Mercury Challenge Prizes and Incentives

Office of the Director of National Intelligence

IARPA
BE THE FUTURE

Prize Calculation Document

MERCURY
CHALLENGE

Total Mercury Challenge Prize Purse
$100,000

Period 1
$21,000

Period 2
$79,000

6 total winners

19 total prizes

During each scoring period, Solvers will be eligible to win one (1) prize. The top scorers **who beat the baseline** will be awarded the Best Overall prize(s). Please review the Challenge Handbook Appendices for a detailed explanation on the base rate models.

Prize Eligibility Requirements

In order to receive a prize, you must do all the following:

+ Achieve a winning score, as defined in the Challenge Handbook and Mercury Challenge Prize document, during the final Period Calculation.

+ Once the final scores are posted and winners are announced, the prize winner candidates have 7 days to submit a report outlining their final algorithm explaining the logic behind and steps to its approach as well as a system diagram of your solution. You will receive a template that helps create your final report. If your report is not sufficient, you will be contacted for clarifying questions. If your report is not complete, or able to sufficiently demonstrate a working system, your submission may be rejected. Reports will be reviewed for technical merit, but will not be judged on the writing style, grammar, or language.

+ If you place in a prize-winning rank but fail to do any of the above, then you will not receive a prize. The prize will instead be awarded to the contestant with the next best performance who completed all of the above requirements.

Bonus Country / Class Category Prize Eligibility

+ After the Best Overall prize(s) have been awarded, the remaining Solvers will be matched against the Country / Class prizes available in the Period based on their overall rank order. (Ex. In Period 1, the 4th place finisher would then be matched against the remaining Country / Class prizes to see if they are eligible)

+ Matching will be based upon the best score in the Country / Class amongst the remaining Solvers who have not yet achieved a prize during the period. During the Country / Class matching, if a Solver has the top score on multiple categories, they will be awarded only one prize, and it will default to the highest prize available to them.

+ If a solver receives the highest score in multiple categories, that top solver will receive a prize for the score with the greatest difference from second place. For example, Solver A gets the highest score for Category 1 and 2. Solver A beat second place by 5 points in Category 1 and beat second place in Category 2 by 10 points. In this scenario, Solver A would win the prize for Category 2, because she won by a greater margin for that category.

Scoring Period 1 Prizes

August 7, 2018* – October 31, 2018

Note: the challenge begins on August 1, but scoring begins on August 7

Scoring Period 1 Distribution

Country	Class Type	# of Winners	Prize Amount
1st Place (All Countries)	Best Overall	1	$7,000
2nd Place (All Countries)	Best Overall	1	$5,000
3rd Place (All Countries)	Best Overall	1	$3,000
Egypt	Military Activity (MA)	1	$2,000
Saudi Arabia	Military Activity (MA)	1	$2,000
Lebanon	Military Activity (MA)	1	$2,000

Scoring Period 2 Prizes
November 1, 2018 – January 31, 2019

Note: the challenge begins on November 1, but scoring begins on November 7

Scoring Period 2 Distribution

Country	Class Type	# Of Winners	Prize Amount
1st Place (All Countries)	Best Overall	1	$21,000
2nd Place (All Countries)	Best Overall	1	$15,000
3rd Place (All Countries)	Best Overall	1	$8,000
Best Undergrad Prize (All Countries)	Best Overall	1	$5,000
Syria	Military Activity (MA)	1	$5,000
Iraq	Military Activity (MA)	1	$5,000
Egypt	Civil Unrest (CU)	1	$5,000
Jordan	Civil Unrest (CU)	1	$5,000
Infectious Disease In Saudi Arabia	Disease (CU)	1	$5,000
Midway Split – Top 10 Overall	Milestone Prize	10 @ $500 each	$5,000

Scoring Calculations

Military Activity (MA) Baseline Calculations

For MA forecasts, the effectiveness of each participant's methods will be judged using the following metrics:

- **Lead time**
  Average number of days between the date the forecast was produced and the date the actual event was reported;

- **Recall**
  The number of forecasts that matched actual events divided by the total number of actual events;

- **Precision**
  The number of forecasts that matched actual events divided by the total number of forecasts issued;

- **Quality Score**
  Average quality score of each valid forecast (ranges from 0 to 4), which is based on the forecast's accuracy with respect to event location, event date and other facets of the actual event.

Military Activity (MA) Baseline Calculations Formula

Ranking and scoring of participant submissions for MA event forecasts will be as follows:

To advance past the first evaluation gate, participant submissions must score higher than the base rate model on Precision AND Recall. This performance will be evaluated using the harmonic mean of Precision and Recall, or "F score"

$$F\ score \equiv \frac{2 \times Precision \times Recall}{Precision + Recall}$$

To advance past the second evaluation gate, participant submissions must beat a preset Lead Time of 3 days and beat the base rate model's quality score by 0.2.

Participant submissions passing the second gate will be ranked by their Quality Score(s).

Count Forecast (CU events & Disease) Baseline Calculations

For count forecasts (CU events & Disease), the effectiveness of each participant's methods will be judged using:

- **Lead time**
  Average number of days between the date the forecasted count was produced and the International Organization for Standardization (ISO) date of the actual event (day, week or month);

- **Quality Score**
  Average quality score of each valid forecast (ranges from 0 to 4), which is based on the accuracy of the count.

Count Forecast (CU & Disease) Baseline Calculations Formula

The Mercury Challenge will compare participant submissions against a "base rate" model. Base Rate models (See Handbook Appendix B) are models that only use information included in the history of observed events. It is expected that participant models will score better than the base rate models.

Ranking and scoring of participant submissions (CU events & Disease count forecasts) will be as follows:

• To advance past the first evaluation gate, participant submissions must meet the lead time requirements specified in the Challenge Handbook.

• To advance past the second evaluation gate, participant submissions must beat the base rate model's quality score.

• Participant submissions passing the second gate will be ranked by their Quality Score(s).

Got Questions or Need help?

Contact us at mercury-challenge@iarpa.gov