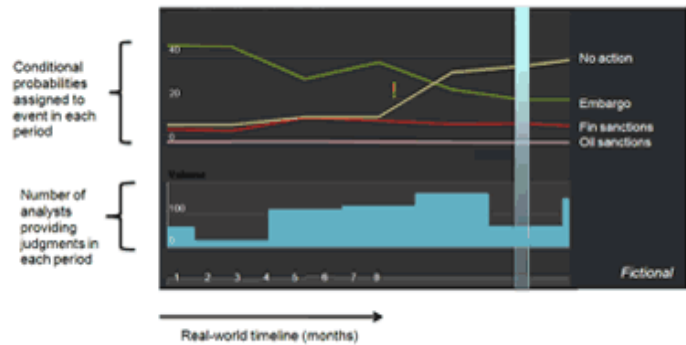


# The Aggregative Contingent Estimation Program: Predicting Global Events Through Crowdsourcing

## Case Study Overview

The goal of the *Aggregative Contingent Estimation Program*, sponsored by the Intelligence Advanced Research Projects Activity, is to enhance the accuracy, precision and timeliness of forecasts for a broad range of global events. The program develops advanced techniques to gather, weight and combine the judgments of people from many backgrounds and fields and in many different locations.



ACE is powered by human judgment, which makes it flexible enough to provide forecasts on just about any type of intelligence-forecasting question. Launched in 2010, ACE is based on the idea that combining forecasts made by an informed and diverse group of people often produces more accurate predictions of future events than those made by a single expert.

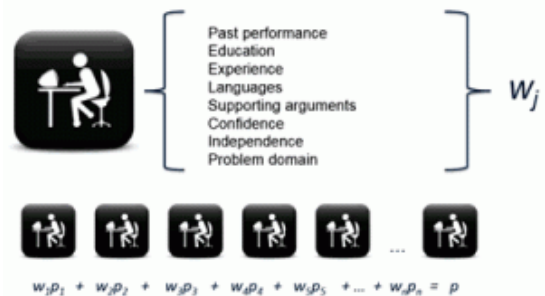
## Project Description

ACE started with a “forecasting tournament.” Five teams of leading scholars from industry and academia competed to forecast events. They recruited thousands of research participants; each year, the participants answered about 100 questions related to social, economic and political events. Every day, the teams sent forecasts to an independent evaluator, who scored them based on actual outcomes. Each research team tried to produce the most accurate forecasts, competing against each other and against a benchmark group that used the unweighted average judgment of a group of forecasters.

After two years, one research team — the Good Judgment Project—substantially outperformed the others. In fact, Good Judgment’s improvement in accuracy was greater than the improvement of the other four research teams combined — about 70 percent over the benchmark. Forecast improvement was measured using Brier scoring, a method originally developed to evaluate weather forecasts.

## Challenges

Four of the five teams had difficulty recruiting and retaining the number of people they needed, because continuous forecasting was somewhat time-consuming (taking about one hour per week). Teams also had to decide how best to use project



resources — and whether to focus most of their effort on finding the best ways to efficiently collect probability judgments; on determining how best to combine and weight those judgments; or on developing training methods for forecasters.

- Collecting judgments involved finding the best way to gather the needed range and number of probabilistic beliefs from a crowd of individuals — whether by surveys, by prediction markets, or by some other technique — and then producing the most intuitive and user-friendly interfaces for these platforms.
- Combining judgments involved developing new algorithms to create the most accurate aggregated forecasts.
- Training involved teaching forecasters the skills that would help them become more accurate and less susceptible to judgmental biases or poor decisionmaking.

In addition, the project's initial concept faced resistance from potential participants and customers. Analysts are not often trained to think in quantitative terms and may be reluctant to provide numerical forecasts that can be scored for accuracy. However, letting forecasters be anonymous made it easier for them to take the risk and to take the time to develop the skills needed.

## **Benefits and Outcomes**

The team that won the *ACE* tournament (the Good Judgment Project) made substantial advances in all three areas:

- **Collecting judgments:** Given the advanced algorithms generated in *ACE*, opinion surveys surpassed prediction market platforms as the best way to elicit probabilistic judgments from forecasters.
- **Combining judgments:** Promising new algorithms weighted individual survey responses based on past accuracy, then pushed up some probability judgments (for example, an average prediction of 70 percent might be pushed to 90 percent if the beliefs of previously accurate forecasters warranted it). This dramatically increased the accuracy of the combined judgments.
- **Training forecasters:** The team created a one-hour online class that improved individual forecaster accuracy by about 10 percent.

*ACE* shows that meaningful geopolitical forecasts can be produced quickly and accurately on topics ranging from violent international confrontations to how long international leaders will stay in power. By better measuring exact levels of uncertainty, the project can also increase the rigor of intelligence analysis more generally. For the first time, we have a quantitative system flexible enough for rapid analysis of almost any subject. Where traditional analysis can take days or weeks, *ACE* forecasts can be obtained in a matter of hours. Consumers of *ACE* forecasts can be confident in their accuracy because the technologies have been validated in a real-world forecasting tournament.

## Tips

The *ACE* case study illustrates the following steps in the Federal Citizen Science and Crowdsourcing Toolkit:

- **Design a Project — Know Your Objectives**  
The objective in this case was to improve forecasting by more than 50 percent over the state-of-the-art forecasts. Choosing a clear, measurable target and having a state-of-the-art control group as a benchmark enabled progress to be clearly gauged. By the program's end, the Good Judgment Project beat the state-of-the-art forecasts by more than 70 percent. Setting specific quantitative performance benchmarks is a hallmark of all IARPA programs, and *ACE* was no different.
- **Build a Community — Engage Your Community**  
The Good Judgment Project recruited and retained an impressive pool of high-quality participants. Participants were highly credentialed (some 60 percent had graduate degrees) and tenacious. Many spent dozens of hours per week forecasting, even though they were paid only a couple hundred dollars per year in Amazon gift cards. The Good Judgment Project understood its pool of participants, providing ongoing feedback on individual accuracy to encourage participation and ongoing effort. The project recognized and rewarded exceptional forecasters as “superforecasters.” The result was a uniquely engaged and loyal group of participants.

## Learn More

- Website: [ACE](#)
- [Principles of Community Engagement, NIH \(2011\)](#)
- [Using Crowdsourcing in Government, IBM Center for The Business of Government \(2013\)](#)
- [Community Engagement Techniques: Best Practices \(2014\)](#)
- [Crowdsourcing: A Geographic Approach to Public Engagement, SSRN \(2014\)](#)

## Contact Information

Dr. Steve Rieber  
Email: [steven.rieber@iarpa.gov](mailto:steven.rieber@iarpa.gov)