HAYSTAC Program

TABLE OF CONTENTS

1.A Program Overview

How humans move has evolved significantly in recent history. With the 20th century advent of the automobile, the train, the commercial airliner, and increasing disposable income, mobility across the world became easier, more efficient, and more prolific. Human movement shifted from hyperlocal property-centric movement to movement across a wide range of activities that include easy transit across cities, rural to urban commutes, and flights between countries. Despite this increase in movement, there was not a similar increase in the data about these movements. It wasn't until the 21st century's explosion in connected Internet-of-Things (IoT) devices and smart-city infrastructures, and particularly the last 10 years, that we see data about actual human movement, and the individual trajectories[1] people create in their daily lives, come into focus. This explosion of recorded movement and generation of human trajectories is likely to continue escalating, offering the opportunity to build new models that understand human dynamics at unprecedented resolution, as well as creating the responsibility to understand the expectation of privacy for those moving through a sensor-rich world.

High fidelity, fine-grained, modeling of human mobility is of critical interest to the Intelligence Community (IC), in particular establishing models of "normal" movement capable of encoding the diversity of human movement present across times, locations, and people. Additionally, how human movement changes due to environmental (e.g., weather, construction) and human (e.g., large gatherings, terrorist attacks) driven events is of interest to the IC both before and after such events. However, today's modeling capabilities can only provide high-level insight into human movement. These techniques can encode sufficient dynamics to study population migration, disease spread, or other highly aggregated properties, but they do not capture the fine-grained activities of human life and transportation logistics that drive daily trajectories of movement. Consequently, identifying what specific types of individual or coordinated human movements are detectable as anomalous within a full population of movement remains an open research question.

The IARPA's Hidden ActivitY Signal and Trajectory Anomaly Characterization (HAYSTAC) program aims to change this by developing new methods capable of encoding human movement at the fine-grained level. By capturing the subtleties of human movements over a broad set of models and rigorously characterizing the level at which some trajectories "stick out from the crowd", HAYSTAC will establish the scientific foundation connecting data, movement, and the expectation of privacy.

The key limitation in achieving this goal is the lack of ground-truthed movement datasets. Without full knowledge of the underlying activities, it is impossible to characterize what movement is practically detectable as anomalous, what movement fits within normal expectation, and what volume and veracity of trajectory data supports those conclusions. HAYSTAC teams will address this by 1) creating a large-scale microsimulation of background activity and associated trajectories 2) inserting specific movement activity into the simulation and 3) attempting to separate inserted activity from the background activity.

By executing this sequence as recurring "trials" in which multiple teams compete to identify what trajectories are generated from activity that was inserted by other teams, HAYSTAC systems will attempt to learn the fine-grained activities of normal movement in order to *both* minimize the separability between normal background activity and their own inserted activity and maximize the separability of another team's inserted activities from the background activity. Robust to data quality and diverse movement activity tested through these trials, successful HAYSTAC systems will be capable of identifying trajectories created by subtle deviations from normal activity and be able to generate activity and corresponding trajectories which are not distinguishable from normal.

[1] Trajectories are defined as multiple temporal measurements of human activity, producing data of the form: identification (ID), latitude, longitude, time.

1.A.1 Technical Challenges and Objectives

The HAYSTAC Program consists of two Technical Areas (TA). TA-1 focuses on the microsimulation of human activity to create robust simulated trajectory datasets that serve as the background activity and trajectories for the trial structure. TA-2 efforts learn normal movement directly from trajectory observations and create automated systems that can both generate normal activity and identify anomalous trajectories. The integration of these activities is represented with the graphic in Figure 1.

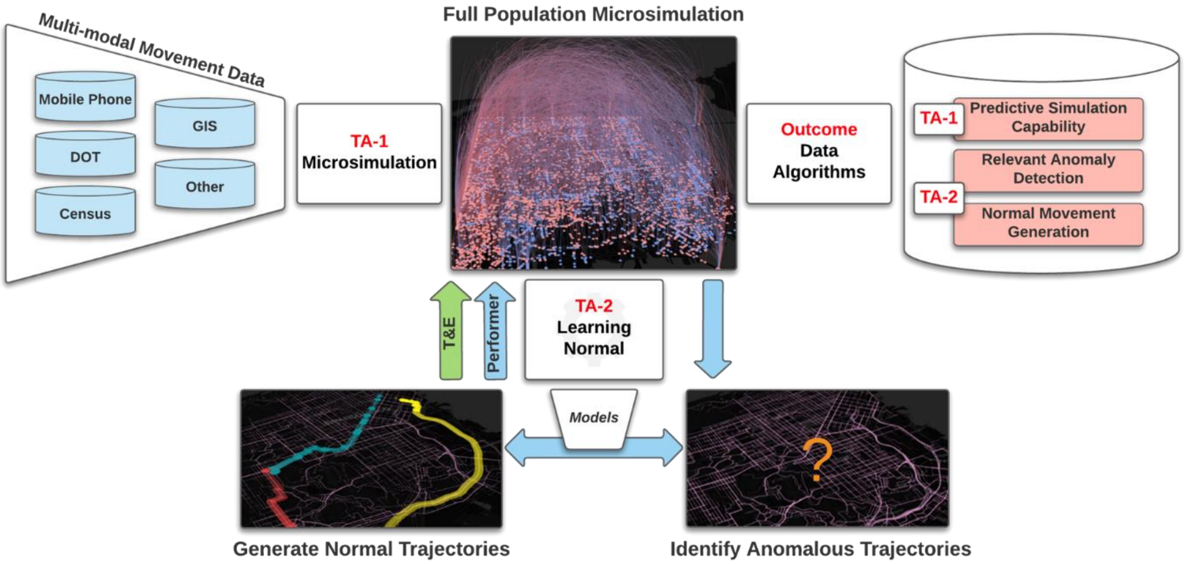

Figure 1: HAYSTAC Program Structure

<u>Offerors may propose to TA-1 only, TA-2 only, or both TA-1 and TA-2.</u> If a proposal is selected for both TA-1 and TA-2, that team's TA-2 will leverage TA-1 simulations developed by a different team.

1.A.1.1 TA-1: Microsimulation

HAYSTAC TA-1 teams will pursue the next-generation of microsimulation capabilities, incorporating realistic human behaviors that encompass the subtleties of fine-grained human activity and travel. These simulation engines will attempt to leverage available data, mobility models, and/or social theory to create trajectory datasets that, while simulated, successfully replicate the real-world dynamics, relationships, and emergent properties visible in human movement. To meet the HAYSTAC goals, these simulations must not only be behaviorally high-fidelity but also overcome key technical challenges:

- Scalability to the populations of mega-cities and surrounding areas[2]
- Adaptability to diverse geographic regions
- Tunability and insight into the behaviors of simulated entities
- Flexibility to handle external inputs (e.g. specific activities to insert for a HAYSTAC trial)

The simulation fidelity will improve through program phases in the components outlined in Table 1. This list is not exhaustive and will be refined by Program Kickoff.

[2] A mega-city is defined as a city with over 10M inhabitants. HAYSTAC aims to simulate urban, suburban, and rural environments with up to 30M total entities.

Table 1: Initial list of TA-1 simulation components

Component	Description	Goal
Number of entities	Quantity of individuals represented in the simulation	All entities in a mega-city and surrounding areas (up to 30M)
Duration	Length of the simulation	One year
Modes of Transportation	The method that enables human movement	Personal or work vehicle, bicycle, walking, ride-sharing, mass transit
Traffic Dynamics	Interactions between vehicles and between vehicles and dynamic environments	Accurate control and dynamic routing
Rare Events	Normal, but infrequent, events such as road blocks, weather disruptions, and large events	Variations from rare events mimic real world data
Realistic Movement	Habits, needs, goals, and social connections create relevant and coherent trips	Entities act with that mimic real world data
Realistic Geography	Relative importance, capacity, flow, and time of day dependencies for various locations	Locations interact with entities in a way that mimics real world data

In addition to the incorporation of realistic fine-grained activity into TA-1 microsimulations, the teams will also pursue capabilities to add realistic noise characteristics to otherwise perfect simulated data that will emulate a variety of sensors in the production of trajectories from the underlying activity. These components should incorporate geospatial error, temporal subsampling, population subsampling, ID dropouts, and ID confusion at a minimum. These degradation modules will be applied to the trajectory data output by the base TA-1 simulation and allow for parameterization and application by the Government.

Offerors should propose additional layers of microsimulation fidelity as necessary to meet HAYSTAC goals. This includes proposing additional metadata and attributes their simulation can include to enrich the overall output in addition to the fundamental components of trajectory data: time, location, and ID. Offerors are also encouraged to propose other forms of noise to incorporate into the data to allow HAYSTAC testing across many types of trajectories. These additional types of noise should be justified with their connection to realistic sensor phenomena. **Additional elements for simulation content or noise characteristics proposed with a strong technical justification will contribute to the assessment of the unique technical capabilities an Offeror possesses.**

TA-1 teams will be expected to design and deploy their own infrastructure for the computational scaling required to meet HAYSTAC program goals of creating large-scale simulated trajectory data. Additionally, teams will provide their own external data as needed in addition to data made available by the Test and Evaluation (T&E) team, as outlined in Section 1.D.

1.A.1.2 TA-2: Learning Normal Movement

Significant information can be learned from human trajectory data, however the patterns of life that are modellable are complex. Movement is highly influenced temporally; time of day, day of week, or the month of the year all impact what defines normal. The modes of transportation and recurrence of various trips can vary significantly across geographic locations as well as subpopulations or demographics within a given location. Moreover, while humans are inherently habitual and many of their daily trajectories are the product of repeated activities such as traveling between home and work, those trips may contain many different stops in between.

The HAYSTAC TA-2 teams will build models capable of encoding these complex signals and dependencies. To accomplish this, TA-2 teams will attempt to learn normal movement from TA-1 simulations, leveraging the fundamental trajectory information of time, location, and id, but not the entire set of attributes a TA-1 simulation may produce. These teams will participate in both sides of the

HAYSTAC trials: recommending activity that attempts to mimic normal so that other teams cannot identify the trajectories as anomalous and automatically identifying trajectories that deviate from normal to be deemed anomalous. While teams will participate in both elements of the trials (although never looking for the same activity that they inserted), they can pursue different technical approaches to for each element.

A TA-2 trial will notionally operate as follows between two TA-2 teams (Team A and Team B):

1. A TA-1 simulation of a region is provided to Team A and Team B over a given period of time
2. Tasking is provided to create movement activity that meets a set of trial objectives within the next period of time
3. Team A's TA-2 system produces activity mimicking normal movement that accomplishes the trial objectives
4. Team B is provided with a new TA-1 simulation which includes the inserted activity and partial information about the trial objectives
5. Team B's TA-2 system identifies trajectories, or groups of trajectories, deemed anomalous and produces an alert

A notional example trial is depicted in Figure 2. These trials will change in complexity and specific objectives over the program, as outlined in Section 1.A.2.

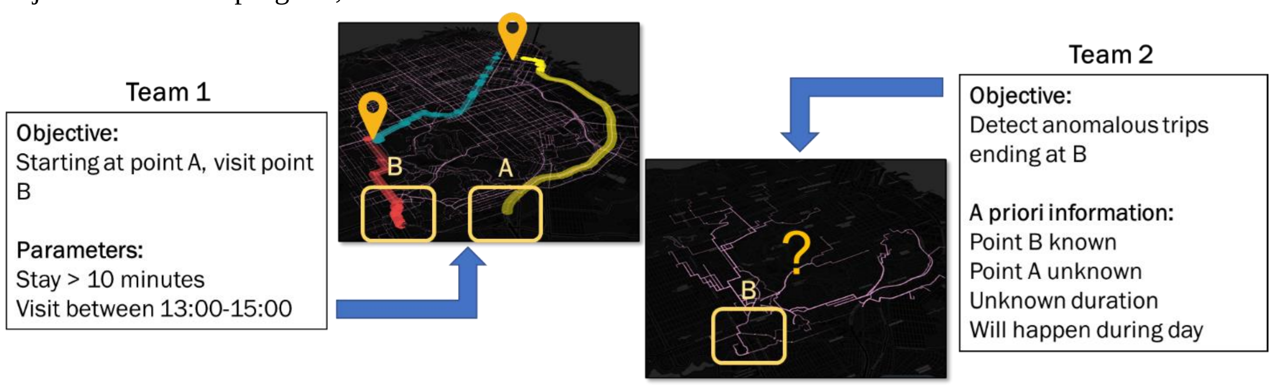

Figure 2: A notional TA-2 trial

The HAYSTAC program will pursue a large variety of these trials, executed through iterative "trial periods" throughout the program, to create a large corpus of anomalous, but ground-truthed, activity. Further examples of potential objectives that may be incorporated are shown in Table 2 and include single entity movements, recurring behaviors from an entity, or multi-entity movements in which movement is coordinated across a social network. The specific objectives will be refined before each trial period, with a plan provided in advance of the trial period start.

Table 2: Initial List of TA-2 Trial Objectives

Movement Type	Objective		
Single	Visiting a friend	Going to a Restaurant	Traversing the city
Recurring	Job	Shopping	Hobbies
Coordinated	Multiple entities executing any of the above		

Trials will be parameterized to create a large set of data and will identify which entities are participants, locations, durations, and the amount of a priori information provided to aid the team searching for anomalous activity. It is expected that the variety of trials created will push teams to build trajectory modeling approaches that are robust to encode diverse activities. The final output of TA-2 will be systems capable of both separating normal from abnormal trajectories as well as generating normal trajectories based on prescribed conditions. This is depicted in Figure 3.

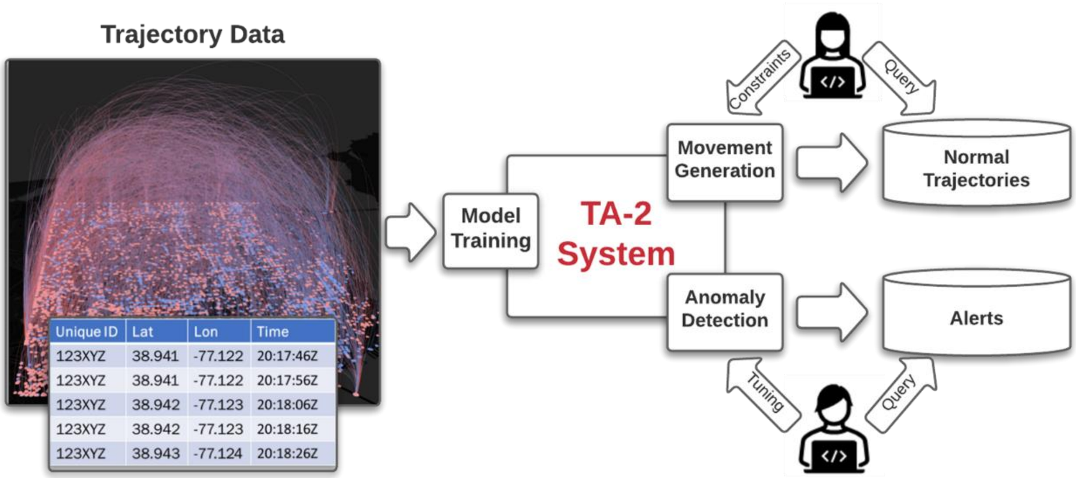

Figure 3: TA-2 System Design

1.A.2 Program Phases

The HAYSTAC program is anticipated to be a 4-year (48 month) effort, comprised of three Phases. All three Phases are being solicited under this BAA. Phase 1 will be 18 months in duration, Phase 2 will be 18 months, and Phase 3 will be 12 months. Each phase will encompass both Technical Areas described above.

Offerors must propose for all Phases. Any submission without all three Phases included will be considered non-compliant.

The high-level goals of each HAYSTAC phase are shown in Figure 4. In Phase 1, the goal is to establish the feasibility of automated characterization of normal movement. Simulations will be created and provided with limited noise and TA-2 trial objectives will be constrained to identifying single trajectories or single entities and significant information about the trial objective will be provided to the team searching for anomalies. In Phase 2, simulations will incorporate noise and sub-sampling and the goal is to build robust systems still able to identify anomalous trajectories with this noise and limited information about the trial objective. In this phase, coordinated activities across multiple actors will be leveraged. In the Phase 3, simulations will fully scale to simulate full mega-cities and their surrounding areas, incorporating up to 30M entities. These final simulations will be created using an "on demand" architecture with limited manual intervention in generating simulations at new locations. For TA-2, full populations will be searched for anomalies using little or no information about trial objectives in this phase.

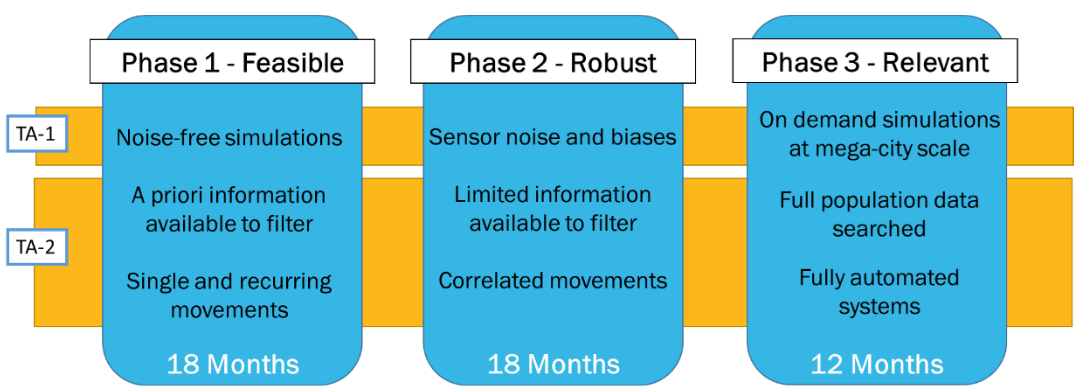

Figure 4: HAYSTAC Phase Goals

Each Phase will consist of multiple trial periods, each beginning with the delivery of a new TA-1 simulation to support the trial period. This schedule is provided in Figure 5 and presented in further detail in Section 1.G.

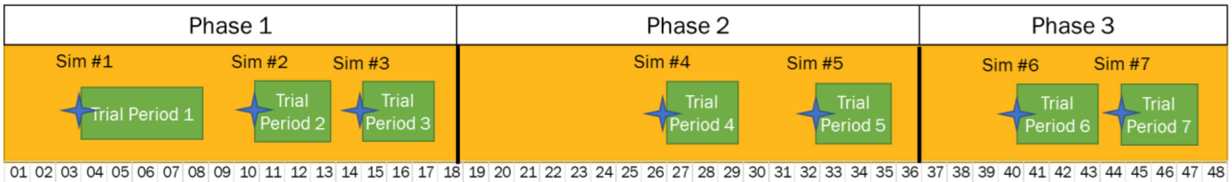

Figure 5: Trial Schedule

The complexity and scale of the TA-1 simulations and TA-2 trials will increase progressively throughout the HAYSTAC Phases. A summary of the TA-1 goals across the delivered simulations is presented as Table 3, and a summary of the TA-2 progression is presented as Table 4.

As described in Section 1.A.1.1, Performers are highly encouraged to propose additional components beyond those specified in Table 3. For those additional elements, the technical progression should be outlined per Simulation Deliveries 1-7, similar to Table 3. Additional elements with strong technical justification will contribute to the assessment of the unique technical capabilities an Offeror possesses.

Table 3: Progression of TA-1 Components

	Phase 1			Phase 2		Phase 3	
	Sim #1	**Sim #2**	**Sim #3**	**Sim #4**	**Sim #5**	**Sim #6**	**Sim #7**
Single Movement	x	x	x	x	x	x	x
Recurring Movement		x	x	x	x	x	x
Coordinated Movement				x	x	x	x
# of Entities	>10,000	200K	500K	2M	5M	10M	30M
Duration	2 weeks	2 mo.	2 mo.	2 mo.	4 mo.	6 mo.	12 mo.
Modes of Trans.	Personal Vehicle	Personal Vehicle	Bike, Walk	Mass trans, ride share		All	
Rare Events	No			Yes		Yes	
Regions	USA	USA	Foreign	Any		Any	
Sensor Noise	No			Yes		Yes	

Table 4: Progression of TA-2 Trials

Phase	Trial Period	Single Movement	Recurring Movement	Coordinated Movement	# of Objectives	Duration of training data	Trial Objective Insight
1	1	x			10	1 week	High
	2	x	x		30	1 months	Medium
	3	x	x		100	1 months	Medium
2	4	x	x	x	50	1 months	Medium
	5	x	x	x	100	2 months	Low
3	6	x	x	x	200	3 months	Low
	7	x	x	x	500	6 months	None

1.A.2.1 Phase 1

TA-1 teams will focus on delivering noise-free simulations in Phase 1, with increasing complexity and scale, resulting in a capability to create simulations with at least 500,000 entities in specified foreign locations by the end of the phase. There will be three trial periods that begin with a simulation deliverable.

Simulation #1 will showcase the Performer team's basic simulation and focus on formatting, standards compliance, and support of trial needs. It is expected that this simulation will be of the scale of at least 10,000 entities and incorporate personal vehicles (POV) only. The Offeror should propose a candidate location for this base simulation, and may leverage previous experience or existing data holdings to produce the simulation. Simulation #2 will be created at this same proposed location, but incorporate recurring behavior of entities, such as preferred shopping locations, restaurants, work, and recreational activities. A scale of 200,000 entities is expected with a duration of 2 months. Simulation #3 will change to a new location selected by the program, which will be provided at Program Kickoff. This location will be outside of the United States, scale up to 500,000 entities at a duration of 2 months, and incorporate additional types of human movement including biking and walking.

TA-2 teams will participate in the first three trial periods during Phase 1. Before each trial period commences, a detailed Trial Plan will be provided. During the first trial period, the trial objectives for a team to achieve with inserted activity will include primarily simple, single movements. For the team seeking to identify which trajectories are associated with inserted activity, a high level of insight will be provided into the trial objectives to aid in narrowing the search space. The second trial period will incorporate changes in recurring behaviors in addition to single movements, and the third trial period will scale to a larger set of data and more objectives. In the third trial period, the teams seeking to identify the anomalous trajectories will be provided with less insight into the trial objectives, and consequently will have to search a larger portion of the overall dataset.

1.A.2.2 Phase 2

In Phase 2, TA-1 teams will deliver the capability to add realistic geospatial error, temporal subsampling, population subsampling, ID confusion or confidence, and other sources of realism to their simulations. The types of noise and strategy for creating these realistic trajectories should be included in TA-1 proposals, as outlined in Section 1.A.1.1.

In addition to noise characteristics, TA-1 teams will add new capabilities to the simulation. Rare events such as roadblocks, concerts, or weather will be incorporated, and the teams will expand to a full range of modes of transportation. The teams will also begin to leverage coordinated movements and social connections, so the movement of individuals is coherent across friends, family, co-workers, or other relationships. These initial new components will be expected for Simulation #4 at a volume of 2M entities over a 2-month period and finalized with an increase in scale and fidelity in Simulation #5 with 5M entities simulated over a 4-month period.

TA-2 teams will be tested in two new trial periods, incorporating coordinated movements and the variety of different noise and subsampling characteristics introduced by TA-1 teams. The teams will need to deploy systems capable of handling varied levels of noise and subsampling while maintaining and increasing performance at modeling the underlying mobility. Additional elements will be available from TA-1 simulations to inform the level of noise to be expected. During the Phase 2 trials, the objectives will now include the identification of trajectories of multiple entities as anomalous, as opposed to the individual entity behaviors (through either single or recurring movements) that were executed in Phase 1.

1.A.2.3 Phase 3

In Phase 3, TA-1 teams will shift from simulations created for locations specified in advance early in a HAYSTAC phase to simulations that can be executed through an on-demand architecture where a new location can be executed quickly. This architecture will take specified input data such as a road network, GPS traces from a portion of the population, specific parameter distributions, and other data as needed, but run successfully with limited manual intervention. Teams should specify in their proposal the input data that their on-demand simulation architecture would require to be successful, and account for situations in which an input data type may not be available globally. The first delivered simulation during Phase 3, Simulation #6, will utilize the original architecture to achieve mega-city scale with 10M entities over a 6-month period. Simulation #7 will be created through the on-demand architecture to achieve 30M entities simulated over an entire year.

TA-2 teams will continue to execute trial periods, adapting to the increasing scale of TA-1 simulations. During this phase, the level of insight into trial objectives will reach zero, requiring teams to assess anomalous trajectories at the scale of the full population of data.

1.B Team Expertise

Collaborative efforts and teaming among Offerors are highly encouraged. It is anticipated that teams will be multidisciplinary and should include expertise in two or more of the disciplines listed below. This list is included only to provide guidance for the Offerors; satisfying all the areas of technical expertise below is not a requirement for selection and unconventional or innovative team expertise may be needed based on the proposed research. Proposals should include a description and the mix of skills and staffing that the Offeror determines will be necessary to carry out the proposed research and achieve program metrics.

- Agent-based modeling
- Anomaly Detection
- Behavioral Science
- Cloud Engineering
- Computational Science
- Deep learning
- Generative machine learning
- Graph Analytics
- Human mobility modeling
- Microsimulation
- Sequence modeling
- Software engineering
- Statistics
- Systems engineering
- Systems integration
- Trajectory analytics
- Transportation science

1.C Program Scope and Limitations

Proposals shall explicitly address all the following:

- **Underlying theory:** Proposed strategies to meet program-specified metrics must have firm theoretical bases that are described with enough detail that reviewers will be able to assess the viability of the approaches. Proposals shall properly describe and reference previous work upon which their approach is founded.

- **Research & Development approach:** Proposals shall describe the technical approach for meeting program metrics.
- **Technical risks:** Proposals shall identify technical risks and proposed mitigation strategies for each.
- **Software development:** Proposals shall describe the approach to software architecture, modularization, and integration.

The following areas of research are **out of scope** for the HAYSTAC program:

- 3D renderings of simulations
- "Black-box" microsimulation capabilities without insight and tuneability
- Creation of trajectories from video or remote sensing
- Tracking algorithms
- Development of hardware
- Approaches that consist merely of integrating currently existing software
- Research that does not have strong theoretical and experimental foundations

1.D Program Data

The HAYSTAC program will leverage data provided by the Government Team and data provided by the Performers. Proposals must specify the data needed to carry out the proposed research and what data characteristics are necessary for the Offeror's approach(es) to be successful at meeting program objectives. These details should be provided for both leveraging the Government Provided data as well as unique Performer provided data.

1.D.1 Government Provided Datasets

As initial sources of mobility insight, the Government Team will provide access to population statistics and geospatial foundation data at Program Kickoff. Teams may, but are not required to, leverage this information in their TA-1 solutions as input data. Documentation will be provided at Program Kickoff outlining these data sources and methods for accessing the data:

- **<u>Population density estimates.</u>** Aggregated estimates at the administrative region level and disaggregated estimates at higher spatial and temporal resolutions, down to building level in certain locations.
- **<u>Geospatial Foundation Data.</u>** Scalable databases of points of interest data, land cover, building footprints and heights, and other geospatially enabled information.

TA-2 teams may leverage these Government provided population estimates and geospatial foundation data as ancillary data in their sense-making but are not required to do so. The primary source of data for TA-2 teams will be the TA-1 Simulations 1-7 outlined in Table 3 and Test & Evaluation team's baseline simulation. However, TA-2 teams will not have access to all metadata attributes that may have been created by the TA-1 simulations for each trajectory, only the foundational information of the trajectory: time, location, and id.

1.D.2 Performer Provided Datasets

Each Performer is anticipated to have a unique technical solution to the HAYSTAC challenges and may require additional data for model training, model running, internal evaluation, or other research needs. Proposals must present a dataset development plan detailing how the team intends to obtain the data required. This documentation should account for any ground-truthing necessary as well as any other associated labor to curate and facilitate usage of data that is acquired.

TA-2 teams may not propose the incorporation of highly temporally sampled trajectory data beyond the TA-1 simulations; however, they may propose leveraging other sources of information, such as additional geospatial foundation data or population statistics. For sources of data that aid TA-2 processing, the data must be non-proprietary and eligible to be made available to all TA-2 teams and for incorporation into the T&E framework at no additional cost to the program.

As part of their proposal, teams shall prepare a HAYSTAC Privacy Plan Version 1.0 that comprehensively describes the efforts the teams will take to protect personally identifiable information and safeguard the security of any personal data collected or services involved in collection, transmission, processing, and storage of this data. Any claims that data are anonymous must be based on evidence and supported with sufficient information regarding how the data have been anonymized.

This version 1.0 of the HAYSTAC Privacy Plan shall be included in the Offeror's proposal as an appendix that covers all external datasets to be leveraged as part of the proposed research approaches. This appendix will not count against page count limits. The HAYSTAC Privacy Plan shall be updated at the beginning of each Phase and when new sources of data or datasets are proposed for use within a Performer's HAYSTAC research activities, including data used for either development or evaluation purposes.

1.E Test and Evaluation (T&E)

T&E will be conducted by an independent team of Government and contractor staff carrying out evaluation and analyses of Performer research Deliverables using program test datasets and protocols. The HAYSTAC Program will pursue rigorous and comprehensive T&E to ensure that research outcomes are well characterized, deliverables are aligned with program objectives, and that algorithm performance is measured across the full range of conditions. T&E activities will inform IARPA and Government stakeholders on HAYSTAC research progress and serve as invaluable feedback to the Performers to improve their research approaches, algorithm training practices, and system development.

Performers will have specific Deliverable Milestones driven by the HAYSTAC trial schedule at which all subcomponent and system algorithms and software will be delivered to IARPA and its designated T&E Team. The T&E Team will then conduct independent evaluations with the objective of characterizing the quality, functionality, and performance of the HAYSTAC systems. In addition to quantitative measurements, T&E will be carried out to establish a thorough understanding of the progress, status, and limitations of the Performer's research.

T&E results and feedback will be provided to Performers at regular intervals to keep them abreast of current independent performance measurements and to inform and improve their R&D approaches and methods. T&E results from all Performers will be shared with all teams to establish an understanding of the current state and progress of HAYSTAC research; T&E results will also be shared with USG external stakeholders, including their contractors, for Government purposes. IARPA may conduct other supplemental evaluations or measurements at its sole discretion to evaluate the Performers' research and Deliverables.

1.E.1 Trial Execution

The HAYSTAC trials will be organized and led by the T&E team. An outline of the sequence of events to accomplish a trial are depicted in Table 5. The first trial will be executed over a longer time period to allow sufficient debugging of interfaces and communication pathways. These interfaces and appropriate standards will be defined and provided to Performers at Program Kickoff.

Table 5: Individual Trial Schedule and Events

Timeline	Event
3 weeks before Trial Period Start (TPS)	Trial Plan provided to teams
TPS	TA-1 simulation delivered to Trajectory Datastore for T&E review
TPS + 1 week	T&E team checkout complete, data access provided to TA-2 teams
TPS + 4 weeks	Container delivered to Processing Framework outputting proposed trajectories to meet each objective
TPS + 5 weeks	T&E Verifies that objects are met and provides to TA-1 team
TPS + 7 weeks	Next period of TA-1 simulation delivered including proposed trajectories
TPS + 7 weeks	T&E team checkout complete, including verification that objectives are met in new TA-1 simulation
Trial Period End (TPE)	Container delivered to Processing Framework outputting flagged trajectories and alerts
TPE + 4 weeks	T&E report on Trial Period

HAYSTAC T&E will be executed using significant automation to aid in the execution of the trials, leveraging a central trajectory datastore for data management and an automated processing framework capable of executing TA-2 submissions.

1.E.2 Trajectory Datastore

The primary integration point between TA-1 simulations and TA-2 trials will be the trajectory datastore deployed by the T&E team for Program Kickoff. The datastore will ingest the output of TA-1 simulations and allow spatial-temporal and attribute-based queries to return data quickly, implementing access controls to ensure that sequestered information is not made available to a team that should not have access, such as any identifiable information as to which trajectories were proposed by a TA-2 trial participant inserting activity. The Datastore will offer access to the full attributes of a TA-1 Performer's simulated output, but return only the fundamental trajectory information (time, location, ID) to TA-2 Performers, with some additional noise characterization data provided in Phase 2. Documentation and standards will be made available to Performers at Program Kickoff.

1.E.3 Processing Framework

The HAYSTAC T&E processing framework will orchestrate software deliverables and enable the operation and scoring of TA-2 trials. The processing framework components include:

- Version controlled software repository
- Containerized workflow orchestration
- Automated scaling based on volume of trajectories

This framework will enable to automation of TA-2 trials, including the integration of Performer algorithms, T&E baseline algorithms, and metrics code as containerized modules. Performer TA-2 systems will be submitted to the software repository and built into containers that are then stored within the container registry and available for T&E execution. This framework will also facilitate parameterized addition of noise to simulations, initially through T&E baseline methods and migrating to TA-1 Performer modules for adding noise in Phase 2.

TA-1 simulations will not be processed at scale in the T&E framework, except for the addition of noise characteristics. Instead, Performers will host TA-1 processing in their own environments and output trajectories to the Trajectory Datastore.

1.E.4 Baseline Algorithms

The T&E team will create baseline algorithms to provide context for Performer innovations and to test and develop the overall HAYSTAC T&E pipeline. The TA-1 baseline will create an example simulation for TA-1 teams to emulate for standards compliance and for TA-2 teams to leverage in their initial model training and testing. TA-1 baseline simulation outputs will be provided at Program Kickoff and at the starts of Trial Periods #1, #2, and #3. An additional baseline module to degrade TA-1 simulations with realistic noise characteristics will be available at the end of Phase 1 and will be replaced with Performer capabilities for adding relevant noise and subsampling in Phase 2 TA-1 systems.

The TA-2 baseline will be leveraged to compete in each HAYSTAC trial, both in generating normal trajectories as well as identifying anomalous ones, in order to provide a stable measure of comparison as TA-2 systems evolve and improve over the program. Teams will be provided with source code of the TA-2 baselines for internal testing and as example code that successfully integrates into the HAYSTAC Processing Framework. An initial version of the TA-2 baseline will be provided at the start of Trial Period #1, and a second version will be provided at the start of Trial Period #2.

1.F Program Metrics

Achievement of metrics is a performance indicator under IARPA research contracts. IARPA has defined HAYSTAC program metrics to evaluate effectiveness of the proposed solutions in achieving the stated program goal and objectives, and to determine whether satisfactory progress is being made. The metrics described in this BAA are shared with the intent to scope the effort, while affording maximum flexibility, creativity, and innovation to Offerors proposing solutions to the stated problem.

The HAYSTAC T&E protocols and evaluation methodology are currently under development; further details will be provided at Program Kickoff through a Metrics Description Document. Program metrics may be refined during the various phases of the HAYSTAC program; if metrics change, revised metrics will be communicated in a timely manner to Performers. The evaluation methodology may be revised by the Government at any time during the program lifecycle to better meet program needs.

The preliminary program metrics and target scores are provided in Section 1.F.1 for TA-1 and Section 1.F.2 for TA-2.

1.F.1 TA-1 Metrics

TA-1 evaluations will assess simulation progress against Phase goals for number of entities, duration, modes of travel, and the level to which rare events and traffic dynamics have been incorporated. Additionally, the realism of movements and locations will be assessed through comparison to Government hold out trajectory datasets across a variety of mobility statistics. The comparison datasets will not be shared with teams, but feedback will be provided in the context of metrics performance. To quantify the similarity between hold out data and simulated data across diverse statistics at the population or subpopulation scale, Jensen-Shannon Divergence (JSD) or similar divergence measure will be leveraged to compare distributions.

$$JSD(P||Q) = \frac{1}{2}\sum_{x\in X}\left[P(x)\log\left(\frac{P(x)}{M(x)}\right)\right] + \frac{1}{2}\sum_{x\in X}\left[Q(x)\log\left(\frac{Q(x)}{M(x)}\right)\right]$$

$$where\ M = \frac{1}{2}(P+Q)$$

This formula computes a symmetric distance measure between two statistical distributions P and Q, where P represents one of the distributions specified from Table 6 computed against the simulated data, and Q represents the distribution computed against real-world data. X represents the full range of probabilities in the metric space of a given mobility statistic. By computing various population level mobility statistics and comparing to real-world data that is held out from the simulations, the JSD metric will capture the fidelity of the simulated movement.

Teams will seek to create simulations with a JSD as close to 0 as possible, as the emergent behaviors in the simulations should match real observations. An initial list of mobility statistics to compare with this divergence metric are listed in Table 6, and will be refined at Program Kickoff. The Phase target values are also included in Table 6, with one target presented as the mean across Movement-based mobility statistics and one target as the mean across the Location-based metrics.

Table 6: Potential Mobility Statistics to Support TA-1 Metrics

Type	Statistic	Definition	JSD Metric Goals		
			Phase 1	Phase 2	Phase 3
Movement	Total Distance Per Day	The total distance traveled by an individual per day	0.02	0.01	0.005
	Radius of Gyration	The level an individual deviates from their geospatial center of mass			
	Number of Locations Visited	The number of unique locations visited by an individual per day			
	Temporal Variability	Deviations in time for recurring departures and visits			
	Level of Exploration	The concentration of an individual's visits within their top-k visited locations			
	Kinematic Profiles	Features derived from velocity, acceleration, and travel time information			
	Pairwise Encounter Statistics	Pairwise frequency and inter-encounter time of multiple individuals			
Location	Origin-Destination Probability	Probability of visiting one region from another	0.20	0.15	0.10
	Importance	Probability of visiting a region at a given time			
	Connectivity	Graph statistics from trips linking locations			

1.F.2 TA-2 Metrics

TA-2 evaluations will leverage the HAYSTAC trials to create true positive sets of trajectories and score teams on their performance at either generating normal trajectories that goes undetected, or by altering to trajectories that were injected by another team. For each trial period, teams will be scored on the probability of detection (PD) and false alarm ratio (FAR) by aggregating the results as shown in Figure 6.

$$PD = \frac{Correct\ Alerts}{\#\ of\ True\ Activities}$$

$$FAR = \frac{Incorrect\ Alerts}{\#\ of\ Alerts}$$

Figure 6: TA-2 Trial Scoring

The T&E team's baseline for trial participation will be leveraged as a basis of comparison to keep a steady state that enables monitoring of program process over time. In the case of generating normal trajectories, a lower probability of detection is desired against the team's injected trajectories at a nominal false alarm ratio operating point, as captured in Table 7. In the case of detecting anomalous trajectories, a higher probability of detection is sought while also minimizing false alarm ratio, as captured in Table 8.

Table 7: Generating Normal Trajectories Metric Goals

Metric	Phase 1 Goal	Phase 2 Goal	Phase 3 Goal
PD	40%	25%	10%
FAR	50%	50%	50%

Table 8: Detecting Anomalous Trajectories Metric Goals

Metric	Phase 1 Goal	Phase 2 Goal	Phase 3 Goal
PD	40%	60%	80%
FAR	50%	35%	20%

1.G Program Waypoints, Milestones, and Deliverables

Waypoints, Milestones, and Deliverables are established from the program's onset to ensure alignment with HAYSTAC objectives, organize research activities in a logical and reportable manner, and facilitate consistent and efficient communication among all stakeholders – IARPA, the HAYSTAC T&E Team, USG Stakeholders, and Research Performers. A schedule of key program Milestones and Deliverables is shown in Figure 7.

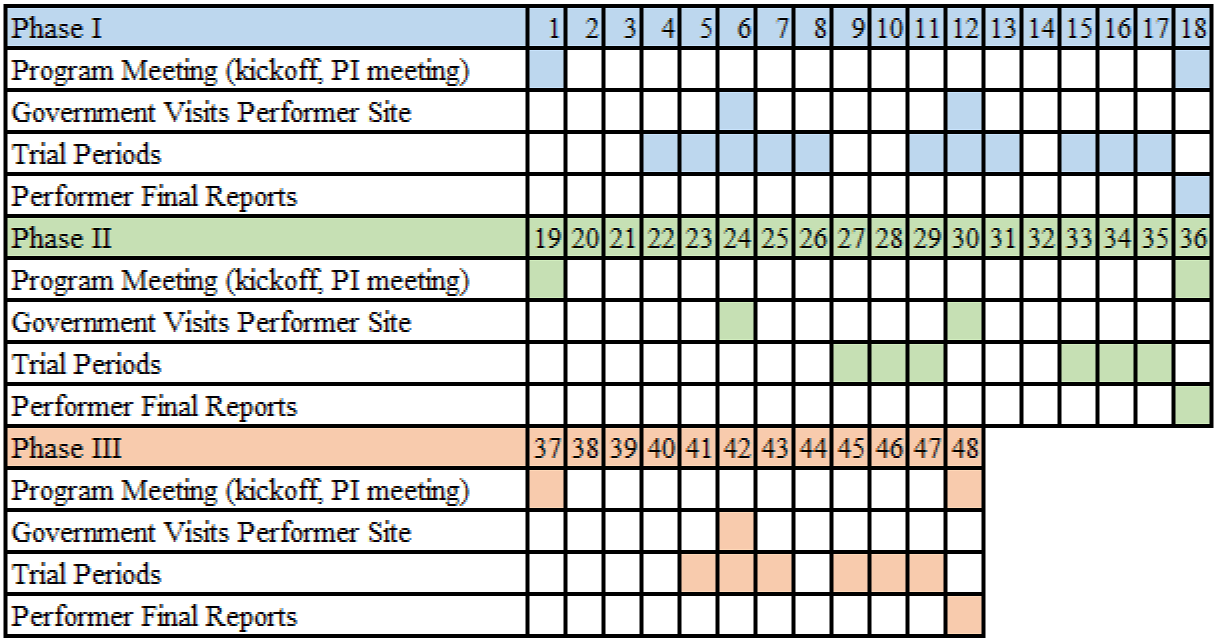

Figure 7: Schedule of Key Milestones and Deliverables

1.G.1 Program Waypoints

Waypoints are the means by which the Performer clearly demonstrates the quantitative and timely progress that must be made for the overall concept to meet end-of-phase Milestones. In other words, the intent of Waypoints is to provide a clear measure of progress towards meeting the program Milestones so the PM and advisors can provide more effective guidance and assistance to the Performers. Performance against these Waypoints will be reviewed throughout the program, and the PM and advisors will use performance against the Waypoints to assess whether course corrections are needed to ensure program success. Table 9, Table 10, and Table 11 show the timeline for the program with defined Milestones, Waypoints and Deliverables.

Table 9: Phase 1 Milestone, Waypoint, and Deliverable Timeline

Month	Event	Description	Comments	Deliverables
All	Waypoint	Monthly Status Report (MSR)	Due on 15th of each month; Both technical and financial reports will be submitted	MSR
All	Waypoint	Progress and Status Meeting	Monthly teleconference with HAYSTAC PM	Presentation; action item list
1	Waypoint	Baseline Simulation	Provided as GFI with format documentation	N/A
1	Waypoint	Population Statistics and Geospatial Foundation Data	Provided as GFI with access documentation	N/A
1	Waypoint	Kickoff Meeting	Location: Washington Metro Area; Metrics Description Documents and Performer Trial Interfaces and Standards Documentation furnished to Performers;	Presentation
2	Deliverable	Privacy Plan v1.1	Submitted for IARPA approval	Report

Month	Event	Description	Comments	Deliverables
3	Deliverable	Simulation #1	Updated for second simulation after TA-2 input trajectories	Simulated data to Trajectory Repository
4-8	Milestone	Trial Period #1	Extended period for first Trial Period; Deliverables updated for bugs throughout	TA2 Software to Processing Framework
6	Waypoint	Site Visit	At Performer site	Presentation; action item list
10	Deliverable	Simulation #2	Updated for second simulation after TA-2 input trajectories	Simulated data to Trajectory Repository
11-13	Milestone	Trial Period #2	Deliverables updated for bugs throughout	TA2 Software to Processing Framework
12	Waypoint	Site Visit	At Performer site	Presentation; action item list
14	Deliverable	Simulation #3	Updated for second simulation after TA-2 input trajectories	Simulated data to Trajectory Repository
15-17	Milestone	Trial Period #3	Deliverables updated for bugs throughout	TA2 Software to Processing Framework
17	Deliverable	Phase 1 Report (Draft)	Any updated or additional data or software also due	Report
18	Waypoint	PI Workshop	Location: Washington Metro Area	Presentation
18	Deliverable	Phase 1 Report (Final)	Any updated or additional data or software also due	Report

Table 10: Phase 2 Milestone, Waypoint, and Deliverable Timeline

Month	Event	Description	Comments	Deliverables
All	Waypoint	Monthly Status Report (MSR)	Due on 15th of each month; Both technical and financial reports will be submitted	MSR
All	Waypoint	Progress and Status Meeting	Monthly teleconference with HAYSTAC PM	Presentation; action item list
20	Waypoint	Kickoff	Location: Washington Metro Area	Presentation
20	Deliverable	Privacy Plan v2.0	Submitted for IARPA approval	Report
24	Waypoint	Site Visit	At Performer site	Presentation; action item list
26	Deliverable	Simulation #4	Updated for second simulation after TA-2 input trajectories	Simulated data to Trajectory Repository

Month	Event	Description	Comments	Deliverables
27-29	Milestone	Trial Period #4	Deliverables updated for bugs throughout	TA2 Software to Processing Framework
30	Waypoint	Site Visit	At Performer site	Presentation; action item list
32	Deliverable	Simulation #5	Updated for second simulation after TA-2 input trajectories	Simulated data to Trajectory Repository
33-35	Milestone	Trial Period #5	Deliverables updated for bugs throughout	TA2 Software to Processing Framework
35	Deliverable	Phase 2 Report (Draft)	Any updated or additional data or software also due	Report
36	Waypoint	PI Workshop	Location: Washington Metro Area	Presentation
36	Deliverable	Phase 2 Report (Final)	Any updated or additional data or software also due	Report

Table 11: Phase 3 Milestone, Waypoint, and Deliverable Timeline

Month	Event	Description	Comments	Deliverables
All	Waypoint	Monthly Status Report (MSR)	Due on 15th of each month; Both technical and financial reports will be submitted	MSR
All	Waypoint	Progress and Status Meeting	Monthly teleconference with HAYSTAC PM	Presentation; action item list
37	Waypoint	Kickoff	Location: Washington Metro Area	Presentation
40	Deliverable	Simulation #1	Updated for second simulation after TA-2 input trajectories	Simulated data to Trajectory Repository
41	Milestone	Trial Period #6	Deliverables updated for bugs throughout	TA2 Software to Processing Framework
42	Waypoint	Site Visit	At Performer site	Presentation; action item list
44	Deliverable	Simulation #1	Created on-demand; Updated for second simulation after TA-2 input trajectories	Simulated data to Trajectory Repository
45	Milestone	Trial Period #7	Deliverables updated for bugs throughout	TA2 Software to Processing Framework
47	Deliverable	Final Program Report (Draft)	Any updated or additional data or software also due	Report
48	Waypoint	PI Workshop	Location: Washington Metro Area	Presentation
48	Deliverable	Final Program Report (Final)	Any updated or additional data or software also due	Report

1.G.2 Software Deliverable Formatting

Performers will be required to provide algorithm and software Deliverables to the software repository that is a component of the HAYSTAC Processing Framework outlined in Section 1.E.3. Software will be provided as source code by the Performer, and the framework will build a container that is registered with the framework's container registry. Teams will be expected to coordinate with the T&E team to ensure successful build of the container from source. For TA-2 models that require training, the expectation is for the initial model training to occur on Performer systems, with the ability for the T&E Team to re-train and test the model with the same and/or other data. Offeror teams that do not include the requisite expertise to conduct such software development should include costs in their proposal to obtain software development support.

Each team is required to include among their key personnel a Lead System Integrator (LSI) who shall be responsible for preparing software Deliverable subcomponents, modules, and systems, performing quality control of Deliverables, and integrating key components into the primary HAYSTAC system(s). The LSI will also oversee communication and coordination across a Performer's research teams including subcontractors, if applicable, to ensure research products are functional and following software coding best practices (e.g., inline comments, documentation). Additional team members and roles are dependent on the proposed research, as such, there is no predetermined or required skill mix.

1.G.3 Program Interfaces and Standards

The HAYSTAC Program will utilize well defined standards for all software Deliverables and evaluations in order to automate significant portions of trial execution. The first version of the HAYSTAC Performer Trial Interfaces and Standards documentation will be provided to Performers at the Phase 1 Kickoff Meeting and updated periodically thereafter. All Performer solutions must be compatible with the HAYSTAC standards.

TA-1 solutions will conform to the defined formatting specifications for simulated trajectories, however TA-1 proposals should include the list of attributes the simulation is capable of including in the output to enrich the data beyond the fundamentals of a trajectory data: time, location, and id.

1.H Meeting and Travel Requirements

Offerors are expected to assume responsibility for administration of their projects and to comply with contractual and program requirements for reporting, attendance at program workshops, and availability for site visits. The following paragraphs describe typical expectations for meetings and travel for IARPA programs as well as the planned frequency and locations of such meetings. In addition to ensuring that all necessary details of developed software, algorithm, and operational instructions are clear and complete, each Performer will be required to be available for questions and troubleshooting from the T&E Team in Performer status meetings.

1.H.1 Kickoff Meetings and PI Workshops

All Performer teams are expected to attend workshops, including key personnel from prime and subcontractor organizations.

The HAYSTAC program intends to hold a program Kickoff Meeting workshop in the first month of the program and the first month of each subsequent program phase. In addition, the program will hold a PI Review Meeting in the final month of Phase 1 and then similar workshops in each Phase thereafter. Kickoff Meetings and PI Review Meetings may be combined for logistical convenience. The dates and locations of these meetings are to be specified at a later date by the Government, but for planning purposes, Offerors should use the approximate times and locations listed in Table 9, Table 10, and Table 11. Both types of meetings will likely be held in the Washington, D.C. metropolitan area, but IARPA may opt to co-locate the meeting with a relevant external conference or workshop to increase synergy with stakeholders. IARPA reserves the right to hold the meeting virtually for logistical or health and safety reasons.

Kickoff Meetings will typically be one day in duration and will focus on plans for the coming Phase, Performer planned research, and internal program discussions. PI Review Meetings will typically be two days in duration and will have a greater focus on communicating program progress and plans to USG stakeholders. These meetings will include additional time allocated to presentation and discussion of research accomplishments as well as interactive system demonstrations for Government stakeholders.

In both cases, the workshops will focus on technical aspects of the program and on facilitating open technical exchanges, interaction, and sharing among the various program participants. Program participants will be expected to present the technical status and progress of their projects to other participants and invited guests. Individual sessions for each Performer team with the HAYSTAC PM and T&E Team may be scheduled to coincide with these workshops. Non-proprietary information will be shared by Performers in the open meeting sessions; proprietary information sharing shall occur during individual breakout sessions with the HAYSTAC PM and the T&E Team.

1.H.2 Site Visits

Site visits by the Government Team will generally take place as outlined in Table 9, Table 10, and Table 11. These visits will occur at the Performer's facility. Reports on technical progress, details of successes and issues, contributions to the program goals, and technology demonstrations will be expected at such site visits. Performers shall participate and provide final meeting documents, to include captured action items, within 15 calendar days following the meeting. Draft materials, for any presentations, are due 5 workdays prior to the meeting. IARPA reserves the right to conduct additional site visits on an as-needed basis. IARPA also reserves to the right to reduce the number of site visits or participate virtually for logistical or health and safety reasons.

1.H.3 Recurring Status Meetings

Remote monthly meetings will be established after the HAYSTAC Program Kickoff, and teams will present the previous month's research activities, review open action items, discuss upcoming research, and identify any concerns or issues which could impact the program. IARPA may establish these remote status meetings biweekly instead of monthly if it is determined that this would be beneficial to maintain progress against HAYSTAC goals.

1.I Period of Performance

The HAYSTAC program is anticipated to be a 4-year (48 month) effort, comprised of three Phases. All three Phases are being solicited under this BAA. Phase 1 will be 18 months in duration, Phase 2 will be 18 months, and Phase 3 will be 12 months. Offerors are to submit a proposal that addresses all three phases.

1.J Place of Performance

Performance will be conducted at the Performers' sites.