Peacock Data talks about the language of names

Written by Australian Business
Tuesday, 22 April 2014 07:00

The new pdNickname and pdGender software covers given names and nicknames from more than 200 languages.

(PRLEAP.COM) April 22, 2014 - According to Barbara Adair, Peacock Data's chief development coordinator, the new language features in their **pdNickname** and **pdGender** software packages **"have never been available before on this scale and required a sizable portion of the nearly five years of research and development."**

With this new software, users can identify the languages associated with first names and nicknames, providing them with critical ethnic and heritage demographics about their clients.

pdNickname is an advanced name and nickname file used by businesses and organizations to merge database records. They can match the data against their lists to determine if two or more records are the same individual. It identifies first names that are the same even when they are not an exact match, but rather equivalent, such as a variation or nickname.

pdGender is gender coding database built on the same set of names. Users can match the data against the first names on their lists to determine male and female identification.

Both packages embrace a host of similar and compatible features including languages of origin and use for each name as well as fuzzy logic so information can be recognized even when lists have typographical errors or stylized spellings.

Peacock Data talks about the language of names

Written by Australian Business
Tuesday, 22 April 2014 07:00

About half the names covered are English and Spanish names, and international names originating and used in over 200 other languages make up the second half. This includes such languages as French, German, Chinese, Japanese, Vietnamese, Korean, Hindustani, Russian, Arabic, Persian, and Yiddish as well as Native American names and ancient Greek, Latin, and Hebrew names.

Plans for the software releases were initially written up in January 2009 and development began in earnest mid-summer of that same year. The products were built during the same development cycle because both are extracted from the same master file. They were first available to the public on December 30, 2013.

According to Barbara, **"Creation of the master name file these new products result from is the biggest venture our company has ever undertaken. There are thousands of sources for names in scores of languages, and our task was to compare and contrast all this data and create the ultimate first name resource."**

"From the start it was essential to identify the languages associated which each name in considerable detail," she added. "This gives users previously unavailable ethnic demographics linked to the names already on their lists."

Barbara showed some of the documents used in construction of the new offerings including a manuscript from 731 AD, written by a monk named Bebe, listing the earliest English names dating from the Anglo-Saxon era of the Early Middle Ages. The still common personal name "Hilda" is an example from the manuscript.

"Because sources often give diverse information and use different spelling conventions, it was crucial not only to gather all the information possible but also to differentiate between the quality of sources," Barbra explained. "Better information became easier to identify after working with the sources over the course of the first year."

According to Barbara, **"Special attention is paid to rare usages of unisex names like Kimberly, Hillary, Valentine, and even Maria. Names like these, while usually associated with one gender, are also occasionally employed by both genders. The new products identify rare usages so they can be considered separately."**

Peacock Data talks about the language of names

Written by Australian Business
Tuesday, 22 April 2014 07:00

"Beyond just identifying the languages of use, we also classify name origins, such as Old English opposed to Middle English opposed to modern English," Barbara noted. "This adds value for those researching personal names or the relationships between languages, such as in the fields of anthroponymy, onomatology, ethnology, and linguistics."

According to the product documentation, both packages identify five basic first name types:

Base Names Variations Short Form Nicknames Diminutives Opposite Gender Forms **"Assigning a type identification to each name was a lengthy part of development, but it is significant because the added information permits more precise filtering and ultimately better results,"**
Barbara said. "Base names are characteristically the oldest because they are the original names all later formations can be traced back to. A lot of time was devoted to these. It is important they are identified as accurately as possible because the remainder of the database is dependent on them."

Both products are available for immediate download from the company's [website](#) . They come with perpetual site licenses allowing installation on all computers in the same building within a single company or organization.

Product informationpdNickname: Name and Nickname Software [Information…](#) **pdGender**: Name and Gender Coding Software
[Information…](#)

About Peacock Data Peacock Data are the makers of unique database products used by business, organizations, churches, schools, researchers, and government.

Their flagship offerings include: **pdNickname**, a highly-regarded name and nickname product recently upgraded to version 2.0;   **pdGender**, a gender coding database also recently upgraded to version 2.0;
pdGeoTIGER
, a precision ZIP+4 and address range GeoCoding package;
pdCensus2010
, with demographic data drawn from 2010 American census tabulations; and
pdACS2013
, unveiled last May, another demographics offering providing American Community Survey (ACS) estimates gathered from the U.S. Census Bureau and summarized at over 100 stratification levels.

Peacock Data talks about the language of names

Written by Australian Business
Tuesday, 22 April 2014 07:00

Peacock Data is a California-based company in business since 2003.