

Personal names around the world

Martin Fenner, Gobbledygook

August 14, 2011

Yesterday I discovered (via a tweet by Owen Stephens) a very interesting document Personal names around the world that discusses the following question:

How do people's names differ around the world, and what are the implications of those differences on the design of forms, databases, ontologies, etc. for the Web?

The document was written by Richard Ishida, Internationalization Activity Lead at the W3C (World Wide Web Consortium). The document was published on July 26, and Richard was seeking comments until August 7 before finalizing the document.

The document is a good summary how names for people differ around the world, e.g. multiple family names (Spain, Latin America), no family name (Iceland), different ordering of names (China, Korea, Japan), non-latin characters in names (many countries), and other issues.

The second part of the document makes a few suggestions of how these variations in names could be handled on the web. The text should be required reading for anybody who is designing databases that handle international names – and there are a lot of them. A form that asks for **first name**, **middle initial** and **last name** will just not be appropriate for a lot of people.

I am lucky that my German name doesn't contain any German umlauts (ä, ö, ü) or the letter ß, so I haven't had any bad (or funny) experiences with my name. But I know that it can be a big problem for a lot of people. The result is that people end up having several spellings for their name, or write their name in ways that were not intended, e.g. a hyphen between the two family names in Spain. The name is something very personal, and I think the least we can do is to allow people to use their name appropriately.

Science is probably no better or worse in this regard than other domains. If we are lucky, we find a journal that prints our name correctly, or have a database that understands that ü should be sorted as ue. But for the most part, this seems to be an unresolved issue. And I don't want everybody to start using a first name and last name in that order and only with ASCII or latin characters. That would be boring. So please start thinking about this issue when you design



Figure 1: Flickr photo by Giant Gingko.

systems that use personal names, and use the W3C document by Richard Ishida as a starting point.

Disclaimer: I sit on the Board of Directors of the Open Researcher & Contributor ID (ORCID) initiative which aims to help solve this and related problems.