

[This is a description of the step-by-step technical process by which we 'scan' each Initial URL. These steps come after the initial process of building the website index.]

First off, before any scans take place, the process of ingesting the Initial URL list into the database populates the following fields: *

Initial URL * Initial Domain * Initial Base Domain *
Initial Top Level Domain * Agency * Bureau * Branch *
Data Source * Public * Filtered

When scanning commences, this core file dictates which scans are run. Due to the nature of the code base, the scans run asynchronously (i.e. they don't necessarily run in the order they are written in the code). Each scan operates separately and don't talk to each other.

The current scans are:

- ▶ **primary** - Loads the Initial URL and analyzes the resulting Final URL, generating most of the Site Scanning data.
- ▶ **dns** - Analyzes the DNS of the Final URL using a Node.js library (instead of Puppeteer).