Statement Of Work: AI Clinical Insights Engine

Introduction

This Statement of Work (SOW) defines the tasks, deliverables, and schedule for the "AI Clinical Insights Engine" project. It establishes the agreement between **HCA Healthcare** (the Client) and **Craig Frint** (the Contractor/Freelance Software Engineer).

Statement of work template

Title

AI Clinical Insights Engine: Natural Language Processing for Physician Notes

Abstract

This project involves the design, development, and deployment of a specialized Artificial Intelligence pipeline to analyze unstructured medical doctors' notes within the HCA Healthcare network. The goal is to utilize Natural Language Processing (NLP) to extract clinical entities, such as diagnoses, medications, and symptoms, to improve coding accuracy and clinical decision support. Craig Frint will be responsible for the end-to-end development of the AI models, data preprocessing scripts, and API integration layers. The project is expected to span six months, resulting in a production-ready model that integrates seamlessly with HCA's existing data environment.

Value

The estimated value of this contract is $150,000. This encompasses all professional services provided by Craig Frint, including architectural design, model training, validation, and post-deployment monitoring for the initial three months. The investment aims to reduce manual administrative overhead by an estimated 30% and improve the identification of billable clinical events. Payments are structured around milestone completions to ensure transparency and project alignment.

Scope

The scope of work includes the development of a Python-based NLP pipeline tailored for medical terminology. Key activities involve:

1. Requirements gathering with HCA medical staff to identify critical data entities.
2. Development of a data anonymization layer to ensure HIPAA compliance.
3. Training and fine-tuning a Large Language Model (LLM) or transformer-based

architecture on de-identified clinical notes.
4. Integration of the AI engine with HCA's secure cloud staging environment.
5. Creation of a RESTful API to allow internal applications to query the engine.
6. Comprehensive documentation and training for HCA's internal IT team.
   The project specifically excludes the development of a front-end user interface, as the engine will serve as a backend utility for existing HCA applications.

Type

This is a work-for-hire agreement as described by U.S. California law. All intellectual property, including custom code, trained model weights, and architectural diagrams developed by Craig Frint during the performance of this contract, will be the sole property of HCA Healthcare. Craig Frint will operate as an independent contractor.

Payment

The total budget of $150,000 will be distributed according to the following schedule:

- **Initial Deposit (10%):** $15,000 upon contract signing.
- **Milestone 1: Prototype/Alpha (30%):** $45,000 upon successful demonstration of entity extraction on a sample dataset.
- **Milestone 2: Integration/Beta (30%):** $45,000 upon successful deployment to the HCA staging environment.
- **Final Delivery (30%):** $45,000 upon final sign-off, delivery of documentation, and completion of the first 30 days of production monitoring.
  Payments shall be made via ACH transfer within 15 days of invoice submission.

Purpose

Objectives

- **Objective 1:** Automate the extraction of ICD-10 codes from narrative doctor notes.
- **Objective 2:** Ensure 100% HIPAA compliance in data handling through robust de-identification.
- **Objective 3:** Provide a scalable API capable of processing 10,000 notes per hour.
- **Key Result:** Achieve an F1-score of ≥ 0.92 on named entity recognition (NER) tasks for primary diagnoses.

Performance

The success of the AI engine will be measured by the following metrics:

- **Accuracy:** The engine must achieve a precision rate of at least 95% in identifying core clinical concepts.
- **Latency:** API response time must be ≤ 500 ms for a standard 500-word clinical note.
- **Uptime:** The engine must maintain 99.9% availability during the testing and production

phase.
- **Scalability:** The system must handle concurrent requests from at least five different internal HCA departments without degradation.

Factors

- **Personnel:** Access to HCA Subject Matter Experts (SMEs) and medical coders for data labeling and validation.
- **Compute Resources:** Provision of GPU-enabled cloud instances by HCA for model training.
- **Data Quality:** Availability of high-quality, diverse clinical notes for training.
- **Security:** Adherence to HCA's strict internal cybersecurity protocols.

Who does what

People

- **Craig Frint:** Lead Software Engineer (Contractor).
- **Dr. Sarah Jenkins:** HCA Chief Medical Information Officer (Stakeholder).
- **Mark Thompson:** HCA IT Project Manager (Work Authority).
- **HCA Data Engineering Team:** Support for data pipeline access.

Roles

- **Contractor:** Responsible for technical execution, model training, and API development.
- **Work Authority:** Responsible for approving milestones, providing access to resources, and handling invoicing.
- **Medical SME:** Provides clinical context and validates the accuracy of AI outputs.

Responsibilities

A RACIO matrix will be maintained by Mark Thompson.

- **Craig Frint:** Responsible for Development; Accountable for Technical Delivery.
- **Mark Thompson:** Accountable for Project Governance.
- **HCA Security:** Consulted on data privacy and HIPAA protocols.

Context

Past

HCA Healthcare previously relied on manual review of physician notes, which was time-consuming and prone to human error. A pilot program using off-the-shelf NLP tools showed promise but lacked the specificity required for HCA's specialized oncology and cardiology notes.

Present

Currently, HCA is modernizing its data infrastructure. The "AI Clinical Insights Engine" is a cornerstone of this modernization, intended to bridge the gap between unstructured text and structured data analytics.

Future

The long-term roadmap includes expanding this engine to analyze patient-submitted data and integrating real-time alerts for physicians based on note analysis.

Planning

Requirements

Craig Frint must deliver:

- A Python source code repository (Git).
- Trained model weights and configuration files.
- A Dockerized API deployment package.
- Performance validation reports using LaTeX-formatted charts and tables.

Specifications

The software must be written in Python 3.10+, utilizing frameworks such as PyTorch or Hugging Face. The API must adhere to OpenAPI 3.0 standards.

Work breakdown structure (WBS)

1. **Phase 1:** Environment Setup and Data Anonymization (Weeks 1-4).
2. **Phase 2:** Model Selection and Initial Training (Weeks 5-12).
3. **Phase 3:** API Development and Staging Integration (Weeks 13-18).
4. **Phase 4:** Validation, Optimization, and Documentation (Weeks 19-24).

Applicable standards

- **HIPAA:** Health Insurance Portability and Accountability Act.
- **HL7 FHIR:** For data exchange standards where applicable.
- **PEP 8:** For Python code style.

Technical, operational, and organizational environment

Work will be performed using HCA's Azure-based cloud environment. Craig Frint will be provided with VPN access and a virtual workstation.

Method and source of acceptance

Acceptance is based on the "Validation Report." Craig Frint will submit a report showing the model meets the ≥ 0.92 F1-score threshold on a "hold-out" test set provided by HCA.

Reporting requirements

A bi-weekly status meeting will be held every Tuesday at 10:00 AM CST via Microsoft Teams. A written progress report must be submitted 24 hours prior to the meeting.

Other terms and conditions

Authorities

Work Authority: Mark Thompson, IT Project Manager. All technical directions must originate from or be approved by the Work Authority.

Client's obligations

HCA will provide access to the internal "Meditech" EHR staging data and necessary GPU computing credits within 48 hours of contract commencement.

Contractor's obligations

Craig Frint shall maintain strict confidentiality regarding all patient data and proprietary HCA algorithms. All work must be conducted within the HCA-provided cloud environment.

Location of work, work site, and delivery point

The work will be performed remotely. Occasional on-site meetings at HCA Healthcare Headquarters in Nashville, TN, may be required with 7 days' notice.

Language of work

All code, documentation, and communication shall be in English.

Security requirements

Craig Frint must undergo and pass a Level 2 Background Check and a HIPAA Security Training module provided by HCA before access to clinical data is granted.

Schedule

Expected start date and completion date

- **Commencement:** March 1, 2024.
- **Expected Completion:** August 31, 2024.

Wordbook

- **EHR:** Electronic Health Record.
- **NLP:** Natural Language Processing.

- **NER:** Named Entity Recognition.
- **HCA:** Hospital Corporation of America.

Sign-off

NOTE: Before signing the Statement of Work, if you have any questions or concerns, please call the Work Authority indicated above to negotiate any issues.

If you agree to the requirements of this Statement of Work, please sign and date the document which will be accepted as your proposal by Client, and return to my attention.

Please return an original signature copy by mail.

HCA Healthcare (Client)

Printed Name: _____

Signature: _____

Date: _____

Craig Frint (Contractor)

Printed Name: _____

Signature: _____

Date: _____