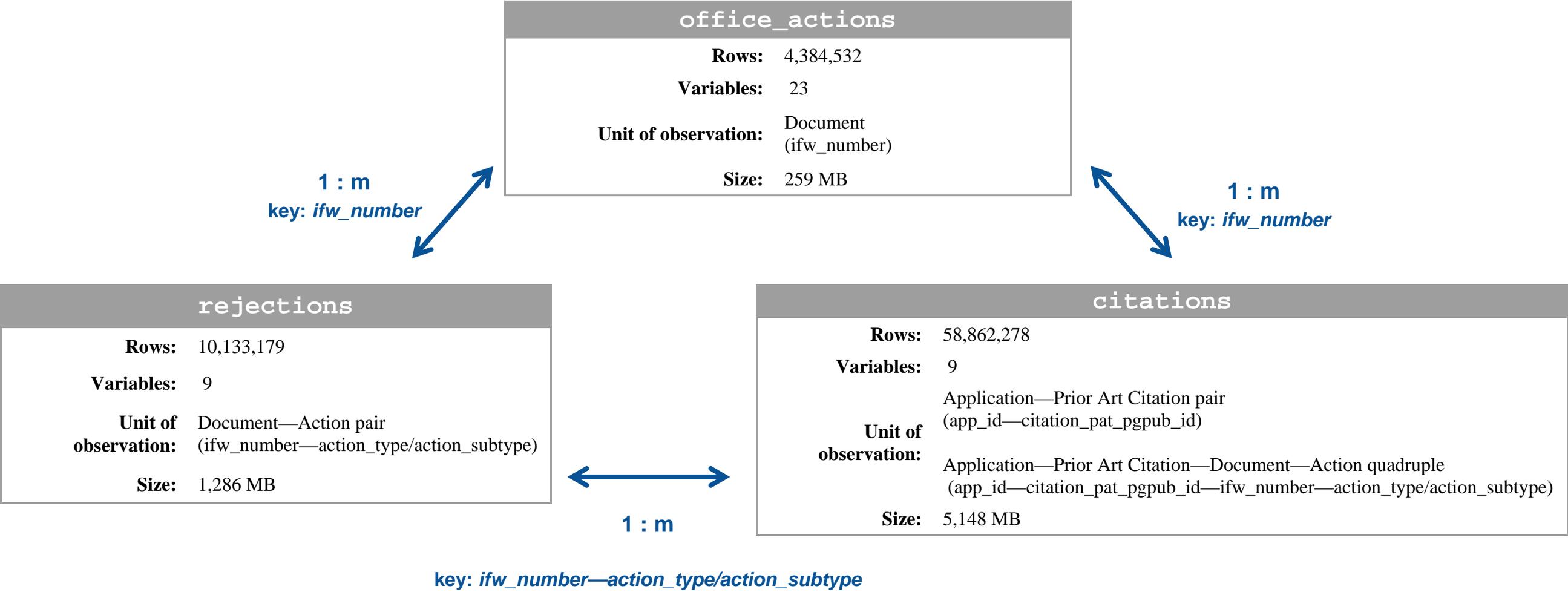


USPTO Office Action Research Dataset for Patents

2017 Release



Notes: Figure depicts the organizational structure of the *Office Action Research Dataset for Patents*. The *Dataset* includes three data files that can be merged with the identified key variables. The **office_actions** file contains Office action document level data. Each Office action is identified by a unique *ifw_number* which serves as the key for performing a one to many join of the **office_actions** table with the other two data files. The **rejections** file contains data at the Office action document-action pair level. A unique pair is identified by the *ifw_number* and action type (*action_type*) and subtype (*action_subtype*) combination. The **citations** file is derived from citations referenced on the Form PTO-892, Form PTO-1449, and text of Office actions. For citations not referenced in an Office action, a unique observation in the **citations** file is the application-citation pair (as identified by *app_id* and *citation_pat_pgpub_id* fields). For citations that are referenced in an action, the **citations** file contains fields to enable linking the application-citation pair back to the document-action pair (via the document *ifw_number* and *action_type/action_subtype* combination) in the **rejections** file.