Will ignorance be the real barrier to interoperability?

B. Kirwan

Submitted in fulfilment of the requirements for the degree of Master of Science in Geographical Information Systems (UNIGIS)

Faculty of Earth and Life Sciences
Vrije Universiteit Amsterdam
The Netherlands

November 2005

Abstract:

The rapid proliferation of spatial data, proprietary software and the implementation of location based technologies to many modern industrial and business activities has created a need to maximise the open exchange of information between geospatial enabled parties. Since the mid nineteen nineties a whole series of international, regional and national spatial data infrastructures (SDI) have been established.

The author has set himself the task of examining the state of knowledge amongst ordinary GIS professionals about these developments. If these developments are not sufficiently understood then the promise of interoperability will be a long time coming. Even where there is coverage one needs to critically examine the information to test to see what constituents of SDI are being treated and which one are not. Much of the advancement is in the area of technology but is the theoretical basis of the discipline suffering from neglect? Are we busy building a road but neglecting to write a road users safety code?

The dissertation attempts to tests these questions and answer the simple question will ignorance be the real barrier to interoperability?

Index

Disclaimer

The results presented in this thesis are based on my own research at the Faculty of Earth and Life Sciences of the Vrije Universiteit Amsterdam. All assistance received from other individuals and organisations has been acknowledged and full reference is made to all published and unpublished sources.

This thesis has not been submitted previously for a degree at any institution.

Place,
Date

Name student

Foreword, acknowledgements and thanks

The author's interest in SDI had been sparked both by the UNIGIS subject material and by what one was reading in the specialised GIS magazines at the time. In the spring of 2004 when I started work on the masters I had decided to build on these interests. Having worked as both a network manager and a GIS operative the width of the subject matter appealed to me. I was keenly aware however, that technical advancements need to be matched with educating people into the possibilities which the technology offers. Of all the constituents of SDI the area which most interested me were the issues of semantics and metadata as they represent the marriage of technology and people. It is this interest which has formed the primary motivation behind the subject material of this thesis and it is the authors hope that in one's own small way one has helped to build on the insights into SDI.

The author would like to acknowledge the assistance and support lent to him from the following persons;
Mr. A. van der Meer, Editor of "Geo-Info", Mrs F.M. van Berkel-Coumans, Editor VI Matrix, Prof. Dr. H. Scholten, VU Amsterdam, Drs. R.van de Velde, VU Amsterdam,Ir.M. van Breda, PZH, Ing. A. Dammhuis, PZH and all my colleagues at the PZH past and present particulary H.Remijn.

On a personal note I would like to take the opportunity to thank the entire staff (past and present) of UNIGIS for their cordial and professional support over the last four years and in particular Mr. Rob van de Velde for his creative and constructive input over the last sixteen months, it was both a pleasurable and educational process for the author.

In addition many thanks go out to Mr. J.Roijen for making this all possible without his foresight and belief in me I would still be a frustrated draughtsman. The drinks are on me Sjef! To Sjaak for the scans and encouragement, it is time to break open a bottle of whiskey young fella. To Kristel, for being prepared to look at the back of my head buried in a computer screen for the last three years and to Jonnick and Tristan for the time I did not spend with them, I will make it up to you all.

Brian Kirwan,
Spijkenisse.
November 2005

Chapter 1
Will ignorance be the real barrier to interoperability?

1.0 Introduction to chapter 1: Defence of the thesis:

The subject of this thesis is the development of what is known as Spatial Data Infrastructure(SDI) and one will attempt to examine how well informed we are, as a professional body, to grasp and implement the developments. The latter with its core principles, their benefits and implications will be dealt with in this opening chapter. The rest of the thesis will be used to establish the method and parameters of the research, generate and examine results and finally to draw as series of final conclusions.

Given the rate of technological change in our world one must be prepared to ask the question are the intellectual and institutional conditions conductive to realising SDI and interoperability? Are we sufficiently well informed or do we need to re-asses the way and the sort of information being disseminated about these developments?

This observer of SDI feels that the time has come to do some difficult soul searching amongst the GIS community. It is the authors belief that we should be at this juncture of the process be asking ourselves the question, will **ignorance** be the real barrier to interoperability?

Speaking from one's own experience, one can say that while there are grounds for technical optimism but also for concern. This thesis proposes to examine the state of knowledge inside a national GIS community in order to be able to answer the questions posed. In the predictions set out in Chapter 2 one will outline predictions ahead of the research and later having examined the data one will be making a series of appraisals and recommendations based on the findings of the research.

1.1 What is interoperability and SDI?

Before moving ahead it is necessary to define and detail the scope of the subject matter. Up to this point one has used the two phrases interoperability and SDI as interchangeable phrases and of course there is a difference between the two. Both need to be defined clearly.

Interoperability:
A quick Internet search demonstrates the ambiguity of the term interoperable; many of articles examined prefaced their remarks by alluding to the degree of confusion which exists about the definition of the word interoperability. A definition offered by the Rand Cooperation in the United States lends us a convenient starting point in that it defines interoperability in a generic light -

"Interoperability would seem to be a straightforward concept. Put simply, is a measure of the degree to which various organizations or individuals are able to operate together to achieve a common goal."

Reference: The Rand Organization (2001).

Alternatively the World Reference.com site defined it simply as -

"Interoperability :(computer science) the ability to exchange and use information usually in a large heterogeneous network made up of several local area networks"

Reference: www.World Refernce.com

A definition from industry, proposed by Panasonic contained the following definition -

"the condition achieved when two or more technical systems can exchange information directly in a way that is satisfactory to the users of the systems."

Reference 3: Panasonic(2002)

The NATO logistics handbook which coupled the concept of interoperability with standardisation contained the following definition-

"Within NATO, standardization is the process of developing concepts, doctrines, procedures and designs to achieve and maintain the most effective level of standardization in the fields of operations, administration and material. The levels of standardization are in ascending order compatibility, interoperability, interchangeability and commonality"

.

Reference: NATO (1997)

What can we summarise from all these disparate definitions?
Well from the Rand Cooperation we learn that at its most fundamental level, interoperability is about different parties being able to work together. This is essential; interoperability is about cooperation between people. In the present age of globalism co-operation is a tactical necessity and a strategic imperative in ensuring our capacity to continue to grow and prosper.

The remaining definitions reveal the technical interpretation of interoperability that it is a goal which allows us to technically interact with one another through whatever engineered system it is that we use. The NATO definition tends to tie the two threads together, that people, their thoughts, their organisations and their hard and software can combine to achieve a state of interoperability.

Interoperability is a property greater than the narrow confines of a technical development, has dimensions other than engineering and if we are to understand it at all we must be prepared to examine it holistically. It has a technological context but ultimately it is an interpersonal dynamic and we need to place the person in the centre of any discussions about it. Interoperability becomes a reality when all the parts – organisational, technological and intellectual work in synch with one another creating a whole, greater than the sum of all the parts.

1.2 SDI- a definition:

It is not the author intention to try and deal with all aspects of interoperability, the subject matter is too broad to deal with in a single piece of research. One shall limit the examination of interoperability to the GIS sphere and define the relationship schematically between interoperability and SDI as follows -

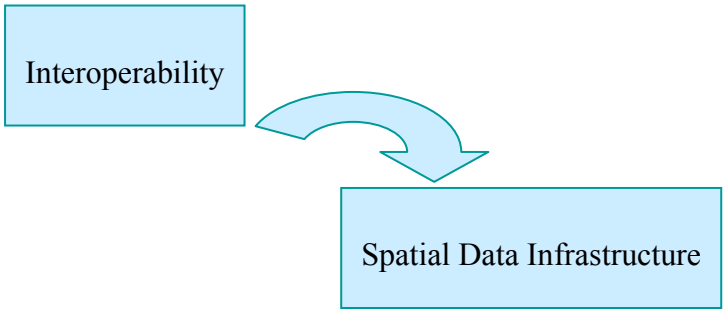


Figure 1.1 "The relationship between interoperability and SDI."

In this schematic SDI is a specific *form* of interoperability, namely one which deals with interoperability in a work sphere where spatial data is the primary sort of data or an environment where spatial data will be used as an organising principle. SDI deals with both institutional and technical aspects of interoperability but in a specific work sphere. The author shall henceforth only deal with interoperability implicitly through SDI.

SDI has become a buzzword in the GIS community in the last few years and like all good buzzwords it means a lot of different things to different people so one is compelled to define what it is exactly the author believes SDI to be. A SDI as defined in the "The SDI cookbook" (version 2.0 January 2004) is,

"often used to denote the relevant base collection of technologies, policies and institutional arrangements that facilitate the availability of and access to spatial data. The SDI provides a basis for spatial data discovery, evaluation and application for users and providers within all levels of government, the commercial sector, the non-profit sector academia and by citizens in general."

It goes on to tell us that -
"*An SDI must be more than a single data set or database; an SDI hosts geographic data and attributes , sufficient documentation (metadata), a means of discover, visualize, and evaluate the data (catalogues and Web mapping), and some method to provide access to the geographic data* "

Reference: Global Spatial Data Infrastructure (2004)

SDI is more than the technical infrastructure (*Reference:* Federal Geographic Data Committee 1997) for the delivery or sharing of spatial enabled data and it also includes the institutional arrangements that facilitate the availability, access and interpretation of the data. It calls for the capacity of the end user to be able to find, evaluate and use the data. Knowledge on personal and institutional levels is the key requirements here. Awareness of the existence of the infrastructure and technical knowledge to be able to use the information delivered in a "meaningful" manner. If interoperability was a wide based concept of institutional co-operation then SDI is a narrower worked example aimed at all users of spatial referenced data. It has specifics aspects (data, metadata, search visualization and evaluation) and it presumes a high level of awareness of end-users and those parties tasked with providing the users with data. It is this dynamic which the author wishes to scrutinize and to try and asses how well we are of the developments and associated issues.

1.3 The background to SDI:

Before detailing the different levels of SDI it is instructive to cite some examples of SDI initiatives and briefly explain their origin. The list of examples in Table 1.1 represents a cross section of government, regional and private SDI initiatives. One can state that a lot of disparate parties, on a whole series of organisational levels, are now investing time, money and energies into creating the conditions where SDI becomes a reality. The importance of this endeavour is underlined by the fact that a special session of the United Nations General assembly (in 1997) identified the need for geographic information to support decision making in trying to realise the goals set out in Agenda 21 of the Rio Earth summit in 1992. The number and the spread of the initiatives clearly indicate that this is an important development.

Table1.1: Sample overview of SDI programmes

Country / Body	**Name of SDI programme**
USA	NSDI (1994, Executive Order 12906)
Australia	ASDI
Malaysia	NaLIS
South Africa	NSIF
Portugal	SNIG
Ireland	ISDI (2004)
EU	INSPIRE (2004)
UK	NGDF (1995)
Netherlands	Ruimte voor Geo (circa 2001)
International Consortium	OGC (1994 as Open GIS Project)
International Consortium	GSDCI (1998)

Some of these projects have been in existence for some ten years or longer (the dates above are indicative in character), thereby proving the durability of concept and this will later be used as an important temporal marker when we come to examine our knowledge of SDI. The concept of SDI has been with us for some time but has yet to mature through all development stages and an examination now of the state of knowledge is both a timely and instructive indicator as to where SDI is in it development stage. Measured shortcomings now will go some way in explaining how SDI develops and may in be instructive identifying where we can expect SDI to go astray.

1.4 The different levels of SDI:

In the preceding pages one established the relationship between SDI and interoperability. One then went on to examine some examples of SDI initiatives; it now befalls one to examine SDI more closely. As with a physical infrastructure, SDI has different distinguishable levels moving form the global to the localised examples. While SDI is still developing one theory regarding the structure is shown in Figure 1.2.

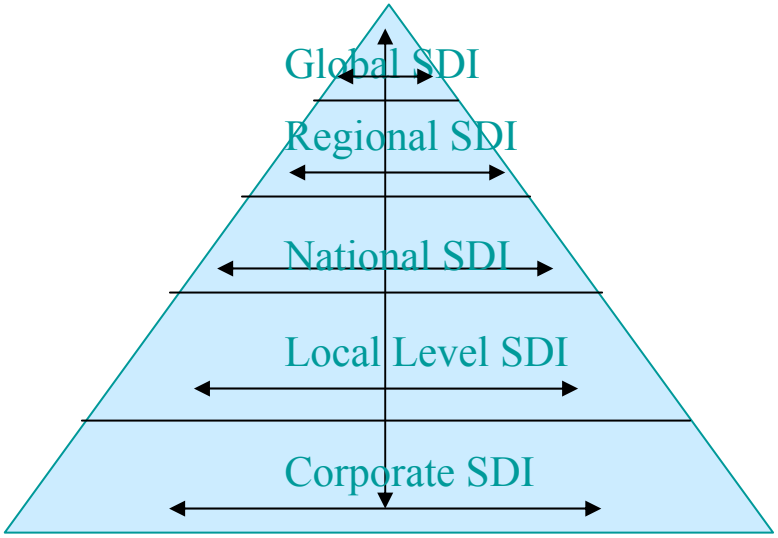


Figure 1.2: The SDI hierarchy

After: Bregt et al (2004)

At the apex of the triangle one see the GSDI, which was defined in the 2nd GSDI Conference in 1997 as

"..The policies, organisational remits, data, technologies , standards, delivery mechanisms and financial and human resources necessary to ensure that those working at the global and regional scale are not impeded in meeting their objectives.."

Reference: GSDI, *(2004) page 141*

This is followed by the regional SDI which is envisioned to allow for cooperation on a regional basis between countries and organisations. The GSDI Cookbook tell us –

"Mindful of the critical social, environmental, and economic issues shared regionally and often globally, the assurance of a Global Spatial Data Infrastructure to enable cooperating nations and organisations to collaborate on issues and solutions is extremely important."

The pyramid continues with national level, moving on to the local level where cooperation occurs between organisations within the same national boarders and lastly the corporate SDI, where employees have access to shared GIS resources. When it comes time later to examine the state of knowledge of SDI, one shall test for the presence or absence of treatment of these different strata.

1.5 The promise of interoperability:

Before going any further it is necessary to ask the question what is the point in doing research into a SDI? Why is SDI worth investigating? In order to answer these questions one needs to examine briefly the economic, political, legal and technical considerations which together form a persuasive defence for the examination of the SDI. Each of these aspects shall be dealt with separately in the coming pages.

Economic benefits of interoperability:

In his testimony to the US senate hearings on "The Role of Standards in the Growth of Global Electronic Commerce"(October 1999), A.B. Whinston (Director, Centre for Research in Electronic Commerce, and Professor of MSIS, Economics and Computer Science The University of Texas at Austin) stated the economic benefits of interoperability were as follows

-Interoperability is one of the key ingredients that allow consumers to substitute one product with another that is manufactured by a different company. This substitutability enhances competition among various manufacturers in the same product market.
-This substitutability and interchange ability implies larger market size, lower unit costs, and lower consumer prices.
-In addition to larger market size, interoperability and standardization enable new market entrants to tap into existing product users. This translates into lowered barriers to entry, further enhancing market competition.
-Interoperability and standardization allow process automation, lowering transaction costs.

Reference: Whinston A.B. (October 1999).

While it is difficult to quantify the benefits of interoperability, research has been carried out to gain insight into what the **lack** of interoperability costs economies. Research carried out in the US by Michael P. Gallagher, Director Technology Economics and Policy Analysis, Research Triangle Institute (NY) for the automotive industry in the United States estimate that

"imperfect interoperability imposes at least $1 billion per year on the members of the U.S. automotive supply chain"
-
Reference: Gallagher M.P.(2001)

The report continues with the following clarification that -
"By far, the greatest component of these costs is the resources devoted to repairing or re-entering data files that are not usable for downstream applications"

Creating a SDI is clearly a costly task but the case for a costs benefit analysis can be ascertained by the EU's own SDI program INSPIRE. While drafting the present EU Directive for INSPIRE the authors estimated that -

"The required investment of the preferred option - a focused framework backed by an EU framework Directive - will to a large extent be borne by the public sector and is estimated at an average of €3.6–5.4 million per annum per EU Member State (EU25). This would represent only 1% of the total expenditure on spatial information."

Reference: EU Commission (2004)

One can conclude from this that the costs of establishing a SDI is a fraction of the existing costs incurred by current (inefficient) practises in spatial data collation. The proposal went on further to state that -

"The benefits include environmental gains, wider social benefits and gains by the private sector. Only the environmental benefits have been quantified. The average annual benefits per Member State (EU25) amount to €27-42 million. Knowing that these elements only represent a partial view of the whole picture, the conclusion is that the benefits outweigh the investment requirements by a considerable amount."

The cost savings outlined so far are only relevant to those parties which **already** use spatial data. What does interoperability mean to current non-users of spatial data? By creating conditions where data is made freely or more cheaply available (including the means to query and view the data) the SDI initiatives create an opportunity to increase the size of the pool of users through the removal of technical and financial barriers. One of the reasons why the use of GIS in the US is ahead of the rest is that the latter has long held the policy that government data should be made freely available to the tax payers who pay for it.*(Reference Albert K. Yeung,1999)*

Interoperability of software and data, for example through internet browsers and GML, coupled with geo-portals are moving GIS outside of the sphere specialized technical section into the IT mainstream and the general populace where it belongs (Google Earth being a good example). This is a benefit of SDI which is very difficult to express in dollars or Euros but the results speaks for themselves.

The political need for interoperability:

The emergence of political and economic globalism has created the need for (inter)national government and non-governmental agencies to co-operate on different levels. In practise this means the exchange information across national borders on a whole range of issues e.g. environmental, social and security, and SDI represents one of these process's to establish the technologies and intellectual infrastructure to facilitate this cooperation. The European Union is an example, through the creation of the INSPIRE programme the Union has sent a clear signal to all member states that in order for the European Union to continue to expand and to achieve it's policy aims there is a need to harmonize national technical spatial data infrastructures. In their proposal to European parliament the authors of the Inspire directive stated the problem as follows

•*" Few Member States have developed a framework for establishing a national infrastructure for spatial information that addresses operational, organisational and legal issues. Where steps have been taken; they have often been restricted to specific regions or specific sectors.*
• *In most Member States where a framework has been adopted, not all problems have been addressed or initiatives are not compatible.*
• *Without a harmonised framework at Community level, the formulation, implementation, monitoring and evaluation of national and Community policies that directly or indirectly affect the environment will be hindered by the barriers to exploiting the cross-border spatial data needed for policies which address problems with a cross-border spatial dimension."*

Reference: EU Commission (2004)

The INSPIRE directive is only one of a series of EU Directives which underline the need to create a situation where public information is being re-used. The relationship between this directive and previous ones was specifically stated as follows -

"This Directive should be without prejudice to Directive 2003/98/EC of the European Parliament and of the Council of 17 November 2003 on the re-use of public sector Information, the objectives of which are complementary to those of this Directive."

The establishment of the EU SDI initiative INSPIRE is an indication of the political need for interoperability. The EU has chosen to given the development an explicit legal context *obliging* member states to establish a standard SDI

"Member States shall bring into force the laws, regulations and administrative provisions necessary to comply with this Directive by [2 years after the date of entry into force]. They shall forthwith communicate the Commission the text of those provisions and a correlation table between those provisions and this Directive."

The motivation behind the legal framework was further underlined as follows -

"Article 175(1) of the EC Treaty is the appropriate legal basis because the spatial data that fall within its scope are needed to support the formulation, implementation, monitoring and evaluation of environmental policies with a view to ensuring a high level of environmental protection. Furthermore, Article 174 requires the Community to take account of available scientific and technical data. INSPIRE contributes to the implementation of this requirement by helping the Community to access and use available spatial data."

It is crucial to understand that SDI is not about creating new spatial data sets but rather an infrastructure for existing heterogeneous spatial data sets from a wide range of policy areas to be coupled to one another.

On a national level the Dutch government have recently sought to confirm it's commitment to a national SDI in it's programme "Ruimte voor Geo"(Space for Geo), on 1 March 2004 the government accepted the findings of the van Wijzen Commission and have honoured the proposal to the tune of €20 million. This couple with another government initiative "Basis Registraties" (a programme to establish six basic data registers for all government activities) hold out the possibility that interoperability and SDI in particular will become a reality in the coming years.

These examples underline the point that not only are these progressive initiatives, they are taking on an increasing **legal context** as the establishment of SDI is gaining (in Europe at least) *(Reference: EUROGI, 2000)* a legal **obligatory** character. Our legislators have now recognized that geographic information is vital to sound decision making process at local, regional, national and international levels and we all have a vested interest in getting SDI right. The infrastructure which we are obliged to build has to contain -
"The component elements of those infrastructures shall include metadata, spatial data sets and spatial data services; network services and technologies; agreements on sharing, access and use; and coordination and monitoring mechanisms, processes and procedures."

Reference: EU Commission (2004)

This means that as GI professional body we need to be knowledgeable in these components and shortcomings in these areas have serious ramifications and this is why the knowledge base needs to be tested.

Technical benefits to SDI:

Having addressed the wider benefits of SDI, one must now focus on the technical advantages of realising a SDI through the use of geo-spatially enriched information. There are a great many advantages of spatial enriched data over traditional alpha- numeric data types. There is nothing new in that but what does SDI offer in terms of technical benefits to GI users? The following is a short overview of some of the more obvious technical benefits

Table1.2. Overview of some of the technical benefits of a SDI

Nr.	**Description**
1.	Offers an organising principle for disparate data i.e. geography.
2.	Scaling up of the Geospatial value chain.
3.	Solving of interoperability in area of supported geometry types.
4.	Solving of interoperability in area of supported data types.
5	Solving of interoperability in area of supported spatial operations.
6.	Making open source GIS programming a reality.
7.	Creates the physical infrastructure for data mining, searching, delivering and interpreting GI.
8	Enable groups to efficiently manage the semantics (or feature schema mismatches) of their own geo-data collections and get maximum benefit from each other's geo-data collections, despite semantic differences.
9.	Encourages software vendors to take into account all of the above when they design software interfaces.

Note: Compiled from different OGC source documents.

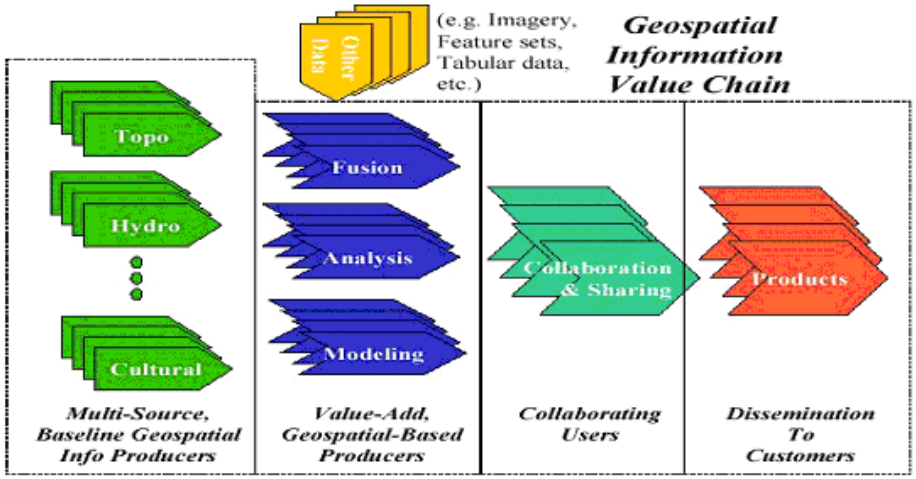
Of the above, number 1 is somewhat of a misnomer, GIS already offers the advantage of an organising principle in the structuring of data so why is it mentioned? Simply put, SDI offers the same old advantage to **more** people and organisations. This offers the potential for a fundamental shift in how ordinary (non-GIS) users can change how they work and it is worth recapping the basic principle of using geography to organise data sets. The OGC in their "OGC Reference Model" (2003, page 5) state that

"Geospatial location is a foundational property for modelling the world in a coherent, intuitive way. Location and time can be exploited as a unifying theme to better understand the context of most real and abstract phenomena. Location is contextually simple and intuitive to most people."

SDI, if successful, can create the conditions whereby the IT mainstream can begin to use data in a way which the GIS has been using for years and because it is a multi-facetted initiative it can help to remove more than any single barrier. It is this characteristic multi dimensional approach which promises an explosion in the users of geo-information in the coming decades. No single technical benefit in itself can do this but the range of solutions at the heart of each SDI can. In a nutshell SDI offers us the same technical benefits but **more** of them as it enhances our capacities to capture, interpret and manipulate (creating new data in the process) geo-information form sources outside our direct work environment. By creating these possibilities it also allows us to integrate alpha-numeric data with spatial data which we might not otherwise be able to do so.

The Open Geospatial Consortium has illustrated this advantage which geospatial data represents through the concept of the Geospatial Information Value Chain. This is represented in here.

Figure 1.3: The Geospatial Information Chain



Geospatial information sources enters an interoperable environment, which then pass through geo-processing chains, creating what the OGC refer to as "intermediate **value-add geospatial-based products "** along the way including **fusion** (combining, correlating, annotating, and
interrelating geospatial information from many sources into a single structure) and **analysis** (operating on geospatial information for the purpose of deriving new information, extracting results or understanding its nature and significance). Lastly **finished products** are created that contain geospatial information, or are derived from geospatial information, for both internal and external customers.

So some of the technical benefits of SDI are in themselves not revolutionary, the Value Chain illustrated above is already available but it does benefit us in the **scale** of expansion of current value chains.

And what of the remaining advantages of SDI , in Table 1.2 , numbers 3-5 are significant in that they represent possible solutions to some of the oldest barriers to interoperability – namely the obstacles of being unable to seamlessly integrate data form various sources (for example CAD and GIS) without incurring data quality penalties. The latter has traditionally

occurred due to lack of interoperability between the basic file types, supported geometry and the operations that an end user can carry out on them. Solve these problems and one has the basis for corporate, local and regional SDI (see Figure 1.2) by sharing existing spatial data and opens the door to coupling these with traditional non-spatial enabled data sets.

Following on form this point, number 7 and 8 (Table 1.2) alludes to the creation of the intellectual and technical infrastructure to support this sort of enterprise. By creating the infrastructure one can (in theory at least) use the internet to search, discover and download data from other parties. One can then find, interpret, integrate and manipulate this data into ones own data set and allowing other to do the same with your data. *(Reference: DGIWG 2000)*

Open source programming in IT has become a reality, Linux being one of the best known examples. SDI is seeking to produce the same improvement in the GIS world (number 6 in Table 1.2 above) – holding out the possibility that one can move away from expensive proprietary software into share-ware environment, reducing costs and lowering the finical barrier to using GI data in business operations.

1.6 Why is SDI important now?

Having laid out the groundwork regarding the developments and the benefits of SDI, one might ask the question why is SDI important now? Why do we need a general knowledge appraisal?
There is no guarantee that SDI will be the hoped for success, however there are signs at this point in time that at least some of the goals have been reached. Apart from the sheer number of SDI initiatives which have been started, there have been a series of tangible results which indicate that the process is on course to achieve significant gains. It is because a series of miles stones have been reached that now is an opportune time to examine SDI.

Examples of these are the incorporation by ISO of parts of OGC initiatives into new standards, ISO 19107 contains much of OGC 's Simple feature abstract, ISO 19115 owns a debt to OGC 's Topic c 11 metadata and Topic 12 – the openGIS service architecture is the same as ISO 19119. Of the original OGC seventeen Abstracts, three now have ISO specified standards. When an international standard body like ISO does this it is indicative of the gains being made in the SDI sphere. It is the intention of OGC to continue to build on this partnership in the future.

Another feature of the developments which needs to be alluded to here is the use by the OGC of fast-tracking of proposals into test environments before completion of the long winded validation procedures. This is of particular relevance to the developments of Abstracts for the use of GIS related internet technologies. The rate of technological developments is considerable faster than the slow pace of the human organisations which use these technologies. By proving the concepts in a technical environment one can move the validation process along more quickly. These are technical milestones (proof of concept) which help to create de facto standards for web mapping which might have not conform to a rigid rational approach but they are practical and in the fast moving world of ICT speed is of the essence.

Another reason why SDI is of current interest is the changes taking place within the broader ICT world. The movement away from PC's back to server side applications means organisations are increasingly seeking to implement corporate SID solutions as opposed to creating yet another island of proprietary based data and software for small group of users. This offers us a golden opportunity to couple and harness the advancements by migrating to an IT architecture whereby the issues of interoperability can be built into the migratory strategy. The author is a member of an inter-provincial work group for GIS and of the twelve provinces, all are at this moment addressing issues of interoperability within their organisations .With the resultant larger group of users, one's raises the stakes even more when considering the need to have constant updates of data. The greater the lag time between collation and distribution places significant pressures on the GIS departments to ensure that the user sees an accurate digital reality. This requires organisations to find cheaper ways of realising updates and cost cutting while increasing the level of service and that can only be made possible through interoperability and cooperation with other organisations.

1.7 Is ignorance the real barrier to interoperability?

Finally we come to the crux of the problem, the author has given a brief overview of the developments pertaining to SDI and where does that leave us?

Simply stated we are at an important technical and organisational crossroads. In the face of significant change it is important to take stock of where one wants to go and what exactly is on offer (in this case in SDI). The professional body needs to be cognitive of these changes and attempt to understand the implications of them.
The goal of this research is to measure how well the GI professional is aware of the changes taking place and attempt to answer the question which this thesis sets itself**, will ignorance be the real barrier to interoperability?**

The above question has been translated into a working hypothesis and the basis thereof have been set out here below -

1.8 Formulation of hypotheses:

This amounts to an examination of SDI and identifying those elements which one thinks needs to be included to see if this hypothesis is borne out in the research phase. The working hypothesis is as follows -

It is this author's contention that one of the barriers to interoperability is the ignorance of GIS professionals of the developments, in understanding the issues involved, in appreciating the costs and benefits and finally the legal framework within which SDI will have to be realised.

This single hypothesis can be further detailed into a series of questions which flow from it

- Are individual GI professionals being **sufficiently well informed** in these issues? Is there an actual communication problem?
- Is there a widespread ignorance of SDI? With all the attention given to the development to SDI in the last few years how **aware is the GIS professional** of the developments in 2005?
- Are there key weaknesses in our intellectual understanding which impair interoperability with particular attention to the two abstract characteristics mentioned i.e. semantics and metadata. Are these been dealt with in the broad professional arena or are they ignored?
- Do GI professionals realize the financial **benefits and costs** to their own organizations and partners in sharing data?
- Do GI professionals have any realization of the possible legal **consequences of failing** to create a SDI? Are they even aware that there may be legal consequences involved in not taking steps towards interoperability?

The hypothesis is based on a series of premises these are as follows -

- That it is possible to measure the degree of intellectual evolution /interest in the developments by measuring the coverage given to the subject in the press.
- This research can reveal the answers to these questions.
- Informal education, through the professional press, as opposed to formal education is the best way to test these questions.

1.9 Extracted predictions:

In addition to the stated hypothesis there are a number of expectations which the researcher feels can be measured and these are -

- That a knowledge profile of the issues can be measured over time.
- The knowledge profile should be increasing over time and the depth of the coverage over time should also be deepening.
- The treatment of SDI materials will tend to emphasise the technical aspects of SDI without alluding enough to the wider institutional context.
- A difference in the amount and type of coverage should make itself felt in the differing target audiences.
- The legal progress will be noted but the positive or negative consequences thereof will not be made explicitly clear.
- The economic costs and benefits will be alluded to, but not detailed.
- Education of professionals (formally and informally) on some kernel points (data quality and metadata) will be lacking thereby negatively affecting the developments.

Chapter 2 will be used to detail how this hypothesis was arrived by detailing the how the research was designed.

Chapter 2
Methodology and research design

2.0 Introduction:

One must now turn one's attention to the subject of how best to design one's research. What this amounts to in fact is the following -

- The establishment of what aspects of SDI one will examine in the research.
- The establishment of the suitability of available methodologies.

The first process forms the initial part of this chapter and then one must clearly set out the research design before proceeding with the actual research. The validity of the entire research is dependant on the logical consistency of this design and the author has detailed the design here. This detailing should enable others to decide if the design is adequate to the purpose or not. In the second process the abstract characteristics and their operational (observable) forms will be set out. The synthesis of a working hypothesis will be detailed with a series of questions which the research needs to be able to answer.

The structure of this chapter reflects the use of research design choice matrix (see Figure 2.1, page2-7). As the latter is a choice matrix only those parts which the author deems relevant will be dealt with here. Please note as the author has used Dutch as well as English language references, all quotes will be given in their original language and thereby remaining true to the original source.

2.1 Scope of the research:

In the previous chapter the case for examining the state of knowledge on SDI was placed firmly in the forefront of the argument. Before proceeding further however it is now necessary to begin to define the scope of the research on a number of different levels.
The hypothesis contains two areas where the scope needs to be detailed. Firstly the scope of the interoperability and SDI need to be narrowed down one needs to identify how the information is being disseminated and to whom.

Interoperability:

One has already defined the relationship between interoperability and SDI, it now is imperative to clearly establish which elements of these are pertinent to the research. The choice of methodology will be influenced by these. It is the authors contention that at this point the examination of interoperability per se was only relevant in establishing the broad framework of the research and as this has now be dealt with it will **not be dealt with further**.

SDI: Metadata and semantics

At the hearth of SDI is the concept of sharing information and this endeavour has two distinct aspects, the first is technological the other is epistemological. It is **not** within the scope of this research to investigate technical (read software) black box interoperability issues. The author has neither the qualifications nor means to deal with these. The research does address the epistemological aspects of SDI namely the areas associated with **metadata and semantics** as these are crucial in determining well we can find, bind and interpret spatial data .The research is primarily designed to examine these two aspects of SDI .This does not exclude the inclusion of other constituents of SDI (see Table 2.1), these will surface in the research data, but they are not the primary area of interest. These areas of interest are leading when choosing the abstract characteristics later on.

Table 2.1.: Common constituents of SDI

Nr.	**Description**
1.	Geospatial data development includes feature definitions, geometry, collections and relationships between features. Covers both vector based as well as raster based spatial data.
2.	Metadata – describing geospatial data.
3.	Geospatial data catalogue and all associated services. Includes semantics and Information Communities.
4.	Geospatial data visualisations. Includes coordinate referencing systems, stored functions and interpolations.
5.	Geospatial data access and delivery. Includes mobile and web services.

After: Global Spatial Data infrastructure, 2004

2.2 Method of knowledge propagation and to whom?

When examining the means of knowledge propagation one has two types of sources, those found in formal education and those in informal channels. Typically the former can be found in academic GI textbooks, course syllabus material and software manuals while the latter is to be found in the specialised GI press, internet e-zines and newsgroups.

As SDI is not just a collection of technologies but also the policies and institutional arrangements that facilitate access to data. The source to be used will have to be able to be dynamic, reflect changing content relative to emerging technical developments, deal with combinations of technologies and reflect the shifting political, economic and institutional context in which SDI develops. It also needs to have a high profile and be known to the GI professional as a source of information on all these aspects.

So SDI has very specific contextual dimensions of which the GI professional needs to aware of. Textbooks tend to have a modular approach to subject matter whereby the context gets overlooked. If SDI were just a combination of technologies and one wanted to measure what GI professionals should know about them then a GI textbook would be a sufficient source, however technologies change rapidly and textbooks do not always change as quickly. An examination of "An Introduction to Geographical Information Systems" by Heywood, Cornelius and Carver, Prentice Hall, 1998, "Principles of geographical Information systems" by Burrough and Mc Donnell, Oxford University Press, 1998 and "GIS – A computing Perspective" by Worboys, Taylor Francis Ltd, 1995 revealed no references to SDI and given their published dates this is not strange as they proceed or coincide with many of the start-up dates of SDI initiatives (see Table 1.1).

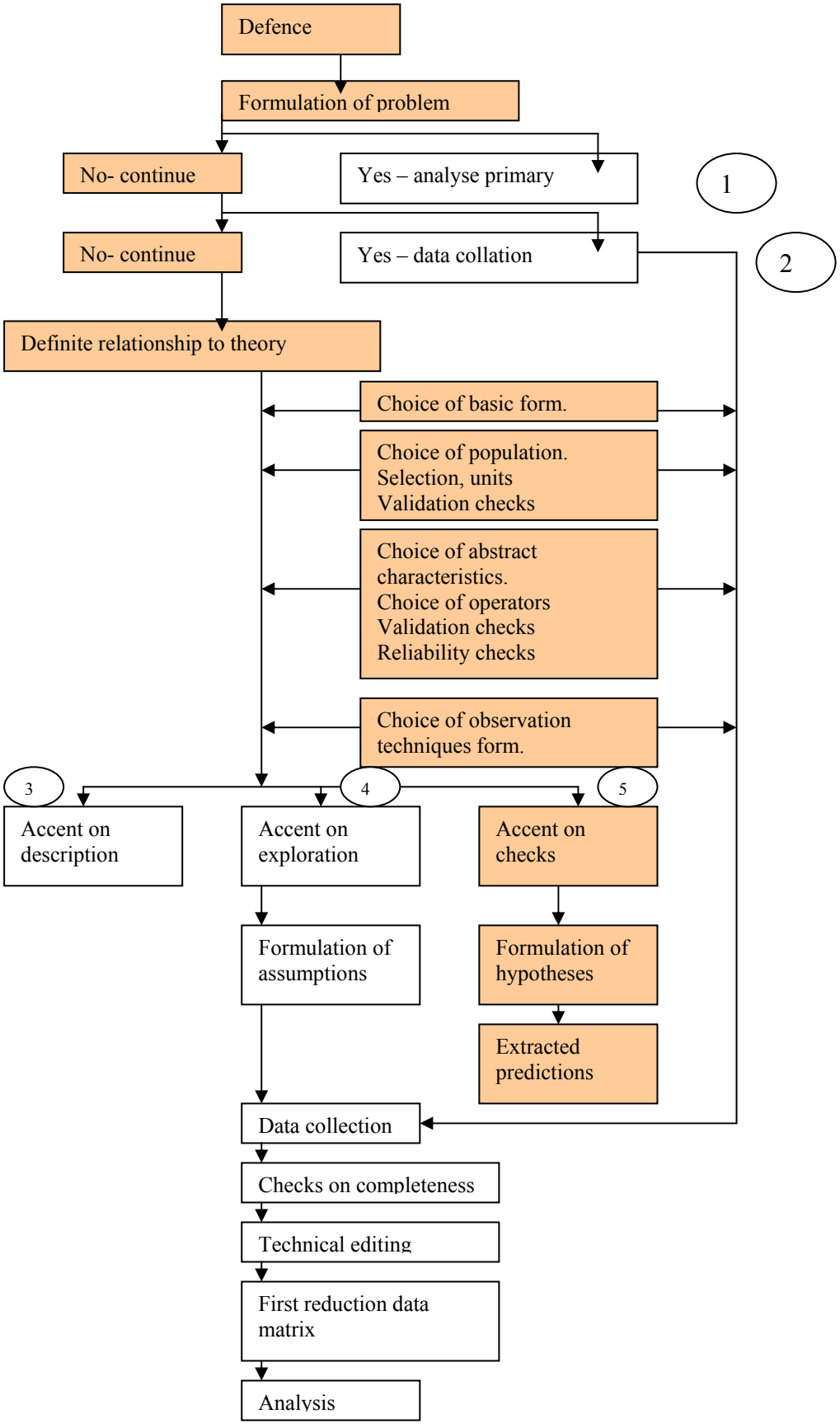
It is the author contention that the specialised GI press is the best source to test for the propagation of knowledge of SDI as it dynamic, multi facetted and is geared towards the GI professional community. As the subject matter is specifically GI biased, one can safely assume that it is the GI professional body which represent the target audience. It is the author belief that the penetration of GI into the general population has not yet reached levels where it would be profitable to examine for the coverage of the subject matter in say, management publications or consumer computing magazines. This situation is likely to change rapidly in the coming few years. In the conclusions to this research one shall come back to this point as a shift in the target audience out of the small circle of GI professionals would be indicative of the success of SDI.

2.3 Choice of methodologies:

As the purpose of an SDI is to increase and improve the capabilities of user groups to exchange information across organisations and IT infrastructures the primary type of research pertains to people, organisations, their information needs and lastly the technical specifications of their data . One must look to sociology for a reference framework for the kinds of methodologies which need to be applied to this research.

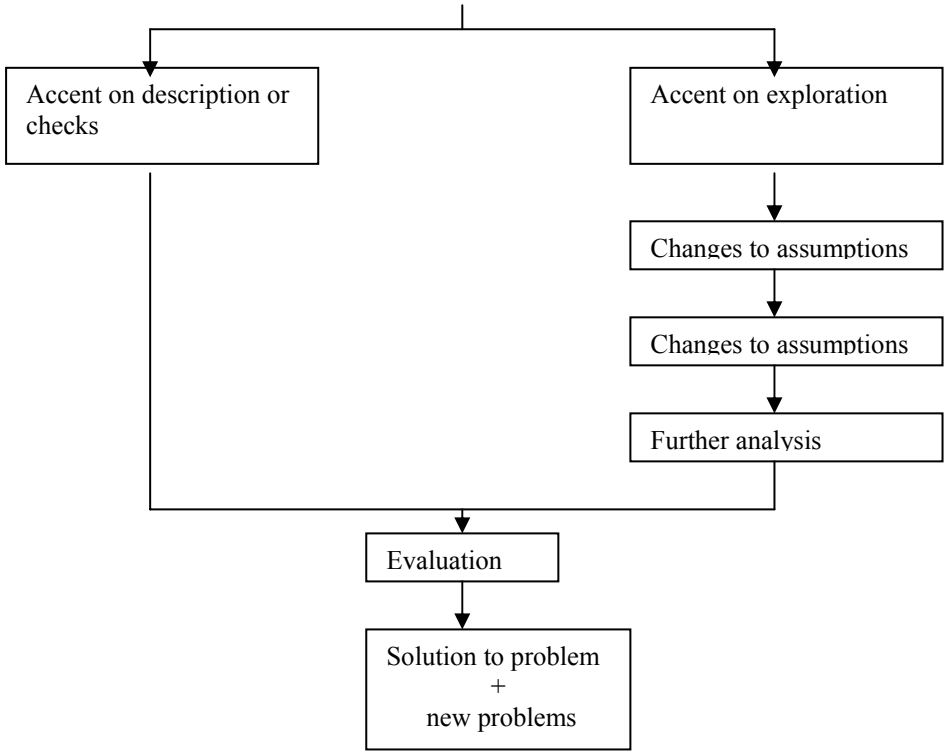
One has examined the structures of sociological research in order to establish a scientific methodology to apply to the survey. One of the best known sociology researchers in the Netherlands is P.G. Swanborn whose works in the field of applicable research methodologies has become a de facto standard. His book "Aspecten van sociologisch onderzoek" (1971), has been used to identify a course to be taken when designing one's research. This methodology has been set out in Figure 2.1 and an examination of these steps and the considerations implicit in them, will form the basis for the defence of the approach applied to the SDI research.

Figure 2.1 Choice matrix for sociological research



(Continued overleaf)

Figure 2.1 Choice matrix for sociological research (continued)



After Swanborn (p.5 ,1971).

Having already formulated the problem and defined the motivation behind the research one must now use Swanborn's matrix in order to arrive at a research methodology and formulate a hypothesis. This means that the next step is to examine the primary sources and assess if they are sufficient to answer the question posed.

2.4 Use of primary sources:

Examination of primary sources can only reveal what GI professional are **supposed** to know. Primary sources cannot tell us what people actually know, but they can tell us what one should be testing for. They will form a theoretical basis for the areas of interest and this point will be developed further in the choice of the abstract characteristics.

Swanborns matrix asks if the primary sources are enough, relative to the problem definition they are not. The primary sources of SDI are not formally stated but Table 2.2 is an overview of some of these and as such they formed the theoretical source the author used. They are a compilation of standards, norms and theoretical guidelines composed by different national and international professional bodies and to the author mind represent a sound theoretical source. However the researcher is interested in assessing what the GI professional body actually knows and the sources are not suitable enough in themselves in establishing that. By examining the means of propagation of these ideas through the specialised press one can get a measure of what is know of these primary sources.

Some of the most relevant primary sources for SDI are set out in Table 2.2 below.

Table 2.2: Some primary references for interoperability

Primary references	***Remarks***
OGC,- abstract specifications	Currently being complied .Various parts are completed .There are instances where OGC specifications have been merged with ISO standards.
TC211 ISO standards,	Currently being complied .Various parts are completed.
GSDI Cookbook	A framework for a global spatial data infrastructure.
EU standards (CEN)	At the moment the INSPIRE project is compiling standards but is not yet complete.
NEN norms	Dutch national standard NEN 3610 (Terrain model) will be superseded by TC211. Reaching a conclusion.
Dutch SDI	"Ruimte voor geo informatie"-in the process of compilation, activities being directed by RAVI.
NSDI	The American federal government's initiative. Active on a wide range of issues .Is an on going project.

If one were to leave aside the reservations about the validity of the approach i.e. equating international SDI initiatives and standards as a form of theory, we can, from primary sources identify the elements which constitute a SDI. The primary sources do help in establishing the dimensions of the subject matter and that it is the treatment of these subjects in the media which will allow one to answer one's hypothesis.

Following on from above one must decide whether or not the problem is definite in concrete terms (Figure 2.1, number 2, page 24). Can we describe the problem in terms of description i.e. surveys, maps and demographics etc.? Yes, because observations can be quantified and to

a degree qualified. One could at this juncture argue that the course to take would be to follow the arrow down from number 2 (Figure 2.1) and begin directly with data collection. However as SDI is a relatively new concept and the theoretical sources not explicitly defined, it would be somewhat irresponsible to simply proceed. It is the author's contention that the findings of the research can better validated if one where to take the time out to clearly define the nature of the research to the theory and to define methodologies applied .i.e. basic form, population, validation checks etc. This exercise will set out the assumptions and preconditions used in the research.

2.5 Define the relationship to the theory:

Before stating the basic form to be used one must first establish a recognised definition of theory and then to define the relationship to the theory. What is a theory?
Swanborn defines a theory as -

"-opgebouwd uit een min of meer samenhangend stel uitspraken over een bepaald gedeelte van de werkelijkheid""

Reference: Swanborn P.G. (1987) 9th edition, page 25

So theory is a collection of coherent pronunciations about an aspect of reality, and these individual pronunciations are called hypothesis or proposition. Hypothesis are linked to characteristics (or variables) and the relationship between them. These characteristics can fit a continuum from abstract to the concrete. These concrete characteristics are called operational characteristics as these are observable and thus quantifiable either through their absence or presence and form thus part of an empirical reality and can be verified by others. An abstract characteristic is a generalised class wherein the operational characteristics fit (Figure 2.2, page 33).

The three goals of theory is to either describe, explain/explore or predict/validate reality It is important at this point to establish which of the three goals are most applicable to the task in hand as these goals have very different research routes which need to be followed.

The goal of the theory of this paper is essentially to predict or validate the levels of knowledge of the constituents of SDI. The research will clarify reality rather than explore it as one has not set out to question the composition of SDI but rather our knowledge thereof. From the schema (Figure 2.1, page 24) this puts this work firmly in path Number 5 and is therefore decisive in determining the path which will be taken from here on out.

2.6 Choice of basic form:

At this juncture having decided the relationship to the theory one must examine the available forms of research and choose which of these are applicable. The available forms are as follows

Table 2.3: Basic forms of research

Type	***Description***
Experiment	Lab, field, natural.
Questionnaire	Single moment uptake photos, multi-moment uptake.
Content analysis	Qualitative analysis.
Case study	Examination of a small number (or individual) cases in detail.
Statistical Analysis	Quantitative analysis using various recognised forms of statistical analysis.
Secondary analysis	

-After P.G. Swanborn, p.47

The choice of the basic form is influenced by the goal of the research. This means from the available forms the experiment can be ruled out as it is a form most associated with exploratory research. Of the remaining forms available a number are suitable either in isolation or in combination with one another.

Before choosing the form one needs to be able to defend the choice based on a number of methodological issues -

- generalisation
- casual or narrative analysis
- the nature of the theory
- authenticity and authority

After: Gomm et al,(2002) page 6

To examine levels of knowledge one must have a basic form which is robust in terms of generalisation, in other words based on the research sample one will have to be able to formulate a series of conclusions which stand up to scrutiny beyond the limit of the sample set. Gomm et al, tells us that -

".. the aim is to draw, or to provide a basis for drawing, conclusions about some general type of phenomena or about members of a wider population of cases."

Reference:Gomm et al,(2002) page 5

These points to a series of options, either the case study or the survey (Swamborn calls the latter a questionnaire).

Casual or narrative analysis addresses a number of limits which should be considered, two of which are the issues of context and temporal limits. Both the survey and the case study have the advantage of investigating casual processes "in the real world" rather than in artificially created settings. This in turn validates the generalisation process as the conditions of testing are the same as those for which we are interested.

The context does not only have to be authentic but there also needs to be attentions paid to the depth of coverage. Is there enough information in the data to allow one to draw conclusions which are valid? The case study is associated with what is called "thick description" where the context is extensively described. However a media survey as proposed has in its article a built in description of the subject matter and given the different sorts of articles (editorials, newsletters and in depth articles) one can refer to this material as the thick description of the theory. The larger the selection set used in the media survey the thicker the description.

Of the two choices available, the case study would appear to offer the best solution to the temporal problem, that the findings need to be validated over time. In a case study one has the options of studying the phenomena in depth over a long period of time and that the findings are not prejudiced by the record of a single point in time. However the proposition is to use the specialised press as a basis for measuring of knowledge and that means there is a written record which can be examined not only for content but for content over time. These time stamps can be used to test for measurable changes over specific periods thereby validating further the findings.

The nature of the theory has been identified as one which will predict reality, the latter will be formalised and the choice of form needs to be able one allow one to make measurable predictions. This latter requirement is important as the thick description of case studies are based on qualitative analysis while the survey, through the use of statistical analysis of the results will allow one to back up one's predictions with a quantitative analysis. The use of the time stamps mentioned above will also further validate the prediction process by quantifying changes over time whereby one can make new predictions based on past performances and trends.

The issues of authentic and authority needs to be considered, -
" ..research is advocated on the basis that it can capture the unique character of a person, situation, group and so on. Here there may be no concern with typicality in relation to a category, or generalizibility to a population. The aim is to represent the case authentically"

Reference: Gomm et al, (2002) page 6

While the above quote refers specifically to the use of case studies it sums up the need to address the authenticity of the research data. While it is a virtue of the case study to represent the authenticity of the case it is not outside the reach of survey to do the same by the simple means of ensuring a broad sampling group, whereby the heterogeneous nature of the group (GI professionals in this case) is fairly represented by a range of magazine titles and authors.

Coupled to this last point is that if one is interested in assessing the knowledge levels of a GI professional body then the specialised press is one of the best ways to do this. In the Netherlands many of the magazine articles are written by or in association with ordinary GI specialists. Examination of this material has the natural authority of a localised group of these specialists.

Conclusions on the choice of the basic form:
It is not uncommon in research to combine research forms in order to establish a method best suited to the task in hand. One has examined the different forms relative to criteria and having considered these the author feels that the form which best suits the research is the survey, in this case a media survey backed up with some statistical analysis of the recorded finding. The extent of use of the latter will be detailed in Chapter 4 where the results will be presented. This choice offers the best compromise between a quantitative and qualitative approach to answering the hypothesis and it has sufficient merit to meet most of the criteria set out above.

2.7 Detailing of the basic form:

The following aspects needs to be assessed when detailing the research form-

- Population – character, sources, temporal dimension and the size of the sampling.
- Choice of abstract and operational characteristics.
- Selection of operational characteristics and their units – these are the observable units which will be measured when testing the hypothesis.
- Choice of observation techniques form.
- Validation checks – the basis upon which the research can examine the data for validity.

Together these will form the intellectual framework for the research which will be worked out in Chapter 3 into the Entity –Relationships model used for the research.

Population – the size of the sample groups:

Given the quantitative basis of the research form it is necessary to state some of the criteria for the sample size and data population characteristics in general.

The fit of the population -

- The author is interested assessing levels of knowledge or ignorance of SDI. This is by definition an international initiative but one which if it is to succeed needs to have definite national and localised dimensions. The media survey could quiet easily be based on English language as well as Dutch sources. This would mean extending generalisations beyond localised conditions and the validity thereof is difficult to prove. On the other hand the examination of the Dutch media will have not only localised coverage but should reflect European and international developments. It is the author's contention that by limiting the population of data to a Dutch language media one can best meet a balance between findings which can be safely generalised while still having enough diversity of influences whereby the raw data is representative of the state of knowledge of the GI professionals.
- The Dutch GI market is sufficiently large enough to sustain a number of independent publications to allow for the examination of four different publications. The author deems this to be sufficient in number to offer a degree of diversity.
- Internet or traditional published – the author made the decision to choose the traditional published press over the Internet. This was based on the premise that the latter still forms the basis for which most professionals use for keeping up to date with developments. Some of the titles used have been long been established and have a higher profile in the GI community than other purely Internet based titles. In addition all of the publications with the exception of one all publish their materials on the Internet after the appearance of the paper version so the Internet dimension is not completely ignored in the choice.
- As already stated temporal aspects are important characteristics of knowledge and ignorance because the latter are subject to change. It is important to be able to not just limit ones findings to a small point in time but to use the written record to extrapolate across time. To that end the author has selected data set which extends back over a period of five years in order to make developments over time appreciable, in effect creating a form of continuous observation over that period. In the conclusions to the research one shall return to the theme of knowledge and time.
- The actual size of the data set is set out in Chapter 4 as the first result. With all research where one carries out a series of experiments the validity of the conclusions rises or falls on the population of the tested data. If the test population is too small

then any conclusions drawn will be seriously flawed .The nature of the research will play an important role in establishing the size of a legitimate population. Research questions which have an explicit quantitative character needs to have a reciprocal large population size. Research which have a qualitative character need to have a population which reflects in terms of character those qualities which the research is interested in.

Choice of abstract characteristics:

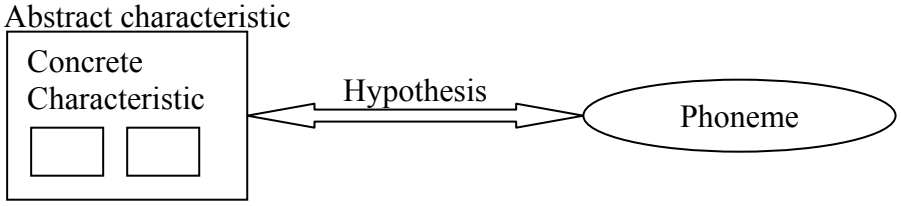
Before going on to establish a hypothesis one needs to re-iterate that a theory deals with characteristics and their relationships with one another. This begins by stating the abstract characteristics, their operational forms and the relationship between these. In the "Selection of operational characteristics units"(see below) the operational characteristics of a SDI will be stated. By re-examining Swanborn's methodology one can establish the relationship between operational and the abstract characteristics, he tell us -

> *"Kenmerken kunnen op een continuüm van meer abstract naar meer concreet geplaatst worden "*

(p.25 P.G. Swanborn, 1971)

In other words all characteristics can be aggregated from their concrete form into an abstract form .Table 2.4 is an overview of the abstract characteristics, they represent the generalities of the worldly phenomena which are the subject of the research thesis. It is worth stating the relationship between these parts –

Figure 2.2.: the relationship between theory, reality and characteristics.



- After Swanborn p.27

Table 2.4 The abstract characteristics and their subservient concrete characteristics

Abstract Characteristics	**Underlying concrete characteristics**
Metadata	Metadata Feature Geometry Spatial Referencing by Co-ordinates Catalog Services
Semantics	Semantics and information communities Features Relationships Between Features

It can be observed that both of the abstract characteristics have concrete characteristics which are the same as themselves. It is the author contention that the other concrete characteristics are needed in order to fully appreciation the *dimensions* of the abstract characteristics .In their concrete form neither of the abstracts are complete.

Selection of operational characteristic units:
In Table 2.1 the common constituents of SDI were set out, one has already stated that the areas of interest to the researchers are metadata and semantics together these from the abstract characteristics which need to be measured in order to carry out the proposed research.
The research will be examining the following themes as the basis of a SDI, each one will have specific subdivisions and the presence or absence of these subdivisions will form the basis of the selection units.
Table 2.5 Overview of main selection units

Title	**Description of sub divisions**
Feature Geometry	Same as ISO 19107 • Symbols • Geometry • Topology
Spatial Referencing by Co-ordinates	Describes modelling requirements for spatial referencing by co-ordinates.
Features	A feature object (in software) corresponds to a real world or abstract entity.
Relationships Between Features	Abstraction for the relationships between entities in the real world. This abstraction is modeled as relationships between the features
Metadata	ISO 19115 • Metadata for geographical data representation • Metadata package and entity relationships • Data dictionary • Hierarchical levels of metadata • Data quality information
Catalog Services	Covers the Geospatial Information Access Services
Semantics and Information Communities	• Model for semantics and information communities. • Geospatail information community • Abstract model for semantics

-After the OGC “Overview of topics”.

Choice of observation techniques form:
The choice of the observation technique as already has been stated as a desktop survey of media literature in this area. This media survey needs to be designed in such as way as to ensure the objectively of the results. To that end, one needs to set out a series of guidelines which needs to be adhered to when designing the observation technique. The goal of the desktop survey is to gain a perspective into how well the core characteristics have been covered. Therefore the (specialised GI) news media will be examined for the prevalence of articles and the depth of treatment thereof; these can be further worked out into the following observations which need to be recorded.

Table 2.6 Desktop survey design criteria

Part	**Description**
Temporal Indicator	Date of the article
Distribution	Number of copies sold
Quantitative treatment	Number of SDI articles
Qualitative treatment	Depth and adherence to concrete characteristics
Prevalence and importance	Is the material re-covered on a regular basis i.e. a fixed feature, and the depth of coverage.
Readership Target audience	The target audience of magazine titles expressed in academic levels.

In Chapter 3 these points will be fully worked out into a coding proposal where the observation technique has been developed into a data model for coding of the articles.

Validation / reliability checks:
The survey shall be carried out on a series of magazine titles which are in the public domain, the validation of the results can be checked by the reader by virtue of reading the articles recorded and checking these with the results set out here. The validity of the results is then based on the combination of the article content, the methodology set out here and in the following chapter and the readers own interpretation of the source material.

The actual coding of the data is subject to a whole series of technical data controls and these were carried out and documented but the lack of space prohibits detailling them here or in the appendix. They represent a check on the validity of the values recorded in all the critical fields for the coding of the articles.

Accent on checks:
This chapter began by stating that one needed to follow a recognised methodology in order to design one's research .The path set out in Figure 2.1 have led to the conclusion that the nature of the research is to clarify theory and that the accents should be on the checking of the theory in the primary sources. This is number 5 in the aforementioned choice matrix. At this juncture on must now fulfil the last two steps before the research can begin, namely the formulation of a hypothesis and the extracted predictions.

2.10 Conclusion:

The purpose of this chapter was to establish the theoretical boundaries of the research, to set out how the research design was arrived at and to prove how the hypothesis was arrived at, even though the latter was set out at the end of chapter 1. By applying a recognised system one has a coherent and defensible position on all three counts.
How does one proceed from here? At this juncture it is necessary to establish exactly where we are at in the process -

1. The defence of the basic subject matter (SDI) has been made.
2. A hypothesis has been formulated around this area of research
3. A series of measurable expectations relative to the research have been established.
4. An area of research pertinent to the subject matter has been identified.
5. The scope of the research has been defined.

The rest of the process which needs to be developed in the following chapters is -

- The research methodology needs to be worked out in detail.
- The actual research needs to be carried out.
- The results of the research need to be detailed relative to the measurable expectations.
- A series of conclusions and recommendations need to be formulated.

Chapter 3
The coding of the magazine survey

3.0 Introduction:

Having decided to carry out a survey of the coverage of the subject matter in the specialised GIS press one must now establish how exactly this survey will be constructed. In this chapter the author will detail the choice of observable attributes, the subject matter one is searching for, the range of the values and measurement scales. This forms the basis of the Entity – Relationship model which was used as a basis to create a full data model for use in MS Access. The description will continue with some worked examples, and finally explain how the results were structured.

Please note a step by step description of how the survey was actually carried out is in Appendix I. It is worth mentioning at this point a pilot survey was carried out with a smaller population data set to test the methodology. The opportunity was used to establish the validity checks and the data controls for the coded data. The data model is set out in Appendix III.

3.1 The choice of observable attributes:

In Chapter 2 (see "Choice of observation techniques form" page 35) the basic attributes which need to be recorded about the articles were named. It falls upon the author to create a coding schematic in which these different attributes will be related to one another and how they will be used in the research. On first viewing it is clear that a clear distinction needs to be made between the attributes which pertain to the magazine and those pertaining to the actual article. From Table 3.1 one can state that criteria 2 and 5 apply to the former while the rest pertain to the latter.

Table 3.1 Observable attributes

Nr.	General attribute	Specific form(s) of the attribute.
1	Temporal Indicator	Date of the article.
2	Distribution	Number of copies sold.
3	Quantitative treatment	Number of articles containing SDI subject material.
4	Qualitative treatment	Character of treatment. Prevalence and importance.
5	Readership Target audience	The academic level which the editorial team approach in their readership.

Each will be treated individually so that the motivation behind the use of that specific attribute will be clear to the reader.

Temporal Indicator:

This attribute will be used to plot potential changes in the coverage of the subject material over time. This reduces the limitations of the survey by increasing the length of the duration of the observation. In addition one will have to establish the relevance of the article relative to developments and changes in SDI.

On first examination it would appear the older the article the less relevant it is to SDI developments. On the other hand one is not judging the articles as such, but one is judging the effect the article have had in raising the potential awareness of GIS professionals to SDI related issues so the older the article the greater the better. How these qualities will be coded and weighted will be dealt with later in this document.

Distribution

Given that the goal of the desktop survey is to establish the potential awareness of SDI related subjects the greater the circulation of the magazine in which the articles appear the better, thus this is a measure which needs to be recognized in the any potential coding. How these qualities will be coded and weighted will be dealt with later in this document.

Quantitative treatment of the subject matter:
With this attribute one shall quantify the number of articles which contain subject material which is of interest to us. The survey is a means of observing the operational forms of the abstract characteristics and it is these which the author needs to quantify. In Table 2.4 one has already defined the relationship between these two. These were further detailed to a level specific enough to be used as magazine article subject material (see Table 2.5). Doing so allows for better categorization of the articles and reduces the subjectivity of the coding process. These descriptions of the sub divisions in the latter table represent the bottom level of the data hierarchy. The following is an overview of the abstract characteristics and their concrete forms, which needs to be identified in the magazine articles. These represent the subject material to be sought out.

Table 3.2: The subject matter sought

SDI subject Criteria	**Abstract Characteristics**	**Operational characteristics**	**Description of sub divisions**
A1	Metadata	Feature Geometry	Symbols
A2	Metadata	Feature Geometry	Geometry
A3	Metadata	Feature Geometry	Topology
A4	Metadata	Spatial Referencing by Co-ordinates	Describes modeling requirements for Spatial referencing by co-ordinates.
B1	Semantics	Features	A feature object (in software) Corresponds to a real world or abstract entity.
B2	Semantics	Relationships Between Features	Abstraction for the relationships between entities in the real world. This abstraction is modeled as relationships between the features
A5	Metadata	Metadata	ISO 19115 Metadata for geographical data representation
A6	Metadata	Metadata	Metadata package and entity relationships
A7	Metadata	Metadata	Data dictionary
A8	Metadata	Metadata	Hierarchical levels of metadata
A9	Metadata	Metadata	Data quality information
A10	Metadata	Catalogue Services	Covers the Geospatial Information Access Services
B3	Semantics	Semantics and Information Communities	Model for semantics and information communities
B4	Semantics	Semantics and Information Communities	Geospatail information community
B5	Semantics	Semantics and Information Communities	Abstract model for semantics
	SDI general		Experience gained during the pilot pointed towards the generalised nature of many articles and the reciprocal need to have a SDI general attribute.

How these qualities will be coded and weighted will be dealt with later in this document.

Qualitative treatment

The desktop survey will not only have to reveal the extent or amount of coverage of the abstract characters of SDI but will have to also have to indicate the quality of the coverage. There are two levels upon which articles can be tested for, namely the character of the treatment and the prevalence and importance given by the editorial team to the subject matter. Each of these two will be dealt with through a series of key words / concepts which depending on their nature will be explicitly or implicitly taken into consideration when carrying out the survey. The former shall be coded as an attribute while the latter will not.

Character of the treatment:

This can be split up into four areas

- Article will be typed as being either reports of developments, editorials or features where subject matter in dealt with in depth.(See Table 3.6)
- Instructive nature of the article: If you read it what do you know about it in comparison to what you knew about it before hand.(implicit)
- Critical tone of the article: this cannot be coded but should be reflected in the former.(implicit)
- Use of references to web site or alternative sources of information, articles containing the latter represent a better source of information than those without references (see Table 3.7).

Prevalence and importance

This quality can be measured based on the whether an article is either a -

- front page and leading article, these will be scored higher than other articles.
- a secondary feature article, in other words neither a front page article nor a regular feature article,
- the article is a regular feature in the magazine title, newsletters and editorial are examples of this. Newsletters typically contain more than a single subject so each will be treated as an article but in order to reduce the statistical distortion that this creates they will be valued lower than either the front page article or the secondary feature article.

How these qualities will be coded and weighted will be dealt with later in this document.

Readership and Target audience:

The subject matter (SDI) is primarily the constituency of GIS and information managers therefore the readership level of the magazine will be influential in the quantitative and qualitative treatment of the subject matter. Magazines which seek to serve the technician as well as the university graduate are less likely to cover the material as often as a magazine title aimed specifically at the academic. How these qualities will be coded and weighted will be dealt with later in this document (see table 3.9).

3.2 Coding model and measurement scales:

Before building a coding model one must reiterate legitimate operations which can be carried out on collected data in relation to one another, if one does not establish these there will be major problems with the coding model in terms of internal consistency. The table here under is an overview of measurement scales and the valid operations which can be carried out on them.

Table 3.3: Overview of measurement scales and their operations

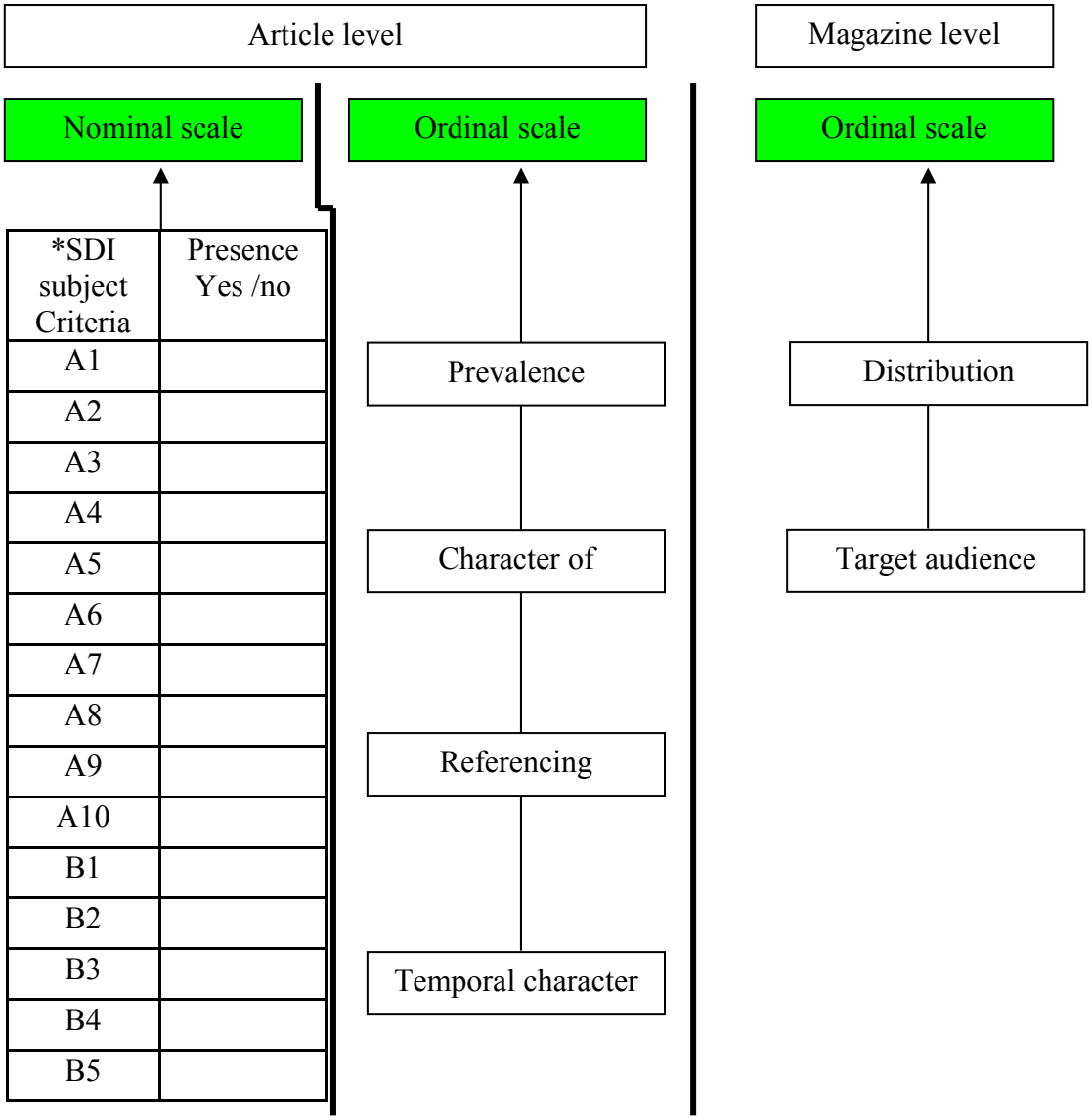
Measurement scales	**Examples**	**Valid Operations**
Nominal	Soil type Owner of a property Acceptable/unacceptable Presence / absence but you record only the presence.	Equals (one= another Frequencies can be calculated for each category Mode
Ordinal	Super profile value (standard of living) League position	All of the above and – < > Rank Median
Interval	Date Elevation Degrees Celsius	All of the above and + - Scaling by a constant Mean Variance
Ratio	Income Length Height Area Volume Population Degrees Kelvin	All of the above and X Division

Source: modified from McGuire and Goodchild. UNIGIS course material Module 6.

Each of the operational attributes set out in the preceding pages will have to be characterised to establish under which measurement scale they fall, then and only then, can one check for logical consistency of the coding system, both individually and collectively.

The relationship between the different criteria and their measurement scale can be demonstrated schematically as follows,

Figure 3.1 The survey attributes and their measurement scales



* See Table 3.2 for the full details

What this in effect means is that while attributes can be scored there are limits to the degree of aggregation of the scores across attributes. The major distinction is between attributes pertaining to an article and those pertaining to a magazine title. These cannot be combined into a single score as they exist on different levels. The end result of the scoring is a series of combinations of values which need to be examined on an individual level (see also the worked examples further up in this chapter). These limits will be referred to in Chapter 4 when presenting the results of the survey.

3.3 Coding of the criteria and their preconditions:

Before detailing the final relationship between these different attributes it is necessary to establish the range of values for each one individually. Each attribute has a series of conditions which will present themselves during the coding. These conditions will be used to assign the relevant value for each record. The descriptions of these are detailed in the tables below. The values will be used later as a means of scoring the raw data.

Table 3.4: SDI subject criteria

Conditi on nr.	Condition description	Value	Remarks
I	One of the 15 subjects **is not** present in the article.	0	The article is ignored. It will only be used in establishing the potential size of the data set in the results.
II	One of the 15 subjects **is** present in the article.	1	The article is registered and subject to further coding.

Note: The first of these is decisive in deciding inclusion or exclusion for consideration in the rest of the desktop survey. Having identified the presence of any of the above subject matter one can register a "hit". Articles which initially score a hit but upon examination are deemed not suitable are discarded and fall under condition I.

Table 3.5: Prevalence

Conditi on nr.	Condition description	Value
I	Leading article including front page (And conditions here under)	5
II	Feature article (2 or more full pages including graphics and tables and excluding advertisement space)	4
III	Feature article (1 page including graphics and tables and excluding advertisement space).	3
IV	Regular column (featured in each issue of the magazine for that calendar year). Length varies from half a page to a whole page.	2

Table 3.6: Character of the treatment

Condition nr.	Condition description	Value
I	*Reporting* – the information contained in the article simply refers to a development without dealing with the material in depth. The publishing of a standard or reporting of a conference. The litmus test is that having read the report the readers knows what event has occurred without having gained any insights in the subject matter. Ranks lower.	1
II	*Examination* – then coverage of the subject matter is extensive and in depth. That the author's intention was to deepen the reader's knowledge of the subject matter. Ranks higher.	2
III	*Editorials* – these are articles either written by the editor or by guest journalists. They are typically commentaries on the subject material.	1

Table 3.7: References

Condition nr.	Condition description	Value
I	Article **contains no** published references (including internet Sources).	1
II	Article **contains** published references (including internet sources).	2

Table 3.8: Distribution bandwidths

Condition nr.	Condition description	Value
I	< 1,000 copies in circulation in the year of publication of the article.	1
II	≥1001 ≤5, 000 copies in circulation in the year of publication of the article.	2
III	≥5001≤10, 000 copies in circulation in the year of publication of the article.	3
IV	≥ 10,001 copies in circulation in the year of publication of the article.	4

Note: The distribution of the magazine is a qualitative indicator in the establishing the potential penetration of the material into the GI community. The ranges of bandwidths reflect those of the four different magazine titles used in the research thereby allowing for sufficient meaningful variation in the results (these differ from the pilot).

Table 3.9: Target audience

Condition nr.	Condition description	Value
I	MBO	1
II	MBO/HBO	2
III	HBO+/ Academic	3
IV	Academic	4

These are based on the typical distinction levels of academic training in the Netherlands.

Table 3.10: Temporal character

Condition nr.	Condition description	Value
I	< 1 year old (0-12 months from the review date)	1
II	1-2 years old (13-24 months from the review date)	2
III	2-3 years old (25-36 months from the review date)	3
IV	3-4 years old (37-48 months from the review date)	4
V	4-5 years old (49-60 months from the review date)	5

Note: Value is higher the older the article is as it has more potential learning value i.e. the GIS manager has had more time to potentially learn about the subject matter. One could argue that the information in newer articles is perhaps more relevant but they should be also more frequent as SDI initiatives gather momentum (thereby be compensated).

3.4 Worked examples of the survey:

In order to help the reader understand how the material in the preceding pages was actually applied a series of hypothetical examples will be detailed here.

In the potential data set there are twelve articles, all were searched for the keyword. Of the twelve there were three found to contain the keywords and whose content are deemed fit for consideration. These were characterised and coded as follows,

Article A

- Article A deals with A1, A4, and A6 -m the development of new NEN 3610.
- It's the leading story for that issue.
- It's a feature article – 3 pages.
- It examines the issues as opposed to reporting the advancement of the project (development of a new 3610).
- It is featured in magazine with a distribution of 10,000.
- The target for the magazine is recognized as Academic, HBO+
- The article is two years old.

Article B

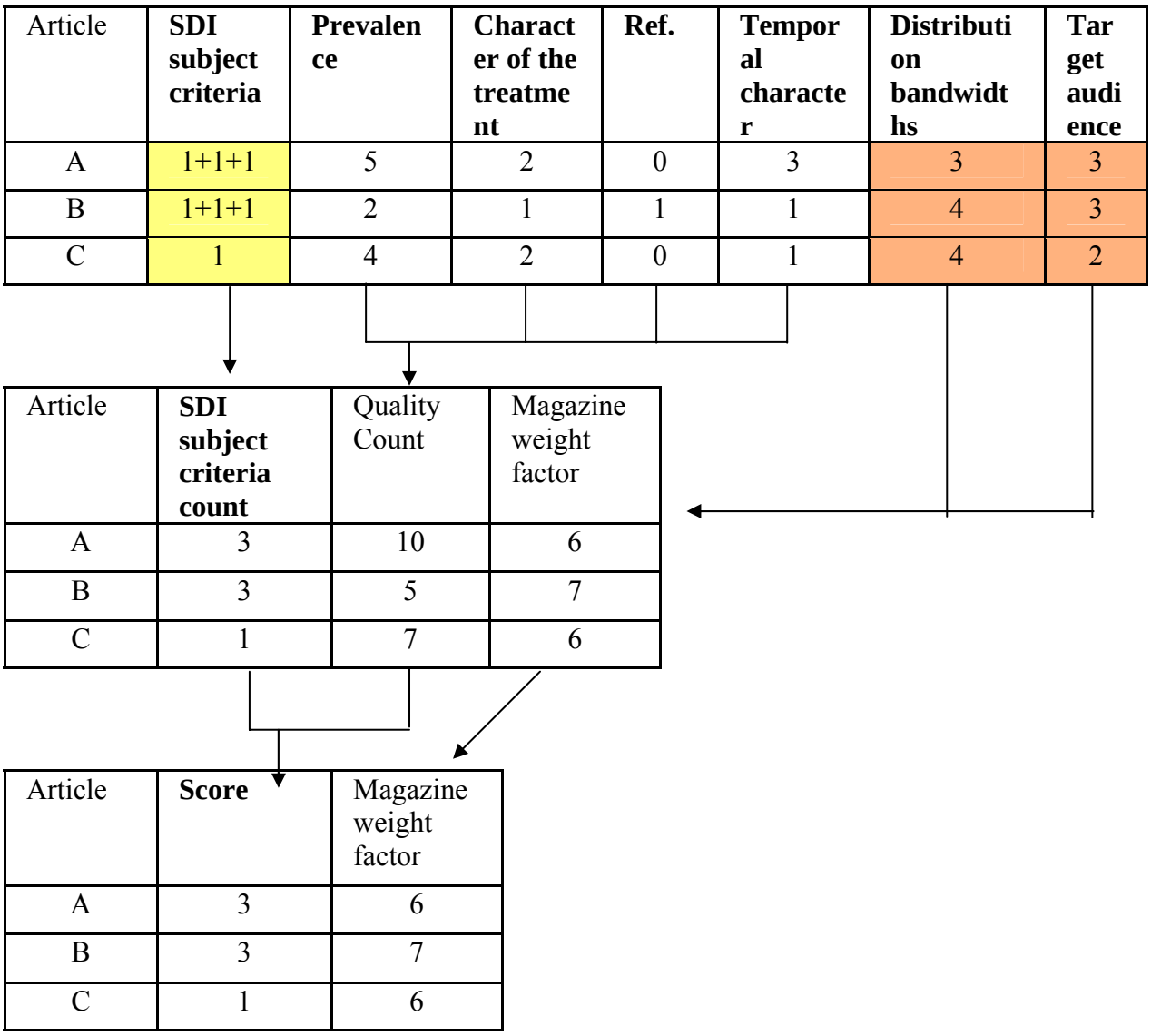
- Article B, deals with A1, A2, B1- the work carried out by OGC and ISO on new semantic translators.
- It's not leading story for that issue.
- It's a column – 1/2 page.
- It reports the advancement of the project.
- It contains references to a public web site of the work group.
- It is featured in magazine with a distribution of 20,000.
- The target for the magazine is recognized as Academic, HBO+
- The article is one year old.

Article C

- Article C deals with A2, the work carried out by researchers into sort of geometry, which is problematic in switching between CAD and GIS environments.
- It's not leading story for that issue.
- It's a feature article – 4 pages (including 1 page of diagrams).
- It is also the first in 4 articles.
- It deals critically with the issues of interoperability of geometry between CAD and GIS.
- It contains no references.
- It is featured in magazine with a distribution of 15,000.
- The target audience for the magazine is recognized as HBO/ MBO
- The article is 6 months old.

How does one rank these three articles relative to one another while using the above-mentioned system? The following is an overview of the raw data scoring of the articles.

Figure 3.2.The relationship between the scoring of the various attributes



Article	**SDI subject criteria**	**Prevalence**	**Character of the treatment**	**Ref.**	**Temporal character**	**Distribution bandwidths**	**Target audience**
A	1+1+1	5	2	0	3	3	3
B	1+1+1	2	1	1	1	4	3
C	1	4	2	0	1	4	2

Article	**SDI subject criteria count**	Quality Count	Magazine weight factor
A	3	10	6
B	3	5	7
C	1	7	6

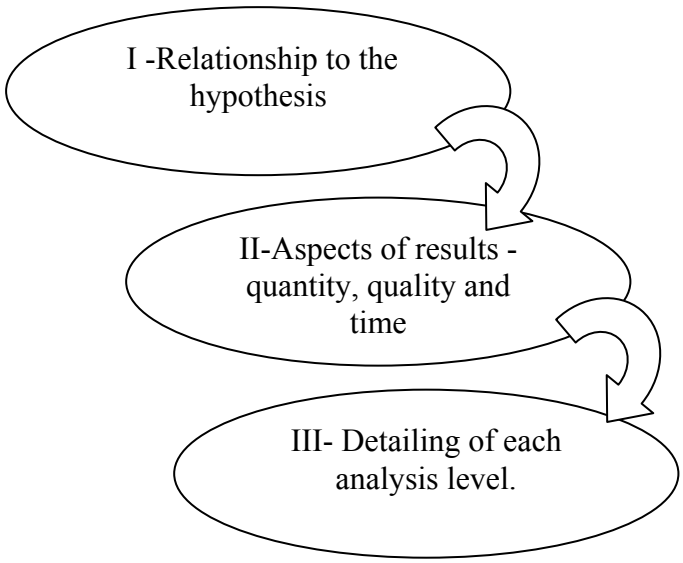
Article	**Score**	Magazine weight factor
A	3	6
B	3	7
C	1	6

This represents the final scoring of the articles.

3.5 Structure of the results analysis:

Before presenting the results in the next chapter one needs first to place the results in their proper context and then detail the structure of the analysis. There are a number issues involved but they can be best represented graphically as follows.

Figure 3.3: Issues involved in the structuring the results of research:



Relationship to the hypothesis:
The first requirement is to draw a clear relationship between the results and the original hypothesis. The relevance of the results to the overall questions being posed needs to be made clear at this juncture. To reiterate the data should allow the author to test the following questions.

Table 3.4: List of questions derived form the hypothesis

Nr.	Questions posed at the outset
1	Are individual GIS operatives being **sufficiently well informed** in these issues? Is there an actual communication problem?
2	Is there a widespread ignorance of SDI? With all the attention given to the development to SDI in the last few years how **aware is the GIS professional** of the developments in 2005?
3	Are there key weaknesses in our intellectual understanding which impair interoperability with particular attention to the two abstract characteristics mentioned i.e. semantics and metadata. Are these been dealt with in the broad professional arena or are they ignored?
4	Do GI professionals realize the financial **benefits and costs** to their own organizations and partners in sharing data?
5	Do GI professionals have any realization of the possible legal **consequences of failing** to create the basis if interoperability? Are they even aware that there may be legal consequences involved in not taking steps towards interoperability?

See: *Chapter 1 (page 17)*

An examination of the list reveals that the questions need to be spilt into those which can primarily be answered based on quantitative analysis of the results and those which require a combined quantitative and qualitative analysis. In the next chapter one shall set out all the available results in order to be able to answer these questions. Those which have a primarily quantitative basis shall form the majority of the results however those qualitative results which are necessary to answer questions will also be set out.

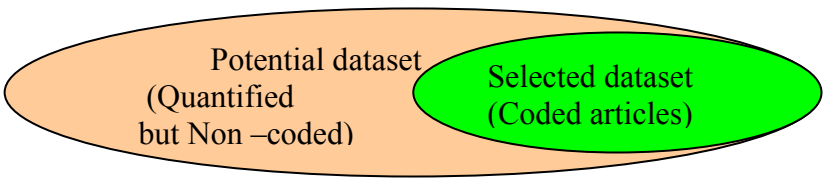
One can say that of the above six questions one can answer with certainty the first three based on a quantitative analysis. These hypotheses have measurable dimensions which through statistical analysis of the data one will be able to answer the questions posed. They will also have a qualitative aspect (many of which are bound up in the coding data model) and these will be dealt with briefly in the analysis.

Of the remaining two questions one can conclude that they have more ambiguous dimensions but they can still be answered based on the notes taken during the survey (stored in the remarks fields in the database). These are qualitative results whereby the comments on the articles content, will be used to answer the questions posed. They do have a quantitative basis in that the articles which contain the answers have a measurable frequency. Ultimately they represent the area where the conclusions are loosely based on the data and the author interpretation thereof.

When organizing the analysis if the results the author took the following three aspects into account for all of the results.

Quantity:
It should be stated clearly that the quantitative analysis has two distinct ranges, namely the results expressed in terms of the potential data set and those expressed in terms of the coded data set. The former is the size of the article test population without heed to their content. The latter represent the articles whose content was deemed relevant to further investigation and it are these articles which were recorded. The relationships between the two data sets are represented graphically as follows.

Figure 3.4



Qualitative aspects:
There are basically three aspects to qualitative analysis of the results

- Those qualitative aspects built into the coding data model and their resultant scores, counts and factors.
- Those qualitative aspects which do not fit the rigid structure of the data model but were recorded or indicated to in the comments fields (which was part of the data model).
- Those qualitative aspects based on the two above and the author's judgment of their value and article content.

Time:
The temporal character of the articles is in itself an important aspect to the overall research in that all the conclusions are based on the state of the data at that point in time. Both within the coding data model and in the time interval over which the total population spans (five years) the author has striven to create the possibility to check the results for changes over time.

One can state that as one moves through the questions the qualitative aspect will increase in domination while the quantitative level will remain stable until one gets into the treatment of the results for questions 4 and 5.

Chapter 4
Results and analysis of the survey data

4.1 Are individual GIS operatives being sufficiently well informed in these issues? Is there an actual communication problem?

In order to answer this question one needs to examine the results along the following lines

- What is the total size of test potential population?
- What is the distribution of articles coded relative to the potential data set?
- What is the distribution of SDI articles per magazine title?
- What is the distribution of magazines weight factors?
- What is the spread of the temporal coverage?

Establishing the size and character of the potential data set

It is imperative at the outset to establish the size of the potential data set. This is a function of the average number of articles per magazine title and the actual number of magazine issues used in the research. It forms the first basis of quantitative analysis as it represents the outer dimensions of what could be the coded data set.

Table 4.1: General details of the potential data set

Magazine title	Number of issues per year	Average number of articles per issue	Number issues tested	# articles
VI Matrix	8	12	24	288
Geo-Info	11	12	14	168
GIS magazine	8	19	14	266
Geo Nieuws	4	18	13	234
Total size of test population (in articles).				**956**

This result is important as all other conclusions drawn in the research are dependant on how representative the potential data set is in size and in character. The composition of this result is worth examination in detail, in terms of both the size and the character of the magazines titles used. A potential data set of nearly a thousand articles represents a substantial potential and is for the purposes of the research a sound statistical foundation. Three of the four titles have near parity in the numbers of issues tested. The fact that the other title (VI Matrix) had more tested is a reflection of the choice to include a five year old interval in the potential data set allowing the possibility to test for changes over time.

The spread of the values of the number of articles does show a spread of nearly 120 articles between the highest (VI Matrix, 288) and the lowest (Geo-info, 168) magazine titles. This represents a difference which needs to be borne in mind when testing other parameters as it will have distorting effect on the performance of a magazine title. However as has already stated the research is not primarily interested in how one title performs against another, the subject of the research is how the press performs in general and thereby acts as a surrogate to the entire GIS community in measuring knowledge levels.

Character:
By restricting ones data set to a single language group necessitated curiously enough the need for a bilingual search word list when searching for suitable articles as much of the technical jargon in articles are a mix of English and Dutch phrases. After the pilot this list, see Appendix II, was extended in an attempt to complete the Dutch variation of the English terms. How complete this list is will have to be judged by the reader as there is no exhaustive source for semantic variations of these subjects. The advantage of a single language situation is that one can see how well international developments are being translated into a regional context. One sees through out the article content American, European, Dutch and Flemish SDI initiatives being treated whereby one gets more diversity than one might otherwise find in say, an American English language data set. This composition reflects the SDI pyramid set out in Chapter 1 (Figure 1.2 page 8) where one moves from international to national, to regional and down to corporate SDI levels.

None of the titles have an obvious or stated affiliation with commercial GI market players and therefore the content is not unduly biased in favor of any software vendor. It should be noted that while some of the titles publish periodically a supplement from commercial vendors these supplements were not included for consideration in the research. This was also a criteria applied when choosing which titles would be included for testing. The market is awash with titles produced by software vendors whose primary purpose is not to inform GIS professionals but to advertise the vendor's product through informative articles.
Having read more than one's fair share of these titles over the years it was the author belief that these sorts of titles could be omitted from consideration as the market still had enough non- affiliated titles to chose from and their impartial nature would offer better source material.

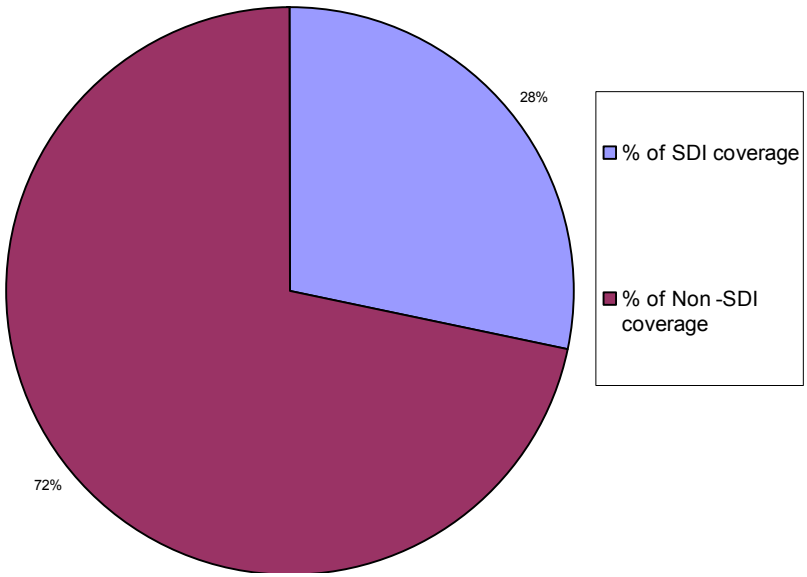
Distribution of articles coded relative to the potential data set
Actual number of articles which returned hits on the Search words and which on examination met the criteria for coding. By using digital copies of the articles (Adobe pdf. file format) one was able to use the Adobe Search routine to create a selection list of articles for consideration .This methodology guarantees uniformity in the selection phase and reduces the amount distortion due to human subjectivity in choosing relevant articles.

These represent the coded data set as opposed to the potential data set. It is on the basis of these articles that most of the results will be generated. It is worth pointing out that the articles which gave no returns for the search words or were found lacking substance were not coded.

Results:
Table 4.2: Distribution of articles coded relative to the potential data set

Total population size	956
Actual number of coded articles	270
% of SDI article coverage	**28**

Figure 4.1 : Overview of SDI coverage from the total article population



This is then a measure of how much GIS professionals are being informed of these issues. As the result is a percentage of the total coverage one can ascertain how important the editorial board of the magazine finds the broad issues of SDI to be for its readership. It allows the researcher to use this result to be able to answer in a quantitative form, the first question posed by the hypothesis.

The presence of 28 % of total number of articles dedicated to SDI related articles is a surprisingly high result. In the original pilot, testing with just one magazine over one year, the result was 30%.The size of the potential data set was then 132 articles with 44 in the coded data set. The result has hardly changed over a much larger data set which points to a general trend which is unaffected when scaled up.

Distribution of SDI articles per magazine title.
Having defined the potential and the actual data set in this overview one details the distribution of SDI articles within a magazine title. In other words the percentages named are relative to the total number of articles for that magazine title in the total data set. It is an indication of the importance lent by editorial teams to the subject of this research.

Table 4.3: Distribution of SDI articles per magazine title

Total # of articles	Total number coded	magazine title	% of SDI coverage relative to the size of the potential data set for that title.
168	50	Geo-Info	30
234	52	Geo Nieuws	22
266	40	GIS magazine	15
288	128	VI Matrix	44

The results presented show the breakdown per magazine of distribution of SDI articles in the different titles. It is instructive as the previous result was a compounded measure across the potential data set. We can see than that the 28% average is based on a range of coverage which varies between 15 - 44 %. Clearly there are significant differences amongst the titles in their coverage of SDI. The lower rate is somewhat disappointing reflecting a more limited editorial scope. The high VI Matrix result can be partially explained by the presence of newsletters in every issue from different organisations tasked with the realisation of either the Authentieke Registraties or the national SDI. This feature guarantees that returns would be high and of course there were more VI Matrix issues tested than any other title.

The distribution of magazines weight factors
Having established a series of parameters relative to the article coverage it is necessary to establish results pertaining to the magazines. In this view one details the distribution of the magazine titles and the target audience.

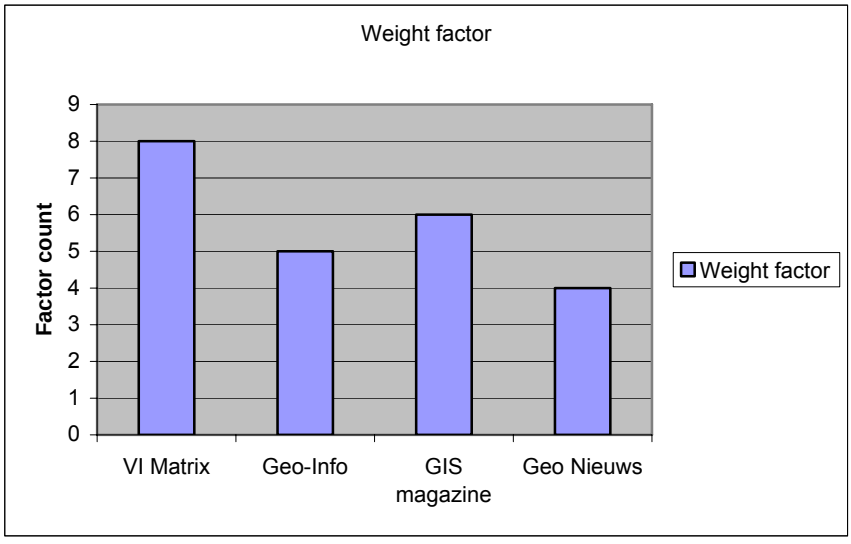
Table 4.4: The distribution per magazine title

Magazine title	Distribution bandwidth description	Target audience description
VI Matrix	Greater or equals 10,001 copies in circulation in the year of publication of the article.	Academic
Geo-Info	greater or equals 2001 less than or equals 5, 000 copies in circulation	HBO+/ Academic
GIS magazine	Greater or equals 5001less than or equals10, 000 copies in circulation.	HBO+/ Academic
Geo Nieuws	Lees than 2,000 copies in circulation in the year of publication of the article.	HBO+/ Academic

These two qualities combine to form the magazine weight factor in establishing the value of the articles. Given that this is a measure on a different level (magazine as opposed to article level) then no single definitive value can be aggregated out of both levels.
While the research is not interested in how well one magazine title performs against each other, in order to establish the value of the coded articles one does need create an overview of the distribution of the weight factors amongst the chosen titles.
Table 4.5.: Distribution of the magazine weight factors

Magazine title	Weight factor
VI Matrix	8
Geo-Info	5
GIS magazine	6
Geo Nieuws	4

Figure 4.2



Magazine distribution:
Obviously a magazine which produces high scoring articles but has a very small circulation is less informative to the general GIS community than a magazine title which scores average results but has a wide readership. The result reflects a large divergence on this account. The attribute value was based on information found in each of the titles and is not a guess by the author. The intervals in the classification where arbitrarily chosen but the subject titles represent one example form each class. This means that there is high level of variance in the results. The data (magazines) can be described as being representative of different prevailing conditions in the market. It is also worth noting that some titles are new to the market and that their distribution levels may well change with time.

Target Audience:
This was an attempt to recognize that the different groups engaged in GIS activities would be served by different magazine titles, that the target audience of a specific title was more geared towards one particular group than the other. The coverage of that title would tell us something about the penetration of the material through the professional community.
Titles which are aimed at the technician and have good SDI coverage would indicate that the material has gained a deep penetration and that SDI has reached the point of demassification. Target audiences were based on levels of academic training in the Netherlands. The latter are typically used when defining the differences in the composition of the workforce. In the GIS community one can expect the same spread of classification to be present from technician (MBO) to the professor (academic), as they are all participants in the GIS workflow.

There are problems with this approach, firstly none of the editors approached to give their definition of their target audience were willing to give an explicit answer to the question. The classification used where therefore the authors own interpretation of what he perceived to be the target audience based on the overall subject matter and the level of treatment thereof, moving from the abstract to the more concrete. In the end, all but one magazine was classed alike, the exception, VI matrix, in the author opinion is clearly geared to the academic. The nature of the articles tends to be less technical and more orientated to process management with an academic undertone. The results therefore have little variance and in fact the homogeneity present has the effect of negating the effect of these values in the weight factor.

The spread of the temporal coverage.
This is an overview of the distribution of the quantitative results as outlined above but with their temporal characteristics detailed. In should be borne in mind that most of the magazine date from the same period and that only two titles cover the temporal period outside that sample. This is an overview of the changes in terms of SDI coverage per magazine per year.

Table 4.7: Temporal distribution of articles

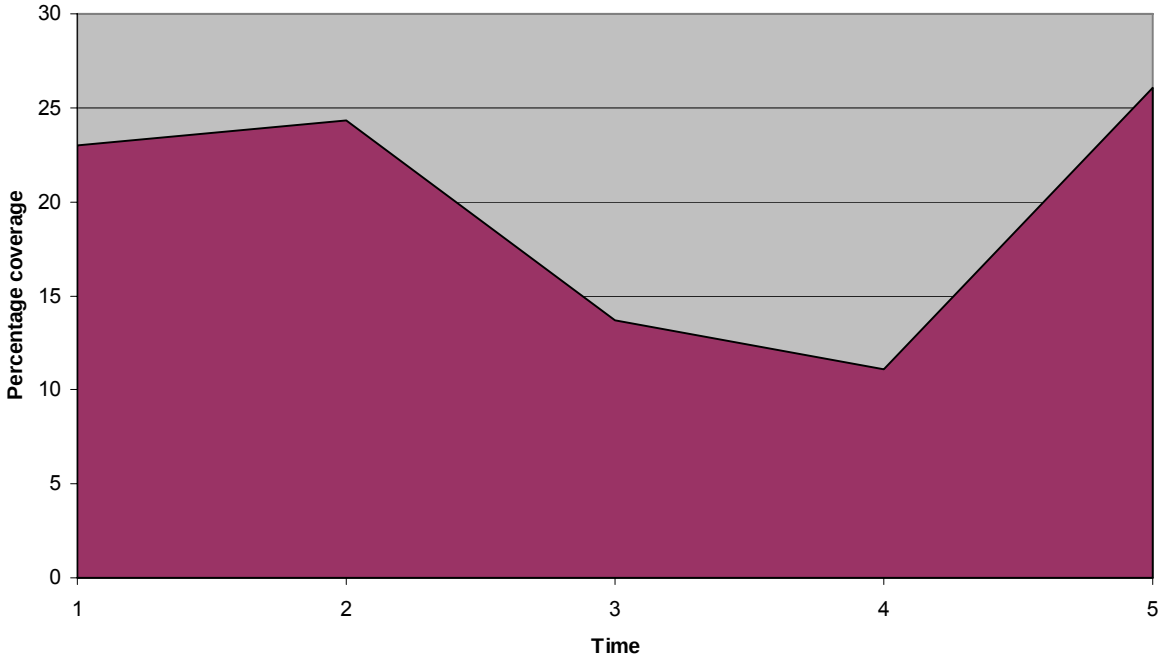
	Magazine title				
Temporal character description	**Geo-Info**	**VI Matrix**	**Geo Nieuws**	**GIS magazine**	**Total**
< 1 year old (0-12 months from the review date)	31	33	18	22	104
1-2 years old (13-24 months from the review date)	19	56	17	18	110
2-3 years old (25-36 months from the review date)		14	9		23
3-4 years old (37-48 months from the review date)			8		8
4-5 years old (49-60 months from the review date)		25			25
Total	**50**	**128**	**52**	**40**	**270**

If we are to express the results of Table 4.7 in terms of the percentage of distribution of coverage relative to the potential data set for that same year then the results are as follows -

Table 4.8: Distribution of actual coverage over time relative to the potential data set

Temporal condition	**Total coded**	**Total potential for that year**	**Percentage distribution**
< 1 year old (0-12 months from the review date)	104	452	23
1-2 years old (13-24 months from the review date)	110	452	24
2-3 years old (25-36 months from the review date)	23	168	14
3-4 years old (37-48 months from the review date)	8	72	11
4-5 years old (49-60 months from the review date)	25	96	26

Figure 4.3. Distribution of actual coverage over time relative to the potential data set



(See Table 4.8)

Examination of these results should lead us to be able to conclude if there was been a change of coverage of SDI over time and thereby indicate the development stage of the subject matter. Only the first two temporal categories can be tested for all the titles outside of that period there was insufficient coverage to test for all titles. This was due to the lack of material for some magazine titles for that period, as "GIS magazine" and "Geo Info" only appeared on the market in the last two and half years. The advantage in using new titles in the selection was to allow for possible innovations in editorial policy when compared to the established competition. It was the possibility of registering a shift in the sort of coverage which ensured the inclusion of these two titles.

Looking at the specific result the second oldest title "GeoNieuws" shows a definite increase over time of its coverage of the subject matter moving from a count of 8 articles three years ago to twice that in the present. For "VI Matrix" the coverage is less consistent from period to period but the fact that five years ago the subject matter was well represented (25 articles compared with 36 in the last year) indicates a **history** of coverage. If one were to use the OGC has a "date of birth" of SDI (1994) then there should be coverage in this temporal category.

At the conclusion of the results one can state that the GI profession is definitely being informed of the issues of SDI and that the press is doing its parts in terms of coverage of the subject.

4.2 Is there a widespread ignorance of SDI? With all the attention given to the development to SDI in the last few years how aware is the GIS professional of the developments in 2005?

In order to answer the question properly one must examine the depth of the coverage quantified in the preceding pages along the following lines -

- The spread of the article scores in general – this is a summation of the quantitative and qualitative values of the articles. Widespread ignorance will produce low scores.
- The spread of the articles across the magazine titles. The general trend that emerges here will be crucial in determining the state of knowledge across the GI community.
- The spread of the quality of treatment of articles scores.
- The spread of treatment over time.

In order to lend this process a degree of objectivity the researcher has detailed in chapter three a whole series of parameters upon which an article can be graded and coded. This will result in a combined score and it is this statistical result which will examined here. We do this by examining the article score followed up by examining the individual attributes which go to make up these scores. Without this detailed analysis we cannot state whether or not the levels of knowledge or ignorance are deep rooted or are temporary.

Article Score:
Before dealing with these it is worth repeating that an article score is composed of the following elements;

Score = SDI count * quality count

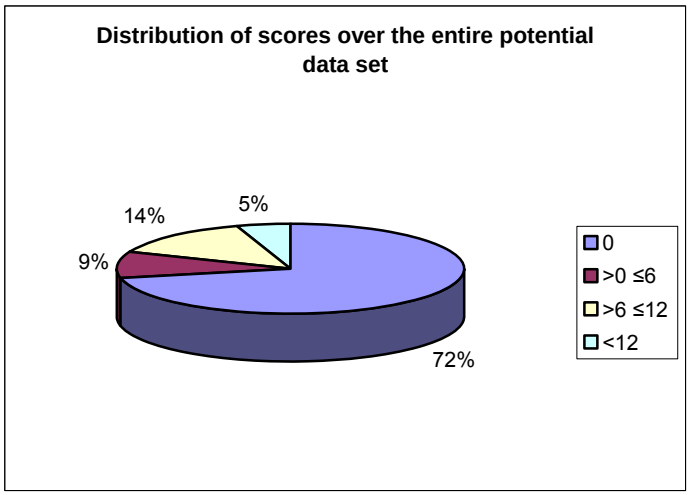
Where SDI count = Number of subject criteria dealt with in the article
Quality count = Prevalence+ Character of treatment + Presence of references + temporal character. The latter have all be given graduating scales so that there differences can be distinguished between the subjective idea of "good" coverage and "bad" coverage.

The articles scores were divided into three classes, the interval of these classifications (6) are based on results of the pilot (where the interval was 10)it was found that six gives a more representative spread of results without artificially distorting the underline trend. The first results presented are for the distribution of scores across the entire potential data set.

Table 4.9 Distribution of article score across the potential data set

Score class	# articles	% of Total potential dataset
0	686	*72*
>0 ≤6	89	9
>6 ≤12	133	14
<12	48	5
Total	**956**	**100**

Figure 4.4: Score distribution over 956 articles

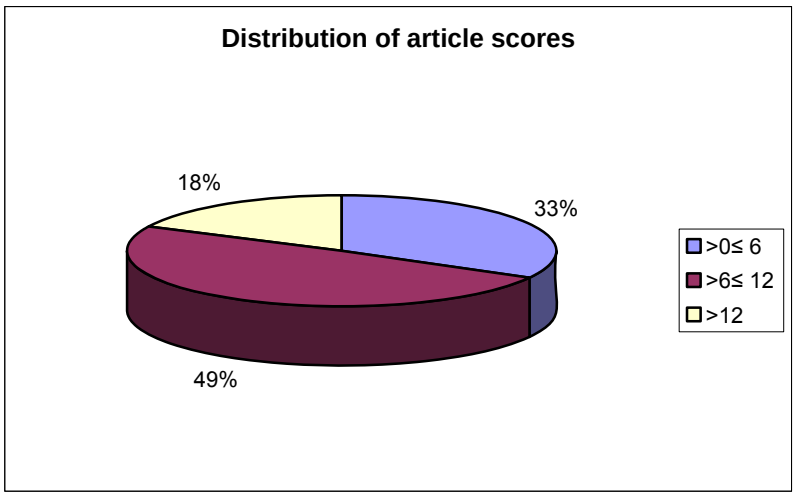


Article Score: continued
These are then same results but expressed in terms of the distribution of the aggregated scores of the *coded data* set.

Table 4.10: Distribution of the aggregated scores of the coded data set.

Score	# articles	% of coded dataset
>0≤ 6	89	33
>6≤ 12	133	49
>12	48	18
	270	**100**

Figure 4.5.: Distribution of the aggregated scores of the coded data set.



Here under one examines the same results but in terms of the distribution of the score within the magazine title. It serves to highlight differences in the way it is being treated through the four magazine titles.

Table 4.11: Distribution of the"Geo Info" scores

Score distribution	# articles	% of database
>0≤ 6	21	42
>6≤ 12	18	36
>12	11	22

Figure 4.6: Distribution of "Geo Info" article scores

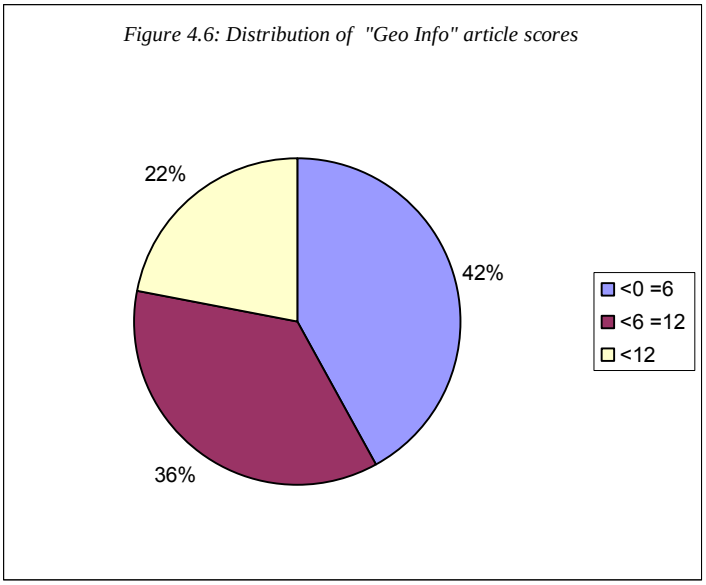


Table 4.12: Distribution of scores for "VI Matrix" articles

Score distribution	# articles	% of database
>0≤ 6	32	25
>6≤ 12	77	60
>12	19	15

Figure 4.7: Distribution of "VI Matrix" article scores

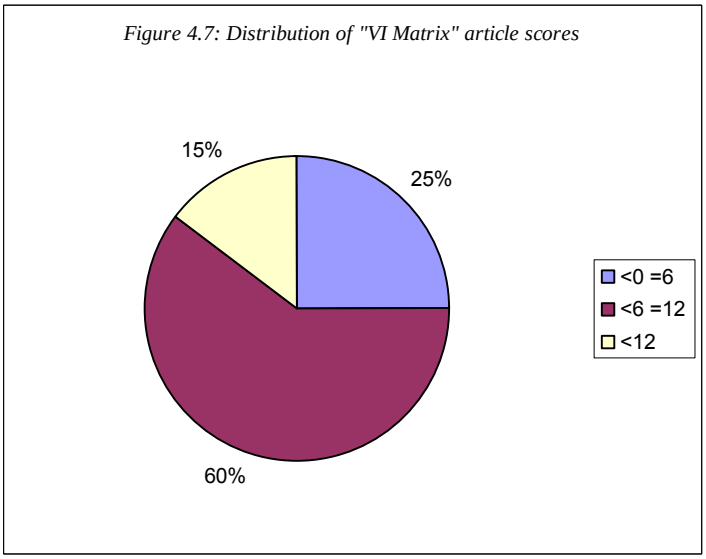


Table 4.13: Distribution of scores for "Geo Nieuws" articles

Score distribution	# articles	% of database
>0≤ 6	17	33
>6≤ 12	23	44
>12	12	23

Figure 4.8: Distribution of article scores Geo Nieuws

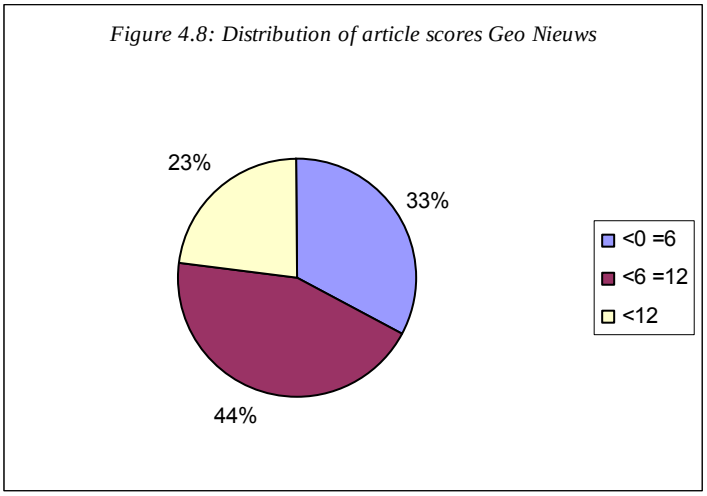
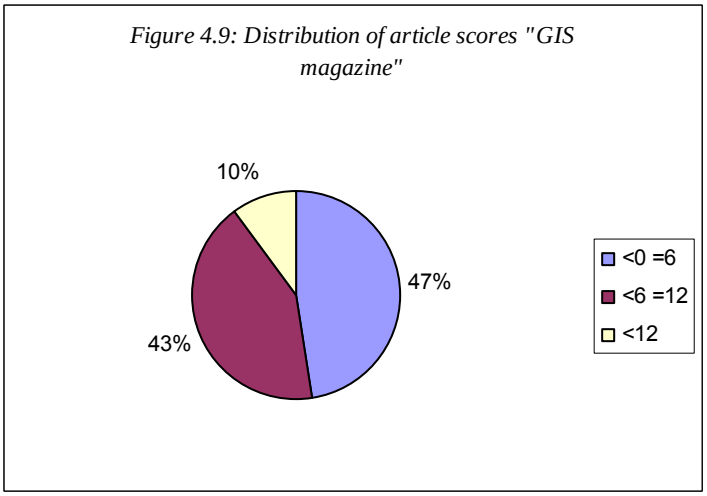


Table 4.14: Distribution of scores for "Gis Magazine" articles

Score distribution	# articles	% of database
>0≤ 6	19	48
>6≤ 12	17	43
>12	4	10

Figure 4.9: Distribution of article scores "GIS magazine"



Instead of taking each result table one by one and making fragmented remarks about the results the author has chosen to combine particular results together to make general conclusions. When 28% of all the articles in the potential data set have a SDI character and these are distributed across three different article score intervals(Table 4.8), we can conclude that there is no institutionalized ignorance with the basic subject matter, the coverage points to interest and curiosity amongst professionals. There is sufficient coverage across the magazine titles (Tables 4.11 through to 4.14) to be able to say that this is not isolated but pervasive phenomena. Of course knowledge has more than a single dimension but the first condition, awareness of the existence of the subject, has been satisfied.

The reasons for this, if one were to go by the subject matter contained in many of the articles are the on-going activities of the OGC, the kick –off of the EU's INSPIRE program in 2004 and on a more local level the Dutch government's "Ruimte voor Geo" program which also opened for tenders in 2004.

The article scores are an indication has to how deep the familiarity with the subject goes. The higher the score the greater the levels of knowledge, one can only expect high scores when the subject matter is sufficiently developed in technical terms which leads us on to the spread of scores in this coded data set (Table 4.10); the articles have a spread of fair to good article score results. Much of the 33 % in the score range of 0-6 represents a lot of "chatter", these are editorials, symposia round-ups and newsletter updates. They reflect interest and knowledge but there is little depth of treatment. To the author's mind this reflects an interest but one which is in development. In the next class of results ($>6 \leq 12$) one commonly came across articles detailing the realization of corporate SDI. These articles detail, in the space of three of four pages the work achieved or plans made to achieve central GIS environments inside companies or organizations. Interestingly enough there seems to be as much movement on this front with state bodies as well within the commercial sector. There are differences in the treatment of these sort of articles some are technical; others are generalized detailing the project management aspects. However the fact that nearly half (49%) of the articles coded fall into this score range shows that the depth of treatment has an upward tendency as GIS professional gain more experience with the subject material.

At the higher end of the scale (<12) in Table 4.10, representing 18% of the total coded data confirms that the knowledge tendency is upwards and not downwards, because it is present on this level as well as the two preceding classes. This is an encouraging trend holding out the possibility that in five years time this percentage should have increased significantly.

In the series of results (Table 4-11 to 4.14) one sought to investigate whether or not the score increase are limited to a single part of the market and thus limited to one target audience by comparing the relative performance of each magazine title to each other. Where there is a broad alignment in the distribution of the score then one can confirm that there is no widespread ignorance of SDI.
"VI Matrix" scored better (Table 4.12) than the average aggregated score, it also the most widely distributed of the titles and that bodes well and certainly refutes the hypothesis that professional body are not being informed in 2005 about developments. The fact that all four titles were producing articles which scored in all three categories proves that there is diversity in the treatment of the material across the board and that enforces the point that there is a range of depth of knowledge which denies ignorance. Had only one title scored highly then one could make a case that the knowledge is still localized and limited to a particular target audience. If the press is a reflection of the GIS profession then as a profession one can say that we are being informed and that the knowledge is not localized as many articles were written either in association with GI professionals or by the professionals themselves. A representative general trend is more apparent in the comparisons between "GeoNieuws"and "Geo info". Together they represent the middle ground in terms of content and quality.

This presentation of the final articles scores continues in order to see if there is a correlation between the article scores and the magazine weight factors. The former is indicative of the quality of the article while the latter compounds the role of target audiences and distribution of the magazine in which the article appeared. Even though the scales cannot be mixed one can examine the data for possible parallels. As there are over two hundred records they will not all be represented here. Without displaying the values graphically one can see that there is no correlation between the article score and the weight factor, this simply means that each magazine title was capable of producing articles which scored both high and low. It is a result which confirms the previous set of results in that there are no prevailing conditions which would allow us to say that one could one magazine to consistently outperform the other.

Table 4.15: A sample of the coded articles and their associated scores and weight factors

article title	magazine title	Score	Weight factor
VGVI nieuwsbrief: Geo DBMS, Certificering en accreditering	VI Matrix	30	8
Jongste OCG-specificaties veelbelovend voor geodata gebruik	VI Matrix	27	8
NCGI: Standaardisatie in de praktijk	VI Matrix	24	8
NCGI nieuwsbrief: Aansluiten bij het nieuwe gedistribueerde concept van NCGI	VI Matrix	22	8
Ravi Nieuwsbrief: Standaardisatie van geo-informatie, Authentieke registratie Geografisch Kernbestand	VI Matrix	22	8
De digitale stroom van bestemmingsplannen	GIS magazine	21	6
Polygonen: het wankele fundament van ruimtelijk modelleren	Geo-Info	21	5
Geoservices: RWS realiseert centrale InternetGIS infrastructuur	Geo Nieuws	21	4
Koepel Architectuur Natte Sector: Nederland, land met water	Geo Nieuws	21	4
RWS naar OpenGIS: OpenGIS Web Catalog Services in de MII	Geo Nieuws	21	4
Rijkswaterstaat start met implementatie OpenGIS	Geo Nieuws	21	4
Implementatie van OpenGIS is startsein voor andere standaardisatiewerkzaamheden	VI Matrix	20	8
Ontbreken semantische meta-informatie bemoeilijkt koppeling ruimtelijke modellen	VI Matrix	20	8
Groningen haalt geo-informatie via warehouse en web uit isolement	VI Matrix	18	8
Zes-dimensiemodel structureert keuzes voor ontwerp geo-loket	VI Matrix	18	8
Het overbruggen van de kloof tussen CAD en GIS- deel 2	Geo-Info	18	5
2004, het begin van een nieuw normentijdperk	Geo-Info	18	5
Gemeente als leverancier van kabel- en leidinginformatie	Geo-Info	18	5
RWS naar OpenGIS: Het Geoservices portaal: naar een nieuwe GII	Geo Nieuws	18	4
Ontwikkelingen bij ESRI	Geo Nieuws	18	4

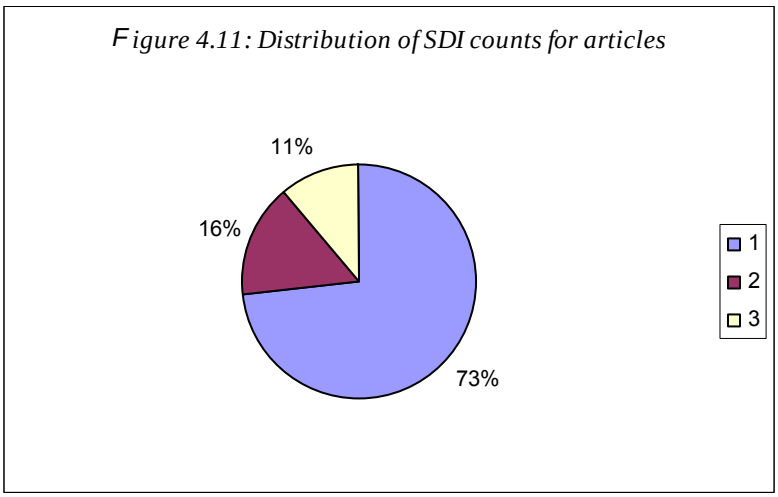
One has established the answer to the question posed however it is instructive to examine how the press is dealing with the subject matter as the results will be instructive in seeing how well the articles are being written. If they are being badly written then their instructive value to the reader diminishes and so to does the appreciation for the subject matter. While not answering an explicit question these results help to establish the state of our knowledge of SDI through the media.
By breaking down the scores into their respective parts to examine the distribution of the various characteristics through the coded data set the underlining trend should become apparent. The results will be presented and collectively dealt with in terms of a single analysis.

SDI counts
This is an overview of the distribution of SDI subjects through the coded articles. It is an indication as to how typically articles would deal with one or more of the SDI subjects in the course of their composition.

Table 4.16: Distribution of SDI counts for articles

SDIsubjectcount	Number of occurrences	% distribution
1	197	73
2	43	16
3	30	11

Figure 4.11: Distribution of SDI counts for articles

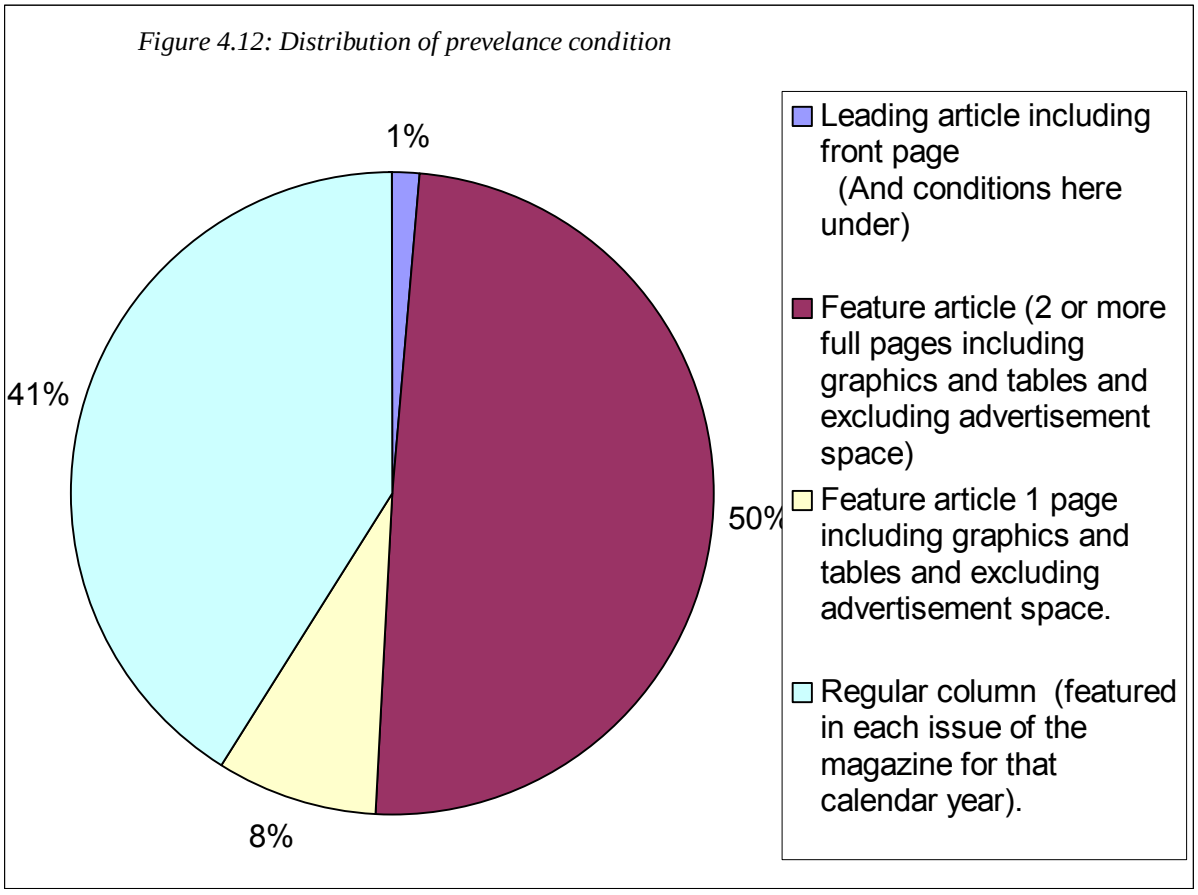


Prevalence of treatment
This is a measure of the distribution of the prevalence of treatment of the coded articles. It helps to clarify the profile attached to the importance of the subject article.

Table 4.17: Distribution of prevalence condition of coded articles

Number of occurrences	Prevalence condition description	% distribution
4	Leading article including front page (and conditions here under)	1
133	Feature article (2 or more full pages including graphics and tables and excluding advertisement space)	49
22	Feature article 1 page including graphics and tables and excluding advertisement space.	8
111	Regular column (featured in each issue of the magazine for that calendar year).	41

Figure 4.12: Distribution of prevelance condition

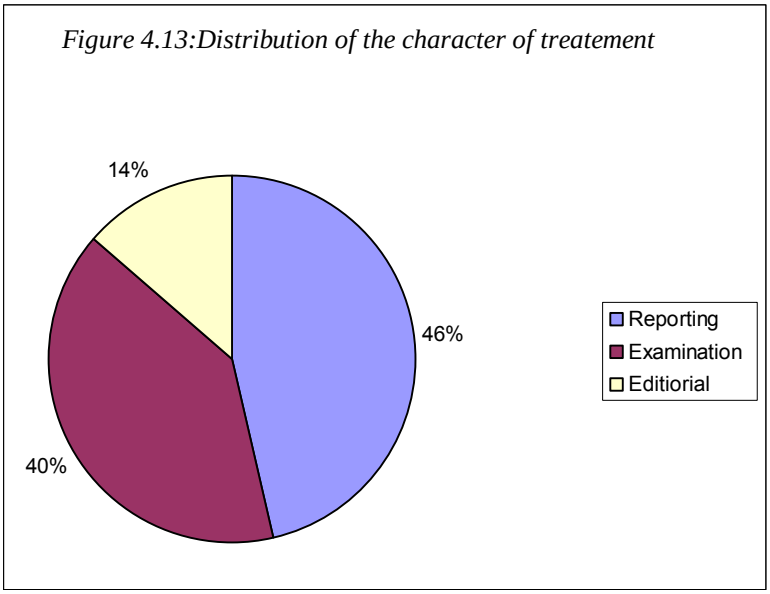


Character of treatment
This is a measure of the distribution of the character of treatment of the coded articles.
It helps to classify the character of treatment into three recognizable forms and gives further generalized insights as to the trend which emerges from the research of the subject articles.

Table 4.18: Distribution of character of treatment

Number of occurrences	Character treatment description	Distribution in %
125	Reporting	46
108	Examination	40
37	Editorial	14

Figure 4.13:Distribution of the character of treatement

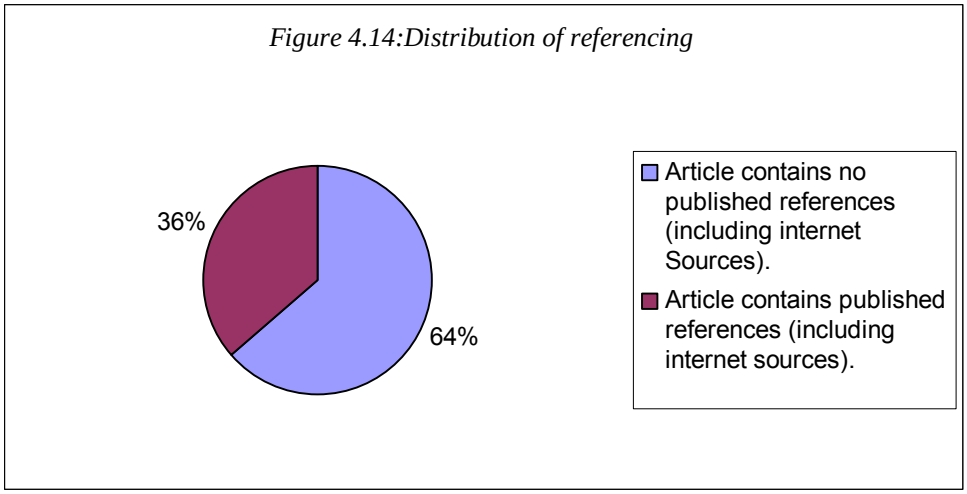


Referencing
These results reveals the extent or otherwise whereby the coded articles have been supported with the use of references. This latter also includes the use of internet references.

Table 4.19: Distribution of use of references in articles

Number of occurrences	Magazine references description	Distribution in %
172	Article contains no published references	64
98	Article contains published references (including internet sources).	36

Figure 4.14:Distribution of referencing



As was pointed out in the last conclusions, each title was capable of producing "good" and "poor" articles. So what can we say about the kinds of articles and what they are telling us about the professional body's level of knowledge. In order to characterize that one must make a series of summations of the results form Table 4.16 through to 4.19.

- The articles have a low SDI subject count (as opposed to a low general score) (Table 4.16) with 73 % of the articles coded not having more than a single SDI subject.
- One third of the articles dealt with more than a single SDI subject 27%, but this result has been negatively affected by the high number of general SDI articles. It is truer to say that many articles dealt with different aspects of SDI but on so generalised a level that the author felt it necessary to code them as general SDI. So while there is a level of awareness of the basic subject matter it has yet to proliferate to a stage where single subject articles are getting detailed treatment.
- the prevalence in the magazine lent to the subject is quiet good – editors are giving these articles plenty of page space (Table 4.17). The front page doesn't seem to have the same pulling power in the speciality press as in the mainstream press.
- A significant proportion of the results (60 %) are either editorials or newsletters. (Table 4.18) so the character of the treatment was fair but the treatment had little technical depth.
- This is reinforced by the distinct lack of references in most cases (Table 4.18) and the interpretation of this rule was liberal, an e-mail address for contact person or their web site was enough for this to be recorded as a reference.

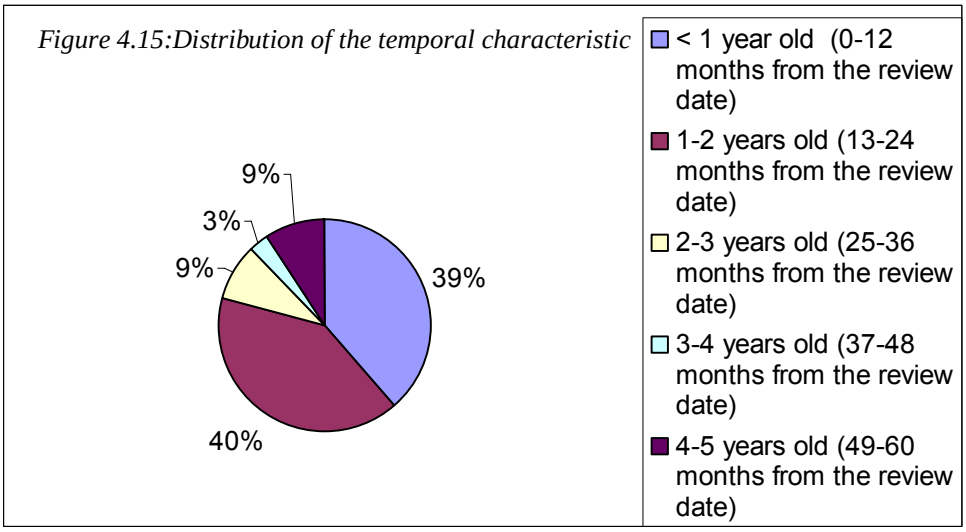
Changes in coverage over time:

Before moving to the next question the results of the history of coverage is worth detailing in some depth as it creates a foundation upon which one can more safely generalise the overall conclusions. Tracing this history helps to reduce the main disadvantage of the survey by moving it through time and helps to support predictions for future trends. The results presented are a more detailed look at the results to be found in Table 4.7 and 4.8 as they trace the coverage of SDI per magazine title and relative to the actual coded data set as opposed to the potential.

Table 4.20: Distribution of temporal changes

Number of occurrences	Temporal character description	Distribution in %
104	< 1 year old (0-12 months from the review date)	39
110	1-2 years old (13-24 months from the review date)	41
23	2-3 years old (25-36 months from the review date)	9
8	3-4 years old (37-48 months from the review date)	3
25	4-5 years old (49-60 months from the review date)	9

Figure 4.15:Distribution of the temporal characteristic



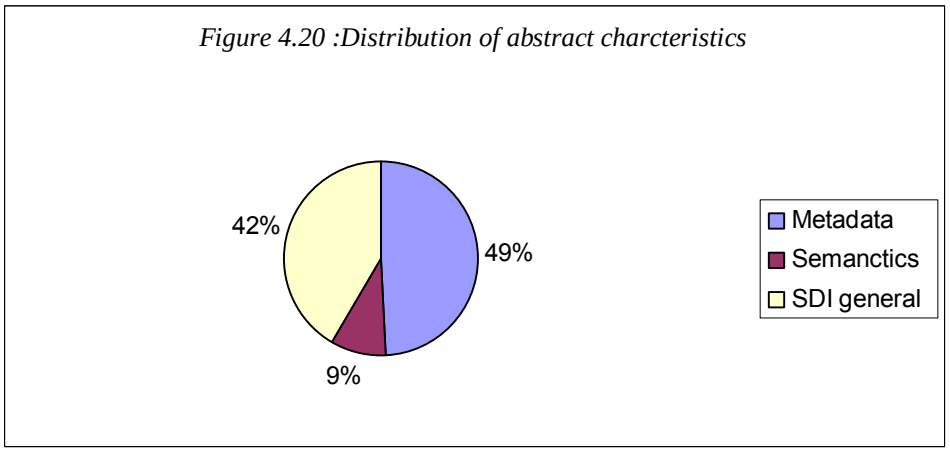
4.3 Are there key weaknesses in our intellectual understanding which impair interoperability with particular attention to the two abstract characteristics mentioned i.e. semantics and metadata. Are these been dealt with in the broad professional arena or are they ignored?

Key weaknesses can be identified by examining what subject matter is being dealt with in the articles and which ones are not. As the areas of interest are semantics and metadata particular attention will be paid to them however as SDI have other subjects the results should lend some insights into these as well. The results will be presented as a whole as will the commentary on them.

Table 4.25: The distribution of abstract characteristics

Abstract characteristic	% of distribution of abstract characteristics
Metadata	49
Semantics	9
SDI general	42

Figure 4.20 :Distribution of abstract charcteristics



Note: While "SDI general" was not originally an abstract characteristic or a concrete form thereof the fact that so many of the articles were generalized in nature was in itself instructive to the general aims of the research. The author therefore decided to include it as not to do so would produce an inaccurate picture of the coverage of the subject matter in the press.

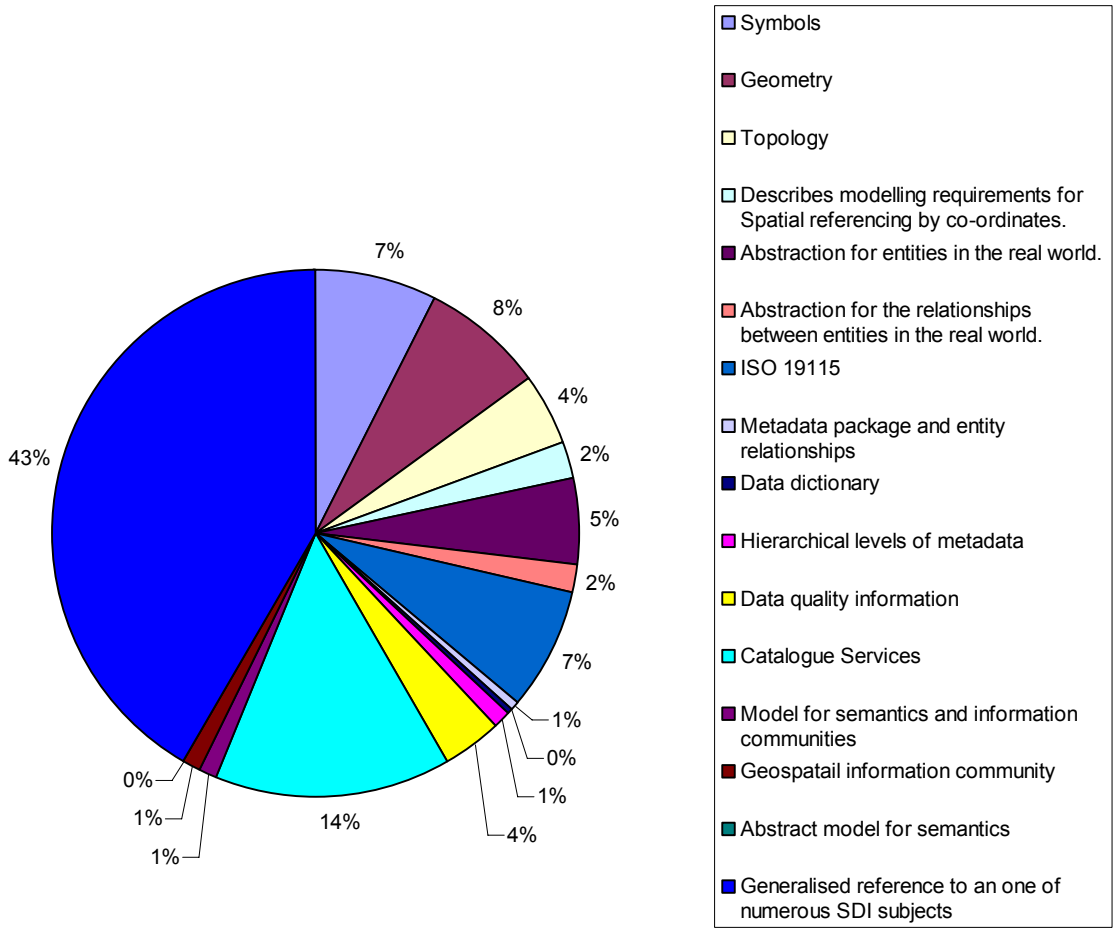
As a follow on from above it is only logical that the distribution of the abstract characteristics should be further detailed into the distribution of their operational /concrete forms.

Table 4.26: Distribution of the concrete characteristics

SDI abstract characteristic	SDI subject description	Total number of occurrence	% of coded data set
Metadata	Symbols	28	7
Metadata	Geometry	29	8
Metadata	Topology	17	4
Metadata	Describes modeling requirements for Spatial referencing by co-ordinates.	8	2
Semantics	Abstraction for the relationships between entities in the real world. This abstraction is modeled as relationships between the features	20	5
Semantics	Abstraction for the relationships between entities in the real world. This abstraction is modeled as relationships between the features	6	2
Metadata	ISO 19115	28	7
Metadata	Metadata package and entity relationships	3	1
Metadata	Data dictionary	1	0
Metadata	Hierarchical levels of metadata	4	1
Metadata	Data quality information	14	4
Metadata	Catalogue Services/ Web services	54	14
Semantics	Model for semantics and information communities	4	1
Semantics	Geospatial information community	4	1
Semantics	Abstract model for semantics	1	0
SDI	Generalized reference to an one of numerous SDI projects	157	42

Note: In the above results all sixteen are represented, in the pilot that was not the case with eleven of the sixteen being treated. The wider selection set has ensured greater variation in the coverage but the general distribution results remained roughly the same.

Figure 4.21 : Disrtibution of the SDI concrete characteristics



These results were compiled further to determine if there is a correlation between the frequency of subject matter and the quality count for the most popular subjects.

Table 4.27 The most popular subjects and the average quality count for these articles

SDI subject	Average quality count / subject
Generalized reference to an one of numerous SDI projects	7
Catalogue Services/Web services.	6
Spatial Referencing by Co-ordinates.	6

Analysis of the results:
Table 4.24 shows the general distribution of the operational characteristics. The reasoning for the inclusion of SDI general as an abstract characteristic has already been stated. The distribution of the characteristics speak for themselves, metadata and the generalised SDI articles accounting for a 91 % of the results and semantics coming in with a very weak 9 %. Here is the first indication of a key weakness, the entire area of semantics is getting very little attention in the press and that this is a reflection of the weakness of our understanding of the subject matter. The inclusion of semantics as an abstract characteristic has been deliberate from the outset as it represents a Esperanto or a lingo franca without which interoperability will be a technical reality but not yet a human one. Semantics, offer us the basis of understanding the dialects of geo information across communities and the absence of attention to this subject reveals a weakness which will certainly have a retarding effect on the realising of a state of interoperability.

The poor treatment of semantics is probably a reflection of the difficulty which the GIS community has in dealing with the subject in general. An example of this is the poor headway which the OGC has made in its own semantic project (since 1999), which unlike many of the fast tracked internet services projects has been put on the back-burner and the proposal is not completed. There seems to be a definite preference to deal with the technological issues rather than the epistemological area of semantics and one can not really fault the popular GIS press in this.

Examination of Table 4.26 reveals that apart from the general SDI that metadata worked out into its concrete characteristics “web catalogue services” was the most popular topic for articles followed by geometry and ISO standards for metadata. The relationship between the catalogue services and the new ISO 19115 standards is not coincidental. The first stage of the EU’s INSPIRE project is aimed directly at creation of metadata catalogues via web services. Many organisations are aware that the first steps to interoperability are being able to find data sets and these two subject address this issue. One can not help wondering that the present activities around metadata standards have been positively influenced by the developments of web services.

Table 4.27 shows us that the most popular subject all had average low articles scores so frequent coverage does not automatically mean in-depth treatment. This is not necessarily a weakness but there is certainly room for improvement in the depth of coverage whereby our knowledge is strengthened.

4.4: Do GI professionals realize the financial benefits and costs to their own organizations and partners in sharing data?

In order to quantify the results for this question the remarks fields were used as a basis to filter out those articles which paid specific attention to the issues of financial benefits of SDI and interoperability. There were thirteen articles which dealt with in some way the issue of the financial benefits of SDI. In this count there has been no recording of the number of articles dealing with either EU or national government funding of SDI programs. It is worth noting there were in the regular updates on the "Ruimte voor Geo" programme references to the financial subsidies available for development of projects for the programme. These were not included in this result except where they specifically addressed the issues of cost benefit analysis. The author was interested in testing in how far the interest in interoperability has been driven by solid business cases.

Have the various SDI initiatives' been supported with the presentation of well argued cost benefit analyses? As was outlined in Chapter 1 the EU's INSPIRE programme was preceded by a cost benefit analysis so there is a much published example which should be familiar to GI professionals. Given that much of the coverage of the articles in the coded data set detail the start up or realisation of corporate or local level SDI's one could assume that there would be reason to expect some coverage of the subject matter. In addition given the developments, articles which remind GI professionals of the need to make Cost benefit analyses would be appropriate and helpful. In reference back to the previous question one could argue that this is a weakness in our approach to SDI. Given the complexity involved in realising a corporate SDI and the number of articles dealing with the latter one would have expected more attention to be paid to the subject.

The results returned thirteen articles out of the two hundred and seventy articles coded latter. This represents coverage of 4.4 % of the total. This is a dreadful result and if it is a reflection of general practise it points to a weakness in the **approach** of GI community in arguing for SDI initiatives. It is also a poor reflection of the fiscal rectitude of management in accepting these projects.

One would have to conclude that, if GI professionals aware of the benefits relative to the costs they are well able to hide it! Every one of us understands vaguely that interoperability is a "good thing", technically and by default financially but apparently we are not used to articulating that in the press. It is because the costs benefit analysis is difficult area for IT in general that the absence of high profile for this subject is disappointing. The GI community would benefit form articles which give insights into quantifying the benefits of SDI in monetary terms.

4.5 Do GI professionals have any realization of the possible legal consequences of failing to create the basis if interoperability? Are they even aware that there may be legal consequences involved in not taking steps towards interoperability?

A distinction needs to be drawn between those article which dealt with legal consequences pertaining to the implementation of SDI and those which (particularly the regular newsletter type of articles) report the progress of government legalization. The latter were generally short pieces and beneficial in the dissemination of information in this area but without the depth of detail that one is looking for. They have not being included in the selection for the results.

The author was specifically interested in the evidence of the effects of INSPIRE directive, the Lisbon guidelines and the national governments "Actieprogramma Andere Overheid", "Authentieke registraties", "Ruimte voor Geo","Wet Persoon registraties", "Wet kenbaarheid publiekrechelijke beperking". While not all of the above are laws, all do have a coercive political function in modernising the digital services which governments offer their citizens. Failure of state and semi state bodies to act on these have legal and political consequences and it is awareness of these which is what the author was testing for.

In contrast to the previous point there seems to be more attention to be given to the legal context of interoperability. Even though there were few articles ten in total, which dealt explicitly with these issues there was a constant thread going through many other articles which referred to the growing legal framework in which one operates. Many of the articles which dealt primarily with the creation of local level or corporate level SDI's referred to changing laws and directives as an impetus to the project.

So while this element was not particularly well represented as explicit article subject material, it was evident from the total data set that this is a subject which is high on the agenda. One can state that GI professionals are quiet aware of the legal and political environments in which they operate and of the consequences of failure to realise their legal obligations and the ambitions of their political masters. This may also serve as an explanation to why the financial aspects (see above) were poorly developed, many of the contributors to these articles work in the state and semi-state sector and the legal catalyser seems to be more effective in affecting change in this environment than the financial one.

Chapter 5
Conclusion, evaluation, and recommendations

5.1 Final conclusions:

By way of reaching a final conclusion the author has decided to reiterate the five questions which were set out at the end of chapter 1

*Are individual GI professionals being **sufficiently well informed** in these issues? Is there an actual communication problem?*
In a word yes, there was 28% coverage of SDI subject matter in all the magazine titles, the coverage varied and had a definite history stretching back some five years.

*Is there a widespread ignorance of SDI? With all the attention given to the development to SDI in the last few years how **aware is the GIS professional** of the developments in 2005?*

One can conclude that there is no institutionalized ignorance with the basic subject matter, the coverage points to interest and curiosity amongst professionals. There is sufficient coverage across the magazine titles (Tables 4.11 through to 4.14) to be able to say that this is not isolated but pervasive phenomena. The various qualative levels which where tested proved that there are strengths and weakness in how the material is been dealt with but the GI professionals should be well aware of the developments in 2005.

Are there key weaknesses in our intellectual understanding which impair interoperability with particular attention to the two abstract characteristics mentioned i.e. semantics and metadata. Are these been dealt with in the broad professional arena or are they ignored?

There are key weakness in the area of semantics as this subject matter gets little or no attention and many of the articles had a much generalised nature however there are reasons to believe the depth of treatment and thus the level of intellectual understanding is increasing.

*Do GI professionals realize the financial **benefits and costs** to their own organizations and partners in sharing data?*

There was little evidence to support the idea that GI professionals are very preoccupied with the issues of cost benefit analysis as there was very little attention paid to the issues in the articles. Apparently everyone believes in the intrinsic value of the SDI undertaking without publicly debating the costs involved.

*Do GI professionals have any realization of the possible legal **consequences of failing** to create a SDI? Are they even aware that there may be legal consequences involved in not taking steps towards interoperability?*

Yes they do, even though the explicit proof in terms of dedicated articles was not overwhelming there was a constant sub-text to many articles of the changing legal context in which we work. There were also much evidence of the constant updates in the newsletters regularly published by RAVI and the like.

5.2 Evaluation of the hypothesis, assumptions and the extracted predictions:

In Swanborn's methodology having completed the research one has to go about evaluating the research (see Figure 2.1, page 24/25) and we do this by looking at the hypothesis, the assumptions upon which it was based and the extracted predictions. Doing so allows the researcher to state the solution to the problem and identify new problems based on the shortcomings of either the research methodology or the results.

Hypothesis:
To reiterate this was the hypothesis set out at the end of Chapter 2 -

"It is this author's contention that one of the barriers to interoperability is the ignorance of GIS professionals of the developments, in understanding the issues involved, in appreciating the costs and benefits and finally the legal framework within which SDI will have to be realised."

The hypothesis was tested through the five questions dealt with in Chapter 4 and from these we can state that GI professionals are certainly aware of the developments, the depth of understanding of the technical issues is increasing though much of the knowledge seems to be of a generalized nature but with an upward trend regarding the depth of knowledge. There are key weakness in the treatment of the issues of semantics which is disquieting, there seems to be little exposure of the cost benefit analysis but a lot of implicit recognition of the burgeoning legal obligations on us as a professional body in the realization of spatial data infrastructures.

This conclusion would lead the author to reject the hypothesis as there are enough proofs that GI professionals are in the throes of a steep learning curve and that this then is the solution to our problem.

Assumptions:
(See page 18 for the original stated assumptions)
The research was based on the presumption that the levels of knowledge or ignorance of the subject matter could be tested by measuring the coverage given to them in the specialized press. It was also assumed that the use of informal education channels were a better source than formal ones and lastly that the scientific methodology of the research method was sufficiently sound for the results of the survey to be able to test the hypothesis.

The scientific basis of the first assumption was not made explicitly clear through reference to theoretical precedents however the survey is an accepted methodology in sociological research in testing a hypothesis. Whether or not informal educational channels are a better source for testing when compared to formal channels was not tested and cannot be proved or disproved. It is up to the reader to decide if the arguments put forward in Chapter 2 are convincing or not. The soundness of the design research is one again which the reader will also have to decide based on the description of the research design in Chapter 2.

Extracted predictions

(See page 18 for the stated predications)

The predictions that knowledge can be measured over time was a sound one but the prediction that knowledge would increase over time was not supported by data, in other words an articles written five years ago were as likely to return a high article score as one written in the last year even when the age of the article was not calculated into the score. What the data does clearly show is that there is an increase of coverage over time.

The prediction that the emphasis would be on technical aspects of SDI without the institutional context was also unfounded, this was more explicit in the qualitative results rather than in the quantitative ones. Many of the generalized SDI articles coded dealt with no single technical issues but the latter mixed with ancillary issues.

The prediction that there would be a difference in the amount and type of coverage could not be tested effectively as there was not sufficient variation in the target audience to test for this accurately. Where one title (VI Matrix) scored better in this regard can be explained by a larger data set and the distorting effect of equating a half page newsletter item as an article.

The treatment of the legal framework was pervasive as a sub theme of many articles but there were few articles which dealt with the matter explicitly and given the developments timelines of SDI it is perhaps too early to expect attention to be paid to this. The prediction is in this regard is premature and thus flawed.

The issue of cost benefits analysis has been dealt with in detail but the prediction has been bourne out.

The prediction that there would be a key weakness in the treatment of some of the subject matter was bourne out and these were detailed in Chapter 4.

5.3 Critical reflections on the research methodology:

Before moving to the final conclusion it is instructive if the shortcomings in the research are set out as they provide a basis for recommendations for further research and help the reader assess the overall quality of the research. For the sake of clarity and brevity they shall be set out in point form.

- The target audience attribute was a failure; there was no hard evidence to back up the attributes attached.
- There was no named precedence's for the assumption that the press could be treated as a representative vassal of the collective knowledge.
- There were no theoretical comparative research alluded to.
- There was no hard evidence that the printed press still enjoys greater sway over GIS professionals when compared to pure Internet titles.
- In the scoring of the magazine articles there was no account taken of the number of "hits" articles received when placed on the Internet – an average of these should be added to the circulation figures for that title.
- Many of the remarks in the coded data recorded if feature articles dealt with, say a project to realize the creation of a corporate SDI, but as it was not a specific attribute it was difficult to measure later. Where applicable code for the specific type of SDI being realized global, region, national, local and corporate.
- The reduction of the all the qualative aspects of an article to a single score is not above reproach, there is a degree of conceit involved namely the mixing of nominal and ordinal scales. However the research sought to offsets this by also examining each of the component parts of these scores whereby the reader can better decide the value of the overall score.

5.4 Recommendations for future research

In order to address some of the shortcomings set out above and to help address the overall theory that knowledge of developments is critical success factor to any undertaking the author has set out a series of recommendations for possible future research.

- Carry out a survey to try identifying what informal channels of education GI professionals use to update their knowledge base.
- Carry out research into how formal education GI materials deal with the subject matter.
- The research should be extended beyond the confines of the specialized GIS press into the mainstream IT magazine market or titles which specifically cater to the IT managers. As pointed out in Chapter 2 this would be a good way to test for the demassification of the SDI into the general population.
- Following on form this point a similar survey could be used for testing where in the evolutionary phase of SDI we are. One would have to make a clearer link between research development theories (Nolan for example) and the subject matter in the media.
- Repeat the research say in three years time again to test for changes both in the depth of coverage and the range of subjects being treated.
- Extend the subject matter into the more technical fields – the subject matter would have to include more than just meta data and semantics.
- Investigate if the new found interest in metadata is being driven by the technology (catalogue web services) rather than an appreciation of the correct theoretical practice. Ideally the technology is supposed to follow where the theory points to and not the other way around. If this is the case and the author feels that this is so, can the same inverted logic be put to use on the issue of semantics at some point in the future?

5.5 Recommendations for the specialised GI press:

Given that the press formed the primary source for data for the research the author has made a series of recommendations which editors of magazines could apply, they are as follows -

- Increase the depth of the treatment of technical articles. When one compares the depth of technical detail of the average GIS articles to say an article in a medical journal the contrast is striking.
- Carrying on from the previous point, the depth of treatment can be artificially increased by the inclusion of references for feature articles. There is too little use of this in the GI press.
- Seek out subject matter which deals with cost benefits analysis.
- Give me a job.
- Seek out more articles which detail methodologies of data quality controls.
- Begin to put the issue of semantics on the agenda through articles which stress diverse user communities and the need for sharing data across professional barriers. Some of the editorials in the data set have addressed this issue but the author knows of no feature article which dealt with the problem in-depth.

Literature list

Author (s)	Date	Title	Page
Http ://			
Arnold Bregt,Joep Crompvoets, Henk Scholten, Peter van de Crommert	2004	The Dutch National Spatial Data Infrastructure past present and future	3
Bradburn N.M., Sudman S.	1979	"Improving interview method and questionnaire design"	General
EU Comission	2004	DIRECTIVE OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL establishing an infrastructure for spatial information in the Community (INSPIRE)	
http://inspire.jrc.it/reports.cfm Feburary 2004			
Gallagher. M.P	2001	“NIST Interoperability Cost Studies”	
http://www.rti.org/publications/cer/7007-3-auto.pdf			
Global Spatial Data Infrastructure	2004	“The SDI Cookbook	8
http://www.gsdi.org/pubs/cookbook/Default.htm			
Gomm, Hammersley and Foster	2002	"Case Study Methods"	5,6
NATO	1997	NATO Logistics Handbook,	Chapter 17:
http://www.nato.int/docu/logi-en/1997/lo-1705.htm			
Open GIS Consortium Inc.	1999	"Topic 14: Semantics and Information Communities"	General
HTTP:/WWW.OPENGIS.ORG/DOCS/99-114.PDF			
Open GIS Consortium Inc.	1999	Topic 13: Catalog Services	General
'HTTP:/WWW.OPENGIS.ORG/DOCS/99-113.PDF'			
Open GIS Consortium Inc.	2001	Topic 11: OpenGIS(tm) Metadata (ISO/TC 211 DIS 19115)	General referen
'HTTP:/WWW.OPENGIS.ORG/DOCS/01-111.PDF'			
Open GIS Consortium Inc.	2003	OpenGIS® Reference Model	5
http://www.opengis.org			
Open GIS Consortium Inc.	2003	FAQ – OGC	
http://www.opengis.org/resources/?page=faq			
Panasonic,Holman L.	2002	“DRM Interoperability”	
whttp://www. europa.eu.int/information_society/eeurope/2005/all_about/digital_rights_man/doc/wg2/interoperability1.ppt			

Author (s)	Date	Title Page	
Http ://			
Rand Corporation	2001	A BROAD DEFINITION OF INTEROPERABILITY	
http://www.rand.org/publications/MR/MR1235/MR1235.chap2.pdf			
Ravi	2004	Proposal Space for Geo-information	General
http://www.ravi.nl/ruimte/Proposal%20SpaceGeoinf.pdf			
Swanborn, P.G.	1987	"Aspecen van sociologisch onderzoek"	10,18,25,29,3
Whinston. A.B	1999	"The Role of Standards in the Growth of Global Electronic Commerce".	
EUROGI	2000	"Towards a strategy for geographic information in Europe"	General
http://www.eurogi.org/geoinfo/eurogiprojects/strategy.pdf			
DGIWG	2000	"The Digital Geographic Information Exchange Standard(DIGEST)"	
CEN	1998	"The Geographic Information European Prestandards and CEN Reports"	
	1998	"ENV 12657: Geographic Information - Data description – Metadata"	
	1998	"ENV 12658: Geographic Information - Data description – Transfer"	
	1998	"ENV 12656:1998 Geographic Information - Data description – Quality"	
http://forum.afnor.fr			
ISO	2003	"Draft Business Plan of ISO/TC 211 - Geographic information/Geomatics" (draft)	
	2000	"704:2000, Terminology Work – Principles and Methods"	
	2002	"ISO/TC 211 N 1320: Text for DIS 19104, Geographic Information – Terminology,	
http://www.isotc211.org.			
Albert K. Yeung	1999	"Public Access to Geographic Information"	
http://www.ncgia.ucsb.edu/giscc/units/u190/u190.html			
Federal Geographic Data Committee	1997	"A Strategy for the NSDI"	
http://www.fgdc.gov/nsdi/strategy/index.html			
Clinton, William	1994	"Coordinating Geographic Data Acquisition and Access: The National Spatial Data Infrastructure, Executive Order 12906"	
ESRI	1997	"Enterprise Spatial Data Management with SDE". *ESRI ARC News, Summer, 8.*	
Laurini R.,Thompson D.	1992	"*Fundamentals of Spatial Information Systems"*, Academic Press : London	
Baker W.,	1997	"Universal Access vs Universal Storage". *Thirteenth International Conference on Data Engineering*, IEEE,Birmingham, April 7-11	

Overview of magazine titles used

magazine title	Jaargang	issue	Date	Number per year
Geo Nieuws				
	2			
		1	March 2002	4
		2	June 2002	4
		3	September 2002	4
		4	December 2002	4
	3			
		1	March 2003	4
		2	June 2003	4
		3	September 2003	4
		4	December 2003	4
	4			
		1	March 2004	4
		2	June2004	4
		3	September 2004	4
		4	December 2004	4
	5			
		1	March 2005	4
Geo-Info				
	0			
		0	October 2003	11
	1			
		1	January 2004	11
		2	February 2004	11
		3	March 2004	11
		4	April 2004	11
		5	May2004	11
		6	June 2004	11
		7/8	July /August 2004	11
		9	September 2004	11
		10	October 2004	11
		11	November 2004	11
		12	December 2004	11

magazine title	Jaargang	issue	Date	Number per year
	2			
		1	January 2005	11
		2	February 2005	11
GIS magazine				
	1			
		3	October/November	8
		4	December 2003	8
	2			
		1	January/February 2004	8
		2	March 2004	8
		3	April/May2004	8
		4	June 2004	8
		5	July /August 2004	8
		6	September 2004	8
		7	October/November	8
		8	December 2004	8
	3			
		1	January/February 2005	8
		2	March 2005	8
		3	April/May 2005	8
		4	June 2005	8
VI Matrix				
	8			
		1	February 2000	8
		2	April 2000	8
		3	May2000	8
		4	June 2000	8
		5	September 2000	8
		6	October 2000	8
		7	November 2000	8
		8	December 2000	8
	11			
		1	February 2003	8
		2	March 2003	8
		3	May 2003	8

magazine title	Jaargang	issue	Date	Number per year
VI Matrix	11			
		4	June 2003	8
		5	September 2003	8
		6	October 2003	8
		7	November 2003	8
		8	December 2003	8
	12			
		1	February 2004	8
		2	March 2004	8
		3	April 2004	8
		4	June 2004	8
		5	September 2004	8
		6	October 2004	8
		7	November 2004	8
		8	December 2004	8

Appendix I:
Description of the coding process and the generation of the results

The following is a step for step description of the actual coding process for the research.

Step 1:

A MS Access database was created, the data model (see Appendix III) of which were based on the methodology from Chapter 3.The primary interface for the coding process was the input form whereby an article was recorded and its characteristics relative to the aforementioned methodology was recorded. This table, when complete formed the basis for the statically analysis of the results.

Step 2:

The data set was gathered, in this case a series of digital articles were collected from the web site of the magazine (or the originals scanned in pdf). A single pfd file per issues was made. All the files were gathered into a single directory, this allowed for easier search routines in Adobe whereby the entire directory (the entire collection of articles) were searched for the presence or absence of the search parameters.

Step 3:

One of the results of the pilot was the realisation that the list of search words was too restrictive – the extension of the list meant that many articles reappeared in the search process but that simply reinforced the validity of the search findings. For the complete list of the search words used see Appendix II.

Step 4:

Having made an initial selection of the articles the registered articles where then examined in detail. The articles were read in detail and coded in terms of the parameters, the subject of the article in terms of the concrete characteristics and a qualitive score based on prevalence, referencing, character of treatment etc. For data controls much use made of mandatory fields, failure to code the record in the required manner would result in the record not being saved. The researcher ensured in this way that the data was recorded in a careful and concise manner. Worthy of note is the presence of the two remarks fields; these were used by the author to record his own impression of the article. These will form a detailed basis for the qualitative analysis of the results in the final research. Without these it will be difficult make these analysis over the larger final data set. What did emerge form the pilot was the viability of using standard phrase in the remarks fields especially for the generic SDI articles.

Step 5:

Having searched the total data set and having made a detailed selection it was now time to start the statistical analysis of the results. This was realised using a combination of MS Access and Excel with an automatic data update from Access into Excel. The graphics were generated in Excel and linked into the final reports.

Appendix II
List of searchwords

English

Metadata
Semantics
- Feature geometry
  - Symbols
  - Geometry
  - Topology
- Spatial referencing Stop
- Features
- Relationship between features
  - ISO 19115
  - ISO
  - Data dictionary
  - Catalogue Services
  - Information communities
  - Data quality information
- Information communities

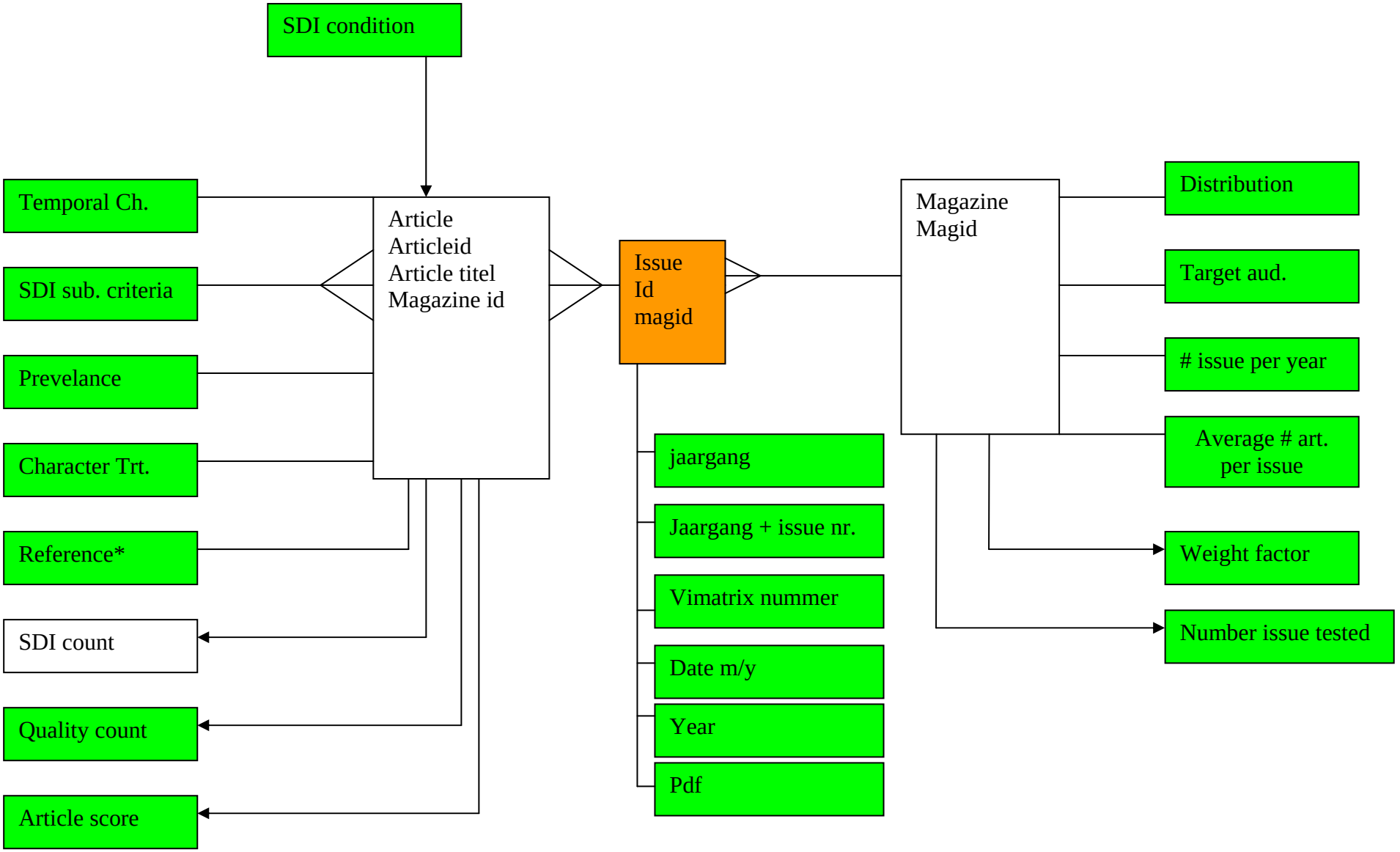
Dutch

Metadata
Semantiek Semantisch
- Kenmerk geometrie
  geometrisch kenmerk
  geometrische kenmerk
  - Symbool
  - Geometrie
  - Topologie topologisch
- Ruimtelijke referentie
- Relatie (verwijderen)
  - Data woordenboek
  - Cataloog diensten
  - Cataloog dienst
  - Catalogus
  - Informatie gemeenschappen
  - Data kwaliteit informatie
  - Data kwaliteit

Generic search words
SDI
GSDI
OGC
INSPIRE
Ruimte voor Geo
OpenGIS
Geo-portal
Geo-portaal
Internet services
NCGI
RAVI
Nationaal pakhuis
Pakhuis
Authentieke registratie
Basis registratie

Appendix III

Datamodel



Appendix IV

Sample of the results

Article title	**magazine title**	**Score**	**Weight factor**
Dataland ligt goed op koers	GIS magazine	4	6
Sherpa moet politie tot GIS-hoogtepunt leiden	VI Matrix	4	8
Jaarlijkse GITA Conferentie	Geo Nieuws	5	4
Eerste tenders Bsik-programma:'Ruimte voor Geo-informatie' van start	Geo Nieuws	5	4
Inspire-Congres in Warschau	Geo Nieuws	5	4
Open Source GIS conferentie en Second Annual MapServer User Meeting	Geo Nieuws	5	4
GIS-productencatalogus RIKZ binnenkort op intranet	Geo Nieuws	5	4
Nederlandse Kernset Metadata	Geo Nieuws	5	4
Eerste projecten Ruimte voor Geo-Informatie van start	Geo Nieuws	5	4
Ruimte voor Geo-informatie van start	Geo-Info	5	5
Geo-Info op bezoek bij Nuon	Geo-Info	5	5
Om de tafel met Jack Dangermond	Geo-Info	5	5
GIS- conferentie 2004 beidt ruimte voor wonen	GIS magazine	5	6
Professor Henri Alders stopt zonder te stoppen	GIS magazine	5	6
Alleen wie de barstjes in de muur ziet, zal klagen	GIS magazine	5	6
Gelderse regio regelt informatie-uitwisseling in 'vredes'- en rampentijd	VI Matrix	5	8
Onderscheid in vermogen van ingenieurs bureau	VI Matrix	5	8
Programma Ruimte voor Geo-Informatie mijn visie ..	VI Matrix	5	8
Er dreigt een tweedeling tussen gewone geodata en Inspire -data	VI Matrix	5	8
Naar een uniforme registratie van het RWS-beheerareaal	Geo Nieuws	6	4
Implementatie centrale infrastructuur: Geoservices in het RWS werkproces	Geo Nieuws	6	4
Een geo-norm(ale) studiedag	Geo Nieuws	6	4
Eén grootschalige topografische kaart voor Nederland?	Geo Nieuws	6	4
Twintig miljoen euro voor geo-informatie beschikbaar	Geo Nieuws	6	4
Kwaliteit van geo-informatie in theorie en praktijk	Geo Nieuws	6	4
Drie opvallende bijeenkomsten in het geo-ICT-werkveld	Geo Nieuws	6	4
Symbolendatabase Topografische Dienst Kadaster	Geo Nieuws	6	4