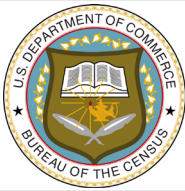


Automated data integration contributes to \$5 billion saving for US Census



Client:

US Census Bureau

Industry:

Government

Key Benefits:

- ▶ Enables a more accurate, more current database
- ▶ Reduces manual field canvassing to 25% of previous figure
- ▶ Contributes to total projected saving of \$5 billion

Conducted every ten years, as required by the Constitution, the US Census is the largest civilian activity in America.

The project, which counts and profiles a population of over 318 million and involves checking 135 million addresses, is used to correctly apportion political representation within the House of Representatives. As well as its political importance, the census is the basis for allocating federal funds worth \$400 billion each year.

The next census will take place in 2020, and plans are fully underway. As Tim Trainor, Geography Division Chief at US Census Bureau, explains,

“This is a census that will have more change than in the last many decades.”

The cost of integrating data from 3,200 partners

Geospatial data underpins everything that the census bureau does, as Tim describes,

“There are 34 different operations in the census and most of those are dependent on geography that we manage in the Geography Division.”

As well as addresses, the bureau requires accurate records of the nation’s full road network and the boundaries between all of the country’s state, local and tribal governments. The United States contains 40,000 separate governments, each with its own boundaries and potential allocation of federal funds.

To build its platform of geographic information, the bureau aggregates data from 3,200 counties and other organisations. Identifying changes, integrating new data and

maintaining the correct relationships between all of these data sets is a critical challenge for the Census Bureau and one which consumed a huge amount of resource and time.

“Integrating the data was a very manual process,” explains Tim. “It took a very long time to deal with that level of data.”

Once the base data set – known as MAF/TIGER (Master Address File / Topologically Integrated Geographic Encoding and Referencing) – is created, it requires validation. For the 2010 census, the bureau hired 140,000 individuals to walk or drive every street in the country and validate the bureau’s address records.

After the census date, the bureau also required over 600,000 people to trace non-respondents, ensuring individual participation as required by law or identifying addresses as vacant. The bureau has only a short time in which to collate and present the results of the census, so this follow-up work had to be completed quickly.

Altogether, the volume of manual work required for integration, validation and post-census follow-up resulted in an enormous expense. The total cost of the 2010 US census was \$12 billion.

If the bureau continued on the same basis, the cost of the 2020 census was forecast to be \$17 billion.

The bureau needed to find more efficient ways to manage the process.

Automating data validation and integration

The bureau identified four innovation areas to improve preparation and completion of the census:

1. Re-engineer the address canvassing operation
2. Improve response collection methods to enable online and telephone collection as well as traditional paper forms
3. Use additional, existing records to improve the base data set
4. Make greater use of technology to manage the field infrastructure

To underpin these, Tim wanted to streamline and automate the processes of managing his geographic data.

“We’ve been looking for tools to help us for a long time,” he explains. “We developed some of our own, but it was still a very manual process.”

Tim and his team started to work with 1Spatial. The company recommended a solution built on 1Integrate which uses rules-based technology to automate data management process based on user-defined, user-managed rules. Together, 1Spatial and the geography division designed an automatic data conflation process to manage the acceptance and integration of data submissions from the bureau’s 3,200 partners.

The solution automatically identifies differences between data in new partner submissions and existing data within MAF/

TIGER. The software generates a Change Proposal Layer containing all of the changes suggested by the new data. Once accepted by bureau users, the Change Proposals are automatically processed by 1Integrate to update the MAF/TIGER database.

This high degree of automation means that the bureau can process more partner files, more quickly. As a result, the MAF/TIGER database is more accurate and up-to-date.

“One of the things we’ve seen from our engagement with 1Spatial is that our approach to data management is different,” explains Tim. “We’re seeing some real improvements in streamlining and automating our processes and we’re managing our data in a more efficient and effective way. At the end of the day that helps in meeting our four objectives.”

More current records mean less field canvassing

A stronger starting point means the bureau will require much less field canvassing for the 2020 census. Tim estimates that only 25% of 2010’s 140,000 field canvassers will be required. The rest of the validation will be done by desk-based operators comparing MAF/TIGER to aerial photography – a much more efficient solution.

Augmenting MAF/TIGER with commercial data

Automating the process of data integration or conflation enables the bureau to easily augment MAF/TIGER with additional data sets, such as “Gone Away” records from the US Postal Service or road network data from

commercial providers. Adding this additional information further reduces the need for field canvassing in advance of the census and for post-census follow-up.

Innovation areas drive cost avoidance

US Census Bureau’s four innovation areas will make the 2020 census more streamlined and efficient than previous censuses.

“We estimate that the cost avoidance from these four innovation areas will be a little over \$5 billion,” says Tim. “That’s close to the cost of the 2010 census, and quite an achievement.”

Tim’s team and 1Spatial continue to explore how the rules-based automation of data management processes could bring savings in other areas for the 2020 census.

To find out more or to view the case study video visit: www.1spatial.com/u-s-census-bureau

For more information about 1Spatial’s solutions visit 1spatial.com