The taxonomy and lexicon project at the US Bureau of Labor Statistics began in summer 2013 with the goal to provide consistent access to BLS data and documents. The taxonomy will support searching for data, while the lexicon will support tagging, cataloging, and searching for documents. In addition, the taxonomy will be incorporated into the user interface for the new DataFinder dissemination tool BLS is building.

A main problem is that technical language used by BLS is sometimes misunderstood by users. For instance, users searching for data using the word "Cities" might not locate or understand the meaning of data labeled with the technical term "Metropolitan Statistical Areas". So, in addition to including BLS technical terms, identifying plain English words, such as "cities", users often employ in their searches for BLS data is necessary for making the taxonomy useful to a broad audience.

The project team includes subject matter experts from the program offices, research, publications, and field operations. First, the team captured the terms and definitions for every measure in the time series database and began the interviewing process to determine the plain English words users commonly employ. In the second phase, terms for measures and terms in classifications used to sub-divide them were separated into hierarchies, called facets; the plain English interviews were completed; and the results were merged into a first draft taxonomy.

Now the team is focused on improving the quality of the classification facet of the taxonomy. This work has uncovered several interesting issues. These will be the main focus of the paper. First, however, some examples from the taxonomy will be presented to illustrate it, then the findings of the plain English examination will be discussed. Finally, there will be a discussion of future work.