

WORK BREAKDOWN STRUCTURE

UC Name	A reusable NLP pipeline for clinical narratives preprocessing and characterization. (OPHDST/I&D)
PROJECT MANAGER	Oscar Rincon Guevara
DATE	5/29/2024

WBS NUMBER	TASK TITLE	PROJECT TEAM	TASK OWNER	START DATE	DUE DATE	DURATION	% of TASK COMPLETE
1	Planning						
1.1	Create WBS	AIX Team	Jaspreet Bains	6/1/2024	6/10/2024		0%
1.2	Get WBS approved	AIX Team	AIX Team	6/10/2024	6/11/2024		0%
2	Project Scope and Requirements						
2.1	Identify Project objectives (including value statements) and deliverables.	Project Team / AIX Team	OG	6/11/2024			0%
2.2	Document Project Success Criteria(s) / Acceptance Criteria(s) Note- this should include metrics which would prove its efficacy compared to all-purpose pipeline that cleans XML, HTML, and RTF markup from clinical narratives that are later passed through a NER task to identify symptoms, medications, and diseases.	Project Team / AIX Team	OG/KB	6/11/2024			0%
2.3	Gather requirements & Dependencies / Constraints (High Level)- including BERT, EDAV Databricks configuration, UC Model tracking, performance requirements, security requirements, etc..	Project Lead / AIX Team / EDAV	OG/KB	6/11/2024			0%
2.4	Data Requirements: Data collection cadence	Project Team / AIX Team	OG/KB	6/11/2024			0%
2.5	End User data consumption requirements/patterns (including any dashboards/reporting requirements)	Project Team / AIX Team	OG				0%
2.6	End user/data consumers identification	Project Team / AIX Team	OG				0%
2.7	Requirements and Scope Review and Signoff	AIX Team					0%
3	Analysis & Design Phase						
3.1	AIX Team Onboarding (Repo access, platform access etc)		CDH Support/KB				
3.2	Develop high-level system architecture	Project Team	KB	6/11/2024			0%
3.2.1	Define System Requirements (INCLUDING data scientist environment needs related to sandbox, etc..)	Project Team	KB				0%
3.2.2	Determine Compliance and Regulatory Needs (ATO)	Project/AIX Team					0%
3.2.3	Define Interactions Between Components (NER from BERT, etc... and Risks)	Project Team	OG				0%
3.2.4	Identify Security Requirements	Project Team /AIX Team / Security (CISP)	CDH Engineering/KB				0%
3.3	Develop Data Architecture	Project Team	KB				0%

3.3.1	Identify data source(s)	Project Team	OG/KB				0%
3.3.2	Data validation/profiling for minimum required data	Project Team	KB/OG				0%
3.3.3	Data Cleaning and preprocessing: how to handle missing data, remove outliers, normalize data, etc..	Project Team	OG				0%
3.3.4	Exploratory data analysis to detect data anomalies, hypothesis testing, etc..	Project Team	OG				0%
3.4	Data Ingestion pipeline design	Project/AIX Team	KB				0%
3.5	Data storage design	Project/AIX Team / EDAV Team	CDH Engineering/KB				0%
3.6	Model Design	Project/AIX Team	OG				0%
3.6.1	Algorithmic selection	Project/AIX Team	OG				0%
3.6.2	Model Architecture	Project/AIX Team	OG/KB				0%
3.6.3	Evaluation metrics to determine model performance	Project/AIX Team	OG				0%
3.7	Document Assumptions, Constraints, and Decisions	Project/AIX Team					0%
3.8	Define Performance Metrics and Monitoring	Project/AIX Team					0%
3.9	Design Review and Signoff	AIX Team					0%
4	Development and Testing Phase						
4.1	Setup data ingestion pipelines	Project Team	KB/OG				0%
4.2	Implement data storage and management	EDAV Team / AIX Team					0%
4.2	Model Development	Project Team	OG				0%
4.2.1	Feature selection and Engineering	Project Team	OG/KB				0%
4.2.2	Model Selection and training	Project Team	OG				0%
4.2.3	Model Validation	Project/AIX Team	OG				0%
4.2.4	Collect and compare Evaluation metrics	Project/AIX Team	OG				0%
4.2.5	Model Optimization	Project Team	OG				0%
4.3	Code Refactor (if applicable)		OG				0%
4.4	Develop User Interface (UI)	Project Team	OG/KB				0%
4.5	Configuration updates for EDAV Services	EDAV Team					0%
4.6	Review & Incorporate standard practices	Project Team / EDAV Team / AIX Team					0%
4.7	Integration Testing	Project Team	OG/KB				0%
4.8	System Testing	Project Team	OG/KB				0%
4.9	Code review and dev signoff	Project/AIX Team					0%
4.10	Testing Signoff	Project/AIX Team					0%

5	Production Deployment						
5.1	Security Approval	Security (CISP)	CDH Data Management				0%
5.1.1	Security Review	Security (CISP)					0%
5.1.2	Security Audit	Security (CISP)					0%
5.1.3	Security Signoff	Security (CISP)					0%
5.2	Prepare and deploy model to prod environment	Project Team					0%
5.3	Production Dry Run (E2E)	Project/AIX Team					0%
5.4	Production RunBook for Ops handover	Project/AIX Team					0%
5.5	Monitoring (including Model performance)	Project Team					0%
5.6	Maintenance and post prod support	EDAV Team / Project Team					0%
5.7	Maintain System Performance	EDAV Team					0%
5.8	Publish & Distribute Required Communication	Project Team					0%
5.9							0%
5.10							0%
5.11							0%