Evaluation Methodology and Award Schedule

Xpress Challenge awards are structured to recognize and reward increasing levels of analytic sophistication. Solver submissions will be evaluated using ODNI's *Rating Scale for Evaluating Analytic Tradecraft Standards* (*RSEATS*). Specifically, solver's submissions will be evaluated based on how they describe the quality and credibility of underlying **sources**, demonstrate national security **relevance**, properly express and explain **uncertainties** associated with major analytic judgments, **distinguish** between underlying information and the assumptions and judgments of analysts, incorporate analysis of **alternatives**, and demonstrate and use clear and logical **argumentation**.

Xpress Challenge Award Schedule

Award Area	Award Criteria	Prize Awards
Literal	Ability to craft sound written material in response to the posed intelligence question	$50,000
Inferential	Ability to discern and characterize how offered reasons[1] support analytic judgements and conclusions	$50,000
Evaluative	Ability to make reasoned assertions and incorporate alternative analysis	$50,000
Creativity	Selection from Xpress Challenge Steering Group members	$50,000
Early STEM Education	Three highest cumulative scores of **Literal**, **Inferential**, and **Evaluative** award areas from high school student teams	$15,000, $10,000 and $5,000[2]
Overall Best Submissions	Five highest cumulative scores of **Literal**, **Inferential**, and **Evaluative** award areas	$100k, $75, $50k, $30k, and $15k
Total		**$500k**

With the exception of the **Creativity** award, Xpress prizes will be awarded in the areas and amounts shown above based on a blind review by ODNI's Analytic Integrity and Standards (AIS)—the IC's established body for reviewing IC-wide analytic products. The **Literal, Inferential, and Evaluative** award areas are derived from AIS' existing RSEATS evaluation criteria and progressively gauge the narrative sophistication of Solvers' submissions. The **Creativity** award area will be based on a selection by a panel of senior USG officials chosen by the Seekers. In an effort to promote early science, technology, engineering, and mathematics (STEM) education, the **Early STEM Education** awards will be awarded to the highest-performing high school team

[1] **Reasons** (including inferential claims) can be direct evidence, assumptions, precedents, or logical inferences.
[2] Awarded directly to the students' represented school.

Solvers that produce the best cumulative score from AIS evaluation of the **Literal, Inferential,** and **Evaluative** award areas. The **Overall Best Submissions** award area will be awarded to the Solver(s) that produces the best cumulative score from AIS evaluation of the **Literal, Inferential,** and **Evaluative** award areas. To be eligible for an **Overall Best Submission** award the submitted Analytic Product must receive a score of **Fair (1)** or above for each evaluation criteria. Winners in the category award areas of Literal, Inferential, and Evaluative are determined by the highest score for the criteria in the respective category regardless of performance in the other categories.

A product may perform well in one or more of the three categories but fall short of standards in one or more of the others. For example, a product may have a strong message that addresses the posed intelligence question, but express weak argumentation that obscures or undercuts its value. Evaluators will do their best not to conflate categories.

A. Literal Response Criteria

Solvers' submissions against the Literal Response Criteria will be scored as the arithmetic sum of the scores for Criteria 1 and 2, as outlined below. AIS will make the ultimate determination of the winner for the Literal award category.

<u>Literal Response Criterion 1</u>: Properly describes quality and credibility of underlying sources, data, and methodologies.

Background: Solvers' submissions should accurately characterize the information in the underlying sources and explain which information proved key to analytic judgments and why. Factors significantly affecting the weighting that the analysis gives to available, relevant information, such as denial and deception, source access, source motivations and bias, or age and continued currency of information, or other factors affecting the quality and potential reliability of the information, should be included in the product.

Poor (0)	Fair (1)	Good (2)	Excellent (3)
(1) Largely lacks sourcing or describes reporting base, data, or methodologies only vaguely; **OR** (2) Misidentifies or misrepresents cited reporting, data, or methodologies.	(1) Contains basic, generic descriptions of cited reporting, data, or methodologies; **BUT** (2) Provides little detail on factors that may affect the quality and credibility of underlying sources, data, or methodologies.	(1) Contains at least basic, generic descriptions of cited reporting, data, or methodologies; **AND** (2) Provides considerable detail on factors that may affect the quality and credibility of underlying sources, data, or methodologies.	Satisfies "good" criteria; **AND** (1) Identifies which sources are most important to major analytic judgments; **OR** (2) Provides additional detail about sources, data, or methodologies that provides insight into their contribution to the analysis.

Note: Source reference citations should be included as endnotes in disseminated analytic products. In rating this standard, the totality of information in a product's source reference citation endnotes, source summary statement, and main text must be taken into account.

<u>Literal Response Criterion 2</u>: Demonstrates customer and addresses implications.

Background: Solvers' submissions should provide information and insight on issues relevant to the products' intended consumers and/or provide useful context. To meet this standard fully, Solvers' submissions should examine and explicitly address direct or near-term implications of the information for the intended audience and/or for U.S. national security interests, and, when possible, also relay longer-term implications or identify potential indirect or second-order effects.

Poor (0)	Fair (1)	Good (2)	Excellent (3)
(1) Provides little or no information or analysis beyond what is generally known; **OR** (2) Does not respond adequately to a specific tasking.	(1) Provides useful information and analysis but does not address implications; **OR** (2) Does not address an important issue or question raised by the analysis; **OR** (3) Satisfies a specific tasking only partially.	(1) Provides useful information and analysis and addresses near-term, direct, or first-order implications; **AND** (2) Adds value by addressing at least one of the following: trends or prospects, appropriate context, insight gained from synthesizing a large volume of information, warning of threats to U.S. interests, or factors affecting opportunities for U.S. actions (without prescribing U.S. policy); **OR** (3) Satisfies a specific tasking fully.	Satisfies "good" criteria; **AND** (1) Assesses longer term, indirect, or second-order implications; **OR** (2) Provides exceptionally expert analysis (e.g., by drawing on multiple disciplines or presenting illuminating comparisons); **OR** (3) Warns of threats in detail (e.g., by discussing specific indicators, likelihood, or imminence); **OR** (4) Analyzes in detail factors affecting opportunities for U.S. action (e.g., by discussing risks, benefits, or possible reactions to potential U.S. actions).

B. Inferential Response Criteria

Solvers' submissions against the Inferential Response Criteria will be scored as the arithmetic sum of the scores for Criteria 1 and 2, as outlined below. AIS will make the ultimate determination of the winner for the **Inferential** award category.

Inferential Response Criterion 1: Properly distinguishes between factual reporting and assumptions and judgments.

Background: For the purposes of this standard, assumptions are defined as explicit or implicit hypotheses that may affect outcomes or that affect the way in which information is interpreted or weighed. They deal with identifying underlying causes and/or behavior of systems, people, organizations, states, or conditions. Assumptions comprise the foundational premises on which the information and logical argumentation build to reach analytic conclusions. Assumptions may also span information gaps that would otherwise inhibit the analysis from reaching defensible judgments. Judgments are defined as logical inferences from the available information or the results of explicit tests of hypotheses. They comprise the conclusions of the analysis.

Solvers' submissions should explicitly identify the critical assumptions on which the analysis is based and explain the implications for judgments if those assumptions are incorrect. As appropriate, Solvers' submissions should identify indicators that would signal whether assumptions or judgments are more or less likely to be correct.

Poor (0)	Fair (1)	Good (2)	Excellent (3)
Does not distinguish among statements that convey underlying information, assumptions, and judgments.	(1) Sometimes distinguishes among statements that convey underlying information, assumptions, and judgments; **OR** (2) Does not explicitly state assumptions that serve as linchpins of an argument or bridge key information gaps.	(1) Consistently distinguishes among statements that convey underlying information, assumptions, and judgments; **AND** (2) Explicitly states assumptions that serve as linchpins of an argument or bridge key information gaps.	Satisfies "good" criteria; **AND** (1) Identifies indicators that, if detected, could validate or refute judgments or assumptions; **OR** (2) Explains the implications for judgments if assumptions are incorrect.

<u>Inferential Response Criterion 2</u>: Properly expresses and explains uncertainties associated with major analytic judgments.

Background: Solvers' submissions should indicate and explain the basis for the uncertainties associated with major analytic judgments. Sources of uncertainty—including information gaps and significant contrary reporting—should be noted and linked logically and consistently to the uncertainty surrounding judgments. As appropriate, solvers' submissions also should identify indicators that would alter the levels of uncertainty for major analytic judgments.

Poor (0)	**Fair (1)**	**Good (2)**	**Excellent (3)**
(1) Does not indicate levels of uncertainty associated with major analytic judgments; **OR** (2) Indicates levels of uncertainty associated with major analytic judgments that are inconsistent with the basis ascribed to them.	(1) Indicates levels of uncertainty associated with major analytic judgments; **BUT** (2) Does not explain their basis (e.g., by reference to strengths and weaknesses of the information base, contrary reporting, assumptions, or the nature of the judgment).	(1) Indicates levels of uncertainty associated with major analytic judgments; **AND** (2) Explains their basis (e.g., by reference to strengths and weaknesses of the information base, contrary reporting, assumptions, or the nature of the judgment).	Satisfies "good" criteria; **AND** (1) Provides especially thorough discussion of nature and sources of uncertainties affecting major analytic judgments; **OR** (2) Identifies indicators that, if detected, would alter levels of uncertainty associated with major analytic judgments.

C. Evaluative Response Criteria

Solvers' submissions against the Evaluative Response Criteria will be scored as the arithmetic sum of the scores for Criteria 1 and 2, as outlined below. AIS will make the ultimate determination of the winner for the Evaluative award category.

Evaluative Response Criterion 1: Uses clear and logical argumentation.

Background: Solvers' submissions should facilitate clear understanding of the information and reasoning underlying analytic judgments. Key points should be effectively supported by information or, for more speculative warning or "think pieces," by coherent reasoning. Language and syntax should convey meaning unambiguously. Solvers' submissions should be internally consistent and acknowledge significant supporting and contrary information affecting key judgments.

Poor (0)	**Fair (1)**	**Good (2)**	**Excellent (3)**
(1) Lacks a main analytic message; **OR** (2) Does not support analytic judgments with relevant evidence or undermines them by using flawed logic; **OR** (3) Often uses unclear language or uses a structure that is not easily understood.	(1) Presents a main analytic message; **BUT** (2) Does not combine evidence, context, and assumptions effectively to support analytic judgments or uses weak logic; **OR** (3) Sometimes uses unclear language or a structure that at times is not easily understood.	(1) Presents a prominent and clear main analytic message; **AND** (2) Presents clear reasoning with no flaws in logic and effectively combines evidence, context, and assumptions to support analytic judgments; **AND** (3) Uses clear language and a structure that displays a logical flow appropriate for the argument being presented.	Satisfies "good" criteria; **AND** (1) Addresses any inconsistent or contrary information in a way that reconciles it with analytic judgments; **OR** (2) Demonstrates notable skill or sophistication in combining evidence, context, and assumptions convincingly to support analytic judgments.

<u>Evaluative Response Criterion 2</u>: Incorporates analysis of alternatives.

Background: Analysis of alternatives is the systematic evaluation of differing hypotheses to explain events or phenomena, explore near-term outcomes, and imagine possible futures to mitigate surprise and risk. Analytic products should identify and assess plausible alternative hypotheses. This is particularly important when major judgments must contend with significant uncertainties, or complexity (e.g., forecasting future trends), or when low probability events could produce high-impact results. In discussing alternatives, products should address factors such as associated assumptions, likelihood, or implications related to U.S. interests. Products also should identify indicators that, if detected, would affect the likelihood of identified alternatives.

Poor (0)	Fair (1)	Good (2)	Excellent (3)
Does not present alternatives when uncertainties, complexity, or low probability/ high impact situations warrant their inclusion.	(1) Presents alternatives when uncertainties, complexity, or low probability/high impact situations warrant their inclusion; **BUT** (2) Does not explain the evidence and reasoning that underpin them or discuss their likelihood or implications related to U.S. interests.	(1) Presents alternatives when uncertainties, complexity, or low probability/high impact situations warrant their inclusion; **AND** (2) Explains the evidence and reasoning that underpin them; **AND** (3) Discusses their likelihood or implications related to U.S. interests.	Satisfies "good" criteria; **AND** Identifies indicators that, if detected, would affect the likelihood of any identified alternatives.

Validation of Potential Winning Submissions

The award is contingent upon evaluation and validation of the submitted Solutions by the Seekers. Solvers with the highest-ranking submissions will be asked to provide source code and documentation in sufficient detail to reproduce the submitted **Analytic Product** and to enable validation of the automated system using a validation question similar to, but different than, the question posed in the Challenge. During the validation effort, Solvers may be asked to assist InnoCentive in its attempts to compile and execute the submitted system. Validation will be performed on commodity hardware running Windows or Linux with no internet access.