## Appendix F: Description of the Address and Attorney/Agent Data Release

## F.1 Datasets Included in this Release

This data release consists of one dataset that provides the information that an analyst would be able to glean from the top panel of the "Address & Attorney/Agent" tab on PTO's Public PAIR website. An example of the "Address & Attorney/Agent" tab is presented in Figure F-1. Note that the tab is broken out into two panels. The top panel presents information on the latest correspondence address for the application along with the customer number. The bottom panel presents information on the attorneys and patent agents who have been granted power of attorney with regard to the subject application. The information includes each attorney's or agent's name and registration and telephone numbers. At this time we are unable to provide the data from the bottom panel, but will include these data in our 2016 release.

The data set that we have included is named CORRESPONDENCE\_ADDRESS and it contains information from the top panel. It includes 8,719,698 observations covering 8,714,076 unique patent applications.

## F.2 Variables Included in CORRESPONDENCE\_ADDRESS

The CORRESPONDENCE\_ADDRESS dataset contains 12 variables, as illustrated in Table F-1. The variable *application\_number* can be used to link information from the CORRESPONDENCE\_ADDRESS dataset to other datasets included in the Public PAIR data release. A handful of the applications included in this dataset have multiple observations, but these account for fewer than 5,500 of the 8,714,076 applications present. The rest of the variables are longer string variables and are fairly self-explanatory. The name of the entity with which USPTO is meant to correspond is given up to two lines of text. In most, but not all, cases the entity is a law firm or the legal department of a commercial enterprise. The next two variables – *correspondence\_street\_line\_1* and *correspondence\_street\_line\_2* – describe the street address of the correspondent. The *correspondence\_city* and *correspondence\_postal\_code* variables are also self-explanatory.

Determining the location (country or U.S. state) of each correspondence address can be done using the *correspondence\_country\_code* and *correspondence\_region\_code* variables. The *correspondence\_country\_code* variable is coded using the ISO 3166

format.<sup>79</sup> We also include a variable *correspondence\_country\_name*, which can also be used to decipher the country codes. For US and Canadian addresses, the *correspondence\_region\_code* variable can be used to determine state or province. States and Canadian provinces are coded using standard US Postal Service 2-digit abbreviations. The final variable is *customer\_number*, which can be used to uniquely identify the correspondent. This variable has a valid value for nearly 5.6 million of the 8.7 million observations in the CORRESPONDENCE\_ADDRESS dataset.

<sup>79</sup> For more information on ISO 3166 country codes see http://www.iso.org/iso/country\_codes.htm.

Table F-1: List of Variables Included in CORRESPONDENCE\_ADDRESS

Variable Name	Description	Type	Formatting
application_number	Application Number	str14	%-14s
correspondence_name_line_1	Entity Name Part 1	strL	%-20s
correspondence_name_line_2	Entity Name Part 2	strL	%-20s
correspondence_street_line_1	Street address Part 1	strL	%-20s
correspondence_street_line_2	Street address Part 2	strL	%-20s
correspondence_city	Name of city	strL	%-20s
correspondence_postal_code	Postal code	str19	%-19s
correspondence_region_code	Region (State)	str4	%-3s
correspondence_country_code	Country code (ISO 3166)	str2	%-2s
correspondence_country_name	Country name	strL	%-20s
customer_number	Customer number	str6	%-6s

Figure F-1: The Address and Attorney/Agent Tab for Application 11/874,690

