

ScribeEMR Systems is a leading healthcare support services company specializing in remote live medical scribing, AI-based medical scribing, medical coding, medical billing, and Virtual Medical Office Solutions (VMOS). With a strong presence across India through our offices in Bengaluru, Coimbatore, Hyderabad, and Mohali, we cater to a wide range of healthcare providers in the United States.

Our mission is to streamline clinical workflows, enhance documentation accuracy, and improve revenue cycle efficiency for healthcare organizations. At the forefront of innovation and service excellence, ScribeEMR Systems empowers medical professionals to focus on patient care while we take care of the rest. We are ranked #1 by KLAS for 2024 and 2025 for best in scribing services.

1. **Official Name of the Designation:** Principal Data Scientist (Healthcare AI, ASR, LLM, NLP, Cloud, Agentic AI)
2. **Reporting location:** Office-based job (Hebbal Ring Road, Bengaluru)
3. **Work timings:** Day shift.
4. **Educational qualifications required:**
  - a) Ph.D. or master's in computer science, AI, Machine Learning, or related field.
  - b) Technical certifications in AI/ML, NLP, or cloud computing would be an advantage.
5. **Prior Experience:** 7+ years of experience solving real-world problems using Natural Language Processing (NLP), Automatic Speech Recognition (ASR), Large Language Models (LLM), and Machine Learning (ML), preferably in the healthcare domain. Experience with Agentic AI, cloud deployments, and fine-tuning transformer-based models would be a significant advantage.
6. **Job Description:** This role will be part of ScribeEMR, a healthcare division of Focus Group, specializing in medical coding and scribing. We are building a suite of AI-powered, innovative, state-of-the-art web and mobile solutions to alleviate the administrative burden of EMR data entry, increase provider satisfaction and productivity, and help improve the quality of care and patient outcomes. Our solutions combine cutting-edge AI technology and live scribes to streamline clinical workflows and enhance clinical decision-making.

The Principal Data Scientist will be responsible for designing, developing, and deploying cognitive AI solutions, including text and speech analytics for healthcare applications. This role requires expertise in deep learning, classical ML, generative AI, cloud-based AI deployments, and agentic AI frameworks. The ideal candidate will be passionate about innovation, automation, and the development of scalable, high-performance AI solutions that provide a seamless user experience for our customers.
7. **Job Summary:**
  - a) Define and develop AI-driven solutions for speech recognition, text processing, and conversational AI.
  - b) Research and implement transformer-based models (Whisper, LLaMA, GPT, T5, BERT, etc.) for speech-to-text, medical summarization, and clinical documentation.
  - c) Architect and deploy AI models on cloud platforms (AWS, Azure, GCP), ensuring scalability and performance optimization.
  - d) Develop and integrate Agentic AI frameworks, enabling multi-agent collaboration in AI-driven automation.
  - e) Fine-tune, train, and optimize large-scale NLP and ASR models to improve accuracy and efficiency.

- f) Collaborate with cross-functional teams, including software engineers, ML engineers, and domain experts, to integrate AI models into real-world healthcare applications.
- g) Design and implement MLOps strategies to streamline AI model lifecycle management.
- h) Ensure compliance with healthcare standards such as HIPAA, HL7, and FHIR, and integrate AI systems with EHRs/EMRs.
- i) Develop rigorous testing and validation criteria for AI models, ensuring high accuracy and performance in clinical settings.
- j) Conduct cutting-edge research in AI, ASR, and NLP and contribute to publications, patents, and AI strategy.

## 8. **Roles and Responsibilities:**

- a) Work with business analysts and healthcare professionals to define software requirements and identify AI use cases.
- b) Develop and optimize robust machine learning algorithms for speech, text, and structured data processing.
- c) Implement LangChain and OpenAI APIs to build modular, flexible AI applications.
- d) Deploy, monitor, and enhance models in cloud environments using containerization, Kubernetes, and serverless architectures.
- e) Lead and mentor junior data scientists and engineers, fostering a culture of AI excellence and innovation.
- f) Conduct performance evaluations and iterative improvements on AI models to enhance efficiency and accuracy.
- g) Stay ahead of emerging AI trends, evaluating the feasibility of new research advancements in NLP, ASR, and GenAI.
- h) Ensure AI solutions adhere to ethical AI principles, regulatory compliance, and bias mitigation techniques.
- i) Design scalable, reusable, and production-ready AI frameworks for speech and text analytics.

## 9. **Skills Required:**

- a) Expertise in Machine Learning, Deep Learning, and Generative AI models.
- b) Strong Python programming skills, with experience in PyTorch and TensorFlow.
- c) Hands-on experience in training and fine-tuning transformer-based LLMs (GPT, BERT, T5, LLaMA, etc.).
- d) Familiarity with ASR models (Whisper, Canary, wav2vec, DeepSpeech) and text embeddings.
- e) Proficiency in cloud computing (AWS, Azure, GCP) and cloud-based AI model deployment.
- f) Experience with LangChain, OpenAI APIs, vector databases (Pinecone, FAISS, Weaviate), and RAG architectures.
- g) Experience with agentic AI frameworks, reinforcement learning, and autonomous AI agents.
- h) Knowledge of FHIR, HL7, HIPAA compliance, and integration with EHR/EMR systems.
- i) Familiarity with containerization (Docker, Kubernetes) and MLOps best practices.
- j) Strong communication, collaboration, and mentoring skills.

## 10. **Reporting hierarchy:** Reports to SVP

## 11. **Compensation offered:** Best in the industry, for suitable candidates

## 12. **Additional Perks:**

- a) Medical insurance of 3 lakhs per annum (coverage for self, spouse, and up to 4 children under 25 years of age).
- b) Group Personal Accident Policy, covering 5 years' CTC in case of accidental death.

13. How to apply: Apply online on our website's career page → <https://www.scribeemr.in/careers>