

# Impact analysis Table Joining Service

Pieter Bresters (Statistics NL)

Hilde Keuning van Oirschot (Statistics NL)

Eric Fokke (Statistics NL)

Joost Venema (Dutch Cadaster)

Thijs Brentjens (Geonovum)

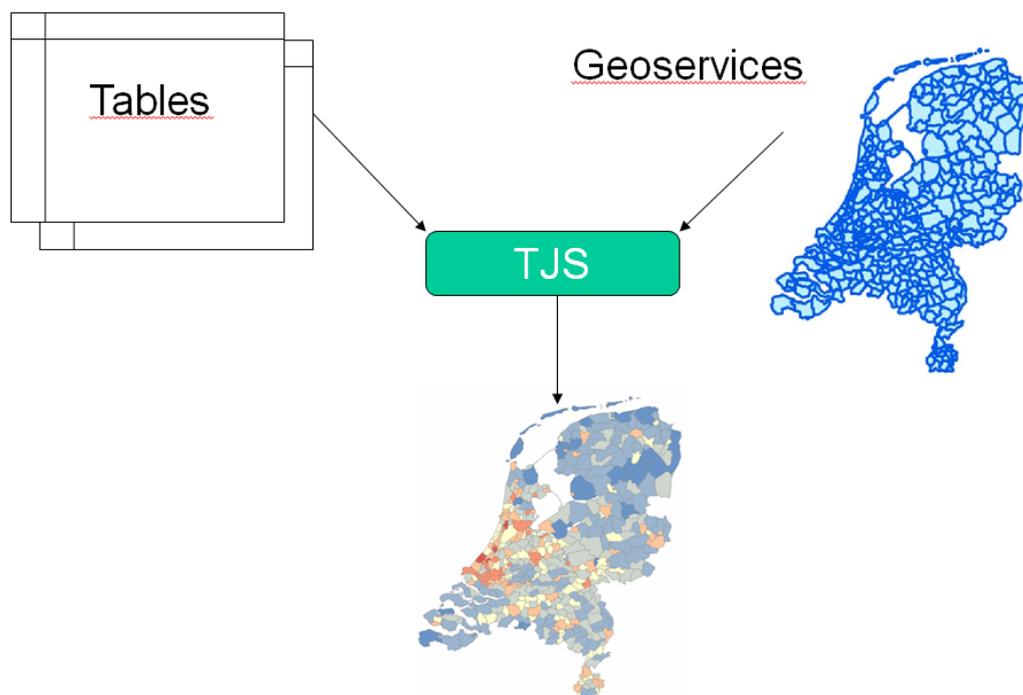
Michel Grothe (Geonovum)

Bart van Pelt (PDOK)

Jeroen Hogeboom (PDOK)

Dorus Kruse (PDOK)

Haico van der Vegt (Dutch Cadaster)



**CBS Den Haag**

Henri Faasdreef 312

2492 JP The Hague

P.O. Box 24500

2490 HA The Hague

+31 70 337 38 00

[www.cbs.nl](http://www.cbs.nl)

project number 301522  
SER/SLO  
September 2016

remarks The views expressed in this paper are those of the author(s) and do not necessarily reflect the policies of Statistics Netherlands.

# Index

<b>1. Introduction</b>	<b>5</b>
1.1 Summary (management)	5
1.2 Background	5
1.3 What is a Table Joining Service	6
1.4 Motivation and benefits	7
1.5 Objective	7
1.6 The Dutch situation	8
1.7 Scope impact analysis	8
1.8 Scenarios	8
1.9 Involved organizations	9
1.10 Versions	12
<b>2. Functionality</b>	<b>13</b>
2.1 Realized in POC	13
2.2 Future needs	14
2.3 Functionality per scenario	15
2.4 Pros and cons per scenario	15
2.5 Preconditions and requirements	16
<b>3. Impact hardware</b>	<b>17</b>
3.1 Servers	17
3.2 Data storage	19
3.3 Client	19
3.4 Governance/Maintenance	19
<b>4. Impact software</b>	<b>20</b>
4.1 Server side software	20
4.2 Conversion to required format	22
4.3 Web Application	22
4.4 Operations / Maintenance	23
<b>5. Impact data</b>	<b>24</b>
5.1 Geo data	24
5.2 Tabular data	24
5.3 Identifier management	25
5.4 Metadata	26
5.5 Maintenance	26
<b>6. Impact organizations</b>	<b>27</b>
6.1 PDOK	27
6.2 Statistics Netherlands	27
6.3 Dutch Cadaster	27
6.4 Geonovum	27

6.5	Other potential users	28
6.6	Other EU organizations	28
6.8	Central EU implementation	29
6.9	Global view	29
<b>7.</b>	<b>Financial impact</b>	<b>30</b>
7.1	PDOK	30
7.2	Statistics Netherlands	31
7.3	Dutch Cadaster	31
7.4	Geonovum	31
7.5	Other Organizations	31
7.6	International	31
<b>8.</b>	<b>Lessons learned and conclusions</b>	<b>32</b>
<b>9.</b>	<b>Recommendations</b>	<b>33</b>
9.1	PDOK	33
9.2	Statistics Netherlands	33
9.3	Dutch Cadaster	33
9.4	Geonovum	33
<b>10.</b>	<b>Acknowledgements</b>	<b>35</b>

# 1. Introduction

## 1.1 Summary (management)

A Table Joining Service (TJS) is an online service that links statistical tables to map services with online thematic map services as a result (See figure 1).

This document describes the impact analysis of a Table Joining Service (TJS) in an environment comparable to the national infrastructure that the Netherlands uses for the European project INSPIRE. It is the final delivery product of the Eurostat Grant awarded to Statistics Netherlands in cooperation with the Dutch Cadaster, PDOK and Geonovum. The involved organizations are described in chapter 1.9. It is created on the basis of a proof of concept of the TJS as executed by these organizations.

This proof of concept has shown how a TJS can work for the machine readable open data from Statistics Netherlands, combined with a map service with open topographic data for statistical units. The findings have resulted in this impact analysis.

One of the strongest motivations behind the TJS is the European project INSPIRE, but other motivations are mentioned.

The impact strongly depends on the amount of functionality that one could wish for. On the basis of the researched functionality wishes, three scenarios have been established (see 1.8)

The impact has been described on the following aspects:

- Hardware
- Software
- Data
- Organizational
- Financial

Technical readers should focus on the hardware, software and data. Managers should look at the organizational and financial impact.

The target group of this document is primarily formed by organizations involved in implementing a TJS in the Dutch PDOK environment, but it can also be interesting for other organizations in different countries. The organizational impact not only looks at the Dutch situation but also gives some thoughts about what impact it might have for the rest of Europe. (See 6.5 until 6.9 ). One of the conclusions is that a central European TJS could be a very cost effective way to realize the INSPIRE goals.

Finally some lessons learned and recommendation are described. In the recommendations one can find contradictions for the choice of scenario. This means that a discussion on how to continue is needed.

## 1.2 Background

The background for executing the proof of concept of the table Joining Service (TJS) in the Netherlands, can be found in the European projects INSPIRE and ELF. It appeared that many INSPIRE themes did not have a direct geometry, like the themes for Population Distribution and Human Health. They only consist of tables. So, a link is needed to create maps like INSPIRE requires. In the year 2013 Statistics Netherlands executed a study, on the basis of a Eurostat Grant that proved a positive business case for the Netherlands.

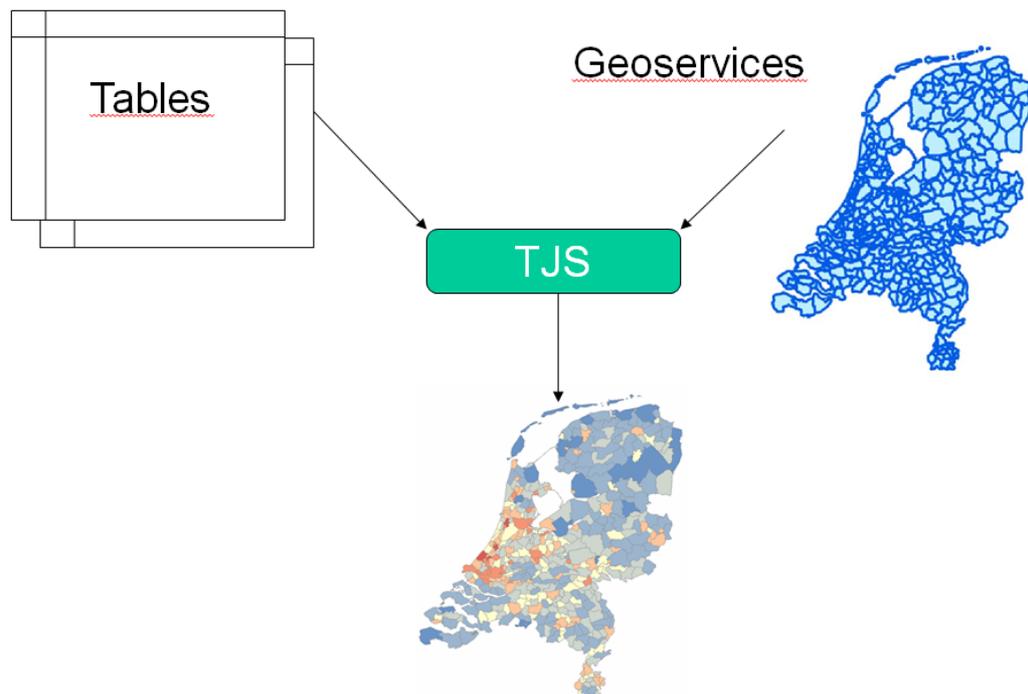
One can distinguish three problems with the current situation:

1. Organizations responsible for these tables do not always have access to the geometry needed for publication as an online map (geoservice). This problem does not only appear between organizations throughout Europe, but also within statistical organizations.
2. The amount of tables is huge. It is larger than one could ever expect the need for maps of these tables. So, they should only be produced when asked for and not in advance.
3. in order to create maps across borders, one should harmonize semantically and technically. This has both already been done by Eurostat by means of machine readable SDMX files. So doing this again according to the INSPIRE data specification would mean that this work is being done all over again.

Table Joining Services could solve these three problems. These SDMX files could be a perfect basis for a TJS. This is also what has been used in the ELF project.

### 1.3 What is a Table Joining Service

A Table Joining Service (TJS) is an online service that links statistical tables to map services. These tables can be table services like the SDMX table service from Eurostat or it can be a table uploaded by users. The geometry can come from existing geospatial information services or map services. The TJS performs an online task that is normally done by a GIS specialist. The TJS is an OGC (Open Geospatial Consortium) standard.



**Figure 1: TJS concept**

The output can be a statistical map service or geoservice, that can be saved as an image, or that can be imported into online applications or GIS systems.

The output services are assumed to be available for a restricted period, for instance one week.

Otherwise, the burden for keeping all the services alive, will be too large. Since the source data will always be available, one can always create the joined output again.

One can make a distinction between the TJS on the server site and the TJS client.

The server site TJS can be used by any application (TJS-client). The TJS-client applications can be simple: just giving the URL of the input and output, or sophisticated with a viewer and a well-designed graphical user interface.

## 1.4 Motivation and benefits

Statistical data have become more and more available as open data. The data have become machine readable, and can be used in several applications, without interference of any statistical officer. They have become table services. Examples are the Eurostat SDMX web services, and the open data portal of Statistics Netherlands.

Also geographical information has become open and machine readable data. They have become geospatial information services (map services or geoservices) . The European project INSPIRE has been a great stimulus for this.

What is missing is an online tool that links the statistical table services to the map services. This would enable us to create online statistical map services. This task could be facilitated with a Table Joining Service.

Other motivations behind the TJS are:

- A strong motivation behind the development of a TJS is the European project INSPIRE. It requires for all EU-offices to publish their geographical data as map services. This also counts for population, human health and sometimes energy statistics. Since these data are in most cases only available as tables and not as maps, this can be a difficult task for the statistical offices in the EU.
- A TJS, operating on existing table services and map services, could be a very cost-effective way to meet the INSPIRE goals: Making European geographical information better available for the public in an harmonized way.
- A TJS creates map services on demand, instead of making them all on forehand while not being sure if they are ever going to be used. Especially for statistical tables, the amount of map service would be huge if we would create map services for all tables with a geocode. It would be an unnecessary and costly duplication of stored and published data.
- Another advantage of a TJS is that it keeps the data at its source. The TJS connects to the source tables and that makes the output more up to date. The output map service remains up to date when a table has been updated. This is mostly not the case with conventional map service made by GIS-specialists, simply because it is too expensive to do it for every table update.
- A final and not unimportant advantage is that when a TJS would be centrally organized in Europe, it would save all the Statistical offices a lot of money, because it would help them for a great part to fulfill their INSPIRE obligations.

## 1.5 Objective

The main objective of the project behind this impact analysis is to show the management that the concept of a TJS can work in the Dutch PDOK environment and to give them insight in the impact. Behind this objective lies a deeper long term objective: that is to realize an online functionality to be able to make online maps on the basis of tables with a geocode, with all the advantages as described in the previous paragraph .

On the global and European level the stronger collaboration between NSIs and NMCA's is a significant development. Important to mention in this respect is the establishment of the UN-GGIM, the United Nations initiative on global geospatial information management. Both NSIs and NMCA's are supporting this initiative as representatives of their countries. One of the

objectives is to show the value and to further develop the use of geospatial information, e.g. in the work processes of statistical institutes.

This project is an example of such a collaboration, aiming at a win-win situation for both Statistics Netherlands, connecting statistical information to a location in a graphical way, and the Dutch Cadaster, whose open topographic data and server hosting facilities are used.

One other side objective on the European level is that other countries might learn from it, although the situation will be different in each country.

## **1.6 The Dutch situation**

In the Netherlands we are in a situation that is quite good for implementing a TJS:

- There is a national hosting service for (INSPIRE) geoservices: PDOK
- Geoservices for the geometry of statistical and administrative units are available as open and free data
- Many tables are available as open, free machine readable data
- There is knowledge on TJS, because of participation in the ELF project
- Geometry is available at many levels, from neighborhood until the lowest NUTS level.
- INSPIRE forces Statistics Netherlands to find a solution for publishing all these tables as geoservices
- Geonovum strongly supports Dutch organizations to fulfill their INSPIRE obligations and to work according to OGC standards

## **1.7 Scope impact analysis**

The basis for this impact analysis is the proof of concept of a TJS implemented in the PDOK environment that was performed with the funding of an Eurostat Grant.

The scope of this impact analysis is the situation in the Netherlands as described in the previous chapter. With the necessary adjustments, it could be projected on other countries. Chapter 6.6 deals with this issue.

This document describes the impact on three scenarios.

## **1.8 Scenarios**

The table in chapter 2.3 describes the functionalities per scenario in detail. The sub paragraphs below describe them globally.

### **1.8.1 Scenario 1**

This would be the extensive scenario with a TJS Plugin on GeoServer, fully supported by the GeoServer community and a client application that is fully maintained. This is the scenario which supports most functionalities.

### **1.8.2 Scenario 2**

This scenario is the same as Scenario 1 but without a fully maintained client application. Both scenario 1 and 2 are with a fully GeoServer community supported official extension since PDOK has indicated that this is a precondition.

### **1.8.4 Scenario 3**

This scenario is an alternative as proposed by PDOK. This alternative would be, using an implementation which fits the PDOK environment. This implementation could still be according to the OGC TJS interface standard.

Based on the requirement “mapping data with a geographic content on a map” PDOK proposes implementing the TJS standard in a way that is in line with the way PDOK implements data processing. This means a OGC TJS call is used to:

- upload the tabular data, and receive input about the connection between tabular data and the geographic data needed for presentation of the data on a map.
- Process the data. This means based on input format, convert to desired format, make a connection with the geographic data.
- Create WFS/WMS and make this services available

Processing data in this way is in line with the intended self-service environment of PDOK. PDOK strongly advises this scenario.

## **1.9 Involved organizations**

### **1.9.1 Statistics Netherlands**

Statistics Netherlands (CBS) provides the project management for this project.

The mission of CBS is to publish reliable and coherent statistical information which responds to the needs of Dutch society. The responsibility of CBS is twofold: firstly, to compile (official) national statistics and secondly to compile European (community) statistics. Statistics Netherlands (CBS) enables people to have debates on social issues on the basis of reliable statistical information.

The information published by CBS deals with subjects directly affecting the lives of Dutch citizens. These include economic growth, consumer prices, crime but also leisure.

The Statistics Netherlands Act constitutes the legal basis for CBS.

The Personal Data Protection Act (Wbp) addresses the handling of personal data. Furthermore, CBS has formulated its own code of conduct.

As of 3 January 2004, CBS has the status of an autonomous administrative body (in Dutch: ZBO). This means that CBS performs public service tasks, but operates independently and not under the direct authority of a Dutch ministry. The Minister of Economic Affairs is politically responsible for relevant legislation, budget and conditions. CBS is financed from the state budget.

More information: [www.cbs.nl/en-gb](http://www.cbs.nl/en-gb)

### **1.9.2 Geonovum**

Geonovum is the National Spatial Data Infrastructure (NSDI) executive committee in the Netherlands. The organization devotes itself to providing better access to geo-information in the public sector. Geonovum develops and manages the geo-standards necessary to implement this goal. The foundation of Geonovum marks an important step to improve coordination of public sector geographic data in the Netherlands. In its role as enabler of the national geo-information infrastructure, Geonovum connects public sector managers with the work floor

professionals. Geonovum's strength is that it discusses, coordinates and monitors arrangements between the parties concerned with regard to how activities are to be carried out and completed. Geonovum is recognizable and professional, result-driven and independent. Its main goals are:

1. To develop and standardize the geo-information infrastructure while also being innovative;
2. To build up and disseminate knowledge in the area of geo-information infrastructure;
3. To make the geo-information infrastructure more accessible to administrative bodies, institutions and departments in the Netherlands and the European Union.

More information: [www.geonovum.nl](http://www.geonovum.nl)

### **1.9.3 Dutch Cadaster (Kadaster)**

The Dutch Cadaster, Land Registry and Mapping Agency – in short Kadaster – collects and registers administrative and spatial data on property and the rights involved. This also goes for ships, aircraft and telecom networks. Doing so, Kadaster protects legal certainty. We are also responsible for national mapping and maintenance of the national reference coordinate system. Furthermore, we are an advisory body for land-use issues and national spatial data infrastructures.

Our information is available predominantly through online web services, including information on energy labels of houses and underground cables and pipelines. Our main customer groups are civil-law notaries, local authorities, businesses, financial institutions and private individuals.

The Dutch Cadaster maintains the Key Registers Cadaster and Topography.

The Dutch Cadaster performs its public tasks in service of society. This is reflected in the way we are organized and the ways in which we publicly account for how we work.

More information: [www.kadaster.nl](http://www.kadaster.nl)

### **1.9.4 PDOK**

PDOK (Public Data on the Map == Kaart in Dutch) is a collaboration of several government departments who want to publish public geographic data. PDOK aims to deliver maximum customer value at low cost. The standard way data is processed consist of: data delivery, data processing and data publishing.

The Dutch National SDI (PDOK) is a central facility for unlocking geo-data sets of national importance. This is actual and reliable information for both the public and private sectors. PDOK makes digital geospatial data available as data services and files. Most PDOK services are based on open data and are therefore available to everyone. The PDOK services meet national and international standards, including the European INSPIRE standards and the Dutch e-government standards.

PDOK was created by a collaboration between the Cadastre, the Ministries of Infrastructure and Environment, Economic Affairs, Rijkswaterstaat and Geonovum. PDOK is an open initiative. Every public organisation that wants to make its geospatial data available for reuse can join PDOK.

The organisation of PDOK consists of two parts. The functional and technical management of PDOK is placed at the Dutch Cadaster in Apeldoorn. The Steering Board is responsible for the control. All partners have a seat on the board.

the Dutch Cadaster is responsible for delivering the PDOK services and for supporting the users. The Steering Board monitors the quality of service and is responsible for the long term development of PDOK. The Steering Board is supported by Geonovum with a small team consisting of representatives of all partner organisations. The Customer Panel represents the users of PDOK and is a sounding board for customer satisfaction.

More information: [www.pdok.nl](http://www.pdok.nl)

## 1.10 Versions

Version number	description	Sent to
0.1	First frame work	All authors
0.2	Second framework after meeting 9-6-2016	All authors
0.3	Addition of 3 scenario's	none
0.9.1	Final version PDOK	All authors
0.9.2	Contribution of all parties	All authors
0.9.3	Assimilation of reviews	All authors
0.9.4	Nearly final	Management SN, PDOK, Geonovum
1.0	Final version	All Authors and Eurostat

## 2. Functionality

The project started with an inventory of possible needed functionalities. In the sub chapters below they are listed, divided in realized and not realized in the POC.

### 2.1 Realized in POC

The following functionality has been realized in the proof of concept:

- ✓ Table input from table services via URL (Odata or SDMX)
- ✓ Geometry input via WFS
- ✓ Join to create output as a temporary geoservice: WFS, WMS
- ✓ Input table format CSV via upload
- ✓ Output in a viewer (identify, classification, legend, zooming, pan, mouse over) (Stage)
- ✓ Output to other formats: geojason, kml (standard in GeoServer)

The online client application with this realized functionality can be viewed at:

<http://INSPIRElab.geonovum.nl/cbs/jstjs/>

We need to emphasize that this application is not fully tested and therefore is not fully ready to be taken into a reliable production environment. There is no maintenance organized and it is unknown how long it will stay online. It is only meant as a proof of concept.

One can consider the result of the POC as something in between scenario 1 and 2 with the difference that it is not ready for a production environment.

#### 2.1.1 Work process

Below are the work process steps needed to realize a join.

1. Select a spatial level
2. Select a table (Content tab) or type in your own csv URL and key
3. If the join is successful it will show a map with 5 classes, which you can change.
4. Click on "join data result" and on the bottom you will find the new WFS URL (for WMS, just change WFS in to WMS in the URL) . The layer name is indicated by the typeName:  
example: `<Parameter name="typeName">F12065887_KERNCIJFERSWIJK_166511</Parameter>`
5. This layer will remain for about 2 weeks and can be used in other applications like QGIS

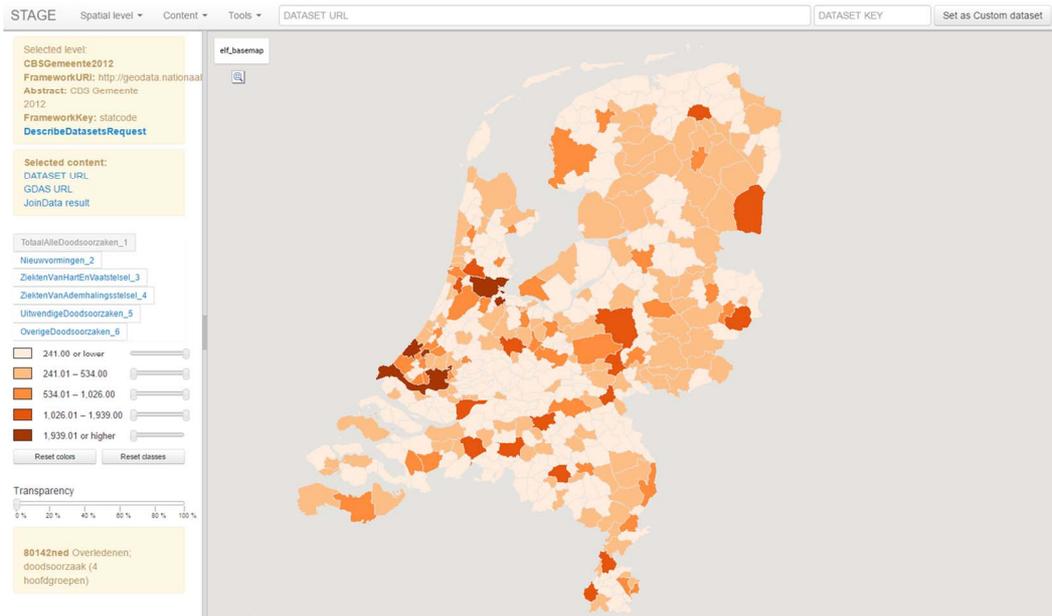


Figure 2: TJS client, STAGE as realized in the POC

## 2.2 Future needs

The following functionality has not been fully realized in the proof of concept:

- Key code selection from available attributes
- Live pick list creation from Odata tables and geoservices
- Warning for mismatches or wrong matches
- Normalisation to area or population
- Options for 1:n relations
- Meta data link to input
- Input from other table formats: xls, sav
- Input from other formats for geometry via upload: shape, geojson, kml
- Support authorisation

## 2.3 Functionality per scenario

Below the functionality is listed per scenario as described in paragraph 1.8.

functionality	POC	scenarios		
		1	2	3
Table input from table services via URL (Odata or SDMX)	+	+	+	+
Geometry input via WFS	+	+	+	+
Join to create output as a temporary geoservice: WFS, WMS	+	+	+	+
Input table format CSV via upload	+	+	+	+
Output in a viewer (identify, classification, legend, zooming, pan, mouse over)	+	+	-	-
Output to other formats: geojson, kml (standard in GeoServer)	+	+	+	+
Key code selection from available attributes	-	+	-	+
Live pick list creation from Odata tables and geoservices	-	+	-	-
Warning for mismatches or wrong matches	-	OGC?	OGC?	+
Normalisation to area or population	-	+	-	-
Options for 1:n relations	-	OGC?	OGC?	-
Meta data link to input	-	+	-	+
Input from other table formats: xls, sav	-	+	-	-
Input from other formats for geometry via upload: shape, geojson, kml	-	+/-	-	+/-
Support authorisation	-	+/-	-	-

“ OGC? ” means the OGC community needs to adjust the OGC standard to make this possible

“ +/- ” means low priority and only if time is left within the budget.

## 2.4 Pros and cons per scenario

Most pros and cons are formed by the functionality and the resulting costs (chapter 7).

Between scenario one and two this is mainly the lack of a viewer in scenario 2. This could lead to less usage, but it is much cheaper. Though it is still conform INSPIRE, since INSPIRE does not force us to create viewers.

Between scenario three and the other two, it is more difficult to define the pros and cons. The difference is defined by the use of a GeoServer plug inn and the stimulation of its community. Scenario three has less danger in problems with the alignment of the versions of GeoServer and its plug inn. It does not need the stimulation of the GeoServer community. So on the short term scenario three seems easier to implement and maintain for PDOK and therefore it is expected to be cheaper. PDOK expects it to be more difficult to organize the acceptance of a TJS plugin within the GeoServer community (scenario one and two) than it is to build something that only fits the PDOK environment (scenario three). Also because PDOK is a high availability SDI with many webservices and users, it is a disadvantage if PDOK needs to maintain two versions of Geo server. This is not only the case for PDOK, but will be of importance for any NSDI that will face the same issues. When Geoserver with TJS plugin is widely accepted and used, this will be a less important issue.

On the other hand, scenario three does not help the rest of the world, since it is not implemented in GeoServer, which is widely used in the rest of the world. So, scenarios one and two would be better for the open source and open data community. This might also be

profitable for the Dutch implementation on the long term and therefore, on the long term these scenarios might be a better choice. A precondition for this profitable situations is that the burden of stimulating the GeoServer community is not only placed on organizations from the Netherlands.

## 2.5 Preconditions and requirements

- A logical precondition is that the geometry and the tables contain the same key to enable joining. INSPIRE might be of help in this issue, because it requires unique identifiers. But still, this can lead to mismatches, ea. when they differ in type (string/integer) or when local-id's and INSPIRE id's are used.
- A major risk with a TJS, is that users try to link tables to the wrong geometries. Especially with administrative units like the Dutch community borders that change every year. The user should be warned for this. Providing all the geometries for the different years and warning the user for mismatches, might help.
- In order to full-fill the INSPIRE requirement, one should make the data findable in meta data registries. It is too much to ask for when this is asked for all the possible outputs one can make with a TJS. Therefore it is suggested to only publish a meta data record for the geometry and for the database tables that are used.
- Another precondition is that the geometry is available as open data map feature services. This is not always the case for each European country. This could be overcome by providing less accurate or generalized borders for these countries.
- The results of a TJS are in principal temporary. This means that a period has to be defined for the resulting geoservices to exist.
- A requirement for PDOK is that we make use of the standard PDOK infrastructure and if a GeoServer plugin is used, it must be fully supported by the GeoServer community.

## 3. Impact hardware

This chapter describes the impact for the hardware. Per chapter, the concerned item is considered per scenario from chapter 1.8.

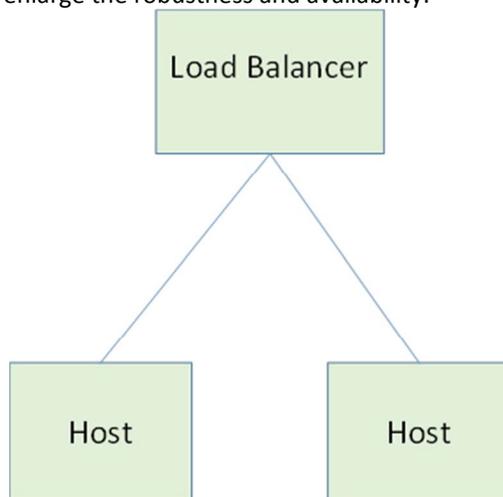
### 3.1 Servers

#### Scenario 1 and 2:

It is assumed that for production purposes the same servers are needed as used in the POC. Within the POC, two servers were used.

- A JAP Container for GeoServer
- A Docker container for the conversion to the GDAS format

A JAP container contains a load balancer and two virtual hosts, divided over data centres, to enlarge the robustness and availability.



**Figure 3: JAP container platform**

A precondition to be allowed to use this platform is that no local state is being occupied. During the POC, it became clear that GeoServer in combination with TJS uses a local memory state. The solution was to disable one of the two hosts. The disadvantage is less robustness and availability.

A Docker container is a virtual machine on which a Docker daemon is active. This enables you to start a Docker image. For the POC an image was started with a python environment. This enables us to run the python scripts for the conversion to the GDAS format as needed for an OGC TJS.

The two servers had the following technical specification:

JAP: per host : 2 CPU's 4 GB Memory 500GB local disc space  
CPU type Intel(R) Xeon(R) CPU E5-2650 v2 @ 2.60GHz  
Docker : 2 CPU's 4 GB Memory 500GB local disc space  
CPU type Intel(R) Xeon(R) CPU E5-2650 v2 @ 2.60GHz

When using the TJS extension plugin and conversion software there are several possibilities for packaging the software and deploying the software.

The following deployment considerations are important:

- 1 The conversion software is provided as a software package that needs to be deployed on a separate server (this was the setup used in the POC)
- 2 The conversion software is provided as part of the TJS extension software
- 3 Using GeoServer with TJS extension has the risk that the lifecycle of the GeoServer software and the TJS extension plugin software does not match. In order to be able to use GeoServer with the TJS extension, this means there always must be a separate instance for TJS usage. This assures that when GeoServer has a new release, adapting the TJS plugin software is not blocking implementing the new GeoServer release. The lifecycles of GeoServer and GeoServer + TJS plugin are separate.
- 4 Since the TJS extension has different usage characteristics then only WMS/WFS services, it needs a dedicated deployment of the GeoServer.

For deployment<sup>1</sup> there remain two options:

- Option 1: A GeoServer + TJS plugin with a separate conversion server, 2 separate servers
- Option 2: A GeoServer + TJS plugin with built in conversion software, 1 server

In the PDOK infrastructure there is a complete OTAP<sup>2</sup> development pipeline, for each of these 4 environments the above mentioned number of servers is needed.

Description	Total number of servers for the complete OTAP development pipeline
Option 1	2 separate servers per environment * 4 environments = 8 servers
Option 2	1 server per environment * 4 = 4 servers

**Scenario 3:**

This scenario has no impact on the hardware for the first period of time. If the TJS will be used extensively, extra hardware is expected to be needed. Costs indication: €18.800,- a year.

<sup>1</sup> As server a PDOK JAP server is used with 4 CPU's an 4 GB memory

<sup>2</sup> OTAP (Ontwikkeling=Development, Test=Test, Acceptatie=Acceptance Test, Production=Production) environments

### 3.2 Data storage

The assumption is that the database storage per dataset is 1MB. Under 50GB there will be no costs. There are no indications that in the first period of time a higher amount of storage is expected. If so, a meeting to deal with this situation will be planned.

This counts for all scenario's.

### 3.3 Client

Client software is only applicable for scenario 1.

The current example client is HTML/JavaScript only: <http://INSPIRElab.geonovum.nl/cbs/jstjs/>

There is no server side processing. So hosting the static files is sufficient. This can be done on any basic webserver, even on the same server as the TJS service because this won't interfere with the server side processes.

### 3.4 Governance/Maintenance

#### Scenario 1 and 2:

The maintenance of the GeoServer + TJS extension would be in monitoring and maintaining the before mentioned option 1 and option 2 (see 3.1).

Besides that the lifecycle of the datasets must be managed, software updates (GeoServer and TJS) must be managed. In order to get TJS as a supported module a developer must be assigned to be able to support migrations to a new GeoServer release.

Estimated cost for total 30 days each year.

PDOK is the infrastructure for provisioning open data. In order to maintain a stable platform the requirement for software is that it is maintainable, and the usage of the software fits in the PDOK architecture vision.

The PDOK architecture uses a distinct separation of concerns for processing data and provisioning data. The TJS extension is not in line with these conventions. This poses problems in governance of GeoServer with a TJS extension.

Therefore PDOK is strongly in favor of **scenario 3**. This alternative has no impact on the hardware for the first period of time, and therefore no hardware maintenance..

## 4. Impact software

This chapter describes the impact for the software on the servers side and the client software. Per chapter, the software is considered per scenario from chapter 1.8.

### 4.1 Server side software

For the POC the following software was installed on the JAP host:

- Het OS is red hat enterprise linux versie 7
- Java 1.7
- Apache Tomcat 7
- Adjusted GeoServer with TJS extension

For the POC the following software was installed on the Docker host:

- OS: Red hat Enterprise Linux version 7
- [peltb@inu432 ~]\$ sudo docker version
- Git commit (client): 786b29d
- OS/Arch (client): linux/amd64
- Server version: 1.7.1
- Server API version: 1.19
- Go version (server): go1.4.2
- Git commit (server): 786b29d
- OS/Arch (server): linux/amd64
- Client version: 1.7.1
- Client API version: 1.19
- Go version (client): go1.4.2
- Git commit (client): 786b29d
- OS/Arch (client): linux/amd64
- Server version: 1.7.1
- Server API version: 1.19
- Go version (server): go1.4.2
- Git commit (server): 786b29d
- OS/Arch (server): linux/amd64
- Python 2.6.6
- Python conversion to GDAS format

#### 4.1.1 GeoServer extension (scenario 1 and 2)

The TJS server is an extension to a default GeoServer installation. It can be installed and will run on recent versions of GeoServer (like 2.7.x, 2.8.x). The TJS extension is originally developed by a Cuban organization and has been modified for the E.L.F. Project and for the CBS application.

In the pilot, GeoServer and the TJS extension have successfully been installed on PDOK's server environment.

Depending on the setup PDOK wishes (e.g. what fits best to maintain the service?), the installation for a TJS could be in a dedicated environment or in the default PDOK environment.

The following requirements and actions have been identified and are considered necessary for PDOK to get the TJS extension from POC to production-level:

1. replace the in-memory database H2 to a PostgreSQL database. This is also necessary to enable load balancing. The database model is straightforward. Replacing the database requires rewriting and extending the TJS code to support another database. (Rough estimate: 5-10 days work including testing)
2. cleanup code and comments, improve build process (rough estimate: 2-5 days work including testing)
3. update technical documentation (rough estimate: 1-4 days work)
4. bug fixing: there are some bugs - none of which have been crucial for the pilot - that are important to fix. For example: fix issues in the web admin pages for the TJS configuration (rough estimate: 3-10 days work).

In addition, there might be functional requirements that require additional changes to the software. For example :

1. improve the report on the results of the Join with more statistics: how many matches, possible mismatches or errors, etc.
2. manage the caching of previous Join Data results. Cached results could for example: be removed automatically after a certain time, or if a storage limit is used or manually. The functional requirements need to be defined for this.
3. add authorization to the TJS, that goes beyond default GeoServer authorization on requests (e.g. on specific Spatial Frameworks for Join Data requests).

TJS as an official GeoServer extension

For GeoServer to become an official extension, following has to be done on the software and documentation (NOTE: also some community building and resources are required, see <http://docs.GeoServer.org/stable/en/developer/policies/community-modules.html> for the process and requirements):

1. cleanup and translate code comments to English (there are some Spanish comments for example, from the original Cuban developers)
2. add automatic tests (test coverage is low at the moment)
3. check IP violations for possible violations on the GPL license
4. write a user manual in the GeoServer docs
5. improve the build process (in Maven)

The GeoServer community also requires a (at least small) community around the plugin (source <http://docs.GeoServer.org/stable/en/developer/policies/community-modules.html>):

***“The module has at least a “handful” of users***

*In order to avoid cluttering the main code base, only those community modules which are of interest to at least 3 users (this may include the maintainer) are promoted.*

***The module has a designated and active maintainer***

*Every core and extension module requires a module maintainer. The job of the maintainer is to fix bugs and address issues which arise with the module. If a community module is promoted and the maintainer “drops off”, the module is in danger of being demoted back to community status.*

***The module is considered “stable” by the majority of the PSC***

*A module will only be promoted if it is deemed “stable” by the majority of the Project Steering Committee (PSC). Those PSC members deeming it “unstable” must provide a reasonable justification for the assertion.”*

*This means that resources have to be available and guaranteed and the module is used (in production environments) at multiple places. If this is not the case, the extension can still be part of GeoServer's code base, but is not supported by the community then.*

#### **4.1.2 Scenario 3**

Global indication of activities for this scenario would be 50 days implementation (only once) for the functionality as described in paragraph 2.3.

A part of the maintenance is included in the standard PDOK maintenance which leaves 20 days a year maintenance.

## **4.2 Conversion to required format**

The assumption is that the conversion software is available as Java software since this fits the PDOK development team.

#### **4.2.1 Scenario 1 and 2**

By specification, TJS Server only supports the file format GDAS for the statistical data. This file format is not very common. To make the use of TJS more user-friendly, a conversion has to take place. In the current implementation this is handled by a stand-alone, stateless web service developed in Python.

The service takes as input CSV, SDMX or ODATA. The data has to be served as a static file or service. The output is a valid GDAS file, which can be used as input for the TJS server.

The service can be deployed on a Docker host (recommended), or installed manually. The instructions for this can be found on Github: <https://github.com/joostvenema/to-gdas>

The requirements and actions necessary depends on the final decision of the implementation (stand-alone or as a part of GeoServer). For the current implementation a rough estimate of 5 days is needed to clean up code, documentation and tests. To translate the conversion software into Java another 10 days is needed.

#### **4.2.2 Scenario 3**

In case of scenario 3, the conversion software has to be rewritten in Java to convert from CSV, SDMX or ODATA into the database format as is in use within PDOK infrastructure.

So in fact the impact for conversion is the same for all 3 scenario's.

## **4.3 Web Application**

#### **4.3.1 Scenario 1**

To demonstrate the functionality of the TJS, a web application has been developed in this project. The application is a fork of the Stage client, developed by the Geodetic Institute of Slovenia: <https://bitbucket.org/ginst/jstjs>.

The current state of the application is for demonstration purposes only. Modifications to the original code have been kept to a minimum. As a result the code doesn't follow any guidelines and maintainability is difficult (e.g. some parts are hard-coded).

<https://github.com/joostvenema/jstjs>

This web application will not be adapted or maintained by the PDOK (PDOK policy). So once we decide that we need a fully maintained client application, it has to be fully redesigned. Costs are estimated roughly at €50.000 once and €10.000 yearly?

#### **4.3.2 Scenario 2**

If no client application is built, app builders could be stimulated to do so by showing them the demonstration version.

Without client applications, one could still call the TJS service via a browser.

#### **4.3.3 Scenario 3**

In this scenario the client software is part of the self-service application that PDOK is developing any way.

### **4.4 Operations / Maintenance**

Both the conversion- as the web application should be straight forward in terms of maintenance. The web application is static code and can be deployed on a standard webserver. The conversion application is stateless and doesn't persist data on disk. Maintenance of the server side software has been described in chapter 4.1.

## 5. Impact data

### 5.1 Geo data

#### 5.1.1 The Dutch Cadaster

The Dutch Cadaster is responsible for the administrative units in the Netherlands. In principle each administrative unit can be used as a statistical unit and therefor used for the TJS.

In theory this is possible, but in practice not yet. The Dutch Cadaster uses nearly the same key code as Statistics Netherlands, only they miss the 2 characters in front of the code. These 2 characters identify the type of code.

Most of the statistical units provided by Statistics Netherlands have been derived from the administrative units from the Dutch Cadaster.

#### 5.1.2 Statistics Netherlands

Statistics Netherlands (SN) provides a geoservice for statistical Units. It is called CBSgebiedsindelingen and can be found via the following Link:

<http://www.nationaalgeoregister.nl/geonetwork/srv/dut/search#|effe1ab0-073d-437c-af13-df5c5e07d6cd>

Five layers from this datasets connected WFS were used for the POC that was the basis for this impact analysis. The dataset contains about 25 region types per year since 2003. They are served in generalized (100m) coordinates based on the Dutch coordinate system RD.

The impact for SN is that the dataset has to be updated twice a year. The first of January and the first of July. Borders of communities are the basis for most regions used. The borders change every year. The borders of the smaller regions (neighborhoods and districts) are changed during the first months of a year, so that is why this is updated on the first of July.

At this moment only generalized borders are served for reasons of performance. In the future SN is considering to also publish the original coordinates without generalization.

Secondly, also a harmonized dataset according to INSPIRE will be served.

Thirdly, SN considers publishing other regions, like grids and postal code zones.

#### 5.1.3 Europe

To be able to create maps from Europe, the Nuts regions should be available to make a TJS work for Europe.

At this moment, not all borders are freely available for all European countries. It should be investigated, if it is possible to set them available at a certain level of generalization.

### 5.2 Tabular data

#### 5.2.1 Statistics Netherlands

SN serves all their tabular data open and machine readable via the CBS Open data Portal:

<https://www.cbs.nl/en-gb/our-services/open-data>

The data is published in such a way that programmers can use them directly for various applications. Many new types of services and apps can be developed based on (combinations of) such data. These will often be apps for smartphones and tablets, but applications for

personal computers or other devices can also be developed. The Table Joining Service can be seen as such an application that uses this tabular data to join with the open geographical data.

The impact of a TJS for SN open tabular data is that it becomes more important to standardize the data. During the experiments with the TJS it was noticed that not all tabular data with geocodes used the same way to specify that they had a geometric dimension. Because of this not all tables could be mapped.

Another issue is that the Odata had a maximum of 10.000 records, where the largest amount of regions is 12.000 (neighborhoods). So hopefully this can be changed.

### 5.2.2 Europe

Eurostat also publishes European tables in SDMX format and a machine readable way:

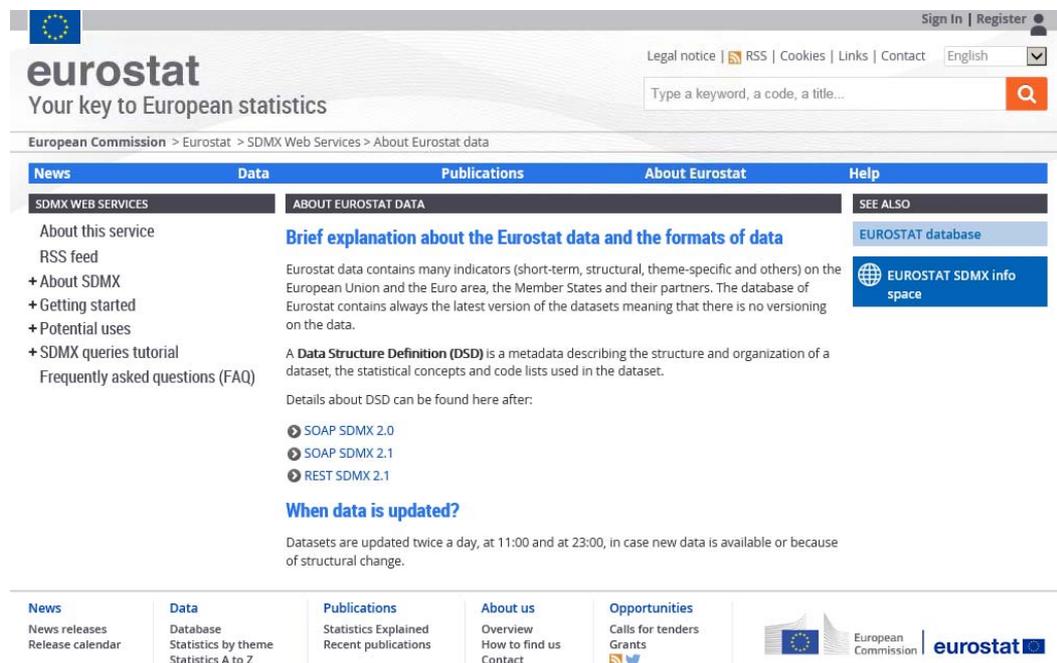


Figure 4: <http://ec.europa.eu/eurostat/data/web-services/>

In the ELF project it has already been shown that it is possible to join these tables with NUTS regions from the Euro boundary map. In this proof of concept, Eurostat data has not been tested.

### 5.2.3 CSV files

Within the current proof of concept it is possible to upload a CSV (comma separated value) table as long as it is positioned on an accessible server. In the case of the Netherlands, one can use the PDOK Datastore (<https://geodatastore.pdok.nl/web/dut/index>) to upload CSV files. Their URL's can be directly copied into the TJS client application with an online map as a result.

## 5.3 Identifier management

Identifiers play a crucial role in making the Joins with TJS. If identifiers are not used consistently, matches may fail. For example: if the statistical data uses lower case identifiers and the spatial framework upper case or if certain prefixes (e.g. "GM0307" vs. "0307") or trailing characters (e.g. 307 vs. 0307), matches might not be successful.

For statistical data, time is also important. If a join is made for data of a certain year (say 2015) to a spatial framework of another year (say 2012), besides some mismatches (because objects may have been changed from 2012 to 2015), the join may lead to misinterpretations of the data too.

Therefore it is important to use consistent and truly unique identifiers, including versions if relevant. Using URIs (including versions) as identifiers across all data (statistical and spatial) could be a very robust solution. Also, INSPIRE is advising to use URIs:  
<http://inspire.ec.europa.eu/index.cfm/pageid/5120>.

## 5.4 Metadata

The only meta data that we need extra is the meta data of the Table Joining Service. This would be meta data of the type of a Spatial Data Service (SDS) as INSPIRE obliges us to publish. It has no use to create metadata for the temporary outcome of a TJS.

Of course the meta data of the input should be available. For the geometry this is already the case in the Netherlands. All metadata of geo-data is stored in the National Geo-register:  
[http://www.nationaalgeoregister.nl/geonetwork/srv/dut/search#fast=index&from=1&to=50&any\\_OR\\_geokeyword\\_OR\\_title\\_OR\\_keyword=cbsgebiedsindeling\\*&relation=within](http://www.nationaalgeoregister.nl/geonetwork/srv/dut/search#fast=index&from=1&to=50&any_OR_geokeyword_OR_title_OR_keyword=cbsgebiedsindeling*&relation=within)

The tabular data also has its own metadata in SDMX format when it concerns the European SDMX service.

The metadata of the Open data portal can be found in a catalogue:  
<http://opendata.cbs.nl/OdataCatalog/>

Both the National Geo-register and the CBS Catalogue are automatically harvested by <http://data.overheid.nl> or specially for CBS: <https://data.overheid.nl/data/dataset?q=cbs>

## 5.5 Maintenance

The maintenance of the data is not influenced by the introduction of the TJS on the short term. On the longer term it could even be reduced, because less geoservices are needed. For instance, the publication of statistics per neighborhoods as a geoservice will not be necessary any more, since users can create it by themselves with the help of a TJS.

## 6. Impact organizations

This chapter describes the organizational impact for the involved organizations, but also for noninvolved organizations like other European countries or other potential users.

### 6.1 PDOK

The organizational impact for PDOK is that it has to organize

- Monitoring
- (re) acting to secure performance and usage
- Customer care
- Extra testing
- Reacting to incidents/bugs
- Possible role in communities
- Additional development
- Creating communication and awareness

### 6.2 Statistics Netherlands

The organizational impact for Statistics Netherlands is that SN has to organize:

- Continuation of the geoservices
- Continuation of the Open data.
- Improvement of standardized geocoding in the open data
  - Increase max records
  - Standardize geokey-codes in line with INSPIRE
- Involvement in PDOK
- Stimulate the TJS usage
- Financial support
- INSPIRE harmonization
- Connection to Eurostat

### 6.3 Dutch Cadaster

The organizational impact for the Dutch Cadaster is that they have to ensure that:

- The geo-services are up to date and available (e.g. boundaries of municipalities, provinces)
- the geo-services have the right content and geokeys
- the geo-services are available as open data
- the geo-services are INSPIRE compliant where needed
- PDOK service management is guaranteed

### 6.4 Geonovum

In the Netherlands (and elsewhere) are many organizations involved in the recording and exchange of information with a geographic component. That is to say, information about objects that are related to a location with respect to the earth's surface. Within these different applications domains can be distinguished, such as cadastral information and information about water management, environment but also statistics. To ensure that the geo-information management of these areas are well connected to each other, and that information can be exchanged between domains, agreements about the standards to be used are needed. The set of Geo-standards maintained by Geonovum provides for this.

As described in paragraph 1.9.2, Geonovum is the national geospatial standards organization responsible for the organization and adoption of these standards in the Netherlands. For this

purpose the Netherlands has a list of approved geospatial standards in the public sector under a 'comply-or-explain' regime. The 'comply-or-explain' geospatial standards are available by: <http://www.geonovum.nl/sites/default/files/Raamwerk%20Geo-Standaarden%20v3.0.pdf>

One of the roles of Geonovum as the executive committee for the Spatial Data Infrastructure (SDI) in the Netherlands is the evaluation and reflection of the use of geospatial standards in the field, as well as exploring new standards developments and performing pilots on testing the adoption of standards in the Netherlands. OGC TJS is one of these new, although a not very well adopted OGC standards, that was worth exploring (see Grothe and Brentjens. 2013). In case of adoption of such a geospatial standard we foresee different activities and tasks from the standardization point of view that Geonovum needs to take care of. The impact of OGC TJS for Geonovum from an organizational point of view consists of the following tasks:

1. Community building around the users of TJS implementations;
2. Providing support to the geo-community using TJS, e.g. by supporting pilots;
3. Awareness creation about the availability of the TJS implementations, e.g. by presentations;
4. Providing feedback to the international standardization organization OGC about the functionalities, performance and use of the standard; and
5. Organization of the community to explore the position of OGC TJS related to the national geospatial standards framework and maintenance of this position.

References: Grothe, M. and T. Brentjens (2013) Joining tabular and geographic data – Merits and possibilities of the Table Joining Service, Geonovum report. Download: <http://www.geonovum.nl/documenten/rapport-joining-tabular-and-geographic-data-merits-and-possibilities-table-joining>.

## 6.5 Other potential users

Other Dutch organizations should also start to use the TJS. Organizations that could benefit are:

- Governmental institutes (RIVM, RVO, PBL, ...)
- Universities, students
- The media
- Civilians

## 6.6 Other EU organizations

In order to be able to project the impact for the Dutch situation to other EU countries, one should consider the differences. This can be done by looking at the following aspects:

- Who is going to host the TJS and the client application?
- Are geoservices for the geometry of statistical and administrative units available as open and free data? If not, one could consider the option for generalized geometries and set up a WFS for the required geometries.
- Are tables available as open, free machine readable data. If not, one could concentrate on the CSV upload functionality
- What organizations are responsible for publishing INSPIRE data with no connected geometry? Try to work together with these organizations.
- In case of larger countries, larger or more servers might be needed.
- One should also consider the amount of statistical units. If it becomes too high, performance can be slow

## 6.8 Central EU implementation

A combined European action by means of a TJS, operating on existing table services (Eurostat SDMX web services) and map services, is a very cost-effective way to meet the INSPIRE goals: Making European geographical information better available for the public in an harmonized way. Of course, Statistics Netherlands is prepared to share its knowledge on this subject.

Since the information that INSPIRE asks for is already available as machine readable data in SDMX files, and since this data is already harmonized in a technical and semantic way, there is no need to do this all over again in another format like suggested in INSPIRE.

Officially, all statistical offices of the member states are responsible for the publication of their map service according to INSPIRE. This would mean that 28 statistical offices would create map services for all their tables that fit the INSPIRE themes for Population and Demography, Human Health and Energy. This is a very large and costly effort, that can be avoided once a combined European action is undertaken by means of a TJS operating on the Eurostat SDMX web services. Europe should also make efforts to set the Euro boundary map available as open data in a central web feature service. Once a country does not want the exact coordinates of their boundaries freely available on the Internet, one could consider less accurate generalized coordinates. This, in combination with the machine readable SDMX-files and a centrally hosted TJS would make all the SDMX files available as maps in an INSPIRE way.

## 6.9 Global view

On a global level, we miss the driving force of INSPIRE, so the need for TJS might be lower.

Although, there is a Canadian implementation of a TJS, in fact, that is where the idea originally came from.

On a global level we could also consider the Statistical Geospatial Framework (SGF) as promoted by the UN-GGIM. The TJS perfectly fits in the 4<sup>th</sup> principle of the SGF: Interoperable data and meta data.

## 7. Financial impact

This chapter describes the financial impact. Below the roughly estimated costs per scenario are listed.

### Scenario 1

Project management	€20.000	Once
Software server side	€29.000	Once
Software conversion	€15.000	Once
Client application	€50.000	Once
Hardware option 1	€37.600	Each year
Hardware option 2	€18.800	Each year
Maintenance hard- and software	€40.000	Each year
Geonovum	€12.000	Each year

### Scenario 2

Project management	€10.000	Once
Software server side	€29.000	Once
Software conversion	€15.000	Once
Hardware option 1	€37.600	Each year
Hardware option 2	€18.800	Each year
Maintenance hard- and software	€30.000	Each year
Geonovum	€8.000	Each year

### Scenario 3

Software	€50.000	Once
Maintenance	€20.000	Each year
Geonovum involvement	€6.000	Each year
Statistics NL involvement	€6.000	Each year

\*Extensive use of TJS will lead to extra hardware with a minimum cost indication of €18.800,- a year.

Scenario	Once	Each year
Scenario 1 option 1	€114.000	€89.600
Scenario 1 option 2	€114.000	€70.800
Scenario 2 option 1	€54.000	€75.600
Scenario 2 option 2	€54.000	€56.800
Scenario 3	€50.000	€32.000

The scenarios have been described in chapter 1.8 and the 2 options have been described in 3.1. Which organizations will pay for this, has not yet been agreed on.

### 7.1 PDOK

PDOK normally charges organizations for publishing geoservices.

In the case of a TJS, usage is supposed to be free and the output services are temporary.

That means that other funding has to be agreed on to realize the TJS implementation.

## **7.2 Statistics Netherlands**

Statistics Netherlands is one of the partners that benefit most by the existence of a TJS. Therefore it is likely that a substantial part of the costs will be provided by Statistics Netherlands. Though, also other organizations might benefit from the TJS. In case of scenario 1 and 2, Statistics Netherlands can provide project management. In case of scenario 3, this is more likely to be within PDOK. SN could still be involved as a member of a user board.

## **7.3 Dutch Cadaster**

For the coming years, extension of available geo-services can be foreseen, joining different geometries with statistical information. This means mostly manpower to extend services and keep expertise up-to-date.

## **7.4 Geonovum**

The impact for Geonovum financially to execute the mentioned activities in paragraph 6.4 is estimated at about approximate 8 to 12 days each calendar year. Financial impact is approximated at 12.000 Euros (ex VAT).

## **7.5 Other Organizations**

Other organizations will also benefit from a TJS. Hopefully they will therefore also start supporting the hosting at PDOK financially. These potential organizations, are mentioned in chapter 6.5.

## **7.6 International**

Whether the cost for the Netherlands can be compared to other countries, is depending on the aspects mentioned in paragraph 6.6.

When a TJS is implemented for the whole of Europe, concentrated on the Nuts regions, one will also need larger or more servers. Still this is a preferable situation, since many countries could save money if there is a central TJS for Europe. This would release many statistical offices from their INSPIRE obligations for the theme of Population Distribution and sometimes Human Health and Energy Statistics.

## 8. Lessons learned and conclusions

1. It will be difficult to stimulate the use of the TJS, just like it was difficult to stimulate the use of WMS and WFS. Even those standards are still considered difficult by app-builders.
2. Building of applications on top of the TJS should be stimulated.
3. Some functionalities are hard to implement, like the way it should deal with mismatches and n:m relations. As far as we know, the current standard for the TJS has no options to deal with that properly. This should be discussed with the OGC.
4. The fact that a WMS or WFS resulting from a TJS is temporally, could be a weak point for the use, because sometimes you want them to be forever. If that is the case we have to think of a way how to pay for the service.
5. Standards for the same key use are of great importance. Even a lower and upper case issue could disturb the process of joining.
6. A major risk is when users try to link tables to the wrong geometries. Especially with administrative units that change every year. Users should be warned for this.
7. Geometry should be available as open data which is not always the case in Europe.
8. The TJS specification is relatively new and not widely used yet. This makes implementation more difficult; customizing software is required at this level. In addition, the specification itself needs some improvements. For example, to support warnings for mismatches.
9. Low performance of WFS for large datasets can be a barrier for the TJS. PDOK has a limit of download of 15.000 records, so at this moment not all data sets are suitable for a TJS. But PDOK is working now on increasing the limit number of downloads as it is also requested internationally to increase this number.
10. On the short term, scenario three seems easier to implement and maintain for PDOK and therefore it is expected to be cheaper. On the longer term it might be a good choice to stimulate the GeoServer community to accept the TJS plugin as suggested in scenario one and two.
11. There are different opinions on the best scenario choice. This means that a discussion is needed on how to proceed.
12. A central European TJS could be a very cost effective way to realize the INSPIRE goals.

## 9. Recommendations

This chapter describes recommendations for future steps as supposed by the involved organizations. One can read contradictions in these recommendations. This is something we will have to discuss in the near future.

### 9.1 PDOK

The PDOK architecture uses a distinct separation of concerns for processing data and provisioning data. The TJS extension is not in line with these conventions. This poses problems in governance of GeoServer with a TJS extension. The alternative (scenario 3 as mentioned before in this document) would be, using the OGC TJS interface with an implementation which fits the PDOK architecture by uploading the dataset and choosing the associated geographic data. This alternative is strongly recommended by PDOK architecture.

In order to be able to publish data for presentation of tabular data on a map this can best be achieved by making use of the standard PDOK infrastructure, instead of introducing alternative implementations of GeoServer with not supported software (the GeoServer TJS plugin). By using the standard infrastructure the load of the new functionality, showing tabular data on a map, can be divided over all available PDOK services. Managing the created datasets is similar to managing all other datasets.

### 9.2 Statistics Netherlands

- The involved organizations should make a choice in the 3 scenarios.
- Funding must be raised for the chosen scenario.
- The demo version should stay online until the final implementation of one of the 3 scenario's.
- INSPIRE compliancy is the next step after implementation.
- It should be investigated, if it is possible to set the Nuts units available at a certain level of generalization for the whole of Europe.

### 9.3 Dutch Cadaster

- Extend the number of possible geometries (postal code areas, etc.) to be combined/joined with statistical information and make these available in the required services .
- Start investigating the possibilities of combining statistical and geospatial information as linked data.
- Start promoting existence of TJS in PDOK within Dutch government and for web developers.

### 9.4 Geonovum

- Develop a strategy for identifiers and implement this in all the datasets involved (see the section Identifier management). URIs could be a good starting point.
- Although TJS is relatively new and not much used so far, to focus on using this specification is worthwhile. In general, implementing standards lowers the risk on a vendor lock-in. The TJS specification could be improved on some points; we recommend to actively provide feedback to OGC on improving TJS and if needed participate in a working group for a new version of TJS;

- If GeoServer is chosen to implement the TJS service on, we recommend to be actively involved in development and maintenance of the TJS plugin. This attracts other organizations (interest has already been shown by several organizations), builds the community and in the end will probably lower the costs because other organizations could contribute too. GeoServer is widely used and has a strong community, so it is a good choice as a platform.

## 10. Acknowledgements

The authors of this document would like to thank the following institutes:

- Eurostat for rewarding the Grant
- Slovenian Geodetic Institute for the open source behind Stage
- All participants in the project
- RIVM, RVO and PBL for testing and input regarding functionality.