**CDC is requesting information about your forecast methodology be provided in a standard format to better capture the types of data and methodologies being used to forecast influenza. Please complete the form for your team’s forecast by November 16, 2018 (or, if this form is being used for the hospitalization challenge, within two weeks of the start of the challenge). If your team submits more than one forecast each week, please complete a separate form for each one. If at any point during the challenge a significant change to the forecast data sources or methodology is made, please complete a new form.**

**1. Date: 11/4/18**

**2. Team name: Fred + Harvard University**

**3. Challenge:**

FluSight

Hospitalization (FluSurv)

State ILI

**4. Model name (if different from team name or if more than one forecast; no more than 25 characters):**

**ARGO**

**5. Brief narrative describing methodology (no more than 50 words):** Our method combines information from electronic health records (EHR) and Internet users’ search activity (Google) with historical flu activity to produce flu forecasts for the United States. Our methodology is based on a regularized multivariate regression model that dynamically selects the most appropriate variables for optimal flu estimates and automatically identifies which variables are important for prediction in every week.

**6. Please provide any references (published article, website, etc.):**

[Yang, S., Santillana, M. & Kou, S. C. Accurate estimation of influenza epidemics using Google search data via ARGO. *Proc. Natl. Acad. Sci. U. S. A.* **112,** 14473–14478 (2015).](http://paperpile.com/b/A85R9i/k0x1)

[Yang, S. *et al.* Using electronic health records and Internet search information for accurate influenza forecasting. *BMC Infect. Dis.* **17,** 332 (2017).](http://paperpile.com/b/A85R9i/25Qm)

[Santillana, M. *et al.* Cloud-based Electronic Health Records for Real-time, Region-specific Influenza Surveillance. *Sci. Rep.* **6,** 25732 (2016).](http://paperpile.com/b/A85R9i/aWfT)

**7. Model components:**

Please indicate which of the following data source(s) your team used in developing your forecast. Check all that apply.

CDC FluSurv-NET

CDC ILINet at the national or regional level

CDC ILINet at the state level

CDC webpage visits

Commute and transportation (e.g., flight) data

Electronic/Cloud-based health records

Google/Internet-search query information

HealthTweets/Twitter

NCHS Pneumonia and Influenza Mortality Surveillance

Online news/outbreak reporting (e.g. ProMed, HealthMap)

School vacation calendars

Weather/climate data

☐ Wikipedia

☐ WHO/NREVSS virologic data

Other. Please specify:

Please indicate the methodological approaches used to create this forecast. Check all that apply.

Crowdsourcing

Machine-learning approaches

Mechanistic model at the individual level (e.g. agent-based model)

Mechanistic model at the population level (e.g. compartmental model)

Metapopulation/Network-based approaches

Statistical/descriptive/phenomenological models (i.e. models that do not directly account for disease transmission dynamics)

Other. Please specify:

Is this forecast the product of an ensemble?

Yes

No

Has this forecast been submitted in previous forecasting challenges?

Yes. Please indicate which influenza season(s): 2016-17

No