

## **The AAPI Database: Data Documentation**

**Purpose:** The TAAF Database Website will display data for AAPI-serving organizations, which will be collected from various sources. The purpose of this document is to outline those data sources, specify collection methodologies, and clarify how the data points between these data sources differ. Broadly, the data sources we use can be split into two categories: (1) data reported from organizations via Give in May registration or the TAAF data intake form (“self-reported data”), and (2) 990 organization data collected from GuideStar (“GuideStar data”).

### **1. Self-reported data**

#### **a. Give in May (GIM) registration:**

- i. **Summary:** GIM is a fundraising campaign that supports the work of AAPI organizations. In order to participate in this campaign, organizations directly report data to APF, and this data is then used to populate that organization’s profile page on the GIM website. APF has agreed to provide 2022 and 2023 GIM registration data to TAAF for those organizations who “opt-in” to the TAAF database.
- ii. **Scope determination:** the number of organizations participating in GIM are largely determined by the outreach done by APF. Specifically, eligible participants must meet the following requirements:

1. IRS registered 501(c)(3) nonprofit organization or fiscally sponsored by a 501(c)(3).
2. 51% of those served are Asian American and/or Pacific Islander.
3. Minimum of two years in operation.

#### **iii. Collected data fields:**

1. Organization Name
2. Employer Identification Number (EIN)
3. Main Contact Name and Email
4. Executive Director Name and Email
5. Mailing Address
6. Year Founded
7. Year Received IRS 501(c)(3) Status
8. IRS Tax Period
9. IRS Revenue Amount
10. Current Operating Budget

11. Populations Served
12. Percent of Population Served that is AAPI
13. Primary Areas of Focus
14. Prior GIM Participation
15. Website Link
16. Number of Employees (to be included in 2023 registration)

iv. Data vetting process:

1. Data vetting and validation is performed by APF during an organization's registration process. Any missing data (i.e., mission statement) may be supplemented using the Candid API or manual retrieval.

b. TAAF data intake form:

- i. Summary: while GIM is intended to be the primary self-reported data source for the first iteration of the TAAF Database Website, we are creating a TAAF data intake form as an option for organizations who would like to either: (1) provide information without participating in GIM, or (2) to directly provide data once GIM registration closes.
- ii. Scope determination:
  1. The number of organizations participating via the intake form is entirely dependent on socialization. TAAF will promote the website in various ways (press releases, influencer promotion, partner networks, etc.) and encourage new organizations to participate in the database.
- iii. Collected data fields:
  1. Data fields will mirror those collected by APF through GIM registration, in addition to the follow data fields:
    - a. Mission Statement
    - b. Donation Link
    - c. Social Media Accounts
    - d. Logo
    - e. Additional Images
- iv. Data vetting process:
  1. Initial Screening

- a. Once an organization contacts TAAF via the website to participate, TAAF will receive a notification and begin an initial screening.<sup>1</sup> When the organization passes TAAF's initial screening, they will receive the data intake form.

## 2. Data Intake Form Review

- a. Verify the organization's EIN and confirm that they are an IRS registered 501(c)(3) nonprofit organization or fiscally sponsored by a 501(c)(3).
- b. Confirm that the organization explicitly and intentionally centers AAPI communities in their mission and work.
- c. Confirm if the organization is a new or existing organization in our database, and either add or update the record accordingly.
- d. Perform a general quality control review of the data received to ensure accuracy, consistency, and completeness.

## 2. Candid/GuideStar:

- a. Summary: Candid is an information service specializing in reporting on U.S. nonprofit companies and maintains a database of information extracted from the IRS Form 990. In the event we are unable to gather data directly from an organization, we will use the data pulled from the IRS Form 990 via the Candid Essentials Application Program Interface (API) to populate the database. Our team developed a Python script to interface with the Candid database and retrieve data that closely matches the self-reported data fields.
- b. Scope determination:
  - i. The Bridgespan Group, a management consulting firm specializing in philanthropies and non-profit organizations, performed a landscape analysis of AAPI-serving organizations in the U.S for TAAF in September 2021 to better understand the ecosystem that supports AAPI communities. Organizations were sourced from a variety of AAPI networks, resource directories, news articles, and targeted GuideStar searches.
  - ii. The scope of this landscape only included entities that explicitly and intentionally centered AAPI communities in their mission and work. Please reference the following for additional details around Bridgespan's landscape scope determination:
    1. While we understand that AAPI communities are served by many community-based organizations, we wanted to understand the

---

<sup>1</sup> The initial screening criteria are still under development.

ecosystem of actors contributing to enhancing AAPI belonging or supporting AAPI communities in culturally competent ways so excluded general organizations.

2. We excluded international development organizations if they did not have programs centered on AAPI communities living in the United States.
  3. We focused on tax-exempt organizations, ranging from 501(c)(3) to 501(c)(6) status, and did not include AAPI-led small businesses, corporations, or employee resource groups within businesses.
- iii. We retrieved data via the Candid API for any of the organizations identified in the landscape analysis that did not self-report data or explicitly opt-out of the database.
- c. Collected data fields:
- i. For the in-depth field and data dictionary for the Candid API, please refer to the [Candid API Data Mapping](#).
  - ii. Fields:
    1. Organization Name
    2. EIN
    3. Address
    4. Main Contact Name, Title, and Email
    5. Executive Director Name
    6. Areas of Focus
    7. Populations Served
    8. Year Received 501(c)(3) Status
    9. Organization Logo
    10. Mission Statement
    11. Website Link
    12. Revenue and Expenses
    13. Number of Employees
- d. Data vetting process
- i. The TAAF Data and Research team mapped the most relevant data fields from the Candid API to the corresponding fields in the self-reported data. The team made these determinations based on data descriptions provided by Candid.

- ii. Candid sources organizational information via publicly available IRS Form 990s and host this information in their database. Via the API, we pulled the relevant data based on the mappings.
- 3. Data Point Mapping: to see data mapping between self-reported data and Candid/GuideStar, please refer to the [Candid API Data Mapping](#).