

Tesi doctoral presentada per En/Na

Elisabet ROSA TRÍAS

amb el títol

**"Conceptes, tècniques i problemes de la
generalització cartogràfica"**

per a l'obtenció del títol de Doctor/a en

GEOGRAFIA

Barcelona, 25 de febrer del 2000.

**Facultat de Geografia i Història
Departament de Geografia Física i Anàlisi Geogràfica Regional**



UNIVERSITAT DE BARCELONA



4. La integració de la informació espacial

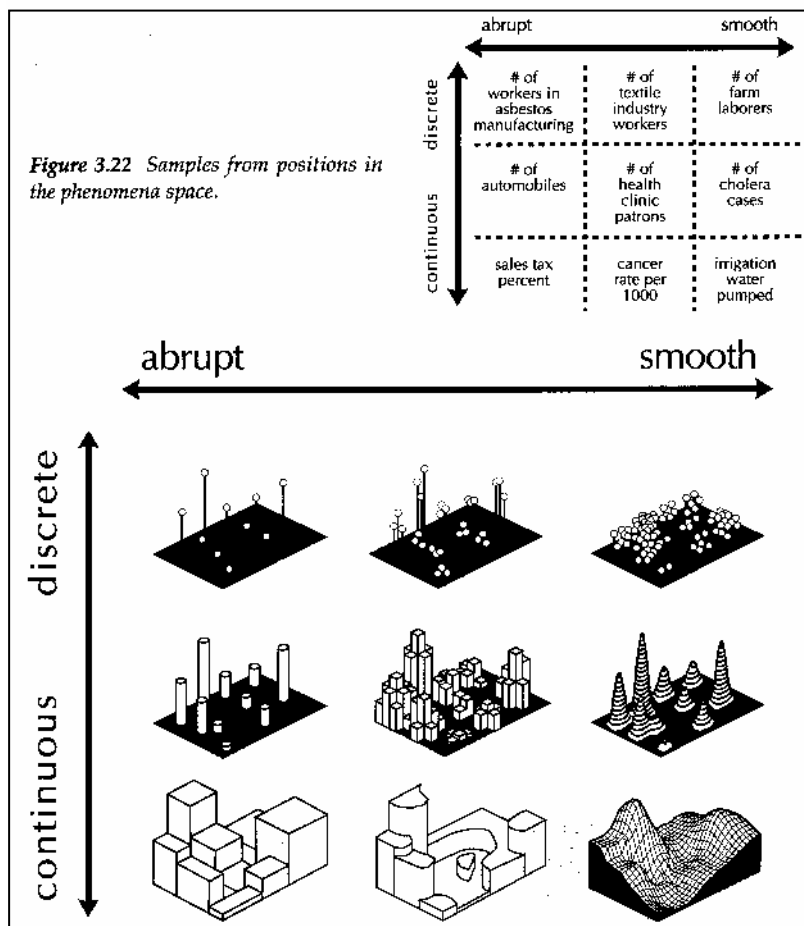
That generalization is a phenomenon which involves large variations in the ways nature is abstracted is nothing new. It would be however of interest to identify these points where a small variation in scale may cause large variations in the geometrical and substantive content of the map (Muller 1991:466)

The multipurpose maps of yesterday, essentially descriptive, static and deterministic are now conceptually challenged by new products which are extremely volatile, single purpose and probabilistic. The realm of cartographic activities has expanded towards the portrayal of highly abstract geographic spaces [...] A challenge for the cartographer will be to extend his activities towards the manipulation of the invisible part of the landscape model, where data are unclassified and ungeneralized, where lines are fuzzy and areas heterogeneous, within the framework of a multidimensional space including time. A stronger involvement of cartographers in "pre-cartography" will emancipate cartographic thinking as dictated by platonic view of a world made of ideal forms and help design cartographic products where "what you get you cannot see" (Muller 1989:679)

En l'entorn digital ampliar el procés de generalització és el que Muller anomena "precartografia", un estadi previ a la representació cartogràfica, en el que el cartògraf s'hi ha d'involucrar, si vol controlar el posterior processament de les dades. És el procés que permet passar del fenomen observat del món real, a la seva integració com a objecte espacial en la base de dades, i on la codificació que donem a les dades originals, com a informació numèrica o discreta, és important de cara al seu posterior anàlisi (MacEachren 1994:40).

El procés tradicional d'elaboració del mapa, comptava amb un preprocessament de la informació semblant al que suggereix com a "precartografia" de Muller o el "pensament visual" de DiBiase (citats per MacEachren 1994:2), que han estat oblidats en l'inici de la cartografia digital. Les feines que han estat deixades de banda, ha estat decisives en el desenvolupament del procés posterior, i són les que corresponen en l'antic procés de compilació de la informació: a la codificació i escalat de les dades i la simbolització.

La "precartografia", en un entorn digital, correspon al preprocessament de les dades que en el procés cartogràfic tradicional ha correspost al procés de compilació de les dades generalment derivades a partir de les seves representacions cartogràfiques a major escala, en els mapes temàtics. Dins del procés d'elaboració del mapa en l'entorn digital, ha permès utilitzar eines de tractament estadístic i numèric espacial de les dades originals, que la teledetecció ja ha vingut utilitzant normalment, des dels seus inicis s'ha. Com ja s'ha vist anteriorment, el tractament de les dades espacials, actualment, no se centra en la manipulació de les representacions cartogràfiques sinó en la manipulació del model de les dades (Mark 1989), i és per tant decisiu el tipus de tractament que donem a les dades en la seva conversió a objecte digital espacial.



Font: MacEachren 1994:59)

En l'inici de la cartografia digital s'havia obviat la importància del preprocessament

de les dades; en el moment actual és evident que la intervenció del compilador és primordial i que ha quedat en mans de l'usuari que moltes vegades en desconeix la seva transcendència.

L'usuari actual d'un SIG, és un usuari fonamentalment multidisciplinar, que té interessos diferents dels tradicionals, principalment en la visualització de les dades, per altres finalitats que no són les de comunicació. I per l'altra banda existeix un usuari que "pensa" en base a les representacions cartogràfiques, interessat en la manipulació de les representacions d'aquestes basant-se en el llenguatge cartogràfic (Muller 1989:680).

En la cartografia tradicional, s'utilitzen les representacions cartogràfiques tant per l'exploració de la informació original com per la seva integració i homogeneïtzació de cara a l'anàlisi visual, és un procés que es guia per unes regles estrictes, com són les de l'escola suïssa, que permeten controlar-ne la qualitat del procés. Aquesta normativa deixa marge d'interpretació al cartògraf que assumeix moltes vegades la importància que té fer prevaler certes formes de representació assumides tradicionalment de certs fenòmens, com pot ser el fet d'associar un fenomen a escales més grans com a tipus d'objecte espacial polígon, i a escales més petites representar-ho com a punt. La representació d'aquesta informació sota altres elements geomètrics com pot ser una superfície estadística comporta dificultats en la seva lectura i per tant en el procés de comunicació. En canvi és de gran utilitat per l'investigador poder veure altres tipus de visualitzacions de les dades amb les que treballa.

En la conversió de les bases de dades per a la seva integració en els sistemes d'informació geogràfica, de cara a la seva anàlisi, ens trobem amb la dificultat de reproduir el procés de compilació per la integració de les dades, ja que només comptem amb la possibilitat de conèixer-ne la seva resolució. Ja s'ha dit que la resolució no supleix a l'escala tradicional i a les seves funcions en la cartografia no digital, i que en aquesta, té un sentit que va més enllà de l'escala entesa com a ratio mètric.

La manca d'escala entorpeix el processament de les dades en l'entorn digital, especialment en el moment actual, en que els SIG comencen a prioritzar les funcions d'anàlisi per sobre de les funcions d'inventari. És per tant fonamental, poder comptar amb automatitzar el procés cartogràfic, i per tant en formalitzar aquest procés.

Com assenyala Robinson i al. en la obra clàssica *Elementos de Cartografia* el preprocessament de la informació es important i decisiu, i ha de poder decidir el model de generalització més convenient en cada situació:

“El cartógrafo actual y el del futuro inmediato deben ser capaces de seleccionar inteligentemente el tipo de generalización más adecuada antes de saber realizar la generalización cartográfica manualmente” (Robinson et al. 1984:248).

La generalització adequa la resolució de les dades de cara al seu ús posterior, o sigui que a través de la homogenització prepara les dades per a la seva integració. En l'entorn manual el procés homogenització es fa durant el procés de compilació tradicional, i la integració es fa mitjançant la manipulació de les representacions dels objectes espacials. Les escales de tractament dels diferents fenòmens suposen una decisió a partir de la resolució i del nivell de detall, sobre quina ha de ser l'escala d'anàlisi. El cartògraf selecciona les entitats espacials, moltes vegades de manera inconscient, fent que les variances d'intraclasse siguin equivalents. En l'entorn digital, hi ha dos processos que es desenvolupen de manera paral·lela: en base a les representacions cartogràfiques dels objectes digitals espacials cartogràfiques; o a partir de les dades originals.

Es parteix en aquest treball de la hipòtesi de que l'escala cartogràfica utilitzada en la cartografia tradicional, no es només una escala de tipus mètric, sinó que recull un concepte menys restringit d'escala: el d'escala conceptual. Quan intentem reproduir en un entorn digital el procés de generalització a través de la formalització de les seves tasques, tal com s'ha fet en la cartografia tradicional, podem observar que si basem el procés cartogràfic en l'escala mètrica, es impossible de dur a terme la rèplica del antic procés, i que necessitem poder utilitzar el concepte d'escala a un altre nivell conceptual.

En la cartografia tradicional, la selecció d'entitats del mon real la fem basant-nos en l'escala conceptual, i escalem i codifiquem, guiats per unes necessitats d'observació i d'anàlisi, on la selecció i codificació d'entitats s'efectua a diferents nivells conceptuals i és, per tant, difícil d'automatitzar el procés que permeti seleccionar, i geocodificar o decidir el tipus d'objecte que correspon a una entitat, si en considerem de manera aïllada basant-nos només en la seva escala mètrica en l'intent de generalització automatitzada.

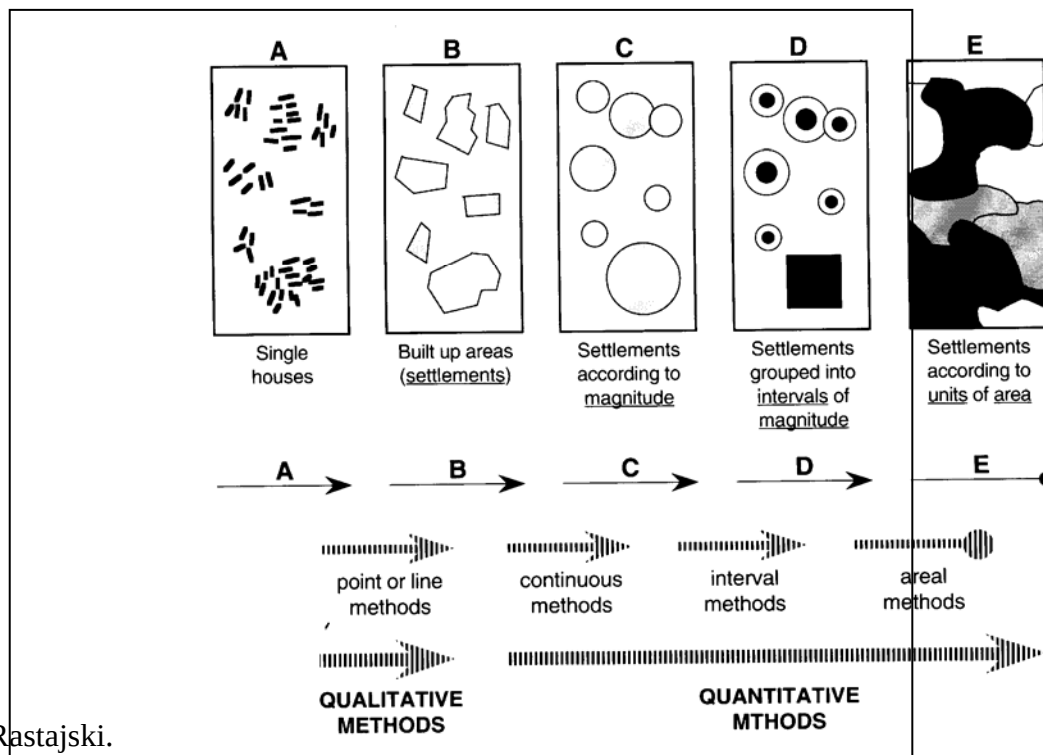
Així doncs, la utilització de l'escala conceptual durant el procés de compilació cartogràfica tradicional es un fet implícit pels cartògrafs, tot i fer-ne només explícit el caràcter mètric de l'escala. L'escala de tipus mètric es fonamenta en la precisió mètrica, mentre que l'escala conceptual ho fa en la precisió semàntica i es trobaria en un nivell més alt d'abstracció. Si considerem que la resolució és l'equivalent de l'escala mètrica, hauríem de poder concloure, com es fa en la cartografia tradicional, de que el procés d'integració de diferents dades es pot recolzar en l'escala tradicional, l'escala mètrica, i que a través de les reconversions successives de les resolucions d'un objecte espacial aconseguir derivar-ne¹ un seguit infinit successiu, de fenòmens. Podríem dir que si representem un eix de les possibles resolucions amb totes els valors possibles que adopta aquesta variable, els plans o espais que conformen, no adquireixen sentit de manera automàtica ni tampoc que tingui sentit la inclusió d'un fenomen determinat en aquest.

" Other categories of space [such as vector and raster], which are in concern to spatial scientists and which represent different levels of generalization [or abstraction], are the metric, topological and graph-theoretic representations. [Note that] these categories of space illustrate various degrees of generalization at the conceptual level, and do not concern the problems of generalization at the display level. They are meant to provide a theoretical foundation for data structuring. Data structures are determined by requirements in accessibility and manipulation of information whose degree of generality may vary according to the type of GIS analysis which must be performed". (Muller 1991:462)

Dit d'una altre manera: a partir d'una base de dades original podem derivar a través de la generalització diferents objectes digitals, i això es el que ha fet la cartografia tradicional i digital; i es cert si considerem un fenomen de manera aïllada, però no si en volem deduir o derivar-ne relacions amb els objectes. Per exemple podem intentar generalitzar un objecte espacial de la base de dades per simplificar-ne o millorar-ne el seu posterior tractament. Ara bé hi ha resolucions a les que els fenòmens o processos del mon real no tenen sentit, per tant es podria recollir el que s'ha anomenat el "catastrophic approach" de Muller que fa referència a l'escala conceptual i el de la cartografia tradicional

¹ Aixó ha estat recullit per DeCola aplicant la teoria de fractals (DeCola 1989)

de Rastajski.

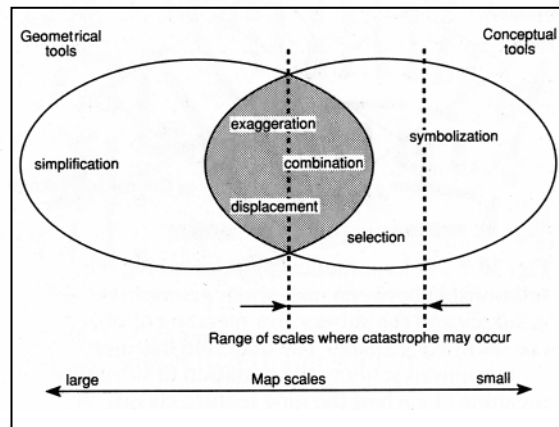


(Font: McMaster, Shea 1992:22)

A nivell de dibuix hi ha elements geomètrics que no es poden representar de manera convenient ja que no permeten la seva inclusió dins d'un espai determinat i el poder oferir unes relacions clares amb els demás elements del mapa, i a nivell conceptual també hi ha objectes que no tenen sentit a tots en tots els espais possibles. El mateix passa amb les possibilitats que apareixen resultat de tractament estadístic, on no tots els resultats calculats o inferits tenen sentit. Per tant no es cert que nosaltres puguem, a través de la derivació de les resolucions d'un objecte digital espacial, derivar-ne objectes que tinguin la seva correspondència amb fenòmens del món real (apareixeria un objecte espacial com si hagués viatjat a través del temps en un espai dins del que no té cap relació amb la resta d'objectes).

"Conceptual models are used to provide a framework in which to express spatial relationships. There are basic categories which are recurrent in the GIS literature, such as vector or raster space, and which reflect the form in which data are basically encoded, organized and manipulated. We have become so entrenched in those categories that now have raised the procedures and techniques of handling spatial data to the status of epistemological models." (Muller 1991:461)

El voler derivar a partir de la resolució original d'un objecte de la base de dades altres objectes espacials, es degut a que s'intenta transmetre informació no geomètrica sinó semàntica entre els diferents nivells d'abstracció, el que evidentment no és possible si partim només de la resolució o dimensions geomètriques d'un fenomen.



*Processos de transformació segons l'escala del mapa
(font: Muller 1991: 467)*

Segons Muller la "catàstrofe" es dona quan es passa de una generalització geomètrica a una de conceptual. I fa notar que algunes eines del processament són tant geomètriques com conceptuales. I això condueix a una àrea a on es desdibuixa el camp de tipus geomètric del conceptual.

La tria de l'escala en l'antic procés cartogràfic, es feia a nivell conceptual; i per tant no es pot homogeneïtzar la base de dades si no és en aquest nivell, i l'escala mètrica o la resolució, no permeten la conversió d'una escala conceptual a una altre. Així ho mostra Muller citant a Muerhke (1969):

" The functional dependence of map utility according to scale, thematic realm, data quality, and map purpose is a topic which requires further study. A simple paradigm may be proposed where utility increases with increasing scales, until a point is reached at which utility starts decreasing. [...] It would be of interest, however, to identify these points where a small variation in scale may cause large variations in the geometrical and substantive content of the map. Those points where a "catastrophic change" may occur would help us to identify more sharply the range of scales which are suitable for a particular type of map use. The term "catastrophic" is used here to indicate that the generalization process is

discontinuous with respect to scale. It is used as an analogy to catastrophe theory (Thom 1973), although this is concerned with description of the operation of whole systems rather than single elements."

El saber del cartògraf en fer la tria dels objectes, es fa en el procés de compilació; i es per tant en aquest estadi on es decideixen les escales possibles de representació dels diferents fenòmens i la resolució o nivell de detall que desitgem.

Pot ser important, recollir els processos d'integració de dades en la compilació cartogràfica tradicional, per que tant en l'antic procés, com en el nou preprocessament hi ha un concepte d'escala lligat al procés cognoscitiu, en base al qual fem la selecció de determinades entitats corresponents al mon real en base, i ho fem moltes vegades en base a la tradició que en base a les necessitats reals d'anàlisi.

El problema, de fet, s'hauria de plantejar a l'inrevés ja que en les base de dades nosaltres tenim resolucions dels objectes espacials, però no sabem dins a quin rang d'escales els hem de fer correspondre. La pregunta que planteja Nyerges seria: ¿com fer la compilació de les dades en l'entorn digital i la tria d'una escala d'anàlisi, quan no ens basem en la seva representació cartogràfica? (Nyerges 1991:61)

En el procés de compilació tradicional la tria de la dimensionalitat de l'element primitiu es fa en funció de la funcionalitat principalment de representació, mentre que actualment en el model de les dades seleccionem i codifiquem en base al context probable d'anàlisi.

"...It seems rather confusing in the case of data coming from surveys of photogrametry it would be more appropriate to speak of precision, accuracy and resolution not scale. The concept of scales would only come when a relationship between object and representation (in some kind of output medium) would be established..." (Land Use and Cover Change 1997).

Aquesta afirmació recull les raons per les quals no ha funcionat considerar a l'escala tradicional cartogràfica, ja que només considera l'escala mètrica o la de representació però no considera l'escala a un nivell conceptual més alt, com pot ser l'escala d'anàlisi.

Tanmateix, el cartògraf sempre és el responsable del procés de compilació en la construcció del mapa, tal com assenyala Robinson:

" El papel del compilador cartográfico es definir los elementos individuales del mapa que se hallan asociados con los procesos cartográficos de selección, clasificación, simplificación y simbolización. Entre una multitud de otros elementos, incluyendo la elección de una proyección, el área de la cobertura, la escala y las características a incluir en el mapa, implica la elección de una técnica de representación cartográfica (puntual, coroplética, isoplética, etc.) y las especificaciones de parámetros de representación cartográfica (intervalos de clase, número de clases, etc.). Es decir, el compilador prepara la hoja de trabajo con ayuda del ordenador, pero tiene en su contenido la misma responsabilidad que en la compilación manual" (Robinson et al. 1984: 401)

Aquest procés previ es pot considerar, com ja s'ha vist, part integrant de la generalització; i en la cartografia tradicional permet que la simbolització de les dades originals i conversió a representació cartogràfica possibiliti el seu anàlisi visual. Sobre el treball de compilació tradicional Robinson assenyala:

"El proceso de compilación requiere a menudo un cambio en la geometría de los materiales fuente [...] La mayor parte del trabajo de compilación de un modo u otro con la escala del mapa. La escala de la minuta de compilación ha de determinarse en base al área a cubrir y al tamaño del mapa final. Una vez se ha determinado la escala de la hoja de trabajo, resulta a menudo que esta difiere de la escala de los mapas fuente de se dispone [...] Todos los mapas están, obviamente, elaborados a una escala. En la práctica muchos mapas temáticos se preparan a un tamaño que encaje con un formato preestablecido, siendo este formato el tamaño y la forma de la hoja sobre la que aparecerá el mapa [...] Cualquiera que sea el formato, el mapa debe encajar en él... " (Robinson et al. 1984: 409)

Respecte al treball de compilació digital el mateix autor:

" Muchos de los factores que fueron críticos en la compilación manual y fotomecánica no son importantes en la compilación digital. Inversamente, la compilación digital introduce factores que no aparecían en la compilación analógica [...] Los métodos digitales generalmente proporcionan al compilador una mayor flexibilidad que los métodos analógicos. El compilador dispone de una mayor libertad para probar diseños alternativos, mezclar datos procedentes de otras fuentes, cambiar la escala y la geometría de

la hoja de trabajo a voluntad..."

Seria interessant poder anar endarrera en el procés d'anàlisi fins al preprocessament, transportant l'escala conceptual d'anàlisi al preprocessament i observar com el cartògraf, en realitat, fa la simbolització de les entitats, ja no en base a la seva representació cartogràfica sinó preveient-ne en el seu processament analític.

Això mateix s'ha fet en els inicis de la cartografia digital. A on s'ha considerat que l'organització de les dades, "... selection of a specific spatial data structure may considerably influence the processing of queries" (Egenhofer, Herring 1991:227), és decisiva pel bon funcionament del sistema. I l'organització de les dades espacials, s'ha basat tradicionalment en l'abstracció del món real, a través de la representació en els objectes geomètrics; així, doncs, la finalitat dels mètodes que s'utilitzen en l'organització de les dades espacials, també anomenades estructures de les dades espacials, ho han fet principalment de cara al processament basant-se en el tipus de tractament que permeten els diferents objectes digitals espacials.

L'estructura de les dades correspon a una estructura de baix nivell, o nivell lògic, a diferència del model de dades, considerat com estructura d'alt nivell. Egenhofer cita, com Kainz en canvi, presenta un marc de treball a on es distingeix l'organització de les dades en el model de dades espacials (considerat estructura de dades d'alt nivell) i estructura de les dades espacials (abans estructura de les dades de baix nivell). Aquesta distinció permet separar el model conceptual o model espacial de les dades de la seva estructuració física o estructura de les dades.

"Els models espacials de les dades descriuen formalment la semàntica dels conceptes presents en el model geomètric de les dades, i que amb l'ajut dels conceptes matemàtiques de topologia, mètrica i ordre poden contestar a diferents qüestions importants en els SIG" (Kainz 1989 cit Egenhofer, Herring 1991:230)

4.1 Model i esquema: diferents nivells en el disseny de la base de dades

Un sistema d'informació s'estableix en relació a un context i a un ús determinats, i com assenyalen Dutton i Kraemer (cit. Man 1990:333), dels dos tipus de sistemes

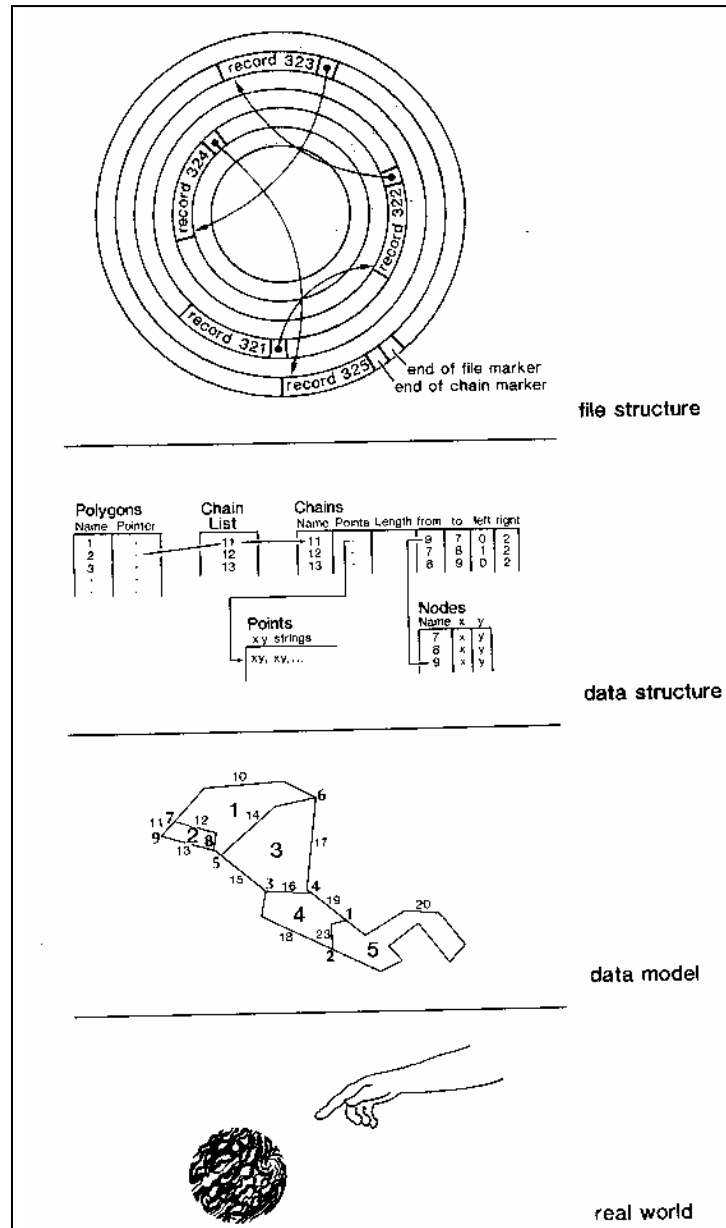
d'informació existents, uns estan basats en les dades, son els que subministren informació mitjançant mètodes no analítics; i els altres sistemes son els que estan basats en la presa de decisions, que necessiten solucions basades en l'anàlisi i serien sistemes mes complets.

Un Sistema d'informació fa la presa de decisions a partir de l'anàlisi permet solucionar el problema cada cop més real del tractament de dades procedents de diferents fonts. Des del d'un punt de vista de l'usuari, aquest és un problema que moltes vegades no és evident fins que en valora els resultats procedents de l'anàlisi que pertany al que en cartografia tradicional corresponia a la feina de compilació o d'intent d'integració de dades heterogènies (Guptill 1991:515) a traves de: el desenvolupament de informació detallada sobre les dades espacials, la seva qualitat i el fluxos associats a l'ús de la base de dades espacial; la creació d'un diccionari tant dels atributs espacials com no espacials, que inclogui les vistes locals dels usuaris de la base de dades; el desenvolupament d'un model lògic de la base de dades que defineixi la estructura que permeti les vistes locals dels usuaris i permeti solucionar les inconsistències del diccionari; reconversió de l'esquema lògic; de quines son operacions addicionals d'integració de les vistes per el soft que farà ús de la base de dades.

Un conjunt de dades representen un determinat model de la realitat; el procés de representació o de conceptualització de la realitat, o model conceptual de les dades, és, doncs, a mig camí dels nivells d'abstracció de les dades que va del pas de la realitat fins l'estructuració física d'aquestes.

La representació dels nivells del model de les dades (*data model*):

Realitat	→	Model de les dades o model conceptual	→	Estructura de les dades o model lògic	→	Estructura dels arxius o model físic
----------	---	---	---	---	---	--



*Nivells d'abstracció en relació en les bases de dades geogràfiques
(segons ([Peuquet, 1984 #338]:254*

Com assenyala McMaster, els conceptes procedents de la literatura de programació, posteriorment adaptada als SIG, pel que fa a una sèrie de definicions com són el model conceptual, el model de les dades, i les vistes dels usuaris, poden diferir en alguns casos. Ara bé, el que és clar és que la gestió de la base de dades a través dels anomenats Sistemes

de Gestió de Base de Dades és un component molt important en la funcionalitat del sistema i comprensió per part de l'usuari, ja que facilita, a través de llenguatges anomenats "d'alt nivell" més propers als llenguatges parlats. Els nivells "alt" o "baix" fan referència a la proximitat en la manera de fer de les persones o de les màquines, o sigui que fa referència al llenguatge parlat o al de programació, respectivament (McMaster, Barnett 1993:388).

4.1.1. Models i esquemes de les dades: intern, extern i conceptuals

El primer nivell d'abstracció corresponen a la realitat, a tots els fenòmens o processos, per tant, els que són percebuts com els que no. En segon lloc, el model de les dades o model conceptual correspon d'una banda a una abstracció del món real que incorpora només les propietats que són importants per resoldre un problema determinat de la realitat. En tercer lloc, l'estructura de les dades, o model lògic, és la representació del model de les dades que reflexa les qüestions de posada a la pràctica, i que normalment s'expressa com a diagrames, llistes o matrius dissenyades per reflectir l'emmagatzematge de les dades en l'ordinador. En darrer lloc l'estructura de l'arxiu, o model físic, al·ludeix a com estan físicament emmagatzemades les dades.(Maguire, Dangermond 1991: 320)

Nivell d'abstracció	Definició
Realitat	Un fenomen en la seva totalitat en el món real.
Model conceptual de les dades	Un conjunt de components i relacions entre els components que pertanyen a un fenomen específic que sembla d'interès per unes necessitats determinades. Un model de dades és independent dels sistemes específics o de les estructures de les dades que organitzen i

	gestionen les dades. Correspon al model de les dades.
Model lògic de les dades	L'organització lògica dels components del model de les dades i la manera en com s'han establert relacions entre els components s'han de definir. Correspon a l'estructura de les dades.
Model físic de les dades	Un conjunt de regles que especifiquen la implementació de l'estructura de les dades dins de diferents entorn (<i>computing system environments</i>). Correspon a l'estructura de l'arxiu.

Aquests son els nivells d'abstracció de les base de dades geogràfiques de Peuquet (1984)

(Guptill 1991:516)

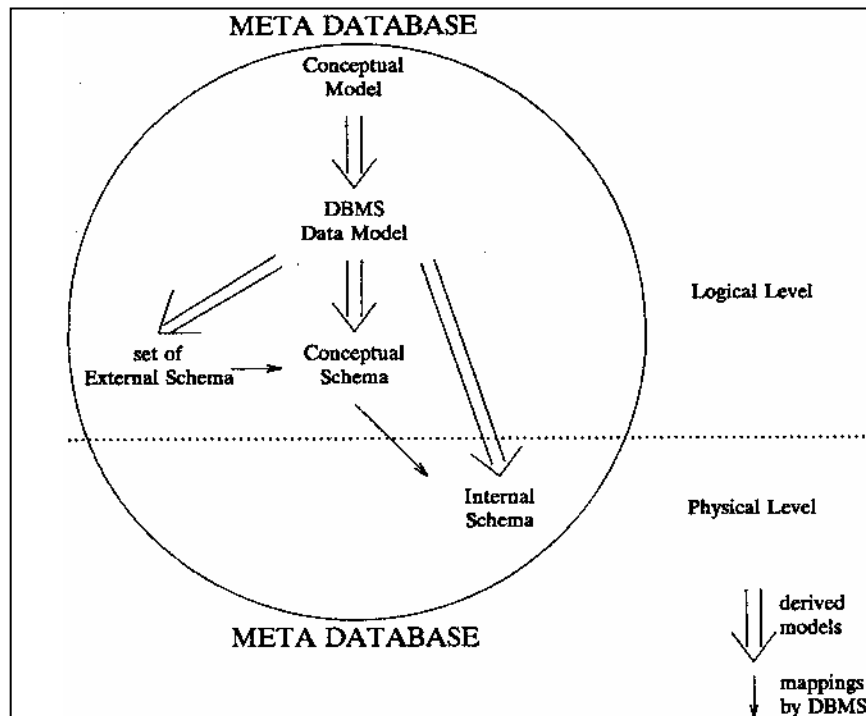
Un model conceptual el podem considerar com una abstracció de la realitat:

"[...]and it is a limited view of reality in the sense that only aspects which are relevant to the application area are reflected in the model" (Oxborrow 1989:57)

El seu contingut és, per tant, decisiu en el resultat final.

A un nivell molt elemental es pot diferenciar: el sistema tradicional d'arxius, on els arxius semblen no tenir cap relació entre si, on el conjunt de dades no té estructura, i per tant, s'ha de construir per cada aplicació; i l'enfoc de la base de dades, on l'usuari percep les dades emmagatzemades en un única base de dades, encara que en la realitat poden estar físicament separades. Aquesta "independència de les dades" permet diferents "vistes" per diferents usuaris. Disposem de "vistes" a nivells altres nivells baixos com: el nivell alt, corresponent al model extern tal com és vist per les aplicacions individuals o els usuaris; el nivell intermedi o nivell lògic, que conté la base de dades sencera i que correspon a l'anomenat model conceptual, i es important de cara a la modelització de les dades. I, en tercer lloc, el nivell més baix que correspon al model intern o model físic de com estan

emmagatzemades les dades.



La relació existent entre el model de les dades i els diferents esquemes és representada en la figura següent (fig 3.17 Oxborrow 1989: 95).

En l'esquema d'Oxborrow, el model conceptual es deriva de la informació obtinguda durant la investigació de les dades inicial; a continuació el model de les dades del DBMS és derivat del model conceptual i d'aquest se'n deriven a la vegada les descripcions lògiques corresponents a l'esquema extern i al model conceptual; i finalment al nivell més baix, trobem el esquema intern a un nivell físic de descripció de l'estructura física de les dades.

La seqüència feta de manera automàtica per el DBMS es fa: de l'esquema extern $\Rightarrow$ esquema conceptual $\Rightarrow$ esquema intern; la meta base de dades conté la informació rellevant tant lògica (semàntica) com física apareix obrint i tancant, ja que es essencial per derivar els esquemes i els models, ja que. Ells diferents esquemes representen diferents vistes de la base de dades tant la vista dels usuaris (*user view = external view*) o de les aplicacions al nivell extern (la més coneguda) com la vista global al nivell conceptual, o la vista de l'emmagatzematge a nivell intern. Es important doncs que la integració es faci al nivell adequat, i per tant més alt possible.

El concepte de model de les dades (*data model*) i el d'estructura de les dades (*data structure*) en el passat s'havien vist com estructures d'alt nivell i baix nivell respectivament (Egenhofer, Herring 1991:228), que es el que fa Oxborrow. Peuquet, en canvi, defineix el model de les dades o (*data model*) de manera diferent com "[data model] is defined as the conceptual data representation scheme"; i el terme estructura de les dades (*data structure*) fa referència exclusivament a "to the programmable implementation of a data model within the context"(Peuquet 1990:68).

4.1.1.1 Disseny eficient de la base de dades espacial

El pas del model de la realitat al model de les dades, en termes de DBMS, es realitza en el procés de disseny de la base de dades; del que se'n deriva un esquema conceptual o descripció del model de les dades.

El sistema de gestió d'una base de dades és un component crucial en un SIG ja que permet emmagatzemar, manipular i consultar les dades; el seu disseny és, per tant, decisiu en l'excel·lència del funcionament d'un sistema d'informació.

Peuquet insisteix en la importància del disseny de la base de dades espacial, que partint del treball de Calkins (1982), encara que el disseny d'aquest no fos per dades espacials Peuquet 1990:55. La metodologia de Calkins per al disseny de bases de dades espacials complexes consisteix en: la formulació de les necessitats; el disseny conceptual basat en el desenvolupament d'un nombre de "vistes" dels usuaris (o subesquemas) de la base de dades, i en la seva integració en una única, coherent i global base de dades final . En aquest procés les "vistes" s'uneixen, en un model conceptual de la base de dades que s'anomena "integració de les vistes"; en una vista global resultant, representa un model conceptual integrat de la base de dades, que conté i gestiona totes les "vistes" locals.

La implementació del disseny consisteix, en primer lloc, en utilitzar el model conceptual de la base de dades com a entrada per poder ser utilitzada per un sistema de gestió de base de dades existent al mercat, i és en la que es basa l'enfoc d'entitat-relació (E-R) de Chen (1976) "The Entity-Relationship Model- Towards a unified View of Data" a on

el més important del procés de modelització, és la construcció d'un conjunt de "vistes" de la base de dades també anomenat "vista integrada".

El procés de disseny per derivar a partir d'una base de dades el model conceptual, és important en el disseny de la base de dades espacial en un estadi inicial del preprocessament i es desenvolupa en quatre fases:

1 - identificació del contingut de les dades corresponents a les entitats espacialment relacionades, i a l'establiment d'una definició específica de cada una d'elles en un diccionari local;

2 - establir els diagrames Entitat-Relació (E-R) per cada una de les entitats;

3 –establir a través dels diagrames individuals E-R, una vista integrada que representi totes les entitats i les seves definicions, per tant conversió dels seus diccionaris en un d'únic;

4 -transformació de la "vista" global resultant en un model conceptual de la base de dades.

Els passos per el disseny d'una base de dades, en un context institucional més ampli, a nivell global i per part d'organismes nacionals ha de tenir en compte a més:

- la identificació dels components potencials d'usuaris;
- el desenvolupament de les necessitats, i establir les consultes o tipus de consultes més freqüents;
- el desenvolupament d'informació detallada sobre les fonts de dades espacials, la qualitat de les dades, i els fluxos associats per a l'ús de la base de dades espacial;
- la creació d'un diccionari de les dades que cobreixi tant els atributs espacials com no espacials, que estan inclosos en les vistes locals dels usuaris, en de la base de dades ;
- el desenvolupament d'un model lògic de la base de dades que defineixi la seva estructura de manera que les vistes locals dels usuaris es vegin suportades, amb la conseqüent solució d'inconsistències del diccionari;
- la reconversió de l'esquema lògic en un pla d'implementació en la base de dades;
- operacions addicionals d'integració de les vistes de cara al programari del sistema

que en farà l'ús.

4.1.1.2. Modelització de les dades: el model conceptual

Els Sistemes d'Informació, i els SIG més concretament, ofereixen, a través de la recollida de dades una selecció de fets del món real; unes dades que s'emmagatzemen i esstructuren en la base de dades, en el que anomenem model de les dades. És important que aquesta estructura reflecteixi de la millor manera els fets del món real, tant a través d'emmagatzemar en les dades informació d'alt nivell, no només de tipus geomètric sinó també semàntic, com a través d'incloure les interrelacions que existeixen entre els diferents fenòmens del món real. Una base de dades ha de permetre la independència de les dades, la separació entre l'estructura física d'emmagatzematge i estructuració de les dades i de les "vistes" corresponents en aquesta, que permeten que els diferents usuaris tinguin la seva pròpia "vista" d'una part, corresponent a totalitat de la base de dades.

La modelització de les dades és un dels processos més importants durant el desenvolupament o construcció de la base de dades, i està formada per els processos de: la investigació de les dades, la modelització de les dades, el disseny de la base de dades, la implementació de la base de dades i el control i ajust d'aquesta. Els primers sistemes d'informació estaven enterament orientats de cara a la funcionalitat sense prestat suficient atenció a les dades, al contrari de com s'intenta fer ara.

Dels processos assenyalats com a necessaris en el desenvolupament d'una base de dades, ens interessa de manera especial: la investigació de les dades i la seva modelització.

Utilitzem el terme "anàlisi de les dades" per descriure la "investigació de les dades", o procés de recerca per la identificació de la naturalesa i l'ús de les dades, això com de la seva organització. La informació que recollim consisteix tant en detalls dels objectes, com en identificar els tipus de condicionaments de domini, integritat i limitacions que ens trobem en la seva representació, com també en regles d'inferència i derivació.

Durant el procés de modelització integrem els diferents fets recollits pel procés de investigació de les dades, en el model conceptual. El model conceptual transforma els aspectes del món real en un model formal, i aquest model derivat és, un model generalitzat

i independent de cap DBMS específic.

El model conceptual ha d'intentar de reflectir les discrepàncies provinents de la percepció de la realitat, i qui dissenya el model de les dades ha de poder oferir una visió àmplia de la realitat. El procés d'investigació de les dades permet obtenir informació, i no té necessàriament que coincidir la visió del model que millor representa la realitat, amb el que serà millor des del punt de vista del sistema d'informació o rendiment del sistema.

" What is this reality which we are trying to model? Reality is effectively boundless [...] But only perceived reality can be modeled and perceived reality is bounded. However, we all have our own, necessarily incomplete, views of reality. It is perceived differently by different individuals, and from different viewpoints. (Oxborrow 1989:)

La literatura procedent del camp de la informàtica, i més concretament, la procedent dels sistemes de gestió de les base de dades, a la que pertany Oxborrow, diferencia entre "vista": física, conceptual i dels usuaris. Aquesta diferenciació assegura la independència de les dades, de la seva implementació en un determinat model, a la vegada que permet que es pugui modificar aquest, sense haver de refer-ne a la resta. Del model conceptual se'n deriven les diferents "vistes" dels usuaris o model extern (*external model*), de manera que aquest i l'estructura d'emmagatzematge o model intern (*internal model*) queden separats.

Els sistemes de gestió de la base de dades proveeixen un limitat recolzament del model semàntic, i això fa que alguns conceptes vinguin representats de manera poc adequada, o bé que estiguin absents, o es trobin dividits entre la base de dades i els programes. No s'ha d'oblidar que és el dissenyar de la base de dades, l'usuari, el responsable de que es compleixin els condicionaments semàntics; i així doncs, la implementació del model conceptual en un sistema de gestió de les dades (DBMS) fa que sigui necessari reconèixer la feblesa en el suport semàntic, a on la única manera la solució, que no seria per altra banda gaire desitjable, es faria a través de la construcció de semàntica addicional:

"This is highly undesirable and violates one of the important objectives of the database approach - that of data independence".

Per a poder considerar que un model conceptual es independent (no es de moment implementable en DBMS existents i ho seria en KBMS) entre d'altres: en primer lloc, hauríem de poder comptar amb metadades, i les dades hauríem de poder integrar en una base de dades que tingués coneixement o saber (*knowledge-base*); en segon lloc, hauria de permetre's que diferents tipus de regles o predicats i el llenguatge intern de la base de dades permetés un llenguatge més proper al nostre.

Les limitacions de tipus semàntic dels actuals models poden ser corregits si aportem informació complementària en les metadades i diccionaris de les dades²; que permetin que els sistemes de gestió de les base de dades DBMS, es transformin en sistemes de gestió amb més saber KBMS (*knowledge base management systems*).

"The semantics of spatial concepts are defined and formalized by spatial data models. Spatial data models are formalizations of the concepts humans use to conceptualize space. Such formalization of spatial concepts are necessary, because computer systems are essentially formal systems that manipulate symbols according to formal rules." (Egenhofer, Herring 1991:229).

El paper que juguen els models espacials "is similar to the conceptual schema in the three-schema view" o sigui que els conceptes estan separats de la seva implementació, permetent la independència de les dades. Els conceptes espacials que venen descrits en el model espacial de les dades ho fan a través dels conceptes matemàtics de topologia, de mètrica i d'ordre. Creiem que probablement els conceptes espacials només han utilitzat en la generalització cartogràfica el concepte espacial mètric deixant de banda l'ordre i la topologia.

² El tipus d'informació que conté un diccionari de les dades es: identificació de l'informació per a els objectes de les dades (nom, sinònim.); informació sobre la representació (tipus d'atribut, llargada.); informació sobre la relació (nom, objectes relacionats.); informació estadística; informació sobre seguretat i informació funcional (Oxborrow 1989:230)

4.2. La integració de la informació

El procés d'integració de la informació espacial com ha assenyalat Nyerges no es basa en la dificultat en la disparitat de formats, sinó en "entendre el procés d'integració" basant-nos en la interpretació semàntica de les entitats per poder assegurar que fem coincidir de manera correcta aquestes. Es basa en la interpretació de la base de dades basant-nos en les classes d'entitats i en les seves descripcions, o sigui en les metadades, enlloc de fer-ho en els valors que hi ha enregistrats en la base de dades. L'altra manera de veure ha estat en base a veure els problemes d'organització de compilació i del seguiment processual que permeti corregir les diferències existents en les bases de dades. Com per exemple la integració que ve representada per la tècnica basada en la superposició de polígons, que seria la més propera a la cartografia tradicional.

Les tendències en la integració de dades heterogènies per poder ser utilitzades en les tasques d'anàlisi, ha estat feta des de fa anys des de diferents enfoc. El primer d'ells, ha consistit en reconciliar un conjunts de dades, per poder-ne fer la seva comparació en un context més general; el segon d'ells es basa en considerar els procediments necessaris en convertir les dades per les necessitats d'una aplicació concreta.

Nyerges basant-se en l'esquema proposat per el "Spatial Data Transfer Specification del The Proposed Standard for Digital Cartographic Data" que descriu tres nivells per la transferència de les dades espacials: un model conceptual de les dades espacials, una forma de transferència i un model de transferència. El terme aquí utilitzat de model conceptual es el que correspon a esquema conceptual en el camp de la enginyeria informàtica, i és important el fet de que en aquest model i/o esquema conceptual especifiquem els fenòmens del mon real segons el sentit en aquest i no en base a les dades i el seu emmagatzematge en la base de dades.

Nyerges es basa l'esquema de Elmasri et al. d'integració per l'anàlisi basat en els següents passos:

1 - Conversió dels esquemes d'implementació en esquemes conceptuais: utilitzant els mdoels de dades de ER; conceptuais que inclogui les representacions de les descripcions temàtiques, locatives i temporals. Hem de poder especificar l'esquema de sortida i el d'arribada.

2 - Esquema d'anàlisi i de les modificacions: identifica els tipus d'entitats i les seves relacions diferencies i les similituts i les diferències i si es necessari transformar els components de l'esquema que difereixen en sentit, es pot fer a traves de la comparació dels seus diccionaris. Un tipus d'entitat pot ser una relació en una altra, i un atribut en un esquema pot ser un atribut en un altre.

3 - Integració dels tipus d'entitats: que consisteix en la identifcació d'atributs, tipus d'entitas i classes de relacions, i la seva agrupació.

4 - Integració de les classes de relacions i d'entitats creats a partir de de les abstraccions de les dades: integrem les relacions una vegada hem identificat les entitats

5 - Creació de l'esquema integrat: a partir de la selecció dels tipus d'entitats que millor representen en aquestes i eliminem les que son redundants.

Cada un d'aquests estadis es poden descriure a tres nivells d'abstracció que son: comprensió conceptual; tècniques que es poden utilitzar per expressar conceptes en la construcció dels termes i les dades; eines del programari/maquinari per l'implementació de la construcció de les dades. Això ho resumeix en el pla d'integració en l'estadi 1; El "que" o quin tipus d'informació existeix i ha d'existir en l'estadi 2; el "com" dissenyar el contingut de la informació per ser emmagatzemat en l'estadi 3; i en com implementar la integració en l'estadi 4.

Un avantatge d'aquest tipus de metodologia es el presentar un marc de treball que permet guiar els estadis d'integració de manera independent de les tècniques i eines utilitzades per la integració, i de manera especial es un treball preventiu de cara la ampliació i posteriors necessitats de integracions successives.

Així doncs, normalment les dades no es passen a un format que faciliti el seu futur anàlisi, sinó que s'utilitzen altres criteris, i en els darrers anys s'han esbossat estratègies destinades a reduir les dades espacials i els seus atributs a un mínim denominador comú. Si volem que tots els atributs de les dades s'escalin en una escala d'ordre, només podrem fer-ho en un sentit partint de la més desagregat a mes gran agregació o menor resolució. En segon lloc, es proposa la conversió de les dades originals a una versió final única (*single target version*), es l'esquema d'integració que presenta Nyerges per múltiples bases de dades (Nyerges 1989).

La possibilitat d'accedir a bases de dades a escales múltiples, provinents de fonts diverses, fa que en la dècada dels 90', una de les àrees que centren l'interès, sigui el de la interoperabilitat entre dades o "activity relating to reliable interoperability among the data is one of the primary issues yet to be successfully resolved". Dues àrees d'investigació són importants aquí: la que busca les maneres de poder mesurar les disparitats entre les dades espacials de tipus heterogeni; i que en canvi en facilita el seu processament i conversió a partir de dades procedents de diferents escales i fonts. Tots dues possibilitats estan, per tant , directament lligades a la generalització i al problema de l'escala en l'entorn digital.

"Resolving heterogeneous data within the image domain requires generalization, but it must be coupled with semantic integration (classification) when dealing with the two different data model environments, namely raster and vector. When different scales or image resolutions are introduced into the integration scenario, spatial and semantic integration must be addressed in the context of scale integration, which involves both statistical and cartographic generalization. These processes are necessary to adjust the different levels of spatial and thematic content." (Richardson 1996:1).

Quan volem relacionar diferents bases de dades entre si, en el que s'ha vingut

definint com a "federació de base de dades heterogènies", la heterogeneïtat es pot definir en base segons sigui: sintàctica, a on cada base de dades correspon a diferents implementacions dels sistemes de gestió dels models de les dades (relacional, orientat a objecte) i de la seva representació espacial (raster o vector); esquemàtica, a on els objectes d'una base de dades poden estar en una i altre i, tenir diferents jerarquies d'agregació; semàntica, si una entitat en el món real pot tenir més d'una descripció en la base de dades. Actualment es tendeix a mantenir una heterogeneïtat esquemàtica que contingui referències o adreces estandarditzades en les bases de dades federades i informació geogràfica que en faciliti el seu intercanvi.

"Most of the current databases and digital geographic-information-exchange standards addresses syntactic and schematic heterogeneity. [...] However, even after users have transferred their geographic data to another database with the aid of these standards, further data structuring is usually undertaken in an effort to incorporate it into their own "schema" and semantics."

I assenyala a continuació el fet que:

" Merely supporting data-sharing with current standards will not make possible a high level of interaction among hierarchies described in the previous section on parallel object hierarchies. It is, therefore, important to support data-sharing on a semantic level." (Bishr et al. 1996:87)

En els sistemes d'informació, un fenomen de la realitat ve representat a través dels tres nivells en l'arquitectura del sistema: al nivell més abstracte o conceptual, a on el usuari identifica aspectes dels fenòmens i selecciona les seves característiques; al nivell de l'esquema lògic on als fenòmens seleccionats es defineixen com entitats; i finalment al nivell on s'emmagatzemen les dades, el que es considera la base de dades física, mitjançant determinades estructures i índexs que en permetin optimitzar tant la capacitat com la rapidesa de resposta. (Kucera, Flaherty 1996:96)

Kucera assenyala a continuació que, als sistemes d'informació com que es la comunicació entre els diferents nivells, aquesta no seria, problemàtica si existissin bases de dades estandaritzades. Però, quan les bases de dades no ho estan o no existeix un diccionari, el que comporten que "...there is no transparent way to move knowledge of real

world features across the interfaces to the application level of the database". Això fa que es tendeixi, a manipular les dades als nivells d'origen i a no moure-les o tractar-les a altres nivells. És precisament, aquesta arquitectura basada en els tres nivells: abstracte, lògic i físic, que ha fet que l'ús de models espacials estigui principalment dominat per les limitacions procedents de la representació dels objectes digitals, més que per les consideracions conceptuals i pels processos d'abstracció dels fenòmens del món real.

4.2.1. La integració de les dades

La integració de les dades és el pas previ a l'anàlisi, i és, per tant, essencial en el desenvolupament de qualsevol posada en marxa d'un sistema d'informació. Flowerdew resumeix les dificultats que ens podem trobar en la integració de les dades, en base a les preguntes de:

¿a quin tipus de dades ens estem referint?, ¿quins sistemes de mesura o d'escalat de les dades s'han utilitzat?; és un procés de selecció o de la presència/absència que fem de certs fenòmens del món real, així com de quina ha de ser la seva categorització, i identificació de quines mesures de quantitat o intensitat (*rank*, discretes i continues) s'han de establir

¿a quin lloc?, ¿a on es refereixen les dades?, ¿a quin tipus d'objecte espacial de punt, línia o àrea o superfícies, s'ha fet correspondre la observació i quin sistema de referenciació s'ha utilitzat?

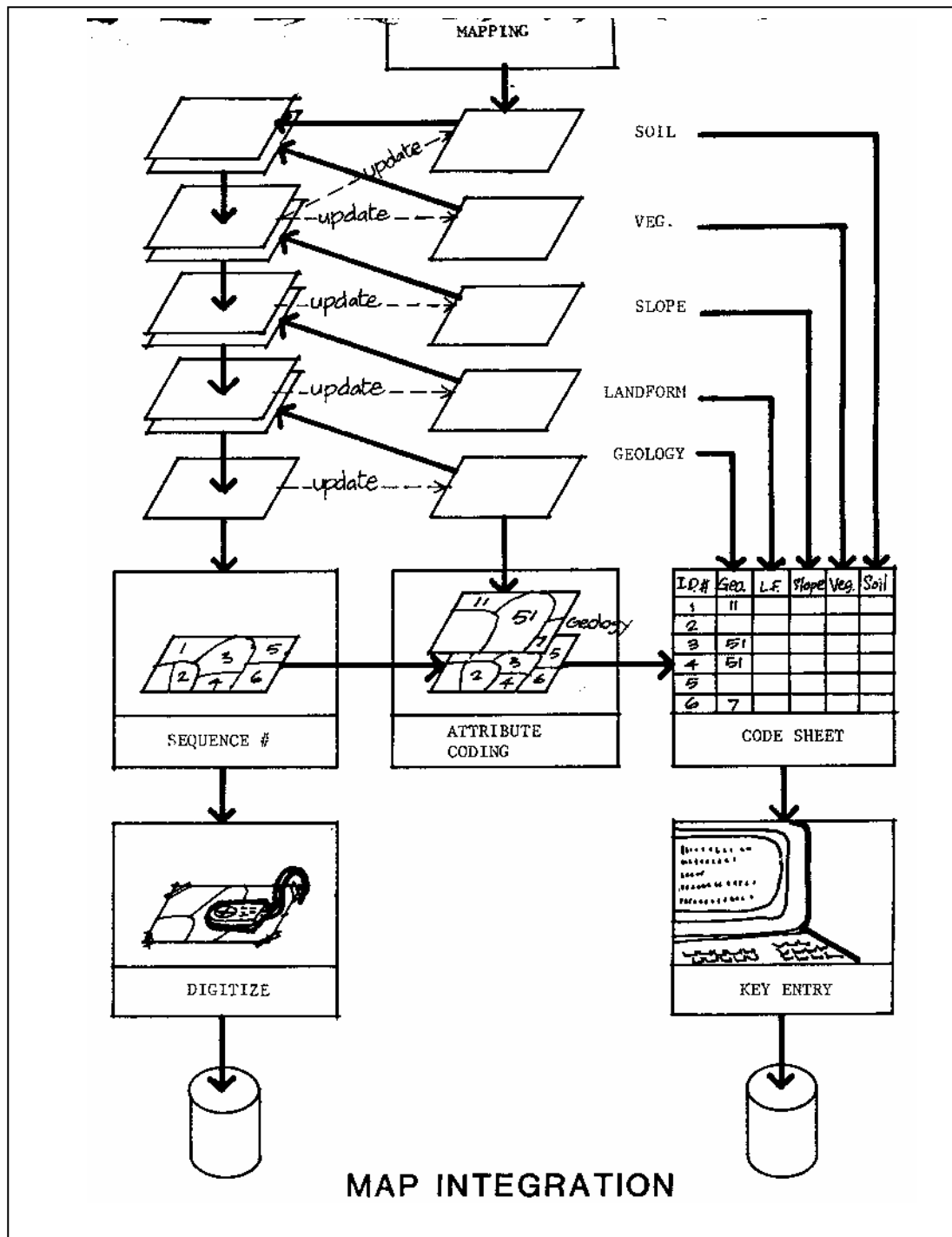
¿a quin moment es refereixen les dades?

¿quin grau de correcció tenen les dades?

En l'intent d'integració de dades espacials cal tenir en compte que la delimitació d'un fenomen pot canviar en funció de la definició de la unitat de tractament bàsica definim, és especialment rellevant en el moment d'analitzar les dades; els resultats de l'anàlisi depèn de quines han estat les unitats utilitzades de les àrees: és difícil integrar àrees que tenen unitats incompatibles (Flowerdew 1991:384).

Els problemes que apareixen al intentar integrar bases de dades (Shepherd 1991:342) consisteixen principalment en: diferències en l'enregistrament o tècniques de

medició, errors en la medició o mètodes d'aixecament topogràfic, variacions en la resolució, vaguetat o imprecisió en les definicions dels atributs espacials i temporals, poca definició (*fuzziness*) dels objectes gràfics espacials, i variacions en l'ús de la terminologia i de la seva nomenclatura. Tot això tendeix a agreujar-se quan diferents organitzacions són les que recullen la informació, quan la informació s'extrau de diferents fonts i a diferents escales, quan la informació és de diversos tipus: mapes, imatges, só, text; quan la informació és disponible a diferents llocs al mateix temps i es troba en perifèrics i sistemes diversos.

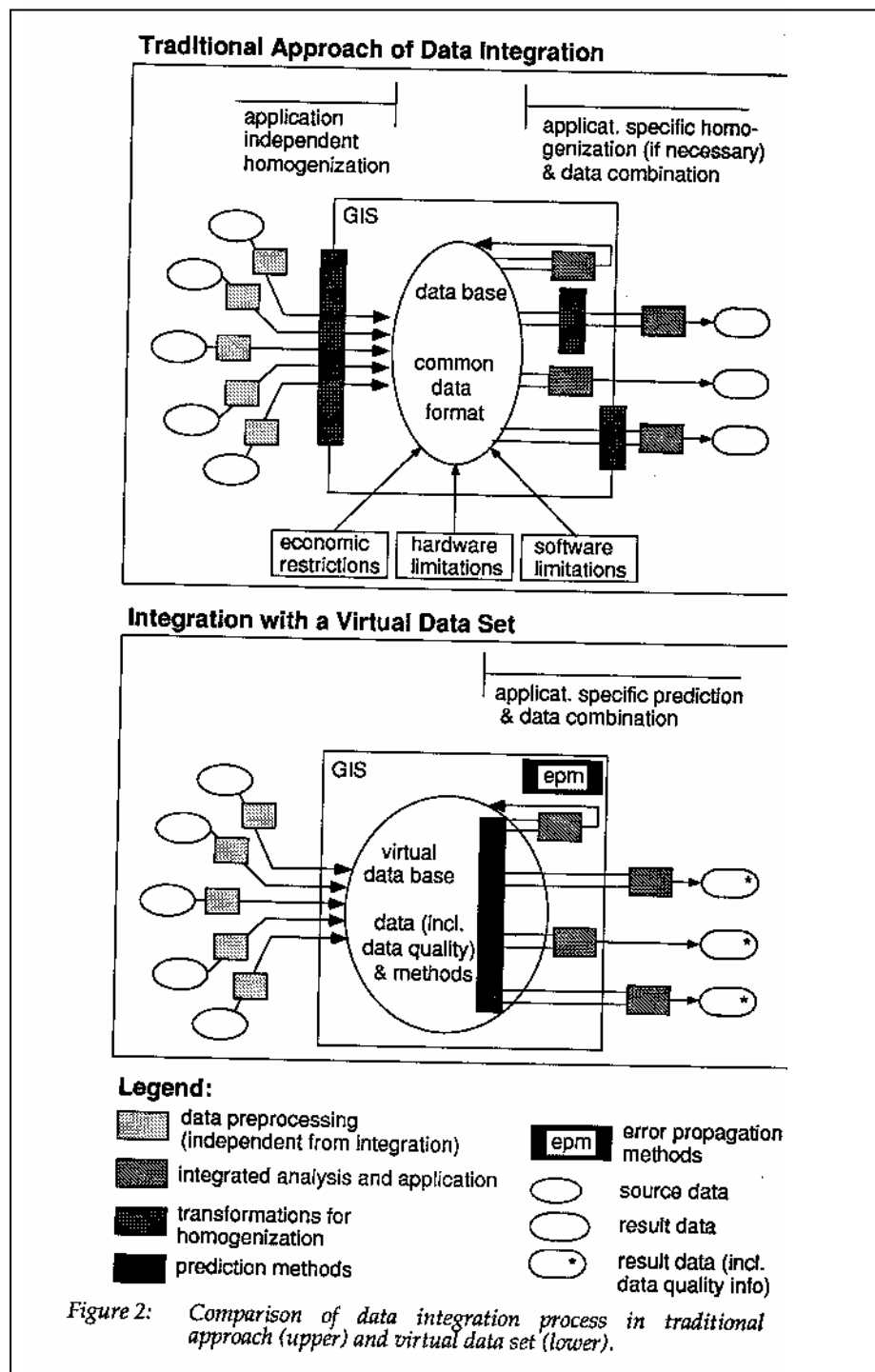


(Font: [Dangermond, 1988 #570])

4.2.2. La integració virtual de conjunts de dades

El SIG virtual permet la integració de dades incompatibles, (Stephan et al. 1993:93) i permet anar més enllà en el que ha estat l'intent d'integració de dades. Segons Stephan, totes les transformacions suggerides d'integració fins el moment actual, introdueixen incertesa, i no permeten fer una valoració de la qualitat amb la que s'han generat les transformacions o metallinatge. Tampoc conserven les dades, la informació addicional o metadades, que permetin millorar-ne el seu processament, i la manca de poder comptar amb els formats adequats fan necessari el seu posterior re-format.

Un conjunt de dades virtuals, en canvi, permet generar informació que no es físicament present, i per tant permet la independència de les dades de les aplicacions. O sigui que poden ser utilitzades en contextos diferents per usuaris varis. La figura de Stephan mostra la diferència entre el concepte tradicional de integració i el de integració virtual.



(Font: Stephan et al. 1993:100)

Establir el procés general d'integració de dades permet veure que les tasques corresponents a la generalització són:

- 1) les tasques menys clares, i subjectes al criteri de l'usuari
- 2) tasques que en la cartografia tradicional es realitzen a través de les representacions cartogràfiques i, per tant, al final del procés de disseny del mapa
- 3) tasques que, en la cartografia digital degut a la possibilitat de separació entre model de dades i de la seva representació, son convenientes de fer en el preprocessament

4.3. Un marc conceptual de treball per la integració de dades heterogènies en un SIG

La posta en marxa de un sistema d'informació geogràfica suposa una sèrie de passos que Aronoff resumeix de la següent manera: la informació espacial es antiquada o de poca qualitat; les dades espacials no estan emmagatzemades en formats estàndard; diversos departaments recullen i gestionen dades espacials similars amb el problema de redundància que suposa, les dades degut a qüestions legals i de confidencialitat no es poden compartir;; les capacitats d'anàlisi i de tractament son inadequades; hi ha demandes que no es poden cobrir amb les dades i tecnologies existents

La integració és un punt clau en la posta a punt de les funcions (principalment d'anàlisi) d'un sistema d'informació geogràfica. Una de les primeres propostes com a marc de treball, que desenvolupa els passos a seguir per la posta a punt d'un SIG que permet integrar dades procedents tant de les ciències de la terra, com dades procedents d'àrees socioeconòmiques, es el que desenvolupa l'ITC International Institute for Aerospace Survey and Earth Sciences, conjuntament amb l'UNESCO (De Man 1984).

És interessant d'aquest projecte preliminar, que pretén analitzar els lligams que hi ha entre les dades basades en els recursos naturals i les dades socioeconòmiques; i permet

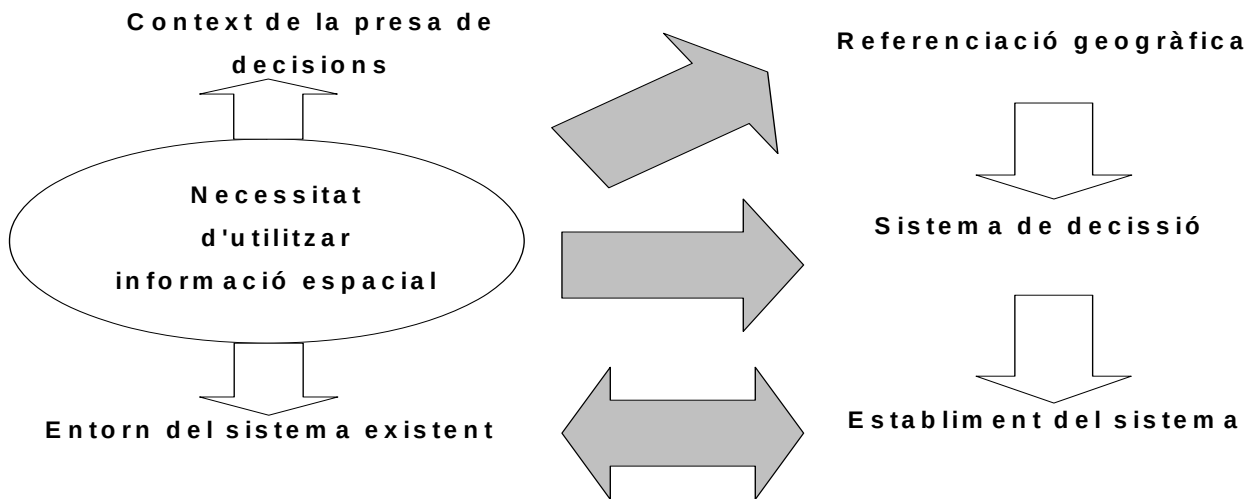
preparar unes línies de treball que poden ser utilitzades com a marc de treball, en el desenvolupament de bases de dades espacials, i que en facilitin la seva futura integració. Aquest treball es presenta el 1984, i pretén facilitar la estructuració de projectes, en un SIG.

Aquest marc de treball enfoca els SIG des de tres punts de vista diferents:

com a sistemes d'informació

la funcionalitat espacial específica que desenvolupen aquests

el permetre la integració de les dades espacials de diferent tipus



Marc conceptual de treball d'integració de dades de tipus heterogeni (físiques i socioeconòmiques) DeMan, 1984, ITC, Holland

4.3.1. La presa de decisions

El procés de presa de decisions és el que estableix quins són els condicionants de partida en funció del tipus d'aplicació.

4.3.1.1 Identificació dels usuaris de la informació

Els usuaris que han d'utilitzar un sistema d'informació tenen probablement una tipificació en un sistema anterior que es convenient de respectar, és convenient de fer inventaris dels organigrames ja existents i utilització de dades/informació, el tipus de plantejament que fem depèn sempre de la finalitat del sistema de informació.

És important de poder definir a l'usuari final i el nivell de dificultat del llenguatge que utilitzarà, tal com es veu en la figura següent que mostra la relació entre els llenguatges de la base de dades i l'usuari final.

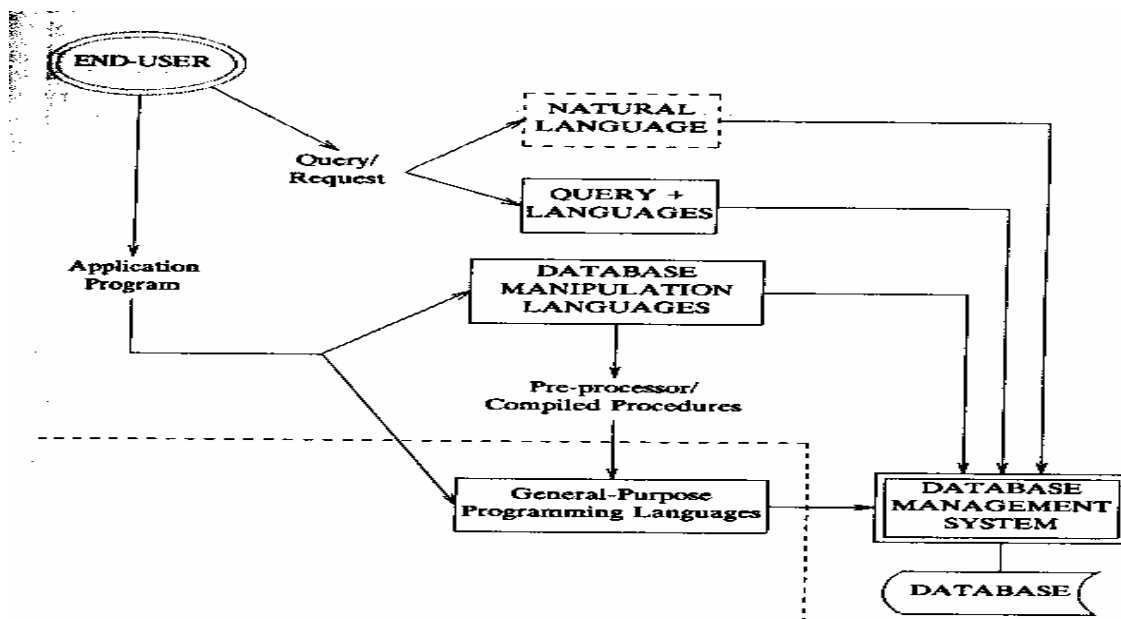


Figure 4.1 Database Languages and the End-User

(Font: Oxborrow 1989:99).

És de cabdal importància la manera com es relaciona l'usuari amb les dades, o sigui quin es el procés d'accés a les dades emmagatzemades en la base de dades; els llenguatges de consulta i els anomenats llenguatges naturals son preferibles a qualsevol llenguatge de programació o funció. Un gran avantatge del llenguatge natural es el fet de que es informal per naturalesa i permet ambigüitat al plantejar les consultes. Disposem de llenguatges naturals de consulta; llenguatges de consulta de tipus gràfic, formal o processual d'alt nivell; i llenguatges de manipulació de la base de dades o processuals de baix nivell. Un

llenguatge es considerat de tipus no-processual si permet a l'usuari especificar el què necessita enlloc de com fer-ho per aconseguir el que necessita.

4.3.1.2. Context d' utilització de la informació i les seves necessitat

Abans de començar el disseny d'un sistema d'informació, s'han de poder resoldre: els tipus d'utilitzacions, el grau de correcció, la precisió i resolució de les dades.

4.3.2 L'entorn del sistema existent

És convenient fer una recollida dels sistemes existents anteriorment de dades i sistemes de gestió abans de començar amb un nou sistema (d'altres sistemes d'informació; habilitats i tecnologies existents; situació existent en la gestió de les dades). És convenient tenir en compte tot tipus d'ajut, tant locals de programari, de maquinari com de recursos humans, així com és convenient que els sistemes de informació vinguin recolzats per les administracions i altres organismes públics

4.3.3. És necessària la informació espacial?

Només podrem considerar que un mètode de referència geogràfica es el apropiat, després de determinar les necessitats en termes de quin serà el tractament espacial de les dades, quin el nivell desitjat de detall espacial, i quines les necessitats específiques d'integració del conjunt de les dades. La tria d'un sistema adequat de geocodificació que ens permeti la presa de decisions fa que puguem imaginar si ens trobem dins dels inventaris regionals de petita escala o si ho fem en les grans escales dels inventaris locals. I per fer-ho ens podem basar en el concepte variació de la informació espacial. El nivell de detall espacial necessari quan dissenyem un sistema d'informació geogràfica acostuma a coincidir amb la unitat mínima que ens interessa a la hora de poder prendre una sèrie de decisions.

4.3.4. Referenciació geogràfica

Determinar el nivell de generalització i nivell òptim de referenciació geogràfica suposa identificar i considerar el disseny del sistema de informació, els tipus de utilització i el grau de correcció, de precisió i de resolució; fa referència a la variació espacial, que formaria part de els problemes que representa la tria del sistema apropiat de geocodificació. La dimensió òptima de les unitats espacials bàsiques d'una base de dades depèn de la precisió espacial que es necessita per analitzar un fenomen o procés. Unitats espacials bàsiques de gran escala són acceptables, “thus simplificant els sampling procedures needed to collect the appropriate data” i normalment, podem dir, que el nivell òptim és la unitat espacial més petita comuna a totes les bases de dades.

Degut al que l'escala tradicionalment s'ha considerat un concepte derivat de l'escala mètrica de les entitats espacials:

- el nivell òptim de detall espacial que hem de adoptar, representa un compromís entre els requeriments de cada un de les bases de dades, i que per millorar-ne la eficiència en el processament i emmagatzematge, les diferents bases de dades es poden considerar com conjunts de dades separats però coordinats.

- la geocodificació en el context de la presa de decisions representa: un mètode de referència que es pot seleccionar només després de determinar els requeriments en termes de quin serà el tractament espacial de les dades, el nivell desitjat de detall espacial i les necessitats d'integració de les dades específiques.

4.3.5. Sistema de decisió

Preveure per referenciació geogràfica idèntiques o mecanismes de transformació per conjunts de dades diferents, fent referència a com fer la integració i si és possible de fer-ho a partir de geocodificacions compatibles o de la seva conversió:

- sempre que es pugui, s'han d'adoptar esquemes de geocodificació compatibles en

la etapa de la recollida de les dades. Sinó es possible s'han de trobar mètodes que ens permetin transformar o ajustar els esquemes de geocodificació que s'han aplicat a un conjunt de dades i que permetin la seva coincidència (*matching*) amb altres bases de dades. Això s'ha de considerar en situacions específiques, ja que els esquemes mixtes de geocodificació, poden ésser menys cars, que tenir-ho tot en un sol esquema amb mètodes de geocodificació bàsics amb més nivell de detall, que tenir-ho amb sistemes relacionats

- alguns tipus de bases de dades, com les de recursos naturals, que tenen un grau de variació espacial gran i una *low decay rate* són bases de dades bàsiques apropiades per la integració en els sistemes de informació

- quan s'integren bases de dades espacials que difereixen en termes d'escala i/o detall, s'ha de tenir en compte de informar als usuaris de les limitacions inherents de les dades

- un sistema integrat de informació geogràfica s'ha de poder dissenyar de manera modular. Pot ésser útil de preveure un "servei de transformació de dades" dels conjunts de dades "bàsics" i "relacionats" per la seva referenciació geogràfica

- un sistema de informació geogràfica integrat, pot ser compost de bases de dades separades però coordinades, tenint cada un dels conjunts un sistema de referència diferents. Aquestes bases de dades, un o més es poden seleccionar com a bàsics mentre que la resta es consideren bases de dades relacionades. Les bases de dades bàsiques tenen una influència més forta en la metodologia per poder integrar les altres bases de dades.

En el disseny de la informació, hem de tenir en compte la relació existent segons el tipus de utilització que fa referència als usuaris i als usos dels sistemes integrats de informació espacial:

- els sistemes de informació que serveixen a un grup heterogeni de clients/usuaris han de fer front a les diferents situacions i han ésser flexibles i adaptables per permetre una fàcil inclusió i consulta de informació detallada. Això fa necessari a vegades algun tipus d'estandardització dels conceptes i formats en la recollida de dades

- els sistemes de informació pot gestionar o avaluar projectes específics i els programes s'han de poder fer en base al disseny d'aquests projectes i programes; de igual

manera els sistemes de informació mes generals han de poder gestionar i avaluar estructures de gestió i presa de decisions dels seus usuaris.

La gestió i estructures de presa de decisions de informació dels usuaris ha de poder:

- gestionar i permetre l'establiment d'un conjunt de indicadors-clau i d'un sistema de informació que els pugui utilitzar de manera ràpida

- quan es dissenya un sistema de informació, la "informació informal" de que es disposa no s'ha de desestimar; aquesta informació s'ha de utilitzar al màxim possible

Quan fem el balanç entre varis components dels sistemes d'informació:

- és convenient un desenvolupament equilibrat de un sistema de informació i dels seus components (base conceptual, recollida de les dades, emmagatzematge i procés, possibilitats analítiques)

- un sistema integrat de informació geogràfica s'ha de dissenyar de manera modular. Pot ésser útil de preveure un "servei de transformació de dades" manual o automatitzat

Un sistema descentralitzat ha de tenir en compte que:

- els sistemes de informació descentralitzats, tenen en general un gran potencial per facilitar una presa de decisions local i sectorial. La descentralització requereix una formulació i és necessària una regulació acurada dels lligams entre la unitat central i els nodes perifèrics

Un canvi d'estratègia representa:

- conjuntament amb la selecció d'un sistema de computació, una tria acurada d'estratègies per el canvi i l'adaptació que s'han de desenvolupar. Procediments per la recollida de dades que han de dependre el mínim en els sistemes de maquinari per facilitar-ne canvis futurs.

La millora dels mètodes manuals cap als sistemes assistits per ordinador :

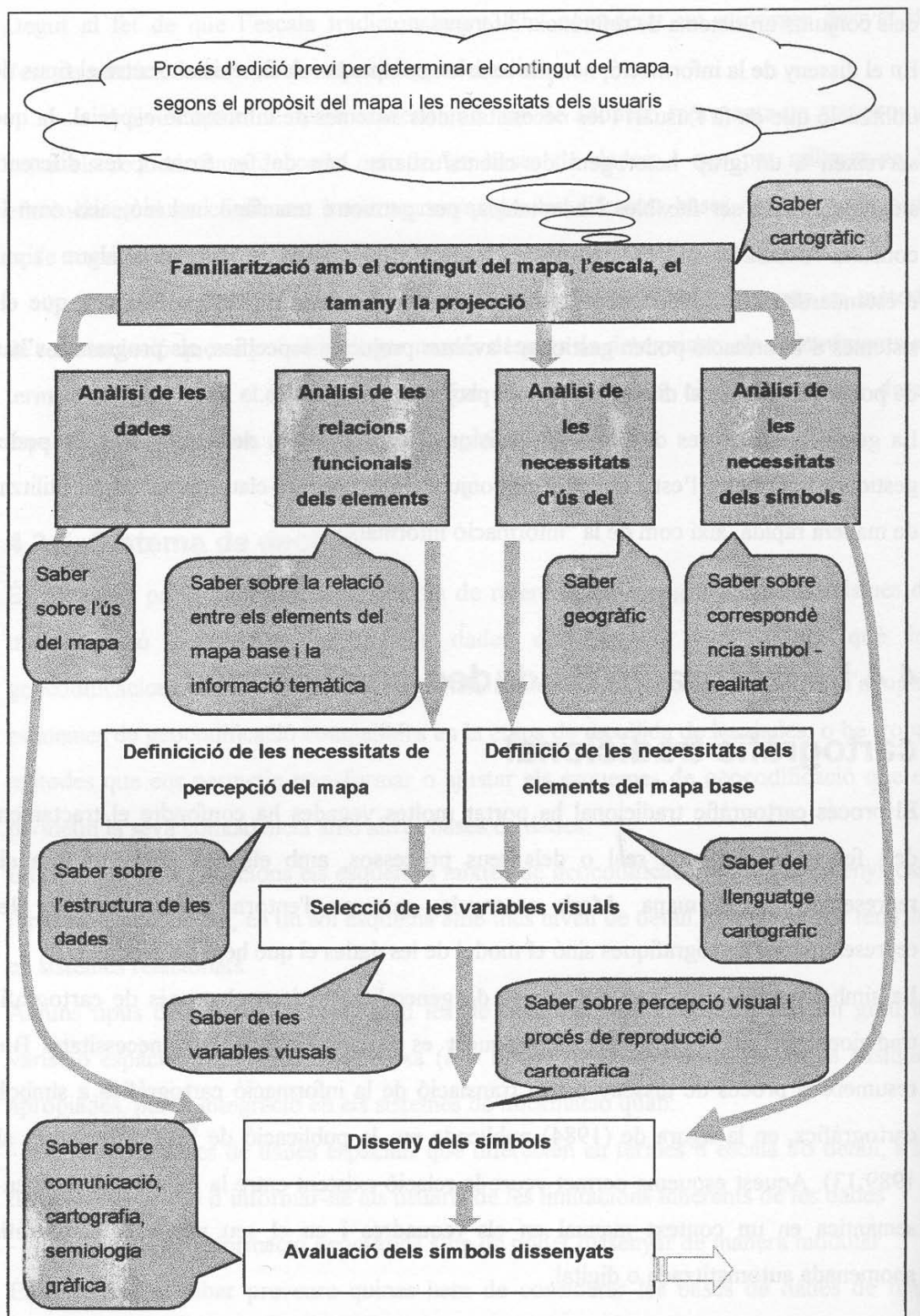
- és millor que els sistemes de informació es desenvolupin dins de contextos administratius i institucionals, per poder així supervisar-ne el pas del mètode manual en sistemes automatitzats.

4.4 La integració de dades en el procés cartogràfic

tradicional

El procés cartogràfic tradicional ha portat moltes vegades a confondre el tractament dels fenòmens del món real o dels processos amb els dels elements que els representen en el mapa. Mark assenyala, que en l'entorn digital, no són les representacions cartogràfiques sinó el model de les dades el que hem de generalitzar (Mark 1991:)

La simbolització forma part del procés de generalització i en el procés de cartografia tradicional permet l'anàlisi visual, i aquest es fa preveient les necessitats d'aquest. Bos resumeix el procés de disseny per la translació de la informació cartogràfica a símbols cartogràfics, en la figura de (1984) publicada en la publicació de l'ITC (Brown et al. 1989:13). Aquest esquema permet veure la relació existent entre la informació mètrica i semàntica en un context manual en els requadres i en el seu pas a la cartografia anomenada automatitzada o digital.



Esquema per la translació de la informació geogràfica a símbols cartogràfics i saber necessari (a partir de Brown et al. 1989)

Els factors que intervenen en el disseny dels símbols del mapa, no son arbitraris: segueixen uns passos clarament definits. En el rerafons de cada un dels factors hi ha la finalitat del mapa, i en son principals, les característiques de les dades geogràfiques, les necessitats perceptives dels nivells visuals. En l'esquema el primer pas correspon a la decisió sobre el contingut del contingut del mapa.

"Usually the map content has been determined by the map editor prior to the design of the symbols. In the map editing process also decisions have to be taken as to the scale of the map, its size and the map projection to be used. The cartographer responsible for the design of the symbols should in the first place familiarize himself with the content of the map, which is so closely related to the content. This first step is so obvious that it needs no further discussion." ([Bos, 1984 #567]:83) A continuació es prenen decisions com son el tipus de codificació, nivell de mesura per aconseguir un bon disseny, "The analysis of the geographical data on its dimensional properties, its organisational structure and on its measurement levels is an essential condition for good symbol design. What perception levels are required for a particular map should be one of the conclusion to be drawn from the data analysis."

Un punt important en el disseny del mapa tradicional es la importància que té en el procés tradicional in que en l'entorn digital acostuma a desaparèixer. Aquesta decisió es trobaria en aquest pas tercer de realització del mapa, i correspon a situar els elements "descontextualitzats" per nosaltres en funció del tipus d'anàlisi que hem decidit i "creem" les relacions amb un pla de referència conegut i familiar per l'usuari. Així ho assenyala Bos:

"A map is no complete if the theme of the map is not related to a topographic reference base, the content of which depends on the purpose of the map and on the thematic information itself. Although in general the base map elements will get a lower level of perception than the thematic elements (figure-ground) there may be some variation in perception level among the base elements. Certain elements need to be stressed more than others, depending on their function in the total map."

En el pas quart podem triar les variables visuals però només després d'haver decidit els nivells de percepció en la finalitat del mapa.

El model de comunicació que ha prevalgut en la cartografia i encara ho fa en la seva incorporació en els SIG pretén que el mapa produeixi una resposta resultant del procés visual de percepció, en la persona que està mirant un mapa. Tal com assenyala Bos a partir de l'anàlisi de la informació, seleccionem les variables visuals, composem uns símbols i combinem aquests símbols per produir mapes.

"We have learned that visual perception is much more than simply understanding the meaning of a symbol. It also includes the total understanding of what is mapped, including the mutual relations between symbols. It also includes the effectiveness of the visual variables and the cartographic symbols to build up meaning."([Bos, 1984#567]:61)

		quantitative Skalen	geordnete Skalen	trennende Skalen	≠	≡
a) <u>Punktförmige Zeichen</u>						
G	Größenverhältnis	●	⊙	⊙	○	
H	Helligkeit	○	●	⊙	○	
C	Farbe	○	○	●		○
O	Orientierung	○	○	●		○
F	Form	○	○	●		○
b) <u>Linienförmige Zeichen</u>						
G	Größenverhältnis	●	⊙	⊙	○	
H	Helligkeit	○	●	⊙	○	
C	Farbe	○	○	●		○
O	Orientierung	○	○	●		○
M	Musterform	○	○	●		○
K	Korn	○	●	⊙		○
R	Rasterwinkelung	○	○	●		○
c) <u>Flächenförmige Zeichen</u>						
H	Helligkeit	○	●	⊙	○	
C	Farbe	○	○	●		○
M	Musterform	○	○	●		○
K	Korn	○	●	⊙		○
R	Rasterwinkelung	○	○	●		○

●	geeignet	●	kaum wirksam
●	etwas weniger geeignet	○	nur Zeichen für Zeichen lesbar
●	wenig wirksam	⊙	gibt höheres Niveau vor

Nur die Anwendung einer Variablen mit einer dem Niveau der Information analogen Niveau erlaubt ein spontanes Erfassen der Information.

Els objectes digitals son enregistrats segons la caracterisitca dimensional i la característica d'atribut. Segons Spiess, en el model cartogràfic basat en la comunicació, les variables visuals s'han de correspondre amb el nivell d'informació (Spiess 1982:37).

En el procés de d'acomodar la percepció a les variables visuals es complica si tenim en compte que els nivells de percepció que definim els construïm a partir de la combinació de varies variables visuals. Les regles tradicionals de combinació d'aquestes variables estan ben establertes, i actuen com a ponderació: segons el tipus de propietat de les variables

visuals, comuns o absents en aquestes, s'han de fer més fortes o s'han de difuminar.

Hi ha per tant un procés hermètic que permet la manipulació de la informació per part del cartògraf, i és evident que el usuari d'un SIG desconeix en la seva totalitat.

La concepció del mapa, durant el procés de compilació es complicat, i junta normes molt estrictes i deixa altres essencials en les mans del cartògraf. I és això el que ha estat durament criticat del model cartogràfic de comunicació ja que representa sota un aire de gran consistència interna una sèrie de decisions externes que son primordials en la posterior presa de decisions.

Una d'aquestes seria el que fa referència a l'"enquadrament" que es fa segons el "tema", i que porta associada l'escala d'anàlisi i de nivell de detall. Spiess suggereix que el procés de compilació es basi en una estructuració exhaustiva de quina es la intencionalitat i finalitat del mapa, que posteriorment, a traves de la simbolització permet adequar fenòmens del món real en un espai d'anàlisi a on definim aquests mitjançant la simbolització. Aquesta permet "ajuntar" en un mateix espai fenòmens o processos que en el món real no tindrien una importància no sempre evident. El poder comptar amb aquesta disposició sistemàtica de les preguntes, i de com representem els fenòmens o processos i de quin tipus de mesura podem donar-los-hi representa moltes vegades un pas difícil del procés cartogràfic. Així es mostra en el quadre de com es dissenya la compilació de la informació inicial del mapa.

Després de clarificar si es possible de cartografiar el que ens interessa podem passar a decidir el format i l'escala i projecció convenients. Podríem dir que l'escala d'anàlisi del mapa mental l'hem decidit en aquest estadi previ, que Bos representa de manera molt clara en el model cartogràfic de comunicació, i que desgrana en el següent diagrama.