United States Patent and Trademark Office

Office of Policy and International Affairs
Office of the Chief Economist

PatEx Technical Documentation: 2022 Addendum

Contact: Richard D. Miller

Date: September 2023

1. Introduction

The 2022 PatEx release includes some information that has not been made available in previous years. The release includes data on patent applicants in addition to inventors (**all_applicants**). It also includes metadata on patent examination-related documents which are sourced from the Content Management System (**cms_documents** and **cms_document_codes**).

2. Information on applicants

One of the provisions of the America Invents Act (AIA) was to allow an assignee to file a patent application as an applicant. Prior to the passage of the AIA, only inventors could apply for patents as applicants. Previous PatEx releases included only inventor information. With the 2022 PatEx release, we include information on non-inventor applicants for the first time.

The applicant information can be found in the **all_applicants** data file. It includes the names and locations of all non-inventor applicants listed on each application if such information exists for an application. For the 2022 PatEx release, there are 6,752,288 observations in this data file. Each observation represents an application/applicant pair and there are 5,171,923 unique patent applications represented in these data. There can be multiple observations for a given application. Table 1 provides a list and brief description of all variables included in the **all_applicants** file.

Table 1: List of variables included in *all_applicants*

Variable Name	Description	Type	Formatting
application_number	Application Number	str14	%-20s
applicant_organization	Applicant Name if Institutional	strL	%-20s
applicant_name_first	Applicant's First Name if Individual	strL	%-20s
applicant_name_middle	Applicant's Middle Name if Individual	strL	%-20s
applicant_name_last	Applicant's Family Name if Individual	strL	%-20s
applicant_city_name	Name of Applicant's City	strL	%-20s
applicant_region_code	Region (State) Code	str3	%-3s
applicant_country_code	Country Code (ISO 3166)	str2	%-2s

3. Information on examination-related documents

The 2022 PatEx data release includes two files on the documents filed by the applicants and examiners during the patent examination process. Although PatEx does not provide access to the documents themselves, they do provide metadata from the Content Management System such as the type of document filed, the length of the document, and the date of filing. Interested parties are directed to USPTO's Patent Center ([Patent Center | USPTO](#)) for access to the documents. The CMS document data provides a useful complement to the transaction history data described in [Technical Documentation for the 2019 Patent Examination Research Dataset (PatEx) Release (uspto.gov)](#).

The **cms_documents** file includes 549,192,929 observations covering 10,653,568 patent applications. Table 2 provides a list and brief description of all variables included in this file. The accompanying **cms_document_codes** file provides a description for each of the 838 unique document codes found in the **cms_documents** file. A list of the variables found in this file is provided in Table 3.

Table 2: List of variables included in *cms_documents*

Variable Name	Description	Type	Formatting
application_number	Application Number	str14	%-20s
mailroom_date	Date of Document Receipt or Mailing	float	%td
document_code	Document Type Identifier	str12	%-12s
number_of_pages	Number of Pages in Document	long	%8.0f

Table 3: List of variables included in *cms_document_codes*

Variable Name	Description	Type	Formatting
document_code	Document Type Identifier from *cms_documents*	str12	%-12s
document_description	Description of CMS Document	strL	%-20s