

**STATEMENT OF WORK
TO ENABLE A
CLINICAL REASONING AND PREDICTION SYSTEM ASSESSMENT
RFI #D13PS00141**

1. *Background*

The terms “clinical reasoning” and “prediction” are over-loaded and potentially mis-understood. In the context of this Statement of Work (SOW), “clinical reasoning” refers to that emerging class of systems able to operate on structured and unstructured data that are developing capabilities to perceive and adapt to their environment. It explicitly does not refer to traditional rule-based systems, though rules may be a component of the system. Similarly, “prediction” systems, used here extends beyond commonly deployed statistical techniques from modeling and data mining and is meant to incorporate advanced natural language processing techniques and machine learning to detect weak and strong patterns that will lead to important risks and opportunities.

For several years the Veterans Health Administration (VHA) Office of Informatics and Analytics (OIA) has been monitoring the maturation and breakthrough successes of reasoning and prediction class systems, in the context of an overwhelming growth of the medical literature corpus. Concurrently, VHA is accumulating significant amounts of clinical data. For example, associated with the roughly 30 million U.S. Department of Veteran Affairs (VA) electronic health records (EHRs) are 3.2 billion clinical orders, 1.8 billion medication prescriptions and 2.3 billion vital sign measurements. This structured data is accompanied by 2 billion clinical text notes with a growth rate of one hundred thousand additional notes per day. The technology advances, relentless increases in medical literature, and the staggering volume of clinical data we collect have motivated VHA to perform an in-depth, first-hand assessment within our environment of the value clinical reasoning and prediction systems may offer.

2. *Primary Assessment Objective*

VHA would like to answer the question: How can leading clinical reasoning and prediction systems help physicians improve patient care, efficiency and outcomes within the VHA clinical environment? We want to determine the value and fitness of this emerging class of systems, including technical acceptability, functional utility, and usability. This work effort has three supporting objectives:

- 2.1. Successfully install and configure, in safe harbor environment, a clinical reasoning and prediction demonstration system(s) at VA’s Austin Information Technology Center (AITC) in Austin, Texas.
- 2.2. In partnership with VHA, technically support the identification of feasible use cases, training the system and ingestion of the corpus of related artifacts.
- 2.3. In partnership with VHA, technically support the assessment potential of the system to improve health outcomes of Veterans.

3. Assessment Areas of Interest

VHA has identified three broad areas of interest they intend to assess: technical acceptability, functional utility, and usability.

3.1. Technical Acceptability Assessment

It is essential for the system to accurately, reliably and swiftly process large amounts of structured, semi-structured and unstructured data. The technical effectiveness with which the system can obtain optimal answers to clinical questions is essential. VHA seeks to assess the system's:

3.1.1. ability to:

- decompose and understand natural language questions from physicians
- perform semantic search in reasoning
- query structured, semi-structured, and unstructured text
- return relevant information (precision and recall)
- present supporting evidence for answers
- detect previously unknown trends
- identify patterns
- develop and use predictive models

3.1.2. initial implementation demands of:

- installing assessment system hardware
- interconnecting to local corpora
- configuring software to meet institutional needs

3.1.3. corpora requirements for:

- domain training: annotation and vetting
- target domain ingestion: acquisition and pre-processing effort
- adding additional domains

3.1.4. infrastructure scale out needs for enterprise deployment, including:

- concurrent processes/nodes
- RAM capacity
- disc storage capacity
- physical footprint
- network bandwidth
- software licensing
- hardware/software maintenance

3.2. Functional Utility Assessment:

Disruptive technologies by definition impact the environments they are introduced to. It is important for VHA to assess what functionally is expected, needed and valued by our physicians. VHA seeks to assess the system's capabilities of:

3.2.1. patient medical management :

- patient problem list management, including automated generation of tentative problem list, identification of new problems and characterization of problem activity
- current standard of care, problem oriented data identification, problem oriented order sets and outcome evaluation
- potential medical error detection, including missed data and adverse events
- standards of care based upon medical literature and clinical practice guidelines

- 3.2.2. diagnostic assistance effectiveness:
 - EHR anomaly detection
 - syndrome identification
 - development of differential diagnoses
 - characterization of prognosis and factors conferring risk of adverse outcomes
 - cognitive error avoidance
- 3.2.3. semantic search across:
 - clinical notes
 - lists including problem, medication, appointment
 - structured and unstructured data results measures
 - patient reported data
- 3.2.4. question answering:
 - context relevance including patient characteristics, problem, location and user
 - interactivity
 - user profiling
- 3.2.5. annotation:
 - human supervision and checking of machine responses or results to check for accuracy
 - store ratings and assessments of machine determinations
 - refine and teach based on corrected findings
- 3.2.6. predictive risk generation in all medical management areas including:
 - general models to evaluate quality of care, increase efficiency and improve outcomes
 - patient re-admission probability
 - drug adverse events (not U.S. Food and Drug Administration (FDA) reported)
 - comparative studies in treatment efficacies
 - new practice guideline detection
 - cohort risk identification
 - general patient risk modeling

3.3. Usability Assessment:

It is important for VHA to understand the user experience and the degree to which the user's wants and needs can be achieved when accessing available functional capabilities. Closely related to this is assessing how well this technology fits into existing or modified clinical processes and workflow. VHA seeks to assess:

- 3.3.1. ease of system use:
 - learnability
 - discoverability
- 3.3.2. task performance:
 - efficiency
 - quickness
 - completeness
- 3.3.3. error risk:
 - susceptibility
 - severity
 - recovery

- 3.3.4. access:
 - workflow (transparency, integration into environment)
 - performance under load (throughput)
- 3.3.5. satisfaction/likability

4. Constraints

- 4.1. The Contractor shall provide a response based on a working product version and published design that best meets the assessment areas of interest. Conceptual or theoretical designs are not of interest.
- 4.2. The Contractor is responsible for provision of hardware, software and services necessary to administer, operate and sustain, as a service, an assessment system.
- 4.3. The duration of this work is expected to last approximately 24 months.
- 4.4. In the run-time assessment phase, the assessment system must be capable of supporting:
 - Up to 100 authorized VA users.
 - 100 complex “patient case” results per hour with 3 minute response time per case –and, non-concurrently- 600 “simple” semantic searches per hour with 1 minute response time.

5. Tasks, Deliverables and Performance Goals

The work segments identified below follow a loose progression but are not intended to be strictly sequential. It is expected that many segments will occur simultaneously with the recognition there is a critical path present (defined as Work Segments 2, 3, 4, and 8).

5.1. Work Segment 1 -Target Use Case Identification

- 5.1.1. Tasks include: Contractor will technically support VHA in the selection and design of target use cases to be exercised and studied. Work includes level of effort, complexity and risk estimates. The use cases will at a minimum include two predictive analytics models and three clinical reasoning domains. Possible examples include patient falls risk, readmission risk, missed clinical data, missed problems, Post-Traumatic Stress Disorder (PTSD) detection, and problem based medical management.
- 5.1.2. Deliverables include: descriptive material on two predictive analytics models and three clinical reasoning domains
- 5.1.3. Performance goals include: the ability to produce a ranked list of optimized merit-based target use-cases.

5.2. Work Segment 2 - Physical system implementation

- 5.2.1. Tasks include: Contractor will install at the VA’s AITC in Austin, TX all hardware, software and other direct infrastructure necessary to achieve a successful assessment. The contractor will provide all system engineering, operation and administration services required for the implementation.
- 5.2.2. Deliverables include: All hardware, software and other direct infrastructure necessary to achieve a successful assessment
- 5.2.3. Performance goals include a running system capable of supporting the users, features, capabilities and computational demands of the target assessment.

5.3. Work Segment 3 - System configuration

- 5.3.1. Tasks include: The Contractor will provide all system configuration services required to optimize the system configuration for execution of the VHA target use

cases. Services include corpora connectivity or other access necessary to receive content such as: EHR data, institutional clinical guidelines, data association pairs, and published articles.

5.3.2. Deliverables include: a detailed log of system parameters, settings, modules and the like used in this specific configuration.

5.3.3. Performance goals include: Provision all system configuration services required to optimize the system configuration for execution of the VHA target use cases.

5.4. Work segment 4 - System training

5.4.1. Tasks include: The Contractor will provide all system training services required to tune core algorithms and optimize candidate response strength for VHA use cases. The VHA will be responsible for providing access to subject matter experts.

5.4.2. Deliverables include: Delivery of training services

5.4.3. Performance goals include: Tuning core algorithms and optimizing candidate response strength so that the correct results are found within the top 3 candidates with 85% precision.

5.5. Work segment 5 - Corpora ingestion

5.5.1. Tasks include: The Contractor will provide services to pre-process the entire corpora available necessary to support the identified use cases. The Contractor will be required to perform an initial bulk-load corpora ingestion for each domain followed by daily sustainment ingestion such that the answer source is current to within 24 hours of new data added to the corpora.

5.5.2. Deliverables include: Ingestion logs and performance metrics.

5.5.3. Performance goals include: Provide adequate services to pre-process the entire corpora available necessary to support the identified use cases.

5.6. Work segment 6 - User Training

5.6.1. Tasks include: The contractor will provide on-line training materials to VHA physician users on system capabilities and use.

5.6.2. Deliverables include: on-line guides, FAQs, webinars

5.6.3. Performance goals include: Student satisfaction scores of 85% or greater.

5.7. Work segment 7 - Predictive modeling service

5.7.1. Tasks include: The contractor will provide a fully operational run-time prediction modeling system for 90 calendar days as a service for VHA's use in exercising use case models to predict probability of an outcome.

5.7.2. Deliverables include: Provide a fully operational run-time prediction modeling system for 90 calendar days.

5.7.3. Performance goals include: Access to fully operational service for 90 calendar days.

5.8. Work segment 8 - Clinical reasoning service

5.8.1. Tasks include: The contractor will provide a fully operational run-time clinical reasoning system for 90 calendar days as a service for VHA's use in exercising use case models and posing complex "patient cases" and "simple" queries.

5.8.2. Deliverables include: Provide a fully operational run-time clinical reasoning system for 90 calendar days.

5.8.3. Performance goals include: Access to fully operational system for 90 calendar days.

5.9. Work segment 9 - Metrics collection

- 5.9.1. Tasks include: The contractor will work with VHA to establish a series of meaningful assessment metrics for each work segment. The contractor will collect and report these metrics electronically to VHA during each active work segment.
- 5.9.2. Deliverables include: Development and delivery of meaningful assessment metrics for each work segment.
- 5.9.3. Performance goals include: Development of meaningful assessment metrics.

5.10. Work segment 10 - Project management

- 5.10.1. Tasks include: The Contractor shall conduct weekly status calls with the VHA Project Manager (VHA PM) to provide immediate notification of any issues that have potential to affect the project schedule or outcomes. The contractor will produce bi-weekly electronic reports documenting progress, issues and risks.
- 5.10.2. Deliverables include: electronic reports documenting progress, issues, and risks.
- 5.10.3. Performance goals include: Successful scope and schedule management.

5.11. Optional Work segment 11 – Evidence profile design and development

- 5.11.1. Tasks include: The contractor will partner with VHA to design and develop a prototype evidence profile consisting of multiple dimensions of clinical evidence for a previously developed use case. (renewable 1x)
- 5.11.2. Deliverables include: Prototype evidence profile.
- 5.11.3. Performance goals include: Successful utility and usability assessment by VA physicians.

5.12. Optional Work segment 12 – Additional clinical reasoning use case

- 5.12.1. Tasks include: Contractor will technically support VHA in the selection and design of an additional use case to be studied. (renewable 2x)
- 5.12.2. Deliverables include: descriptive material on a new clinical reasoning domain
- 5.12.3. Performance goals include: Sufficient detail to effectively execute a subsequent assessment.

5.13. Optional Work segment 13 – Clinical reasoning service extension

- 5.13.1. Tasks include: The contractor will provide a fully operational run-time clinical reasoning service for 90 calendar days for VHA's use in exercising exercise use case models and posing complex "patient cases" and "simple" queries. (renewable 3x)
- 5.13.2. Deliverables include: Fully operational run-time clinical reasoning service.
- 5.13.3. Performance goals include: Access to fully operational system for 90 calendar days.

5.14. Optional Work segment 14 – Additional predictive modeling use case

- 5.14.1. Tasks include: Contractor will technically support VHA in the selection design and development of an additional use case model to be studied. Work includes estimating the level of effort, complexity and risk estimates. (renewable 3x)
- 5.14.2. Deliverables include: descriptive material on a new predictive modeling use case
- 5.14.3. Performance goals include: Sufficient detail to effectively execute a subsequent assessment.

5.15. Optional Work segment 15 – Predictive modeling service extension

- 5.15.1. Tasks include: The contractor will provide a fully operational run-time prediction modeling service for 90 days calendar days for VHA’s use in exercising use case models to predict probability of an outcome. (renewable 3x)
- 5.15.2. Deliverables include: Fully operational run-time prediction modeling service.
- 5.15.3. Performance goals include: Access to fully operational system for 90 calendar days.

6. Schedule

6.1. The total duration for this requirement is not expected to exceed 24 months. However, the specific milestones for the Work Segments have yet to be determined.

+++++

Note: pre-award RFP requirement: 4 hour in-person “show me” capability demonstration of any domain (doesn’t have to be health) that can operate at a scale equivalent to that found within the VHA clinical environment.

