# SDI development to lower the impedance of the data exchange network

Glenn Vancauwenberghe
KU Leuven/SADL
Celestijnenlaan 200E
Leuven, Belgium
glenn.vancauwenberghe@sadl.kuleuven.be

**Abstract**

Network theory provides a valuable framework and appropriate tools to study the use and exchange of spatial data in an SDI environment. From a network perspective, an SDI can be defined as the collection of arrangements and initiatives that give shape to a network of spatial data relationships. Previous research demonstrated that the impact of these arrangements and initiatives can be observed in the realization of data flows between producers and users. In this paper, it will be argued that SDI arrangements also shape the network by determining the characteristics of the data flows. Several of these characteristics are related to barriers that hinder or obstruct the exchange and use of data, and can thus be considered as indicators of the impedance of a data flow. The analysis of these characteristics provides valuable insights on the strengths and weaknesses of SDI arrangements and helps us understand why different arrangements coexist next to each other.

*Keywords*: network perspective, SDI, spatial data flows, network impedance

## 1 Introduction

The concept of a Spatial Data Infrastructure (SDI) entails a shift away from the commonly isolated and scattered approaches to the production, management, dissemination and use of spatial information towards the use and sharing of spatial information in an open environment characterized by collaboration and interoperability [5]. While an SDI can be considered as the collection of technological and non-technological components and arrangements intended to create such an open environment, the SDI concept also reflects the increasing importance of interconnectivity in working with geographic information.

The need to better understand and conceptualize social and societal developments has led to the introduction of the network perspective in several scientific disciplines. While traditional theories and approaches assume that actors think and act without concern for others, theories based on the network perspective stress the importance of an actor's position within a broader social network populated by other actors [4]. In network theory, the structural relations with other actors are more significant than individual characteristics in determining actor behaviour. These relations, and their dimensions and content in particular, are the main research object in network analyses. As such, network theory represents a shift, away from individualistic explanations, towards a more relational, contextual and systematic approach [3].

The objective of this paper is to further demonstrate the value of a relational, contextual and systematic approach for analyzing the use and exchange of spatial data in the context of an SDI. The network perspective on SDI was introduced as a perspective to characterise and underpin the evaluation of Spatial Data Infrastructures that makes it possible to describe how data flow between actors in the network and how individual actors behave within the network [8]. An application of the network perspective on the exchange of spatial data in Flanders demonstrated how different – formal and informal - SDI arrangements give shape to a network of spatial data relationships between users and producers of spatial data. This paper examines how different arrangements also have an impact on the characteristics of these data relationships.

## 2 Network perspective on SDI

Any network encompasses two indispensable elements: actors and relations [2]. Applied to the use and exchange of spatial data, the organizations which use, produce, maintain and/or distribute spatial datasets constitute the actors, while the ties are formed by flows of spatial data between these organizations. From a network perspective, an SDI can be defined as the set of arrangements and initiatives that shape the network. These arrangements and initiatives can intervene in and shape the network in various ways. They can lead to the introduction of new actors in the network, the creation of new data flows between these actors or the removal and replacement of existing data flows. In addition, an SDI can also include measures aimed at modifying the behaviour of actors. Revising organizational structures, introducing new tools and applications and providing training and support are some examples of how an SDI can impact the behaviour of actors and the configuration of the network.

SDI arrangements and initiatives also intervene in the network by determining the characteristics of the data relationships. Changes in pricing, in the terms of use or in the transfer method modify the characteristics of the network. Many of these changes are related to barriers that might impede the use and exchange of data. One of the key objectives of an SDI should be to remove all barriers that hinder or prevent the exchange of data, and reduce the impedance of the network of spatial data flows. Measuring the network impedance can help us understand the strengths and weaknesses of different arrangements and initiatives within an SDI.

# 3 Methodology

## 3.1 Network analysis

Network theory provides a new framework to study the use and exchange of spatial data, as well as the impact of SDI's. By providing specific tools and measurements, network theory also offers a methodological contribution. These tools and measurements can be used to analyze the structure of existing networks. The methodological contribution of a network perspective for analyzing spatial data infrastructures is demonstrated by a network analysis of spatial data exchanges in Flanders. This analysis makes use of empirical data that were collected through on online questionnaire among public organizations in Flanders [1]. All regional, provincial and local authorities in Flanders were invited to participate in this survey. In addition, a selection of inter-municipal organizations was also invited. Private organizations and non-profit organizations were left out of the research population, as the focus was on the exchange of spatial data within the public sector. In total 189 organizations were included in the network analysis.

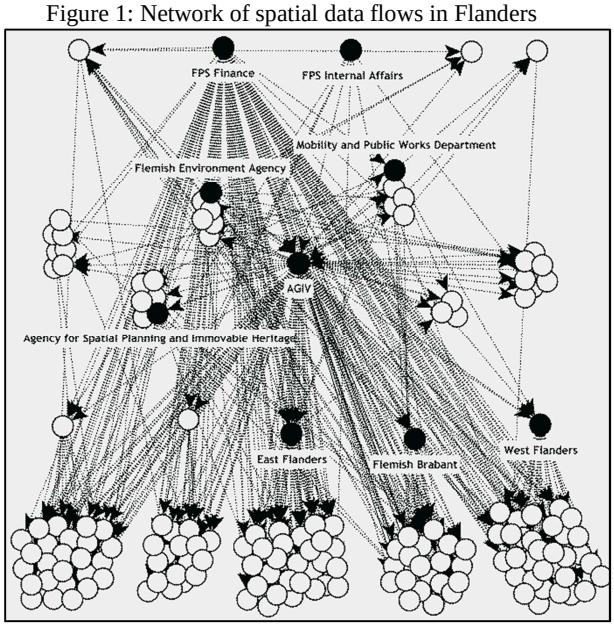
The primary objective of this analysis was the identification of spatial data flows among public organizations in Flanders. In order to identify these data flows in a clear and unambiguous manner, four specific types of spatial data were selected: parcel data, address data, road data and hydrographic data. The resulting network of spatial data exchanges in Flanders was analyzed both in a graphical and a mathematical manner [6].

## 3.2 SDI network in Flanders

Figure 1 presents the graphical representation of the network of spatial data exchanges in Flanders. In this figure, all organizations are represented as nodes, while their relational ties are represented by lines. The actors in this network are positioned by their administrative levels. Four levels can be distinguished: 1) the Belgian federal organizations (at the top), 2) the organizations of the Flemish public administration, 3) the five provinces and 4) the municipalities (at the bottom row). Between these organizations, 362 lines can be discerned. These lines refer to the exchange of at least one type of spatial data. Even organizations that exchange more than one type of data, are linked with only one line in Figure 1.

The graphical representation and the centralization degree scores of the network demonstrated that the re-use of spatial data in Flanders is strongly scattered around all the involved actors, while the distribution of data is centralized around a limited number of actors [6]. The centrality measures indicated that the AGIV, the coordinating body of the central SDI initiative in Flanders, can be seen as the most central actor in the Flemish spatial data network. The AGIV is responsible for 45,9% of registered data flows. Other central actors in the network are the Federal Public Service Finance (21,3% of the data flows) and the Flemish provinces (together responsible for 14,2% of the data flows). The residual 18,6% of the data flows are governed by different – formal and informal – arrangements.

Figure 1: Network of spatial data flows in Flanders




## 3.3 Data flow characteristics

Based on the network analysis of spatial data exchanges in Flanders, four groups of data flows can be distinguished: 1. the *AGIV* data (302 data flows), 2. the *FPS Finance* data (141 data flows), 3. the *Provincial* data (94 data flows); and 4. the *other* data (123 flows). Three of these groups can be linked to an SDI arrangement in Flanders [7]. The *AGIV* operates as a central data distribution hub of the partnership 'SDI-Flanders', which is the central SDI arrangement in Flanders. The *FPS Finance* data flows are the result of the provision of cadastral data by the Federal Public Service Finance. The *Provincial* data flows are mainly arranged by different partnerships between the provincial administration and the municipalities in the province.

While the primary aim of the survey was the identification of spatial data relationships in Flanders, information was also collected on several characteristics of these relationships. For each data flow that was identified in the survey, information is available on the price, the legal base, the transfer method, the need for additional preparations and the existence of use restrictions. It is interesting to see how these characteristics are related to barriers that can hinder or prevent the exchange and use of spatial data. In that manner, these characteristics can be considered as indicators of the impedance a certain data flow is confronted with. A high level of impedance means that the flow of data between a producer and a user is obstructed by many barriers.

# 4 Results

To deal with the heterogeneity between the number of answer options in the survey, all information related to the data characteristics was converted to a 3-point scale, with

values 0 (low impedance), 0,5 (medium impedance) and 1 (high impedance). Table 1 uses the mean values for the five characteristics to make a comparison between the four groups of data. The higher the mean score in this table, the higher the level of impedance. In addition, table 2 presents for each characteristic the distribution of all groups of data over the three levels of impedance. Chi-square tests revealed that there were significant differences between the four groups on all five characteristics ($p < 0.000$).

### 4.1 Price

A first difference between the four groups of data is related to the price the receiver/user has to pay for the data. In the analysis, the level of impedance is considered as low (0) when data are fully free and no price needs to be paid. Impedance is high (1) when the full price has to be paid for the data. If data are acquired by paying a (small) delivery cost or in exchange for other data, the level of impedance is considered as medium. Table 1 shows that the price is especially a barrier for the exchange of *AGIV* data and *other* data. In both cases, the impedance caused by the price that has to be paid, is mainly the result of a small delivery cost that the users has to pay. Data for which users need to pay the full price, are rather limited. For what concerns *FPS Finance* data and *Provincial* data, the impedance as a result of the price is relatively low, as most of these data can be acquired for free. The comparison of the mean scores for the different characteristics demonstrates that impedance caused by the price is lower compared to impedance caused by the four other barriers in the analysis. However, caution is needed when comparing the mean scores for different characteristics, as these mean scores are influenced by methodological choices.

### 4.2 Legal base

A second characteristic of the data flows is related to the legal basis under which the data flow is arranged. Here, a distinction can be made between data flows that are informal or that are fully governed by regulations (low level of impedance), data flows that are governed by a simple exchange contract or a registration procedure (medium level of impedance) and data flows that require an individual purchase contract (high level of impedance). The mean scores in table 1 demonstrate that barriers in terms of the legal base of the data flow are especially significant for the exchange of *AGIV* data. This is mainly due to the need to complete a registration procedure or to sign an exchange contract for each data flow (55,9%) and in less extent to the use of data purchase contracts (7,9%). The same goes for the other three groups of data. For what concerns *AGIV* data, *FPS Finance* data and *Provincial* data, more than half of the registered data flows are governed by a simple registration procedure or an exchange contract. For what concerns the group of *other* data, the majority of data flows are informal data exchanges or are governed by regulations (57,7%).

### 4.3 Transfer method

The transfer method of the data can also be seen an indicator of a particular type of impedance that hinders or slows the flow of spatial data. Like for the other indicators, three levels of impedance are discerned in the analysis. Impedance is low if data are obtained by download. Impedance is considered as medium if data are transferred on a flash device, DVD, CD-ROM or other media. If data are consulted online in a web viewer, the level of impedance is considered as high. Of all four data groups, the *AGIV* data are exposed to the lowest level of impedance. Although most of the *AGIV* data are transferred on DVD or CD-ROM (78,3%), in some cases these data can be downloaded (17,3%). Compared to the *AGIV* data, the so-called *other* data are even more often acquired by download (23,7%). However, a large amount of the *other* data can only be viewed online (17,1%), which leads to a higher mean score for this group of data. Impedance as a result of the transfer method is the highest for *FPS Finance* data.

### 4.4 Need for preparations

While some spatial data can be used immediately in different processes, other data require additional preparation and processing before they can be used. In some cases these additional preparations and processing can be relatively limited and easy, for other data the preparation of data might be more difficult and time-consuming. The need for additional preparations leads to an higher impedance of data to flow. In the analysis, a distinction is made between data flows that do not need any preparations (low impedance), data that require some minor preparations (medium impedance) and data for which major preparations are needed before they can be used (high impedance). Table 1 shows the differences between the four groups of data flows in Flanders that were identified, and shows that the impedance caused by the need for additional preparations is the highest for *FPS Finance* data. More than $3/4^{th}$ of the *FPS Finance* data require additional minor (54,6%) or even major (20,6%) preparations or processing. For the other groups of data, the need for additional

Table 1: Mean values of different groups of data flows

| | ***Price*** | ***Legal base*** | ***Transfer method*** | ***Need for preparations*** | ***Use restrictions*** | ***Total*** |
|---|---|---|---|---|---|---|
| AGIV data | 0.21 | 0.36 | 0.43 | 0.34 | 0.42 | *1.76* |
| FPS Finance data | 0.08 | 0.28 | 0.50 | 0.48 | 0.56 | *1.9* |
| Provincial data | 0.04 | 0.33 | 0.49 | 0.25 | 0.37 | *1.48* |
| Other data | 0.13 | 0.23 | 0.47 | 0.29 | 0.41 | *1.53* |
| **All data** | 0.14 | 0.31 | 0.46 | 0.34 | 0.44 | *1.71* |

preparation is significantly lower. For instance, less than 3% of the *Provincial* data require major preparations before being used.

### 4.5 Use restrictions

The fifth and final data flow characteristic is related to the data use conditions and, in particular, to the use restrictions these conditions impose. Examples of use restrictions are restrictions on the use for publication on the internet, restrictions on distribution to third parties and restrictions that cover the personal use of the data.. In the analysis, a distinction is made between data flows without any use restrictions (low impedance), data flows with only a few use restrictions (medium impedance) and data flows with multiple use restrictions (high impedance). Based on the mean scores presented in table 1, it appears that the impedance as a result of the use restrictions is the highest for *FPS Finance* data. While for the other three groups of data the percentage of data flows with a high level of impedance is lower than 15%, more than one fourth of the *FPS Finance* data flows has several use restrictions and is thus confronted with a high level of impedance (26,2%).

### 4.6 Discussion

A more general examination of table 1 reveals how data flows that are facilitated and organized by different SDI arrangements are confronted with different types and different levels of impedance. The impedance that hinders the flow of *AGIV* data is mainly caused by the price and legal base of the data exchange. The main barriers for the exchange of *FPS Finance data* are related to the transfer method, the need for preparations and the use restrictions, which all cause impedance to the flow of spatial data. Compared to the *AGIV* data and the *FPS Finance* data, the *Provincial* data and the so-called *other* data are confronted with generally lower levels of impedance.

The mean score for the total level of impedance, which is based on the sum of all five types of impedance, indicates that the *Provincial* data are confronted with the lowest level of impedance. It is interesting to notice how the *Provincial* data flows but also the *other* data flows are confronted with a lower level of impedance than the *AGIV* data flows, which are arranged by the central SDI initiative in Flanders. The lower levels of impedance for both groups of data can be seen as an important reason why these data flows still exist, despite the presence of a central SDI initiative in Flanders. The *provincial* SDI arrangements as well as different arrangements that give shape to the exchange of the so-called *other* data are more

Table 2: Distribution of data groups over different levels of impedance

| | | **Low impedance** | **Medium impedance** | **High impedance** |
|---|---|---|---|---|
| **Price** | AGIV data | 67,8% | 22,7% | 9,5% |
| | FPS data | 87,2% | 9,9% | 2,8% |
| | Provincial data | 91,5% | 8,5% | 0,0% |
| | Other data | 79,7% | 13,8% | 6,5% |
| **Legal base** | AGIV data | 36,2% | 55,9% | 7,9% |
| | FPS data | 45,4% | 53,2% | 1,4% |
| | Provincial data | 33,0% | 67,0% | 0,0% |
| | Other data | 57,7% | 39,0% | 3,3% |
| **Transfer method** | AGIV data | 17,1% | 78,3% | 4,6% |
| | FPS data | 0,7% | 99,3% | 0,0% |
| | Provincial data | 13,8% | 74,5% | 11,7% |
| | Other data | 23,6% | 59,3% | 17,1% |
| **Need for preparations** | AGIV data | 43,4% | 46,1% | 10,5% |
| | FPS data | 24,8% | 54,6% | 20,6% |
| | Provincial data | 52,1% | 45,7% | 2,1% |
| | Other data | 49,6% | 41,5% | 8,9% |
| **Use restrictions** | AGIV data | 28,3% | 59,5% | 12,2% |
| | FPS data | 12,8% | 61,0% | 26,2% |
| | Provincial data | 30,9% | 64,9% | 4,3% |
| | Other data | 26,8% | 65,0% | 8,1% |

successful in removing the barriers for data use and data exchange and minimizing the network impedance. Of all four groups of data, the total impedance is the highest for *FPS Finance* data. These data flows mainly involve the exchange of cadastral data, and exist because the FPS Finance is the authentic source for these data.

## 5 Conclusion

The network perspective on SDI offers a theoretical framework as well as methodological tools to analyze the use and exchange of spatial data in an SDI environment. This perspective can be used to analyze the complex interactions between different arrangements within a network of spatial data exchanges. The concept of network impedance and the measurement of this impedance helps us understand the nature and impact of different SDI arrangements. The analysis in this paper demonstrates how different SDI arrangements remain in place because they are more successful in minimizing the impedance of spatial data flows than the central SDI initiative. In view of this, it is interesting to notice how the further development of the central SDI initiative in Flanders is clearly focused on removing several of the barriers that were discussed in this paper, in order to lower the impedance of the network.

## References

[1] Joep Crompvoets, Ezra Dessers, Tessa Geudens, Katleen Janssen, Glenn Vancauwenberghe, Danny Vandenbroucke and Matthias Van hoogenbemt. *Het GDI-netwerk in Vlaanderen. Een kwantitatieve verkenning van het gebruik en de uitwisseling van geodata in Vlaanderen*. Spatialist, Leuven, 2009.

[2] David Knoke and James H. Kuklinski. *Network Analysis*. Sage, Newbury Park, 1982.

[3] David Knoke and Song Yang. *Social Network Analysis*. Sage, Los Angeles, 2008.

[4] Evelien Otte and Ronald Rousseau. *Social network analysis: a powerful strategy, also for the information sciences.* Journal of Information Science, 28(6): 441–453, 2002.

[5] Mauro Salvemini, From the GIS to the SDI: a design path. In *Proceedings of 7th AGILE conference on Geographic Information Science*, Heraklion, 2004.

[6] G. Vancauwenberghe, J. Crompvoets, G. Bouckaert and D. Vandenbroucke. Social network analysis of the SDI in Flanders. In Z. Nedovic-Budic, J. Crompvoets and Y. Georgiadou, editors, *Spatial Data Infrastructures in Context: North and South*, pages 121-135. CRC Press Taylor & Francis Group, Boca Raton/London/New York.

[7] Glenn Vancauwenberghe. *Coördinatie binnen de Geografische Data Infrastructuur. Een analyse van de uitwisseling en het gebruik van geografische informatie in Vlaanderen.* Faculteit Sociale wetenschappen, Leuven, forthcoming.

[8] Danny Vandenbroucke, Joep Crompvoets, Glenn Vancauwenberghe, Ezra Dessers and Jos Van Orshoven. A network perspective on Spatial Data Infrastructures: Application to the sub-national SDI of Flanders (Belgium). *Transactions in GIS*, 13(1): 105-122, 2009.