NATIONAL SPATIAL DATA CLEARINGHOUSES

WORLDWIDE DEVELOPMENT AND IMPACT

Promotor
Prof. dr. ir. A.K. Bregt
Hoogleraar in de Geo-informatiekunde
Wageningen Universiteit, Nederland

Co-promotor
Dr. ir. G.J. Hofstede
Universitair hoofddocent bij Leerstoelgroep Informatietechnologie
Wageningen Universiteit, Nederland

Promotiecommissie
Prof. dr. ir. A. van den Brink, Wageningen Universiteit, Nederland
Prof. dr. I. Masser, University College London (UK), The University of Melbourne (Australia), Universiteit Utrecht (Nederland)
Dr. A. Rajabifard, The University of Melbourne (Australia)
Prof. dr. H.J. Scholten, Vrije Universiteit Amsterdam (Nederland)

Dit onderzoek is uitgevoerd binnen de C.T. de Wit onderzoekschool
Production Ecology and Resource Conservation

NATIONAL SPATIAL DATA CLEARINGHOUSES

WORLDWIDE DEVELOPMENT AND IMPACT

Joep Crompvoets

PROEFSCHRIFT

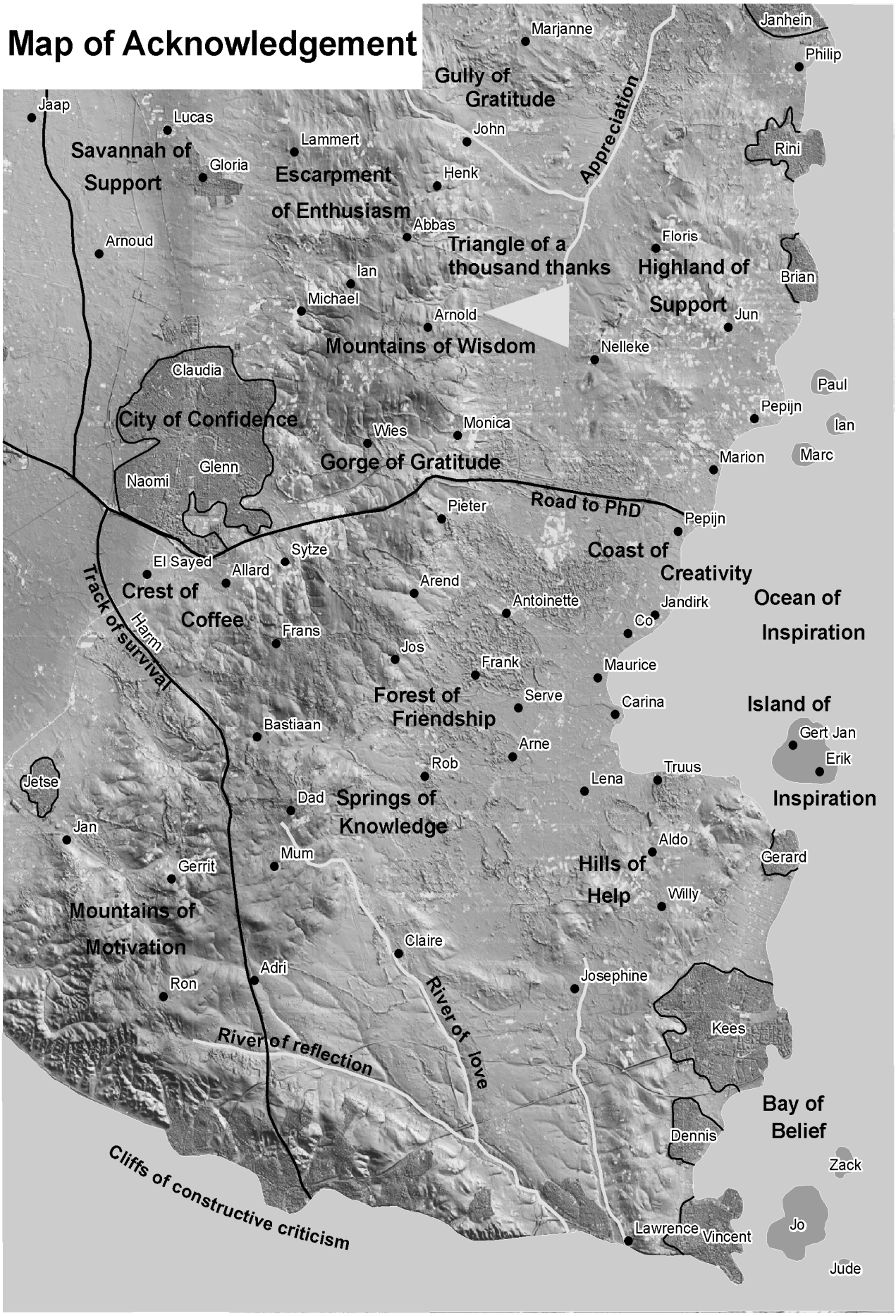
ter verkrijging van de graad van doctor
op gezag van de rector magnificus
van Wageningen Universiteit,
Prof. dr. M.J. Kropff,
in het openbaar te verdedigen
op vrijdag 20 januari 2006
des namiddags te vier uur in de Aula

Joep Crompvoets

National spatial data clearinghouses, worldwide development and impact
PhD Thesis Wageningen University, Wageningen 2006. – with references
- with summaries in English and Dutch

ISBN: 90-8504-341-7

Map of Acknowledgement



Design cover: Jandirk Bulens

Contents

Chapter 1	General introduction	1
Chapter 2	World Status of National Spatial Data Clearinghouses	9
Chapter 3	Assessing the worldwide developments of national spatial data clearinghouses	21
Chapter 4	Worldwide impact assessment of national spatial data clearinghouses	47
Chapter 5	Establishment of national spatial data clearinghouses, exploring worldwide the impact of society	65
Chapter 6	Exploring worldwide the impact of society on the success of national spatial data clearinghouses	79
Chapter 7	Main conclusions and recommendations for further research	97
References		103
Summary		111
Samenvatting		115
Appendix 1		119
Curriculum Vitae		125
List of selected papers written by author		127

Chapter 1

General introduction

1.1 BACKGROUND

We live in an information age. Geographic information in particular is one of the most critical elements underpinning decision-making for many disciplines (Clinton, 1994; Gore, 1998; Longley et al., 1999; Williamson et al., 2003; Morales, 2004). In the past we used maps to show where people and assets were located. This has now evolved into a complex digital environment with sophisticated technology.

Over the last decades, many governments and the private sector have invested tens of billions of Euros in the development of geographic information, largely to serve specific communities (forestry, agriculture, urban/rural planning, land records management, military, security service, health care, development aid, emergency services, retail, etc.), within a local, national, international, and even global framework (Groot and McLaughlin, 2000). At present, the focus is increasingly shifting to the challenges associated with integrating broadly sourced geographic information, so as to create a manageable framework. This has led to the creation of Spatial Data Infrastructure (SDI). This infrastructure facilitates access to the spatial data and services, improving on the existing complex and multi-stakeholder decision-making process (Feeney, 2003). Moreover, it facilitates (and coordinates) the exchange and sharing of spatial data (at a local, national and international level), between stakeholders within the geo-information (GI) community. This community includes mainly those employed by mapping agencies, universities, governmental and non-governmental organisations, and both public and private institutions.

Over the last few years many countries have spent considerable resources on developing their own National Spatial Data Infrastructure, in order to manage and utilise spatial data assets more efficiently, reduce the costs of data production and eliminate duplication of data acquisition efforts (Groot and McLaughlin, 2000; Bernard et al. 2005a; Williamson et al., 2003; Masser 2005). National SDIs have, according to Masser (2005), three common characteristics:

1. They are explicitly national in nature.
2. Refer either to geographic information or spatial data.
3. Imply the existence of some form of coordinating mechanism for policy formulation and implementation purposes.

National SDIs are also facilitating in nature, possess features, which make data accessibility and sharing ‘easier’, and are set up with a view to long-term development.

A key feature of a national SDI is the national spatial data clearinghouse (Clinton, 1994; FGDC, 1997; Coleman and McLaughlin, 1998; Onsrud, 1998; AUSLIG, 2001; Crompvoets and Bregt, 2003). A spatial data clearinghouse can be defined as an electronic facility for searching, viewing, transferring, ordering, advertising and/or disseminating spatial data from numerous sources via the Internet. Such a facility usually consists of a number of servers, which contain information (metadata) about available digital data (Crompvoets, 2002; Crompvoets and Bregt, 2003). It provides complementary services and improves the exchange and sharing of spatial data between suppliers and users.

The concept of a clearinghouse originates from the financial world. With respect to the financial transactions between banks, the clearinghouse keeps the data on mutual indebted amounts. At the end of each day, banks are informed about the final amounts to be transferred between banks. Every day there is a 'clearing' between them (Bogearts, 1997). The first clearinghouse was the London Banker's Clearinghouse, which was established in 1773. The New York Clearinghouse Association described its clearinghouse role in 1853 as to simplify the chaotic exchange between New York City banks (The Clearinghouse Payments Company, 2005). Even today, this clearinghouse regards itself as the place where payments meet, mix and move expeditiously to their final destination. In 1994, the US Federal Geographic Data Committee (FGDC) established the National Geospatial Data Clearinghouse. This aimed to facilitate efficient access to the overwhelming quantity of spatial data (from federal agencies) and coordinate its exchange, with the objective of minimizing duplication (in the collection of expensive spatial data) and assisting partnerships where common needs exist (Rhind, 1999; FGDC, 2000; Crompvoets et al., 2004).

A national clearinghouse for spatial data can be considered as the access network of a national SDI, which focuses on the facilitation of spatial data discovery, access and related services. It is not a national repository where datasets are simply stored. It can be seen as a one-stop shop for all national spatial data, sourced from governmental agencies and/or industrial bodies (Crompvoets et al., 2004). National clearinghouse implementation can vary enormously. The way in which a national clearinghouse is set up depends on technological, legal, economic, institutional, and cultural factors within the territory. These factors determine to what extent the clearinghouse retains control over data. In addition to the national clearinghouse, clearinghouses at a local, state, international, and even global level exist. However, the national clearinghouse differs due to the fact that it is embedded in the National SDI. In April 2005, 83 national clearinghouses were established on the Internet. A few examples of current national clearinghouses are: MIDAS (MetaInformacni Databazovy System), Czech Republic; geodata-info.dk, Denmark; India NSDI Portal, India; Spatial Data Catalogue, Malawi; Russian GIS Resources, Russia; Geocat.ch, Switzerland; and the Clearinghouse Nacional de Datos Geograficos del Uruguay, Uruguay. Those listed, share the same objective, that of discovering and accessing spatial data, through the available metadata.

National clearinghouses are evolving worldwide. These developments have contributed to the realisation of national SDIs. A body of literature has been compiled on national experiences (e.g. Spatial Applications Division, Catholic University of Leuven 2003, conference papers of Global Spatial Data Infrastructure Association 2002-2005). So far, the majority of this literature focuses on the technical aspects of clearinghouses, and does not take into account the evolutionary nature of these electronic facilities. It is important to have a longitudinal perspective when establishing and maintaining clearinghouses. A detailed study of developments of all national clearinghouses worldwide could be an appropriate starting point. This could identify the critical factors behind the success or failure of a clearinghouse. In this way, knowledge could be used for the support of future implementation strategies. Factors for consideration could be societal, for instance legal, economic, technological, historical, cultural, demographic, environmental and institutional characteristics of a country, or clearinghouse-internal, such as the network architecture, availability of view services, type of search mechanisms and funding stability. However it is worth noting that, simply consolidating the best practices of a few well-operating national spatial data clearinghouses (Australia, Canada and USA), gives no guarantee of sustainability for other national clearinghouses. Such best practices cannot necessarily be applied equally in other countries

due to societal differences (Crompvoets and Kossen, 2002; Crompvoets and Bregt, 2003; Delgado et al., 2005). Knowledge collected through a worldwide impact assessment of national clearinghouses in terms of relevance, efficiency and effectiveness could also support strategies. This impact assessment is considered a key tool for improving 'strategy-making' and implementation.

The focus of this thesis is on the development and impact studies of national spatial data clearinghouses worldwide, in order to provide support for the development of strategies for establishing and maintaining clearinghouses. This study identifies the critical factors for future implementations and help overcome current obstacles. At this moment, no comprehensive study exists. This thesis attempts to fill the gap. Since national clearinghouses form such a key feature of national SDIs, the results of such research could enhance and innovate national SDIs by improving the availability and accessibility of spatial data and services.

1.2 AIM AND SCOPE

The main objective of this thesis is to analyse the development and impact of national spatial data clearinghouses worldwide as well as the impact of society on these facilities. The results obtained provide support for the development of strategies for establishing and maintaining clearinghouses. To achieve this objective, the following five sub-objectives have been formulated:

1. To assess the worldwide status of national spatial data clearinghouses.
2. To assess the worldwide developments of national spatial data clearinghouses.
3. To assess the impact of national spatial data clearinghouses on society in particular the GI-community.
4. To explore the societal impact on the establishment of national spatial data clearinghouses.
5. To explore the societal impact on the success of national spatial data clearinghouses.

In the context of this thesis, society is considered to be the sum of human conditions and activity regarded as a whole, functioning interdependently (at national level) (Rogers, 1995). The above presented assessments and explorations provide a basis for identifying critical factors that affect national clearinghouses. This knowledge can in turn be used to improve the use, content and management of (future) clearinghouses, and so provide support for the development of clearinghouse strategies.

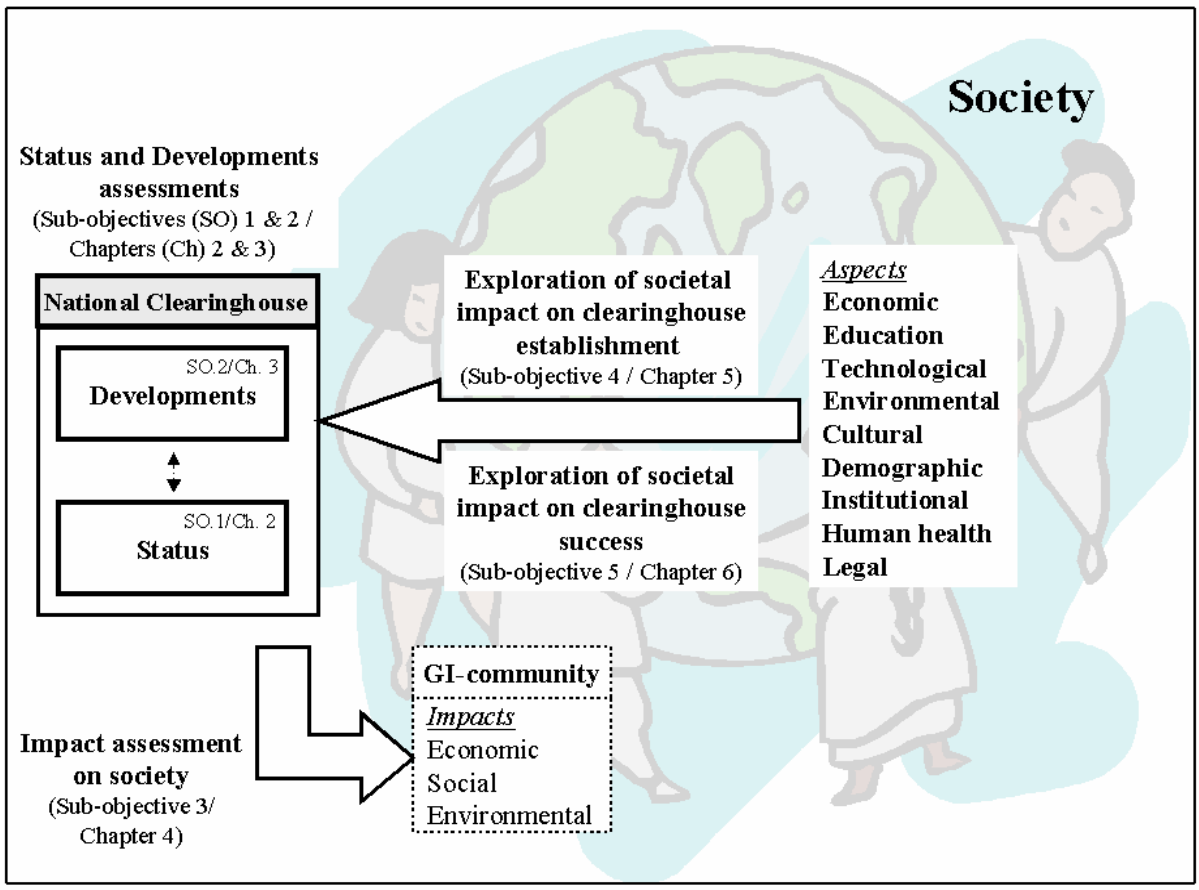


Figure 1.1 Main research steps.

Figure 1.1 represents the relationship between sub-objectives (as well as chapters). The status assessments (sub-objective 1) describe, analyse and evaluate the worldwide status, the spatial distribution of all national clearinghouses, and the similarities and differences between them. Monitoring the status systematically forms the basis to assess the developments (sub-objective 2). This longitudinal study includes the description, analysis and evaluation of the worldwide developments, in order to identify the critical (clearinghouse-internal) factors. The status and development assessments provide the required information for the impact assessment of clearinghouses on society and, in particular the GI-community, e.g. clearinghouse inventory and questions for survey undertaken (sub-objective 3). The explorations of societal impact on national clearinghouses (sub-objective 4 and 5), strongly related to the status description of a specific moment, identify the critical factors from society that could impact on the establishment or success of clearinghouse implementation.

The national clearinghouse coordinators, who require strategic support in establishing and maintaining national spatial data clearinghouses, are regarded as the main target group of this thesis. In addition, this thesis can assist policy makers in the decision-making process regarding whether investments in establishing and maintaining (national) clearinghouses are justified. It is hoped that this thesis will encourage a greater public awareness of the need to pursue education, training and research in the development and maintenance of clearinghouses, and to ensure that clearinghouses will continue to evolve and be relevant to users of spatial data and services.

Several choices have been made in order to keep the research in this thesis within manageable proportions, the most important of which are listed below.

- The thesis focuses on national clearinghouses, which can be considered to be one of the main features of national SDIs. These national SDIs have a high impact on the other levels of the SDI hierarchy (global, regional (international), state and local) through their components (Rajabifard, 2002; Williamson et al., 2003; Bernard et al., 2005b). For example, in terms of policy, national SDIs have an important effect on the management of the higher and lower levels. In terms of core datasets, a national SDI has an important role in establishing a data framework for a country. In terms of technical standards, a national SDI has a direct influence on the state and local SDIs, and its position is important for higher SDI levels to decide on their strategies and standards. Therefore, a national SDI has a crucial role in building the other levels of SDI. Other reasons for focussing on national clearinghouses are due to the fact that they are measurable, identifiable, and sustainable in time.
- This research does not include case studies, in order to analyse more deeply business requirements that could have shaped the purpose, scope, design, implementation and technical aspects of national clearinghouses.
- The emphasis of this thesis is slightly more towards the use, management and content of national clearinghouses, than on the technology. Technology has clearly been an important driving factor in influencing current clearinghouses. Ongoing changes and improvements in technology will ensure that clearinghouses continue to evolve for many years to come. When looking to the existing literature on spatial data clearinghouses, the vast majority focuses on technological developments and impacts. Examples are, the PhD-theses (Morales, 2004; Nogueras Iso, 2004), conference papers (proceedings of Global Spatial Data Infrastructure Association, EC GI&GIS Workshop, Association of Geographic Information Laboratories in Europe) and books (Groot and McLaughlin, 2000; Nebert, 2004; Bernard et al., 2005a). However, to fully analyse the development and impact of these clearinghouses, a clear understanding about the use, management and content of these facilities, is also required.
- Finally, this thesis focuses on the impact assessment of clearinghouses on the GI-community. This was chosen rather than society since society as a whole has too high a degree of uncertainty. Uncertain elements are the assessments of the proportion of the impacts, the identification of the affected stakeholders, and the numerous developments.

1.3 OUTLINE OF THE THESIS

The core of this thesis (chapters 2–6) is based on a series of five papers that have been published in, or submitted to internationally reviewed journals. The content is the work of the author of this thesis Joep Crompvoets. In a few cases, the analyses have been assisted by the co-authors given the extensive nature of the study.

The chapters themselves cover the worldwide development, impact assessment of national spatial data clearinghouses, and explorations of societal impact on national clearinghouses. Each chapter focuses on a sub-objective as earlier discussed. Figure 1.1 outlines the relationship between the chapters.

Chapter 2 assesses systematically and presents the worldwide status of national spatial data clearinghouses in December 2001. The examination of the clearinghouse status and the spatial distribution is based on a web survey, which comprises an inventory of all established clearinghouses and measurements of several key characteristics. The main differences in clearinghouse implementation are presented and briefly discussed using the clearinghouse characteristics.

Chapter 3 assesses and presents the development of all national clearinghouses throughout the world, with reference to the concepts, definitions and history of both SDI and clearinghouses. The development assessment is mainly based on a longitudinal web survey, undertaken in April 2000, 2001, 2002 and December 2000, 2001, 2002. The main results are presented using the main SDI-components. Additionally, these developments are discussed and critical (clearinghouse-internal) factors are identified.

Chapter 4 assesses and presents the impact of clearinghouses on society, and in particular the GI-community, with reference to the economic, social and environmental dimensions of sustainable development. The comprehensive and systematic impact assessment is based on a survey, undertaken among coordinators of (almost) all known clearinghouses of the world, using indicators to assess the relevance, efficiency and effectiveness of clearinghouses (November 2003 – April 2004). Complementary analyses are performed as a means of understanding the significance of these impacts. Additionally, the main impact results are discussed.

Chapter 5 explores and presents the societal impact on the establishment of national clearinghouses with reference to the economic, educational, technological, environmental, cultural, demographic, institutional, health care and legal characteristics of a country in 2002. This societal impact assessment is based on ANOVA and data mining techniques. The main result is the identification of critical (societal) factors that could impact on the establishment of national clearinghouses. Additionally, the significance of these critical factors for establishment strategies is discussed.

Chapter 6 explores and presents the societal impact on the success of national clearinghouses for the situation of 2002. This societal impact study is primarily based on a clearinghouse suitability index and statistics (e.g. partial least squares regression). The main result is the identification of critical (societal) factors for success. In addition, the significance of these critical factors for implementation strategies is discussed.

Finally, Chapter 7 concludes the thesis with the main conclusions and recommendations for further research.

Kindly note that the figures presented between chapters 2 and 3 are slightly inconsistent. The main reason is that a couple of clearinghouses, not yet discovered in December 2001, were included at a later stage of the study. Moreover, Timor-Leste became an independent nation in 2002.

Chapter 2

World Status of National Spatial Data Clearinghouses

J. Crompvoets and A. Bregt

URISA Journal (2003) 15, APA I: 43-50
Special Issue on Public Participation GIS

2.1 INTRODUCTION

At present (2001), 120 of the 192 countries in the world are working on their National Spatial Data Infrastructure (NSDI) in order to create an efficient environment for the access of spatial data. One of the main features of a NSDI is a national spatial data clearinghouse (McLaughlin, 1991; Clinton, 1994; FGDC, 1997; Onsrud, 1998). A spatial data clearinghouse can be defined as an electronic facility for searching, viewing, transferring, ordering, advertising and/or disseminating spatial data from numerous sources via the Internet. Such a clearinghouse usually consists of a number of servers that contain information (metadata) about available digital data. A national clearinghouse aims to become a kind of shopping mall for all national available spatial data as acquired by governmental agencies and/or industrial bodies.

The first national clearinghouse was established in 1994 in the United States (USA). Since then there has been much development within the field of national clearinghouses. Few studies exist with information about the worldwide status of national clearinghouses. Onsrud (1998) and Lance and Hyman (2001) presented a list of the existing national clearinghouses; however, no detailed description was included. To the best of our knowledge, no systematic periodical research has taken place with regard to the status of national clearinghouses. In order to fill this gap, we started a survey on the World Wide Web in 2000, which was repeated twice a year. The web survey's main objectives are to assess the worldwide progress, the spatial distribution, and the similarities and differences between national clearinghouses. Additionally, this web survey can be considered to be a starting-point to gather information necessary for the analysis of the legal, economic, cultural, technological and institutional impacts on the development of clearinghouses. Moreover, since clearinghouses are a key feature of spatial data infrastructures, the evaluation of the findings of this web survey might improve the planning and investing of spatial data infrastructures in a more strategic way. This article presents only the results of data collected in the month of December 2001.

2.2 WEB SURVEY METHODOLOGY

In order to assess national clearinghouse developments around the world, a web survey was undertaken. The main intention of this survey was to examine the condition (status) of national clearinghouses. Added to this survey is the word 'web' to emphasize that the needed information was mainly collected on the web. Because of the easy access to the clearinghouse sites, the web is an excellent means by which to gather the needed information quickly and objectively on a regular temporal basis. In this case, a traditional research survey, which would collect information by asking a set of pre-formulated questions in a predetermined sequence in a structured questionnaire to individuals (Hutton, 1990), would not be a suitable approach to collect the needed information quickly, objectively and easily.

The web survey began in April 2000 and was conducted in a systematic and periodical manner. The procedure consisted of the following two steps:

1) Making an inventory of all existing national clearinghouses on the Internet.
2) Measuring several characteristics to describe each clearinghouse.

The inventory (step 1) was compiled by extensive browsing of the Internet, reading related literature, and contacting experts and several webmasters. Clearinghouses were characterized (step 2) based on the following criteria: ease of measurement, objective character, and clear presentation of history, content, use and management of the clearinghouse. The following twelve characteristics were measured: 1) the year of first implementation; 2) the number of data suppliers; 3) the level of (meta)data accessibility; 4) the metadata-standard applied; 5) the number of datasets; 6) the most recently produced dataset; 7) the number of web references (AltaVista and Google); 8) the number of monthly visitors; 9) the frequency of web updates; 10) the languages used; 11) the use of maps for searching; and 12) registration-only access. Almost all of the above information was sourced from clearinghouse web pages. Additionally, in cases of uncertainty or missing data, the webmaster was contacted. The history of the clearinghouse is described by characteristic 1. The content is described by characteristics 2, 3, 4, 5 and 6. The use is described by characteristics 2, 7, 8, 10, 11 and 12. Finally, the management is described by characteristics 4, 6, 9 and 12. In the next section, each characteristic is explained and discussed in greater depth. Additionally, where available, information about funding and clearinghouse strategy has been incorporated.

2.3 RESULTS AND DISCUSSION

The main results of this 'December 2001' web survey for each of the 12 mentioned characteristics are presented and discussed. Some of the results are presented by region. This division in regions is derived from Dorling Kindersley´s World Atlas (1997).

1. The year of first implementation.
This characteristic describes the history of the clearinghouse. Beginning in 1994, the number of national clearinghouses has been steadily increasing with time (Figure 2.1). Currently, 59 countries have an implemented version on the web (the year of first implementation is known for 52 of the clearinghouses). Additionally, it is known that nine countries have projects for implementation. However, the variety in number between the different regions is considerable (Table 2.1). For example, in Europe, North America, and South America, more than 50% of the countries have established a national clearinghouse, whereas in Africa this number is less than 5%. It is important to note that 124 countries have not conducted any initiative to build such a national facility.

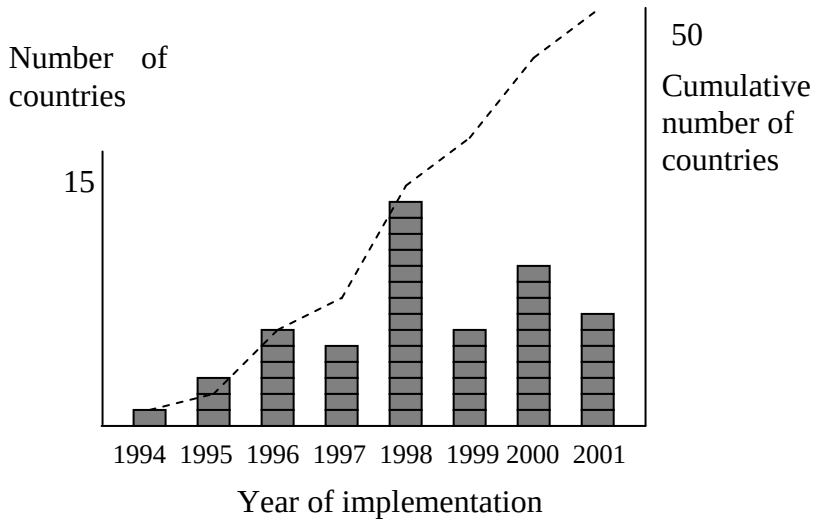


Figure 2.1 The first year of national clearinghouse implementation: per year (columns) and cumulative (dashed line).

Table 2.1 The first year of national clearinghouse implementation (distributed per region).

First Year of implementation	Total No. of countries	No. of African countries	No. of Asian countries	No. of Austra-lian countries	No. of European countries	No. of North-American countries	No. of South-American countries
1994	1					1	
1995	3				3		
1996	6	1	3	1	1		
1997	5		2		2		1
1998	14	1	3	1	6	2	1
1999	6				3	1	2
2000	10		1		3	4	2
2001	7				3	1	3
Date unknown	7				3	3	1
1994 – 2001	**59**	**2**	**9**	**2**	**24**	**12**	**10**
Building phase	9	4	2	0	1	2	0
No implementation	124	45	38	12	18	9	2

In Figure 2.2, the global distribution of implementation status of national clearinghouses is presented. Here, we can see that the main ‘hotspots’ of implementation are the American continent, Europe (except Eastern Europe), Southeast Asia and Australia. On the other hand the main ‘holes’ are the African continent and the Middle East.

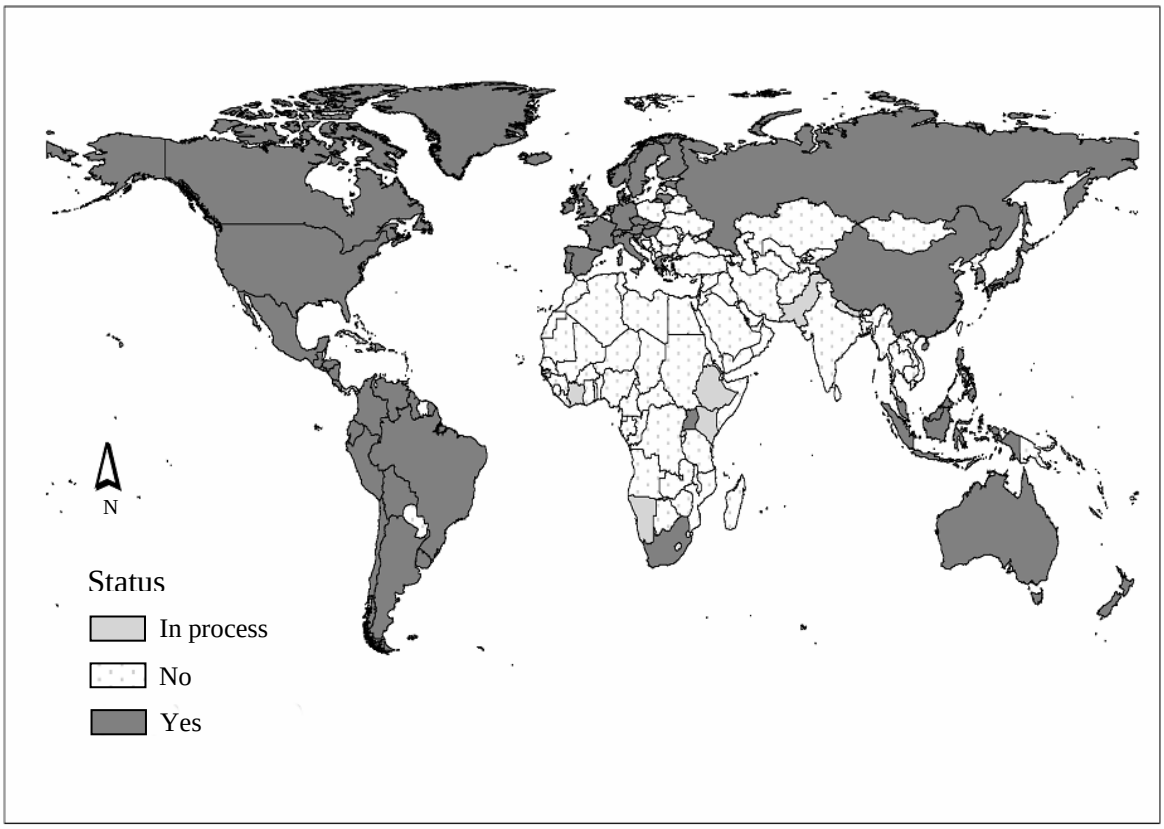


Figure 2.2 Global distribution of status of national clearinghouses (December 2001).

2. The number of data suppliers.
This characteristic describes the number and diversity of data suppliers. The power of a clearinghouse is that several data suppliers can disseminate their products via this facility. The average number of data suppliers participating in a clearinghouse is high; however, there is great variety between the clearinghouses (Table 2.2). For Austria, the Czech Republic, Slovenia and USA, the number of data suppliers exceeds 100. In Canada, there are 1758 data suppliers. This contrasts with the 35 clearinghouses with fewer than ten suppliers (notably in South America and Asia with their powerful national mapping agencies).

Table 2.2 The number of data suppliers per region. The number in parentheses is the number of clearinghouses analysed per region.

Region	Average	Standard Deviation	Median	Maximum number
Africa (2)	11	11	10	19
Asia (9)	7	6	5	16
Australasia & Oceania (2)	14	17	14	26
Europe (24)	33	49	12	133
North America (12)	204	551	6	1758
South America (10)	4	3	2	8
WORLD (59)	**54**	**239**	**6**	**1758**

3. The level of (meta)data accessibility.
This characteristic describes the presentation of the content. Not all-existing clearinghouses give access to data or metadata. For example, in some cases the clearinghouse presents only a simple (not standardised) description of the datasets. For this reason, three classes of accessibility are distinguished: 1) abstract (simple/short description about the databases without using any formal meta-data description); 2) metadata; and 3) data (+ metadata). In most clearinghouses, the user has access to metadata (Table 2.3). However, in fifteen countries the user has directly access to the data.

Table 2.3 The level of (meta)data accessibility of national clearinghouse per region. The number in parentheses is the number of clearinghouses analysed per region.

Region	Abstract	Metadata	Data (+ metadata)
Africa (2)	1	1	0
Asia (9)	1	3	5
Australasia & Oceania (2)	0	1	1
Europe (24)	8	12	4
North America (12)	1	7	4
South America (10)	2	7	1
WORLD (59)	**13**	**31**	**15**

4. The metadata-standard applied.
This characteristic describes the metadata-standard used. With the diverse sources from which spatial databases are built, it is extremely important to maintain information about the content, quality, source and lineage of the data. A number of organizations have developed (or are in the process of developing) standards for storing and maintaining metadata. The most mature of these metadata-standards are developed by the Federal Geographic Data Committee (FGDC, 1995) and the European Committee for Standardisation (CEN/287, 1996). These metadata-standards form the backbone of national clearinghouses. The FGDC metadata-standard is the most applied and distributed standard around the world (Table 2.4). The CEN standard is only applied in Europe. Recently, the International Organization of Standardisation has created the ISO19115 standard (ISO/TC-211, 2001). Currently, ten countries have started a project to apply this last-mentioned standard for their national clearinghouse.

Table 2.4 Metadata-standards applied per region. Between brackets, the number of clearinghouses analysed per region.

Region	CEN	FGDC	National	Others
Africa (1)		1		
Asia (8)		2	2	4
Australasia & Oceania (2)			1	1
Europe (16)	9		3	4
North America (11)		10	1	
South America (8)		7	1	
WORLD (46)	**9**	**20**	**8**	**9**

5. The number of spatial datasets.
A means to quantify the content of a clearinghouse is the number of datasets. However, it does not represent the importance of the accessible datasets to the economic and social development of the country. The variety in the number of datasets is significant (Table 2.5). For example, US federal clearinghouse gives access to almost 100,000 datasets (December 6, 2001), while the average of the 24 European clearinghouses is 440. The difference in the total number of accessible datasets between US and Europe is easily noticed (100,000 vs. 10,000). In total, the clearinghouses describe 170,000 spatial datasets together. Ten clearinghouses have more than 1000 datasets described (Australia, Austria, Canada, the Czech Republic, Japan, Mexico, South Africa, Switzerland, Uruguay and USA).

Table 2.5 Number of datasets described per region. The number in parentheses is the number of clearinghouses analysed per region.

Region	Average number	Standard deviation	Total number of data sets	Median	Mini-mum	Maxi-mum
Africa (2)	1561	2198	3122	1561	6	3115
Asia (6)	676	857	4056	260	9	1782
Australasia & Oceania (2)	15,031	21,249	30,062	15,031	5	30,056
Europe (22)	440	867	9768	118	8	3011
North America (10)	11,802	31,089	118,020	211	8	99,649
South America (8)	721	1646	5768	38	7	4735
WORLD (50)	**3616**	**14,618**	**170,796**	**111**	**5**	**99,649**

6. The most recently produced dataset.
This characteristic describes the up-to-date nature of content and the management of content in the clearinghouse. It is the difference in months between the date of web survey and the date of the most recently produced dataset described in the national clearinghouse. On average, the time of the production of the most recent dataset is more than two years (Table 2.6). However, 22 national clearinghouses describe spatial datasets, produced within one year from the date of web survey. However, for twelve national clearinghouses, this duration is longer than three years (mainly countries located in South America or Asia).

Table 2.6 The duration in months between date of the web survey (December 2001) and the date of the most recently produced dataset. The number in parentheses is the number of clearinghouses analysed per region.

Region	Average duration	Standard deviation	Median	Mini-mum	Maxi-mum
Africa (2)	31	41	31	2	60
Asia (6)	47	19	55	23	66
Australasia & Oceania (2)	2	1	2	1	2
Europe (22)	21	31	9	1	126
North America (10)	15	12	17	1	29
South America (8)	73	106	43	1	257
WORLD (50)	**28**	**44**	**15**	**1**	**257**

7. The number of web references.
This number can be interpreted as a means to measure the popularity (use) of the clearinghouse site within the Internet network. The 'Free Link Popularity Service' http://www.linkpopularity.com (The PC Edge, Inc.) is used, which measures the number of links to the homepage of the national clearinghouse that can be checked by the following search engines: AltaVista, and Google. A well-linked popularity can dramatically increase traffic to the specific web site. The link popularity of national clearinghouse is high, which means that they are an excellent source of consistent and targeted web traffic. However, the variety is significant (Table 2.7). The number of web references does not differ that much between the regions and so the popularity of a national clearinghouse can be considered as universal. The following national clearinghouses have high link popularity: Australia, Canada, Colombia, Finland, New Zealand, Norway, USA and Venezuela.

Table 2.7 The number of web references by AltaVista and Google search engines per region (STD = Standard deviation). The number in parentheses is the number of clearinghouses analysed per region.

	Number of web references AltaVista				Number of web references Google			
Region	Ave-rage	STD	Me-dian	Maxi-mum	Ave-rage	STD	Me-dian	Maxi-mum
Africa (2)	50	56	50	89	41	43	41	71
Asia (9)	151	148	105	477	59	112	52	175
Australasia & Oceania (2)	3084	2851	3084	5100	1315	913	1315	1960
Europe (24)	320	792	42	3642	123	129	67	502
North America (12)	96	168	40	480	146	351	55	1080
South America (10)	112	140	48	428	76	78	53	213
WORLD (59)	**312**	**857**	**50**	**5100**	**145**	**309**	**54**	**1960**

8. The monthly number of visitors.
This characteristic describes the use of national clearinghouses for accessing spatial datasets. This amount relates to the monthly number of visitors who have visited the homepage of the clearinghouse. The average number of visits of this page exceeds the 5000 visitors. It is worth noting that the variety between the implementations is significant due to some particularly popular clearinghouses (Table 2.8). The following national clearinghouses are visited the most: Canada, Finland, Portugal, Slovenia and USA; Portugal's clearinghouse has approximately 60,000 visits per month.

Table 2.8 The monthly number of visitors per region. The number in parentheses is the number of clearinghouses analysed per region.

Region	Average	Standard deviation	Median	Minimum	Maximum
Africa (1)	423		423		
Asia (5)	1055	382	1120	618	1576
Australasia & Oceania (1)	4378		4378		
Europe (10)	10,521	18,571	1743	410	60,000
North America (6)	5384	7492	1973	328	18,700
South America (3)	1684	944	1517	835	2700
WORLD (26)	**5871**	**12,337**	**1334**	**328**	**60,000**

(note: the number of clearinghouses analysed is much lower as presented in the other tables due to the fact that not all clearinghouses are able to count the number of visitors).

9. The frequency of web updates.
This characteristic describes the management in the clearinghouse. It is the duration (days) between the day of the last web update and the date of measurement. High number of days refers to a low frequency of web update. One possible indication of a well-managed clearinghouse can be seen by the high frequency of updated information. The average number of days of last update is high for the whole population of clearinghouses due to instances of poor management (with some updates exceeding 100 days) in Europe and Asia (Table 2.9). The variety between clearinghouses is significant as, alongside the poorer managed clearinghouse, numerous excellently managed facilities operate (update within one day).

Table 2.9 The frequency of web updates per region (days). The number in parentheses is the number of clearinghouses analysed per region.

Region	Average duration	Standard deviation	Median	Clearing-houses updated within 1 day	Clearinghou-ses updated more than 100 days ago
Africa (2)	902	1270	902	0	1
Asia (9)	410	723	7	4	3
Australasia & Oceania (2)	12	13	12	0	0
Europe (22)	170	312	27	6	7
North America (9)	3	3	2	4	0
South America (10)	37	42	26	3	1
WORLD (54)	**179**	**440**	**15**	**17**	**12**

10. The Languages used.
This characteristic refers to the language used at the ‘search page’ of the national clearinghouse. It describes the number and diversity of users able to access data because of their familiarity and knowledge of the given language. 30 clearinghouses do not have a search mechanism written in English (in addition five of these are written in Arabic, Chinese, Greek, Japanese or Korean script). 29 clearinghouses use only their home language. These language problems reduce the accessibility to data (for English-speaking people).

11. The use of maps for searching.
This characteristic refers to the use of maps when searching. Maps can be used for locating an area of interest or by clicking on an area with predefined boundaries. These maps can improve the accessibility to data. In eighteen clearinghouses, maps can be used as an option to search for (meta)data. This relatively advanced alternative for searching is popular in Europe and Asia.

12. The registration-only access.
This characteristic describes the management and possible limitations of use. Sometimes, before accessing the data, users have to register themselves by entering personal details. This characteristic could have a negative impact on accessibility. For nine national clearinghouses, the user is required to register him/herself to access metadata or data (Canada, El Salvador, Finland, Hungary, Malaysia, Singapore, Spain, Canada and Uruguay).

2.4 CONCLUSIONS

Since 1994, the number of national clearinghouses has steadily increased to a total of 59. Looking at the trend of implementation, it is expected that more countries will establish clearinghouses in the future. In fact, building clearinghouses is a global activity (with the exception of Africa and the Middle East (as well as Australasia & Oceania)). Most clearinghouses are established in Europe, Southeast Asia, North and South America. The main initiatives for establishment come from Anglo-Saxon countries, such as USA, South Africa and Australia. The USA in particular, which is supported by the FGDC, has stimulated many countries (of the American continent) to build a clearinghouse. However, 124 countries have still not shown any initiative to build one. There are several reasons for this. For example, a country may not have appropriate network architecture or there may be institutional bottlenecks for implementation. The differences in content, use, and management between the clearinghouses are broad. An example of such broad difference in content is the total number of accessible datasets described in a clearinghouse. In the US clearinghouse this number is ten times as high as the total number of all 24 European clearinghouses. The reason for such difference is due to each country's unique historical, institutional, economic, legal, technological and cultural setting. Especially in Europe, there are great contrasts in suppliers, web references, visitors, and frequency of web updates, probably as a result of the high institutional, economic, legal, technological and cultural diversity within this region.

The most applied metadata-standard is the FGDC. However, looking to the numerous projects to apply the ISO-standard, it is likely that ISO19115 will be the most applied standard in the future. This international consensus standard reflects FGDC, CEN and other inputs. It provides detail that goes beyond FGDC and CEN metadata, including special coverage of raster and imagery information. Currently, there are several initiatives to create workable subsets and extensions of ISO 19115 so that conversion of FGDC-support tools and implementations to meet ISO conformance requirements is facilitated (FGDC Metadata Staff Coordinator, 2001).

Looking to the average number of data suppliers, web references, and visitors, we can conclude that national clearinghouses are a popular facility to distribute and access spatial data.

Finally, in the future, it is highly probable that many national clearinghouses will give access to spatial data itself and provide complementary services such as online mapping. However, a concern could be the low frequency of web updates of several clearinghouses due to poor management. Therefore, special attention has to be given to keep clearinghouse managers motivated for having a well-managed clearinghouse. Based on the twelve characteristics used, we can conclude that Australia, Canada, Portugal and USA have the best existing national clearinghouses. Additionally, this web survey shows that not only the richest countries have good clearinghouses. Examples of relatively poorer countries with suitable national clearinghouse are El Salvador, Nicaragua and Uruguay. Based on the above research, for all countries, it seems that one of the keys for successful clearinghouse implementation is high political support and interest by means of funding and long-term strategy.

ACKNOWLEDGEMENTS

The authors would like to thank Dr. Gert Jan Hofstede and Dr. Monica Wachowicz of Wageningen UR for their critical comments and suggestions on earlier versions of this article. Additionally, the authors appreciate very much the numerous responses of the webmasters to all our questions.

Chapter 3

Assessing the worldwide developments of national spatial data clearinghouses

J. Crompvoets, A. Bregt, A. Rajabifard and I. Williamson

International Journal of Geographical Information Science (2004), 18(7): 665-689

3.1 INTRODUCTION

Spatial Data Infrastructure (SDI) is about the facilitation and coordination of the exchange and sharing of spatial data between stakeholders in the spatial data community. With this objective in mind, many countries are developing SDI to manage and use their spatial data assets more efficiently. These countries are finding it necessary to develop SDIs to assist in decision-making that has an important impact within their national boundaries. Over the past few years, many countries have spent considerable resources on debating optimal National SDI (NSDI). One of the main features of a NSDI is the national spatial data clearinghouse (McLaughlin, 1991; Clinton, 1994; FGDC, 1997, 2002; Coleman and McLaughlin, 1998; Onsrud, 1998; Groot and McLaughlin, 2000; AUSLIG, 2001). The national clearinghouse is the access network of an NSDI that facilitates access to the spatial data. It provides complementary services and improves the exchange and sharing of spatial data between suppliers and users.

Based on an overall assessment, the average cost of a spatial data clearinghouse (including some services) is around €1.5 million a year (INSPIRE Architecture and Standards working group, 2002). This money is spent in management and coordination costs, GIS and Internet application development, training, hardware, network server, standardisation activities, legal environment creation, and metadata preparation. In December 2002, 67 national clearinghouses have been implemented and thirteen countries have projects for implementation. Based on these 80 initiatives, it means that globally around €120 million worldwide is spent yearly for clearinghouse management. Up to now, this amount of money has never been audited or evaluated (certainly not globally).

Since 1994, a fast development in national clearinghouses implementation has taken place throughout the world. Not many studies exist about the worldwide development of these national clearinghouses. To the best of the authors' knowledge, no systematic periodical survey has taken place with regard to the development of national clearinghouses (Crompvoets, 2002). The purpose of the present paper is to fill this gap, its main objectives being to assess worldwide developments, to analyse and describe these developments, to understand the reasoning behind them and to determine the critical factors for success.

This paper assesses and presents the developments of all national clearinghouses throughout the world, with reference to the concepts, definitions and history of SDI and clearinghouses. The assessment methodology has been described based on a longitudinal survey and the main development results are presented using the main SDI-components: policy, access network, standards, people and data. Additionally, these development results are discussed and some key factors for success are determined.

3.2 SPATIAL DATA INFRASTRUCTURES AND CLEARINGHOUSES

3.2.1 Spatial Data Infrastructures

Viewing the core components of SDI as policy, access networks, (technical) standards, people (including partnerships) and data, Rajabifard et al. (2002) suggested that different categories could be formed based on the different nature of their interactions within the SDI framework. Considering the important and fundamental role between people and data as one category, a second can be considered consisting of the main technological components: the

access networks, policy and standards. The best example of access network is the clearinghouse. The nature of both categories is very dynamic due to the changes occurring in communities (people) and their needs, as well as their ongoing requirement for different sets of data. Additionally, with the rapidity with which technology develops, the need for the mediation of rights, restrictions and responsibilities between people and data are also constantly subject to change (Figure 3.1). This suggests an integrated SDI cannot be composed of spatial data, value-added services and end-users alone, but instead involves other important issues regarding interoperability, policies and networks. This in turn reflects the dynamic nature of the whole SDI concept. It is an issue also highlighted by Groot and McLaughlin (2000).

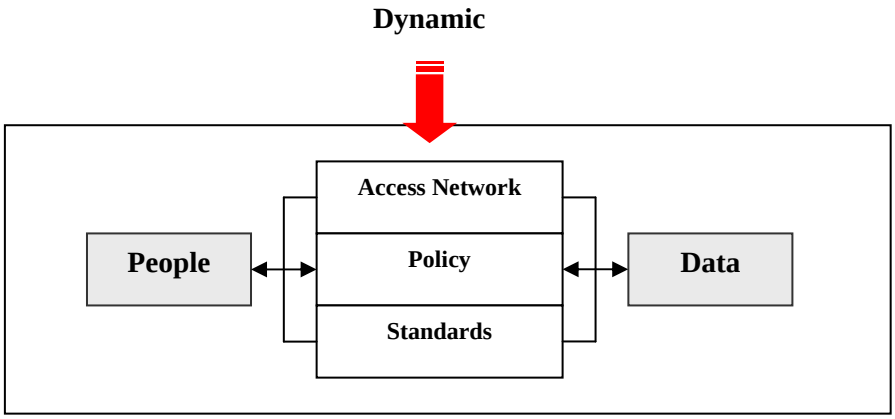


Figure 3.1 Nature of and relations between SDI components (adopted from Rajabifard et al., 2002).

The concept of SDI can be defined as an integrated, multileveled hierarchy of interconnected SDIs based on partnerships at corporate, local, state/provincial, national, regional (international) and global levels. This enables users to save resources, time and effort when trying to acquire new datasets by avoiding duplication of expenses associated with the generation and maintenance of data and their integration with other datasets.

With this in mind, every nation undertakes to some extent the development of strategic national mapping and spatial data activities to meet their national planning and management needs. The accumulation of these activities over time has resulted in the identification of key linkages between institutional and technical aspects similar in many respects to other forms of infrastructure, and occurring in a continuum of development strategies. Based on this, Rajabifard et al. (2003) distinguished and reported on two generations of SDIs, the first and the second generations.

The first generation of SDIs development has emerged since mid-1980s when USA and Australia, for example, started to develop the data access relationships, which became the precursor to the development of NSDI initiatives. At this time, countries developing SDI on any jurisdictional level had only very limited ideas and knowledge about different dimensions and issues of the SDI concept, and rather less experience of such development.

Within this generation, each country designed and developed SDI based on their specific requirements and priorities and nationally specific characteristics. The ultimate objectives of the SDI initiatives in this generation as summarised by Masser (1999) were to promote economic development, to stimulate better government and to foster environmental sustainability. A significant milestone overcome by the first generation, for whom there were few experiences and existing SDI developments from which to learn, was the documentation of researchers' and practitioners' experiences and status reports on their SDI initiatives and as part of that report on their clearinghouse activities which facilitated their SDI initiatives. This achievement not only gave countries a knowledge-base from which to learn and/or develop their initiatives, providing exposure to the developmental strengths and weaknesses of different SDI initiatives, but also provided social capital to share and foster SDI developments in other countries. Consequently, many countries involved in SDI development over the first generation took a product-based approach, which became the dominant model for SDI justification and development partially through a lack of awareness of other options.

However, the transition to the second generation can be marked by a change in focus on SDI development by several countries (Australia, USA, Canada) involved in developing the concept from the beginning. This led to a rapid increase in the number of countries becoming involved in SDI development, fostered by the definition of an SDI community where experiences could be shared and exchanged. This shows the continuum of strategic spatial data development.

The second generation started around 2000 when some of the leading countries on SDI development changed their development strategies and updated their conceptual models. In second-generation SDI, the strategy for SDI development is changing towards a more process-based approach (Rajabifard et al., 2003). This approach focuses on the creation of a suitable infrastructure to facilitate the management of information assets instead of the linkage to existing and future databases.

The second generation of SDI developments characteristically falls into two groups: those countries that started to develop an SDI initiative during the period of the first generation and are gradually modifying and upgrading the initiative, as well as those countries that have recently decided to design and develop an SDI for their respective countries and/or have just commenced doing so (Lance and Hyman, 2001; Wehn de Montalvo, 2001).

The distinguishing features of the second generation include leverage of the experiences, expertise, social capital of SDI development and the development of clearinghouse systems derived from the first generation. For the first generation, data were the key driver for SDI development and the focus of initiative development. However, for the second generation, the use of that data (and data applications) and the need of users are the driving force for SDI development. Introduction of web services is the main technological indicator of second-generation SDI because such services are partly able to fulfil the needs of users and improve the use of data. In summary, second- generation SDI development has been relatively quick due to the concept gaining momentum and because of the existence of early prototypes, clarification on many initial design issues, increased sharing and documentation of experiences to facilitate implementation and face the complexity of decision-support challenges.

3.2.2 Spatial Data Clearinghouses

A spatial data clearinghouse can be defined as an electronic facility for searching, viewing, transferring, ordering, advertising, and/or disseminating spatial data from numerous sources via the Internet and, as appropriate, providing complementary services. Such a clearinghouse usually consists of a number of servers that contain information (metadata) about available digital data.

A clearinghouse is based on a distributed network of people (spatial data suppliers, managers and users) linked electronically (Clinton, 1994; FGDC, 2002). The term 'distributed system' refers to a distributed collection of users, data, software and hardware, whose purpose is to meet some predefined objectives (Bishr and Radwan, 2000). The clearinghouse allows suppliers to make known what spatial data exist, the condition of these data and instructions for accessing these data. Each data supplier describes available data in an electronic form and provides these descriptions (or metadata) to the network using a variety of software tools. Additionally, the data supplier can offer access to his produced data. Users can discover who has what spatial data and their type and quality (Radwan, 2002).

The main reason for creating a clearinghouse is the desire of users to have a single source for accessing all the available resources. A spatial data clearinghouse is a system to provide this capability serving as a central point for sharing data among data producers and users (Phillips, 1998; Noori-Bushehri and Rajabifard, 2001; Rajabifard, 2002).

To make a clearinghouse as the access network operational within an SDI, it has to be strongly linked to the other SDI-components: people, data, policy and (technical) standards (Rajabifard and Williamson, 2001). A clearinghouse can only be a success within an SDI when there is a strong cohesion between these components. The use, management, supply of data and the content, and/or quality of these data determine its success subject to the quality of the standards, response time of the clearinghouse and legal/economical policy.

There are different understandings of the definition of clearinghouses. The Australian and New Zealand Land Information Council (2000) defines clearinghouses in a much wider context. It incorporates: discovery, transfer and access services; legal arrangements including supporting custodians' ability to control access to their data; co-ordination and management functions; and the spatial information commercial market place in which data are value-added and integrated to produce products, services and solutions. The definition of the US Federal Geographic Data Committee (FGDC, 1997) includes besides the technical mechanism also the institutional aspect of clearinghouses. However, mainly out of practical reasons, this paper focuses on the most tangible product of the clearinghouse; the electronic facilities for discovery, transfer and access.

The latest definitions of the clearinghouse play more emphasise on the inclusion of services. For example, the position paper of INSPIRE Architecture and Standards working group (2002) describes a clearinghouse (Portal) as a site featuring a suite of commonly used services, serving as a starting point and gateway to the Web for a user community. This service-oriented approach is in accord with the objectives of the second generation of SDI. Examples of services which could be included within a clearinghouse environment are: map service, coverage service, feature service, gazetteer service, coordinate transformation service, authentication service, analysis/spatial data fusion service, web pricing and ordering service. These web services will change the implementation of clearinghouse functional capability, as well as the way in which users use them in their own applications (Bernabé et al., 2002).

A national clearinghouse for spatial data is a central web-portal at national/federal level that focuses on the facilitation of spatial data discovery, access and services recognised (in legal or institutional sense) by the national government (mainly through the national council for geo-information) as the country's main clearinghouse for spatial data. It is not a national repository where datasets are stored. It aims to be a kind of shopping mall for all national-wide available spatial data as acquired by main (national) governmental agencies and/or industrial bodies. It is important to know that many differences in implementations exist. The way a national clearinghouse is set up depends on its technological, legal, institutional, cultural, commercial environment and management. This environment determines to what extent the clearinghouse management retains control over their products. Besides national clearinghouses, there are clearinghouses at the local, state and regional (international) level. The national clearinghouse is different than the others in the sense that it is embedded in the nation's institutions and legal framework. Examples of national clearinghouses are: USA, National Geospatial Data Clearinghouse; Australia, Australian Spatial Data Directory; UK, GIgateway; and The Netherlands, Nationaal Clearinghouse Geo-Informatie.

The US Federal Geographic Data Committee (FGDC) established the first national (federal) clearinghouse in 1994. The implementation of this clearinghouse among federal agencies was motivated by a desire to minimise duplication of effort in the collection of expensive digital spatial data and foster partnerships where common needs exist (Rhind, 1999; FGDC, 2000). At this moment, the FGDC focuses on the extension of the clearinghouse network that provides 'one stop' access to standardised spatial data, applications, programs and products from all federal agencies and incorporates similar non-federal information, and it establishes web mapping and online data services to meet general requirements of government and citizens users. After initial deployment and testing of the comprehensive web portal, reusable commercial replication services will be required (FGDC, 2002).

3.3 SURVEY METHODOLOGY

To assess national clearinghouse developments around the world, a longitudinal web survey was undertaken whose main intention was to examine the developments of all existing national clearinghouses. Added to this survey is the word 'web' to emphasise that the required information was mainly collected on the Web. The Internet is an excellent means to gather regularly the needed information quickly and objectively, because of the easy access of the clearinghouse sites.

The survey was conducted systematically and periodically (April 2000, December 2000, April 2001, December 2001, April 2002, December 2002). The intention was to have a survey each half year because developments in information communication technology are difficult to monitor and to keep up with. However for practical reasons it was chosen to have a survey in the months of April and December.

The methodology used in this survey consisted of the following two steps:

Step 1. Making an inventory of all existing national clearinghouses on the Internet.
Step 2. Measuring several characteristics to describe each of the clearinghouses.

The inventory (Step 1) was periodically compiled by extensive browsing of the Internet, reading related literature, and contacting experts and several webmasters. Clearinghouses were characterised (Step 2) based on the following criteria: ease of measurement, objective character, and clear presentation of history, people (suppliers, coordination, users), data, access network, policy, and standards of the national clearinghouse environment. The following twelve characteristics were periodically measured and recorded:

(1) Year of first implementation.
(2) Number of data suppliers.
(3) Monthly number of visitors.
(4) Number of web references (AltaVista and Google).
(5) Languages used.
(6) Frequency of web updates.
(7) Level of (meta)data accessibility.
(8) Number of datasets.
(9) Most recently produced dataset.
(10) Use of maps for searching.
(11) Registration-only access.
(12) Metadata-standard applied.

Almost all the above information was sourced from clearinghouse web pages. Additionally, in cases of uncertainty, language problems or missing data the webmaster/system administrator was contacted. The history of the clearinghouses is described by characteristic (1); people by (2-6); data by (7–9); the access network by (10); policy by (11); and standards by (12). In following sections, some of the characteristics identified above will be explained and their developments will be discussed in greater depth.

Additionally, for analytical purposes, to improve the understanding of the development processes, extra information (especially about access network) was collected during January/February 2003 as below:

(a) Status of national clearinghouse.
(b) Name of national clearinghouse.
(c) (De)centralised network architecture.
(d) Mechanisms for searching.
(e) Availability of view (web mapping) services.
(f) Type of coordination body.
(g) Percentage of periodically changed web addresses.
(h) Funding stability.

The Access Network is additionally described by characteristics (a-e); people by (f,g); and policy by (h). When referring to these characteristics below, the number or letter related to the characteristic is placed in parentheses.

3.4 RESULTS

The presentation of the results is grouped using SDI component classification, which has been mentioned above (access network, people, data, policies and standards). First, the history of national clearinghouses is presented, followed by the results related to each of the SDI-components. To present some of the developments in figures, both the average and median are used (Figures 3.5, 3.6, 3.8 and 3.11). Due to the highly skewed distribution of some of the clearinghouse characteristics, the median is less sensitive to extremes than the average. In these cases, the median is more informative than the average. Some of the results are presented by region (this division in regions is derived from Dorling Kindersley (1997)).

3.4.1 History

The history of national clearinghouses is characterised by the year of first implementation (1) on the web. When implementing a clearinghouse, there are several activities that interact: design and implementation of discovery, transfer and access services, metadata preparation/validation/publication, legal/institutional environment creation and standardisation definitions. The most tangible product derived from all these activities is the placement of a national clearinghouse on the Web. .

From 1994, the number of national clearinghouses has been steadily increasing (Figure 3.2). Now 67 countries have an implemented version on the web. Additionally, it is known that thirteen countries have projects for implementation (with all its related activities). Based on these developments, it can be expected that increasingly more clearinghouses will be established. Implementing national clearinghouses is a global activity; however, the variety in number between the different regions is considerable (Table 3.1). For example, in Europe and America, more than 50% of the countries have established a national clearinghouse, whereas in Africa and Asia-Pacific, this is less than 20%. The many project initiatives in Africa are promising.

Note that 113 countries did not conduct any initiative to build such a national facility. This may be due to several reasons, e.g. a country may not have appropriate network architecture or have institutional/legal bottlenecks preventing implementation (Crompvoets and Bregt, 2003).

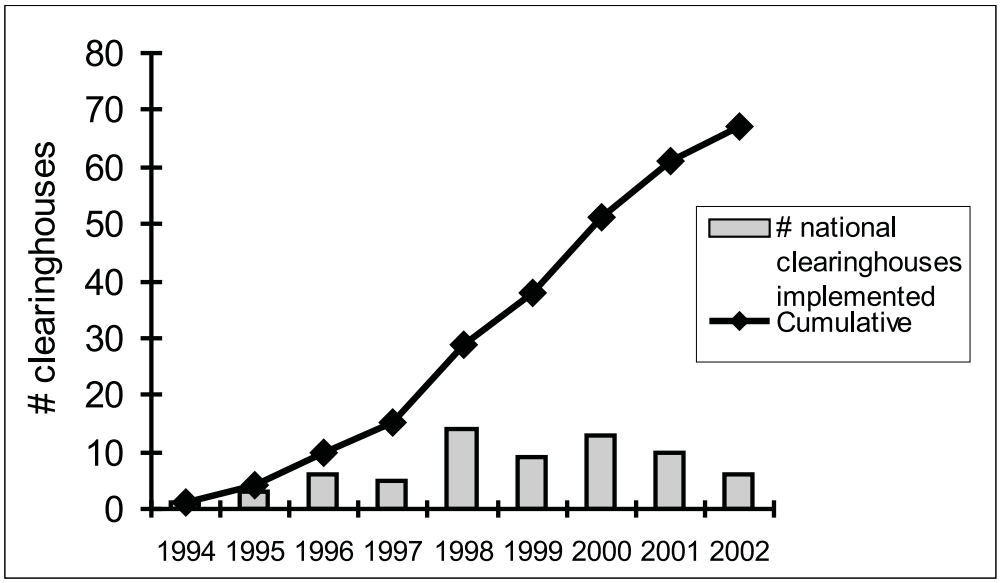


Figure 3.2 First Year of national clearinghouse implementation: per year (columns) and cumulative (dashed line).

Table 3.1 First year of national clearinghouse implementation (distributed per region).

First year of implementation	Total No. of countries	No. of African countries	No. of Asian-Pacific countries	No. of European countries	No. of American countries
1994	1				1
1995	3			3	
1996	6	1	4	1	
1997	5		2	2	1
1998	14	1	4	6	3
1999	9			5	4
2000	13		2	3	8
2001	10	1		4	5
2002	6	2	2	1	1
1994 – 2002	**67**	**5**	**14**	**25**	**23**
2003/2004?	13	5	3	3	2
No implementation	113	41	47	15	10

3.4.2 Access network

The access network component of an SDI is critical from a technical perspective to facilitate the use of data by people. This component seeks to facilitate access to relevant data sources and spatial information services by anyone, anywhere. The best example of an access network at national level is the national clearinghouse. The clearinghouse as access network is characterised by the status of national clearinghouse (a), by the name of a national

clearinghouse (b), by (de)centralised network architecture (c), by mechanisms for searching (d), by use of maps for searching (10), and by the availability of view (web mapping) services (e). The differences in implementation status, network architectures, search mechanisms, and view service provisions between the clearinghouses are broad.

According to the status of national clearinghouse implementation (a), three classes are distinguished: project, product-portal, and clearinghouse. An implementation is considered a project when no clearinghouse is established on the Web where a project plan to establish one is already available. It is still in the stage of designing, preparing metadata, creating a legal/institutional environment or defining standards. For example, countries as Latvia, Poland, Botswana, Madagascar, Nigeria, Israel and India have set up projects for implementation. A national clearinghouse has the status of a product-portal when the spatial data are sourced from only one supplier. According to the clearinghouse definition, a clearinghouse should have spatial data from numerous sources. However, in several countries there is only one dominant national spatial data supplier (national (military) geographic institute). These countries are mainly located in South America and Europe. Most of these suppliers present their (meta)data as products. That is the reason why these types of national clearinghouses are called product-portals. This difference in the number of sources is the only difference between product-portals and clearinghouses. Examples of product-portals are: FÖMI-Products (Hungary, Institute of Geodesy, Cartography and Remote Sensing), Clearinghouse-SNIT (Chile, Military Geographic Institute), National Geospatial Data Clearinghouse (Peru, National Geographic Institute), and Geospatial catalogue (Venezuela, Geographic Institute of Venezuela Simon Bolivar). A national clearinghouse is classified as a clearinghouse when it completely fulfils all the criteria of the clearinghouse definition. The majority of the existing national clearinghouses have the status of Clearinghouse (72%). Examples of classified clearinghouses are: MIDAS, Meta-Information DAtabase System (the Czech Republic), Slovenian National Data Catalogue, Spatial Data Discovery Facility (South Africa), GeoConnections Discovery Portal (Canada), Geospatial Data Clearinghouse of Dominican Republic, Geographic Data Clearinghouse of El Salvador, and National Clearinghouse of Geographic Data (Uruguay).

Within the spatial community, various names have been assigned to this national facility (b); examples are catalogue services (OpenGIS Consortium (OGC)), Spatial Data Directory (Australia) and Clearinghouse (USA). Although they have different names, the goals of discovering, and accessing spatial data through the metadata properties they report are the same. As shown in Figure 3.3, the most popular names are Clearinghouse and (Meta)data information system. Clearinghouse is especially used within the American continent.

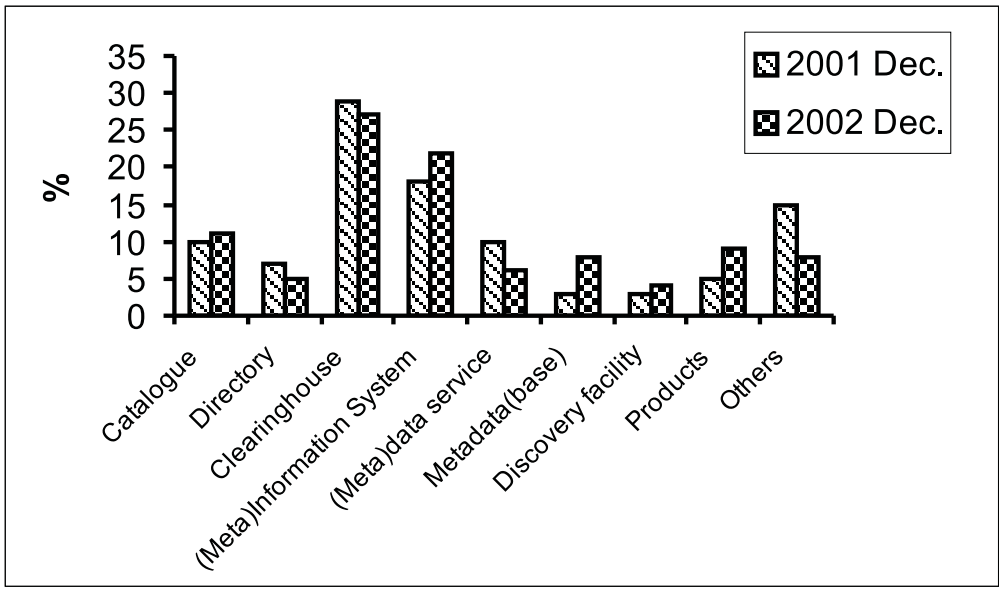


Figure 3.3 Distribution (%) of national clearinghouses based on naming.

A national clearinghouse could consist of a number of servers on the Internet that contain information about available data. A clearinghouse is an example of a client-server network architecture (c). The server machines hold the metadata and services. The clients request metadata or services by visiting the server(s), usually through a web browser. A clearinghouse with a decentralised network architecture (c) means that the (meta)databases and services are distributed over numerous servers installed at different suppliers interconnected through a network with each server running autonomously. In total, nine decentralised national clearinghouses exist, mainly located in countries of Anglo-Saxon origin (USA, UK, South Africa, Australia). However, the majority has implemented a centralised version wherein all the metadata and services are stored on servers installed at the main supplier or coordination body.

When searching on a clearinghouse, the user has the ability to choose different mechanisms for spatial data searching (d), like predefined search terms (hypertext links), location (spatial search), maps with predefined boundaries (index maps), free keyword search, place name (gazetteer) and production time. The most popular mechanisms are predefined terms and free keyword (Figure 3.4). Searching by means of predefined terms is mainly done by name database (text index), theme, organization, and geographical name. The use of digital maps when searching (10) is very popular. Maps can be used for locating an area of interest or by clicking on an area with predefined boundaries. These maps improve the data discovery. Nevertheless, the trend is that few national clearinghouses are using this mechanism. Fifteen national clearinghouses have made a link to the FGDC's search page by using the search-and-retrieve protocol known as ANSI Z39.50 or ISO 10163. Recently, OGC Catalog Service version 1.0 came on the market, which is an improved version to query, search and present search results to the user. It was recommended in *The SDI Cookbook* (Nebert, 2001) that clearinghouses should use this new OGC Catalog Services Specification.

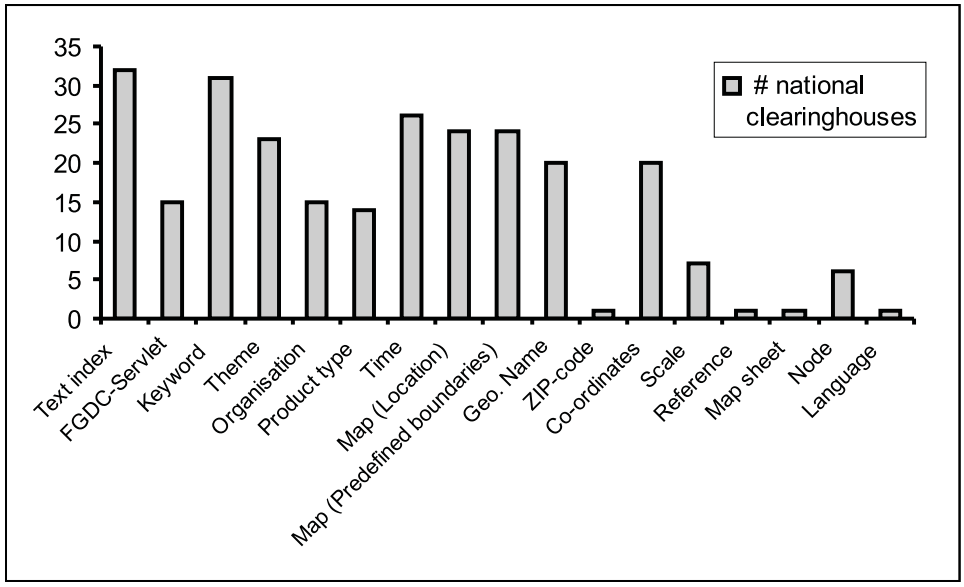


Figure 3.4 Distribution (%) of national clearinghouses based on mechanism for searching.

A relatively new facility within the domain of national clearinghouses is the standardised view services for documented data (INSPIRE Architecture and Standards working group, 2002). This viewing of geographic information over the Internet is also referred to as web mapping (e). This service is a typical tool for second-generation SDIs. It includes the presentation of general-purpose maps to display locations and geographic backdrops, as well as sophisticated interactive and customisable mapping tools. Viewing is supported by simple queries. Related to the viewing is the requirement that the geographic information presented must be given a context, made easy understandable and of value to end-users. Understanding thematic geographic information coming from different sources requires standardisation of the way in which the information is portrayed. The technology behind web mapping is aimed at portraying spatial information quickly and easily for most users, requiring only basic map reading skills. At this moment, fourteen existing clearinghouses have already implemented services to view for documented data. They are mainly located in Europe, North America and Asia. In Asia-Pacific especially this view service is very popular.

Other examples of services that are already implemented in some of national clearinghouses are download and e-commerce services. It is expected that these web services-based developments will dominate the field for the coming years.

3.4.3 People

Through increased use and awareness of spatial information, a dramatic growth has occurred in the user base. With the proliferation of online web mapping, and navigation/direction information, an increasing number of people are using GIS, this is obviously of importance in the development of an SDI to facilitate spatial data activities.

The people involved in a clearinghouse environment can be categorised into three groups: data suppliers, managers (service administrators) and end-users. The data suppliers represent the spatial data providers and developers. The behaviour of people is characterised by the number of data suppliers (2), by the number of monthly visitors (3), by the number of web references (4), by the languages used (5), by the type of coordination body (f), by the frequency of web updates (6), and by the number of periodically changed web addresses (g).

Characteristic number of data suppliers (2) describes the variety and quantity of data providers and the spatial data developers of the national clearinghouse. The power of a clearinghouse is that several data suppliers can disseminate their products via this facility. The number of data suppliers was measured four times (December 2000, 2001, 2002, April 2002). As shown in Figure 3.5, the average number of data suppliers participating in a clearinghouse is high (in particular when compared with the median), however the variety between the clearinghouses is extremely high. For example, more than 1500 suppliers use the clearinghouses of Canada or the USA to disseminate their data. This is in contrast with the more than 45 clearinghouses, which have less than ten suppliers. The median is very low because of the high number of product portals. The recent trend is that the average and median are decreasing. Based on this development, it seems that a national clearinghouse is losing its popularity to participate as a data supplier and to supply spatial data. From a spatial distribution context around the world, the very low number of data suppliers in South America is remarkable. The reason may be the dominant role of the leading mapping agencies.

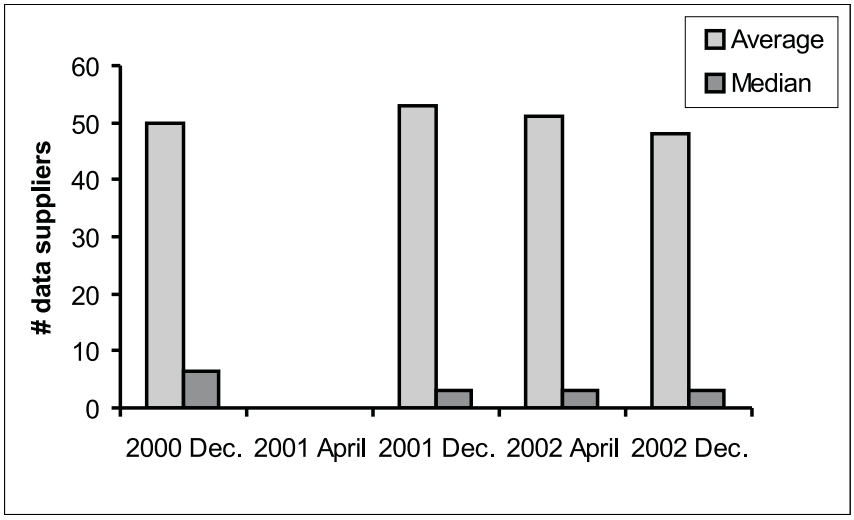


Figure 3.5 Number of data suppliers during time (average and median). Due to practical reasons, this characteristic was not measured in April 2000 and April 2001.

Characteristic monthly number of visitors (3) indicates quantitatively the use of national clearinghouses to access spatial datasets by end-users. This number relates to the monthly number of visitors who have visited the homepage of the clearinghouse. It does not really present the usability of the clearinghouses for searching to spatial data, because it does not identify the behaviour of the users within the clearinghouse and the number of different visitors. This characteristic is measured six times. The numbers relate to the month just before the measurement. This is the reason that March and November are presented (Figure 3.6). For November 2002, the average number of visits exceeds 5000 visitors (however, the median is just above 1000). Note that the difference in visiting numbers between the clearinghouses is high. The most visited are the ones of USA, Portugal, Finland, Canada and Slovenia. In general, the highest numbers of visitors are found in Europe and North America. The recent development is that the numbers are stabilising and not increasing. This could be an indication that the clearinghouse is losing its popularity to access spatial data.

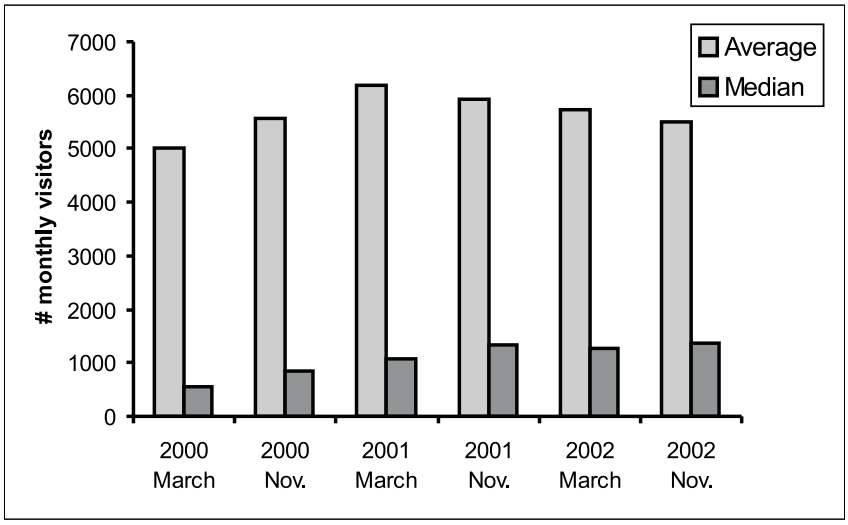


Figure 3.6 Number of monthly visitors during time (average and median).

Based on the gathered data, it seems that the implementation of web mapping facilities within the national clearinghouse environment increases the number of visitors. The average and median of the clearinghouses with web mapping are much higher than those that have not such a web mapping facility (Table 3.2). Therefore, web-mapping facilities could increase use. However, an important limitation of web mapping is the costs of implementation: €150,000 summed with €50,000–100,000 for each participating organization (INSPIRE Architecture and Standards working group, 2002).

Table 3.2 Impact of web mapping facilities on the number of monthly visitors (data in parentheses are the number of measured national clearinghouses).

Web mapping facilities	Average No. of visitors	Median No. of visitors
Yes (10)	13,324	5000
No (22)	1750	1280

Another aspect that could increase the use of a national clearinghouse is the number of suppliers. This is based on the correlation coefficient *(r)* between the number of suppliers and the monthly visitors of December 2002: 0.73.

The characteristic number of web references (4) can be interpreted as a means to measure the popularity (use) of the national clearinghouse site within the Internet network. Used is the 'LinkPopularity.com, Free Link Popularity Service ©' (http://www.linkpopularity.com) (The PC Edge, Inc.), which measures the number of links to the homepage of the national clearinghouse that can be checked by the following search engines: AltaVista, and Google. A well-linked popular web site can dramatically increase traffic to the specific web site. The link popularity of national clearinghouse is high, which means that they are an excellent source of consistent and targeted web traffic. This characteristic is measured six times. The differences in numbers between the national clearinghouses are huge. The national clearinghouses with the highest number of web references are: Finland, South Africa, Australia, Canada, USA and Colombia. Compared with

other regions, the national clearinghouses of Africa have lower numbers than the other regions. Since April 2002, Google has generally a higher number of web references than AltaVista. The correlation coefficient *(r)* of 0.71 between the number of web references (Google) and the number of monthly visitors, and $r = 0.75$ between the number of web references (Google) and the number of data suppliers show that the number of web references is an appropriate indicator of the use of national clearinghouses (Franzblau, 1958). The development is that both numbers of web references on AltaVista and Google are decreasing. The number of web references searched by AltaVista is already decreasing since April 2001 and the one searched by Google since April 2002. This is another indication that the national clearinghouses are losing their popularity.

Characteristic languages (5) refers to the language used at the 'search' page of the national clearinghouse. It indicates the ease for searching, the number and diversity of end-users able to access data, because of their familiarity and knowledge of given language. The percentage of clearinghouses in which data can be searched using the English language is slightly increasing. In December 2002, 60% of the clearinghouses have search facilities written in English. The percentages of 'Spanish' clearinghouses are stabilising (around the 25%), and the ones of the multilingual clearinghouses are increasing. This means that people capable of understanding English have access to more spatial data. However, the metadata are mainly written in one language, mostly in one of a country's official languages.

The organization that provides the national clearinghouses as a service to society is characterised by the type of coordination body (f). It is the intermediary between the data users and the suppliers for the clearinghouse. It facilitates the integrity of access to the required data by ensuring system technical services as well as the administrative, data security, and financial services necessary to broker between data suppliers and data users within the information policies governing the SDI. The main coordination of the national clearinghouses is mainly in the hands of the national mapping agencies (Figure 3.7). In addition, the environmental agencies are of relatively high importance. This coordination characteristic has already been measured twice (December 2001, 2002). Additionally, a prediction for 2003 is made based on the existing information of projects for implementation. The relative trends (%) are generally very stable. The percentages related to national mapping and environmental agencies are slightly increasing, meanwhile the percentages related to cadastre, geological services, national councils for geo-information and commercial companies are slightly decreasing. The contribution of environmental agencies is particularly remarkable. It seems that environmental policy triggers focus on the need for SDI around the world.

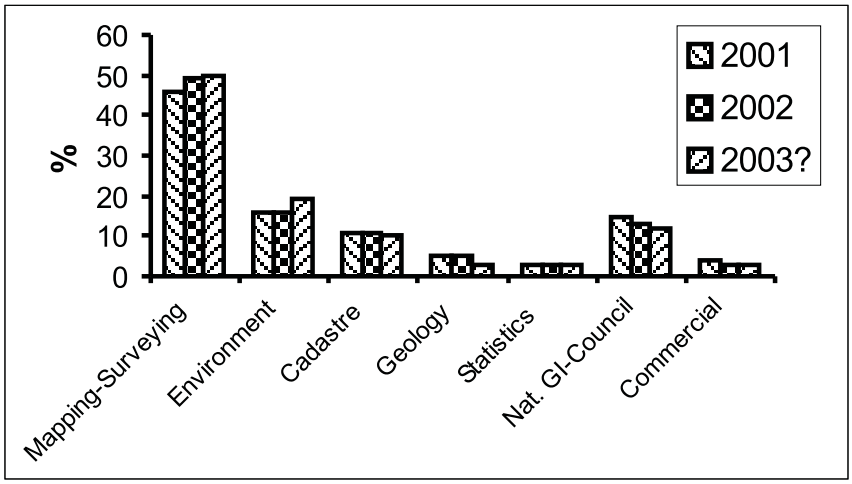


Figure 3.7 Distribution (%) of national clearinghouses based on type of coordination body.

Characteristic frequency of web updates (6) refers to the management in the national clearinghouse. It is the duration (days) between the day of the last web update and the date of measurement. High numbers of days refers to a low frequency of web update. One indication of a well-managed clearinghouse could be seen by the high frequency of updated information. It does not directly refer to the updates of the data itself, but to the updates of the clearinghouse web site. This characteristic is measured six times (Figure 3.8). For consistency, it was measured each time on Fridays in the first week of the months of April and December. The average number of days from the last update is high for the whole population of clearinghouses due to instances of poor management (with some updates exceeding one year). However, the median value is low. The variation between clearinghouses is high as, alongside the poorer managed clearinghouse, numerous excellently managed facilities operate (update within one day). The update frequency is generally high in North America and low in Africa. The development is an increase of the duration of the days of web updates, and so a lower frequency of these updates. The median value shows an increase in trend. This could relate to a decrease of interest by clearinghouse managers.

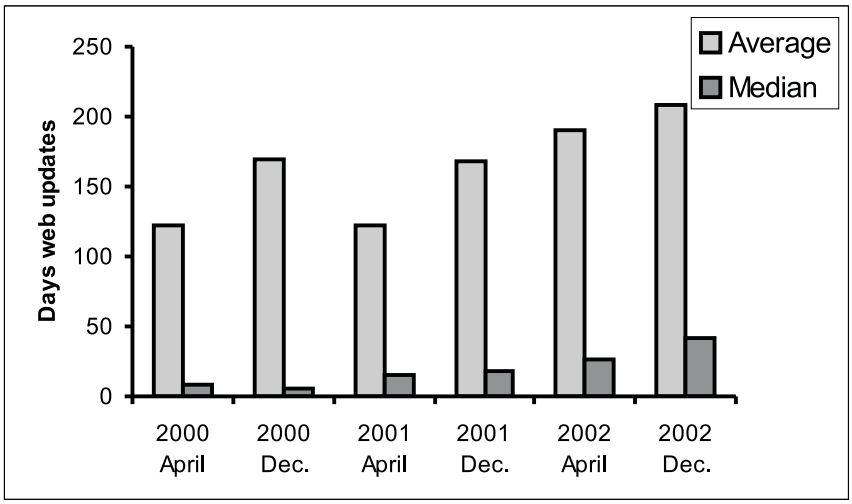


Figure 3.8 Frequency of web updates during time (average and median).

Another characteristic to express management of national clearinghouses is by percentage of periodically changed web addresses (g). This refers to the change of URL address of the homepage of the same clearinghouse site during the period between the two times of measurements. This change is measured for five periods. During period April–December 2002, 17% of all existing national clearinghouse in April 2002 changed their address in another one (Figure 3.9). Most of the changes were in Europe. The development is an increase of the percentage of periodically changed web address. This dynamism could be an indication that the clearinghouse managers are struggling to create the right (technological) environment.

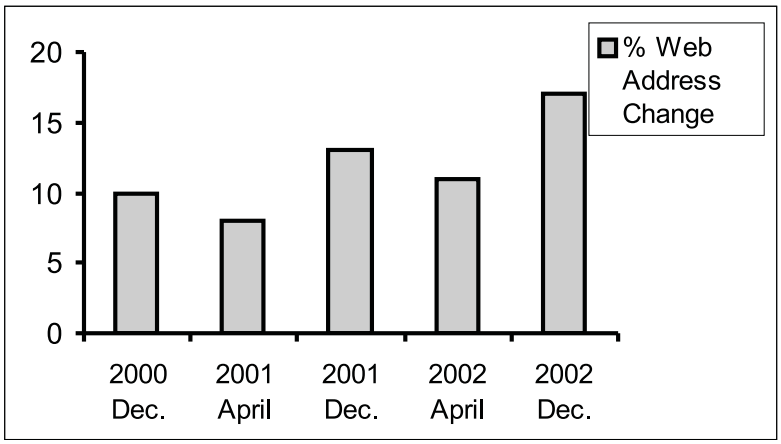


Figure 3.9 Percentages of periodically changed web addresses during time.

3.4.4 Data

Interoperability is a key consideration of both the standards and data component of an SDI. Data within an SDI should be compatible in terms of format, reference system, projection, resolution and quality. Data are described by the following characteristics: the level of (meta)data accessibility (7), the number of spatial (meta)datasets (8) and the most recently produced dataset (9). The difference in quantitative and qualitative content of accessible data between different clearinghouses is very high.

Characteristic level of (meta)data accessibility (7) describes the presentation of the data content within national clearinghouses. In most cases clearinghouses provide access to standardised metadata. However, in a few cases these metadata are not standardised. In other cases, the user has the ability to access the data directly by means of its metadata. Delivery of data over the Internet can be realised in various ways. Some clearinghouses transmit data via e-mail; others serve the data through what OGC calls a Web Feature Server. Four classes are distinguished: prototype, non-standardised metadata, standardised metadata and data (+ standardised metadata). Prototype refers to the level when final metadata security checks have to be arranged. Data (+ standardised metadata) refers to direct access to data through the clearinghouse without contacting the producer. In spatial context, Asia-Pacific has many clearinghouses that allow the user an opportunity to download spatial data directly. This is in contrast with the European and American region. In a relative sense, the percentages of standardised metadata access over time are very stable, and the percentages of data (+ standardised metadata) access are only slightly increasing (Figure 3.10). Once a clearinghouse is established, the level of accessibility is unlikely to be changed.

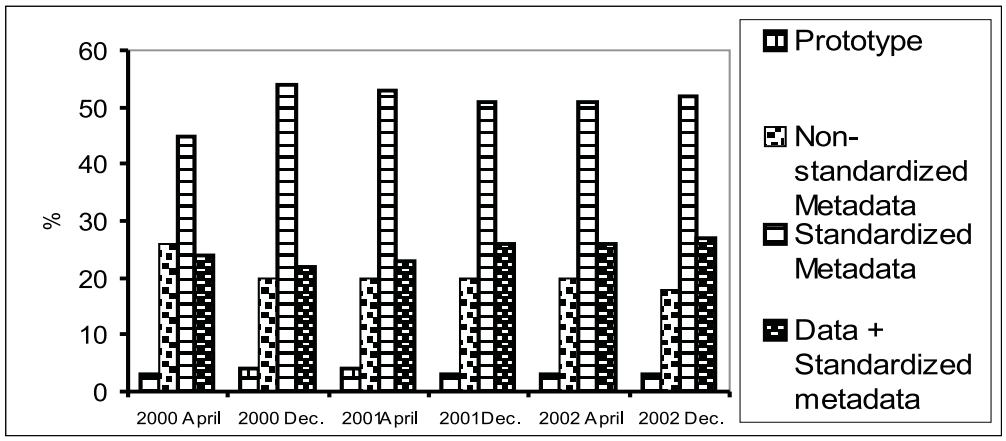


Figure 3.10 Percentages of the level of (meta)data accessibility during time.

A special form of non-standardised metadata is a graphical presentation (e.g. jpeg or bitmap format) of a subset of described dataset. This alternative is partly based on the concept that 'a picture tells more than a thousand words'. It is particular popular in Europe. The recent development is a continuous increase of this type of metadata presentation.

A possible means to quantify the content of a national clearinghouse is characterised by the number of datasets (8). This number refers to the metadata records, where each describes one dataset. This characteristic does not represent the importance of the accessible datasets to the economic and social development of the country. This characteristic is measured five times. The variety in the number of datasets is enormous. For example, the US Federal clearinghouse allows the user access to more than ±139,000 datasets, while the 25 European clearinghouses together give access to only ±10,000 datasets. This is the reason that the average and median are so different (Figure 3.11). It is remarkable that the average and median are decreasing over time. This trend could relate to the decreasing trend of the number of data suppliers. This observation is based on the very strong $r = 0.85$ between the number of suppliers and the number of datasets (December 2002). In the past, the total summed number of datasets of all the clearinghouses was steadily increasing. However, recently this trend has stopped and is stabilised around the 210,000. The trend of a decreasing number of monthly visitors could also partly relate to the decreasing number of datasets presented in the national clearinghouse. An r between the number of spatial datasets and the number of monthly visitors is 0.64.

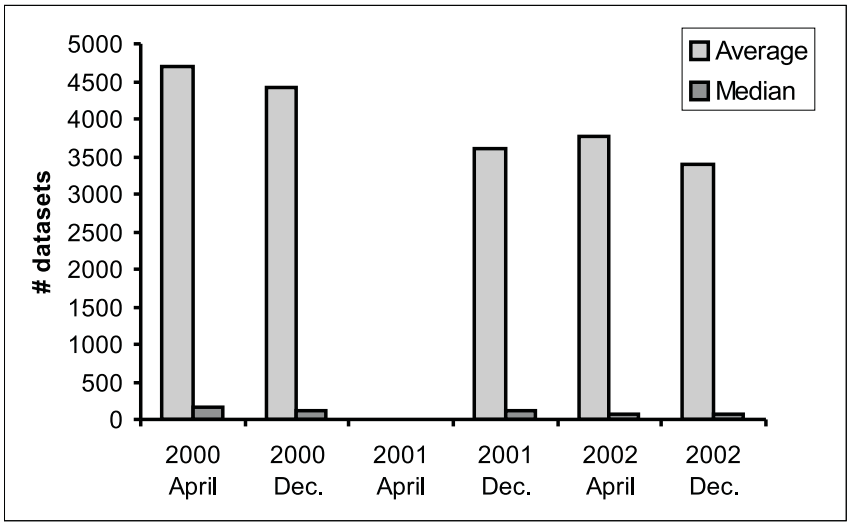


Figure 3.11 Number of datasets during time (average and median). Due to practical reasons this characteristic was not measured in April 2001.

The characteristic most recently produced dataset (9) describes the up-to-date nature of content and partly the management of content in the national clearinghouse. It is the duration (months) between the date of the most recently produced dataset described in the national clearinghouse and the date of measurement. It is measured six times. In general, the most recently produced datasets are in Europe and the least in South America and Asia-Pacific. The average time of most recently produced dataset is 33 months and the median is 20 months (December 2002). The recent development is a slight increase of the average and median duration. Therefore, the data of the national clearinghouses are becoming less up to date. This trend could have a negative impact on usage (*r* between the most recently produced dataset and the number of visitors is -0.45).

3.4.5 Policy

The policy and administrative component of the SDI definition is critical for the construction, maintenance, access and application of standards and datasets for SDI implementation. In general, policies and guidelines are required for SDI that incorporate: spatial data access and pricing; funding; spatial data transfer; custodianship; metadata; and standards. SDI component policy is only described by funding stability (h) and registration-only access (11).

Funds are mainly used to design and to establish discovery, transfer and access services, metadata preparation/validation/ publication, and legal/institutional environment creation. The most easily identified funding of clearinghouses comes from governmental agencies. Nevertheless, some signs suggest that other parties, mainly other public agencies, make enormously important contributions. Identification can be difficult because some are indirect, e.g. some are provided in kind through the time of staff already on payrolls, which are not properly attributed, or in returns to government through the tax system (Rhind, 2000). The identifiable contribution to the national clearinghouse from the private sector is modest. An example that reflects partly this statement is that the government or other public agencies directly control 94% of the national clearinghouses. Funding of the national clearinghouses is mainly piecemeal (72%). Additionally, even a few clearinghouses exist that were never

funded (6%)! A small number were continuously funded (22%). It is interesting that 25% of existing national clearinghouses were initially funded by foreign agencies such as the US Agency for International Development (USAID).

It seems that the stability of funding has a positive impact on people's use and management, and the quantity of datasets in the national clearinghouses (Table 3.3). Stable funding refers to the continuously funded clearinghouses; not stable to the piecemeal and never funded clearinghouses.

Table3.3 Impact of funding stability on the number of monthly visitors, the frequency of web updates, the number of suppliers and datasets (average and median).

Funding stability	Visitors average	Visitors median	Updates average (days)	Updates median (days)	Suppliers average	Suppliers Median	Datasets average	Datasets Median
Stable	6487	5093	107	28	176	16.5	12,563	207
Not stable	5076	1323	239	68.5	15	2	467	42

It was expected that personal registration in a clearinghouse could have a negative impact on end-users due to privacy. For this reason registration-only access (11) was introduced as a characteristic to be measured. Sometimes, before accessing the data, end-users have to register themselves by entering personal details. In 15% of all clearinghouses, the user should register him/herself by filling in a 'registration template' (December 2002). The percentage of clearinghouses with registration obligations fluctuates around 15 % over time. Based on acquired information, it seems that this registration has no impact on the number of users and so it is no limitation for the users.

3.4.6 Standards

To ensure interoperability amongst the datasets and access mechanisms defined by an SDI, standards are essential. Standards can be applied at many different levels within an SDI. In terms of data, Australia's former national mapping organization, the Australian Land Information Group (AUSLIG, 2001) identified that standards are required 'in reference systems, data models, data dictionaries, data quality, data transfer, and metadata'. This component is represented by the application of metadata-standard (12) in the national clearinghouse. With the diverse sources from which spatial databases are built, it is extremely important to maintain information about the content, quality, source, and lineage of the data. A number of standard organizations have developed, or are in the process of developing, standards for storing and maintaining metadata. The Federal Geographic Data Committee and the European Committee for Standardisation (CEN/Technical Committee 287) have developed the main metadata standards. These metadata-standards form the backbone of national clearinghouses. The FGDC metadata-standard (Content Standard for Digital Geospatial Metadata, Version 2.0, 1998 (FGDC, 1998)) is the most applied and distributed one around the world (26 clearinghouses, December 2002). Application of this standard is stabilising between the 40 and 50% of all the national clearinghouses (Figure 3.12). In Europe CEN/287 Env 12657 (CEN/TC287, 1996) is widely accepted. Recently, the International Organization of Standardisation TC/211 has created the ISO19115 standard (ISO/TC-211, 2001). Many countries have set up projects, that deal with the harmonisation of their metadata standard with the ISO standard 19115 for Geographic Information – Metadata in order to adopt the ISO standard as national standard (27 countries, December 2002). Many countries are creating their own profile of ISO 19115. That is the reason why the percentages of

national metadata standards are increasing over time. The scope of the metadata standard ISO 19115 specifies that it defines metadata for services as well as data. ISO/TC 211 Project Team 19, dealing with ISO 19119 – Services, is developing more detailed service metadata. This ISO 19119 Services standard is at the Draft International Standard stage (Østensen, 2001).

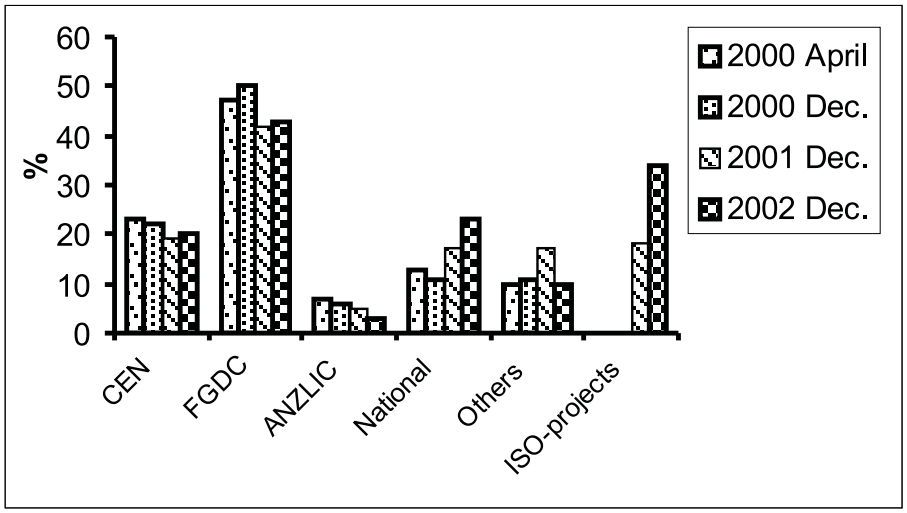


Figure 3.12 Distribution (%) of national clearinghouses based on the application of metadata-standards.

3.5 DISCUSSION OF RESULTS

The main objectives of the present paper are to analyse the worldwide developments, to describe these developments, to understand the reasoning behind these developments, and to determine the critical factors for success. Each objective will be discussed in turn.

One of the survey's objectives was to analyse the worldwide developments. Through this web survey, it was possible to examine the developments of all national clearinghouses on a worldwide scale. Using this approach, it was possible to gather the needed data quickly and objectively and on a regular-temporal basis. A traditional survey research described as a method for collecting information by asking a set of pre-formulated questions in a predetermined sequence in a structured questionnaire to individuals would not have been a very suitable approach to collect quickly, objectively and easily the required information. Of great importance was the back up of the numerous webmasters/system administrators in case the necessary information could not been collected directly from the Web. The response of these webmasters to the e-mailed questions was overwhelming. More than 80% of the emails sent were answered. A frustrating aspect of the followed approach was the high change of web addresses for the same national clearinghouse site.

A complementary study to this web survey could be to analyse some case studies more deeply to determine the business requirements that have shaped the purpose, scope, design and implementation of the process and technical aspects of the national clearinghouses. An appreciation of these business requirements could improve the explanation of the trends in usage and in investment in the clearinghouse.

Another objective of the survey was to describe the developments. The number of national clearinghouses implemented is steadily increasing over time. In the near future, it is very likely that increasingly more countries will establish their national clearinghouse. From an implementation point of view, the introduction of national clearinghouse as a concept for the dissemination and access of spatial data can be considered as a success. This implementation of national clearinghouses is becoming almost a global activity. Not only is it rich countries that have built effective clearinghouses. Examples of relatively poor countries with suitable national clearinghouses are Chile, Colombia and Uruguay. However, currently, a decrease of interest in managing national clearinghouses is observed, which is a concerning development. This observation is based on the decreasing number of data suppliers, frequency of web updates, and the stabilising number of datasets. These trends are almost all related to poor management of data providers and (especially) clearinghouse managers. Another concerning trend is the decline in use. This development refers to the decreasing number of monthly visitors, suppliers and web references. Additionally, another negative trend is the lowering quantity and quality of data content. This development is based on the decrease of the number of datasets and the less up-to-date nature of the produced dataset. Of concern is also the way national clearinghouses are funded. Only a small number are funded continuously.

Although the differences in management, use, content and technology between the clearinghouses are high, the developments are very similar. After an initial peak, figures gradually enter a decline in the management, use and data content (only a few show a continuous increase in use and content).

In the future, many national clearinghouses are likely to give more access to spatial data and provide complementary services such as online mapping. These web services will change the functional capability of clearinghouses, as well as the use of them to applications. At this moment, many clearinghouses are in kind of ‘identity crises’. During the period of first-generation SDI, data were (the only) key driver for clearinghouse development. During the transition to second-generation SDIs, the use of data (data application) and the need of users is becoming the main driving force for national clearinghouses. These trends are reflected by the variety of search mechanism alternatives, the introduction of web services, and the increase of multilingual clearinghouses.

The dynamic nature of SDI and clearinghouses is well reflected by the application of metadata standards to improve the interoperability amongst the datasets and access mechanisms. Many countries have set up projects to harmonise their metadata standards with the new ISO standard 19115.

Another objective of the survey was to understand the reasoning behind these developments. The reason that more national clearinghouses are implemented is that more countries realise the need and potential power of having one. Most of the existing clearinghouses are established in Europe, South East Asia, and North and South America. Nevertheless, numerous countries have still not shown any initiative to build one. There are several reasons for this. For example, a country may not have appropriate network architecture or might have institutional, legal, cultural or economic bottlenecks.

The declining trends in use, management and content reflect the changing demands on clearinghouses. It seems that after a stage of orientation to clearinghouse implementations, SDI people are becoming dissatisfied with the existing clearinghouse functional capabilities. This dissatisfaction could be the consequence of the fact that the existing clearinghouses do not fit anymore the current expectations of the second- generation SDI people. This difference

in demands relates to the dynamic nature of SDI, which has been highlighted by Rajabifard et al. (2002). Therefore, the implementations and concepts behind clearinghouses to share data between suppliers and data users are likely to be changed and resolved.

Another possible reason of these declining trends could be that the majority of clearinghouses are only funded in a piecemeal way. Management of clearinghouses is directly affected by funding. A consequence of bad management could be worse data content, which again could have a negative impact on use. These trends could also have a negative impact on the data suppliers who are no longer willing to disseminate their data through the clearinghouse. This again could have a negative impact on use. These observations can be justified by the high correlation coefficients between the number of suppliers, the number of datasets, and the monthly number of visitors.

As mentioned above, the focus on the need of the users is becoming a crucial aspect for implementation. At this moment, users have the right to expect more of these types of facilities. A point of concern is that clearinghouses are not always user-friendly. It seems that the concepts of metadata and clearinghouses are too complicated; the terminology used is too discipline specific and too focused on the data alone. Therefore, it seems that clearinghouses do not fulfil the current demands of the users.

Several external developments are also impacting on the development of national clearinghouses and they will continue to influence the evolution of their application of this clearinghouse: expanding technologies, market-demand, changing business models, sustainable development, e-government, and participatory democracy (Williamson et al., 2003). The introduction of web services within the clearinghouse environment is partly the technological answer to some of these developments.

The last objective was to determine the critical factors for success. Because of each country's unique historical, institutional, economic, legal, technological and cultural setting, no single best solution or recipe exists. However, the following are some indications of critical success factors:

- Be specific about the purpose of the clearinghouse under consideration. To be successful, there has to be a direct need to share data and services. When the context for implementation is missing, people will become frustrated to implement clearinghouses only out of fashion.
- Provide good communication channels for the community for sharing and using datasets instead of aiming only toward the linkage of available databases.
- Create stable funding. Stability of funding is needed to build a suitable framework that facilitates the management of information assets.
- Create trust in the clearinghouse. Stability of funding could support this process.
- Create more user-friendly interfaces with less discipline-specific terminology.
- Introduce web services to clearinghouse. It is very likely that the software for web services will become cheaper in the future.
- Motivate data suppliers and web service providers to participate within the clearinghouse. The more data and web service providers, the more data and services are available. This improvement of the content will attract end users.
- Motivate the clearinghouse managers to update their environment regularly.

3.6 CONCLUSIONS

Through the methodology followed, it was possible to assess the developments of all national clearinghouses on a worldwide scale. The strength of the web survey was the ease, speed and objectivity to measure the required data regularly.

The number of national clearinghouses is steadily increasing over time. It is expected that more countries will have implemented clearinghouses in the future. From this perspective, the implementation of national clearinghouses can be considered a success.

This contrasts with the decline in the use, management, and content of national clearinghouses. One of the main reasons for these concerning trends could be the dissatisfaction of the spatial data community with the functional capability of SDIs and clearinghouses. It seems that with the transition to the second-generation SDI, the demand on the efficiency of clearinghouses as well as how they are used is changing. Clearinghouses during the first-generation SDI were only data-oriented, while in the second-generation SDI they are becoming more user and application oriented. Another reason for these declining trends could be that the majority of clearinghouses are only funded in a piecemeal way, which means that no suitable framework can be built to facilitate the management of information assets.

The main success factors that have a positive impact on the development of national clearinghouses are the inclusion of web services within clearinghouses, stable funding, the clarity of purpose of the clearinghouse, the provision of good communication channels, the creation of user-friendly interfaces with clear terminology, and trust in the management environment. Addressing these factors will go a long way towards meeting the immediate needs of the current users. The full implementation is a major challenge for the future of the clearinghouse.

Another challenge is to keep spatial data suppliers, clearinghouse managers, and end-users motivated for and informed about this mechanism**.** This is really needed when one considers their role in all the worldwide activities related to the implementation of national SDIs wherein national clearinghouses form the key feature.

Based on the demands of the people to the second-generation SDI, the usability of spatial data and web services are the main research topics that have to be investigated to improve the functional capability of national clearinghouses.

ACKNOWLEDGEMENTS

The authors thank Dr. Wies Vullings of Wageningen UR, Ir. Bastiaan van Loenen of Delft University of Technology and Brian Goodsell for their comments and suggestions on earlier versions of this article. The authors also acknowledge the support of all the webmasters of the national clearinghouses in the preparation of this paper. Their contributions were crucial for the results.

Chapter 4

Worldwide impact assessment of spatial data clearinghouses

J. Crompvoets, F. de Bree, P.A.J. van Oort, A. Bregt, M. Wachowicz, A. Rajabifard and I. Williamson

Submitted to Computers, Environment and Urban Systems

4.1 INTRODUCTION

Many international regions, countries, states and counties throughout the world have spent considerable resources over the past few years implementing and managing Spatial Data Clearinghouses (SDCs). These SDCs can be considered to be a prominent feature of Spatial Data Infrastructures (SDI) (Clinton, 1994; FGDC, 1997; Onsrud, 1998; Crompvoets et al., 2004), because they are the facilities for making spatial data accessible to the general public, and promoting data sharing. SDCs facilitate the searching, viewing, transferring, ordering, publishing and/or disseminating of spatial data and services from numerous sources via a web site (interface) on the Internet, and as appropriate, providing complementary services. Such SDCs contain data catalogues, which are access systems that use metadata (INSPIRE Architecture and Standards working group, 2002; Maguire and Longley, 2005; Tait, 2005).

The access service for spatial data on the web is known variously within the spatial community as clearinghouse, catalogue services, spatial data directory, geoportal and geospatial one-stop portal. Although different names are used it is obvious that the goals of accessing spatial data through the metadata remain the same (Crompvoets et al., 2004; Beaumont et al., 2005). The enhancement of data/service accessibility, and the sharing of spatial data and related services between suppliers and users are considered to be the main reasons to build these electronic facilities (Bernard et al., 2005c; Beaumont et al., 2005; Maguire and Longley, 2005).

Based on an overall assessment, the average cost of a SDC is around € 1,500,000 a year (Southern California Association of Governments, 1998; INSPIRE Architecture and Standards working group, 2002; Pasca et al., 2004). This money is spent on management and coordination costs, GIS and Internet application development, training, hardware, standardisation activities, legal environment creation, and metadata preparation. Currently, around 500 (non-corporate) SDCs have been established and it is expected that many more SDCs will be set up in the future. This indicates that on a global scale hundreds of millions are spent yearly on SDC activities. Up to now this large investment has rarely been audited or evaluated. A study conducted by the Urban and Regional Information Systems Association (Gillespie, 2000) cited that while the costs of SDC projects may be relatively easy to assess and highly 'front-loaded', the benefits are often very difficult to measure and may not emerge until well into the life of the SDC and depend on other factors coming into play (FGDC, 2002; Commission of the European Communities, 2004).

SDCs could be developed at different administrative levels ranging from local to state/provincial, national, and international levels, to a global level, to better access and share spatial data and related services. There is a need to address politicians and decision-makers to demonstrate the benefits of such a system. One of the difficulties in selling the benefits to decision-makers has been the paucity of systematic evidence of the full economic, social and environmental impacts. This was highlighted in the context of Geospatial One-Stop (FGDC, 2002), and the Extended Impact Assessment of the INSPIRE-initiative (Commission of the European Commission, 2004). However, it has been difficult to extrapolate impacts from these individual cases to reach more generalised conclusions. In addition, it is critical to move away from a narrow focus on the technical considerations of SDCs to their potential contribution to area competitiveness, innovation, productivity, job creation, etc. (Craglia et al., 2003).

The focus of this paper is on the worldwide impact assessment of the current SDCs with the main objective of providing this information to policy makers, in order to assist them in their task of evaluating whether or not investment in setting up and maintaining these SDCs is justified. In this context, the term 'impact' is described as the (positive or negative) effect that SDCs could have on society. Few studies exist about the worldwide impact of these facilities. To the best of the authors' knowledge, no comprehensive and systematic impact assessment has taken place. The purpose of the present paper is to fill this gap.

This paper presents and assesses the impacts of current SDCs throughout the world with reference to the economic, social, and environmental dimensions. The impact assessment presented here is based on a survey undertaken among coordinators of known SDCs of the world using indicators to assess the relevance, efficiency and effectiveness. Complementary analyses are implemented in order to interpret the significance of the impacts.

4.1.1 Introduction to impact assessment

Impact assessment is a key tool for improving policy-making and implementation, and promoting sustainable development (Long and Alastair, 1997; Commission of the European Communities, 2002; Bråthen, 2003). Many techniques can be used to assess the impacts (Jorgenson, 1998; Environmental Protection Agency, 2000), but whatever method is used the results need to be transparent, reproducible and robust. To make comparison accurate as possible, it is recommended that impacts are expressed in quantitative and monetary terms (e.g. cost – benefit analysis) in addition to a qualitative appraisal.

Impact assessment identifies and assesses problems arising from the pursued of the objectives, and the available options to achieve those objectives. It also highlights the positive and negative impacts with their respective advantages and disadvantages, including synergies and trade-offs (Commission of the European Communities, 2002; Bråthen, 2003). Any assessment should be based on the following criteria:

- Relevance for solving the problem.
- Efficiency in the use of human and financial resources.
- Effectiveness in achieving the defined objectives.

These assessments of impact are difficult mainly because of the degree of uncertainty in the reliability of the data, the assessments of the proportion of the impacts, the range of affected stakeholders, the short and long-term developments, and the efficacy of the assessment method.

Systematic assessment of impacts should also consider sustainable development. Sustainable development is based on the idea that in the longer run economic growth, social inclusion and environmental protection should go hand in hand. At this moment, many governments regard these economic, social and environmental dimensions as the main driving force behind their policies (Williamson et al., 2003). The economic, social and environmental impacts should be identified and cover all positive and negative effects, including costs and benefits. Economic, social and environmental impacts have been identified by the report of the European Communities (2002).

4.1.2 Existing impact assessment studies

Several studies assess the impact of SDIs including SDCs (Renong Berhad, 1995; PriceWaterhouse, 1995; Canadian Council of Land Surveyors, Canadian Institute of Geomatics, Geomatics Industry Association of Canada, 2000; Berends and Weesie, 2001; Fornefeld and Oefinger, 2001; FGDC, 2002; Pasca et al., 2004; Commission of the European Communities, 2004). These studies encountered difficulties in estimating the costs, while the estimation of benefits appeared to be even more difficult.

Previous assessment research focused mainly on the impact of one SDC and was neither comprehensive nor systematic (Price Waterhouse Nederland, 1996; FGDC, 2002; Commission of the European Communities, 2004; Pasca et al., 2004; Tait, 2005; Walther, 2005). As with many SDI-initiatives, the majority of impacts were qualitative in terms. The main findings of these six studies are that SDCs:

- Improve the availability, accessibility, usability and 'downloadability' of data supplied.
- Are cost effective and efficient. For example, the Benefit Cost ratio, related only to the reduction of time to access data, ranges from 1.1 to 4.
- Widen the range of users with different levels of education and technical skills.
- Increase the awareness of spatial data amongst the general public.
- Enhance the performance and productivity of (publicly funded) organizations.
- Improve metadata quality.
- Increase government participation.
- Support better decision making.
- Are catalytic to innovation and new ways of working.
- Improve partnerships.

These initial assessment results and literature (e.g. Groot and Sharifi, 1994; Askew et al., 2005, Maguire and Longley, 2005, Beaumont et al., 2005), suggest that SDCs are a relevant means to enhance data accessibility as well as data sharing, effective and efficient in the use of human and financial resources.

4.2 METHODOLOGY

This paper focuses on the development and implementation of a procedure to assess the impacts of currently existing international, national, federal, interstate, state, county and local SDCs of the world. The 'pre-clearinghouse situation' was considered to be the baseline against which to assess the current impact of SDC-development. The 'pre-clearinghouse situation' refers to when no electronic facility existed on the Internet to access spatial data using metadata. To undertake the assessment it was important to take into account developments over time, to use existing knowledge and experience, to consult interested parties and relevant experts, to be transparent, and to compare negative with positive impacts.

Assessment difficulties have circumscribed the very few studies containing quantitative and qualitative information on the impacts of SDCs. Therefore, the approach chosen in the study was to determine impacts by referring to the expert knowledge and experiences of SDC coordinators as their perceptions are sensitive indicators for changes as well as impacts. These coordinators organise activities as management, marketing, technical and legal environment creation, and human resources so that their SDCs operate well. Other reasons to focus on SDC coordinators were their intermediate role between data/service suppliers and users, their awareness of the historical, institutional, cultural, legal, economic and technological context, and their ability to provide accurate data about the development, use, management, content and technology of their SDC. Moreover, they were relatively easy to contact. This was not the case with the data users as well as suppliers of SDCs. In addition, the expertise and experiences of a selected number of European SDC practitioners (users and data/service suppliers) were used to evaluate the objectivity of coordinators' perceptions. The availability of this expertise meant that the impact in terms of economic, social and environmental context could be described fairly comprehensively.

The procedure used in this assessment study consisted of the following steps:

- Undertaking extensive literature research (see section 1.2 existing impact assessment studies).
- Determining assessment indicators in order to evaluate the relevance, efficiency and effectiveness.
- Designing and conducting survey in order to collect information about the perceptions of coordinators.
- Analysing results by categorisation of the SDCs in order to facilitate the interpretation of these results.
- Assessing the objectivity of coordinators' responses.

4.2.1 Determining assessment indicators

The assessment was confined to the use of number of economic, social and environmental impact assessment indicators, because a full implementation of a quantitative assessment study was proscribed by cost considerations. These indicators were measurable and illustrative (Taylor et al., 1990). They could measure the relevance, efficiency and effectiveness of SDCs and provide insight into how economic and social structure and environment alter when SDCs are implemented. The selection of indicators was based on expert knowledge, literature and direct relevance for SDCs.

The economic indicators used were:

- Consumption of data/services.
- Data market transparency.
- Duplication of data collection.

The social indicators were:

- Spatial data/service awareness.
- Social cohesion between citizens.

The only environmental indicator was:

- Data delivery for environmental policy formulation.

4.2.2 Designing and conducting survey

The survey was undertaken (November 2003 – April 2004) to collect information about the perceptions of coordinators. A questionnaire was distributed to all known coordinators of SDCs. This survey was strongly supported by the INSPIRE expert group (a group composed of representatives of the European Commission, and member states' Environmental and GI-communities) and the Executive Board of the Permanent Committee of GIS Infrastructure for Asia and Pacific (PCGIAP).

It was important that as many SDC coordinators as possible completed the survey to provide a full and reliable impact assessment. For this reason an inventory of identified SDCs was compiled by extensive browsing on the Internet (using several search engines), reading literature, contacting experts and SDC coordinators. Where possible the e-mail address (and name) of the SDC coordinator was collected.

A questionnaire was used to collect the relevant information. The questions were based on current literature as well as expert knowledge, so that the coordinators' perceptions of their SDC could be analysed. Most questions could be answered by selecting the appropriate option boxes; none of the questions were 'open'. The questions were framed in a way that they described the impacts of SDCs as well as the future developments. The questions were:

1) On which administrative level listed is your SDC mainly operating? (In section 4.2.3 the administrative levels listed are presented).
2) Which of the countries listed does your SDC cover (partly) metadata? (193 countries were listed).
3) Which of the options listed are the main benefits of your SDC? (Figure 4.3 presents the benefits listed).
4) Which of the options listed are the main drawbacks of your SDC? (Figure 4.4 presents the drawbacks listed).
5) Which of the options listed is likely to take place with your SDC within the next five years? (In section 4.3.3 the future options are partially presented).

Moreover, fourteen statements were formulated to assess what SDC coordinators considered to be the impacts of their SDC on a scale from strongly agree to strongly disagree. Examples of such statements were:

a) Your SDC increases the consumption of spatial data and services.
b) Your SDC improves data market transparency.
c) Your SDC reduces data duplication.
d) Your SDC improves the awareness of spatial data.

e) Your SDC strengthens the social cohesion between citizens. This statement refers to the solidarity and social bonding between people within state, country or international region.
f) Your SDC improves the appropriate data delivery for environmental policy formulation.
g) Establishment and maintenance of your SDC is economically beneficial.

In addition, there were supplementary statements designed to check the face validity of the responses.

The questionnaire was distributed using e-mail and was addressed personally to the coordinators. The main advantages of the use of e-mail are that it is fast, easy and cheap for distribution. In total, 428 coordinators were contacted.

4.2.3 Analysing results

The analyses were carried out to process the answers and to interpret the results better. The worldwide answers were aggregated. However, as the world is very diverse in historical, institutional, legal, cultural, technological and economic respects, and different Geographical Information (GI) processes take place at different administrative level, the variability of the answers between regions and administrative levels were categorically analysed. The classification by region was based on the division of Dorling Kindersley (2002). Eight administrative levels were identified: 1) worldwide, 2) continental, 3) international, 4) national (federal), 5) interstate, 6) state, 7) county and 8) local. The Chi square and Fisher exact tests (Agresti, 1990) were used to test whether respondents at different regional areas and administrative levels reacted differently to the questions and statements of the questionnaire. Throughout, test results with a (1-sided) p-value of less than 0.1 were considered significant.

4.2.4 Assessing the objectivity of coordinators' responses

As the results of the questionnaire were all based on the response from the SDC coordinators it was expected that their views could be biased. To mitigate this, a comparison of responses from the European SDC coordinators with those of the European user community was made, assuming that the objectivity of European coordinators' responses represent well the objectivity of all SDC coordinators' responses. To facilitate this procedure a short version of the questionnaire was distributed to 75 European representatives of the GI-user community (Summer 2004). These practitioners were member of the INSPIRE Expert Group, and were considered to be important stakeholders who could use SDCs to access or supply spatial data (e.g. ministries, municipalities, mapping agencies, cadastres, universities, public/private institutions, utilities, etc.). The Chi square and Fisher exact tests were also used to test the differences of the views between the European SDC coordinators and these practitioners.

4.3 RESULTS AND DISCUSSION

The inventory resulted in a list of 456 SDCs (of 80 countries) of which 428 had personal e-mail addresses of their SDC coordinator. Figure 4.1 indicates the worldwide distribution of all identified SDCs by country. It appears that the establishment of SDCs has become a global activity as recorded by Crompvoets and Bregt (2003), and Crompvoets et al. (2004). Most SDCs are established in Europe, Southeast Asia, North and South-America. The countries with the highest number of SDCs are USA and Canada. The areas with few implementations are Africa and The Middle East.

A total of 105 coordinators from 31 countries completed the survey; 25% of the population of coordinators. This percentage is in line with the responses to similar types of surveys (Hamilton, 2003). This sample size was adequate in respect to the SDC population in the developed world since the respondents were mainly coordinating SDCs in North America (USA/Canada) (41%), Europe (32%) and Australia (8%) (only 19% in total were African, South American and Asian ones (Figure 4.2)). In order to obtain reliable results, the regional analysis included only the North American, European and Australian ones. The other regions were excluded from the regional analysis due to limited number of responses.

As mentioned earlier, the survey identified eight administrative levels (question 1). To achieve reliable statistical analysis, several levels were reclassified. Finally, three classes were considered: 1) (inter)state, 2) national (including federal), and 3) international. Classes interstate and state were reclassified into (inter)state (41%); class national was unchanged (31%); classes worldwide, continental, and international were reclassified into international (20%); classes county and local were excluded from the administrative level analysis (8%).

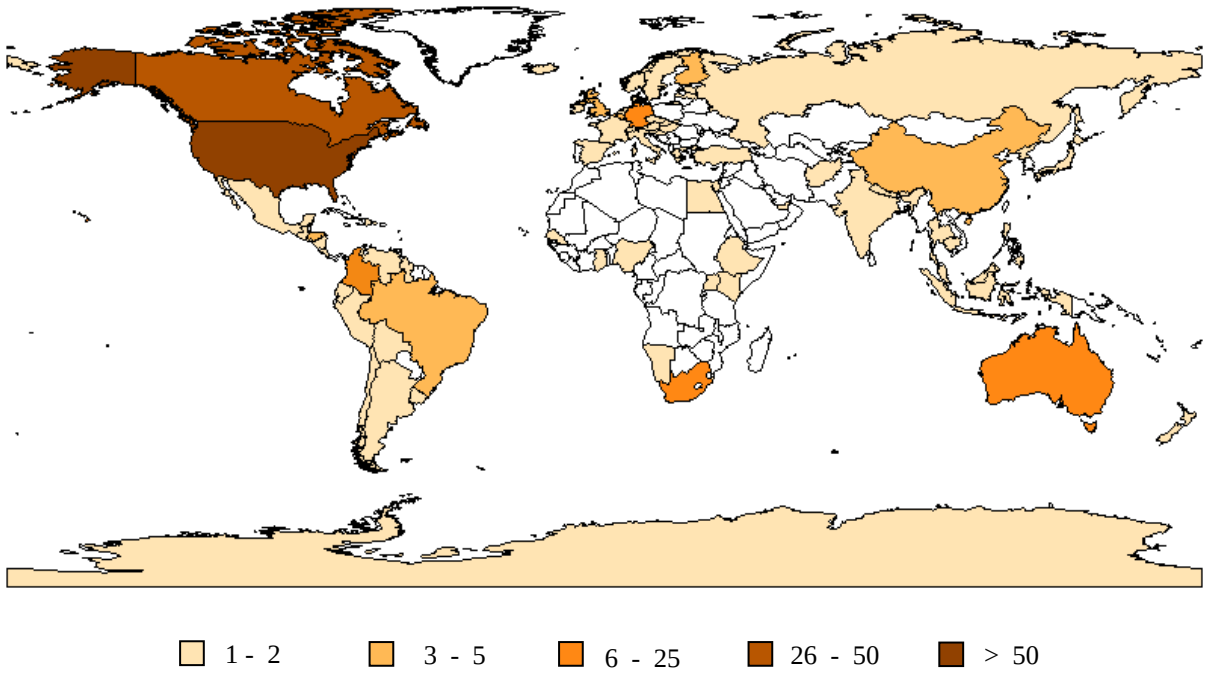


Figure 4.1 Worldwide distribution of spatial data clearinghouses (456) by country.

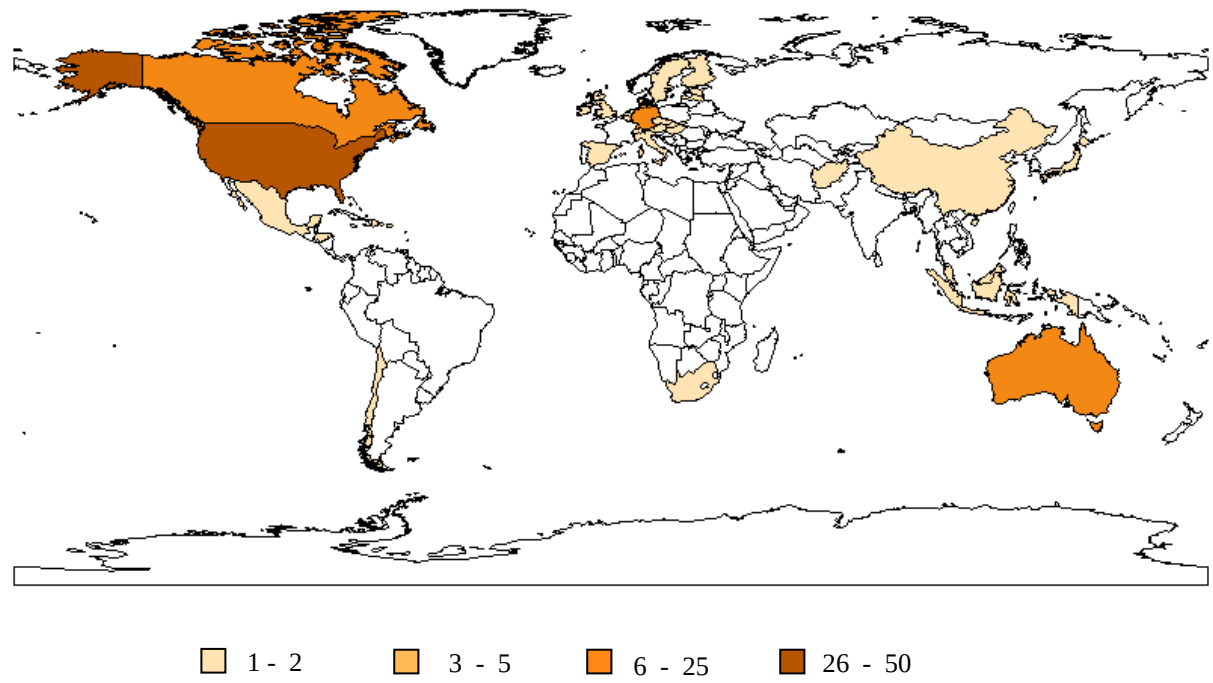


Figure 4.2 Worldwide distribution of survey responses (105) by country.

4.3.1 Benefits and drawbacks

The enhanced access to spatial data, and the improved data sharing and distribution are regarded as the main benefits (question 3) of the current SDCs (Figure 4.3). This confirms the results derived from the previous studies and literature (section 1.2). On the basis of this result, it seems that overall SDCs are relevant facilities to access data/services and to promote sharing. However, many SDCs still lack integration among suppliers and users. This could result in inefficient use of resources, potential duplication, inconsistency, incompatibility, and the inability to maximise the value of data and services. The main benefits appear to be mainly economic in nature. Minor benefits are the more effective use of available data, the improved spatial data awareness and the reduction of spatial data duplication. Cost savings are not really seen as a benefit, which could be an indication that SDC coordinators are not very cost conscious.

Coordinators of North-American SDCs regard the reduction of data duplication, the improved data sharing and distribution significantly more as benefits (this is in contrast with European SDCs).

In addition, coordinators of international SDCs see the reduction of data duplication significantly less as a benefit. This is in contrast with (inter)state coordinators who also look upon cost savings significantly more as a benefit.

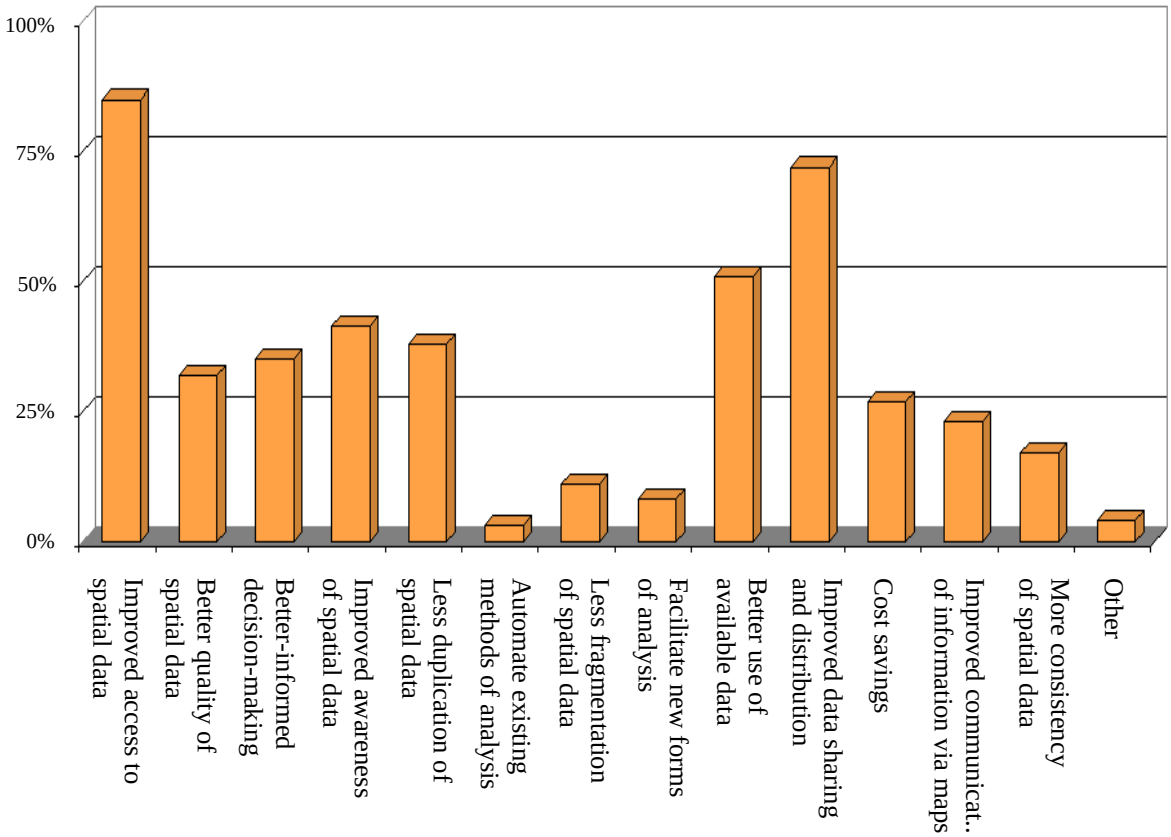


Figure 4.3 Worldwide distribution of SDC coordinators' responses (%) relating benefits of spatial data clearinghouses.

It appears that besides costs and funding (80%) not one single drawback (question 4) could be identified as another important obstacle for SDC implementations and maintenance (Figure 4.4). Institutional problems (33%), lack of specialised data managers (25%), and data standardization (23%) can however be considered as significant drawbacks. The lack of harmonised reference systems (3%), and liability problems (12%) and inadequate Internet bandwidth (16%) are less significant as drawbacks for SDC implementation. This result is in line with literature (INSPIRE Architecture and Standards working group, 2002; FGDC, 2002; Wehn de Montalvo, 2004; Askew et al., 2005). None of the main obstacles are (directly) technology-related. It seems that the challenges to be faced are more likely to be organizational than technical.

North-American coordinators consider lack of specialised managers significantly more as a drawback and problems with data pricing less. On the other hand, the European SDC-coordinators look upon problems with data pricing and commercialisation of data significantly more as a drawback.

The high degree of correspondence in coordinators' views with respect to the perceived benefits and drawbacks is significant in so far it gives a clear indication that SDCs worldwide function within a broadly similar operating environment.

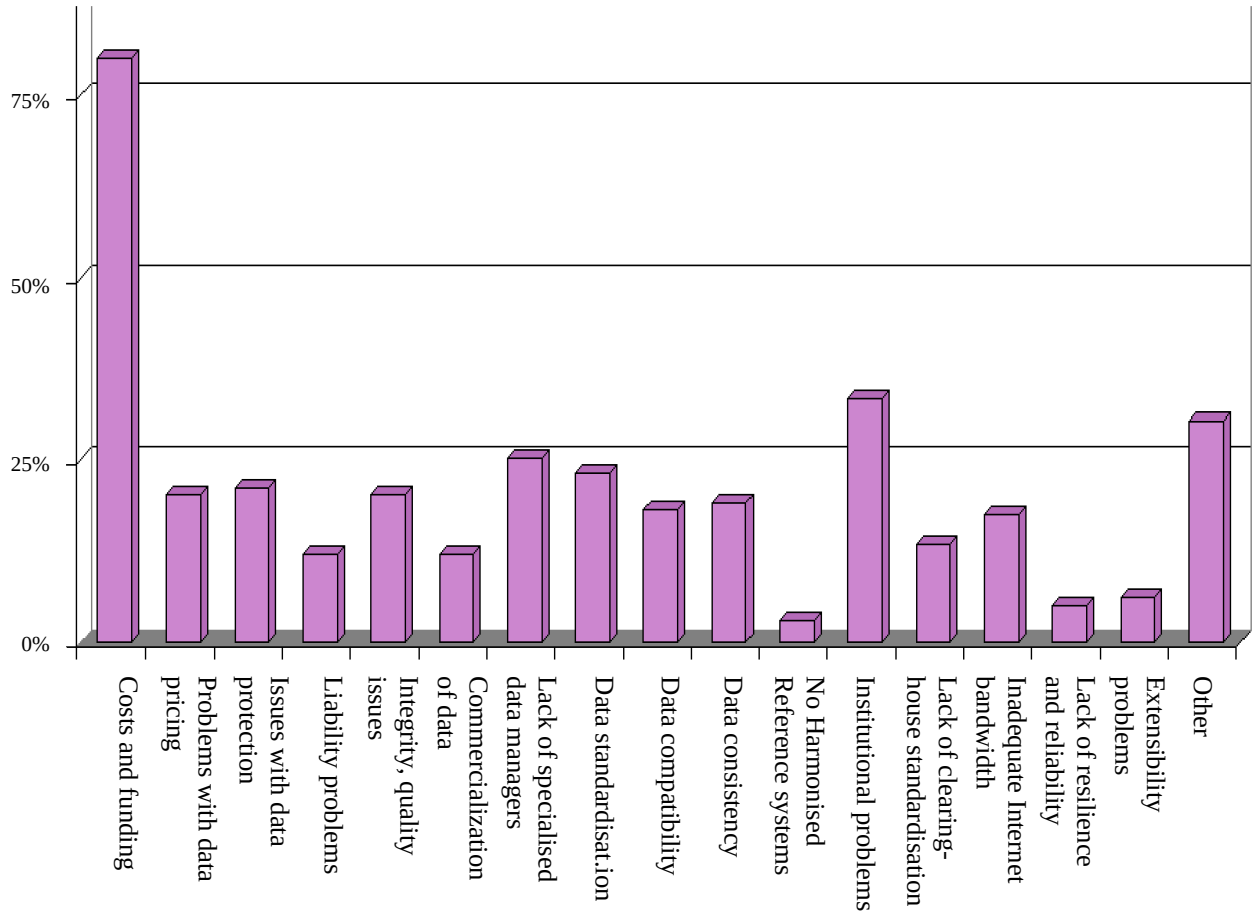


Figure 4.4 Worldwide distribution of SDC coordinators' responses (%) relating drawbacks of spatial data clearinghouses.

4.3.2 (Economic, Social and Environmental) impacts

Economic impact

The economic impact is primarily assessed by the use of economic indicators. Several statements in the questionnaire refer to these economic indicators. The survey results show the likelihood of higher consumption of spatial data and services, as well as the reduction of data duplication as the main economic impacts. This impact result is illustrated in Figure 4.5, which presents the responses of SDC coordinators to three economic indicators: consumption of data and services (statement a), data market transparency (statement b), and duplication of data collection (statement c). On the basis of these results, it is apparent that the vast majority of respondents agree with the statement that their SDC increases the consumption of spatial data and services. This implies that this increase of consumption could be regarded as the most important economic impact. Additionally, a majority also agrees with the statement that their SDC reduces duplication of spatial data. The result related to the statement that SDC improves data market transparency is not clear (the majority neither agrees nor disagrees). On the basis of the responses related to these three economic indicators it could be deduced that SDCs have a significant (positive) impact on the economic dimension.

Looking from regional perspective, evidence can be found that more North-American coordinators agree with the statements that their SDC increases the consumption of spatial data and services, and reduces duplication of spatial data.

Evidence exists that national SDCs agree less that their SDC increases the consumption of spatial data and services while (inter)state SDCs agree more that their SDC reduces duplication of data.

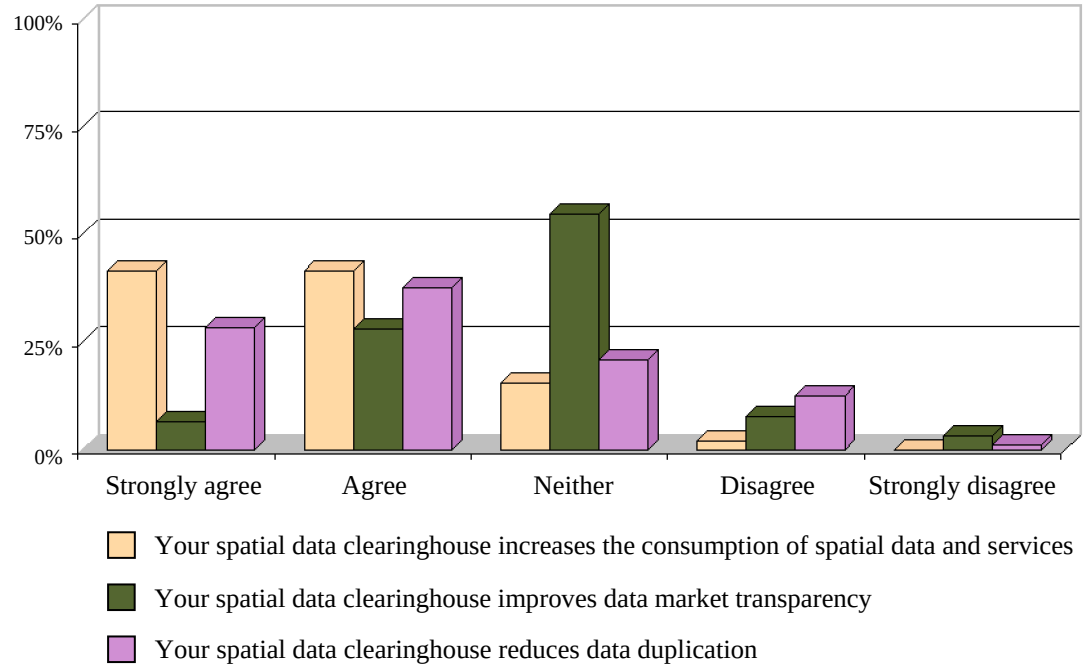


Figure 4.5 Worldwide distribution of SDC coordinators' responses (%) to statements relating economic indicators.

Besides the statements directly related to the indicators the coordinators could also respond to the statement that establishment and maintenance of their SDC is economically beneficial (statement g). 70% of the coordinators agree and only 11% disagree with this statement. Since the main benefits and drawbacks are likely to be economic in nature, this result indicates that SDC coordinators perceive that the positive impacts more than counterbalance the negative impacts.

Both data users and suppliers could gain economically by the implementation of SDCs. Data users benefit from the improved efficiency to access spatial data, and data suppliers from the increased effectiveness to distribute their spatial data and the improved efficiency to collect data by reducing data duplication. It seems that the establishment and maintenance costs of these facilities are economically justified, although the cost savings for the SDC coordination organisations appear a less important impact.

Social impact

The social impact is primarily assessed by the use of social indicators. Two statements in the questionnaire refer to these indicators: spatial data/service awareness (statement d), and social cohesion between citizens (statement e). These impact results are illustrated in Figure

4.6. From the responses of SDC's coordinators, it is apparent that the vast majority agrees that their SDC improves spatial data awareness. This implies that this improvement of spatial data awareness could be regarded as the most important social impact. It appears that SDCs could change the way society is using this spatial data. In many decision-making processes the role of spatial data is increasing. SDCs improve (indirectly) these processes in a way that enable stakeholders to become better informed. Additionally, a majority also agrees that their SDC strengthens the social cohesion. It appears that SDCs are, for example, able to provide equal spatial information access to rural, urban and remote communities, which will support local decision making capacity development and new socio-economic activities in these communities. In view of these social results it is reasonable to deduce that SDCs have a significant impact on the social dimension.

From regional perspective, evidence exists that North-American coordinators agree more with the statement that their SDC improves the awareness of spatial data.

From administrative level perspective, no differences in agreement exist.

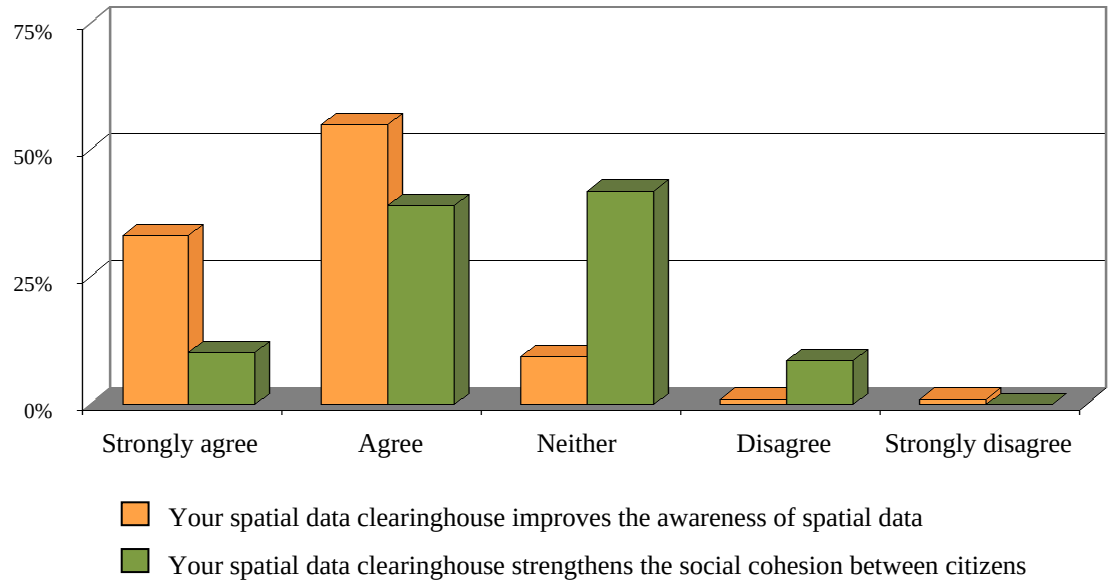


Figure 4.6 Worldwide distribution of SDC coordinators' responses (%) to statements relating social indicators.

Environmental impact

The environmental impact is assessed by the use of one environmental indicator: data delivery for environmental policy formulation (statement f). The coordinators expect little impact on the environment. From the response it appears that the majority of the coordinators neither agree nor disagree (60%) with statement f. SDCs do not seem to deliver the data appropriately for environmental policy formulation. Nevertheless, some environmental policy-makers make use of SDCs to access the needed spatial data and services (Williamson, 2004).

From regional perspective, the evidence indicates that North-American coordinators do not consider that this impact is important.

From administrative level perspective, no differences in agreement exist.

Examining assessment indicators in combination with the benefits, it appears that the main positive impact of implementing SDCs is economic. The high degree of correspondence in coordinators' views with respect to the economic, social and environmental impacts is significant confirming that SDCs worldwide function within a broadly similar operating environment.

4.3.3 Future developments

The coordinators were asked to select what they expect will happen with their SDC in the next five years (Question 5). A subset of their response was that:

- The use of spatial data will increase (89%).
- More (new) services will be provided (55%).
- The data quality will improve (50%).
- The use by governments will increase (49%).
- More datasets will be provided (35%).
- More specific datasets will be needed (34%).
- The metadata standards applied will be changed (31%).
- New expertise will be needed (26%).

The coordinators expect mainly that the spatial data consumption as well as the range of service provision of their SDC will increase. These developments are in line with literature (Maguire and Longley, 2005; Beaumont et al., 2005), and link strongly to the gradual shift in focus of SDC development: from data-centric to user-centric. In the nineties of previous century, data and technology were the main driving forces for SDCs. At the present moment, the use of data (and services) and the needs of the users are becoming the main forces for SDC development (Reeve and Petch, 1999; Williamson et al., 2003; Crompvoets et al., 2004).

The similarity in development views of the coordinators is significant showing that the coordinators have the same future objectives probably created by such external developments as expanding technologies, market-demand, changing business models, sustainable development, e-government and participatory democracy. The few differences are that more North-American coordinators expect that more datasets will be provided, and new expertise will be needed.

4.3.4 Assessment of the objectivity of coordinators' responses

A total of 41 European practitioners completed a short version of the questionnaire. The high degree of correspondence between the responses of these European practitioners and the (34) European SDC coordinators with respect to the questions and statements is significant. This result implies that the coordinators' perceptions are not unduly biased (at least the European coordinators' perceptions), and justifies the choice to focus on SDC coordinators as reliable sources of information to assess the impacts. Furthermore, the practitioners look upon cost savings as a more significant benefit, and consider the improved awareness of spatial data as a less important impact. This indicates that the coordinators underestimate the efficiency of SDCs and overestimate the improved awareness.

4.3.5 Methodology used

The implementation of the assessment procedure was appropriate to measure the impact of SDCs on a worldwide scale in order to assist policy makers to decide whether investments in the establishment and maintenance of SDCs are justified. The strength of this impact assessment was that it was systematic, reproducible, robust, expert knowledge based, and that it identified significant economic and social impacts. Through the survey it was possible to gather the perceptions of the coordinators in a fast, cheap, and easy way. The complementary analyses were needed to interpret the results of the survey. The main limitation of this study was that only qualitative impacts could be assessed and it was not possible to determine quantitative measures such as financial impacts. The current experiences of the SDC operations are limited by the fact that they are still at an early stage of their development. There is a need to refine methodology so that more precise records of numerical and financial data can be recorded. In this way, a better and a more accurate grasp of financial and operational impacts could be delivered. Nevertheless, the usage of indicators gave some insight into how economic, social structure, and environment alter when SDCs are implemented.

4.4 CONCLUSIONS

The main conclusions of this impact assessment referring primarily to SDCs of the developed world are:

- SDCs are likely to have a positive impact on society.
- The main (positive) impacts are of an economic nature, but social impacts are obviously important as well.
- SDCs have likely little impact on the environment.
- SDCs could be considered as relevant facilities in order to enhance spatial data/service accessibility and to promote the sharing of these resources.
- SDCs could be considered as efficient facilities in order to enhance data/service accessibility and to reduce data duplication.
- SDCs could be considered as effective facilities in order to increase the use and distribution of spatial data/services, to improve the awareness of spatial data/services, to strengthen social cohesion between citizens, and to improve potentially better-informed decision-making.
- Costs and funding could be regarded as the main obstacle for SDC-implementation.
- In the near future, it is expected that the use of spatial data resources of SDCs will increase as well as the range of service provisions.
- Coordinators have similar views towards the benefits, drawbacks, impacts as well as future developments of SDCs. These similarities could form a perfect basis to ensure interoperability between datasets and access mechanisms, and to create a culture of sharing as well as a shared language amongst coordinators.
- North-American SDCs are considered to be the most efficient and effective facilities, and have substantial acceptance within the community. This is in line with Maguire and Longley (2005), who mention that many US as well as Canadian SDCs already in the nineties of previous century were able to promote awareness of spatial data, create community involvement, and build capacity to access this data (Maguire and Longley,

2005). The Australian SDCs form the intermediate in efficiency and effectiveness between North-American and European SDCs.

- The diversity in benefits, drawbacks, impacts, and future developments between the different administrative levels appear to be low. This could imply that the GI-processes relating to spatial data/service accessibility do not vary much at different administrative levels.

The results obtained could be used to justify present and support future investments in SDCs. However, the authors observe that in spite of these positive results in terms of relevance, efficiency and effectiveness, the SDC concept to share resources continues to be resisted which leads to unnecessary inefficiencies resulting in duplication of data collection and storage, and consequent costs (Nedovic-Budic and Pinto, 2000; FGDC, 2002; Askew et al., 2005). To utilise these SDCs effectively there must be a clear understanding of how they influence and justify their costs, and overcome institutional problems. It therefore appears that more impact assessment research is needed (e.g. case studies).

ACKNOWLEDGEMENTS

The authors would like to thank Hugh Buchanan, Watse Castelein, Michael Gould, Louis Hecht, Günther Pichler, Gerda Schennach, Paul Smits and Marc Vanderhaegen for their contributions to the generation of the questions of the questionnaire, and the members of the INSPIRE Expert Group and the Permanent Committee on GIS Infrastructure for Asia & Pacific (PCGIAP) to disseminate the questionnaire to potential SDC coordinators. Moreover, the authors gratefully wish to acknowledge the support of all the SDC coordinators and the European practitioners in preparation of this paper. Their contributions by answering the questions and statements were crucial for the results presented in this paper. Finally, the authors thank Gert Jan Hofstede, Brian Goodsell and Rini Kools for comments and suggestions on earlier versions of this paper.

Chapter 5

Establishment of national spatial data clearinghouses, exploring worldwide the impact of society

J. Crompvoets, M. Wachowicz, A. Bregt, G.J. Hofstede, J. Wang, P. Vereijken, E. de Man

Submitted to Transactions in GIS

5.1 INTRODUCTION

Over the last decade, many countries have spent considerable resources on a National Spatial Data Infrastructure (NSDI), in order to improve the accessibility and sharing of spatial data (Groot and McLaughlin, 2000; Williamson et al., 2003; Bernard et al., 2005b; Masser, 2005). One of the key features of a NSDI is the national spatial data clearinghouse (Crompvoets et al., 2004). This national clearinghouse is the access network of an NSDI that facilitates access to spatial data and promotes data sharing.

A spatial data clearinghouse can be defined as an electronic facility for searching, viewing, transferring, ordering, advertising and/or disseminating spatial data from numerous sources via the Internet, and as appropriate, providing complementary services. Such a clearinghouse usually consists of a number of servers containing information (metadata) about available digital data (Crompvoets et al., 2004). The challenge in establishing clearinghouses lies in the management of a substantial number of different disciplines and the examination of a large numbers of factors and issues. For example, creating community participation, reaching agreement between stakeholders, building a common access network, etc. (Rajabifard, 2002; Williamson et al., 2003; Tait, 2005).

Since 1994, there has been a steady increase in the number of established national clearinghouses worldwide. The US Federal geographic Data Committee established the first in 1994. By April 2005, 83 national clearinghouses were operational (Crompvoets et al., 2005). This number is expected to increase in the near future (Crompvoets and Bregt, 2003; Crompvoets et al., 2004). Considering that 193 countries exist in the world, more than 50% do not, as yet, have an established national clearinghouse. Concerning the high percentage of countries without a clearinghouse, interesting questions to be investigated are: What are the critical factors influencing such establishment of national spatial data clearinghouses? And, why have some countries established a national clearinghouse, and why have other countries not done so?

Answering these questions is not easy although various authors have addressed them. Some mention as differentiating factors, political vision, institutional leadership, human capital, web connectivity, telecommunication infrastructure or financial resources (Tosta, 1997; Streudler, 2003; Reece, 2004; Crompvoets et al., 2004; Delgado et al., 2005). Since each country has its own unique societal conditions, it is likely that society plays an important role. Moreover, there is common agreement in the literature that societal issues, including those which are institutional, legal, political, and organizational, are more crucial for clearinghouse implementation than technical issues (Bregt, 2000; Groot and McLaughlin, 2000; Crompvoets and Kossen, 2002; Williamson et al., 2003; Tait, 2005; Masser, 2005; Bernard et al., 2005c). However, detailed and systematic knowledge about these critical factors is limited. The aim of the research presented in this paper is to explore societal factors that are critical for the establishment of clearinghouses and to provide answers as to why some countries have a national spatial data clearinghouse established, and why have other countries not done so. Besides these societal factors, clearinghouse-internal factors such as expenses for project management, staff, marketing, training, network server, software, standardisation activities and metadata preparation, could have an impact on the clearinghouse establishment as well. In the context of this paper, society is considered to be a set of societal attributes describing human conditions and activity, regarded as a whole, functioning interdependently at a national level. As we are dealing with a complex system,

with many unknown relations and interactions, an explorative and empirical analysis approach was followed.

5.2 METHODOLOGY

This study aims to explore societal factors that could be critical for the establishment of national clearinghouses. Results of a web survey conducted, and datasets, including values of numerous societal attributes, were the starting point of this study.

In December 2002, a web survey dealing with the worldwide status of national clearinghouses was conducted. In that month, 67 countries had established a national clearinghouse on the Internet, while 126 countries had no established clearinghouse. In addition, several clearinghouse characteristics were measured and recorded (Crompvoets et al., 2004)

In order to carry out a thorough analysis of society, nine societal aspects were investigated: 1) Economy, 2) Education, 3) Technology, 4) Environment, 5) Culture, 6) Demography, 7) Institution, 8) Health care and 9) Jurisdiction. These aspects cannot be regarded as mutually independent. In particular culture has an impact on the other aspects. Each societal aspect is described by a set of underlying societal attributes. The values of numerous societal attributes around the year 2002 were collected from different data sources.

Figure 5.1 presents the main methodology steps. These steps are: 1) Collecting societal data, and 2) Analysing societal factors for establishment. Each step is described in more detail.

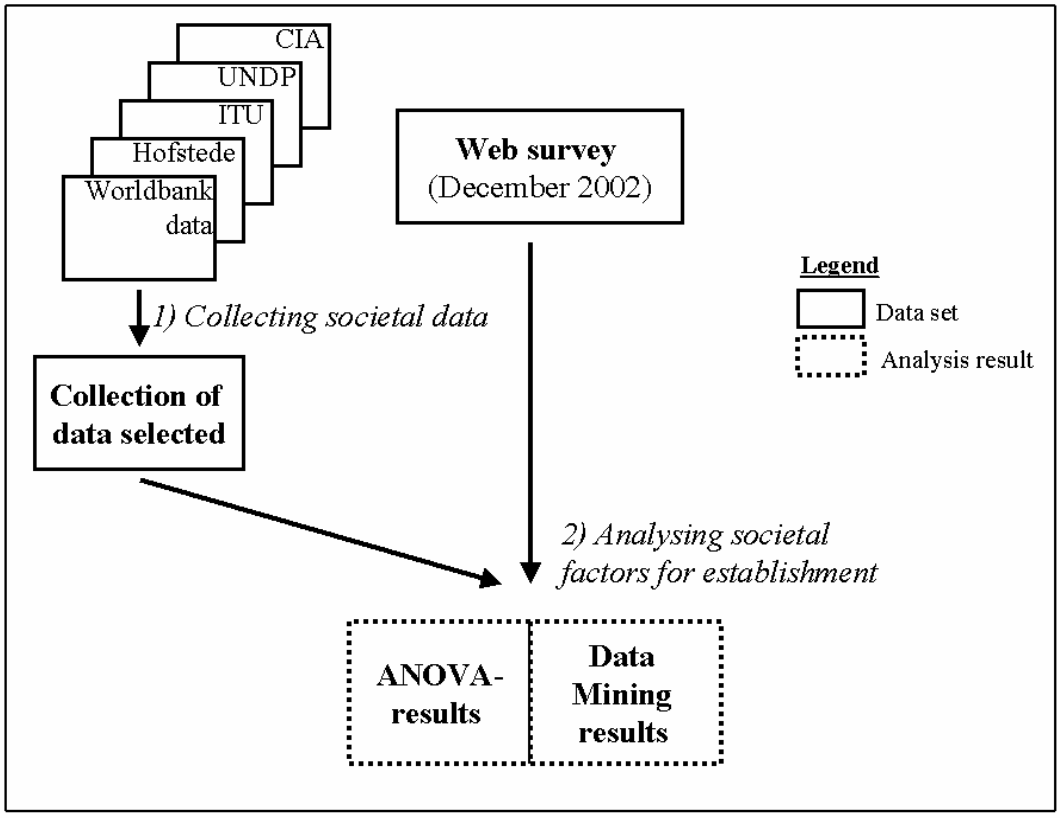


Figure 5.1 Main methodology steps.

5.2.1. Collecting societal data

Data were collected for 193 countries, as existing in 2002, according to United Nations. The World Development Indicators published by The World Bank (2003; 2004), the World Telecommunication Indicators published by International Telecommunications Union (2003), data in The World Factbook published by Central Intelligence Agency (2003; 2004), and data in Human Development Reports published by United Nations Development Program (2003; 2004), as well as a few data from Dorling Kindersley's Great World atlas (2002), were collected. The data on national culture were taken from Hofstede (2001). All attributes together describe the state of a society. These attributes collected were mainly classified into societal aspects in accordance with the different sections of The World Bank's World Development Indicators books (2003; 2004) and UNDP's Human Development report (2003; 2004). The selection of attributes was based on the knowledge of four experts in the domain of spatial data infrastructures. The main criteria used for selection were:

- Having a potential link with the use, management, technology and/or content of national clearinghouses.
- Not being outdated (most data before 2000 was disregarded).
- Not missing too many values.
- Representing one of the nine societal aspects.

Prior to analyses, pre-processing of data took place. This pre-processing included:

- Classifying countries into two establishment classes for a national clearinghouse (using the results of the web survey conducted): 1) countries with an established national clearinghouse, 2) countries with no established national clearinghouse.
- Handling missing values. Countries as well as attributes with more than 50% of missing values were removed. In addition, the missing values of remaining attributes were replaced by the mean or mode of each establishment class; mean for the quantitative attributes, mode for qualitative attributes.

5.2.2. Analysing (societal) factors for establishment

Two complementary methods were used in order to analyse societal factors for the establishment of a national clearinghouse: analysis of variance (ANOVA) and decision trees of data mining. Together they examine the factors for clearinghouse establishment more comprehensively and strengthen the findings. Both methods focus on (significant) relations between variables. The relations between the establishment class and societal aspects, as well as attributes, were analysed.

ANOVA (Snedecor and Cochran, 1980) was used to test, per societal attribute, the differences between the means of the two establishment classes. ANOVA is a statistical method used to determine the significance of the differences among the means of two or more groups on a variable. The independent variables are usually nominal, and the dependent variable is quantitative. In this study, the establishment classes are the independent variables; the societal attributes are the dependent ones. Throughout, test results with a p-value of less than 0.01 were considered to be significant. In order to compare the results at society level, the number of significant different attributes was summarized as a percentage of all attributes for each societal aspect. The attributes of the societal aspect jurisdiction were excluded from these analyses, because most attributes collected were nominal in nature.

A data mining method (Dilly, 1995; MacEachren et al. 1999; Wachowicz, 2002) was applied to select the societal attributes, which can infer the reasons for the establishment of national clearinghouses. Several methods have been proposed in the literature (Murthy, 1998) in pursuit of one of the general tasks of clustering, classification, generalisation and prediction. The classification mining tasks using the decision tree were used to examine the societal attributes and places into establishment classes. A decision tree contains the mapping between the attributes and the class. This is achieved by computing internal nodes by partitioning the data recursively by selecting one attribute each time. The model contains splits, which are a set of statements about the dependencies among attributes in a rule form using entropy and information gain. Entropy is a measure applied to determine the disorder in the population, according to Information Theory (Shannon, 1948). In this study, entropy represents the degree of randomness of the classification of national clearinghouses into two establishment classes; the greater the randomness (disorder), the greater the entropy. Tree construction is achieved by expending tree nodes that contribute to the largest gain in information of the whole tree. The aim is to select recursively an attribute, which results in the largest information gain. Information gain is also a concept sourced from the Information Theory (Shannon, 1948). It defines the increase in information, which results from adding a new attribute node to a rule. It is equal to the total entropy for an attribute if for each of the attribute values a unique classification can be made for the result attribute. In order to compare the results at society aspect level, nine decision trees were created for each society aspect, and the number of attributes that contribute to the largest gain in information was summarized as a percentage of all attributes for each society aspect. In addition, out of the nine decision trees, a final decision tree was created for the whole of society using all attributes that contribute to the largest gain in information. Wang (2005) presents a detailed description of the data mining method applied, as well as the set of parameters used.

5.3 RESULTS AND DISCUSSION

The presentation of the results follows the two methodology steps.

5.3.1 Collecting societal data

Through the data collected of 234 societal attributes, an indicative description covering the wide scope of society was achieved. Table 5.1 presents the number of attributes collected and examples of attributes per aspect of society. These examples indicate only the numerous attributes that were collected. Appendix 1 shows the full list of attributes collected (including units, attribute scales, and sources). From the sources available, it appeared that more economic, technological, environmental, and demographic attributes could be collected than cultural, institutional, and legal attributes.

Many attributes collected affect data availability, and/or reliability. This lack of quality is primarily due to the weak statistical systems in many countries with the consequence that statistical methods, coverage, practices, and definitions differ widely (International Telecommunication, 2003; UN Development Programme, 2003; 2004; The World bank, 2003; 2004). Data coverage may not be complete for countries experiencing problems (such as those stemming from internal or external conflicts) with the collection of data.

No data about the geo-information (GI) community were collected. Current literature suggests that the GI-community could have a high impact on clearinghouses (Steudler, 2003; Ravi, 2003; Reece, 2004; Delgado et al., 2005). It could also be argued that both the existence of spGI-communities and the establishment of national clearinghouses could depend on common causal factors. Unfortunately, collecting data referring specifically to the GI-community on a worldwide scale appeared to be impossible.

In total, 23 countries with too many missing values were removed (Andorra, Bahamas, Barbados, Brunei, Democratic Republic Korea, Dominica, Guyana, Holy See, Iceland, Iraq, Liberia, Libya, Liechtenstein, Luxembourg, Monaco, Myanmar, Nauru, Qatar, San Marino, Sierra Leone, Somalia, Timor-Leste, Tuvalu).

Table 5.1 Number of (societal) attributes collected and examples of attributes per societal aspect (in parentheses the units used).

Societal aspect	Number of attributes	Examples of attributes
Economy	77	Gross national income ($ billions), purchasing power parity per capita ($), gross domestic product (average annual % growth, 1990-2001), agricultural productivity, value added per worker ($), tax revenue (% of gross domestic product), listed domestic companies, household final consumption expenditure ($ millions)
Education	15	Adult literacy rate (% of age 15 and above), education expenditure (% of gross national income), average years of schooling, primary pupil teacher ratio (pupils per teacher), primary completion rate (% of relevant age group)
Technology	44	Internet users (thousands), Internet secure severs, personal computers per 1000 people, mobile phone subscribers (thousands), ICT-expenditures per capita ($), high technology exports ($ millions)
Environment	37	Surface area (thousand km²), arable land (% of land area), carbon dioxide emissions per capita (metric tons), energy use per capita (kg of oil equivalent), freshwater resources (total renewable resources per capita meter³)
Culture	5	Individualism (index), uncertainty avoidance (index)
Demography	26	Population density (people per km²), average annual population growth rate (%, 2001-2015), total fertility rate (births per woman), life expectancy (years), rural population (% of total population), infant mortality rate per 1000 live births
Institution	6	Type of government, institutional investor credit rating, year of independence
Health care	13	Health expenditure per capita ($), prevalence of HIV (% of adults), births attended by skilled health staff (% of total)
Jurisdiction	11	Type of legal system

5.3.2 Analysing (societal) factors for establishment

Through the complementary use of ANOVA and data mining decision trees, the exploration of societal factors for national clearinghouse establishment was achieved.

In total, 164 quantitative attributes were analysed by ANOVA to test differences between means of the two establishment classes. After pre-processing, 60 countries with an established national clearinghouse, and 110 countries without an established national clearinghouse, were involved in the analyses. Table 5.2 presents the percentage of attributes, with significant differences between means of the two establishment classes for each societal aspect.

Table 5.2 Number of attributes analysed using ANOVA and percentages of attributes with significant differences between means of the two establishment classes for each societal aspect (Clearinghouse, No Clearinghouse), as well as overall ($p < 0.01$).

Societal aspect	Number of attributes	% of attributes with significant differences between Clearinghouse – No Clearinghouse
Economy	48	65
Education	12	92
Technology	30	90
Environment	29	50
Culture	4	25
Demography	26	65
Institution	4	100
Health care	11	91
Overall	**164**	**71**

From Table 5.2, we see that 71% of the attributes analysed differ significantly between countries with an established clearinghouse, and countries with no established clearinghouse. Each of the societal aspects has significant attributes. The aspects education, technology, institution and health care show very high percentages of significant attributes (higher than the overall percentage). On the other hand, societal aspects, economy, environment, culture and demography show relatively low percentages. In general, the high percentages of significant attributes can indicate a high correlation among the societal attributes. Table 5.3 presents examples of significant attributes, including the means for both classes. From this table it appears that countries with a high standard of living have an established national clearinghouse, e.g. citizens of these countries are more educated, more technologically advanced, healthier, have higher incomes, and live in a more institutional stable environment. A higher standard of living also seems to be connected to higher carbon dioxide emissions, figures of which are also included. In line with the findings of Craig et al. (1992) and Hofstede (2000, 2001), this high standard of living is more likely to cause individualism, society's way of accommodating the interests of the individual. Individualist cultures comprise calculating citizens: what is in it for me, one's life is one's own, individual views matter, group views are necessary unavoidables to be challenged whenever one feels like it (Hofstede, 2001).

On the basis of the very high number of societal attributes (and aspects) with significant differences, it appears that the societal impact on the establishment of national clearinghouses is very strong. A reason that society could have such a strong impact on the establishment is that clearinghouses should facilitate the need of society to spatial data accessibility as well as data sharing. In addition, having such facility could promote society's economic development, (e)governance, and environmental sustainability (Masser, 2005). It seems that the relation between having a clearinghouse and a high standard of living is cyclic in nature; the one promotes the other.

Table 5.3 Examples of significant attributes sorted by society aspect (including the means for establishment classes countries with an established national clearinghouse, and countries with no established clearinghouse).

Attribute name	Societal aspect	Mean Clearing-house	Mean No Clearing-house
Gross national income ($ billions)	Economy	455	11
Agricultural productivity, value added per worker ($)	Economy	15,894	1766
Listed domestic companies	Economy	541	149
Household final consumption expenditure ($ millions)	Economy	298,772	10,155
Adult literacy rate (% of age 15 and above)	Education	7.1	23.4
Average years of schooling	Education	7.6	3.7
Primary pupil-teacher ratio (pupils per teacher)	Education	20	35
Internet users (thousands)	Technology	7820	189
Internet, secure servers	Technology	2044	9
Personal computers per 1000 people	Technology	190	18
Mobile phone subscribers (thousands)	Technology	15,190	939
Surface area (thousand km²)	Environment	1291	346
Carbon dioxide emissions per capita (metric tons)	Environment	8.2	2,3
Individualism	Culture	48	26
Average annual population growth rate (2001-2015) (%)	Demography	0.77	1.89
Total fertility rate (births per woman)	Demography	2.3	3.8
Life expectancy (years)	Demography	72.3	61.0
Institutional investor credit rating	Institution	60	24
Year of independence	Institution	1864	1955
Health expenditure per capita ($)	Health care	875	66
Prevalence of HIV (% of adults)	Health care	1,0	4.2
Births attended by skilled health staff (% of total)	Health care	87	63

In total 179 (societal) attributes were involved in the data mining method (including 15 qualitative attributes). The same number of countries was involved as in the ANOVA analyses. Table 5.4 shows the number of attributes for each of the nine society aspects (including jurisdiction), as well as the percentages of attributes that contribute to the largest gain in information. From this table, we see that overall 35% of the attributes contribute to the largest gain in information. In addition, each aspect of society has attributes that contribute to the largest information gain. Society aspects education, technology, culture, institution and health care show the highest percentages (higher than the overall percentage). Society aspects economy, environment, demography and jurisdiction show relatively low percentages. This result confirms the main society aspects of Table 5.2 (education, technology, institution and health care), except that culture has a higher percentage. Moreover, the results of Table 5.4 confirm that many societal factors could be critical for the establishment of national clearinghouses. This is in line with literature results concerning Spatial Data Infrastructures (Groot and McLaughlin, 2000; Williamson et al., 2003; Spatial Application Division, Catholic University of Leuven, 2003; Craglia et al., 2003; Delgado et al., 2005; Masser, 2005).

Table 5.4 Number of attributes, and percentages of attributes that contribute to the largest gain in information per society aspect.

Societal aspect	Number of attributes	% of attributes that contribute to the largest gain in information
Economy	48	29
Education	12	50
Technology	30	37
Environment	35	23
Culture	4	50
Demography	26	31
Institution	6	83
Health care	11	55
Jurisdiction	7	29
Overall	**179**	**35**

From the decision tree for the whole of society, it appeared that the following attributes could be critical for the clearinghouse establishment:

- Agricultural productivity, value added per worker ($) (Economy).
- Average years of schooling (Education).
- Primary pupil / teacher ratio (Education).
- Internet, secure servers (Technology).
- Average annual population growth rate, % (2001-2015) (Demography).
- Year of independence (Institution).
- Births attended by skilled health staff (% of total) (Health care).

All these attributes appeared to be significant attributes from the ANOVA-analyses as well (Table 5.3). These attributes presented confirm that countries whose community has a high standard of living are likely to be in a position to establish national clearinghouses rather than be concerned with daily problems of survival. Standard of living links strongly with the national well-being that includes all aspects of community life which influence the physical and mental health of the members of a nation (Moriaty, 1996; Henderson et al., 2000, Centers for disease control and prevention of US department of health and human services, 2005).

Agricultural income refers indirectly to the standard of living of farmers who produce the needed food and other agricultural products of a country. In many countries agriculture is the main source of employment. The World Bank (2003) estimates that every $1 earned by a farmer, raises incomes in other sectors by as much as $2.60. The educational attributes refer to the individuals who construct knowledge; and participate as informed citizens in society (The World Bank, 2003). The Internet attribute refers to the information and communication technology that has the potential to offer better service delivery. The attribute population growth rates refer to country's long-term sustainability (United Nations Department of Economic and Social Affairs, 1992). A rapid population growth rate can place a strain on a country's capacity for handling a wide range of economic, social and environmental issues (European Commission, 1999). The year of independence could be linked to the institutional stability and security of a country. Finally, attribute births attended by skilled health staff refers to the service quality of medical care.

As an illustration, Figures 5.2 and 5.3 indicate spatially the strong relations between the establishment classes and two attributes from the decision tree for the whole of society. Figure 5.2 presents the spatial relationship between countries with established national clearinghouses and the attribute average years of schooling. Almost all countries with an established national clearinghouse (except three) appear to have an average year of schooling above 4.6 years. Figure 5.3 presents the spatial relationship between countries with no established clearinghouse and the attribute agricultural productivity, value added per worker ($). Almost all countries with no established national clearinghouse (except two) appear to have a low agricultural productivity (lower than $3,590).

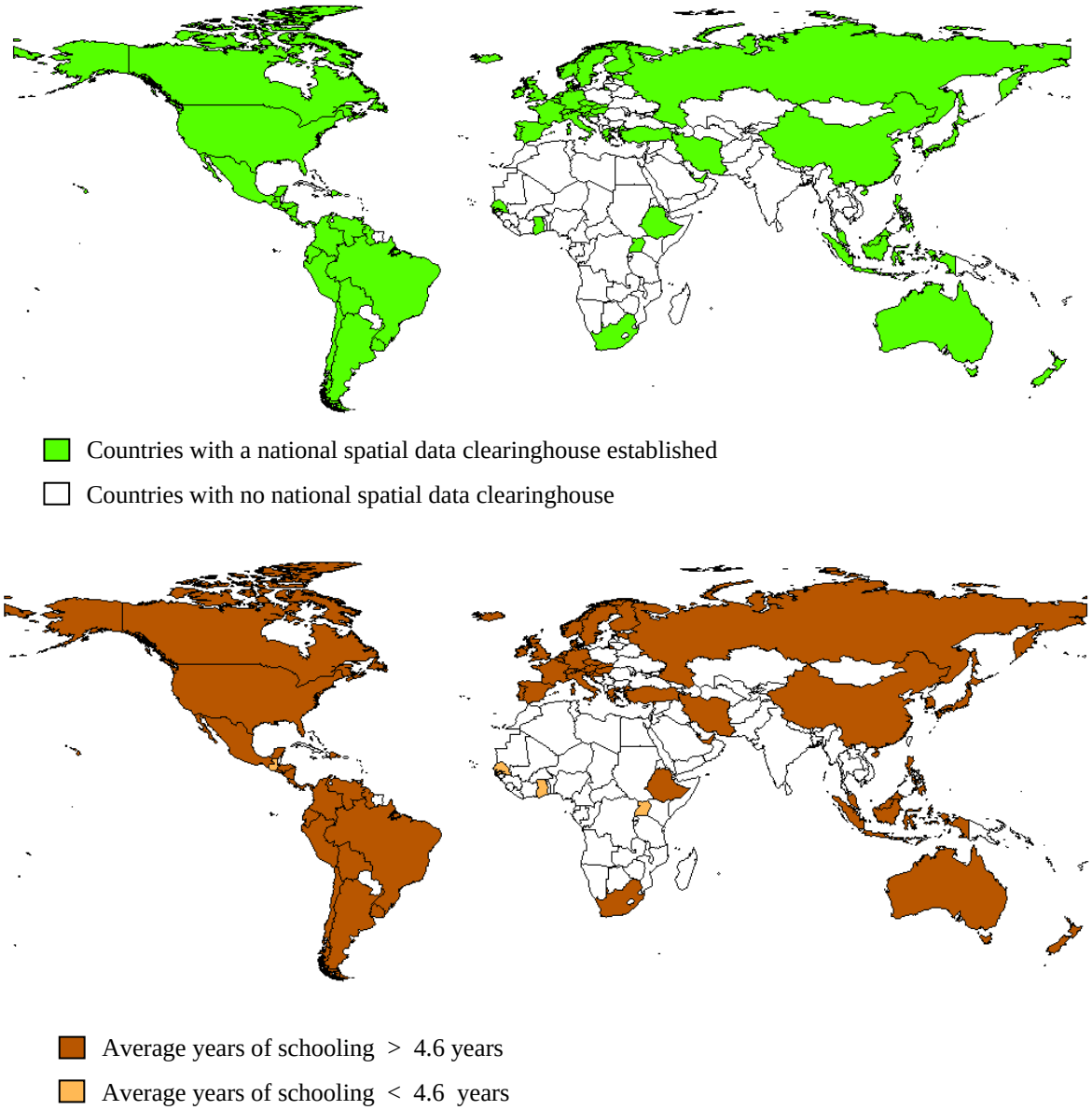


Figure 5.2 Countries with an established national clearinghouse in December 2002 (dark-coloured) compared with average years of schooling. Countries with no established clearinghouse presented in white.

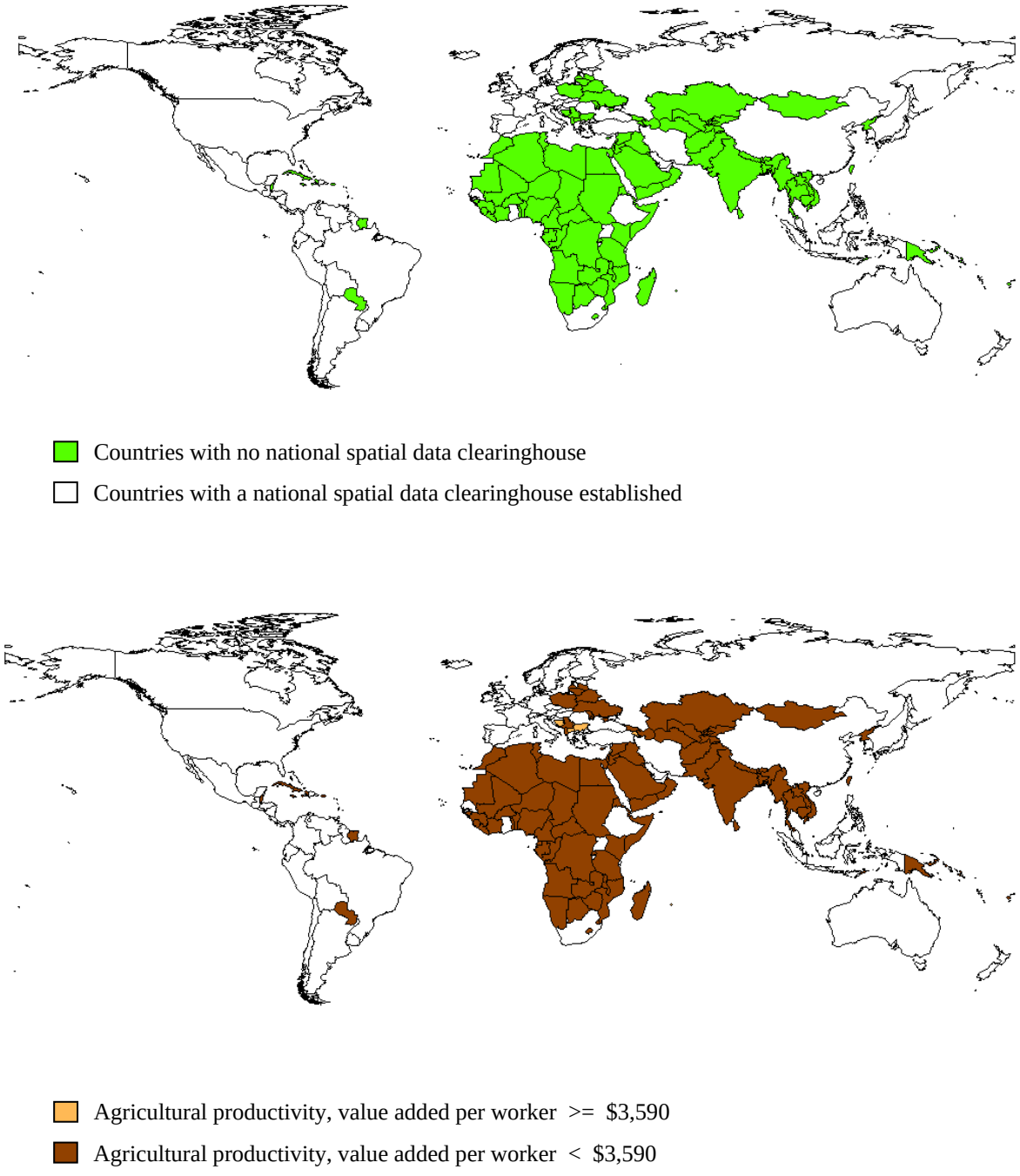


Figure 5.3 Countries with no established national clearinghouse in December 2002 (dark-coloured) compared with agricultural productivity, value added per worker ($). Countries with established clearinghouse presented in white.

5.4 CONCLUSIONS

The purpose of this paper is to explore the questions; what are the critical factors influencing the establishment of national spatial data clearinghouses? Why have some countries established a national clearinghouse, and why have other countries not done so? It appears that all aspects of society could be considered as critical, particularly education, technology, institution and health care. There is however, no single, dominant aspect. This may be due to the strong interdependence of human conditions and activity within society. Consequently, many societal attributes could be critical. Examples of critical attributes collected are: agricultural productivity (value added per worker ($)), average years of schooling, primary pupil / teacher ratio, Internet secure servers, average annual population growth rate (% (2001-2015)), year of independence, and births attended by skilled health staff (% of total). From these attributes, it appears that the standard of living of a national community could be an important factor affecting whether a country establishes a national clearinghouse. Countries whose citizens have a high standard of living are more likely to be in a position to establish national clearinghouses. On the other hand, countries whose citizens are primarily concerned with the daily problems of survival are less likely to be in a position to establish national clearinghouses. Despite these findings, wealth, defined as a large amount of money that a country owns, is not the dominating factor for clearinghouse establishment.

The analysis methods, ANOVA and decision tree method of data mining, appeared to complement each other. Both methods showed similar results at the level of societal aspects, and all attributes obtained from the decision tree for the whole of society appeared also as significant attributes from the ANOVA-analyses.

In order to better identify the societal factors, it is recommended to involve data describing the spatial data community, to implement a comprehensive classification system for the societal attributes into societal aspects, and to use Partial Least Squares Discriminant Analysis (PLS-DA) (Barker and Rayens, 2003). This would make it possible to analyse, more rigorously, the underlying relationships between societal attributes. In addition, complementary case studies would allow for a more in depth analysis of the purpose and scope of specific national clearinghouses.

Moreover, it is strongly recommended to explore the societal factors periodically. Critical (societal) factors may change over time. Exploring them should take their dynamic character into account. It is likely that the critical (societal) factors for the establishment of the first clearinghouses were different from the factors for the situation in December 2002. This is in line with the diffusion of innovation model developed by Rogers (1995). For example, the pioneering countries of national clearinghouses appeared to be in a position to handle the high uncertainties associated with these innovations (Masser, 2005). This is unlikely to have been the case in later years, since the later adopting countries are less likely to have managed such high uncertainty.

ACKNOWLEDGEMENT

The authors gratefully acknowledge the financial support by the Bsik innovation programme ‘Space for Geo-information’.

Chapter 6

Exploring worldwide the impact of society on the success of national spatial data clearinghouses

J. Crompvoets, A. Bregt, M. Wachowicz, G.J. Hofstede, P. Vereijken,
E. de Man

Submitted to International Journal of Applied Earth Observation and Geoinformation

6.1 INTRODUCTION

In April 2005, 83 countries had established a national spatial data clearinghouse on the Internet. In the near future, it is expected that more countries will establish a national clearinghouse. In order to facilitate the access of spatial data and related services, these clearinghouses are regarded as key features of a National Spatial Data Infrastructure (Crompvoets and Bregt, 2003; Crompvoets et al., 2004). However, at present only a few clearinghouses are highly functional, in the sense that they provide efficient facilities for spatial data/service accessibility as well as effective facilities for data/service use and dissemination (Crompvoets et al., 2005).

It appears that socio-technical issues including those which are economic, educational, cultural, institutional, legal, political and organisational, are currently the major impediments to the success of national spatial data clearinghouses, rather than technical issues alone (Bregt, 2000; Groot and McLaughlin, 2000; Williamson et al., 2003; Tait, 2005; Bernard et al., 2005c). Implementing clearinghouses appears to be a complex task, fraught with difficulties in sustaining a shared language, a shared sense of purpose, and reliable financing. Because of these impediments, spatial data accessibility and dissemination in many countries is not optimal.

In order to improve this situation, it is essential that clearinghouse coordinators and practitioners are made more aware of the factors, which could determine the success of a national clearinghouse. These factors could be societal or clearinghouse-internal. Examples of such clearinghouse-internal factors could be: the introduction of web services, stability of funding, and creation of user-friendly interfaces (Crompvoets et al., 2004).

A growing body of literature reflects an increasing interest in societal issues as factors that are critical to successful implementation of national spatial data infrastructure and/or spatial data clearinghouses (Tosta, 1997; Masser, 1999; 2005; Groot and McLaughlin, 2000; De Man, 2000; 2002; Crompvoets and Kossen, 2002; Rajabifard et. al., 2002; Ravi, 2003; Williamson et al., 2003; Spatial Application Division, Catholic University of Leuven, 2003; Craglia et al., 2003; Kok and Van Loenen, 2005; Wehn de Montalvo, 2004; Reece, 2004; Delgado et al., 2005; Bernard et al., 2005b,c; Tait, 2005). However, comprehensive and systematic knowledge in this domain is currently limited.

A web survey focussing on the worldwide status of national clearinghouses was conducted in December 2002. It appeared that 67 countries had established a national clearinghouse on the Internet, thirteen countries had projects to initiate such an electronic facility, and 113 countries had taken no action to establish one (Crompvoets et al., 2004). To the best of the authors' knowledge, no systematic and comprehensive assessment has taken place with regard to the societal impact on the success of national clearinghouses. The purpose of the present paper is to fill this gap. Its main objectives are to analyse societal impact worldwide on the success of national clearinghouses based on the situation in December 2002, and specifically to identify the critical societal factors for success, and to predict clearinghouse success in countries where it has not previously been established.

6.2 METHODOLOGY

This study can be best described as empirical and explorative. The research methodology is presented in Figure 6.1. The main methodological steps are: 1) Indexing suitability, 2) Collecting societal data, 3) Analysing societal factors for success, and 4) Predicting clearinghouse success in countries with no clearinghouse established. Each of these steps is described in more detail.

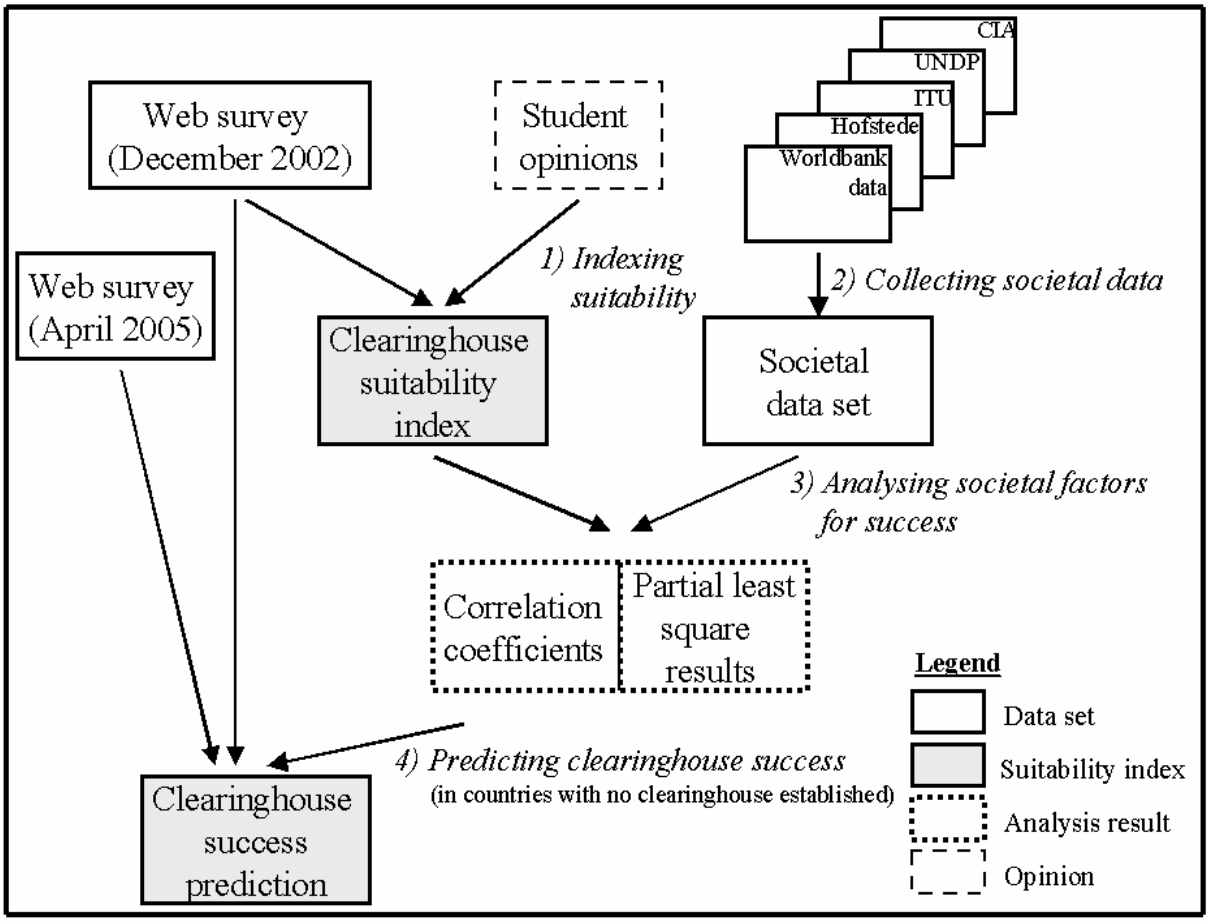


Figure 6.1 Main methodology steps.

6.2.1Indexing suitability

A suitability index for national clearinghouses was developed in order to indicate a measure for the success of a national clearinghouse implementation, using the SDI-component classification: access network, people, data, standards, and policy (Rajabifard et al., 2002; Williamson et al., 2003; Crompvoets et al., 2004). Figure 6.2 presents the steps of this indexing; a) Determining suitability criteria, b) Classifying criteria by SDI-components, c) Weighting SDI-components, d) Weighting (clearinghouse) characteristics by SDI-components, e) Indexing suitability of SDI-components, and f) Indexing clearinghouse suitability.

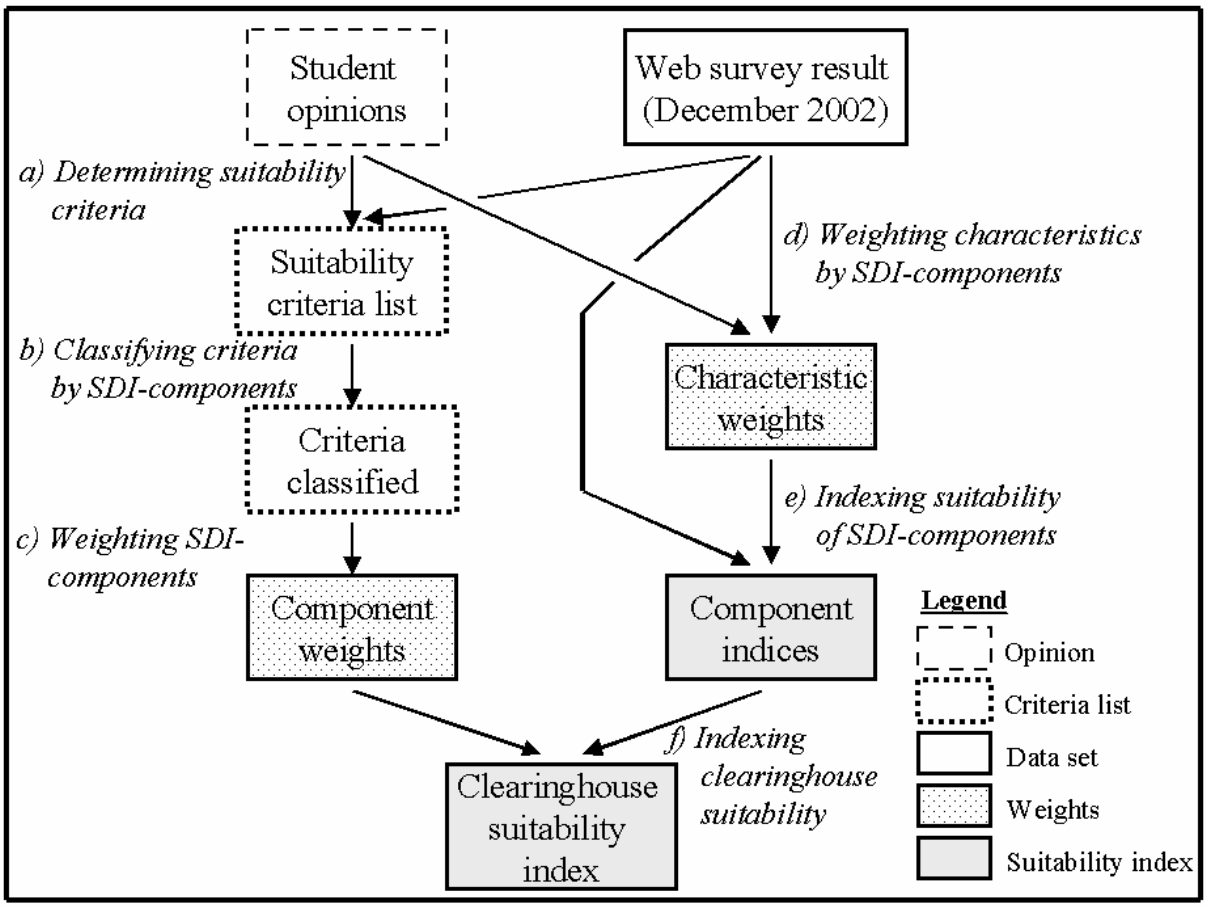


Figure 6.2 Steps for developing Clearinghouse Suitability Index (CSI).

The aim of steps a-c is to determine weights for each of the five SDI-components taking into account their contribution to facilitating access to spatial data and related services.

Around 200 MSc-students (of Wageningen University, Delft University of Technology, Utrecht University, International Institute for Geo-information Science and Earth Observation, Enschede, and Universidad Politécnica de Madrid) were asked to determine the main criteria for assessing the suitability of clearinghouses, in terms of spatial data accessibility and related services. In order to answer this question, a list of 67 national clearinghouses was provided based on the December 2002 web survey (step a). These MSc-students (40 different nationalities) had the added advantage of having no vested interest in the results. The result of this step was a list of suitability criteria. Examples of suitability criteria were: user-friendliness of interfaces and type of search mechanism.

The list of criteria was subsequently classified using the SDI-component classification (step b). For example, the two aforementioned criteria on user-friendliness and search mechanisms were classified as suitability criteria under SDI-component access network. The result of step b therefore was the classification of suitability criteria.

The number of suitability criteria classified under one component was measured as a percentage of all classified criteria (step c). This percentage was changed into a weight value for each SDI-component ranging between 0 and 1. The results were weights for each of the five SDI-components.

The aim of steps d-e is to determine suitability indices of the five SDI-components (for all national clearinghouses).

The MSc-students were also asked to determine weights of nineteen clearinghouse characteristics by SDI-components in order to assess the suitability of each characteristic within a component (step d). These characteristics were measured and recorded during the web survey conducted in December 2002, and classified into components. For more detailed information of the characteristics used, see Crompvoets et al. (2004). The characteristics classified under component data were: number of datasets, level of (meta)data accessibility, and most recently produced dataset. Results were weights between 0 and 100 for each characteristic. The summation of all weights relating to the characteristics that were classified under one component resulted in 100. For example, the weights of characteristics number of datasets, level of (meta)data accessibility, and most recently produced dataset, were respectively 40, 55 and 5. These weights could be considered as maximum values assigned to the most suitable class of each characteristic. Several characteristic classes were distinguished by expert judgement each with a different (suitability) value. In case of characteristic number of datasets, the following classes were distinguished: <100, 100-500, 500-1500, 1500-5000 and >5000 datasets. Since class >5000 datasets was considered to be the most suitable, the maximum value of 40 was assigned. Lower values were assigned (by expert judgement) to other classes, respectively 0, 10, 20 and 30.

As a next step, all data from the original web survey were classified into a characteristic class, and assigned a (suitability) value. All values of characteristics that describe a particular component were added together, resulting in a component suitability index (step e). The range of component index values varies between 0 and 100; 0 meaning the lowest suitability, 100 the highest suitability. Each of the five components was indexed in a similar way.

Finally, the SDI-component weights (step c) were matched with the five component indices (step e), resulting in the Clearinghouse Suitability Index (CSI) (step f). In total, the suitability of all 67 national clearinghouses was indexed.

6.2.2 Collecting societal data

Data were collected for 193 countries, as existing in 2002 according to United Nations. Data was collected from the following sources: The World Bank (2003; 2004), International Telecommunications Union (2003), Central Intelligence Agency (2003; 2004), United Nations Development Program (UNDP) (2003; 2004) and the Dorling Kindersley's Great World Atlas (2002). The data on national culture were taken from Hofstede (2001). All attributes together describe the societal conditions of a country. These attributes collected were classified into nine societal aspects using primarily the different sections of The World Bank's World Development Indicators book and UNDP's Human Development report: 1) Economy, 2) Education, 3) Technology, 4) Environment, 5) Culture, 6) Demography, 7) Institution, 8) Health care and 9) Jurisdiction (including legal and judicial aspects). Each societal aspect describes a particular part of society and consists of a set of underlying societal

attributes. The selection of attributes was based on expert judgement. The main criteria used for selection were:

- Having a potential link with the SDI-components.
- Not being outdated (most data before 2000 was disregarded).
- Not missing too many values.
- Representing one of the nine societal aspects.

Prior to analyses, pre-processing of data took place. This pre-processing included:

- Classifying countries into two establishment classes for a national clearinghouse using the results of the web survey conducted: 1) countries with a national clearinghouse established, and 2) countries with no national clearinghouses established.
- Transforming continuous attributes with boundaries, in order to stretch these boundaries. Attributes with non-negative measurements were logarithmically transformed. Attributes with measurements in a bounded interval were logistically transformed.
- Handling missing values. Countries as well as attributes with more than 50% of missing values were removed. In addition, the missing values of remaining attributes were replaced by the mean or mode of each establishment class; mean for the quantitative attributes, mode for qualitative attributes.

6.2.3 Analysing societal factors for success

The Clearinghouse Suitability Index (CSI) was used as the main indicator for success. In order to analyse comprehensively the societal factors for success, correlation coefficients, analysis of variance (ANOVA) and Partial Least Squares regression (PLS) were used. The analysis was restricted to 60 countries since seven countries (Barbados, Brunei, Dominica, Guyana, Iceland, Luxembourg and Qatar) had too many missing values.

Correlation coefficients between the CSI and the quantitative societal attributes (transformed as well as untransformed) were calculated. These coefficients (r) measure the degree of linear relationship between two variables.

For the qualitative attributes referring to jurisdiction, the ANOVA-method (Snedecor and Cochran, 1980) was used to test differences between CSI-means of several legal systems, e.g. civil and common law (p-value = 0.01).

In this research, numerous societal attributes were collected; many more than the number of countries. It was expected that many societal attributes were strongly related to each other giving multi-collinear data, and that no single attribute alone would make a clearinghouse successful (Nebert, 2004; Reece, 2004). Therefore, the PLS regression method of analysis was chosen (Martens and Næs, 1989; Geladi and Kowalski, 1996; Massart et al., 1997; Tobias, 1997; Kooistra, 2004). This method made it possible to explore relationships between the CSI and attributes of each societal aspect, as well as all attributes of society. PLS is one of the most widely used methods for multivariate analysis applied to a broad range of fields (econometrics, chemistry, education, marketing, and the social sciences). It is the method for constructing predictive models when data are multi-collinear. The strength of PLS is on predicting responses rather than analysing the underlying relationship between variables (Tobias, 1997). PLS is based on latent attribute decomposition using two blocks of variables, matrices X and Y, which contained (transformed) societal data and suitability indices respectively. The objective of the method was to find a small number of latent PLS-factors

that are predictive for Y and use X efficiently. All PLS-analyses were carried out in Matlab™. Some standard techniques were taken from the PLS_Toolbox of Matlab™.

Due to the limited number of existing national clearinghouses, the dataset was not split into a training set and a separate test set. Instead 10-fold cross-validation was used to calibrate and assess the prediction capability of PLS-models.

The parameters used to assess the quality of fitting PLS-models (Kooistra, 2001) to CSIs were:

- The root-mean-square error of calibration (RMSEC), used to estimate the 'average' deviation of the model from the data.
- The percentage variance captured by the regression model (R^2).
- The root-mean-square error of cross-validation (RMSECV) used to measure 'average' predictive error. The optimal number of latent PLS-factors was determined by the lowest value of RMSECV.

6.2.4 Predicting clearinghouse success in countries with no clearinghouse established

The strength of applying PLS-models is that they are predictive. The best PLS-model was applied to predict the CSI in countries where no clearinghouse had been established. The results are an indication of society's capability of establishing successful national clearinghouses. The PLS-model was applied to 110 countries.

For evaluating the quality of prediction results, a web survey was conducted in April 2005. The prediction results (primarily based on data of 2002) were compared with the status of national clearinghouses in April 2005.

6.3 RESULTS AND DISCUSSION

The presentation of the results is organised according to the methodology steps as presented before. The main results are discussed in turn.

6.3.1 Indexing suitability

Through indexing the clearinghouse suitability, an indication of clearinghouse success was achieved in an easy, and transparent way. The main limitation of this indexing is that only MSc-students were approached. Approaching other clearinghouse practitioners could have improved the quality of the index. The main results of this step are the component weights (step c), the suitability indices of the five components (step e) and the CSI (step f). Table 6.1 presents the component weights as well as the percentages of criteria classified for each of the five SDI-components (including examples of criteria determined by the MSc-students). It appears that the components access network, people and data have the highest weights, while standards and policy have the lowest. One reason for these differences could be that the meaning of access network, people and data is much more transparent and visible. This could have made the MSc-students more aware of the need for these components. Alternatively, the suitability of access network, people and data may simply have been easier to assess.

Table 6.1 Component weights and percentages of classified criteria for each SDI-component (including examples of criteria mentioned).

SDI-component	Component weight	% of criteria classified	Examples of criteria mentioned
Access network	0.4	40	Mechanisms for searching, availability of view services, availability of download services, user-friendliness of interfaces, use of maps for searching, type of network architecture
People	0.3	30	Language used, number of visitors, number of suppliers, link to contact webmaster, frequency of web updates
Data	0.2	20	Number of datasets, level of (meta)data accessibility, use of maps for searching
Standards	0.07	7	Type of metadata standard applied
Policy	0.03	3	Registration-only access, payment-only access, funding stability

Table 6.2 presents the suitability indices of the five SDI-components as well as the CSI of each national clearinghouse. Matching the component weights (Table 6.1) with the suitability indices of the five components resulted in the CSI. This index was defined as follows:

CSI = (0.4 * access network) + (0.3 * people) + (0.2 * data) + (0.07 * standards) + (0.03 * policy)

where access network, people, data, standards and policy are the component indices.

From this assessment, it appears that the national clearinghouses of Australia, Canada, Finland, UK and USA were assessed to be the most suitable facilities for accessing spatial data and related services (Table 6.2).

The low values of the mean (40) and median (24) of CSI indicate that the suitability of the clearinghouses in December 2002 was still not high. These low values could be the consequence of the fact that the functional capabilities of clearinghouses did not fit the expectations of the MSc-students, in particular to the access network. MSc-students expected more web services to have been provided and user-friendly interfaces. This is in line with the findings of Crompvoets et al. (2004).

The high diversity in values of the CSI is significant (standard deviation 22). It is important to be aware of the differences between countries like Denmark, Iceland, Ireland, The Netherlands and Uruguay. From a technical (access network) perspective, these clearinghouses are similar (e.g. have similar network architectures and searching mechanisms). However, their suitability was assessed differently due to differences in other components (people, data, standards and policy). Since these components appear to be influenced by prevailing societal conditions, it is argued that the main reason for such CSI diversity is probably due to the fact that these clearinghouses were embedded within different societies, each with unique conditions.

Table 6.2 Suitability indices of the five SDI-components (AN: Access Network; PE: People; DA: Data; ST: Standards; and PO: Policy), and Clearinghouse Suitability Index (CSI) for 67 national clearinghouses.

Country	AN	PE	DA	ST	PO	**CSI**	Country	AN	PE	DA	ST	PO	**CSI**
Argentina	10	63	56	80	100	**43**	Ireland	30	57	44	80	85	**46**
Australia	65	97	100	70	100	**83**	Italy	0	21	13	0	30	**10**
Austria	40	73	31	20	30	**46**	Japan	40	33	44	100	30	**43**
Barbados	10	49	31	80	30	**31**	Luxemburg	29	0	13	0	30	**12**
Belgium	0	60	31	40	30	**28**	Malaysia	65	77	31	70	10	**61**
Bolivia	10	59	31	80	30	**34**	Mexico	40	25	81	80	30	**46**
Brazil	10	67	31	80	30	**37**	Netherlands	30	44	44	100	30	**42**
Brunei	10	59	44	20	30	**33**	N. Zealand	55	85	50	70	10	**63**
Canada	95	94	100	70	85	**94**	Nicaragua	10	59	44	80	30	**37**
Chile	60	43	56	100	30	**56**	Norway	55	51	13	70	10	**45**
China	60	18	31	70	100	**44**	Panama	0	17	13	0	30	**9**
Colombia	40	49	44	50	100	**46**	Peru	10	31	56	80	30	**31**
Costa Rica	10	55	31	80	30	**33**	Philippines	10	59	31	100	30	**36**
Croatia	0	67	50	80	30	**37**	Portugal	30	44	75	100	30	**48**
Czech Rep.	35	44	69	100	100	**51**	Qatar	40	24	13	0	30	**27**
Denmark	30	88	44	100	100	**57**	Russia	0	18	13	50	30	**12**
Dominica	40	61	31	80	85	**49**	Senegal	10	25	31	80	30	**24**
Dom. Rep.	10	53	31	80	30	**33**	Singapore	40	79	13	20	10	**44**
Ecuador	0	33	13	0	30	**13**	Slovak Rep.	0	23	31	80	100	**22**
El Salvador	10	63	44	80	30	**38**	Slovenia	10	72	63	80	30	**45**
Estonia	55	41	13	20	100	**41**	Sth.-Africa	40	87	69	100	30	**64**
Ethiopia	10	57	31	100	30	**35**	Sth.-Korea	0	28	13	70	30	**17**
Finland	70	93	63	50	10	**72**	Spain	30	57	13	30	10	**34**
France	10	41	31	80	30	**29**	Sweden	10	45	44	80	30	**33**
Germany	65	73	31	100	30	**62**	Switzerland	10	69	44	20	30	**36**
Ghana	10	55	31	100	30	**35**	Trinidad	10	49	31	80	30	**31**
Greece	0	16	13	0	30	**8**	Turkey	0	23	13	50	30	**14**
Guatemala	10	27	44	80	30	**27**	Uganda	0	49	13	0	30	**18**
Guyana	0	49	13	0	30	**18**	UK	70	75	56	70	100	**70**
Honduras	10	51	44	80	30	**35**	United Arab Emirates	0	28	31	20	30	**17**
Hungary	10	69	31	50	15	**35**							
Iceland	30	35	31	80	100	**37**	Uruguay	30	69	88	80	80	**58**
Indonesia	60	77	50	50	30	**62**	USA	95	98	100	100	100	**97**
Iran	20	4	63	70	30	**28**	Venezuela	10	61	63	80	30	**41**

6.3.2 Collecting societal data

Through data collected from 234 societal attributes, an indicative description considering the wide scope of society was achieved. Table 6.3 presents the number of attributes collected and some examples of attributes per societal aspect. Appendix 1 shows the full list of attributes collected (including units, attribute scales and sources). From the sources available, it appeared that more economic, technological, environmental and demographic attributes could be collected than cultural, institutional, and legal attributes.

Overall, collecting societal data was negatively affected by limited data availability and reliability (International Telecommunication, 2003; UN Development Programme, 2003; The World bank, 2003; 2004; 2004).

Table 6.3 Number of attributes collected and examples of attributes per societal aspect (between parentheses the units used).

Societal aspect	Number of attributes	Examples of attributes
Economy	77	Taxes on income, profits, capital gains (% of revenue), listed domestic companies, agricultural productivity, value added per worker ($), household final consumption expenditure ($ millions), gross national income (GNI, $ billions), gross domestic product (GDP, $ millions), armed forces personnel (% of labour force), military expenditure (% of GDP), net national savings (% of GNI)
Education	15	Education expenditure (% of GNI), net enrolment ratio (primary % of relevant age group), primary completion rate (% of relevant age group), average years of schooling, education index
Technology	44	Internet hosts per 10,000 inhabitants, Internet users per 10,000 inhabitants, Internet secure servers, personal computers per 1000 people, mobile phone subscribers, annual growth rate (%), air, aircraft departures (thousands), high technology exports ($ millions)
Environment	37	Energy use per capita (kg of oil equivalent), carbon dioxide emissions per capita (metric tons), arable land use (hectares per capita), surface area (thousand km²), irrigated land (% of cropland)
Culture	5	Individualism, uncertainty avoidance, power distance
Demography	26	Population density (people per km²), labour force gender parity index, average annual population growth rate (%, 2001-2015), total fertility rate (births per woman), life expectancy (years)
Institution	6	Institutional investor credit rating, type of government
Health care	13	Health expenditure per capita ($), prevalence of HIV (% of adults), hospital beds per 1000 people
Jurisdiction	11	Type of legal system

No data describing the geo-information (GI) community were collected. Judging by the available literature, it is expected that the GI-community in particular could have a high impact on clearinghouses (Steudler, 2003; Ravi, 2003; Reece, 2004; Delgado et al., 2005). Unfortunately, collecting data referring specifically to this community on a worldwide scale appeared to be impossible.

From the analysis, 23 countries were excluded due to too many missing values.

6.3.3 Analysing societal factors for success

Through the complementary use of correlation coefficients with PLS regression models, an exploration of how society could affect the success of national clearinghouses was undertaken.

Correlation coefficients between the CSI and the 164 (societal) attributes (transformed as well as untransformed) were calculated. The attributes of societal aspect jurisdiction were excluded from these analyses, because most attributes collected were nominal in nature.

Table 6.4 presents only those attributes with coefficient values above 0.40. A classical interpretation of coefficient values is that when these values are ranging between 0.40 and 0.60, they may be regarded as an indication of a moderate degree of correlation, and when higher than 0.60, as an indication for a marked degree of correlation (Franzblau, 1958).

The results presented suggest that several economic, environmental, educational and technological attributes could be important factors affecting success of national clearinghouses success. On the other hand, demographic, institutional and health care attributes are likely to be less important. The attributes with the highest correlation coefficients are related to taxes, energy use, Internet and agriculture. Taxes refer to the main source of revenue for many governments with the potential of influencing incentives and thus economy's competitiveness (The World Bank, 2003). The use of energy refers indirectly to people's wealth even though energy consumption has environmental consequences (e.g. carbon dioxide emissions). The Internet attributes refer to the information and communication technologies that offer vast opportunities for economic growth and better service delivery. The agricultural attributes refer to the agricultural productivity. The relatively high number of cultural attributes indicates that culture could also have an impact on the success of a national clearinghouse, in particular individualism.

Table 6.4 Correlation coefficients between CSI and societal attributes (untransformed as well as transformed) with values higher than 0.400.

Name attribute	Societal aspect	Correlation coefficient Untransformed	Correlation coefficient Transformed
Taxes on income, profits, capital gains (% of revenue)	Economy	0.665	0.490
Energy use per capita (kg of oil equival.)	Environment	0.632	
Internet, hosts per 10,000 inhabitants	Technology	0.622	
Carbon dioxide emissions per capita metric tons	Environment	0.589	
Arable land (hectares per capita)	Environment	0.546	0.419
Personal computers per 1000 people	Technology	0.542	0.404
Education expenditure (% of GNI)	Education	0.540	0.430
Listed domestic companies	Economy	0.526	0.453
Surface area (thousand km²)	Environment	0.525	
Agricultural productivity, value added per worker ($)	Environment	0.516	
Individualism	Culture	0.512	0.479
Household final consumption expenditure ($ millions)	Economy	0.490	
Gross national income ($ billions)	Economy	0.489	
Gross domestic product ($ millions)	Economy	0.488	0.415
Internet, users per 10,000 inhabitants	Technology	0.478	0.453
Internet, secure servers	Technology	0.476	0.517
Air, Aircraft departures (thousands)	Technology	0.469	0.422
Internet, hosts total	Technology	0.453	0.465
Gross national income per capita ($)	Economy	0.443	
Personal computers (thousands)	Technology	0.438	0.408
Internet, users thousands	Technology	0.430	0.405
High technology exports ($ millions)	Technology	0.429	
Uncertainty avoidance	Culture	-0.428	
Power distance	Culture	-0.427	
Purchasing power parity, GNI ($ billions)	Economy	0.426	
Purchasing power parity, per capita ($)	Economy	0.425	
Armed forces personnel (% of labour force)	Economy	-0.417	-0.401
Roads, total road network (km)	Technology	0.415	
Health expenditure per capita ($)	Health care	0.414	
Foreign direct investment ($ millions)	Economy	0.410	
Average years of schooling, total	Education	0.405	
Education index	Education		0.410
Institutional investor crediting rating	Institution		0.400

Table 6.4 shows three attributes with a negative correlation coefficient. It appears that attribute armed forces personnel (% of labour force) could be critical for the failing of clearinghouses. Although national defence is an important function of government (with security from external threats contributing to economic development), high levels of defence spending burden the economy and may impede growth (The World Bank, 2003), and consequently investments in clearinghouses. In addition, cultures with a high power distance and/or strong uncertainty avoidance are likely to have failing national clearinghouses. These (negative) cultural impacts on clearinghouses are supported by similar observations in other literature (e.g. De Man, 2002).

The main result of the application of ANOVA when testing for differences in CSI between several legal systems is that the mean of countries with a common law (52) is significantly higher than the mean of countries with a civil law (35).

PLS-regression models were applied to predict the CSI from the (transformed) societal attributes of each societal aspect. Since they were nominal in nature, attributes of societal aspect jurisdiction were excluded from these PLS-analyses. The percentages of variance captured by the best eight PLS-regression models (R^2) appear to be low for each societal aspect (Table 6.5). This indicates that none of the societal aspects alone could explain the success of a clearinghouse. This is in line with literature on spatial data infrastructures (Groot and McLaughlin, 2000; Williamson et al., 2003; Spatial Application Division, Catholic University of Leuven, 2003; Craglia, 2003; Masser, 2005; Delgado et al., 2005). Societal aspects with relatively high percentages are: economy, education and technology. On the other hand, societal aspects with relatively low values are: demography, institution and health care.

Table 6.5 Number of attributes used, percentages of variance captured by the PLS- models (R^2), optimal number of latent PLS-factors, and attributes with the highest PLS-regression coefficients of each societal aspect.

Societal aspect	Number of attri-butes	R^2	Number of PLS-factors	Attributes with highest PLS-regression coefficients
Economy	48	0.43	2	- Taxes on income, profits, capital gains (% of revenue) - Armed forces personnel (% of labour force) - Listed domestic companies
Education	12	0.42	3	- Net enrolment ratio (primary % of relevant age group) - Primary completion rate (%)
Technology	30	0.39	2	- Internet, secure servers - Mobile phone subscribers, annual growth rate (%)
Environment	29	0.31	1	- Arable land, hectares per capita
Culture	4	0.30	1	- Individualism - Uncertainty avoidance
Demography	26	0.26	2	- Population density (people/km²) - Labour force gender parity index
Institution	4	0.14	1	- Institutional investor crediting rating
Health care	11	0.11	1	- Health expenditure per capita ($)

PLS-regression was also applied to predict the CSI from societal conditions of each country considering all 164 (transformed) attributes, covering the eight societal aspects. The percentage variance captured by the PLS-model (R^2) with the optimal number of three latent PLS-factors is 0.77, which is considerably higher than those from PLS-models of each societal aspect. This suggests that a combination of several attributes from different societal aspects form a better explanation for the success. The attributes with the highest PLS-regression coefficients are mainly economic, educational and technological (Table 6.6). In addition, this R^2-value (0.77) also suggests that the CSI cannot be predicted from societal attributes alone (assuming that the quality of determining CSI is high). The value of RMSEC is 10.5, and that of RMSECV 15.0 (section 6.2.3). These values confirm that this index cannot be fully predicted from societal attributes alone.

Table 6.6 Attributes with the highest PLS-regression coefficients (in order of importance) using the PLS3-model of society considering 164 societal attributes.

Name variable	Societal aspect
Taxes on income, profits, capital gains (% of revenue)	Economy
Net enrolment ratio (primary % of relevant age group)	Education
Armed forces personnel (% of labour force)	Economy
Gross capital formation, average annual growth (%)	Economy
Household final consumption expenditure ($ millions)	Economy
Internet, secure servers	Technology
Education expenditure (% of GNI)	Education
Military expenditure	Economy
Individualism	Culture
Labour force gender parity	Demography
Listed domestic companies	Economy
Gross domestic product, average annual growth (%)	Economy
Arable land (hectares per capita)	Environment
Irrigated land (% of cropland)	Environment
Population density (people/km²)	Demography
Surface area (thousand km²)	Environment
Average annual change in Consumer price index (%)	Economy
Domestic credit to private sector (% of GDP)	Economy
Net national savings (% of GNI)	Economy
Uncertainty avoidance	Culture

Figure 6.3 presents the CSI-values, the predicted values of PLS3-model considering 164 societal attributes, and the PLS3-residual values of each national clearinghouse. From this figure, it appears that the PLS3-model overestimates slightly the values of national clearinghouses with a low CSI and underestimates the values of clearinghouses with a high CSI. This supports the suggestion that the CSI cannot be predicted from societal attributes alone.

A reason that societal conditions might not have a full impact on clearinghouse success could be due to the fact that critical clearinghouse-internal factors also need consideration. Regarding Figure 6.3, the underestimation of the PLS3-model for national clearinghouses with a high CSI is likely to be caused by to the implementation of clearinghouse-internal success factors. On the other hand, the slight overestimation of the clearinghouses with a low CSI could be due to the absence or partial implementation of these success factors.

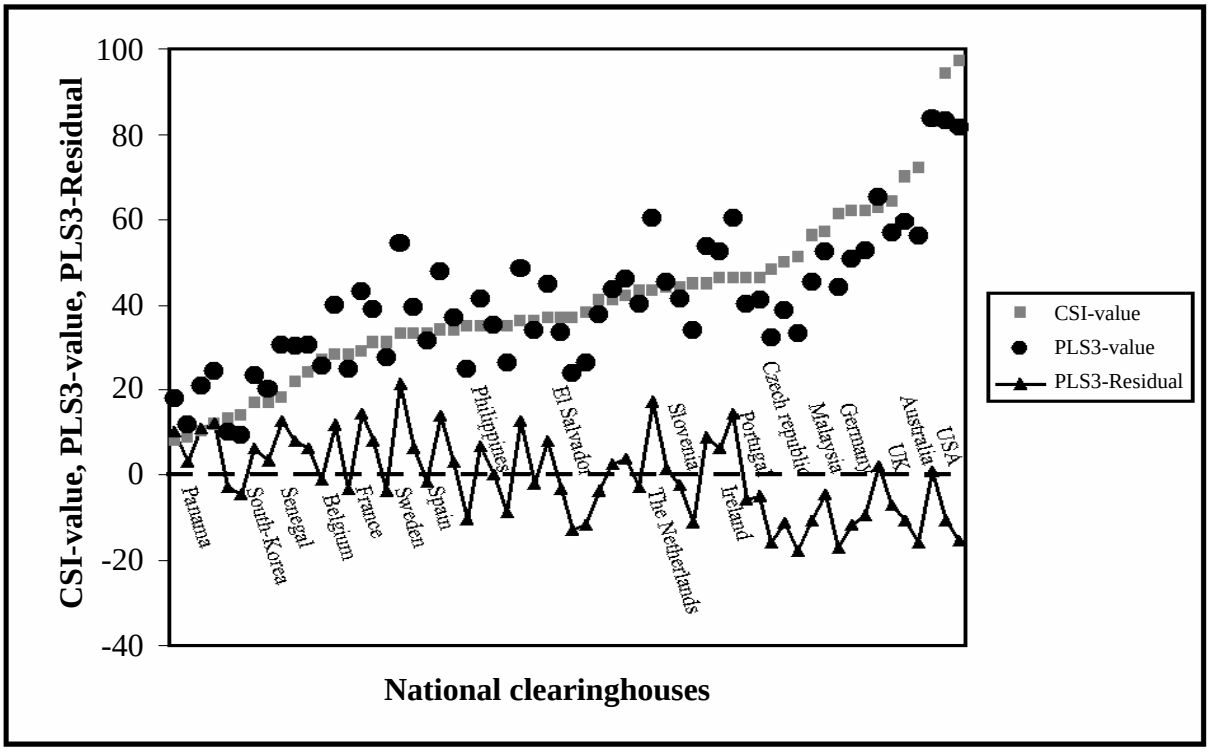


Figure 6.3 Clearinghouse Suitability Index value (CSI-value), PLS3-value and PLS3-residual values of each national clearinghouse. Samples ordered by increasing CSI-values.

Another PLS-regression model was applied to predict the CSI from societal conditions taking only into account the attributes of aspects economy, education and technology (90 attributes). The percentage variance captured by this PLS-model (R^2) with 3 PLS-factors is 0.70. The RMSEC and RMSECV values of this PLS-model are slightly worse than the one considering all 164 societal attributes, respectively 11.5 and 16.0. These results suggest that a combination of economic, educational and technological attributes only, predicts the CSI reasonably well.

Finally, a PLS-regression model was applied to predict the CSI using the latent PLS-factors of the eight societal aspects. In total, the thirteen latent PLS-factors presented in Table 6.5 were used. The percentage variance captured by this PLS-model (R^2) with three latent PLS-factors is 0.67. The RMSEC and RMSECV values of this PLS-model are similar to the PLS-model considering all 164 societal attributes, respectively 11.0 and 15.0. The latent PLS-factors with the highest PLS-regression coefficient values are (in order of importance): economic, technological and educational. This result seems to confirm the findings that the success of national clearinghouses is likely to be influenced by a combination of several factors from different societal aspects, in particular economy, technology and education. On the other hand, the latent PLS-factors with very low values are: institutional and health care.

Examples of economic, educational and technological attributes that could be critical for clearinghouse success are: taxes on income, profits and capital gains (% of revenue), listed domestic companies, gross domestic product ($ billions), education expenditure (% of GNI) and Internet secure servers (Table 6.4 and Table 6.6). In addition, arable land per capita (environment) as well as individualism (culture) could also be critical.

The suggestion of these explorative analyses is that wealthier countries, where communities are comprised of calculating citizens focussed on self-interest, are more likely to have a successful clearinghouse. These countries have a sound investment climate with good macro-economic management, trade and investment policies that promote openness, and high quality (technological) infrastructure and services. In addition, they have good education that enables people to take advantage of new opportunities. Moreover, they have supportive legal and regulatory systems (e.g. common law) that support the day-to-day operations of governments and firms by protecting property rights, promoting access to credit, and ensuring tax, customs, and judicial services. Finally, these countries still have the opportunity to increase their agricultural productivity by expanding arable land or by using their land more intensively. Investing in successful national clearinghouses could improve the spatial data/service accessibility, as well as data/service dissemination, in order to raise the level of wealth or respond to the demand of citizens. But whether the successful implementation of a national clearinghouse indeed contributes to increasing wealth – in other words, the causal direction of the association – is not ascertained by the present exploration

6.3.4 Predicting clearinghouse success in countries with no clearinghouse established

Through the application of the PLS3-model, which takes all 164 societal attributes into consideration, the success of establishing a clearinghouse could be predicted in countries where it has not previously been established. Countries with PLS-values that were predicted to be higher than the mean of CSI-values determined (40) were regarded as countries that could potentially have a successful clearinghouse. As a result of the web survey conducted in April 2005, it appeared that 83 countries had a national clearinghouse. This means that 16 countries established a national clearinghouse between December 2002 and April 2005. In addition, 25 countries set up projects to establish a national clearinghouse in the same period. Table 6.7 presents only the nine countries that were predicted to be successful, and includes their predicted PLS-values as well as their status in April 2005. It appears that five of the nine countries actually established a national clearinghouse, and four countries initiated projects to set up a national clearinghouse between December 2002 and April 2005. The PLS-model did not predict any clearinghouse success in countries that had not initiated any steps in April 2005! In addition, it predicted no clearinghouse success for 101 countries. This could indicate that new national clearinghouses established in the near future may eventually fail, unless the clearinghouse-internal factors for success are implemented. Between December 2002 and April 2005, eleven countries not predicted to be successful had established a national clearinghouse, and 21 countries had started projects. Table 6.8 presents the 32 countries that were not predicted to be successful, but which had established a national clearinghouse or set up projects in April 2005. The predicted PLS-values as well as their status in 2005 are also included in the table. In addition, 69 countries that were not predicted to be successful had taken no national clearinghouse initiative in April 2005.

The strength of the PLS-model appears to be in predicting potential clearinghouse success in countries that have no clearinghouse. It also indicates where, in addition to societal conditions, clearinghouse-internal factors need to be implemented to compensate for the lack of societal support.

Table 6.7 Countries with predicted clearinghouse success, their predicted PLS-value, and status in April 2005.

Country	Predicted PLS-value	National clearinghouse status 2005
Botswana	46.3	Established
India	49.9	Established
Israel	43.1	Established
Mali	41.2	Project
Namibia	54.5	Established
Paraguay	40.2	Project
Poland	42.9	Project
Saudi Arabia	40.7	Project
Thailand	40.7	Established

Table 6.8 Countries with no predicted clearinghouse success that had established a national clearinghouse or set up projects for clearinghouse establishment in April 2005, their predicted PLS-value, and their status in April 2005.

Country	Predicted CSI-value	National clearinghouse Status 2005	Country	Predicted CSI-value	National clearinghouse Status 2005
Afghanistan	12.0	Established	Madagascar	20.6	Established
Belarus	4.6	Established	Malawi	29.5	Established
Belize	30.1	Project	Mozambique	22.7	Project
Benin	28.8	Project	Nepal	23.8	Established
Bulgaria	7.1	Project	Nigeria	21.7	Project
Burkina Faso	33.6	Established	Oman	21.5	Established
Cambodia	29.4	Established	Pakistan	16.7	Project
Cuba	31.8	Project	Romania	12.8	Project
Cyprus	27.6	Project	Seychelles	22.7	Project
Egypt	18.7	Established	Swaziland	27.1	Project
Jamaica	22.1	Project	Tanzania	32.3	Project
Kenya	34.5	Project	Togo	29.9	Established
Kuwait	17.9	Project	Vietnam	26.3	Project
Latvia	28.0	Project	Yemen	22.1	Established
Lesotho	33.9	Project	Zambia	35.0	Project
Lithuania	34.9	Project	Zimbabwe	25.2	Project

6.4 CONCLUSIONS

The findings of these explorative analyses highlight the importance of societal conditions on the success of national spatial data clearinghouses. However, society alone does not fully explain the success of national clearinghouses, since clearinghouse-internal factors are also an important consideration. It is likely that the success of national clearinghouses depends on a combination of critical attributes from different societal aspects, in particular economy, education and technology. Critical factors for clearinghouse success are likely to be those that raise the level of wealth in the respective country (e.g. gross domestic product, education expenditure, Internet servers, and taxes on income, profits, and capital gains). In this exploration, it appeared to be difficult to identify the critical factors due to the complex interaction of societal (and clearinghouse-internal) factors. Subsequent research may help us understand the causal hierarchies between these societal factors. To conclude, the results of

the present exploration suggest that there is no simple solution or uniform approach for setting up a successful clearinghouse. However, being aware of the most significant societal factors for success as detailed in this paper can enhance (future) clearinghouse coordinators' and practitioners' understanding of national clearinghouse advancement.

The strength of applying the PLS-model appears to be the prediction of potential clearinghouse success in countries where it has not previously been established. In the near future, newly established national clearinghouses are likely to face serious obstacles, if not failure, due to unfavourable societal conditions. This could however be reversed by additional efforts to implement the critical clearinghouse-internal factors.

ACKNOWLEDGEMENT

The authors would like to thank dr. Wies Akkermans, dr. Cajo ter Braak, dr. Lammert Kooistra of Wageningen UR, ir. Harm Kossen and Judith Uwimana MSc. for comments and suggestions on earlier versions of this paper. The authors also acknowledge the support of the MSc-students of Wageningen University, Delft University of Technology, Utrecht University, International Institute for Geo-information Science and Earth Observation, Enschede (The Netherlands) and Universidad Politécnica de Madrid (Spain) in the preparation of the CSI. Their contributions were crucial for the results. The authors also acknowledge the financial support by the Bsik innovation programme 'Space for Geo-Information'

Chapter 7

Main conclusions and recommendations for further research

7.1 MAIN CONCLUSIONS

Many countries have invested resources in developing national spatial data clearinghouses. However, only a few of these (Australia, Canada, USA) are highly functional in the sense that they provide efficient facilities for spatial data/service accessibility, as well as effective facilities for data/service dissemination to the GI-community. Consequently, the use of spatial data assets of many countries is still not optimal. In order to improve this situation coordinators as well as policy makers involved in national clearinghouse development need to be supported in their strategies for establishing and maintaining successful facilities. This support can be effected by identifying critical factors behind the success (or failure) of a national clearinghouse, and by assessing the impact of a national clearinghouse in terms of relevance, efficiency and effectiveness. Identification of critical factors can be accomplished by analysing longitudinal developments of national clearinghouses and the societal impact on clearinghouses.

The main objective of this thesis is to analyse the development and impact of national spatial data clearinghouses worldwide as well as the impact of society on these facilities. The results obtained provide support for the development of strategies for establishing and maintaining clearinghouses. To achieve this main objective, five sub-objectives were formulated. The main conclusions relating to these sub-objectives, as well as recommendations for further research, are presented in this chapter.

1) To assess the worldwide status of national spatial data clearinghouses.

If judged only by the number and spatial distribution of national clearinghouses established, national clearinghouses can be considered a worldwide success. To date, 83 countries have an implemented version posted on the World Wide Web, and 25 countries have initiated projects for implementation. National clearinghouse implementation is a worldwide activity. However, numbers vary considerably between regions. More than 60% of the countries in Europe and America have established a national clearinghouse, whereas in Africa this is less than 20%.

There is significant diversity among national clearinghouses in use, content, technology, and management, which is largely due to the societal conditions of each nation.

2) To assess the worldwide developments of national spatial data clearinghouses.

Since 1994 the number of national clearinghouses has been steadily increasing. In the near future, more countries are likely to establish their own national clearinghouse.

This increase in the number of implementations contrasts with the slight decline in the use, management and content of national clearinghouses. Downward trends observed between 2000 and 2002, continue in 2005. Possible reasons for these trends are the dissatisfaction of the spatial data community with the functional capability of clearinghouses and the piecemeal funding of the majority of clearinghouses, which means that no appropriate long-term framework can be built to facilitate the management of information assets.

The main critical (clearinghouse-internal) factors identified for the success of national clearinghouse development are: introducing web services to clearinghouses (e.g. view services), continuous funding, having a clear vision of the clearinghouse function, providing good communication channels, building user-friendly interfaces with clear terminology, creating trust in the management environment, motivating data suppliers and web service providers to participate within the clearinghouse, and ensuring that a motivating environment is created so that clearinghouse coordinators update the clearinghouse more regularly. Addressing these clearinghouse-internal factors will go a long way towards meeting the immediate needs of the current users.

In the near future it is expected that national clearinghouses will provide more web services. These services could increase spatial data consumption and also widen the range of users. These expectations are in line with current literature (Maguire and Longley, 2005, Beaumont et al., 2005, Foust et al., 2005)

3) To assess the impact of national spatial data clearinghouses on society in particular the GI-community.

The impact of national clearinghouses is mainly positive. This impact is mainly economic in character, but social impact is obviously important as well. National clearinghouses are relevant facilities for enhancing spatial data accessibility, for providing an efficient means of accessing spatial data, and the effective promotion of data use and distribution. The results obtained from this impact assessment justify ongoing investment in establishing and maintaining national clearinghouses.

4) To explore the societal impact on the establishment of national spatial data clearinghouses.

Societal conditions of a country have a strong impact on the establishment of national clearinghouses. It is likely that the standard of living of a national community is a critical factor affecting whether a country has established a national clearinghouse or not. This standard of living encompasses all aspects of community life, which influence the physical and mental health of its members. Countries with a high standard of living will probably be able to establish national clearinghouses since they are not concerned with daily problems of survival. In countries where they have not previously been established, the newly established national clearinghouses are likely to face serious obstacles in the near future (e.g. cost and funding, institutional problems, lack of specialised human resources) or may even fail due to unfavourable societal conditions. This outcome could, however, be reversed if critical clearinghouse-internal factors, as presented above, are implemented.

5) To explore the societal impact on the success of national spatial data clearinghouses.

Societal conditions also affect the success of national clearinghouses. However, societal factors alone do not fully determine the success, as clearinghouse-internal success factors may also have a strong impact. Many societal factors for success are likely to be critical, in particular those that could raise the level of wealth in the country (e.g. gross domestic product, education expenditure, Internet accessibility, and taxes on income, profits and capital gains).

To conclude, when considering the strategies for the establishment and maintenance of the national clearinghouses, no single best solution or recipe exists since each country has its own unique society. However, the critical factors identified as well as the result of the impact assessment will support clearinghouse coordinators and policy makers in the development of successful strategies for establishing and maintaining national clearinghouses. In this way, the results of this thesis can contribute to the enhancement of national clearinghouses and NSDIs in many countries.

7.2 RECOMMENDATIONS FOR FUTURE RESEARCH

The following main recommendations for further research can be made:

- Focus on Spatial Data Infrastructures (SDIs). This thesis focuses on clearinghouses, because they are measurable, and easily identified. Given the extent to which clearinghouses are expected to change over time it will be necessary to focus strongly on the interdependence between SDI-components (Williamson et al., 2003; Masser, 2005). In future, clearinghouses should not be analysed as single entities, but treated as integrated facilities within SDIs. Therefore, SDIs are a better level of analysis for future studies than clearinghouses.
- Implement case studies that focus on single national clearinghouses. The scope of this thesis was on national clearinghouses worldwide. Consequently, many results were aggregated to world level. Complementary case studies could analyse more thoroughly business requirements and driving forces that have shaped the purpose, scope, design, implementation and technical aspects of single national clearinghouses. An appreciation of these business requirements/driving forces could improve the explanation of the trends and impacts presented.
- Analyse clearinghouses that are not national. This thesis focuses on clearinghouses at national level. However, clearinghouses at local, state, international and even global level exist. Coordinators and policy makers of these other clearinghouses also need support in the development of their strategies. Results obtained from the analyses could provide this support. Moreover, the results could also be used to improve the identification of critical factors affecting national clearinghouses.
- Investigate the usability of spatial data and web services. Based on people's demands of second-generation SDI, the usability of spatial data and web services should be investigated to improve the functional capability of clearinghouses.

References

Agresti, A., 1990. Categorical Data Analysis. Wiley.

Askew, D., Evans, S., Matthews, R. and Swanton, P., 2005. MAGIC: a geoportal for the English countryside. Computers. Environment and Urban Systems, 29: 71-85.

Australian and New Zealand Land Information Council (ANZLIC), 2000. Outcomes of the ANZLIC Clearinghouse Workshop. Adelaide.

AUSLIG, 2001. Australian Spatial Data Infrastructure. Australian Land Information Group.

Barker, M. and Rayens, W., 2003. Partial Least Squares for Discrimination. J. Chemometrics, 17(3):166-173.

Beaumont, P., Longley, P.A. and Maguire, D.J., 2005. Geographic information portals – a UK perspective. Computers. Environment and Urban Systems, 29: 49-69.

Berends, J. and Weesie, E., 2001. Kosten en effecten Ruimte voor Geo-informatie. Rotterdam: Twynstra Gudde & NEI (in Dutch).

Bernabé, M.A., Gould, M., Muro-Medrano, P.R., Nogueras, J. and Zarazaga-Soria, F.J., 2002. A Spatial Data Catalogue Based Initiative to Launch the Spanish SDI. Proceedings of 6th Global Spatial Data Infrastructure Conference, From Global to Local. Budapest, Hungary.

Bernard, L., Fitzke, J. and Wagner, R.M. (Editors), 2005a. Geodaten-infrastruktur, Grundlagen und Anwendungen, Herbert Wichmann Verlag, Heidelberg, Germany (in German).

Bernard, L., Crompvoets, J. and Fitzke, J., 2005b. Geodateninfrastrukturen – ein Überblick. In: L. Bernard, J. Fitzke and R.M. Wagner (Editors), Geodateninfrastruktur, Grundlagen und Anwendungen, Herbert Wichmann Verlag, Heidelberg, Germany, pp. 4-8 (in German).

Bernard, L., Kanellopoulos, I., Annoni, A. and Smits, P., 2005c. The European geoportal –one step towards the establishment of a European Spatial Data Infrastructure. Computers, Environment and Urban Systems, 29: 15-31.

Bishr, Y. and Radwan, M., 2000. GDI Architectures. In: R. Groot and J. McLaughlin (Editors), Geospatial Data Infrastructure: Concepts, Cases and Good Practice. Oxford University Press, Oxford, UK, pp. 135–150.

Bogaerts, Th.J.M., Aalders, H.J.G.L. and Gazdzicki, 1997. Components of Geo-Information Infrastructure. Proceedings ELIS'97, Prague.

Bråthen, S., 2003. Economic Impact Assessment: Description.

Bregt, A., 2000. The Dutch clearinghouse for geospatial information: cornerstone of the national geospatial data infrastructure. In: R. Groot and J. McLaughlin (Editors), Geospatial Data Infrastructure: Concepts, Cases and Good Practice. Oxford University Press, Oxford, UK, pp. 262-267.

Canadian Council of Land Surveyors, Canadian Institute of Geomatics and Geomatics Industry Association of Canada. 2000. Geomatics Sector - Human resources study.

CEN/TC287, 1996. Geographic information-data description-metadata, Technical Report, prEN12657. CEN, Brussels.

Centers for disease control and prevention of US Department of Health and Human Services, 2005. Healthy places terminology.

Central Intelligence Agency, 2003, 2004. The World Factbook.

Clinton, W., 1994. Coordinating geographic data acquisition and access to the National Spatial Data Infrastructure. Executive Order 12096, Federal Register 59, 17671-4 Washington, DC, US.

Coleman, D.J. and McLaughlin, J., 1998. Defining Global Geospatial Data Infrastructure (GGDI): components, stakeholders and interfaces. Geomatica, 52(2): 129-144.

Commission of the European Communities, 1999. A European System of Environmental Pressure Indices, First Volume of the Environmental Pressure Indices Handbook: The Indicators, Part I: Introduction to the political and theoretical background. Joint Research Centre, Institute for Systems, Informatics and Safety.

Commission of the European Communities, 2002. Communication from the Commission on Impact Assessment. Brussels.

Commission of the European Communities, 2004. Proposal for a Directive of the European Parliament and the Council establishing an infrastructure for spatial information in the Community (INSPIRE), Extended Impact Assessment.
Commission Staff Working Document. Brussels.

Craglia, M., Annoni, A., Klopfer, M., Corbin, C., Hecht, L., Pichler, G. and Smits, P. (Editors), 2003. Geographic Information in the Wider Europe. Geographic Information Network In Europe (GINIE).

Craig, C.S., Douglas, S.P. and Grein, A., 1992. Patterns of convergence and Divergence among Industrialized Nations: 1960-1988. Journal of International studies 23: 773-785.

Crompvoets, J., 2002. Developments of national clearinghouses for geo-information. Proceedings of 6th Global Spatial Data Infrastructure Conference, From Global to Local, Budapest, Hungary.

Crompvoets, J. and Kossen. H., 2002. The impact of culture on national spatial data clearinghouses. Proceedings of GISDECO-conference, Governance and the use of GIS in Developing Countries. Enschede, The Netherlands, pp. 9.1–9.3.

Crompvoets, J. and Bregt, A., 2003. World status of National Spatial Data Clearinghouses. URISA Journal, Special Issue on Public Participation GIS, 15, APA I: 43-50.

Crompvoets, J., Bregt, A., Rajabifard, A. and Williamson, I., 2004. Assessing the worldwide developments of national spatial data clearinghouses. International Journal of Geographical Information Science, 18(7): 665-689.

Crompvoets, J., Bregt, A. and van Adrichem, M., 2005. Disappointing NSDI National Clearinghouse Survey, International Developments, Status, Suitability and Spatial Distribution. GIM International, 19 (9): 35-37.

Delgado Fernández, T., Lance, K., Buck, M. and Onsrud, H.J., 2005. Assessing an SDI Readiness Index. Proceedings From Pharaohs to Geoinformatics, FIG Working Week 2005 and 8th International Conference on Global Spatial Data Infrastructure, Egypt, Cairo, April 2005.

De Man, W.H.E., 2000. Institutionalization of geographic information technologies: unifying concept? Cartogr. Geogr. Inf. Sci. 27(2): 139-151.

De Man, W.H.E. and van den Toorn, W.H., 2002. Culture and the adoption and use of GIS within organisations. International Journal of Applied Earth Observation and Geoinformation, 4: 51-63.

Dilly, R., 1995. Data Mining, An Introduction, Student Notes, Queens University Belfast, UK.

Dorling Kindersley, 1997. World atlas. Dorling Kindersley Limited, London, UK.

Dorling Kindersley, 2002. Great World Atlas. Dorling Kindersley Limited.
Environmental Protection Agency, 2000. Guidelines for preparing economic analysis.
Feeney, M-E. F., 2003. SDIs and Decision Support. In: I. Williamson, A, Rajabifard and M.E. Feeney (Editors), Developing Spatial Data Infrastructures: From Concept to Reality. Taylor & Francis, London, UK, pp. 195-210.
FGDC, 1995. Content Standards for Digital Geospatial Metadata Workbook, Federal Geographic Data Committee, Washington DC, USA.
FGDC, 1997. Metadata to Clearinghouse Hands-on Tutorial. Federal Geographic Data Committee, Washington DC, US.
FGDC, 1998. Content Standards for Digital Geospatial Metadata Workbook, Version 2.0, Federal Geographic Data Committee, Washington DC, USA.
FGDC, 2000. Questions and Answers about clearinghouses. Federal Geographic Data Committee, Washington DC, USA.
FGDC (Metadata staff coordinator), 2001. Proposal for a National Spatial Data Infrastructure Standards Project.
FGDC, 2002. Geospatial One-Stop; Office of Management and Budget Capital Asset Plan (Exhibit 300).Version 3. Federal Geographic Data Committee, Washington DC, USA.
Fornefeld, M. and Oefinger, P., 2001. Marktstudie: Aktivierung des Geodatenmarktes in Nordrhein Westfalen. market study, MICUS GmbH, Germany (in German).
Foust, J., Tang, W.S.M., Selwood, J., 2005. Evolving infrastructure: Growth and Evaluation of Spatial Portals. Proceedings From Pharaohs to Geoinformatics, FIG Working Week 2005 and 8th International Conference on Global Spatial Data Infrastructure, Egypt, Cairo, April 2005.
Franzblau, A., 1958. A Primer of Statistics for Non-Statisticians. Harcourt, Brace & World, New York.
Geladi, P. and Kowalski, B.R., 1986. Partial least-squares regression: a tutorial. Analytica Chimica Acta, 185: 1-17.
Gillespie, S. (Editor), 2000. Determining, Measuring, and Analyzing the Benefits of GIS. Urban and Regional Information Systems Association.
Gore, A., 1998. The Digital Earth: understanding our planet in the 21st century, The Australian Surveyor, 43(2), 89-91.
Groot, R. and Sharifi, M.A., 1994. Spatial data infrastructure, essential element in the successful exploitation of GIS technology. EGIS/MARI '94, Fifth European Conference and Exhibition on Geographical Information Systems. Utrecht, The Netherlands.
Groot, R. and McLaughlin, J. (Editors), 2000. Geospatial Data Infrastructure: Concepts, Cases and Good Practice. Oxford University Press, Oxford, UK.
Hamilton, M., 2003. Online Survey Response Rates.
Henderson, H., Lickerman, J. and Flynn, P. (Editors), 2000. Calvert-Henderson quality of Life indicators: a new tool for assessing national trends. Calvert group.
Hofstede, G., 2001. Culture's consequences, Comparing Values, Behaviors, Institutions, and Organisations Across Nations. Sage Publications, Thousand Oaks.
Hofstede, G.J., 2000. The Information Age across cultures. Actes du 5ème colloque AIM, Montpellier, France.
Hutton, 1990. Survey Research for Managers: How to Use Surveys in Management Decision-making, 2nd edition. Basingstoke: Macmillan.

INSPIRE Architecture and Standards working group (Editor P. Smits), 2002. INSPIRE Architecture and Standards Position Paper. Infrastructure for Spatial Information in Europe. Commission of the European Communities.

International Telecommunications Union, 2003. World Telecommunications Indicators. International Telecommunication Union, Geneva.

ISO/TC-211, 2001. Geographic Information: Metadata, International Standard 19115.

Jorgenson, D.W., 1998. Growth, Volume 2: Energy, Environment, and Economic Growth. Cambridge, MA:MIT Press.

Kok, B. and van Loenen, B., 2005. How to assess the success of National Spatial Data Infrastructures? Computers, Environment and Urban Systems, 29: 699-717.

Kooistra, L., 2004. Incorporating spatial variability in ecological risk assessment of contaminated river floodplains, PhD-thesis. Catholic University Nijmegen, The Netherlands.

Kooistra, L., Wehrens, R., Leuven, R.S.E.W. and Buydens, L.M.C., 2001. Possibilities of VNIR spectroscopy for the assessment of soil contamination in river floodplains. Analytica Chimica Acta, 446: 97-105.

Lance, K. and Hyman, G., 2001. Adoption and implementation of national spatial data infrastructure in Latin America and the Caribbean. 5th Global Spatial Data Infrastructure Conference, Cartagena de Indias, Colombia,

Long, M. and Alastair, I., 1997. Assessing Climate Change Impacts: Co-evolution of Knowledge, Communities, and Methodologies. BCSIA Discussion Paper, Discussion Paper E-97-09, Cambridge, MA: Kennedy School of Government, Harvard University.

Longley, P.A., Goodchild, M.F., Maguire D.J. and Rhind, D.W. (Editors), 1999. Geographic Information Systems, Second Edition. John Wiley & Sons, New York.

MacEachren, A.M., Wachowicz, M., Haug, D., Edsall, R. and Masters, R., 1999. Constructing Knowledge from Multivariate Spatiotemporal Data: Integrating Geographic Visualization with Knowledge Discovery in Database Methods. International Journal of Geographic Information Science, 13: 311-334.

Maguire D.J. and Longley, P.A., 2005. The emergence of geoportals and their role in spatial data infrastructures. Computers, Environment and Urban Systems, 29: 3-14.

Martens, H. and Næs, T., 1989. Multivariate calibration. John Wiley & Sons, Chichester, UK.

Massart, D.L., Vandeginste, B.G.M., Buydensm L.M.C., De Jong, S., Lewi, P.J. and Smeyers-Verbecke, J., 1997. Handbook of Chemometrics and Qualimetrics: Part A. Elsevier Science, Amsterdam.

Masser, I., 1999. All shapes and sizes: the first generation of National Spatial Data Infrastructures. International Journal of Geographical Information Science, 13: 67-84.

Masser, I., 2005. GIS Worlds, Creating Spatial Data Infrastructures. ESRI Press, Redlands, California.

McLaughlin, J., 1991. Towards National Spatial Data Infrastructure. Proceedings of the 1991 Canadian Conference on GIS. Canadian Institute of Geomatics, Ottawa, Canada, pp. 1-5.

Moriarty, P., 1996. “Can Urban Density Explain Personal Travel Levels?”. Urban Policy and Research, 14(2): 109-117.

Morales, J., 2004. Model-driven Design of Geo-information Services. PhD-thesis, ITC Enschede.

Murthy, S.K., 1998. Automatic Construction of Decision Trees from Data: A Multi-Disciplinary Survey. Data Mining and Knowledge Discovery, 2(4): 345-389.

Nebert, D.D. (Editor), 2001. Developing Spatial Data Infrastructures: The SDI Cookbook, Version 1.0. GSDI-Technical Working Group.

Nebert, D.D. (Editor), 2004. Developing Spatial Data Infrastructures: The SDI Cookbook, Version 2.0. GSDI-Technical Working Group.

Nedovic, Z. and Pinto, K., 2000. Information sharing in an interorganisational GIS Environment. Environment and Planning B: Planning and Design, 27 (3): 455-474.

Nogueras Iso, J., 2004. Contributions to the problem of collections, interoperability and information retrieval of geographic information metadata for spatial data infrastructures. PhD-thesis, University of Zaragoza, Spain.

Noori-Bushehri, S. and Rajabifard, A., 2001. A Proposal for a Workshop on APSDI-Clearinghouse. PCGIAP-WG2, 7th PCGIAP Meeting, Tsukuba, Japan.

Onsrud, H.J., 1998. Compiled responses by question for selected questions, Survey of national and regional spatial data infrastructure activities around the globe. Global Spatial Data Infrastructure.

Østensen, O.M., 2001. The expanding agenda of geographic information standards. ISO Bulletin, pp. 16-21.

Pasca, M., Artioli, G., Ciardi, G., Asunis, G., Vinelli, F., Salvemini, M. Cipriano, P. Garretti, L., Vico, F., Oggiano, A. and Castelein, W., 2004. Lessons learnt from Italian NSDI. INSPIRE Infrastructure for Spatial Information in Europe.

Phillips, A.,1998. A Metadata Management System for Web-based SDIs. MSc Thesis. The University of Melbourne, Australia.

Price Waterhouse, 1995. Australian Land and Geographic Infrastructure-Benefits Study. Canberra: Australia Government Publishing Service.

Price Waterhouse Nederland BV, 1996. Beheerorganisatie National Clearinghouse Geo-Informatie. (in Dutch)

Radwan, M., 2002. The Development of Geographic Information Infrastructure 'GDI/SDI' to Support Access to Spatial data in Distributed Environment. Enschede, The Netherlands.

Rajabifard, A. and Williamson, I.P., 2001. Spatial Data Infrastructures: an initiative to facilitate spatial data sharing. In: R. Tateishi and D. Hastings (Editors) Global Environmental Databases – Present Situation and Future Directions. Volume 2. International Society for Photogrammetry and Remote Sensing (ISPRS-WG IV/8), Hong Kong, pp. 108-136.

Rajabifard, A., 2002. Diffusion for Regional Spatial Data Infrastructures: particular reference to Asia and the Pacific. PhD-thesis, The University of Melbourne, Melbourne, Australia.

Rajabifard, A, Feeney, M.E. and Williamson, I.P., 2002. Future directions for SDI Development. International Journal of Applied Earth Observation and Geoinformation, 4(1): 11-22.

Rajabifard, A., Feeney, M.E., Williamson, I. and Masser, I., 2003. National SDI-initiatives (Chapter six). In: I. Williamson, A, Rajabifard and M.E. Feeney (Editors), Developing Spatial Data Infrastructures: From Concept to Reality. Taylor & Francis, London, UK, pp. 95-109.

Ravi, 2003. Space for Geo-information, Bsik knowledge project proposal. Amersfoort, The Netherlands.

Reece, K., 2004. Defining and evaluating the criteria for effective implementation of national geospatial data clearinghouse in Jamaica. MSc-thesis, International Institute for Geo-Information Science and Earth Observation, Enschede, The Netherlands.

Reeve, D. and Petch, J., 1999. GIS Organisations and People, A Socio-Technical Approach. Taylor & Francis.

Renong Berhad, 1995. Feasibility study for the National Land Information System. Final Report.

Rhind, D., 1999. National and internal geospatial data policies (Chapter 56). In: P. Longley, M. Goodchild, D. Maguire and D. Rhind (Editors), Geographical Information Systems: Principles, Techniques, Management and Applications. Wiley, New York, pp. 767-787.

Rhind, D., 2000. Funding an NGDI. In: R. Groot and J. McLaughlin (Editors) Geospatial Data Infrastructure Concepts, Cases and Good Practice. Oxford University Press, Oxford, UK, pp. 39-55.

Rogers, E.M., 1995.Diffusion of Innovations. The Free Press, New York.

Shannon, C., 1948. Mathematical Theory of Communication. Bell Systems Technical Journal, 27: 379-423, 623-656.

Snedecor, G.W. and Cochran, W.G., 1980. Statistical methods. 7th edition. Iowa State University Press, Ames.

Southern California Association of Governments, 1998. Access Project. US Department of Transport.

Spatial Applications Division, Catholic University of Leuven, 2003. Spatial data infrastructures in Europe: state of play during 2003, Summary report.

Steudler, D., 2003. Developing evaluation and performance indicators for SDIs. In: I. Williamson, A. Rajabifard, and M.E. Feeney (Editors), Developing Spatial Data Infrastructures: From Concept to Reality. Taylor & Francis, London, pp. 235-246.

Tait, M.G., 2005. Implementing geoportals: applications of distributed GIS. Computers, Environment and Urban Systems, 29: 33-47.

Taylor, C., Nicholas, C. Hobson, B. and Goodrich, C.C., 1990. Social Assessment: Theory, Process and Techniques.

The Clearinghouse Payments Company, 2003. New York Clearing House, Historical Perspective.

The World bank, 2003; 2004. World Development Indicators. The World bank, Washington.

Tobias, R.D., 1997. An introduction to partial least squares regression. SAS Institute, Cary, NC.

Tosta, N., 1997. Data revelations in Qatar: why the same standards won't work in the United States. Geo Info System, 7:5.

United Nations Department of Economic and Social Affairs, 1992. Agenda 21.

United Nations Development Programme, 2003; 2004. Human Development Report. Oxford University Press, Oxford.

Wachowicz, M., 2002. Data mining in environmental sciences. In: J. Meij (editor), Dealing with data flood, mining data, text and multimedia. Study centre for technology trends, The Netherlands, pp. 183-202.

Walther, J., 2005. GEOmis.bund – 1. Stufe der GDI-DE. In L. Bernard, J. Fitzke and R.M. Wagner (Editors), Geodaten-infrastruktur, Grundlagen und Anwendungen. Herbert Wichmann Verlag, Heidelberg, pp. 163-169 (in German).

Wang, J., 2005. Using data mining to determine the societal impact on national spatial data clearinghouses. MSc thesis, Wageningen University, The Netherlands.

Wehn de Montalvo, U., 2001. Strategies for SDI implementation: A survey of national experiences, 5th GSDI Conference, Cartagena de Indias, Colombia.

Wehn de Montalvo, U., 2004. Chapter Eight, Outreach and Capacity Building, In: Nebert, D.D. (Editor), Developing Spatial Data Infrastructures: The SDI Cookbook, Version 2.0. GSDI-Technical Working Group, pp. 96-116.

Williamson, I., Rajabifard, A., and Feeney, M-E. F. (Editors), 2003. Developing Spatial Data Infrastructures: From Concept to Reality. Taylor & Francis, London, UK.

Williamson, I., 2004. Report on inter-jurisdictional coordination and benchmarking of spatial data infrastructures for the Office of Spatial Data Management Geoscience Australia.

Summary

Over the last few years many countries have spent considerable resources on developing their own National Spatial Data Infrastructure (NSDI) in order to manage and utilise spatial data assets more efficiently, reduce the cost of data production, and eliminate duplication of data acquisition. One of the key features of an NSDI is the national clearinghouse for spatial data, which can be regarded as the access network of an NSDI facilitating access to a nation's spatial data and related services. A clearinghouse can be defined as an electronic facility for searching, viewing, transferring, ordering, advertising and/or disseminating spatial data from numerous sources via the Internet. Such a facility usually consists of a number of servers, which contain information (metadata) about available digital data. A fully operational national clearinghouse enhances and innovates the NSDI of a country

The main objective of this thesis is to analyse the development and impact of national spatial data clearinghouses worldwide as well as the impact of society on these facilities. The results obtained provide support for the development of strategies for establishing and maintaining clearinghouses. In order to achieve this main objective, the following five sub-objectives have been formulated:

1) To assess the worldwide status of national spatial data clearinghouses.
2) To assess the worldwide developments of national spatial data clearinghouses.
3) To assess the impact of national spatial data clearinghouses on society, and in particular the GI-community.
4) To explore the societal impact on the establishment of national spatial data clearinghouses.
5) To explore the societal impact on the success of national spatial data clearinghouses.

The results of these assessments and explorations provide a basis for identifying critical factors that affect national clearinghouses. This knowledge can be used to improve the use, content and management of (future) clearinghouses. Each of the chapters 2-6 focuses on a sub-objective as presented earlier.

Chapter 2 systematically assesses and presents the worldwide status of national spatial data clearinghouses in December 2001. The examination of clearinghouse status was based on a web survey comprising an inventory of all established clearinghouses as well as measurements of twelve key characteristics: 1) year of first implementation; 2) number of data suppliers; 3) level of (meta)data accessibility; 4) metadata-standard applied; 5) number of datasets; 6) most recently produced dataset; 7) number of web references; 8) number of monthly visitors; 9) frequency of web updates; 10) language used; 11) use of maps for searching and 12) registration-only access. In December 2001, 59 countries had established their own national clearinghouse with significant differences in content, use and management.

Chapter 3 systematically assesses and presents the development of all national clearinghouses throughout the world. Development assessment was primarily based on a longitudinal web survey, undertaken in April and December of 2000, 2001 and 2002 respectively. By December 2002, 67 countries had already established their own national clearinghouse. This quantity alone is significant. Of concern however, are the declining trends in use, management and content. One of the main reasons is likely be the dissatisfaction of the

GI-community with the functional capability of national clearinghouses. Functional capabilities should perhaps be shifted from a data-oriented into a user and application-oriented focus; this is in line with the objectives of second-generation spatial data infrastructures (Rajabifard et al., 2003). Therefore, the main indications of critical (clearinghouse-internal) success factors are: the introduction of web services, the stability of funding, the provision of good communication channels, the formulation of a clear vision of the national clearinghouse function, the creation of user-friendly interfaces with less discipline-specific terminology, as well as trust in the clearinghouse.

Chapter 4 systematically assesses and presents the worldwide impact of spatial data clearinghouses on society, and in particular the GI-community. The aim of this assessment is to assist policy makers in their task of evaluating whether or not investment in setting up and maintaining these facilities is justified. In order to achieve this objective a procedure was devised for the systematic evaluation of sustainable development within the worldwide clearinghouse population. The assessment procedure entailed a survey undertaken by clearinghouse coordinators (November 2003 – April 2004). A range of economic, social and environmental indicators was chosen to evaluate the relevance, efficiency and effectiveness of clearinghouses. This chapter also presents the results of complementary analyses, which were carried out to assess the significance of the impacts recorded. They were also used to assess the objectivity of the coordinators' responses. The results of these assessments reveal that clearinghouses (of the developed world) have mainly positive (economic) impacts. In addition, the results also indicate the significance of clearinghouses as relevant facilities for enhancing spatial data accessibility, for providing efficient means of accessing spatial data, and the effective promotion of data use and distribution. Finally, it is argued that results obtained can be used to justify present and support future investments in the clearinghouse system.

Chapter 5 explores and presents the impact of society on the establishment of national clearinghouses. The aim of this exploration is to answer the following questions from a societal perspective: what are the critical factors influencing the establishment of national spatial data clearinghouses; why have some countries established a national clearinghouse, and why have other countries not done so? With these questions in mind, explorative and empirical analyses were carried out using the analysis of variances (ANOVA) and the decision tree method of data mining. The starting point was the worldwide status of established national clearinghouses in December 2002. The societal conditions of each country were characterised by attributes relating to nine societal aspects (economy, education, technology, environment, culture, demography, institution, health care, and jurisdiction). The main findings of this exploration indicate that all nine aspects can be considered as critical for such establishment, especially education, technology, institution, and health care aspects. Consequently, many societal attributes can be considered to play a critical role. Examples of critical attributes collected are: agricultural productivity, years of schooling, number of Internet secure servers, year of independence, and births attended by skilled health staff. These critical attributes give an approximate indication of the standard of living of a nation's society. It is plausible that the standard of living (rather than wealth) plays an important role in determining whether a country has an established national clearinghouse or not.

Chapter 6 explores and presents the impact of society on the success of national clearinghouses for the situation of December 2002. It also identifies the societal factors behind the success of such a clearinghouse. A clearinghouse suitability index was developed in order to measure the success of clearinghouses. The society of each country was described

by the same attributes relating to the nine societal aspects of Chapter 5. Several explorative analyses were performed based on statistics (e.g. partial least squares regression). Moreover, the success of potential clearinghouses was predicted for countries with no previously existing facility. These prediction results were tested against the results of the web survey of April 2005. The findings of the analyses highlight the importance of certain societal conditions for the success of national clearinghouses. However, society alone does not fully determine success, since clearinghouse-internal factors are also important. From a societal perspective, success is likely to be dependent on a combination of attributes from different societal aspects, in particular economy, education and technology. Critical factors for success could be those that raise the level of wealth in the country (e.g. taxes on income, gross domestic product, education expenditure and Internet accessibility). In April 2005, 83 countries had established their own national clearinghouse. In the near future newly established national clearinghouses, in countries where they have not previously been established, are likely to face serious obstacles or may even fail as a result of unfavourable societal conditions. This could, however, be reversed if critical clearinghouse-internal factors are implemented.

Through the results obtained, it appears that no single best solution or recipe exists since each country has its own unique society. However, the results of this thesis will support clearinghouse coordinators and policy makers in the development of successful strategies for establishing and maintaining national clearinghouses. In this way, the results of this thesis can contribute to the enhancement of national clearinghouses and NSDIs in many countries.

Samenvatting

De laatste jaren is in veel landen een aanzienlijke hoeveelheid middelen geïnvesteerd in een eigen Nationale Geo-Informatie Infrastructuur (NGII). Het doel hiervan is ruimtelijke gegevens efficiënter te gebruiken en te beheren, de kosten van gegevensproductie te verlagen en duplicatie van gegevensinwinning te voorkomen. Eén van de belangrijkste onderdelen van een NGII is een nationaal clearinghouse voor geo-informatie. Dit kan beschouwd worden als het toegangsnetwerk van een NGII waarmee de ontsluiting van landelijke ruimtelijke gegevens en gerelateerde diensten gefaciliteerd wordt. Een clearinghouse kan gedefinieerd worden als een elektronische faciliteit voor het zoeken, kijken, overzetten, ordenen, adverteren en/of verspreiden van ruimtelijke gegevens van verschillende bronnen via het Internet. Een dergelijke faciliteit bestaat gewoonlijk uit verschillende servers die informatie (metadata) bevatten over de beschikbare digitale gegevens. Een goed functionerend nationaal clearinghouse voor geo-informatie kan een wezenlijke bijdrage leveren aan de verbetering en innovatie van een NGII.

De hoofddoelstelling van dit proefschrift is het wereldwijd analyseren van de ontwikkeling en invloed van nationale clearinghouses, evenals de invloed van de samenleving op deze clearinghouses. De gepresenteerde resultaten bieden hulp bij de ontwikkeling van strategieën voor het opstellen en onderhouden van clearinghouses. Om deze doelstelling te bereiken zijn de volgende vijf deeldoelstellingen geformuleerd:

1) Het vaststellen van de wereldwijde stand van zaken omtrent nationale clearinghouses.
2) Het vaststellen van de wereldwijde ontwikkelingen van nationale clearinghouses.
3) Het bepalen van de invloed van clearinghouses op de samenleving, met name op het geo-informatie werkveld.
4) Het verkennen van de maatschappelijke invloed op het tot stand komen van nationale clearinghouses.
5) Het verkennen van de maatschappelijke invloed op het succes van nationale clearinghouses.

De resultaten van deze deeldoelstellingen verschaffen een basis voor de identificatie van kritieke succesfactoren voor nationaal clearinghouses. Deze kennis kan gebruikt worden om het gebruik, de inhoud en het beheer van (toekomstige) clearinghouses te verbeteren. Deze deeldoelstellingen zijn in hoofdstukken 2-6 verder uitgewerkt.

Hoofdstuk 2 beschrijft systematisch de stand van zaken van nationale clearinghouses in december 2001. Deze inventarisatie is gebaseerd op een ‘web survey’, waarbij alle bestaande clearinghouses zijn gekarakteriseerd aan de hand van twaalf kenmerken: 1) jaar van aanvang; 2) aantal dataleveranciers; 3) niveau van (meta)data ontsluiting; 4) metadata standaard; 5) aantal datasets; 6) meest recent aangemaakte dataset; 7) aantal internetverwijzingen; 8) aantal bezoekers per maand; 9) paginaverversing; 10) gebruikte taal; 11) gebruik van kaarten bij het zoeken; 12) gebruikersregistratie. In december 2001 hadden 59 landen een nationaal clearinghouse met grote verschillen in inhoud, gebruik en beheer.

Hoofdstuk 3 bevat een systematische analyse van de ontwikkeling van alle clearinghouses wereldwijd. Deze analyse is hoofdzakelijk gebaseerd op een longitudinale ‘web survey’ uitgevoerd in resp. april 2000, 2001, 2002 en december 2000, 2001, 2002. De ontwikkeling van het aantal nationale clearinghouses kan beschouwd worden als een wereldwijd succes. In december 2002 hadden 67 landen een nationaal clearinghouse. De

afnemende trends in gebruik, beheer en inhoud van deze faciliteiten is echter zorgwekkend. Een van de belangrijkste redenen hiervan is waarschijnlijk de ontevredenheid van geo-informatie gebruikers over de functionaliteit van nationale clearinghouses. Deze zou moeten verschuiven van een gegevensgeoriënteerde naar een meer gebruikers- en toepassingsgeoriënteerde focus, hetgeen overeenkomt met de doelstellingen van de tweede generatie Geo-Informatie Infrastructuren (Rajabifard et al., 2003). De resultaten van deze ontwikkelingsanalyse verschaffen een basis voor de identificatie van kritieke succesfactoren die een nationaal clearinghouse positief kunnen beïnvloeden. De belangrijkste clearinghouse-gebonden succesfactoren zijn: de introductie van webservices, de continuering van financiering, het verstrekken van goede communicatiekanalen, het creëren van een duidelijke visie, het bouwen van gebruikersvriendelijke interfaces met minder jargon en het vertrouwen in het clearinghouse.

Hoofdstuk 4 bevat een systematische analyse van de invloed van clearinghouses op de samenleving en in het bijzonder op het geo-informatie werkveld. Het doel van deze analyse is beleidsmakers te ondersteunen in hun taak om vast te stellen of de investering in opzet en onderhoud van een clearinghouse al dan niet gerechtvaardigd is. De analyse is gebaseerd op een wereldwijde 'survey' die door de clearinghouse-coördinatoren is uitgevoerd in de periode November 2003 – April 2004. In deze 'survey' is gebruik gemaakt van diverse economische, sociale en milieu-indicatoren om de relevantie, efficiency en effectiviteit van clearinghouses vast te stellen. De resultaten laten zien dat clearinghouses (in ontwikkelde landen) voornamelijk positieve (economische) gevolgen hebben. Clearinghouses zijn relevante faciliteiten voor de verbetering van de ontsluiting van ruimtelijke gegevens, efficiënte middelen voor gegevensontsluiting en effectieve bevordering van ruimtelijke gegevensgebruik en -verspreiding. Hiermee wordt aangetoond dat de huidige en toekomstige investeringen in clearinghouses verantwoord zijn.

In hoofdstuk 5 wordt de invloed van de samenleving op het tot stand komen van nationale clearinghouses verkend. Het doel van deze verkenning is om vanuit een maatschappelijk perspectief antwoord te geven op de volgende vragen: wat zijn de kritieke factoren die ontwikkeling van een clearinghouse beïnvloeden, en waarom hebben sommige landen een clearinghouse ontwikkeld en andere landen niet? Met deze vragen in het achterhoofd zijn verkennende en empirische analyses uitgevoerd, gebruikmakend van 'kwantitatieve multivariatie' analyses (ANOVA) en de 'Decision Tree' methode van 'Data Mining'. Hierbij is uitgegaan van de clearinghouses zoals vastgesteld in 'web survey' van december 2002. De samenleving van een land werd gekarakteriseerd door attributen gerelateerd aan negen maatschappelijke aspecten: 1) economie; 2) onderwijs; 3) technologie; 4) milieu; 5) cultuur; 6) demografie; 7) instituties; 8) gezondheidszorg en 9) rechtspraak. Een belangrijke conclusie van deze verkenning is dat alle negen maatschappelijke aspecten kritieke aspecten zijn, waarvan onderwijs, technologie en gezondheidszorg de belangrijkste. Als gevolg hiervan moeten veel maatschappelijke attributen als kritiek worden beschouwd. Voorbeelden hiervan zijn: de landbouwproductiviteit, het aantal jaren opleiding, het aantal Internet servers, het jaar van onafhankelijkheid en het aantal geboorten onder begeleiding van een verloskundige. Hiermee kan een indicatie van de levensstandaard gegeven worden. Het is daarom aannemelijk dat de levensstandaard (en niet alleen de welvaart) voor een belangrijk deel bepalend is of een land een clearinghouse heeft opgezet of niet.

In hoofdstuk 6 wordt de invloed van de samenleving op het succes van een clearinghouse verkend (situatie december 2002) en de maatschappelijk succesfactoren vastgesteld. Hiervoor is een clearinghouse geschiktheidsindex ontwikkeld welke een indicatie geeft voor het succes van het clearinghouse. De maatschappelijke context van een land is beschreven door dezelfde attributen die gerelateerd zijn aan de negen maatschappelijke aspecten van Hoofdstuk 5. Verschillende verkennende analyses zijn uitgevoerd gebruikmakend van statistische methoden zoals 'Partial Least Square' regressie. Bovendien is het potentiële succes voorspeld voor landen die nog geen clearinghouse hebben. Deze voorspellingen zijn getoetst aan de werkelijke situatie in april 2005. De resultaten van de analyse benadrukken het belang van de samenleving op het succes van nationaal clearinghouses. Echter, maatschappelijke aspecten alleen verklaren niet volledig het succes, omdat clearinghouse-gebonden factoren (zoals die van Hoofdstuk 3) ook belangrijk blijken. Maatschappelijk gezien is het waarschijnlijk dat het succes afhankelijk is van een combinatie van attributen met economie, onderwijs en technologie als belangrijkste. Welvaartsbevorderende factoren (zoals bijvoorbeeld inkomstenbelasting, bruto nationaal product, onderwijsuitgaven en Internet-toegang) blijken belangrijke kritische succesfactoren te zijn. In april 2005 hadden 83 landen een nationaal clearinghouse ingericht. De meeste landen die in de nabije toekomst een nationale clearinghouses willen opzetten zullen ernstig belemmerd worden in de uitvoering, zoniet falen, door de ongunstige maatschappelijke situatie. Dit kan echter omgekeerd worden als de kritieke clearinghouse-gebonden factoren goed opgepakt worden.

Uit de verkregen resultaten blijkt er niet één de beste oplossing of het beste recept te zijn om het gebruik, de inhoud en het beheer van (toekomstige) clearinghouses te verbeteren. Echter, de resultaten van dit proefschrift kunnen clearinghouse-coördinatoren en beleidsmakers ondersteunen bij de ontwikkeling van strategieën voor de opzet en onderhoud van clearinghouses. Door deze ondersteuning kan dit proefschrift een wezenlijke bijdrage leveren aan de verbetering van nationale clearinghouses en Nationale Geo-Informatie Infrastructuren in veel landen.

Appendix 1

A list of attributes (with units) is presented below. These attributes were used in Chapters 5 and 6.

The attributes are classified into nine societal aspects: 1) Economy, 2) Education, 3) Technology, 4) Environment, 5) Culture, 6) Demography, 7) Institution, 8) Health care and 9) Jurisdiction.

The attribute scale (nominal, ordinal, binary, interval, ratio) as well as the source code (see Table below) of each attribute are presented in parentheses. The source code refers to the source material used.

Table Explanation of Source code used.

Source code	Source description
CIA	Central Intelligence Agency, 2003; 2004. The World Factbook.
Hofstede	Hofstede, G., 2001. Culture consequences, Comparing Values, Behaviors, Institutions, and Organisations Across Nations. Sage Publications, Thousand Oaks.
Worldbank	The Worldbank 2003; 2004. World Development Indicators. The World Bank Group.
UNDP	United Nations Development Programme, 2003; 2004. Human Development Reports,
ITU	International Telecommunications Union, 2003. World Telecommunications Indicators
DK	Dorling Kindersley, 2002. Great World atlas. Dorling Kindersley Limited.

Attribute list

1. Economy

1. Gross national income, $ billions, 2001 (ratio, Worldbank)
2. Gross national income per capita $ (calculated using the Worldbank atlas), 2001 (ratio, Worldbank)
3. Purchasing power parity (PPP) gross national income, $ billions, 2001 (ratio, Worldbank)
4. Purchasing power purchasing (PPP) gross national income, $ per capita, 2001 (ratio, Worldbank)
5. Purchasing power parity (PPP) conversion factor, local currency units to international $, 2001 (ratio, Worldbank)
6. Ratio of PPP conversion factor to official exchange rate, 2001 (ratio, Worldbank)
7. Real effective exchange rate, 1995 = 100, 2001 (interval, Worldbank)
8. Gross domestic product, $ millions, 2001 (ratio, Worldbank)
9. Gross domestic product, average annual % growth, 1990- 2001 (ratio, Worldbank)
10. Gross domestic product, % growth, 2000-01 (ratio, Worldbank)
11. Gross domestic product, per capita, % growth, 2000-01 (ratio, Worldbank)
12. GDP implicit deflator, average annual % growth, 1990-2001 (ratio, Worldbank)
13. Unemployment: Total, % of total employment, 1998-2001 (ratio, Worldbank)
14. Long-term unemployment, % of total unemployment, total, 1998-2001 (ratio, Worldbank)
15. Average hours worked per week, 1995-2000 (ratio, Worldbank)
16. Minimum wage, $ per year, 1995-2000 (ratio, Worldbank)
17. Labour cost per worker in manufacturing, $ per year, 1995-2000 (ratio, Worldbank)
18. Value added per worker in manufacturing, $ per year, 1995-2000 (ratio, Worldbank)
19. Agricultural productivity, value added per worker, 1995 $, 1999-2001 (ratio, Worldbank)
20. Population below $2 a day, %, 2001 (ratio, Worldbank)
21. Poverty gap at $2 a day, %, 2001 (ratio, Worldbank)
22. Gini-index, 2001 (ratio, Worldbank)
23. Gross domestic savings, % of GDP, 2001 (ratio, Worldbank)
24. Gross national savings, % of GNI, 2001 (ratio, Worldbank)
25. Net national savings, % of GNI, 2001 (ratio, Worldbank)

26. Adjusted net savings, % of GNI, 2001 (ratio, Worldbank)
27. Energy depletion, % of GNI, 2001 (ratio, Worldbank)
28. Trade in goods, % of GDP, 2001 (ratio, Worldbank)
29. Trade in goods, % of goods GDP, 2001 (ratio, Worldbank)
30. Change in trade, % of GDP, 1990-2000 (ratio, Worldbank)
31. Gross private capital flows, % of GDP, 2001 (ratio, Worldbank)
32. Net private capital flows, $ millions, 2001 (ratio, Worldbank)
33. Domestic credit to private sector, % of GDP, 2001 (ratio, Worldbank)
34. Gross foreign direct investment, % of GDP, 2001 (ratio, Worldbank)
35. Foreign direct investment, $ millions, 2001 (ratio, Worldbank)
36. Aid dependency ratio, Aid as % of GNI, 2001 (ratio, Worldbank)
37. Aid dependency ratio, Aid as % of gross capital formation, 2001 (ratio, Worldbank)
38. Aid dependency ratio, Aid as % of imports of goods and services, 2001 (ratio, Worldbank)
39. Euromoney country credit-worthiness rating, September 2002 (ratio, Worldbank)
40. Value traded, % of GDP, 2001 (ratio, Worldbank)
41. Listed domestic companies, 2002 (ratio, Worldbank)
42. Tax revenue, % of GDP, 2001 (ratio, Worldbank)
43. Taxes on Income, profits and capital gains, % of total taxes, 2001 (ratio, Worldbank)
44. Taxes on Income, profits, and capital gains, % of total current revenue, 2000 (ratio, Worldbank)
45. Social security taxes, % of total current revenue, 2000 (ratio, Worldbank)
46. Domestic taxes on goods and services, % of value added in industry and services, 2001 (ratio, Worldbank)
47. Taxes on goods and services, % of total current revenue, 2000 (ratio, Worldbank)
48. Taxes on international trade, % of total current revenue, 2000 (ratio, Worldbank)
49. Central government expenditures, Goods and services, % of total expenditure, 2000 (ratio, Worldbank)
50. Wages and salaries, % of total (governmental) expenditure, 2000 (ratio, Worldbank)
51. Subsidies and other current transfers, % of total (governmental) expenditures, 2000 (ratio, Worldbank)
52. Military expenditure, of % of GDP, 2001 (ratio, Worldbank)
53. Military expenditure, % of central government expenditure, 2001 (ratio, Worldbank)
54. Armed forces personnel, Total thousands, 2001 (ratio, Worldbank)
55. Armed forces personnel, % of labor force, 2001 (ratio, Worldbank)
56. Services, average annual % growth, 1990-2001 (ratio, Worldbank)
57. Services value added, % of GDP, 2001 (ratio, Worldbank)
58. Net barter terms of trade, 2000 (interval, Worldbank)
59. Fuels, % of total (national) merchandise exports, 2001 (ratio, Worldbank)
60. Fuels, % of total (national) merchandise imports, 2001 (ratio, Worldbank)
61. Consumer price index, average annual % growth, 1990-2001 (ratio, Worldbank)
62. Food price index, average annual % growth, 1990-2001 (ratio, Worldbank)
63. Household final consumption expenditure, % of GDP, 2001 (ratio, Worldbank)
64. Household final consumption expenditure, $ millions, 2001 (ratio, Worldbank)
65. Household final consumption expenditure, average annual, % growth, 1990-2001 (ratio, Worldbank)
66. Household final consumption expenditure per capita, average annual % growth, 1990-2001 (ratio, Worldbank)
67. General government final consumption expenditure, % of GDP, 2001 (ratio, Worldbank)
68. General government final consumption expenditure, average annual % growth, 1990-2001 (ratio, Worldbank)
69. Gross capital formation, average annual % growth, 1990-2001 (ratio, Worldbank)
70. Central government finances, Current revenue, % of GDP, 2000 (ratio, Worldbank)
71. Central government finances, Total expenditure% of GDP, 2000 (ratio, Worldbank)
72. Central government finances, Overall budget balance (including grants), % of GDP, 2000 (ratio, Worldbank)
73. Total external debt, $ millions, 2001 (ratio, Worldbank)
74. GDP index, 2001 (ratio, UNDP)
75. Services: male, % of male employment, 1998-2001 (ratio, 130 Worldbank)
76. Services: female, % of female employment, 1998-2001 (ratio, Worldbank)
77. Average annual change in Consumer price index, %, 2000-2001 (ratio, UNDP)

2. Education

1. Adult literacy rate % age 15 and above, 2001 (ratio, UNDP)
2. Adult illiteracy rate, Male, % ages 15 and above, 2001 (ratio, Worldbank)
3. Adult illiteracy rate, Female, % ages 15 and above, 2001 (ratio, Worldbank)
4. Public expenditure on education % of GDP, 2000 (ratio, Worldbank)
5. Public expenditure on education, per student % of GDP per capita, 2000 (ratio, Worldbank)
6. Public expenditure on education, % of total government expenditure, 2000 (ratio, Worldbank)
7. Primary pupil / teacher ratio, pupils per teacher, 2000 (ratio, Worldbank)
8. Combined primary, secondary and tertiary gross enrolment ratio, %, 2000-2001 (ratio, UNDP)
9. Gross enrolment ratio, Primary % of relevant age group, 2000 (ratio, Worldbank)
10. Gross enrolment ratio, Tertairy % of relevant age group, 2000 (ratio, Worldbank)
11. Net enrolment ratio, Primary % of relevant age group, 2000 (ratio, Worldbank)
12. Primary completion rate, Total, 1995-2001 (ratio, Worldbank)
13. Average years of schooling, Total, 2000 (ratio, Worldbank)
14. Education expenditure, % of GNI, 2001 (ratio, Worldbank)
15. Education index, 2001 (ratio, UNDP)

3. Technology

1. Electricity production, billion kwh, 2000 (ratio, Worldbank)
2. Access to electricity, % of population, 2000 (ratio, Worldbank)
3. Electric power, consumption per capita kwh, 2000 (ratio, Worldbank)
4. Electric power, Transmission and distribution losses, % of output, 2000 (ratio, Worldbank)
5. Motor vehicles, per 1000 people, 2000 (ratio, Worldbank)
6. Motor vehicles, per kilometre of road, 2000 (ratio, Worldbank)
7. Fuel prices, Super $ per litre, 2002 (ratio, Worldbank)
8. Fuel prices, Diesel $ per litre, 2002 (ratio, Worldbank)
9. Telecommunication $million, 2001 (ratio, Worldbank)
10. Roads, Total road network km, 1995-2000 (ratio, Worldbank)
11. Roads, % Paved roads, 1995-2000 (ratio, Worldbank)
12. Roads, Goods hauled million ton-km, 1995-2000 (ratio, Worldbank)
13. Railways, Rail lines, Total km, 1996-2001 (ratio, Worldbank)
14. Railways, Traffic Density traffic units per km, 1996-2001 (ratio, Worldbank)
15. Air, aircraft departures thousands, 2001 (ratio, Worldbank)
16. Total telephone subscribers, Total (k), 2002 (ratio, ITU)
17. Fixed line and mobile phone subscribers per 1000 people, 2001 (ratio, Worldbank)
18. Main telephone lines per 100 inhabitants, 2002 (ratio, ITU)
19. Main telephone lines per 100 inhabitants, Annual growth %, 1997-2002 (ratio, ITU)
20. Main telephone lines, (k), 2002 (ratio, ITU)
21. Telephone mainlines, Annual growth %, 1997-2002 (ratio, Worldbank)
22. Mobile phone subscribers, (k), 2002 (ratio, ITU)
23. Mobile phone subscribers, Annual growth rate, %, 1997-2002 (ratio, ITU)
24. Mobile phones per 100 inhabitants, 2002 (ratio, ITU)
25. Mobile phone subscribers, % Digital, 2002 (ratio, ITU)
26. Mobile phone, as % of total telephone subscribers, 2002 (ratio, ITU)
27. International telecommunications, Outgoing traffic minutes per subscriber, 2001 (ratio, Worldbank)
28. Personal Computers, thousands, 2002 (ratio, ITU)
29. Personal computers per 1000 people, 2001 (ratio, Worldbank)
30. Personal computers in education, 2001 (ratio, Worldbank)
31. Internet, Hosts Total, 2002 (ratio, ITU)
32. Internet, hosts per 10000 people, 2002 (ratio, ITU)
33. Internet, Users thousands, 2001 (ratio, Worldbank)
34. Internet, Users per 10000 people, 2002 (ratio, ITU)
35. Internet, Monthly off-peak access charges, Service provider charge, $, 2001 (ratio, Worldbank)
36. Internet, Monthly off-peak access charges, Telephone usage charge, $, 2001 (ratio, Worldbank)
37. Internet, Secure servers, 2001 (ratio, Worldbank)

38. ICT-expenditures, % of GDP, 2001 (ratio, Worldbank)
39. ICT-expenditures per capita, $, 2001 (ratio, Worldbank)
40. Scientists and engineers in R&D, per million people, 1990-2000 (ratio, Worldbank)
41. Technicians in R&D, per million people, 1990-2001 (ratio, Worldbank)
42. Expenditures for R&D, % of GDP, 1989-2000 (ratio, Worldbank)
43. High-technology exports, $ millions, 2001 (ratio, Worldbank)
44. High-technology exports, % of manufactured exports, 2001 (ratio, Worldbank)

4. Environment

1. Surface area thousand sq. km, 2001 (ratio, Worldbank)
2. Access to an improved water source, % population, 2000 (ratio Worldbank)
3. Access to improved sanitation facilities, % of population, 2000 (ratio, Worldbank)
4. Land Area, Thousands sq. km, 2000 (ratio, Worldbank)
5. Land use, Arable land, % of land area, 2000 (ratio, Worldbank)
6. Land use, Permanent crop land, % of land area, 2000 (ratio, Worldbank)
7. Land use, Other, % of land area, 2000 (ratio, Worldbank)
8. Arable land, hectares per capita, 1998-2000 (ratio, Worldbank)
9. Irrigated land, % of cropland, 1998-2000 (ratio, Worldbank)
10. Crop production index, 1999-2001 (ratio, Worldbank)
11. Food production index, 1999-2001 (ratio, Worldbank)
12. Livestock production index, 1999-2001 (ratio, Worldbank)
13. Cereal yield, kilograms per hectare, 1999-2001 (ratio, Worldbank)
14. Forest area, % of total land area, 2000 (ratio, Worldbank)
15. Average annual deforestation, %, 1990-2000 (ratio, Worldbank)
16. Mammals, Threatened species, 2000 (ratio, Worldbank)
17. Nationally protected area, % of total land area, 2002 (ratio, Worldbank)
18. Freshwater resources, Internal flows billion cu. M, 2000 (ratio, Worldbank)
19. Freshwater resources, Total renewable resources per capita cu. M, 2000 (ratio, Worldbank)
20. Annual freshwater withdrawals, billion cu. M, 2000 (ratio, Worldbank)
21. Annual freshwater withdrawals, % of total renewable resources, 2000 (ratio, Worldbank)
22. Emissions of organic water pollutants, kilograms per day, 2000 (ratio, Worldbank)
23. Emissions of organic water pollutants, kilograms per day per worker, 2000 (ratio, Worldbank)
24. Commercial energy use, thousand metric tons of oil equivalent, 2000 (ratio, Worldbank)
25. Commercial energy use, average annual % growth, 1980-2000 (ratio, Worldbank)
26. Commercial energy use per capita, kg of oil equivalent, 2000 (ratio, Worldbank)
27. Commercial energy use per capita, average annual % growth, 1980-2000 (ratio, Worldbank)
28. Carbon dioxide emissions, Total million metric tons, 2000 (ratio, Worldbank)
29. Carbon dioxide emissions per capita metric tons, 2000 (ratio, Worldbank)
30. Carbon dioxide emissions damage, % of GNI, 2001 (ratio, Worldbank)
31. Continent, 2002 (DK)
32. Year, Environmental strategy or action plan, 2002 (interval, Worldbank)
33. Year, Treaty, Climate Change, 2002 (interval, Worldbank)
34. Year, Treaty, Ozone layer, 2002 (interval, Worldbank)
35. Year, Treaty, CFC control, 2002 (interval, Worldbank)
36. Year, Treaty, Law of the Sea, 2002 (interval, Worldbank)
37. Year, Treaty, Biological diversity, 2002 (interval, Worldbank)

5. Culture

1. Power Distance, 20011 (ratio, Hofstede)
2. Uncertainty Avoidance, 2001 (ratio, Hofstede)
3. Long-term Thinking, 2001 (ratio, Hofstede)
4. Individualism/Collectivism, 2001 (ratio, Hofstede)
5. Masculinity/Feminity, 2001 (ratio, Hofstede)

6. Demography

1. Population, million, 2001 (ratio, Worldbank)
1. Population density, people/sq,km, 2001 (ratio, Worldbank)
2. Life expectancy at birth, years, 2001 (ratio, Worldbank + UNDP)
3. Life expectancy at birth, Male years, 2001 (ratio, Worldbank)
4. Life expectancy at birth, Female years, 2001 (ratio, Worldbank)
5. Gender-related development index (GDI), 2001 (ratio, UNDP)
6. Gender empowerment measure (GEM), 2001 (ratio, UNDP)
7. Promote gender equality, Ratio of female to male enrollments in primary and secondary school, 2000 (ratio, Worldbank)
8. Labour force gender parity index, 2001 (ratio Worldbank)
9. Average annual population growth rate, %, 1980-2001 (ratio, Worldbank)
10. Average annual population growth rate, %, 2001-15 (ratio, Worldbank)
11. Median Age, years, 2002 (ratio, CIA)
12. Population under age 15, %, 2001 (ratio, UNDP)
13. Population between ages 15-64, %, 2001 (ratio, Worldbank)
14. Population ages 15-64, millions, 2001 (ratio, Worldbank)
15. Crude death rate per 1000 people, 2001 (ratio, Worldbank)
16. Infant mortality rate, per 1000 live births, 2001 (ratio, Worldbank)
17. Under-five mortality rate per 1000, 2001 (ratio, Worldbank)
18. Crude birth rate per 1000people, 2001 (ratio, Worldbank)
19. Total fertility rate, births per woman, 2001 (ratio, Worldbank)
20. Adolescent fertility rate, births per 1000 women, ages 15-19, 2002 (ratio, Worldbank)
21. Labour force, total millions, 2001 (ratio, Worldbank)
22. Labour force, average annual growth rate, %, 1980-2001 (ratio, Worldbank)
23. Rural population, %, 2001 (ratio, Worldbank)
24. Urban population (% of total population), 2001 (ratio, Worldbank)
25. Human development index, 2001 (ratio, UNDP)
26. Life expectancy index, 2001 (ratio, UNDP)

7. Institution

1. Composite International Country Risk Guide (ICRG) risk rating, September 2002 (ratio, Worldbank)
2. Institutional Investor credit rating, September 2002 (ratio, Worldbank)
3. Type of Government, 2002 (nominal, DK)
4. Federation, 2002 (binary, DK)
5. Year Independence, 2002 (interval, CIA)
6. Year Current borders established, 2002 (interval, CIA)

8. Health care

1. Private health expenditure, % of total, 2000 (ratio, Worldbank)
2. Public expenditure on health, % of GDP, 2000 (ratio, Worldbank)
3. Health Expenditure, Total % of GDP, 1997-2000 (ratio, Worldbank)
4. Health expenditure per capita, $, 1997-2000 (ratio, Worldbank)
5. Hospital beds, per 1000 people, 1995-2000 (ratio, Worldbank)
6. Child immunization rate, % of children under age one, Measles, 2001 (ratio, Worldbank)
7. Child immunization rate, % of children under age one, DTP, 2001 (ratio, Worldbank)
8. Tuberculosis treatment success rate, % of registered cases, 2000 (ratio, Worldbank)
9. Incidence of tuberculosis, per 100000 people, 2000 (ratio, Worldbank)
10. Prevalence of HIV, % of adults, 2001 (ratio, Worldbank)
11. Improve maternal health, Births attended by skilled health staff, % of total, 2000 (ratio, Worldbank)
12. Prevalence of under nourishment, % of population, 1998-2000 (ratio, Worldbank)
13. Population with sustainable access to affordable essential drugs, %, 2000 (ratio, UNDP)

9. Jurisdiction

1. Legal system, 2002 (nominal, CIA)
2. Royalty and license fees, Receipts, $ millions, 2001 (ratio, Worldbank)
3. Royalty and license fees, Payments, $ millions, 2001 (ratio, Worldbank)
4. Patent applications, filed, residents, 2000 (ratio, Worldbank)
5. Patent applications filed, non-residents, 2000 (ratio, Worldbank)
6. International Convention on the Elimination of all Forms of Racial Discrimination 1965 (nominal, UNDP)
7. International Covenant on Civil and Political Rights 1966 (nominal UNDP)
8. International Covenant on Economic, Social and Cultural Rights 1966 (nominal, UNDP)
9. Convention on the Elimination of All Forms of Discrimination Against women 1979 (nominal, UNDP)
10. Convention Against Torture and Other Cruel, Inhuman or Degrading Treatment or Punishment 1984 (nominal, UNDP)
11. Convention on the Rights of the Child 1989 (nominal, UNDP)

Curriculum Vitae

Joep Crompvoets was born in Oss (The Netherlands) on 28th of August 1968, and finished his secondary education (VWO) at Titus Brandsmalyceum (Oss) in 1988. In 1993 he obtained his MSc. degree in soil science from Wageningen Agricultural University. In 1997 he obtained an MSc. degree in Geo-Information Science at the same university.

From 1993 to 1995, he worked as a researcher in Seville, Spain, where he participated in an EU-project on the development of a spatially distributed soil, agro-climatic, and soil hydrological model to predict the effects of climate change on land use within the European Community (IRNA-CSIC).

Since 1997 he has been a lecturer at the Laboratory of Geo-Information Science and Remote Sensing of Wageningen University and is mainly involved in GIS-related education and research. One of his main educational contributions was the implementation of a course in 'Spatial Data Infrastructures (SDI)'. Moreover, he is a lecturer in the MSc. 'Geographical Information Management and Applications' programme (a joint partner programme of Wageningen University, Delft University of Technology, International Institute for Geo-Information Science and Earth Observation, Utrecht University).

From 1999 until 2001 he participated as a researcher in an EU-project on an integrated approach for sustainable management of irrigated lands susceptible to degradation/desertification. In 2003 he spent three months at the Centre for Spatial Data Infrastructures of The University of Melbourne (Australia) for Dutch-Australian collaboration on SDI research. Since 2005 he has been leading the 'Development of Framework to Assess National Spatial Data Infrastructures' project of Bsik Innovation Programme 'Space for Geo-information'. Additionally, he was one of the main contributors of the EU INSPIRE Impact Analysis Working Group. INSPIRE (INfrastructure for SPatial InfoRmation in Europe) is an initiative of the European Commission to develop a European SDI.

He wrote more than 50 scientific publications in the field of soil science and geo-information science (spatial data infrastructures).

List of selected publications written by author

Bernard, L., Crompvoets, J. and Fitzke, J., 2005. Geodateninfrastrukturen – ein Überblick, In: L. Bernard, J. Fitzke and R.M. Wagner (Editors), Geodateninfrastruktur, Grundlagen und Anwendungen, Herbert Wichmann Verlag, Heidelberg, Germany, pp. 4-8. (in German).

Bregt, A. and Crompvoets, J., 2000. Geo-informatie in de netwerksamenleving, een tweeluik, VI Matrix 55(3): 16-19. (in Dutch)

Bregt, A., Crompvoets, J. and Wachowicz, M., 2002. INSPIRE Impact Analysis Working Group Position Paper. INSPIRE, Infrastructure for Spatial Information in Europe. November 2002. 49 pages.

Bregt, A., Crompvoets, J., Scholten, H. and Crommert, H. van den, 2005. Die Niederländische Nationale Geodateninfrastruktur – Rückblick, Gegenwart, und Zukunft. In: L. Bernard, J. Fitzke and R.M. Wagner (Editors), Geodateninfrastruktur Grundlagen und Anwendungen, Herbert Wichmann, Verlag, Heidelberg, Germany, pp.29-37. (in German).

Crompvoets, J., 1998. Method to assess the spatial suitability of SOTER-units to water Erosion. In: I. Masser and P. Burrough (Editors), Proceedings of the first meeting of the Association of Geographic Information Laboratories in Europe, Enschede.

Crompvoets, J., Booltink, H.W.G., 2001. Predesodi, an integrated approach for sustainable management of irrigated lands susceptible to degradation/desertification. Final report EU-project.

Crompvoets, J., 2002. Developments of national clearinghouses for geo-information. Proceedings of 6th Global Spatial Data Infrastructure Conference, From global to local, Hungary, Budapest, September 2002.

Crompvoets, J. and Kossen. H., 2002. The impact of culture on national spatial data clearinghouses. Proceedings of GISDECO-conference, Governance and the use of GIS in Developing Countries. Enschede, The Netherlands, pp. 9.1–9.3.

Crompvoets, J. and Bregt, A., 2003. Wereldwijd stagneert ontwikkeling nationale clearinghouses, ondanks groeiperspectief. VI Matrix 83(7): 28-31. (in Dutch).

Crompvoets, J. and Bregt, A., 2003. World status of National Spatial Data Clearinghouses. URISA Journal, Special Issue on Public Participation GIS, 15, APA I: 43-50.

Crompvoets, J., Bregt, A., Rajabifard, A. and Williamson, I., 2004. Assessing the worldwide developments of national spatial data clearinghouses. International Journal of Geographical Information Science, 18(7): 665-689.

Crompvoets, J., Wachowicz, M., Bree, F. de, and Bregt, A., 2004. Impact assessment of the INSPIRE Geo-Portal. In: Proceedings of ESDI: State of the Art, 10th edition of the EC GI & GIS Workshop, Warsaw, Poland, 7 pages.

Crompvoets, J., Bree, F. de, and Rajabifard, A., 2004, Status and Impact Assessment of spatial data clearinghouses/geoportals in Asia-Pacific region. Proceedings of Permanent Committee Geo-Information of Asia and Pacific (PCGIAP) WG1 and WG2 Workshop and PCGIAP Executive Board Meeting, September 7-9, 2004, Chengdu, China. 8 pages.

Crompvoets, J., Bregt, A., Bree, F. de, Oort, P.A.J., Loenen, B. van, Rajabifard, A. and Williamson, I., 2005. Worldwide impact assessment of geoportals. Proceedings From Pharaohs to Geoinformatics, FIG Working Week 2005 and 8th International Conference on Global Spatial Data Infrastructure, Egypt, Cairo. 6 pages.

Crompvoets, J., Stuiver, H.J. and Bulens, J.D., 2005. Worldwide impact assessment of geoportals. In: AGILE2005 Conference Proceedings, Lisbon, pp. 565 – 568. Winner of best Poster AGILE2005.

Crompvoets, J., Bregt, A. and Adrichem, M. van, 2005. Disappointing NSDI National Clearinghouse Survey, International Developments, Status, Suitability and Spatial Distribution. GIM International, 19 (9): 35-37.

De Bree, F. and J. Crompvoets, 2004, Clearinghouses hebben vooral economische effecten. VI-Matrix, 86(2): 20-21. (in Dutch).

De la Rosa, D. and Crompvoets, J., 1996. Land Vulnerability evaluation and climate change impact in Andalucia, Spain. Soil erosion and contamination. Journal of International Agrophysics, 10: 225-238.

De la Rosa, D., Crompvoets, J., Mayol, F., Moreno, J.A., Loveland, P., Rounsevell, M., Mayr, T. and Simota, C., 1996. Evaluación de impacto de cambio climático sobre capacidad productiva y riesgo de erosión en suelos agrícolas de Andalucía. Edafología. (in Spanish).

De la Rosa, D. and Crompvoets, J., 1997. Evaluating Mediterranean Soil Contamination Risks in Selected Hydrological Scenarios. Agriculture, Ecosystems & Environment, 1997, 1: 23–28.