

Merging statistics and geospatial information, 2012 projects - the Netherlands

Statistics Explained

This article forms part of Eurostat's statistical report on *Merging statistics and geospatial information: 2019 edition*

Final report 9 September 2013

Problem

No institutional framework for linking statistics to locations (due to an absence of standardisation and interoperability).

Objectives

The general aim of this study was to explore the merits and possibilities of a table joining service (TJS), a geospatial standard developed by the [Open Geospatial Consortium \(OGC\)](#). A TJS offers a standardised web interface for automatically joining tabular statistical data to geographical information (administrative boundaries, postal codes, statistical units).

The study examined the concept and functional possibilities of using a TJS, existing TJS implementations and client applications. The implications, roles and tasks, as well as a cost-benefit analysis of the adoption and implementation of a TJS were also considered. As such, the main goal of the study was to gain an insight into the merits and possibilities of using a TJS for merging or linking tabular statistical data and geographical information.

Method

A TJS can be seen as an alternative method (to SQL and databases, GIS-based, WFS-based methods) in the context of the adoption and use of data services that are based on spatial infrastructures. A TJS has several potential advantages, such as its use in service-oriented architectures, as it is an open standard that promotes interoperability; it is also relatively simple and powerful, using the power of networked computers, while it is relatively easy to find public datasets through registries. At the same time, some potential disadvantages of TJS were assessed, such as the adoption of another service specification and another specific format for encoding data, or the lack of a mechanism within a TJS for uploading and creating tabular (or geographic) data.

A TJS can be considered as a supporting concept within spatial data infrastructures for joining data from various, distributed sources or infrastructure nodes. The concept of table joining in a distributed service-oriented

architecture (over networks) requires that organisations implement their services according to a common spatial (identifier) framework. A common spatial and identifier framework consists of several aspects:

- data models for spatial frameworks based on a generic conceptual model;
- a generic approach for handling and encoding unique geographic identifiers;
- a registry for the maintenance of data models and identifiers.

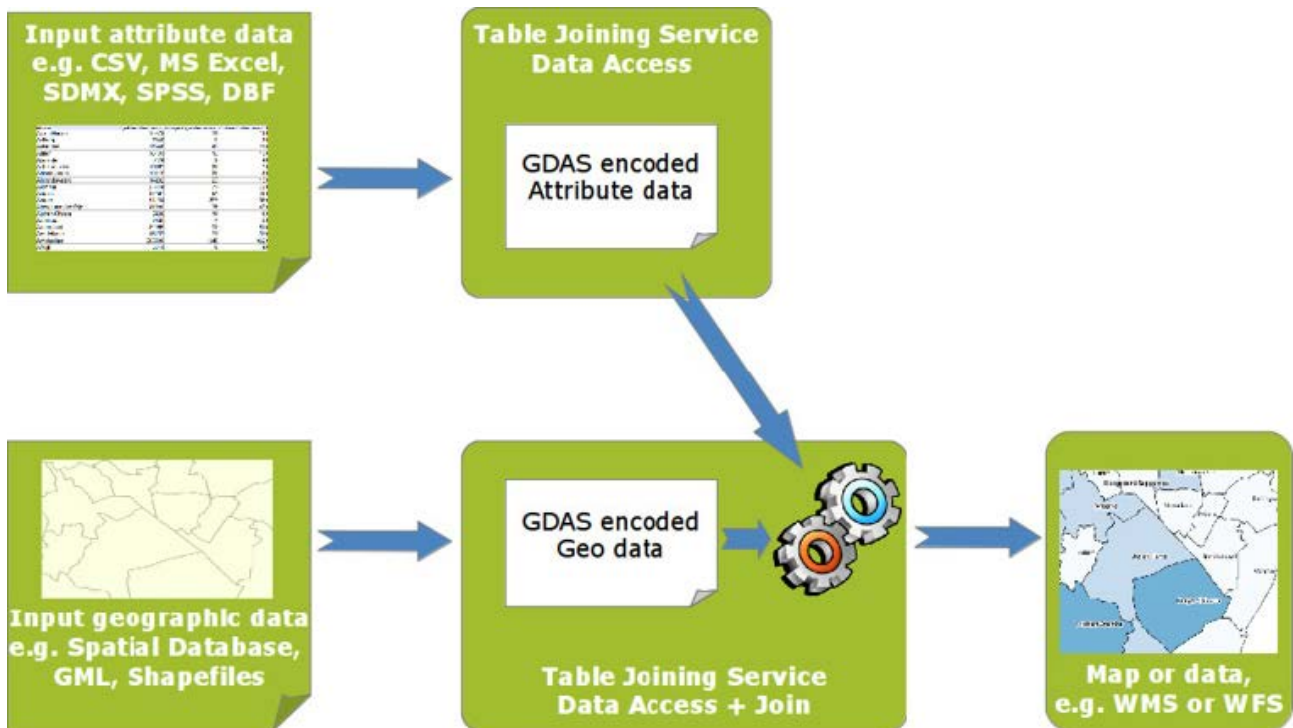


Figure 1: Input of data to create GDAS encoded data for Table Joining Services (TJS)

The project conducted a cost-benefit analysis of using a TJS across a large number of public organisations in the Netherlands. These looked at a variety of costs, including the costs of software development (client and server), infrastructure set-up and maintenance, hosting and training; these were compared with the costs of manually joining datasets. Three scenarios were considered: the first assumed continuing to manually join datasets; the other two looked at a small and a large-scale implementation of a TJS.

Results

The study concluded that a TJS has the potential to replace manual data joining operations in the daily data management practices that focus on thematic mapping and spatial statistics. The TJS standard offers a web-service interface to enable the automatic, service-oriented joining of tabular and geographic datasets across the web, while keeping original source data stored at the location of each individual data provider.

The cumulative cost of a small-scale implementation was estimated to be below the costs associated with manually joining datasets from the fifth year onwards, whereas for a larger scale implementation the cumulative costs of implementation were estimated to be lower by the eighth year.

Besides the monetary advantages, additional non-monetary benefits were also expected if a TJS was implemented:

- indirect efficiency benefits and higher quality output — the development of a framework requires a common agreement between many organisations offering spatial and tabular data, which may increase the quality and coherency of data management in several ways;
- societal benefits through better use of spatial statistics — more users would have access to spatial data through adopting a TJS solution, which would likely lead to an increased use of spatial statistical data with improved policy and decision-making;

- increasing and improving the quality of spatial statistical data exchange between organisations through the use of an international open standard — a TJS was expected to provide faster data delivery and a shorter 'time-to-market'; the use of open standards was also thought to have a positive impact on the independence of client software suppliers;
- benefits for organisations insofar as the use of TJS service-oriented architectures provide possibilities for monitoring data access and data use, thereby offering organisations the possibility to improve their services and service delivery.

The small number of TJS implementations found based on OGC standards and the lack of large vendors specialised in geographic information system (GIS) and TJS software implementations means that additional investment will be required before it is possible to test existing TJS software and client applications that make use of TJS.

Methodology

- [Final report](#)
- [Total cost of ownership model](#)
- [Presentation](#)
- [Geonovum](#)

External links

- [Statline](#)