- LinkedIn
- Self-Coded Portfolio
- Professional Link Hub
- YouTube Channel
- EDAV Community Contribution
- GitHub - 47 Public Repositories
- Credly Credential Wallet
- Quora - Popular Writer Profile
- Columbia Data Science Institute Feature