

DPA26BZ04-DV015
Art of Novel Signals: Predicting and Forecasting with High Confidence
Frequently Asked Questions (FAQs)

1. What event types or forecasting domains are most important for evaluating prediction confidence?

A: The topic Phase I feasibility criteria is seeking “Forecast precision at the baseline (days-long) horizon at or above 80% fidelity (metrics validated by an independent team/organization) event-type data such as Armed Conflict Location and Event Data (ACLED) style political and conflict events.”

2. Which novel signal sources are most valuable: multilingual radio, local news, community reporting, multimodal media, or other sources?

A: Proposers are responsible for sourcing, generating, and justifying all data and testing materials required to execute their proposed effort. The end goal of the Phase II effort is to demonstrate a commercially deployable, validated early-warning forecasting capability built on multilingual signal sourced from data-sparse environments, extending reliable prediction from five days to a two-week horizon at higher precision.

3. What geographic or language coverage is expected for initial demonstrations?

A: The topic Phase I feasibility criteria is seeking “deployment and evaluation in following strategic regions of interest: Central and Southeast Asia, East and Northeast Africa, and South America”

4. How should precision, false-positive reduction, and forecasting lead time be balanced?

A: Phase II fixed milestones for this program should include a Base Period where the performers are expected to produce precision across target regions with a precision of approximately 80%, and the novel-signal approach that would raise this precision level up to 90%, while reducing within-country false-positive rates by roughly half.

A successful solution will be evaluated on its ability to produce precision across target regions with a precision of approximately 80%, and the novel-signal approach that would raise this precision level up to 90%, while reducing within-country false-positive rates by roughly half. The performers must justify their choice in performance metrics with references, data and/or theory.

5. What level of explainability is expected for forecasts generated from temporal knowledge graphs and in-context learning?

A: Proposals will be considered for DP2 funding based on the ability of the proposing team to build a theoretical framework-based forecasting system that leverages novel multilingual, multimodal open-source signal with in-context learning over temporal knowledge graphs to anticipate events in data-sparse environments.

6. Direct-to-Phase-II readiness evidence: Because this topic is soliciting Direct-to-Phase-II proposals only, what specific prior evidence should offerors provide to demonstrate Phase-I-equivalent feasibility? For example, would the Government prefer documented evidence of a working in-context-learning forecasting pipeline, prior temporal knowledge graph demonstrations, multilingual/multimodal ingestion results, independent validation of baseline forecasting precision, or all of these elements?

A: This topic is soliciting Direct to Phase II (DP2) proposals. To be considered for a DP2 award, proposers must demonstrate that they have developed a theory for identifying critical knowledge and a framework for testing the theory by determining knowledge requirements and assessing technologies that facilitate multi-modal data ingestion and analyses in complex, real-world settings. Performers should plan to operationalize and evaluate their frameworks during Phase II.

7. Forecasting mission and event taxonomy: What event types should the proposed system prioritize for prediction and early warning? Should offerors focus on Armed Conflict Location and Event Data-style political and conflict events, force-protection indicators, instability and unrest events, humanitarian assistance and disaster-response indicators, or another Government-defined event taxonomy?

A: The topic Phase I feasibility criteria is seeking “Forecast precision at the baseline (days-long) horizon at or above 80% fidelity (metrics validated by an independent team/organization) event-type data such as Armed Conflict Location and Event Data (ACLED) style political and conflict events.”

8. Target regions, languages, and data-sparse environments: The topic identifies Central and Southeast Asia, East and Northeast Africa, and South America as strategic regions of interest. Are there priority countries, languages, dialects, or operational environments that should drive the Phase-II technical approach, data-collection plan, and demonstration design?

A: No. The topic Phase I feasibility criteria is seeking “deployment and evaluation in following strategic regions of interest: Central and Southeast Asia, East and Northeast Africa, and South America”

9. Novel signal sources and collection assumptions: What types of novel signals are considered most valuable and acceptable for this effort, including multilingual radio, local news, community audio, messaging-platform content, software-defined radio collection, online radio streams, and other open-source data? Are offerors expected to deploy or operate in-region radio receivers, or is a software-only ingestion and analytics approach acceptable if it can demonstrate access to relevant signal?

A: Proposers are responsible for sourcing, generating, and justifying all data and testing materials required to execute their proposed effort. The end goal of the Phase II effort is to demonstrate a commercially deployable, validated early-warning forecasting capability built on multilingual signal sourced from data-sparse environments, extending reliable prediction from five days to a two-week horizon at higher precision.

10. Data policy, privacy, and operational constraints: Should the proposed capability rely exclusively on publicly available information and open-source intelligence, or may Government-furnished, partner-provided, or restricted datasets be used during Phase II? Are there specific privacy, legal, operational security, or host-nation constraints that should be reflected in the proposal and data-management plan?

A: The proposers are responsible for sourcing, generating, and justifying all data and testing materials required to execute their proposed effort. Additionally, proposers are responsible for adhering to privacy, legal, operational security, or host-nation constraints for their respective proposed efforts.

11. Performance metrics and independent validation: How will success be measured across precision, accuracy, recall, false-positive rate, false-negative rate, forecasting horizon, and warning value to an operational user? In particular, how should offerors interpret the target of improving forecasting precision toward 90 percent, reducing within-country false-positive rates, and extending warning from days to weeks?

A: Phase II fixed milestones for this program should include a Base Period where the performers are expected to produce precision across target regions with a precision of approximately 80%, and the novel-signal approach that would raise this precision level up to 90%, while reducing within-country false-positive rates by roughly half.

A successful solution will be evaluated on its ability to produce precision across target regions with a precision of approximately 80%, and the novel-signal approach that would raise this precision level up to 90%, while reducing within-country false-positive rates by roughly half. The performers must justify their choice in performance metrics with references, data and/or theory.

12. Automatic speech recognition and language onboarding: What level of automatic speech recognition performance is expected for noisy, multi-speaker, low-resource, predominantly oral languages? Should offerors propose explicit word-error-rate targets by language and acoustic condition, and how should the Government evaluate rapid onboarding of a new crisis language using cross-lingual transfer, human correction, and synthetic data?

A: The proposers are responsible for sourcing, generating, and justifying all data and testing materials required to execute their proposed effort.

13. Temporal knowledge graph and in-context-learning architecture: Does the Government require forecasting over temporal knowledge graphs without retraining graph embeddings, as stated in the feasibility criteria, or are hybrid approaches allowed if they preserve explainability, traceability, and rapid adaptation? Are there preferred schemas, entity/event ontologies, uncertainty representations, or interfaces for temporal knowledge graph construction and query?

A: The proposers are responsible for sourcing, generating, and justifying all data and testing materials required to execute their proposed effort.

14. Multimodal and paralinguistic forecasting features: The topic mentions progression from textual event forecasting toward paralinguistic features such as pitch, energy,

speaking rate, emotional arousal, valence, and agitation. Should paralinguistic analysis be treated as a required Phase-II capability, an option-period enhancement, or a stretch goal? How should cultural and language-dependent bias in affective-signal interpretation be mitigated and evaluated?

A: The Phase II fixed milestones of the effort should include:

Month 15

Milestone: Extend the audio model beyond transcription to paralinguistic signal. Extract prosodic and affective features from broadcast audio, including pitch, energy, speaking rate, and indicators of emotional arousal, valence, and agitation; stand up this feature pipeline across the target languages; and establish a forecasting baseline that combines these features with the textual event stream.

Proposers are responsible for sourcing, generating, and justifying all data and testing materials required to execute their proposed effort.

15. Deployment model, user workflow, and transition pathway: What deployment model and transition path should offerors assume for Phase II and the option period: cloud-hosted analytics, disconnected or austere edge deployment, a ruggedized field device, or integration into an existing Government indications-and-warning workflow? Who is the intended operational user or transition sponsor, and what outputs, dashboards, application programming interfaces, or alert formats would be most useful for an operational pilot?

A: Proposers must identify, describe, and justify their own deployment model and transition strategies, including any technical, ethical, regulatory, or operational challenges applicable to their proposed approach. The end goal of the Phase II effort is to demonstrate a commercially deployable, validated early-warning forecasting capability built on multilingual signal sourced from data-sparse environments, extending reliable prediction from five days to a two-week horizon at higher precision that will feed into Phase III. Phase III will be oriented towards transition within DoW/Military ecosystem and further commercialization of the technology.

16. The topic states that the desired approach has 4 interlocking parts, would DARPA consider a proposal responsive that addresses a subset of the parts, but not all four? For example, a proposal addressing integration of audio transcription into forecasting systems to provide early warnings, but not development of new ASR models?

A: No. The end goal of the Phase II for this effort is to demonstrate a commercially deployable, validated early-warning forecasting capability built on multilingual signal sourced from data-sparse environments, extending reliable prediction from five days to a two-week horizon at higher precision.

17. The linked reference for forecasting with LLMs using in context learning and temporal knowledge graphs are constrained to inductive settings, i.e. prediction questions where the answer MUST be from observable history. Is DARPA interested

in transductive approaches which are capable of predicting events that are outside of observable history?

A: Proposals will be considered for DP2 funding based on the ability of the proposing team to build a theoretical framework-based forecasting system that leverages novel, multilingual, multimodal open-source signal with in-context learning over temporal knowledge graphs to anticipate events in data-sparse environments.

18. For this topic, is the question of what to forecast thought to be known beforehand? That is, is the question about what to predict known ahead of time?

A: The topic Phase I feasibility criteria is seeking “Forecast precision at the baseline (days-long) horizon at or above 80% fidelity (metrics validated by an independent team/organization) event-type data such as Armed Conflict Location and Event Data (ACLED) style political and conflict events.”

19. The topic references 80% state of the art forecasting precision across several regions of interest, are you able to provide a reference for this state-of-the-art precision?

A: Phase II fixed milestones for this program should include a Base Period where the performers are expected to produce precision across target regions with a precision of approximately 80%, and the novel-signal approach that would raise this precision level up to 90%, while reducing within-country false-positive rates by roughly half.

A successful solution will be evaluated on its ability to produce precision across target regions with a precision of approximately 80%, and the novel-signal approach that would raise this precision level up to 90%, while reducing within-country false-positive rates by roughly half. The performers must justify their choice in performance metrics with references, data and/or theory.

20. Are SBCs expected to procure, deploy, and maintain the physical SDR hardware infrastructure in the regions of interest, or will the Government provide access to existing data feeds?

A: Proposers are responsible for sourcing, generating, and justifying all data and testing materials required to execute their proposed effort.

21. Does the Government have a prioritized list of low-resource or predominantly oral languages it wishes to target?

A: The topic Phase I feasibility criteria is seeking “deployment and evaluation in following strategic regions of interest: Central and Southeast Asia, East and Northeast Africa, and South America”

22. How will the benchmark of “Forecast precision... at or above 80% fidelity” be evaluated during the DP2 effort?

A: Phase II fixed milestones for this program should include a Base Period where the performers are expected to produce precision across target regions with a precision of approximately 80%, and the novel-signal approach that would raise this precision level up to 90%, while reducing within-country false-positive rates by roughly half.

A successful solution will be evaluated on its ability to produce precision across target regions with a precision of approximately 80%, and the novel-signal approach that would raise this precision level up to 90%, while reducing within-country false-positive rates by roughly half. The performers must justify their choice in performance metrics with references, data and/or theory.

23. Please confirm that the topic's precision criteria will be adjudicated against the evaluation methodology the offeror proposes and registers, and that the independent validation named in the feasibility criteria applies to those registered terms. If you have set values for any of these terms, please state them.

A: Proposers are responsible for sourcing, generating, and justifying all data and testing materials required to execute their proposed effort. The end goal of the Phase II effort is to demonstrate a commercially deployable, validated early-warning forecasting capability built on multilingual signal sourced from data-sparse environments, extending reliable prediction from five days to a two-week horizon at higher precision.