

# IARPA GEOPOLITICAL FORECASTING CHALLENGE

## RULES & REGULATIONS

### Contents

Background .....	2
Overview .....	2
Challenge Process .....	2
Stage 1: Registration & Configuration .....	2
Stage 2: Development, Forecasting, and Scoring .....	2
Stage 3: Solution Evaluation and Prize Awards .....	3
Overall Performance Prizes .....	3
Additional Incentives/Milestone Prizes .....	4
Payment Terms .....	4
Rules .....	5
Judging & Scoring Criteria .....	6
Intellectual Property .....	7
Eligibility .....	7
General Eligibility Requirements .....	8
Companies / Teams .....	9
Foreign Nationals and International solvers .....	9
General Liability Release .....	9
Warranties / Indemnification .....	10
Appendix: Brier Scores .....	12

## BACKGROUND

Can you create a method to forecast the future? The Geopolitical Forecasting (GF) Challenge invites Solvers from around the world to develop innovative solutions and methods for integrating crowdsourced forecasts and other data into accurate, timely forecasts on worldwide issues. The challenge presents an opportunity for individuals and teams to earn prizes by creating methods that successfully demonstrate a forecast of a wide variety of geopolitical events, such as political elections, disease outbreaks, and macro-economic indicators.

## OVERVIEW

Existing methods of geopolitical forecasting include human judgment-intensive methods, such as prediction markets, and data-intensive approaches, such as statistical models. GF Challenge Solvers will develop solutions that produce probabilistic forecasts in response to numerous closed-ended forecasting questions that concern specific, objectively verifiable geopolitical events containing timeframes with deadlines and locations. The effort will run in parallel to IARPA's most current geopolitical forecasting research program-- Hybrid Forecasting Competition (HFC). Challenge Solvers will be competing on a largely overlapping set of Individual Forecasting Problems (IFPs) as HFC research teams, and given access to the same human forecaster data stream. In addition to the provided data stream, Solvers may use other data streams and their own data and models for the challenge (in accordance with applicable laws).

## CHALLENGE PROCESS

The following three (3) stages are anticipated as part of the GF Challenge:

### Stage 1: Registration and Configuration

Solvers will register on the HeroX platform (<https://herox.com/IARPAGFChallenge>) and acknowledge their understanding and intent to follow the rules of the challenge. Following registration, instructions will be sent to access the application programming interface (API) within the Cultivate Labs platform. Solvers will be granted an API token and documentation for interacting with the challenge platform.

Stage 1 registrations will be accepted on a rolling basis, so Solvers may register at any point. However, solvers will be rewarded by accuracy scores that incentivize participation on their ability to forecast on more IFPs. Eligibility for prizes (documented below) have minimum participation requirements for the resolved IFPs included in each prize.

### Stage 2: Development, Forecasting and Scoring

Forecasting questions, known as IFPs, will be released periodically and will be open for a specified period of time. Solvers should check daily for new IFPs. While an IFP is open, Solvers may update their forecasts daily until the IFP closes. Please see the Judging and Scoring section below for the full-time periods associated with the IFP forecasting period. Each Solver will be permitted to submit forecasts using up to 25 different methodological "slots" as described in the Rules section below. Once the IFP closes, the forecasts will be locked and scores for the IFP for each method will be calculated. See the Judging and Scoring section below for further information. Scores will be tracked on the leaderboard within the Cultivate Labs platform.

In addition to the forecast submission activities, Solvers are encouraged to work to improve their forecasting methods throughout the competition.

Throughout the Challenge, there will be three (3) Milestone Check-Ins at which Milestone Prizes will be awarded. The Milestone Prizes will be awarded based on all IFPs that resolve during the Milestone duration. The Milestone schedule is as follows (milestone prizes will be validated within 30 days of the milestone end):

- Milestone 1: March 7, 2018 – April 30, 2018
- Milestone 2: May 1, 2018 – June 30, 2018
- Milestone 3: July 1, 2018 – September 7, 2018

### Stage 3: Solution Evaluation and Prize Awards

At the end of the challenge, the grand total scores will be confirmed and determined for each of the prize categories. Solvers will need to provide a write-up of their solution to the Government for final review and evaluation. This final review will determine if the solutions are sound and meet the criteria of the challenge. Solvers will provide a short (4 page) explanation for their solution. Before final prize award, IARPA may ask for additional documentation.

Upon conclusion of the Government evaluation of the documentation provided by the Solvers from the Evaluation stage, prizes will be awarded as follows.

#### Overall Performance Prizes *(Based on overall score)*

Prize	Amount	Criteria
<b>1<sup>st</sup> Place Overall</b>	\$20,000	<ul style="list-style-type: none"> <li>• Must participate in at least 70% of the IFPs. Must have positive NBPs on more than half of the IFPs you attempt</li> <li>• Uses the single best scoring method from all of the Solver's available methods across all IFPs</li> </ul>
<b>Bonus: Ultimate Forecaster</b>	\$40,000	<ul style="list-style-type: none"> <li>• Must achieve the 1<sup>st</sup> Place Overall Prize</li> <li>• Must beat the GF Challenge Baseline</li> <li>• Must beat the HFC Top Score</li> </ul>
<b>2<sup>nd</sup> Place Overall</b>	\$12,000	<ul style="list-style-type: none"> <li>• Must participate in at least 70% of the IFPs. Must have positive NBPs on more than half of the IFPs you attempt</li> <li>• Uses the single best scoring method from all of the Solver's available methods across all IFPs</li> </ul>
<b>3<sup>rd</sup> Place Overall</b>	\$10,000	
<b>4<sup>th</sup> Place Overall</b>	\$7,000	
<b>5<sup>th</sup> Place Overall</b>	\$4,000	
<b>Bonus: Star Forecasters</b>	\$30,000 split amongst all eligible Solvers/teams	<ul style="list-style-type: none"> <li>• Must achieve 2<sup>nd</sup> – 5<sup>th</sup> Place Prizes</li> <li>• Must beat the GF Challenge Baseline</li> <li>• Must beat the HFC top score</li> <li>• Bonus will be divided among the qualifying winning Solvers / teams according to the overall place as follows: <ul style="list-style-type: none"> <li>○ 4 winners: 50%, 25%, 15%, 10%</li> <li>○ 3 winners: 50%, 30%, 20%</li> <li>○ 2 winners: 70%, 30%</li> <li>○ 1 winner: 100%</li> </ul> </li> </ul>

#### Definitions of prize criteria:

- **Net Brier Points (NBPs)** – See the Judging and Scoring section point 3 for a description of Net Brier Points.
- **GF Challenge Baseline:** Challenge solvers will have access to a stream of de-identified individual-level human forecast data, and an aggregate stream of human forecast data. This aggregate will be constructed using a method based on the state-of-the art from on the IARPA Aggregative Contingent Estimation (ACE) Program. The aggregate

measure will serve as the baseline against which NBPs will be calculated. This baseline will be available to solvers via the Cultivate Labs platform.

- **HFC Top Score:** The top scoring method coming out of the parallel HFC Perform teams will be visible on the leaderboard during the challenge.

### *Additional Incentives/Milestone Prizes*

Prize	Amount	Criteria
<b>Best in Domain / Region Pair</b>	\$25,000 (5 awards of \$5,000 each)	<ul style="list-style-type: none"> <li>• Top score on each of the five (5) domain/region pairings at the end of the challenge</li> <li>• Minimum participation in 80% of the total IFPs in the domain/region pairing</li> <li>• Uses the best scoring method to achieve the domain/region pairing, regardless of the Solver's overall method used in final scoring</li> </ul>
<b>Best Undergraduate</b>	\$4,000	<ul style="list-style-type: none"> <li>• Best score by an undergraduate, or all-undergraduate team who did not place in the Top five (5)</li> <li>• Minimum participation in more than half of the total challenge IFPs</li> <li>• Certification requirements (transcript, certified letter) from current institution must be presented in order to claim this prize</li> </ul>
<b>Milestone 1</b>	\$7,500 (10 awards of \$750 each)	<ul style="list-style-type: none"> <li>• Top ten Solvers with the highest score in the Milestone will each get \$750</li> </ul>
<b>Milestone 2</b>	\$10,000 (10 awards of \$1,000 each)	<ul style="list-style-type: none"> <li>• Top ten Solvers with the highest score in the Milestone will each get \$1,000</li> </ul>
<b>Milestone 3</b>	\$12,500	<ul style="list-style-type: none"> <li>• The specific breakdown for distribution of Milestone Prize funds will be announced prior to the start of the Milestone 3 start date</li> </ul>
<b>Election Forecaster</b>	\$10,000	<ul style="list-style-type: none"> <li>• Top Solver in election-related IFPs at the end of the challenge</li> <li>• Minimum participation in 80% of the total IFPs in the Elections topic</li> <li>• Based on Solver's best scoring method for Elections IFPs, regardless of Solver's best overall method</li> </ul>
<b>Spring Forecaster</b>	\$2,400 (3 awards of \$800 each)	<ul style="list-style-type: none"> <li>• Top three Solvers with the highest score on IFPs that open during the month of April and resolve by June 20</li> <li>• Minimum participation in 80% of the total IFPs in the relevant interim category</li> <li>• In the event of a tie, winning Solvers will split the prize purse evenly among the top scores</li> </ul>
<b>Interim Prize</b>	\$2,600	<ul style="list-style-type: none"> <li>• Additional incentives will be provided throughout the challenge and at the discretion of the challenge team. Prize amounts could be applied to milestone, domain/region pairings, or additional bonus incentives to the Top 5 Overall Performance Prizes. These Interim Prizes will be announced prior to the start of the Prize Calculation.</li> <li>• Minimum participation in 80% of the total IFPs in the relevant interim category</li> </ul>

### *Payment Terms*

Solvers will need to submit a W-9 tax form, or a W8-BEN form in order to receive payment. Solvers are responsible for payment of all taxes incurred from the acceptance of Prize funds. Solvers will be given 30 days to submit their paperwork upon notification of award. IARPA and Booz Allen are not responsible for lost or stolen prize payments, or incorrect routing and payment information provided by the winning Solvers.

## RULES

- All Solvers will need to fill out a registration form providing information on their team and asserting eligibility to participate through the HeroX GF Challenge page. (<https://herox.com/IARPAGFChallenge>)
- All daily system forecast submissions must be made via the Cultivate Labs GF Challenge API and authenticated using the API token provided during the registration process. All submissions must match the format described in the API document available on the Cultivate Labs Platform.
- IFPs and the human forecast stream will only be provided to Solvers via the Cultivate Labs GF Challenge API.
- GF Challenge forecast days begin at 14:01 U.S. Eastern Time (Standard or Daylight Savings as appropriate). The daily reporting deadline is 14:00:59 U.S. Eastern Time.

Each day's final/official forecast (due 14:00:59 ET) will be treated as a prospective forecast that covers the immediately following 24-hour period, ending at 14:00:59 ET the next calendar day. The final forecast for an IFP is defined to be the last forecast submitted on the last forecasting day before the IFP is resolved.

Suppose IFP closes as scheduled at 23:59:59 ET on 31 December, then a forecast at 14:00:58 ET that day would count but a forecast at 14:01:01 ET would not count. If an IFP closes early, say 09:00:00 ET on 30 December then a forecast at 14:00:58 ET on 29 December would count but a forecast at 14:01:01 ET on 29 December would not count.

- Each Solver is permitted to submit daily forecasts for up to 25 official competition forecasting methods or permutations (aka "slots") at any given time. Solvers shall provide names of 50 characters or fewer for their forecasting methods, which must be used consistently throughout the challenge period. These names may not include profane or derogatory terms.
- Solvers may submit more than one forecast update per IFP per day. For each forecasting method, only the last forecast for an IFP submitted before the daily submission deadline of a given forecast day will be treated as the Solver's official forecast for each IFP for that forecast day.
- Each Solver's forecasting method will be scored across all IFPs for which forecasts are submitted, and an overall forecast accuracy score for that method will be determined. For each Solver, the Testing and Evaluation (T&E) Team will identify the single method with the best score across all IFPs as the Solver's "best" method. In other words, each Solver's best method will be a unitary, prospectively defined method selected from among the Solver's 25 official methods: We will not "cherry-pick" per-IFP most accurate methods to formulate composite, per-Solver "best" methods.
- Once a Solver submits their first forecast on an IFP they should provide daily forecast updates. If they do not, their previous forecast will be carried over each forecasting day until a new forecast is submitted.
- For forecast days prior to the first submission by a Solver on an IFP, the assumed forecast is the aggregated human forecaster stream forecast for that IFP.
- For Milestone Prizes, only the IFPs that are resolved during that Milestone's time period will count for that Milestone's prize. IFPs that are opened, but not resolved, during a Milestone will not count. They will be counted towards the Milestone period in which the IFP resolves.



- In order to be eligible domain/region or interim prizes a Solver must submit forecasts for at least 80% of the IFPs considered for that prize. In order to be eligible for one of the Top five (5) overall performance prizes, a Solver must submit forecasts for at least 70% of the total number of IFPs in the challenge.
- Solvers are responsible for ensuring that submitted forecasts are compliant with the standards documented within the API documents on the Cultivate Labs Platform. Forecast submissions must be proper in that the probabilities assigned to the several outcomes for an IFP must sum to 1.0. Improper, malformed or non-compliant forecast submissions will be rejected. IARPA reserves the right to change its policies with reasonable notice, including disqualification of non-compliant data.
- Solvers may use external data streams in their solutions. However, these external data sources will need to be documented. The Solver is responsible for ensuring that they have the necessary rights and licenses for the use of such external data sources for the purposes of this challenge.
- The Human Forecaster data stream provided through the challenge is available for Solvers to use, free of charge, only for the purposes of this challenge. All other uses are prohibited without the express consent of IARPA. However, Solvers are not required to use this data stream as part of their solution.
- Solvers are solely responsible for ensuring they have valid licenses or rights to any software libraries or data sources used in their forecasting systems. Any method that infringes on another party's intellectual property will be disqualified from the competition.
- Solvers may only register once for the challenge and can compete either as an individual or as part of a team; Solvers cannot participate as both an individual and as part of a team.

## JUDGING AND SCORING

1. Your competition score is based on how accurate your forecasts are when compared with the stream of Government-furnished human forecasts. The Government-furnished forecast stream (GFFS) is a "crowd wisdom" aggregation of asynchronously submitted human forecasts for each IFP. Given an IFP we compute the consensus forecast as a function of the weighted average of the latest forecasts from a portion of the unique forecasters who most recently provided forecasts for this IFP. The weight assigned to each forecaster is computed from three factors. The first two are based on their performance across all IFPs: historical forecast accuracy for that forecaster and frequency with which that forecaster submits updates. The third factor is whether or not they have completed a brief, self-administered forecast training course. The weighted average of the recent forecasts is then extremized toward 0 or 1 using a tuning parameter. This method for computing the consensus forecast was determined to have good performance during the IARPA ACE Program and is a suitable benchmark against which to compare solver performance.
2. We measure absolute forecast accuracy using a formula called the Brier score, which reflects the squared difference between the probabilities a solver or the GFFS assigns to each outcome and reality (see appendix section A.1). Lower Brier scores reflect greater accuracy: A perfect score is 0 (for instance, on a binary question, you assigned 100% to the correct outcome), the worst possible score is 2 (for instance, on a binary question, you assigned 100% to an incorrect outcome). For some IFPs the possible answers have an ordering such that some answers are closer to being correct than others. In this case, we use a modified version of the Brier score that awards more credit when assigning higher forecast probabilities to answers that are closer to the correct answer (see Appendix: A.2 Application of Brier Scoring to Ordered Outcomes). The appendix provides a complete description of generic and ordered Brier scoring.

3. For a given IFP, we compare the accuracy of your forecasts to those of the GFFS using Net Brier Points (NBPs), defined as the difference between the GFFS's Brier score and your Brier score for each day that the IFP is open. Negative NBPs means the GFFS was more accurate than you, positive NBPs means you were more accurate than the GFFS, and more NBPs is better. For days that pass before your first forecast on the IFP, your forecast is assumed to be the same as the GFFS consensus forecast, meaning you do not accrue (or lose) NBPs for a given IFP until you submit your first forecast. Once you submit an initial forecast, that forecast holds each day thereafter until/unless you update it. Your score for the IFP is then the sum of your NBPs for all days that an IFP is open.
4. For most of the prize categories your score is the sum of your NBPs over all germane IFPs for that category. However, the bonus Ultimate Forecaster and Star Forecasters Prizes are scored slightly differently. These bonus prizes feature a comparison of your method to the best of the performers in the IARPA HFC program on the same IFPs. For the purposes of this bonus prize category we will compute your Mean Daily Brier Score (see Appendix: A.3 Mean Daily Brier Score ) for all IFPs and take the mean of those scores across those IFPs. If your mean Mean Daily Brier score is *lower* than the best HFC program method you will be eligible for this bonus prize category.
5. To be eligible for overall prizes you must attempt at least 70% of all IFPs and you must have positive NBPs on more than half of the IFPs you attempt. For example, suppose there are 150 IFPs for the overall prize category. Suppose that you submit non-trivial forecasts on 105 of these IFPs. Furthermore, let us say that you have positive NBPs on at least 53 IFPs. You will meet all of the eligibility criteria for overall prizes.
6. In order to be eligible for a prize the Solver must document the solution method in a short document (no more than four pages) using the provided format (coming soon). Please note that IARPA will be examining the substance and clarity of the technical details documented, not the quality of the writing.
7. Solutions competing for prizes will be verified by the Solver or Team that all data sources and software libraries can be identified through the solution with proper attribution where necessary.

## INTELLECTUAL PROPERTY

Solvers will retain the Intellectual Property (IP) rights to their solutions. IARPA will use information submitted through the challenge and application process to understand the technology.

- IARPA will have the rights to utilize all data, forecasts, and method explanations provided for U.S. Government purposes.
- Challenge participant names, titles, general technology descriptions, photographs, and abstracts for their submissions may be utilized in challenge-related media and promotional materials or for other internal Government uses. No sensitive intellectual property information will be shared in this manner, and must be clearly marked prior to submission.
- Challengers will need to sign the IP Agreement from within the HeroX GF Challenge page prior to participating in the challenge.

## ELIGIBILITY

Some individuals or organizations may not be eligible for prizes for reasons listed below. However, these Solvers may, upon IARPA approval, participate in the challenge and be eligible for ranking on the leaderboard. Approval will need to be obtained in advance of participation by emailing [gfcchallenge@iarpa.gov](mailto:gfcchallenge@iarpa.gov).

Any teams who forego prizes will be publicly listed on the HeroX GF Challenge page to ensure full transparency of their competition status.

## General Eligibility Requirements

To be eligible to win a prize under this competition, an individual or entity:

1. Must have completed and submitted a registration form at HeroX GF Challenge page
2. Must be (1) an individual or team each of whom are 18 years of age and over, or (2) an entity incorporated; and
3. May not be a federal entity or federal employee acting within the scope of their employment. An individual or entity shall not be deemed ineligible because the individual or entity used federal facilities or consulted with federal employees during a competition if the facilities and employees are made available to all individuals and entities participating in the competition on an equitable basis.

Federal grantees may not use federal funds to develop challenge applications unless consistent with the purpose of their grant award. Federal contractors may not use federal funds from a contract to develop challenge applications or to fund efforts in support of a challenge submission.

Employees of IARPA, Booz Allen Hamilton, Cultivate Labs, HeroX, Good Judgment, Inc., and MITRE are not able to compete in the challenge. Companies and contractors who are currently supporting, or who have previously supported, the above (or other) entities in efforts related to the IARPA HFC program are also not able to participate in the challenge. This restriction extends to members of such persons' immediate families (spouses, children, siblings, parents), and persons living in the same household as such persons, whether or not related.

Federally Funded Research & Development Centers (FFRDCs) and (DoD) University Affiliated Research Centers (UARC)s not mentioned above may be eligible to submit forecasts and receive leaderboard recognition, but are not eligible to win prize dollar awards. In order to compete, an email must be sent to [gfchallenge@iarpa.gov](mailto:gfchallenge@iarpa.gov) with the team name, organization represented, and team member information.

IARPA's Hybrid Forecasting Competition (HFC) performer team members and HFC Good Judgment Forecasters, as well as members of such persons' immediate families (spouses, children, siblings, parents), and persons living in the same household as such persons, whether or not related, are not eligible to participate in the competition.

Federal employees, employees of FFRDCs, and employees of UARC)s can participate in the challenge as a private citizen, not affiliated with their organization, as long as (a) the Solver is not employed by an organization that is actively involved in the challenge or closely related programs; (b) the Solver has received approval from their organization, if applicable, to participate in their personal capacity (note, it is the Solver's responsibility to verify this); and (c) Solvers cannot use government or their organizations' resources, computers, or access to government information in aid of their participation, except for those resources available to all other participants on an equal basis.

Entrants must agree to assume any and all risks and waive claims against the Federal Government and its related entities, except in the case of willful misconduct, for any injury, death, damage, or loss of property, revenue, or profits, whether direct, indirect, or consequential, arising from their participation in a competition, whether the injury, death, damage, or loss arises through negligence or otherwise.



Entrants must also agree to indemnify the Federal Government against third-party claims for damages arising from or related to competition activities. Entrants are not required to obtain liability insurance or demonstrate financial responsibility in order to participate in the competition.

By participating in the competition, each entrant agrees to comply with and abide by these rules and the decisions of IARPA and/or the individual judges, which shall be final and binding in all respects.

By participating in the competition, each entrant agrees to follow all applicable local, state, federal and country of residence laws and regulations.

## Companies/Teams

Companies, universities, individuals, and Teams of Solvers are able to participate in this challenge. Companies / Teams will need to elect a Team Captain, who will be the main point of contact (POC) for communications. Individuals who have registered on the HeroX GF Challenge page as an individual competitor, will not be eligible to participate in a team. All Team Members must meet the General Eligibility Requirements, self-register, and acknowledge these rules through the HeroX GF Challenge page (<https://herox.com/IARPAGFChallenge>).

- Upon registering for the challenge, the Team Captain will need to provide a breakdown of all Team Members along with the percentage allocation between all Team Members
- Each Team Member will need to sign an agreement that they agree to distribution of prize funds
- A Team Captain may provide information for a company or organization instead of their personal information for tax purposes.
- One representative from the Team, will need to set up their Team's API. There will only be one API allowed per Team for the entire duration of the challenge
- Once a Team has submitted their first forecast, the team members cannot change. This means that no new Solvers can be added to a Team and no Solvers can leave the Team.
- Individual Solvers who have submitted a forecast on their own may not form a Team after their first forecast has been submitted.

## Foreign Nationals and International Solvers

Except for residents of Iran, Cuba, North Korea, Crimea Region of Ukraine, Sudan or Syria or other countries prohibited on the U.S. State Department's State Sponsors of Terrorism list, Solvers residing in any country are able to participate. In addition, Solvers are not eligible to participate if they are on the Specially Designated National list promulgated and amended, from time to time, by the United States Department of the Treasury. It is the responsibility of the Solver to ensure that they are allowed to participate in this challenge and allowed to export their technology solution to the United States.

## GENERAL LIABILITY RELEASE

By participating in the competition, each entrant hereby agrees that:

1. IARPA and the U.S. Government, Booz Allen Hamilton, Cultivate Labs, HeroX, and MITRE shall not be responsible or liable for any losses, damages, or injuries of any kind (including death) resulting from

participation in the competition or any competition-related activity, or from entrants' acceptance, receipt, possession, use, or misuse of any prize; and

2. Entrants will indemnify, defend, and hold harmless IARPA and the U.S. Government, Booz Allen Hamilton, Cultivate Labs, HeroX, and MITRE from and against all third-party claims, actions, or proceedings of any kind and from any and all damages, liabilities, costs, and expenses relating to or arising from entrant's participation in the competition.

Without limiting the generality of the foregoing, IARPA and the U.S. Government, Booz Allen Hamilton, Cultivate Labs, HeroX, and MITRE are not responsible for incomplete, misdirected, prize notifications; or for late, lost, inaccessible, network connections, or for issues with Internet Service Providers, websites, or other connections and third-party sites that the entrants choose to use for the competition; or for any technical malfunctions, failures, difficulties, or other errors of any kind or nature.

These rules cannot be modified except by IARPA. All decisions by IARPA regarding adherence to these rules are final. The invalidity or unenforceability of any provision of these rules shall not affect the validity or enforceability of any other provision. In the event that any provision is determined to be invalid or otherwise unenforceable or illegal, these rules shall otherwise remain in effect and shall be construed in accordance with their terms as if the invalid or illegal provision were not contained herein.

## WARRANTIES / INDEMNIFICATION

By participating in the competition, each Solver represents, warrants, and covenants as follows:

1. The Solver – whether an individual, team or entity – is the sole author, creator, and owner of the submission;
2. The submission is not the subject of any actual or threatened litigation or claim;
3. The submission does not and will not violate or infringe upon the intellectual property rights, privacy rights, publicity rights, or other legal rights of any third party;
4. The entrant will not interfere with or attempt to influence, the events that are the subject of the IFPs in the challenge. If a Solver has a question about an activity related to their job role, please email [gchallenge@iarpa.gov](mailto:gchallenge@iarpa.gov) for clarification.
5. The Solver will not engage in activities that could cause political unrest, or lead to a violent event or civil unrest that could cause the outcome of the IFP to change.; and
6. The Submission, and Solvers' use of the Submission, does not and will not violate any applicable laws or regulations, including, without limitation, applicable export control laws and regulations of the U.S. and other jurisdictions.

If the Submission includes any third party works (such as third-party content, data sets, algorithms, or open source code), entrant must be able to provide, upon the request of IARPA, documentation of all appropriate licenses and releases for such third-party works. If entrant cannot provide documentation of all required licenses and releases, IARPA reserves the right to disqualify the applicable Submission, or seek to secure the licenses and releases for the

benefit of IARPA, and allow the applicable Submission to remain in the Competition. IARPA also reserves all rights with respect to claims based on any damages caused by participant's failure to obtain such licenses and releases.

Solvers – whether an individual, a team or an entity – will indemnify, defend, and hold IARPA, Booz Allen Hamilton, Cultivate Labs, HeroX and MITRE from and against all third-party claims, actions, or proceedings of any kind and from any and all damages, liabilities, costs, and expenses relating to or arising from entrant's submission or any breach or alleged breach of any of the representations, warranties, and covenants of entrant hereunder.

IARPA reserves the right to disqualify any submission that IARPA, in its discretion, deems to violate these Rules. IARPA reserves the right to disqualify any Solver that is operating in a manner that disrupts the competition, does not comply with the spirit of the competition, or is found to be manipulating the competition. IARPA also reserves the right to amend these rules throughout the duration of the contest should extenuating circumstances arise.

## APPENDIX: BRIER SCORES

This appendix provides the details for scoring a single daily forecast against a resolved IFP.

### A.1 Generic Brier Score Formula

We will employ the Brier score (aka Mean Quadratic Score or MQS) as our primary forecast accuracy scoring rule. This proper scoring rule is applicable to probabilistic forecasts offered in discrete choice scenarios (i.e., How likely is each specified outcome?). It can be calculated as:

$$\sum_{j=1}^r (f_j - o_j)^2$$

where:

- $f_j$  = the forecast probability for outcome option  $j$  and  $o_j = 1$  (if outcome option  $j$  comes to pass) or 0 (otherwise).
- $r$  = the number of possible forecast outcomes

These scores can range from 0 to 2, with lower scores indicating superior accuracy (i.e., minimal deviation from observed outcomes).

Readers will note that for binary forecasting problems this score is equivalent to a widely-used version of the Brier score that ranges from zero to one (0-1), but unlike that version, Brier's (1950) original formula is also applicable to unordered multinomial probabilistic forecasts.

#### A.1.1. Example

Consider the forecasting question "Will the incumbent party retain control in the next parliamentary elections in Freedonia? (Yes/No)". You issue a forecast with a probability for "Yes" = 0.7 and the probability for "No" = 0.3. If the incumbent party retains control (Reality = "Yes") your score is  $(1 - 0.7)^2 + (0 - 0.3)^2 = .09 + .09 = .18$ . However, if the reverse occurs (Reality = "No") your score is  $(0 - 0.7)^2 + (1 - 0.3)^2 = .49 + .49 = .98$ . Recall that for core Brier scores, smaller values are better.

### A.2 Application of Brier Scoring to Ordered Outcomes

Some forecasting questions require the assignment of probabilities to ordered categories. Consider, for example, a question whose outcome possibilities are temporal intervals:

- a. 20 days or less
- b. Between 21 and 40 days
- c. Between 41 and 60 days
- d. 61 days or more.

We call these ordered individual forecasting problems, or oIFPs.



In this example, if the event occurs during period B, the following distribution

D1: A – 25%, B – 25%, C – 50%, D – 0%

is in principle more accurate than

D2: A – 25%, B – 25%, C – 30%, D – 20%

because D1 is uniformly closer to the truth than D2, even though both assigned  $p=0.25$  to the correct interval.

Jose, Nau and Winkler (2009) identify extensions of binary and unordered multinomial scoring rules to ordered categorical forecasts. Their approach satisfies the following criteria.

1. If a distribution is uniformly closer to ground truth than a second distribution, then the first distribution receives a better score.
2. The scoring rule is strictly proper.
3. The scoring rule is commensurate with the Brier score in terms of scaling and interpretation.

The basic procedure is as follows.

Step 1 – Take the original categories (A-B-C-D) and break them up into a set of cumulative, binary categories (A-BCD; AB-CD; ABC-D)

Step 2 – Apply the scoring rule to each binary pair

Step 3 – Take an average across the binary pair scores.

In the above example, the standard unordered Brier scores for D1 and D2 are 0.88 and 0.76, respectively<sup>1</sup>. Note that even though D1 was closer to ground truth than D2 it received a worse score. This occurred because the Brier score increases as the distribution over the false categories deviate from uniform. In general, we would expect that any forecast system that assigns a high probability to a near miss will be penalized using the original Brier score.

In contrast to the unordered results, the 'ordered' Brier scores for D1 and D2 are 0.21 and 0.24 respectively<sup>2</sup>. The closer distribution receives a better score, which is what we would desire in this scenario.

### A.3 Mean Daily Brier Score

The procedures in A.1 and A.2 provide the score for a single day's forecast against the IFP outcome. However, most IFPs are open for more than one day. The Mean Daily Brier (MDB) score is the average of the daily Brier scores across the life of the IFP. For those days prior to the first affirmative forecast submission there is an implicit forecast, imputed according to a predefined rule. For this challenge, the imputed forecast is the GFFS. Once you have submitted your first forecast for an IFP that forecast is assumed to apply to subsequent forecast days until you submit a new forecast. If you never submit a forecast on an IFP then your MDB will be the same as the MDB for the GFFS.

<sup>1</sup> For D1, Brier =  $((1-0.25)^2) + ((0-0.25)^2) + ((0-0.50)^2) + ((0-0.00)^2) = 0.875$ . For D2, Brier =  $((1-0.25)^2) + ((0-0.25)^2) + ((0-0.30)^2) + ((0-0.20)^2) = 0.755$ .

<sup>2</sup> For D1, Brier for A-BCD =  $((1-0.75)^2) + ((0-0.25)^2) = 0.125$ , Brier for AB-CD =  $((1-0.50)^2) + ((0-0.50)^2) = 0.5$ , Brier for ABC-D =  $((1-1.00)^2) + ((0-0)^2) = 0.00$ . Ordered Brier is the average of these: 0.208. For D2, Brier for A-BCD =  $((1-0.75)^2) + ((0-0.25)^2) = 0.125$ , Brier for AB-CD =  $((1-0.50)^2) + ((0-0.50)^2) = 0.5$ , Brier for ABC-D =  $((1-0.80)^2) + ((0-0.20)^2) = 0.08$ . Ordered Brier is the average of these: 0.235.