STATEMENT OF WORK - 04/26/2023

PROPOSED START DATE 10/01/2023

Site 1:	University of Chicago 5801 S. Ellis Ave. Chicago, IL 60637 PI: Ishanu Chattopadhyay	Site 2:	University of Iowa 51 Newton Road Iowa City, IA 52242 Site PI: Balaji Manicassamy

Specific Aim 1: Formulate sequence similarity metric E-distance	**Timeline (Months)**	**Site 1**	**Site 2**
Major Task 1.1: Emergenet Development			
Subtask T1.1.1: Precisely formulate the Emergenet inference platform, that puts together ML algorithms capturing maximally predictive patterns of change and mutational dependencies	1-3	✓	
Subtask T1.1.2: Provide uncertainty quantification for the inferred patterns represented in the Emergenet models	2-6	✓	
Milestones Achieved: M1) Emergenet software beta release M2) Uncertainty quantification of inferred models			
Major Task 1.2: Sample-complexity for Emergenet inference			
Milestones Achieved: M3) Sample-complexity estimates complete			
Major Task 1.3: Event Timeline Estimation			
Subtask T1.3.1: Map mutational change dynamics to "wall-time" to forecast *when* future variants will show up	3-9	✓	
Subtask T1.3.2: Computationally validate timeline predictions using records of past emergence events	9-12	✓	
Milestones Achieved: M4) Software update for timeline estimation released			

Specific Aim 2: Validate E-distance as a similarity metric on strain space identifying biologically valid sequence variations	**Timeline (Months)**	**Site 1**	**Site 2**
Major Task 2.1: Quantify asymmetric transition probabilities between strains			
Subtask 2.1.1 Develop analytical framework to identify probabilistic movement direction between strains	9-12	✓	
Subtask 2.1.2 Develop analytics to chart multi-step probabilistic trajectories from observed strains	12-15	✓	
Milestones Achieved: M5) Software update for future trajectory calculation			
Major Task 2: Assessment of fitness of potential emerging zoonotic IAV variants in cell cultures			
Subtask 2.2.1 Shortlist HA variants with maximal emergence probability of H3N2 and H1N1 subtypes	6-9	✓	✓
Subtask 2.2.2 Generate predicted HA variants using reverse genetics in human lung epithelial cell line (A549) and primary human lung cells	9-22		✓
Subtask 2.2.3 Evaluate generated variants for replicative fitness	15-24		✓
Milestones Achieved: M6) Emergenet experimental validation complete			

Specific Aim 3: Develop a working implementation of BioNORAD	Timeline (Months)	Site 1	Site 2
Major Task 3.1 IRAT score replication			
Subtask T3.1.1 Investigate how each of the ten dimensions of IRAT comparison map to our Emergenet based risk	12-18	✓	
Subtask T3.1.2 Evaluate sensitivity of IRAT scores to risk-assessment timepoint	15-18	✓	
Subtask T3.1.3 Incorporate event timeline estimation in BioNORAD prototype to predict time to emergence	18-24	✓	
Milestones Achieved: M7) IRAT score-based validation of BioNORAD prototype			
Major Task 2: Results for Current/Recent Surveillance Data			
Subtask T3.2.1 Demonstrate we can analyze collected sequences at scale, by enumerating the risk profiles of all sequences collected recently within teh last few years	18-24	✓	
Subtask T3.2.2 Set up an automated pipeline that pulls in sequence data of for new submissions, and publish a risk score automatically	18-24	✓	
Milestones Achieved: M8) Working version of BioNORAD platform demonstrated M9) Final project report submitted			

DELIVERABLES

☐ Emergenet software

☐ BioNORAD implementation

☐ Experimental protocols for fitness assessment of Emergenet predicted strains

TIMELINE

The project timeline, indicating various milestones and tasks is illustrated in Fig. 4

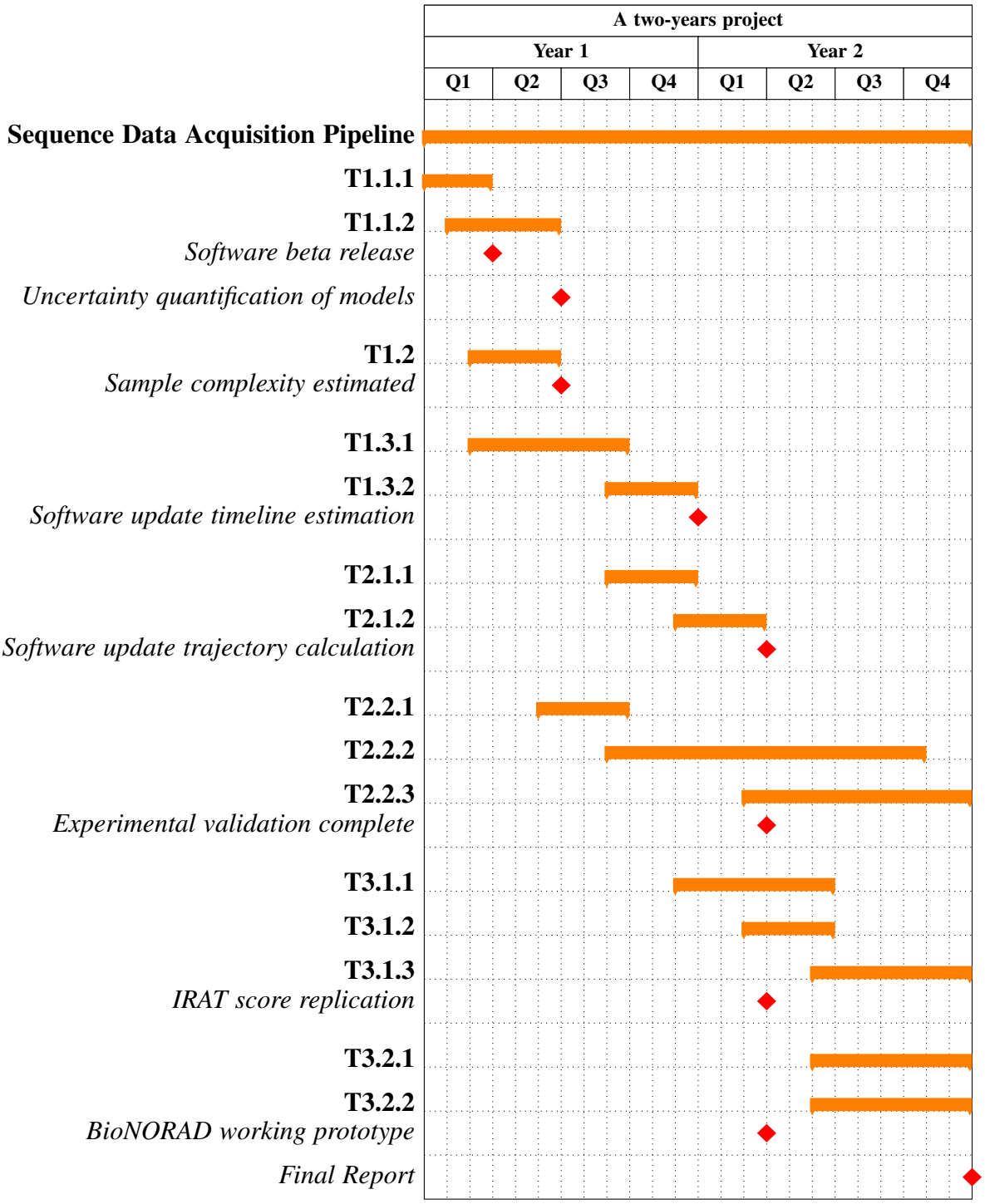

Fig. 4. Gantt chart of project timeline