



Representação de dados geoespaciais

Édipo Henrique Cremon

edipo.cremon@inpe.br

SER-350-3 & CAP395
Introdução à Geoinformática

O que são dados geoespaciais?

Dados geoespaciais são informações que possuem uma **referência explícita à localização** no espaço geográfico. Diferente de dados convencionais, eles não apenas descrevem um fenômeno — eles respondem também a pergunta fundamental: *onde?*

Localização

Todo dado geoespacial está associado a coordenadas ou referências espaciais no mundo real.

Atributos

Além do "onde", carregam informações descritivas sobre o fenômeno observado.

Tempo

Muitas vezes incluem uma dimensão temporal, registrando como algo muda ao longo do tempo.



Mapa de John Snow – A importância da localização



Em 1854, uma grave epidemia de cólera devastou o bairro de Soho, em Londres. O médico John Snow foi pioneiro ao cartografar cada óbito por endereço, revelando um padrão espacial que mudaria para sempre a história da epidemiologia e da saúde pública.

Snow plotou manualmente centenas de casos de cólera nas ruas ao redor da Broad Street, identificando uma concentração inequívoca de mortes ao redor de uma única bomba d'água. Sem microscópios ou testes laboratoriais, a análise espacial foi sua única ferramenta — e foi suficiente.



Mapa de John Snow – A importância da localização



Em 1854, uma grave epidemia de cólera devastou o bairro de Soho, em Londres. O médico John Snow foi pioneiro ao cartografar cada óbito por endereço, revelando um padrão espacial que mudaria para sempre a história da epidemiologia e da saúde pública.

Snow plotou manualmente vários de casos de cólera nas ruas ao redor da Broad Street, identificando uma concentração inequívoca de mortes ao redor de uma única bomba d'água. Sem microscópios ou testes laboratoriais, a análise espacial foi sua única ferramenta e foi suficiente.

Mapa de John Snow – A importância da localização



→ Dados geospaciais conectam eventos a coordenadas

Cada morte foi registrada com sua localização exata, transformando registros individuais em padrão coletivo visível no mapa.

→ A visualização espacial revelou a fonte da epidemia

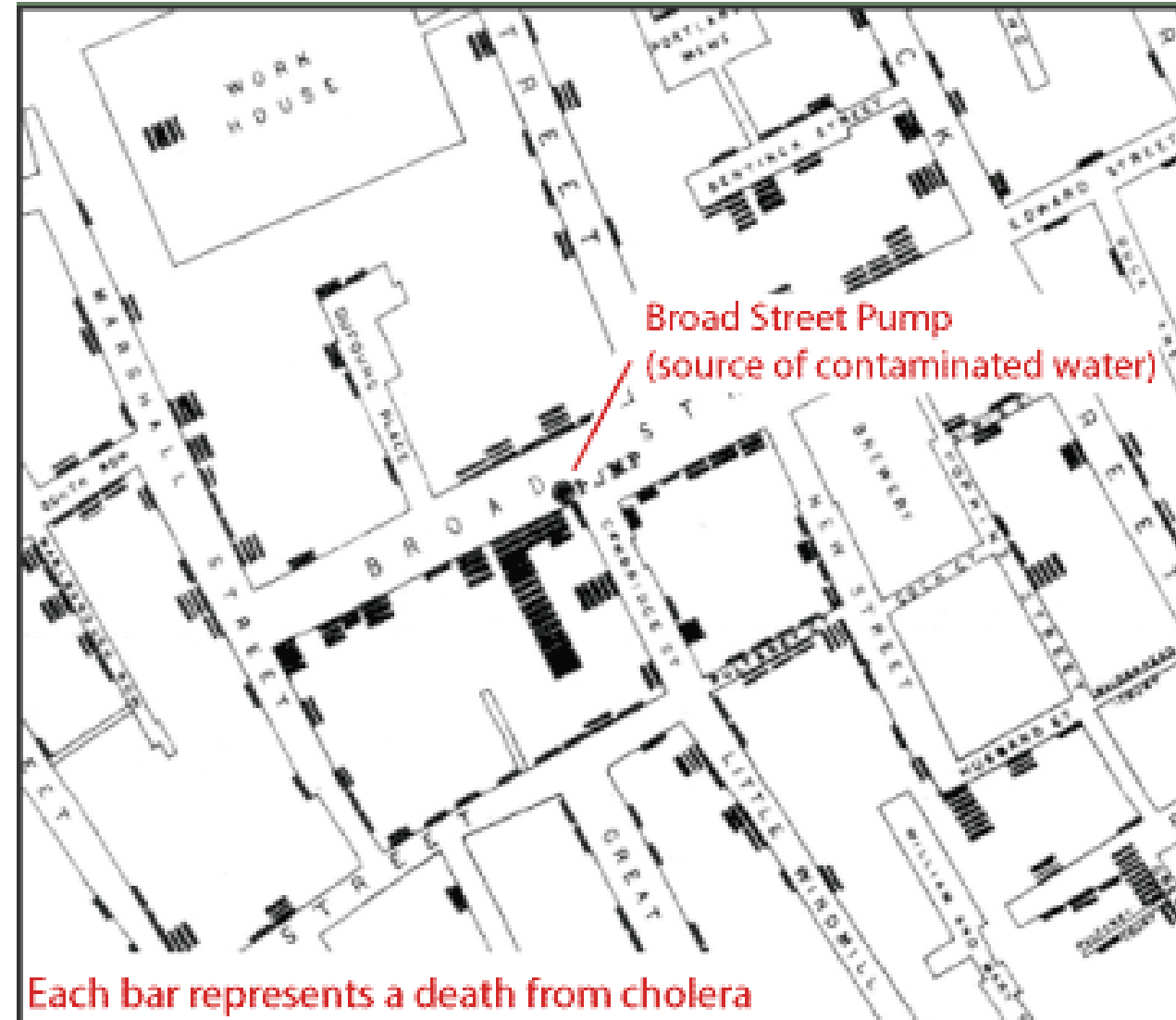
O agrupamento de casos ao redor da bomba da Broad Street apontou inequivocamente para a água contaminada como origem do surto.

→ Análises espaciais orientam políticas públicas

A remoção do cabo da bomba interrompeu a epidemia — prova de que decisões baseadas em dados geospaciais salvam vidas.



John Snow found that cholera deaths were clustered near the Broad Street water pump



Source: John Snow 1965 (originally 1854)

Mapa de John Snow – A importância da localização



Os dados geoespaciais do mapa de John Snow estão disponíveis em vários repositórios e formatos:

```
install.packages("HistData")  
library(HistData)  
data(Snow.deaths)
```



<https://blog.rtwilson.com/john-snows-famous-cholera-analysis-data-in-modern-gis-formats/>

http://www.rtwilson.com/downloads/SnowGIS_v2.zip

<https://geodacenter.github.io/data-and-lab/snow/>

Mapa de John Snow – A importância da localização



Te enganaram! O Mapa de John Snow não é dele!

<https://youtu.be/xdtH7CXuv28>

<https://www.esri.com/arcgis-blog/products/arcgis-pro/mapping/something-in-the-water-the-mythology-of-snows-map-of-cholera/>

Dados Convencionais vs. Dados Geoespaciais

A diferença entre dados convencionais e geoespaciais vai além do formato — trata-se de uma mudança na **capacidade analítica** e no potencial de responder perguntas sobre o território.

Dado Convencional

Temperatura: **32 °C**

- Descreve um valor isolado
- Sem contexto espacial
- Não permite análise territorial
- Exemplo: registro em planilha simples

Dado Geoespacial

Temperatura: **32 °C em 15°S, 47°W**

- Valor associado a uma localização
- Permite mapeamento e interpolação
- Habilita análise de padrões regionais
- Exemplo: estação meteorológica georreferenciada

📄 A adição da dimensão espacial transforma um simples número em um dado com poder analítico territorial.

A Importância da Localização nas Decisões



Estima-se que aproximadamente **80%** dos dados utilizados em processos de tomada de decisão possuem algum componente espacial (seja explícito ou implícito).

<https://www.gislounge.com/80-percent-data-is-geographic/>

<https://povesham.wordpress.com/2010/02/22/the-source-of-the-assertion-that-80-of-all-organisational-information-is-geographic/>

Espacialização de dados não-espaciais

Nem todo dado nasce com coordenadas geográficas, mas muitos podem ser **vinculados ao espaço** por meio de processos de espacialização. Endereços, CEPs, nomes de municípios e códigos administrativos são exemplos de referências que permitem converter dados convencionais em geoespaciais.

1

Dado Original

Endereço textual, CEP, nome de cidade ou código IBGE de município

2

Geocodificação

Processo de associar a referência textual a coordenadas geográficas precisas

3

Dado Geoespacial

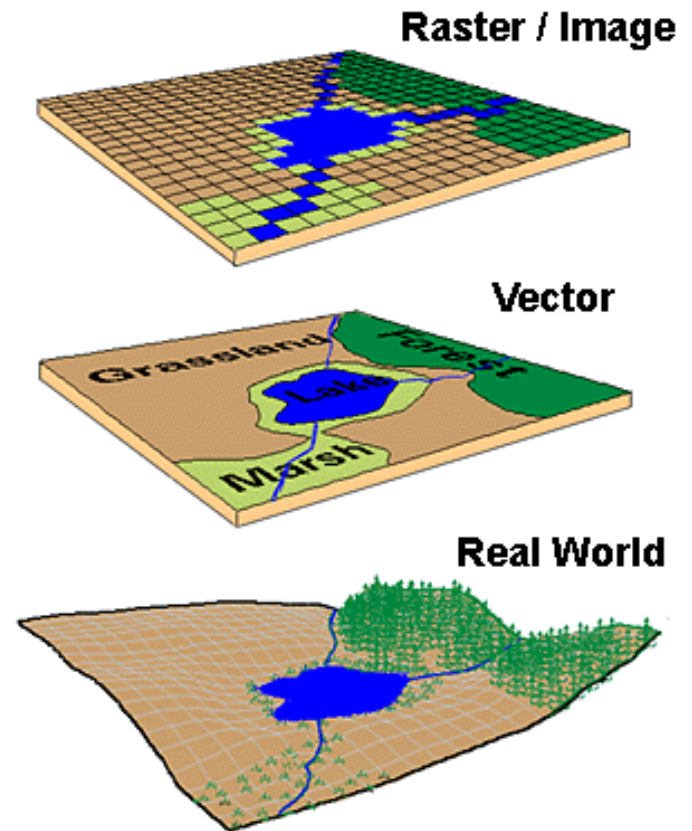
Ponto, polígono ou linha com atributos descritivos e localização georreferenciada

Dados Geoespaciais

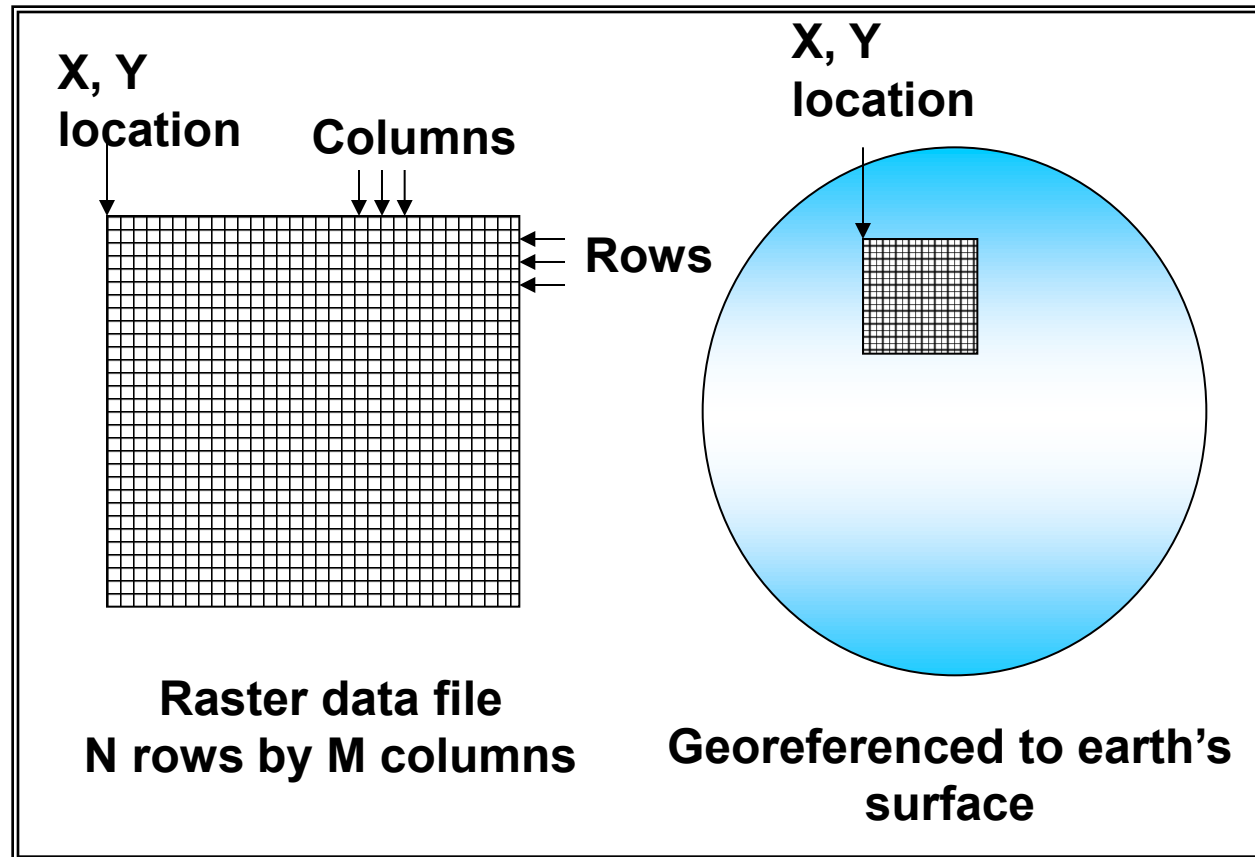
- Representação Matricial (raster)
- Representação Vetorial

Matricial (raster): refere-se a representação gráfica do mundo real através de células ou pixels (picture elements) que são definidos pela suas posições em relação as colunas e linhas de uma malha.

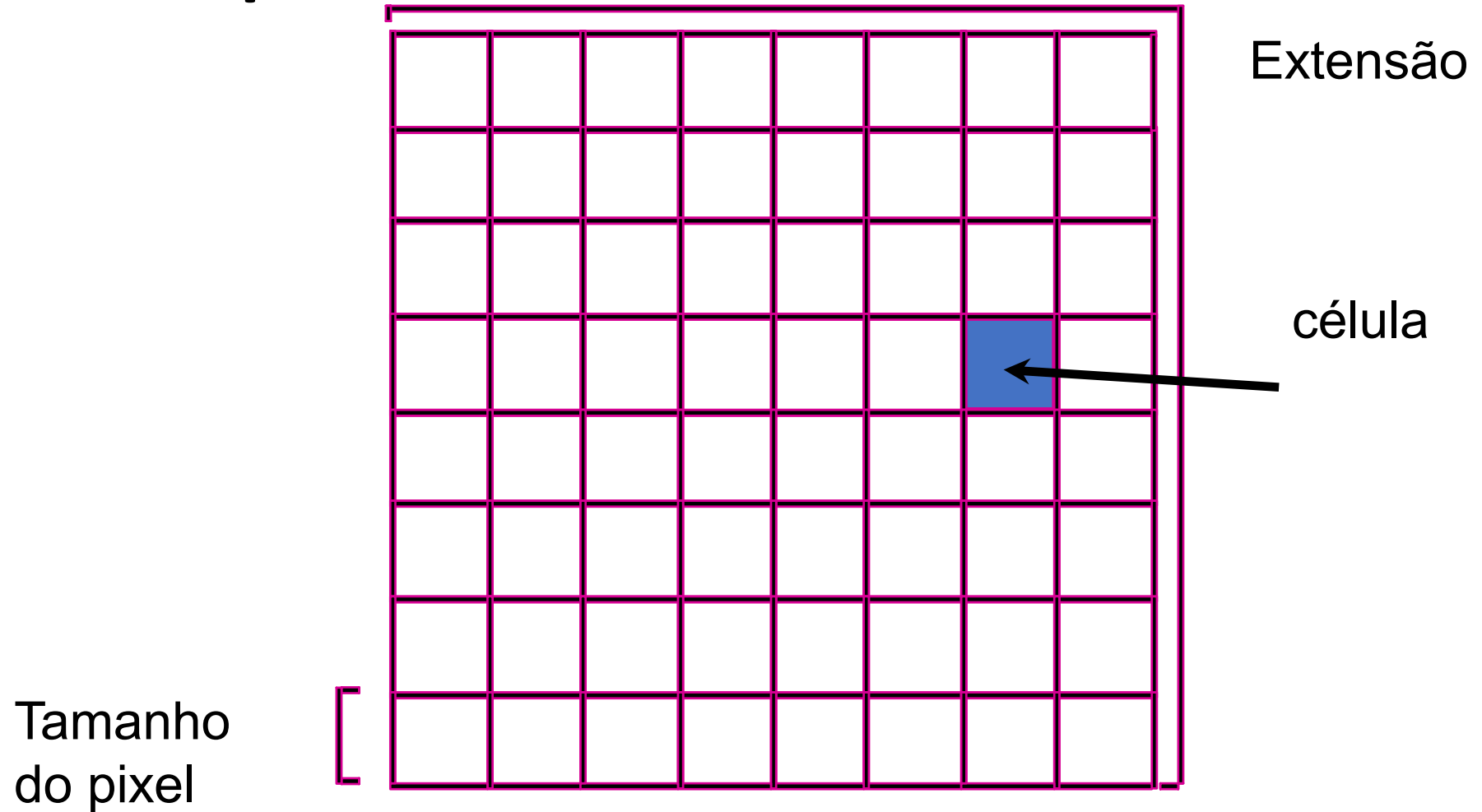
Vetorial: é a representação gráfica do mundo real através de sistemas de coordenadas. Dessa forma a unidade fundamental é o par de coordenadas x, y .



Dados Geoespaciais - Raster

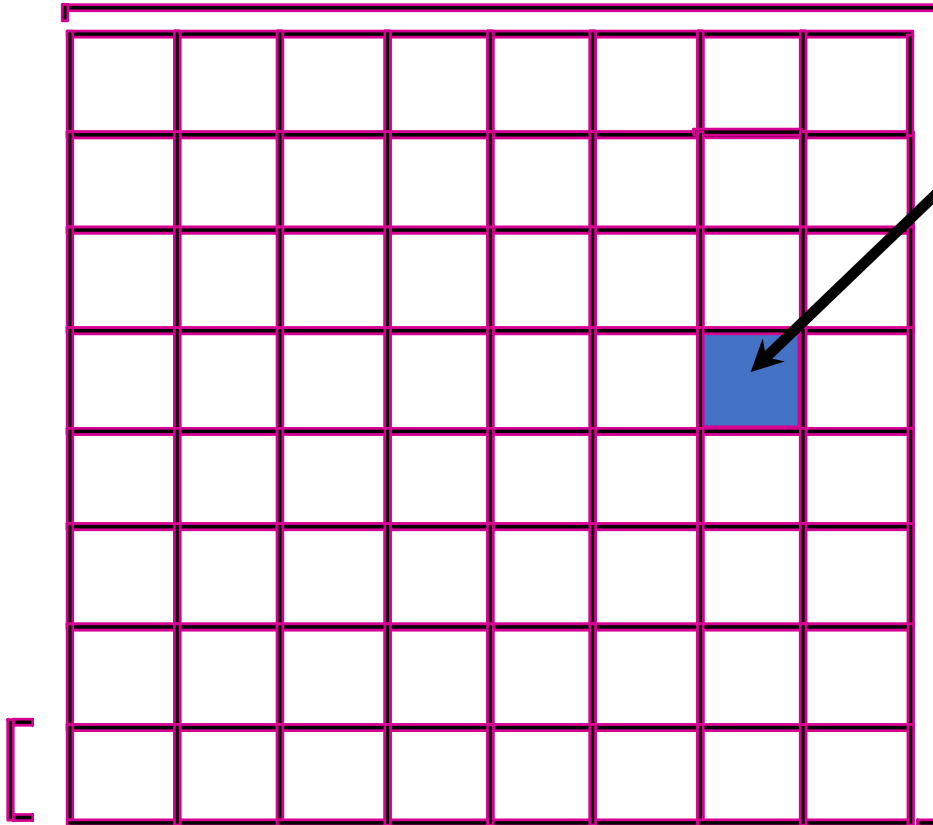


Dados Geoespaciais - Raster



Fonte: Mohamed Yagoub

Dados Geoespaciais - Raster



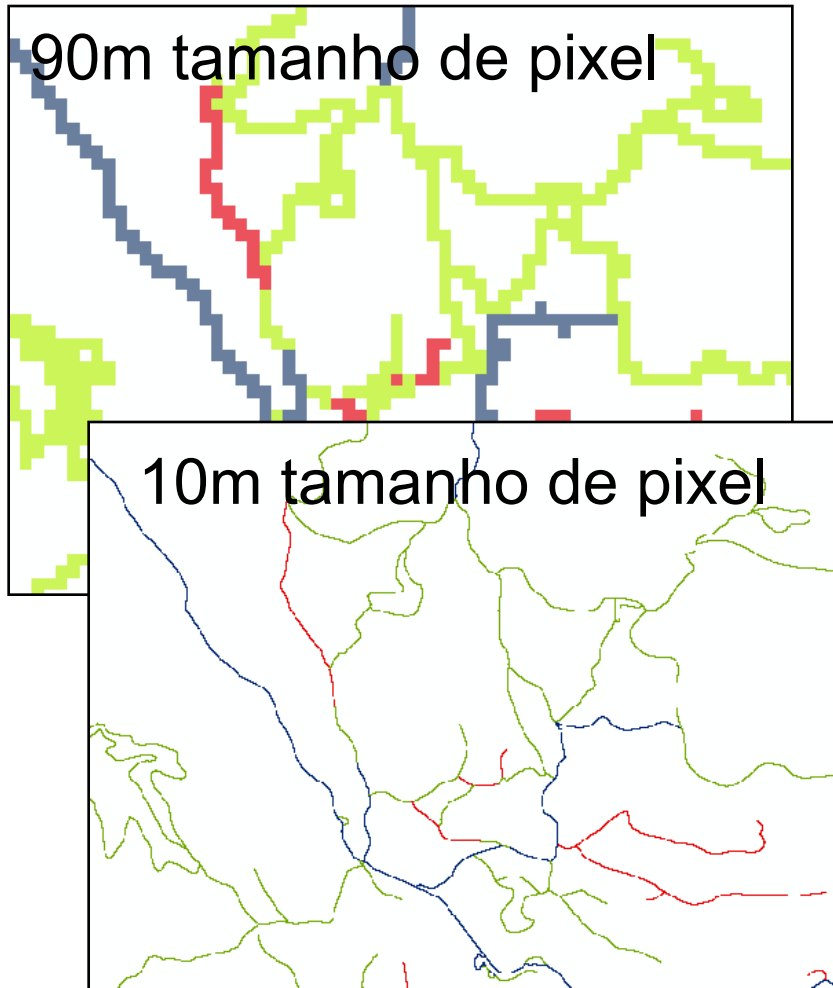
Célula

Qualidades: “3 - Alto, 1 - baixo”
(temático)

Quantidades: teor de cobre
(numérico)

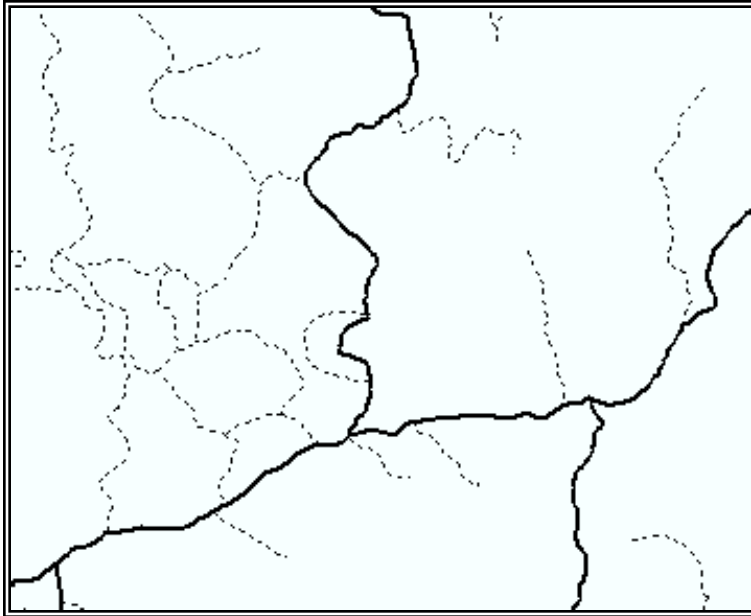
Fonte: Mohamed Yagoub

Dados Geoespaciais - Raster

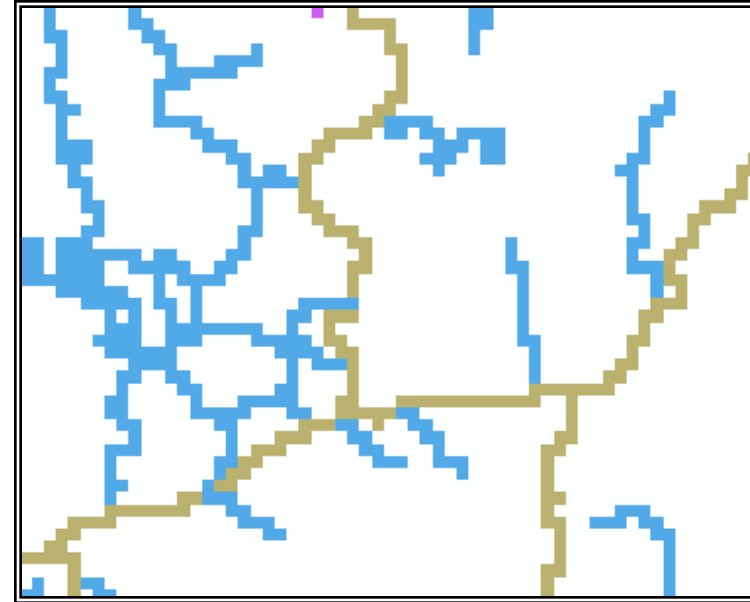


- Espaço de armazenamento aumenta pelo tamanho do pixel;
- Representar grandes áreas com detalhes é problemático.

Dados Geoespaciais - Raster



Modelo vetorial

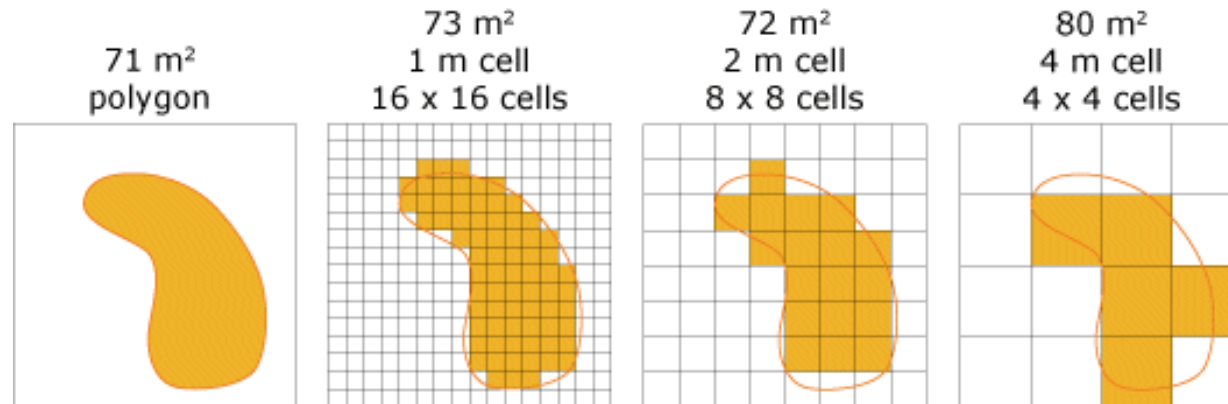


Modelo matricial

Dados Geoespaciais - Raster

Tamanho do pixel

O tamanho do pixel está relacionada ao tamanho da célula no raster e a área por ela coberta no terreno.



- Smaller cell size
- Higher resolution
- Higher feature spatial accuracy
- Slower display
- Slower processing
- Larger file size

- Larger cell size
- Lower resolution
- Lower feature spatial accuracy
- Faster display
- Faster processing
- Smaller file size

- Superior-> áreas e distâncias mais precisas
- Armazenamento maior

- Inferior -> áreas e distâncias menos precisas
- Armazenamento menor

Dados Geoespaciais - Raster



Água domina

A	A	G
A	A	G
A	A	G

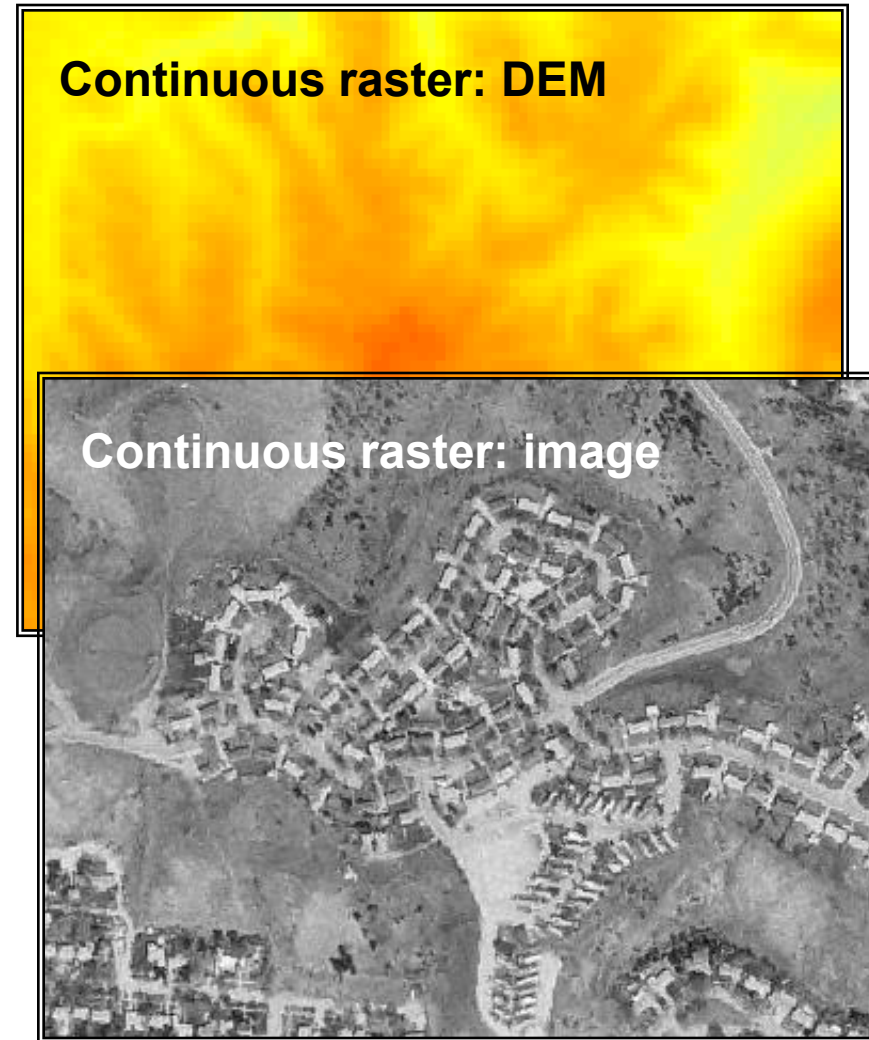
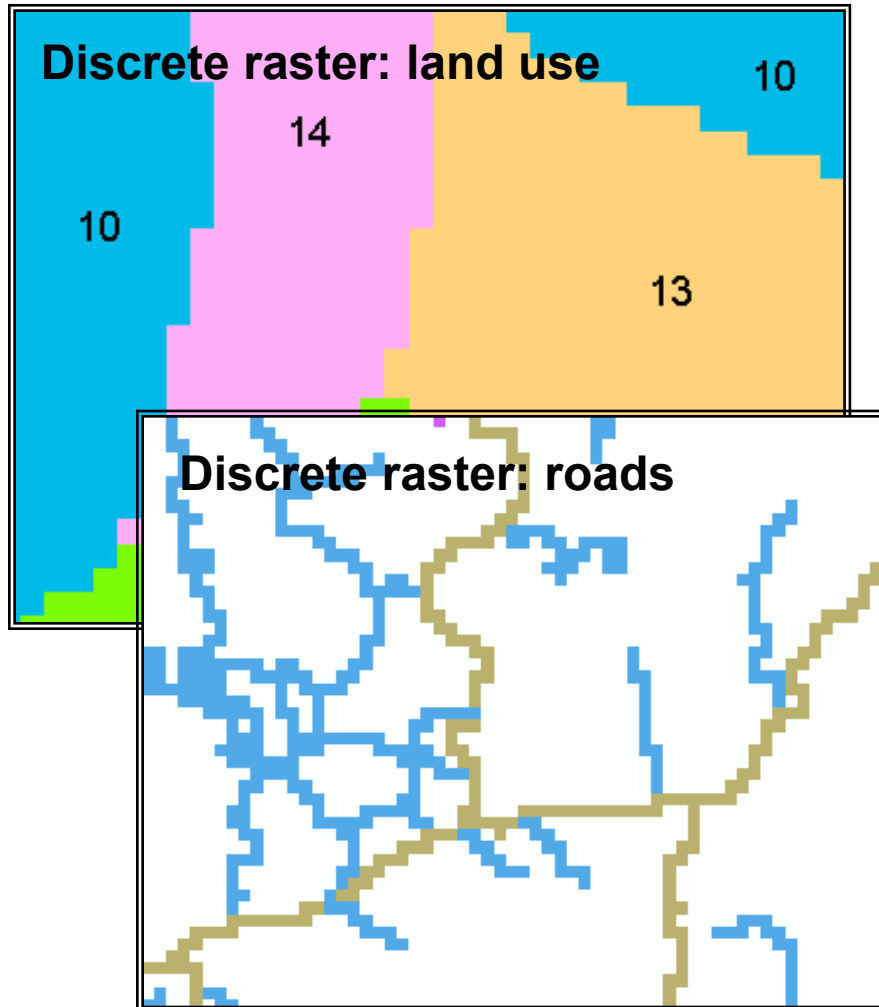
Maioria

A	G	G
A	A	G
A	G	G

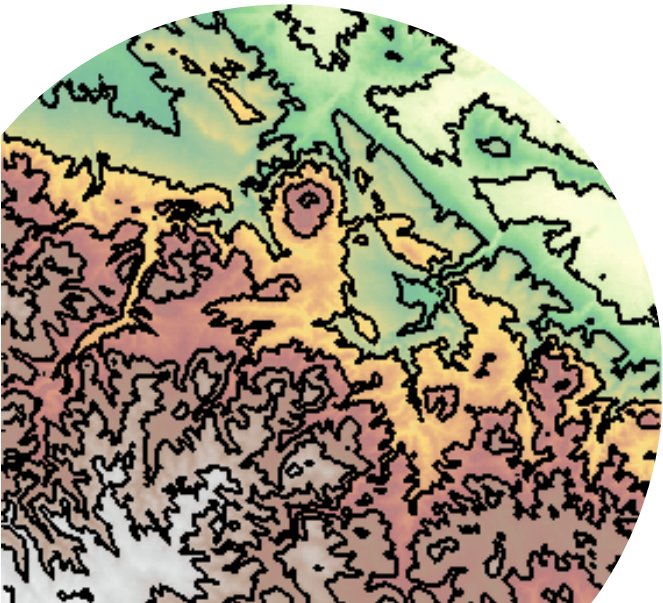
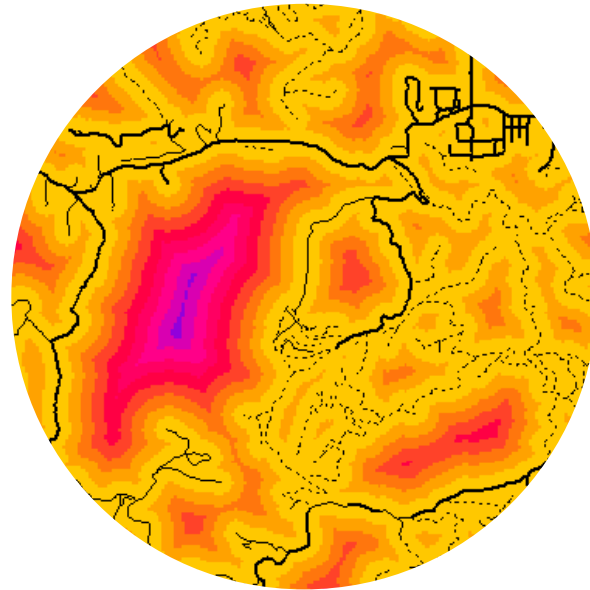
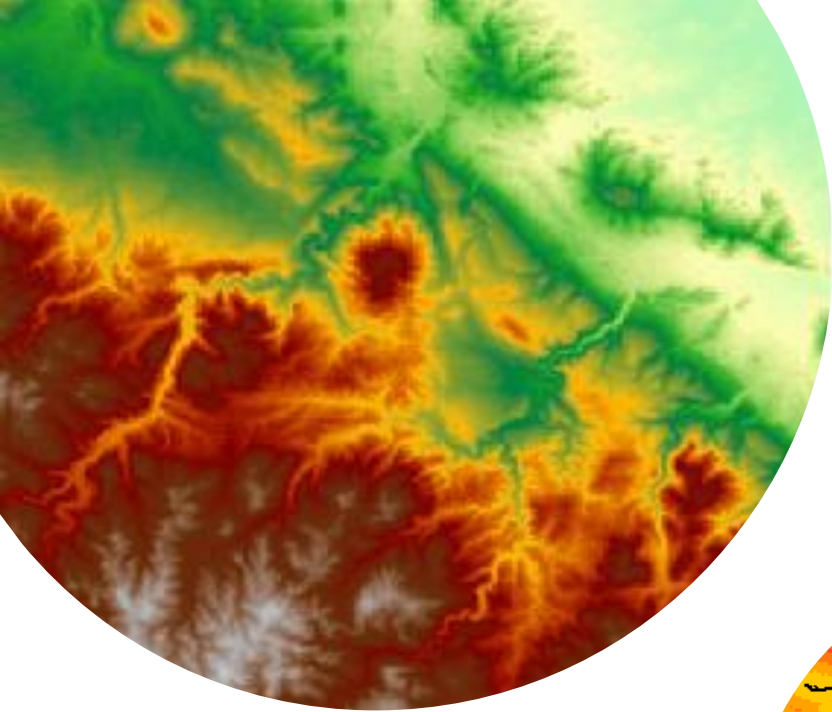
Bordas

A	B	G
A	B	G
B	B	G

Dados Geospaciais - Raster



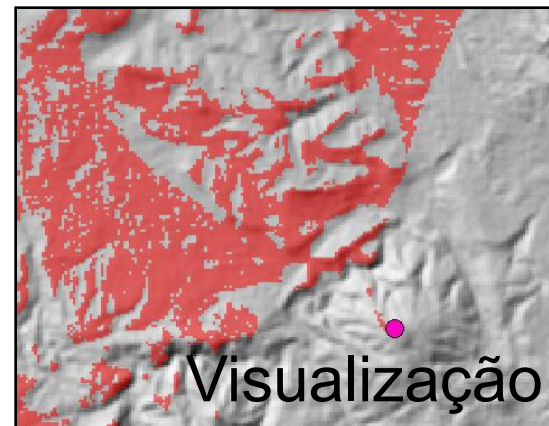
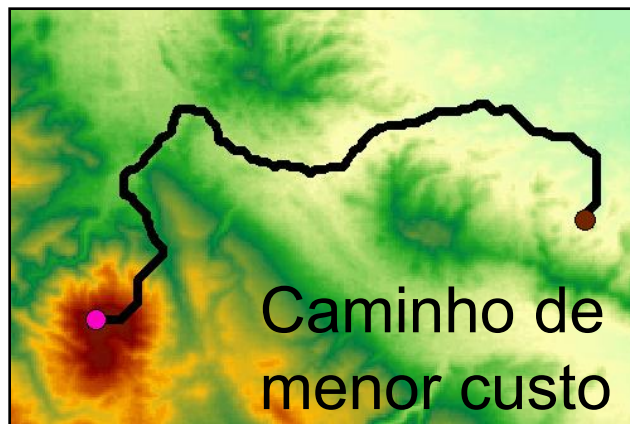
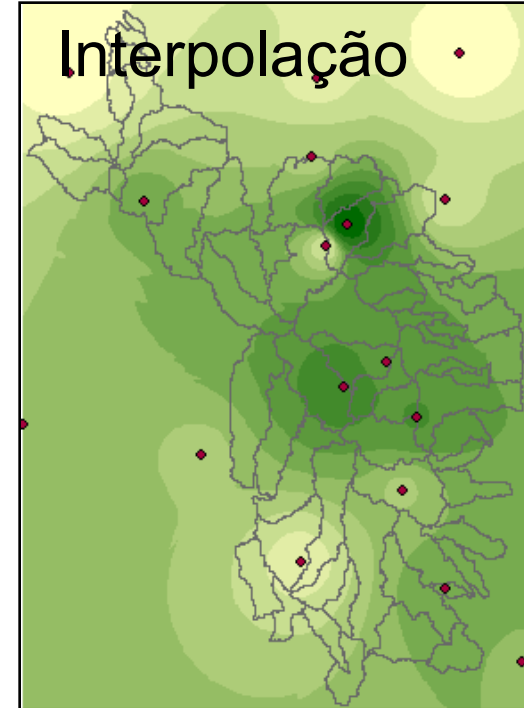
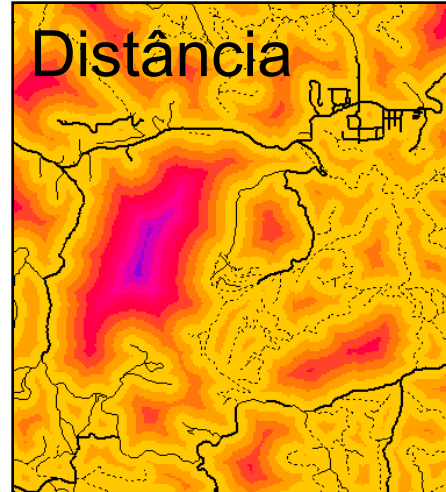
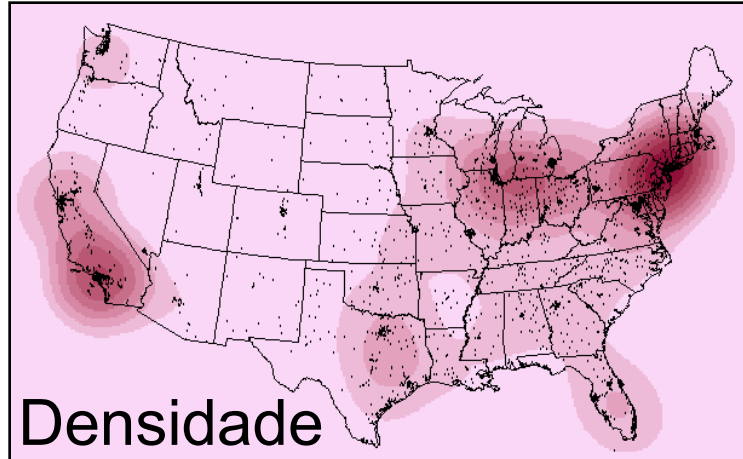
Dados Geospaciais - Raster



- Raster é a melhor maneira de armazenar valores que mudam continuamente como elevação;
- Análise mais rápida e flexível do que vetores para muitas aplicações;
- Algumas análises só é possível usando raster

Dados Geoespaciais - Raster

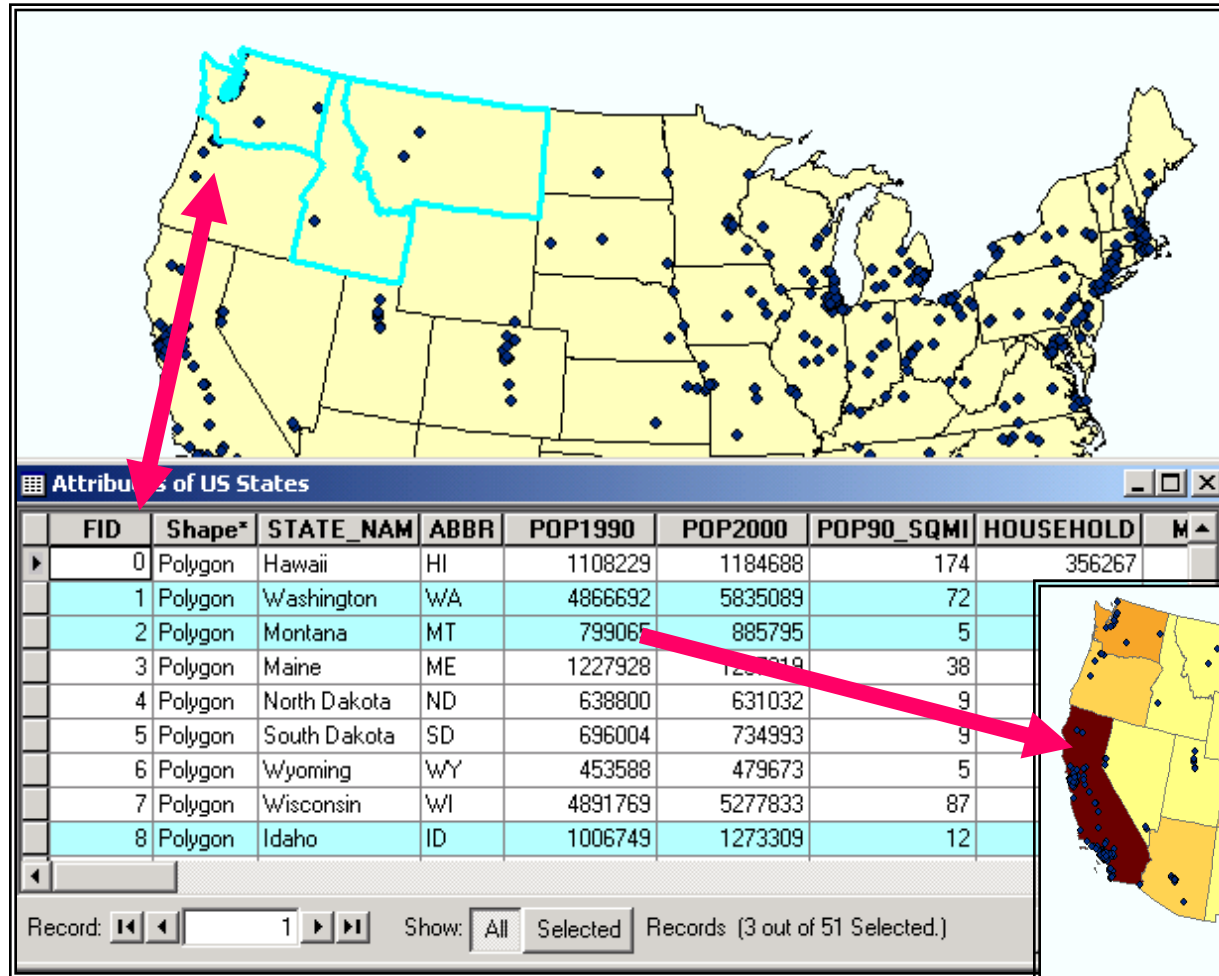
Análises



Principais formatos - raster

Formato	Principais aplicações	Vantagens	Desvantagens	Lançamento
GeoTIFF	Desktop GIS	Formato raster padrão na maioria das aplicações GIS Mantém alta qualidade mesmo após reamostragens ou reprojeções	Utiliza muito espaço de armazenamento comparado a outros formatos	1992
JPEG 2000	Desktop GIS, WebGIS	Relativamente leve, sendo ideal para aplicações que exigem escalabilidade e acessibilidade	Perda de qualidade das imagens após compressão	1999
IMG	Desktop GIS	Boa performance de acesso e escalabilidade Desenvolvido e otimizado para imagens de satélite	Formato proprietário da ERDAS, não sendo ideal para transferência de dados	1990
ECW	Desktop GIS	Boa performance de acesso e escalabilidade Compressão dos dados sem perder qualidade Desenvolvido e otimizado para imagens de satélite	Formato proprietário da Hexagon Geospatial, não sendo ideal para transferência de dados	1998

Dados Geoespaciais - Vetor



Cada vetor é relacionado a uma tabela de dados que contém informações sobre a feição.

ISO 19125



As **Simple Features** (Feições Simples) é um conjunto de padrões que especifica um modelo comum de armazenamento e acesso de feições geográficas compostas principalmente por geometrias bidimensionais (ponto, linha, polígono, multiponto, multilinha, etc.) usado por bancos de dados geográficos e sistemas de informações geográficas.

É formalizado tanto pelo Consórcio Geoespacial Aberto (OGC) quanto pela Organização Internacional de Normalização (ISO).

O padrão ISO 19125 é dividido em duas partes:

Parte 1 - define um modelo de arquitetura comum para **Simple Features** bidimensionais, com interpolação linear entre vértices, definidos em uma hierarquia de classes; esta parte também define a representação da geometria em formas de texto e binária.

Parte 2 - visa definir um esquema padrão de SQL que suporte o armazenamento, recuperação, consulta e atualização de coleções de feições. Uma feição (feature) possui atributos tanto espaciais quanto não espaciais. Os atributos espaciais possuem valores geométricos, e as *simple features* são baseadas em entidades geométricas bidimensionais ou tridimensionais (ponto, curva e superfície) com interpolação linear ou planar entre vértices. Este padrão depende dos componentes arquitetônicos comuns definidos na Parte 1 deste padrão.

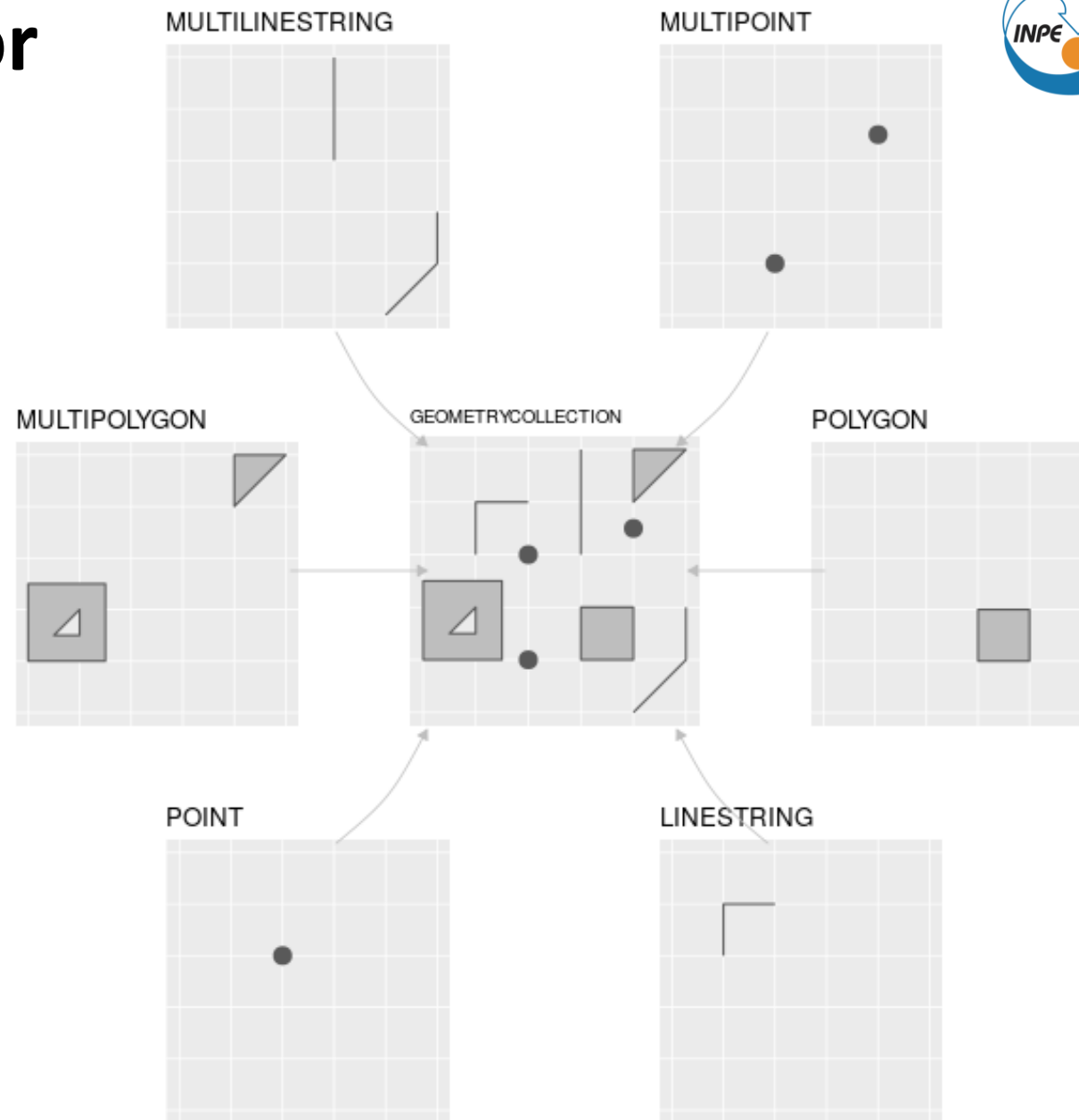
Dados Geoespaciais - Vetor

Em SIG, os dados vetoriais correspondem a basicamente 3 tipos de geometrias:

- Ponto(s)
- Linha(s)
- Polígono(s).

Mas ainda é possível ter dados com geometrias em:

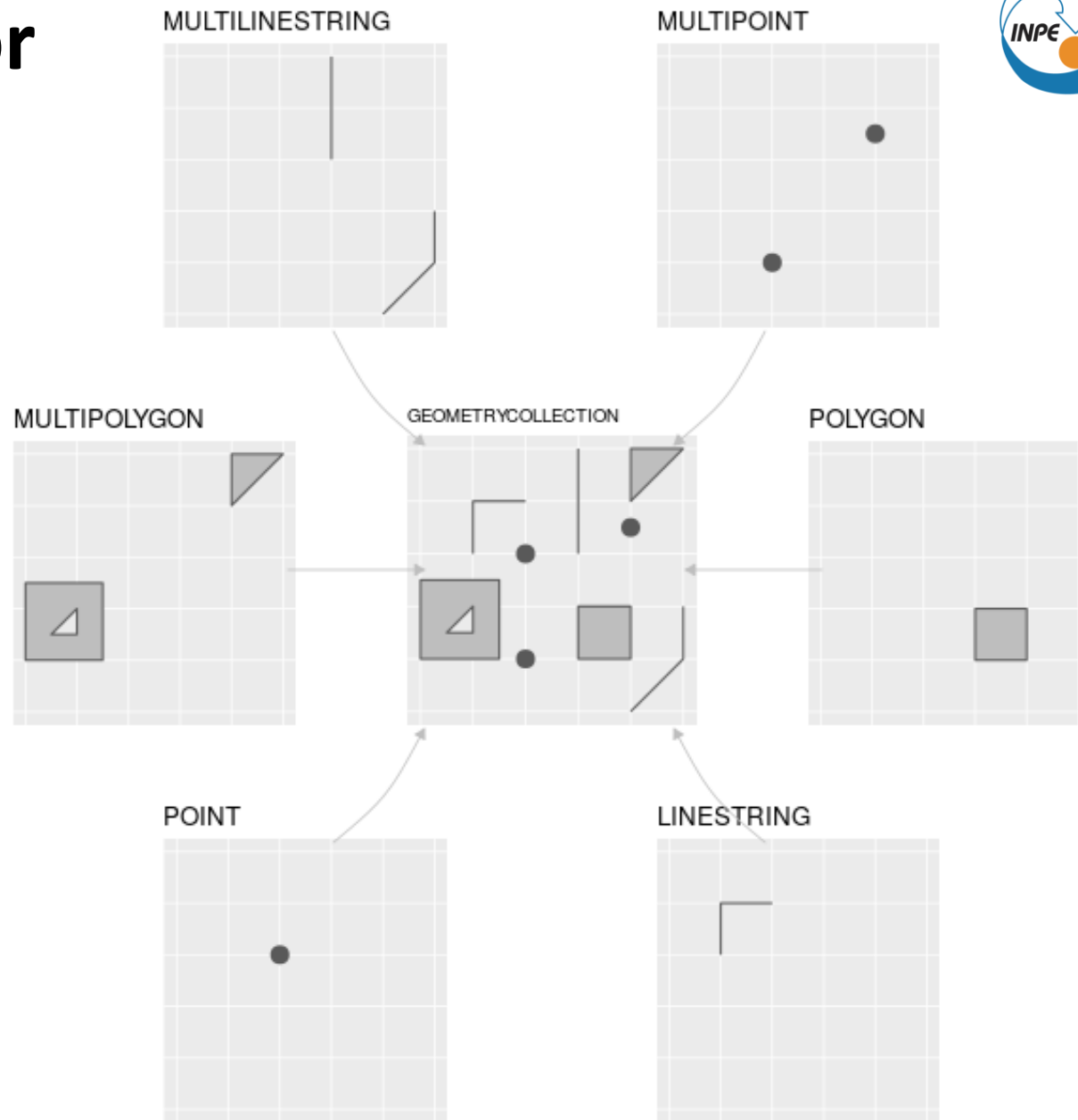
- Multi-pontos
- Multi-linhas
- Multi-polígonos
- Coleção de geometrias



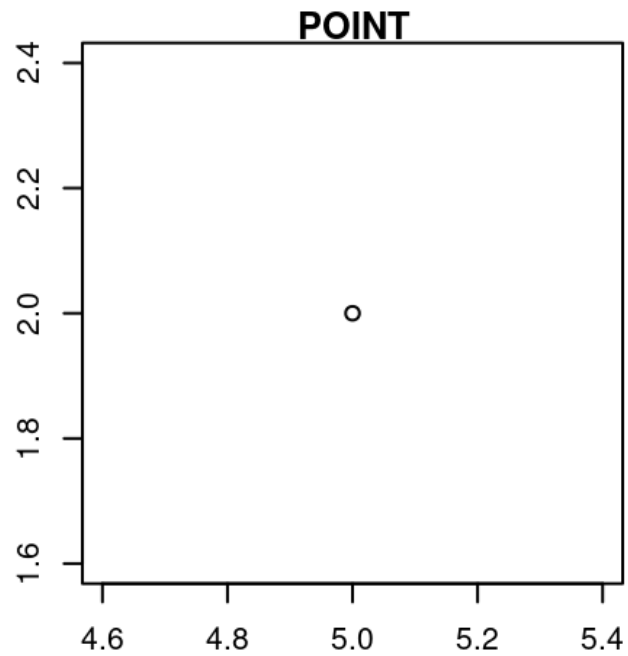
Dados Geoespaciais - Vetor



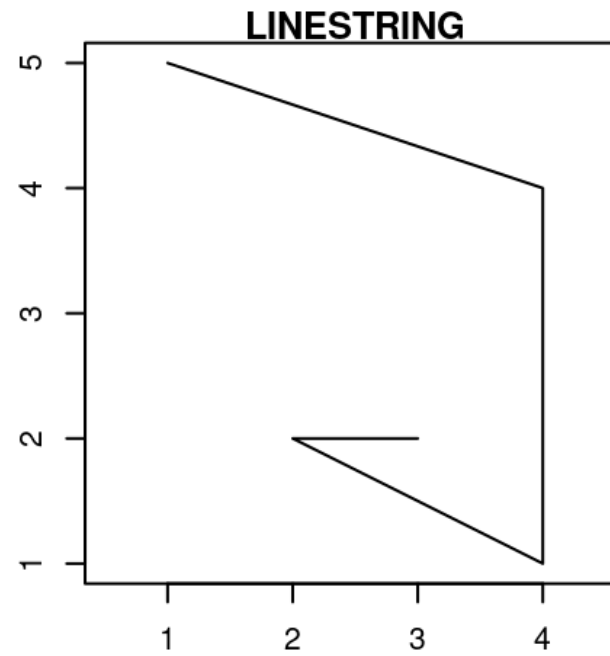
Tipo	Descrição
Ponto	geometria de ponto único
Multiponto	conjunto de pontos
Linha	linha única (dois ou mais pontos conectados por linhas retas)
Multilinha	conjunto de linhas
Polígono	Polígono sem “anéis internos”, denotando furos
Multipolígono	conjunto de polígonos
Coleção de geometria	conjunto das geometrias acima



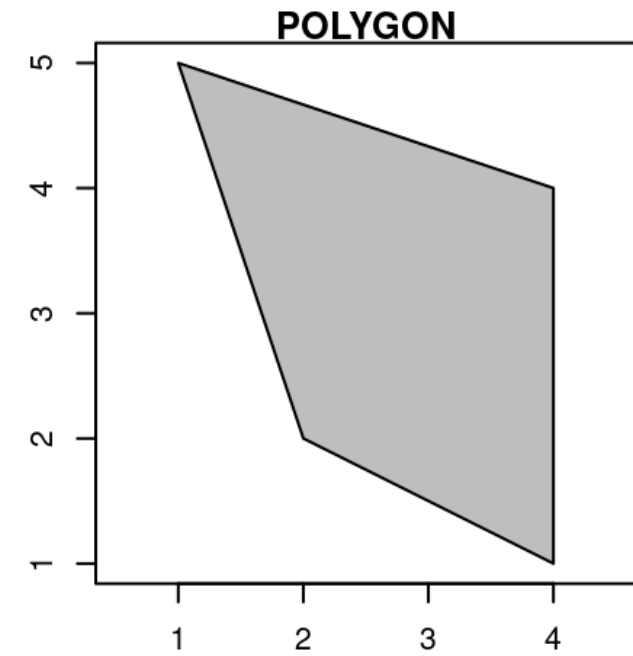
Dados Geoespaciais - Vetor



POINT (5 2)

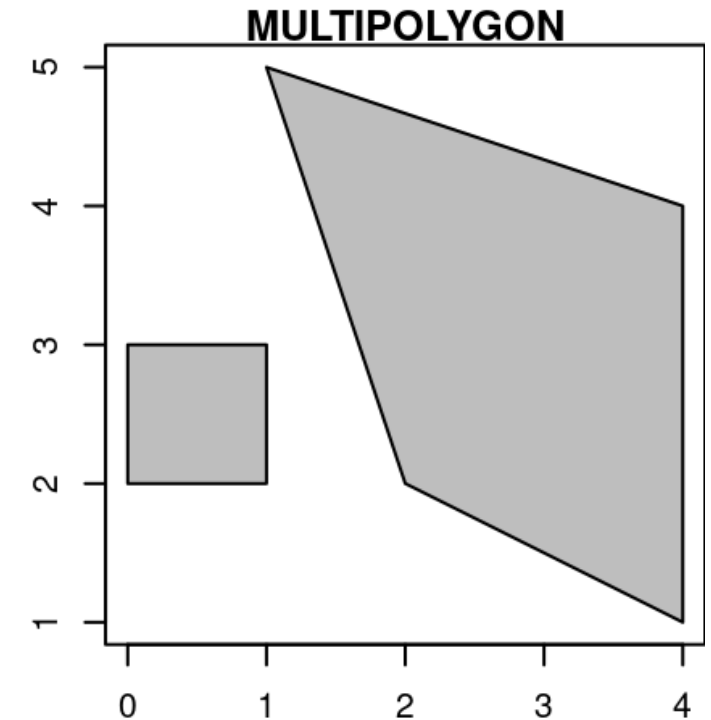
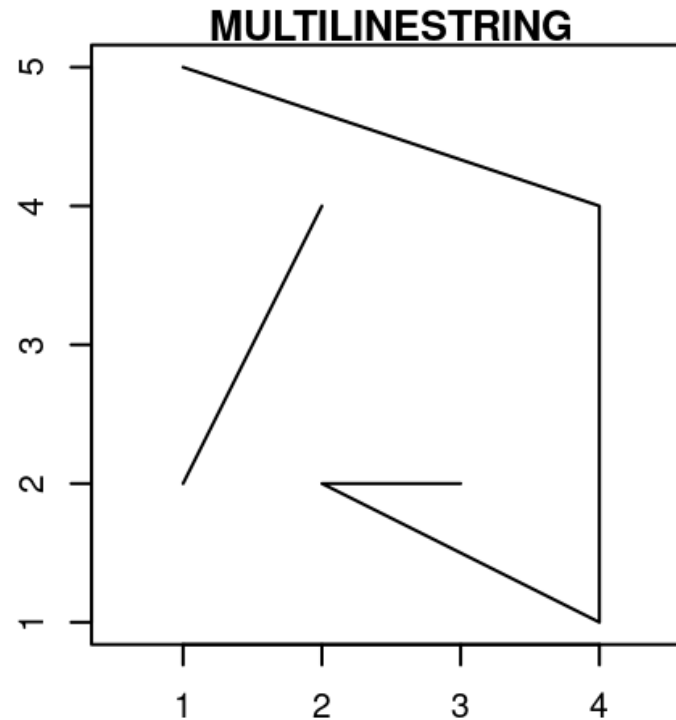
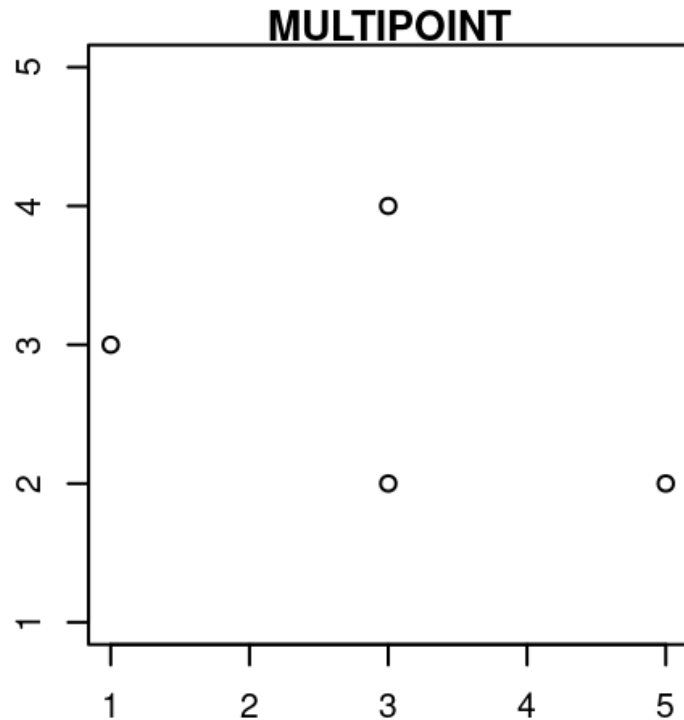


LINestring (1 5, 4 4, 4 1, 2 2, 3 2)



POLYGON ((1 5, 2 2, 4 1, 4 4, 1 5))

Dados Geoespaciais - Vetor

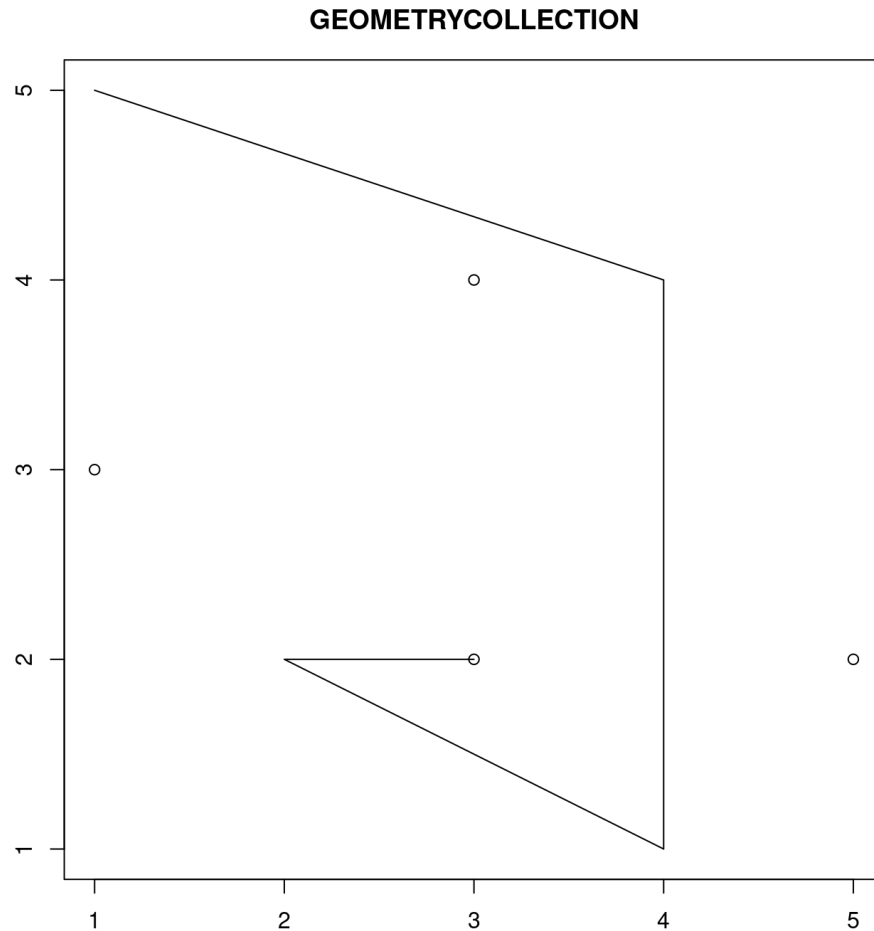


Multipoint: **MULTIPOINT (5 2, 1 3, 3 4, 3 2)**

Multilinestring: **MULTILINESTRING ((1 5, 4 4, 4 1, 2 2, 3 2), (1 2, 2 4))**

Multipolygon: **MULTIPOLYGON (((1 5, 2 2, 4 1, 4 4, 1 5), (0 2, 1 2, 1 3, 0 3, 0 2)))**

Dados Geoespaciais - Vetor



Geometry collection:

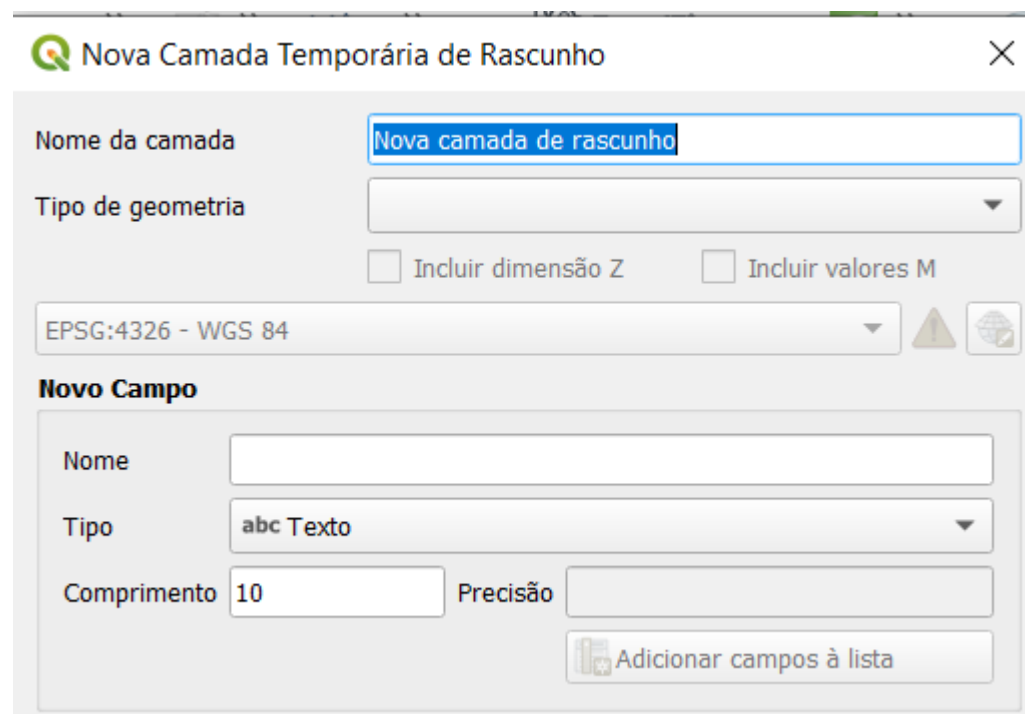
GEOMETRYCOLLECTION (MULTIPOINT (5 2, 1 3, 3 4, 3 2), LINESTRING (1 5, 4 4, 4 1, 2 2, 3 2))

Dados Geoespaciais - Vetor

Além das coordenadas X e Y, os vértices de geometrias de feições simples podem ter:

- uma coordenada **Z**, denotando altitude, e/ou
- um valor **M**, denotando alguma "medida"

O atributo Z e M deve ser uma propriedade do vértice. Eles não são comuns e o suporte de software de SIG para fazer algo útil com eles é (ainda) raro. O QGIS suporta.



Nova Camada Temporária de Rascunho

Nome da camada: Nova camada de rascunho

Tipo de geometria: [dropdown]

Incluir dimensão Z Incluir valores M

EPSG:4326 - WGS 84

Novo Campo

Nome: [input]

Tipo: abc Texto

Comprimento: 10 Precisão: [input]

Adicionar campos à lista

Dados Geoespaciais – Vetor

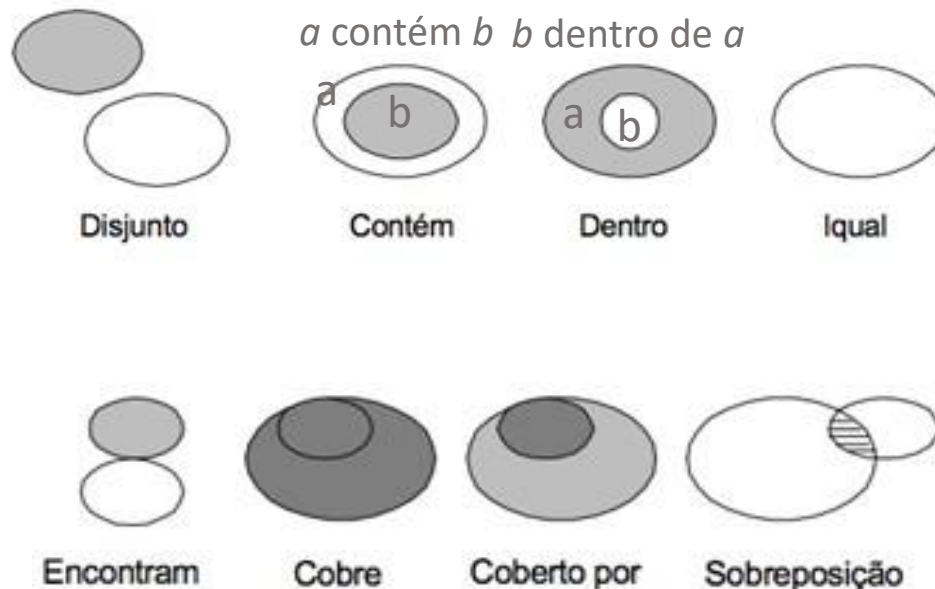
Além da geometria

Além da geometria, SIG permitem armazenar **RELACIONAMENTOS ESPACIAIS** entre feições.

Exemplos: adjacência, interseção, contenção e conectividade.

Essas relações são fundamentais para análises espaciais e consultas geográficas.

Exemplo: municípios que fazem fronteira, rios que cruzam estradas, bairros dentro de cidades.



SIG vs CAD

- **SIG**
 - Modelam o espaço geográfico com geometria + atributos + relacionamentos espaciais.
 - Permitem análises espaciais: proximidade, redes, interseção, vizinhança.
- **CAD (*Computer-Aided Design*)**
 - Focado principalmente na representação geométrica de objetos.
 - Normalmente não armazena explicitamente relações topológicas entre feições.
 - É voltado para desenho e projeto, não para análise espacial.

Dados Geoespaciais – Vetor

Além da geometria

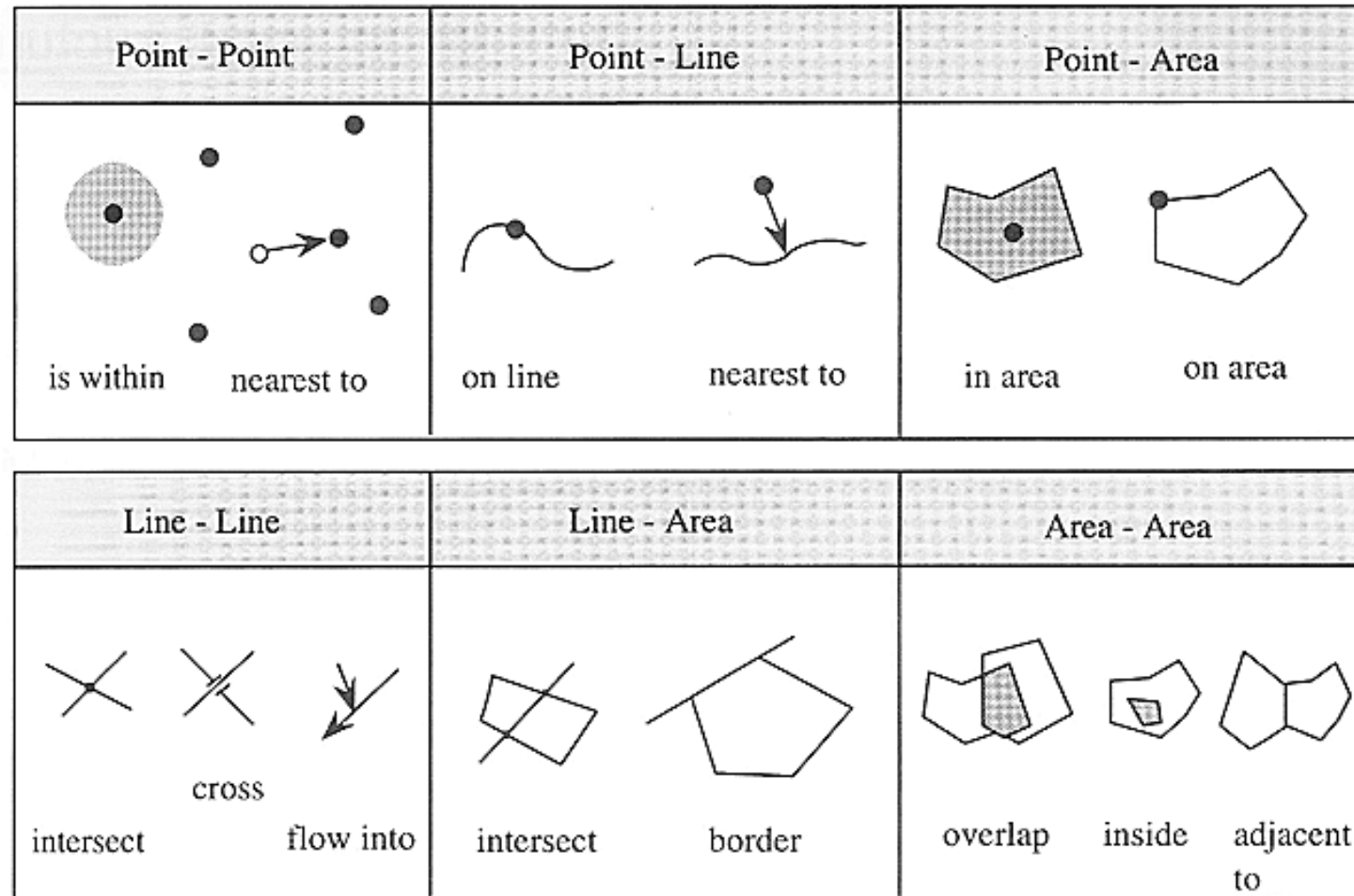


Figure 2.4 Topological Relationships Between Spatial Objects

Dados Geoespaciais - Vetor

Há mais 10 tipos de geometria que são menos usuais, mas que cada vez mais encontram implementação:

Tipo	Descrição
CIRCULARSTRING	CIRCULARSTRING é o tipo de curva básica, semelhante a um LINESTRING no mundo linear. Um único segmento requer três pontos, os pontos de partida e chegada (primeiro e terceiro) e qualquer outro ponto na curva. A exceção a isso é para um círculo fechado, onde os pontos de partida e chegada são os mesmos. Neste caso, o segundo ponto DEVE ser o centro do círculo, ou seja, o lado oposto do círculo. Para encadear arcos juntos, o último ponto do arco anterior se torna o primeiro ponto do próximo arco, assim como no LINESTRING. Isso significa que um CIRCULARSTRING válido deve ter um número ímpar de pontos maiores que 1.
COMPOUNDCURVE	Uma curva composta é uma única curva contínua que possui segmentos curvos (circulares) e segmentos lineares. Isso significa que, além de ter componentes bem formados, o ponto final de cada componente (exceto o último) deve ser coincidente com o ponto de partida do componente seguinte.
CURVEPOLYGON	Exemplo de curva composta em um polígono de curva: CURVEPOLYGON(COMPOUNDCURVE(CIRCULARSTRING(0 0,2 0, 2 1, 2 3, 4 3),(4 3, 4 5, 1 4, 0 0)), CIRCULARSTRING(1.7 1, 1.4 0.4, 1.6 0.4, 1.6 0.5, 1.7 1))
MULTICURVE	Um MULTICURVE é uma coleção de geometrias de 1 dimensão cujos elementos são curvas, podendo incluir strings lineares, strings circulares ou strings compostas.
MULTISURFACE	Uma MULTISURFACE é uma coleção de geometrias de 2 dimensões cujos elementos são superfícies, todas usando coordenadas do mesmo sistema de referência de coordenadas.

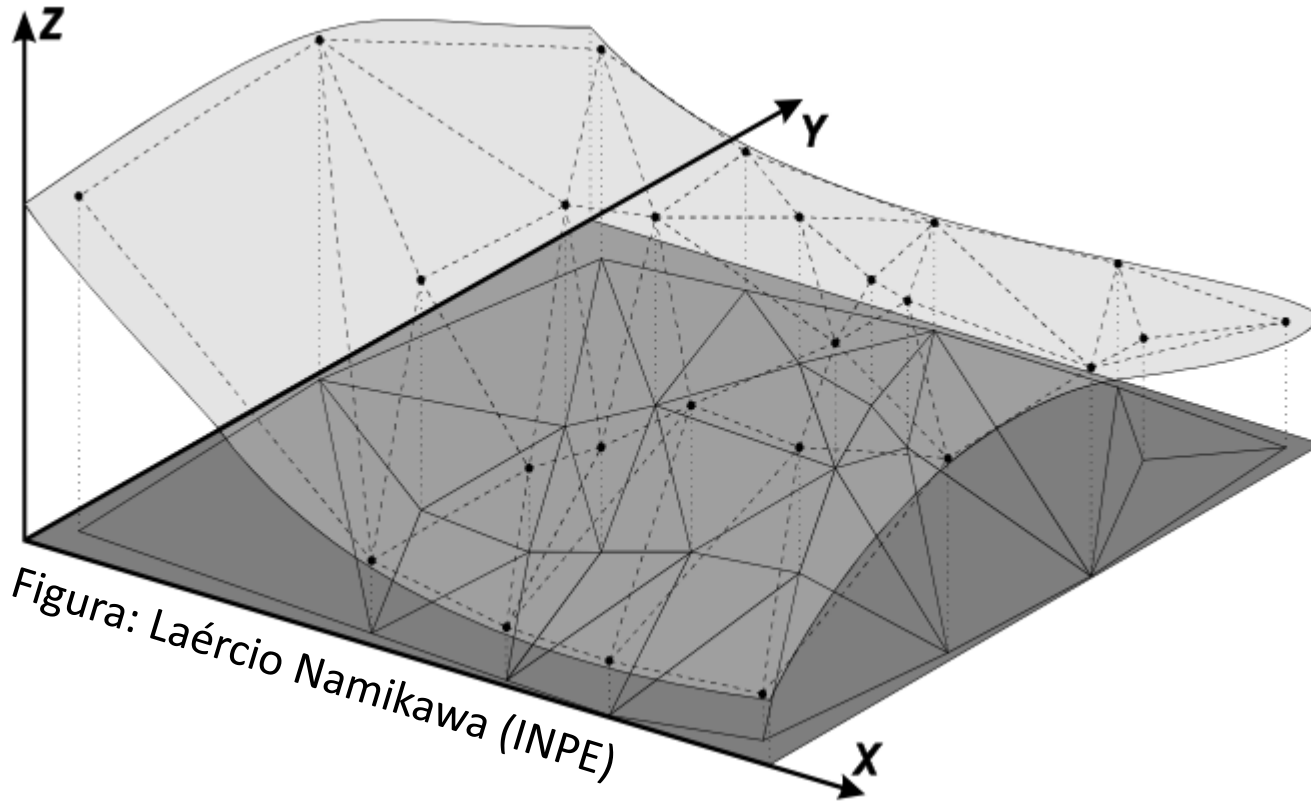
Dados Geoespaciais - Vetor

Há mais 10 tipos de geometria que são menos usuais, mas que cada vez mais encontram implementação:

Tipo	Descrição
CURVE	Uma Curva é um objeto geométrico de 1 dimensão geralmente armazenado como uma sequência de Pontos, com o subtipo de Curva especificando a forma da interpolação entre os Pontos.
SURFACE	Uma Superfície é um objeto geométrico de 2 dimensões.
POLYHEDRALSURFACE	Uma Superfície Poliédrica é uma coleção contígua de polígonos que compartilham segmentos de borda comuns.
TIN	Um TIN (rede irregular de triângulos) é uma Superfície Poliédrica que consiste apenas em patches de triângulos.
TRIANGLE	Um TRIANGLE é um polígono com 3 vértices distintos e não colineares e sem borda interior.

Dados Geoespaciais – Vetor

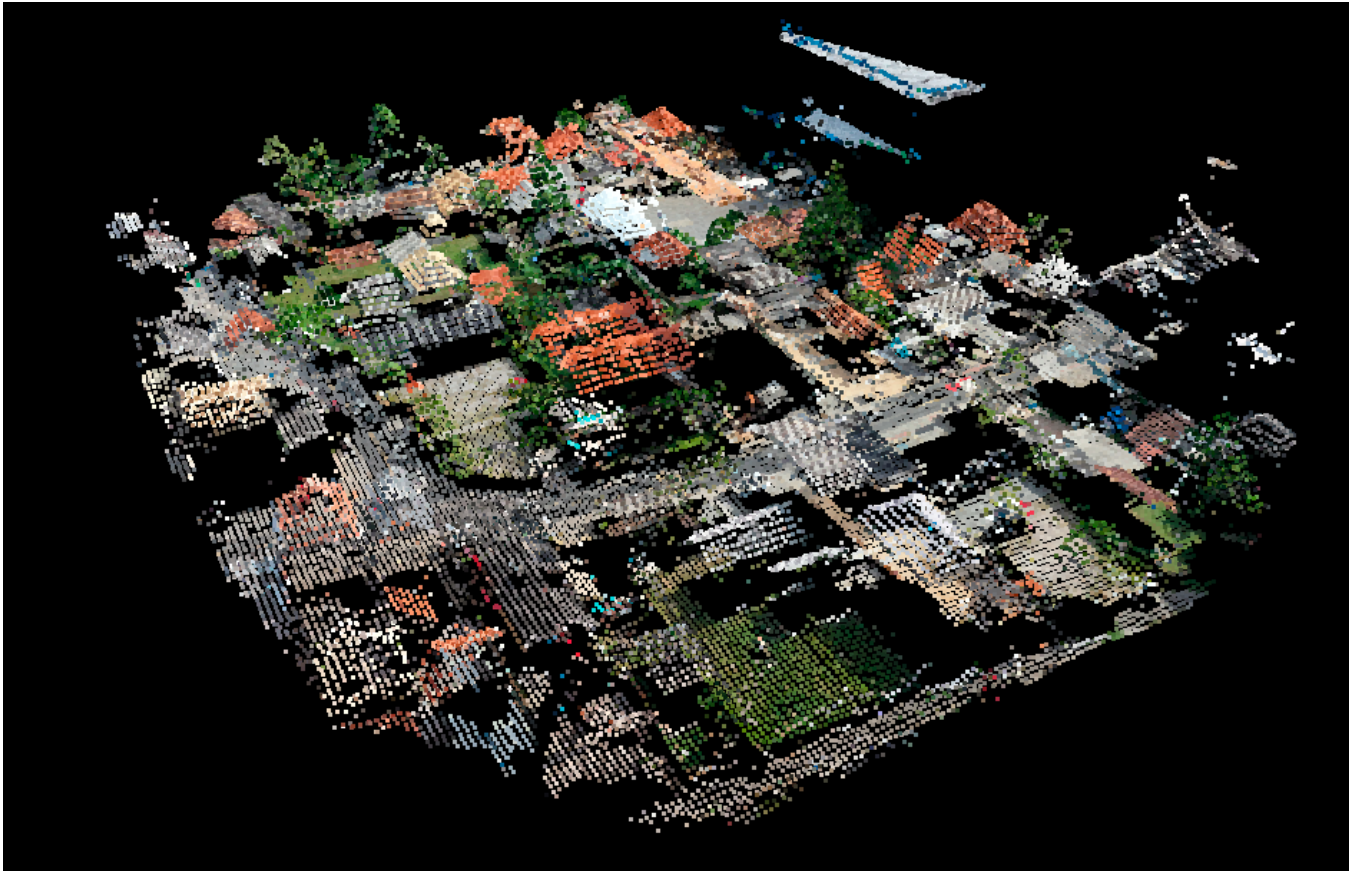
TIN - triangulated irregular network



- Estrutura de dados vetoriais que representa superfícies por meio de uma **rede irregular de triângulos conectados**, construída a partir de pontos amostrais com coordenadas (x, y, z) .
- Também mencionada como “Mesh”.
- O TIN é amplamente utilizado em aplicações de **engenharia, hidrologia e modelagem 3D**, estando presente em softwares como **BIM (ex.: Civil 3D, Revit)**, **HEC-RAS**, **ArcGIS**, **QGIS** e **ferramentas de modelagem de terreno e LiDAR**. Essas superfícies triangulares são geralmente armazenadas em formatos como **LandXML (.xml)** e **ArcGIS TIN (.tin)** em SIG, **2DM (.2dm)**, **SELAFIN (.slf)**, **UGRID (.nc)** e **MIKE21 mesh (.mesh)** em modelagem hidrodinâmica, além de formatos de malha 3D como **OBJ (.obj)**, **PLY (.ply)** e **STL (.stl)**.

Dados Geoespaciais – Vetor

Nuvem de pontos (*point cloud*)



- Nuvens de pontos (ex.: LiDAR em formato .las ou .laz) podem ser vistas como grandes coleções de pontos vetoriais com coordenadas 3D e atributos. Entretanto, devido ao volume massivo de dados e estruturas de armazenamento específicas, elas são frequentemente tratadas como uma **categoria própria de dados geoespaciais** em SIG modernos.
- Apesar da semelhança conceitual, nuvens de pontos normalmente são tratadas como **uma categoria própria de dados geoespaciais**, porque:
 - possuem **bilhões de pontos**
 - têm **estrutura otimizada para streaming**
 - usam **formatos específicos (.las, .laz, .copc)** requerem **algoritmos especializados**

Dados Geoespaciais - Vetor

Vantagens do dado vetorial

- Localização mais precisa de feições;
- Armazenamento de muitos atributos;
- Flexibilidade para cartografia;
- Armazenamento compacto de informações;
- Idealmente projetado para certos tipos de análises envolvendo áreas, comprimento, conexões.

Principais Formatos - vetoriais



Formato	Principais aplicações	Vantagens	Desvantagens	Lançamento
ESRI Shapefile	Desktop GIS	Padrão geoespacial na maioria das aplicações GIS	Limite de armazenamento (2 GB) Limite de caracteres por coluna e atributos Muitos arquivos (.shp, .shx, .dbf) para uma geometria	1999
KML	Google Earth, WebGIS, Desktop GIS	Formato comprimido (KMZ) capaz de armazenar diversas geometrias em um mesmo arquivo Suporte nativo a 3D	Impossibilidade de edição dos atributos via Desktop GIS, sendo necessário conversão para outro formato	2004
GeoJSON	WebGIS, DesktopGIS	Estruturado em JSON, otimizado à aplicações na web Ideal para geometrias simples	Ausência de indexação Limitações de escalabilidade Suporte limitado para diferentes projeções	2008
SpatialLite	Desktop GIS	Baseado em SQLite – baixo armazenamento em disco Limite de 281 TB Suporte em boa parte das aplicações GIS	Armazenamento de <i>rasters</i> depende de extensões	2008
GeoPackage	Desktop GIS	Limite de 140 TB Baseado em SQLite – baixo armazenamento em disco Capaz de armazenar dados vetoriais e raster Padrão de armazenamento do QGIS	Relativamente recente, tendo pouco suporte na maioria das aplicações GIS Monousuário	2014
File Geodatabase	Desktop GIS	Limite de 256 TB Atualização por multiusuário Capaz de armazenar dados vetoriais e raster Padrão de armazenamento do ArcGIS	Edição e criação somente possível em softwares ArcGIS	2010

Como os softwares conseguem ler tantos formatos diferentes?



O **GDAL/OGR** (Geospatial Data Abstraction Library) é uma biblioteca C++ open-source mantida pela **OSGeo** que fornece uma camada de abstração unificada para leitura e escrita de dados geoespaciais. É o motor por trás de ferramentas como QGIS, ArcGIS, PostGIS, Python (`rasterio`, `geopandas`), R (`terra`, `sf`) e muitos outros.

GDAL

Componente para dados **raster** — imagens, grades, modelos de elevação e formatos como GeoTIFF e NetCDF.

OGR

Componente para dados **vetoriais** — pontos, linhas e polígonos em formatos como Shapefile, GeoPackage e GeoJSON.

Uso na linha de comando

```
gdalinfo meu_raster.tif  
ogrinfo meu_vetor.gpkg
```

Esses dois comandos retornam metadados completos dos arquivos: sistema de referência, extensão espacial, tipo de dado e muito mais.

📄 **Site oficial:** <https://gdal.org>



Tem garantido a interoperabilidade no contexto da Geoinformática

Principais formatos vetoriais

Shapefile

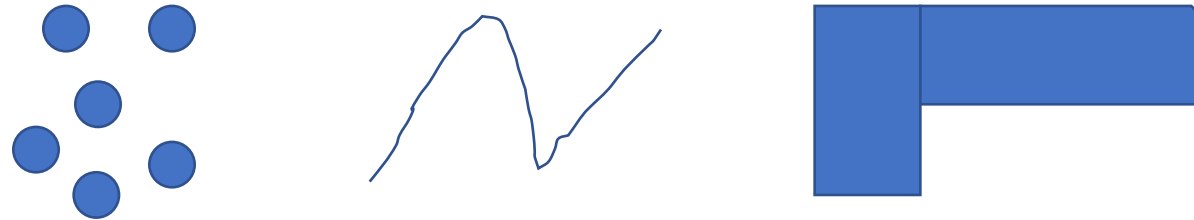
É um formato de armazenamento de dados de vetorial da Esri para armazenar a posição, forma e atributos de feições geográficas.

É **formado por um conjunto de arquivos** relacionados, contendo os arquivos básicos: .shp, .shx, .dbf no mapa que você criará com o visualizador de mapa.

Principais formatos vetoriais

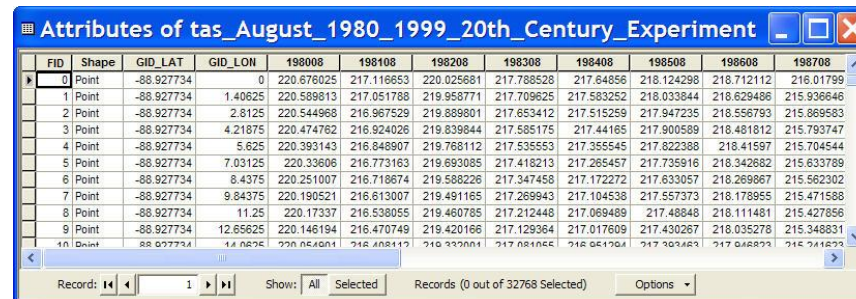
Shapefile

.shp (geometria)



.dbf (tabela de atributos)

Tamanho máximo do nome das colunas = 10 caracteres



FID	Shape	GID_LAT	GID_LON	198008	198108	198208	198308	198408	198508	198608	198708
0	Point	-88.927734	0	220.676025	217.116653	220.025681	217.788528	217.64856	218.124298	218.712112	216.01799
1	Point	-88.927734	1.40625	220.589813	217.051788	219.958771	217.709625	217.583252	218.033644	218.629486	215.936646
2	Point	-88.927734	2.8125	220.544968	216.967529	219.889801	217.653412	217.515259	217.947235	218.556793	215.869583
3	Point	-88.927734	4.21875	220.474762	216.924026	219.839844	217.585175	217.44185	217.900589	218.481812	215.793747
4	Point	-88.927734	5.625	220.393143	216.848907	219.768112	217.535553	217.355545	217.822388	218.41597	215.704544
5	Point	-88.927734	7.03125	220.33606	216.773163	219.693085	217.418213	217.265457	217.735916	218.342682	215.633789
6	Point	-88.927734	8.4375	220.251007	216.718674	219.588226	217.347458	217.172272	217.633057	218.269867	215.562302
7	Point	-88.927734	9.84375	220.190521	216.613007	219.491165	217.269943	217.104538	217.557373	218.178955	215.471588
8	Point	-88.927734	11.25	220.17337	216.538055	219.460785	217.212448	217.069489	217.48848	218.111481	215.427856
9	Point	-88.927734	12.65625	220.146194	216.470749	219.420166	217.129364	217.017609	217.430267	218.035278	215.348831
10	Point	-88.927734	14.0625	220.064801	216.408413	219.333001	217.081955	216.951704	217.393483	217.948893	215.261822

.shx (arquivo de vínculo entre a geometria e tabela de atributos)

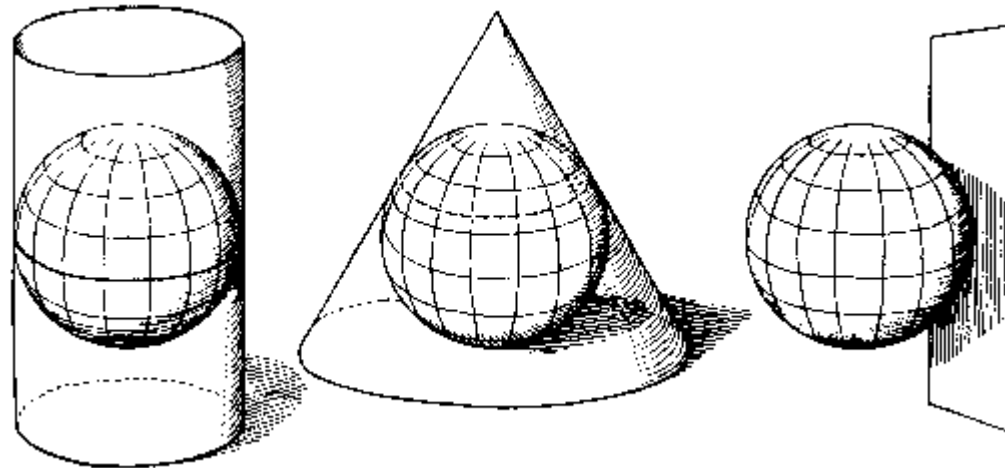
Principais formatos vetoriais

Shapefile

Pode haver outros arquivos opcionais, um importante é o:

.prj

Vai conter o sistema de referência do dado geográfico



Geopackage



- Código aberto, baseado em banco de dados SQLite
- Muito leve, mas altamente compatível entre ambientes (esp. em dispositivos móveis onde a conectividade e largura de banda é limitada)
- Os geopackage são geralmente ~1,1-1,3x mais leves em tamanho de arquivo em comparação com os shapefiles.
- Não há limite para o tamanho do arquivo e ele pode lidar com um grande número de feições.
- Em comparação com os shapefiles, os cabeçalhos das colunas podem ser nomes completos e corretos, fornecendo o contexto correto para cada coluna.
- Você verá uma execução mais rápida e saídas de algoritmos em geopackage comparados com shapefiles.
- Um único arquivo geopackage pode ter múltiplas camadas vetoriais com cada camada tendo um tipo de geometria diferente.
- É possível armazenar a simbologia do arquivo dentro do geopackage.

Principais formatos vetoriais



FlatGeobuf

Performance

Preliminary performance tests has been done using road data from OSM for Denmark in SHP format from download.geofabrik.de, containing 906602 LineString features with a set of attributes.

	Shapefile	GeoPackage	FlatGeobuf	GeoJSON	GML
Read full dataset	1	1.02	0.46	15	8.9
Read w/spatial filter	1	0.94	0.71	705	399
Write full dataset	1	0.77	0.39	3.9	3.2
Write w/spatial index	1	1.58	0.65	-	-
Size	1	0.72	0.77	1.2	2.1

<https://github.com/flatgeobuf/flatgeobuf>



Estrutura de Dados - Representação

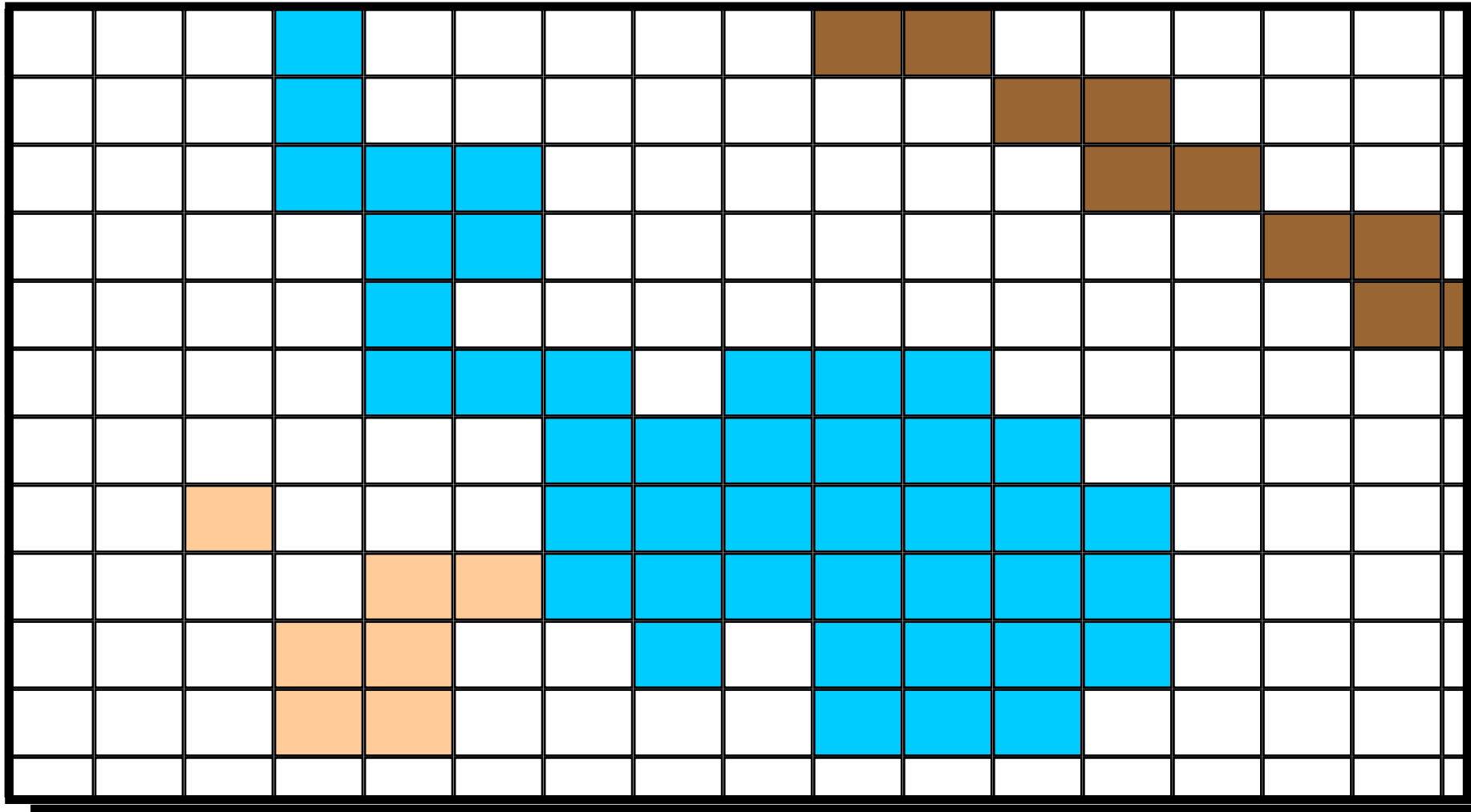
VETOR (Vector)

Formatos

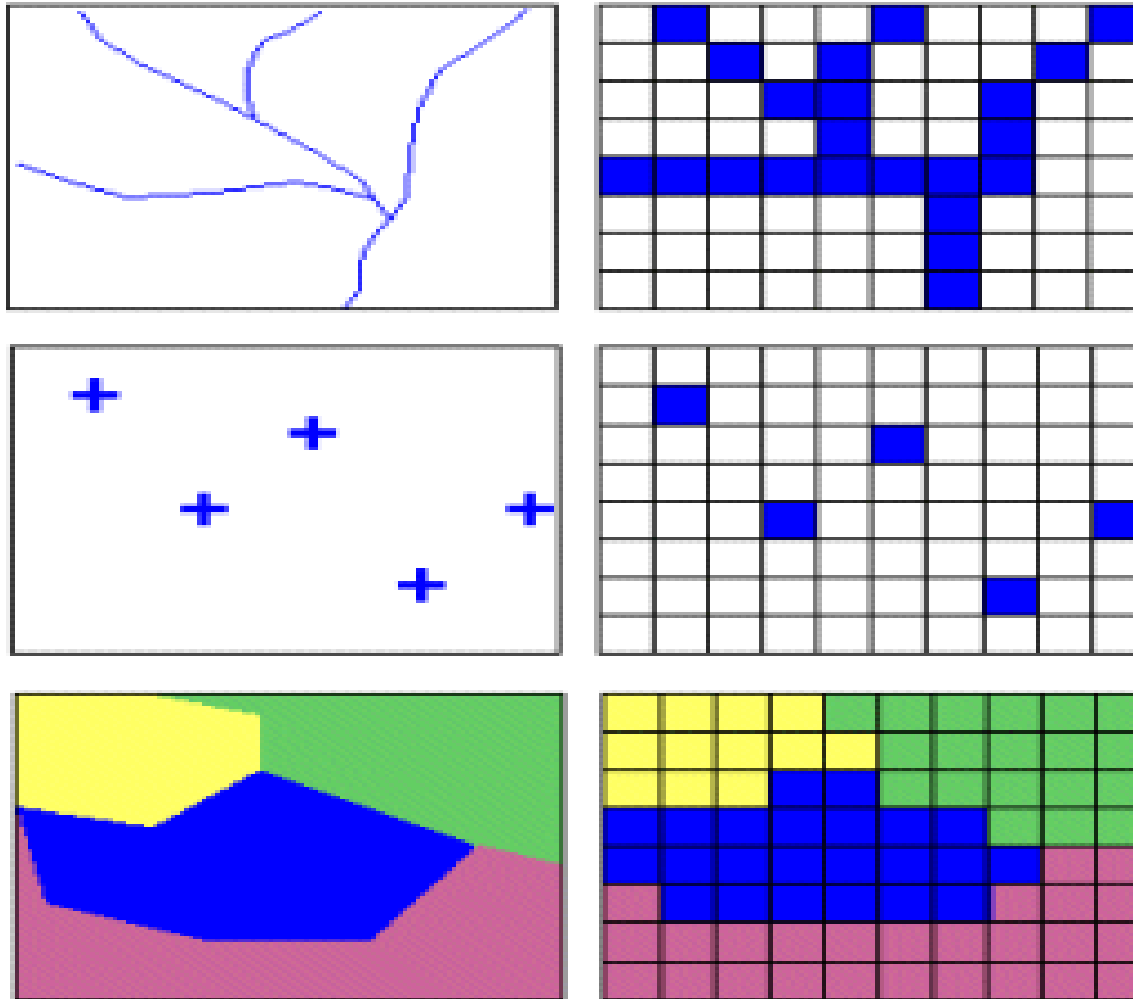
As performance is highly data dependent I've also made similar tests on a larger dataset with Danish cadastral data consisting of 2511772 Polygons with extensive attribute data.

	Shapefile	GeoPackage	FlatGeobuf
Read full dataset	1	0.23	0.12
Read w/spatial filter	1	0.31	0.26
Write full dataset	1	0.95	0.63
Write w/spatial index	1	1.07	0.70
Size	1	0.77	0.95

Conversão Vetorial → Matricial



Conversão Vetorial → Matricial



VECTOR X RASTER

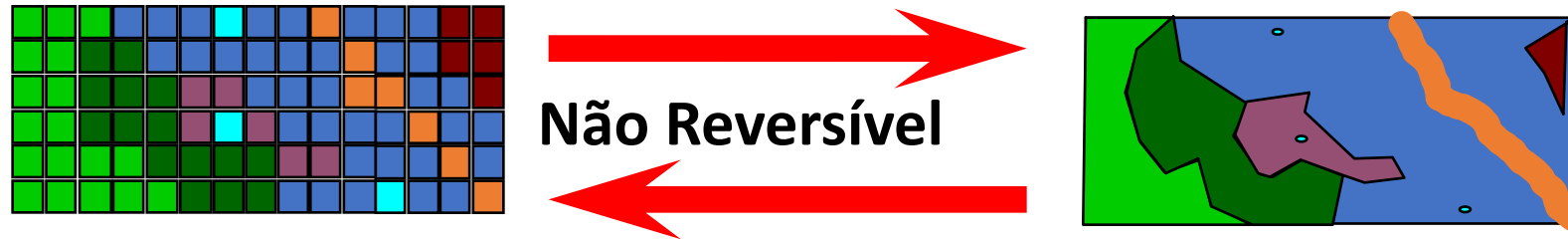
Comparando os dois modelos

Tipo de dado	Vantagem	Desvantagem
Raster	<ul style="list-style-type: none"> • Estrutura de dados simples (matriz regular de células), o que facilita implementação e processamento computacional • Operações algébricas entre camadas (map algebra) são diretas e eficientes devido à estrutura matricial regular • Adequado para representar fenômenos contínuos no espaço, como elevação, temperatura e precipitação 	<ul style="list-style-type: none"> • Pode demandar grande espaço de armazenamento, especialmente em alta resolução ou grandes extensões territoriais • A reprojeção (transformação de SRC) introduz reamostragem, o que pode gerar perda de informação ou artefatos nos dados • Representação de relações topológicas é limitada.
Vetor	<ul style="list-style-type: none"> • Armazenamento eficiente para feições esparsas ou com geometria complexa, ocupando menos espaço do que um raster equivalente em muitos casos • A reprojeção é geometricamente exata, pois consiste na transformação das coordenadas dos vértices sem necessidade de reamostragem • Permite a representação explícita de relações topológicas 	<ul style="list-style-type: none"> • Estrutura de dados mais complexa (geometrias, atributos, índices espaciais), exigindo maior esforço de implementação e manutenção • Operações sobre grandes volumes de feições podem ser computacionalmente custosas, especialmente sem índices espaciais adequados • Menos adequado para representar fenômenos continuamente variáveis no espaço

Componentes do SIG

VETOR X RASTER

Conversão Vetor x Raster



Os softwares SIG realizam tal conversão;

Alguns problemas esperados:

- Vetor para raster – Perde a precisão do vetor;
- Raster para vetor – Perde a variação dos dados;
- Raster para vetor (polígonos) – Distorções nas bordas.

Componentes do SIG

VETOR X RASTER

Conversão Vetor x Raster



Os softwares SIG realizam tal conversão;

Alguns problemas esperados:

- Vetor para raster – Perde a precisão do vetor;
- Raster para vetor – Perde a variação dos dados;
- Raster para vetor (polígonos) – Distorções nas bordas.

Processo de Representação Computacional



Editado - <http://naahdolls.blogspot.com>

Modelagem de Dados Geográficos

A modelagem de dados geográficos é o processo de **definir formalmente como fenômenos do mundo real serão representados digitalmente** em um SIG. Envolve a identificação de entidades, seus atributos e as relações espaciais entre elas.

1 Identificação de Entidades

Definir quais elementos do mundo real serão representados: rios, estradas, municípios, parcelas de terra, edificações.

2 Definição de Atributos

Especificar as características descritivas de cada entidade: nome, área, população, tipo de pavimento, data de criação.

3 Relações Espaciais

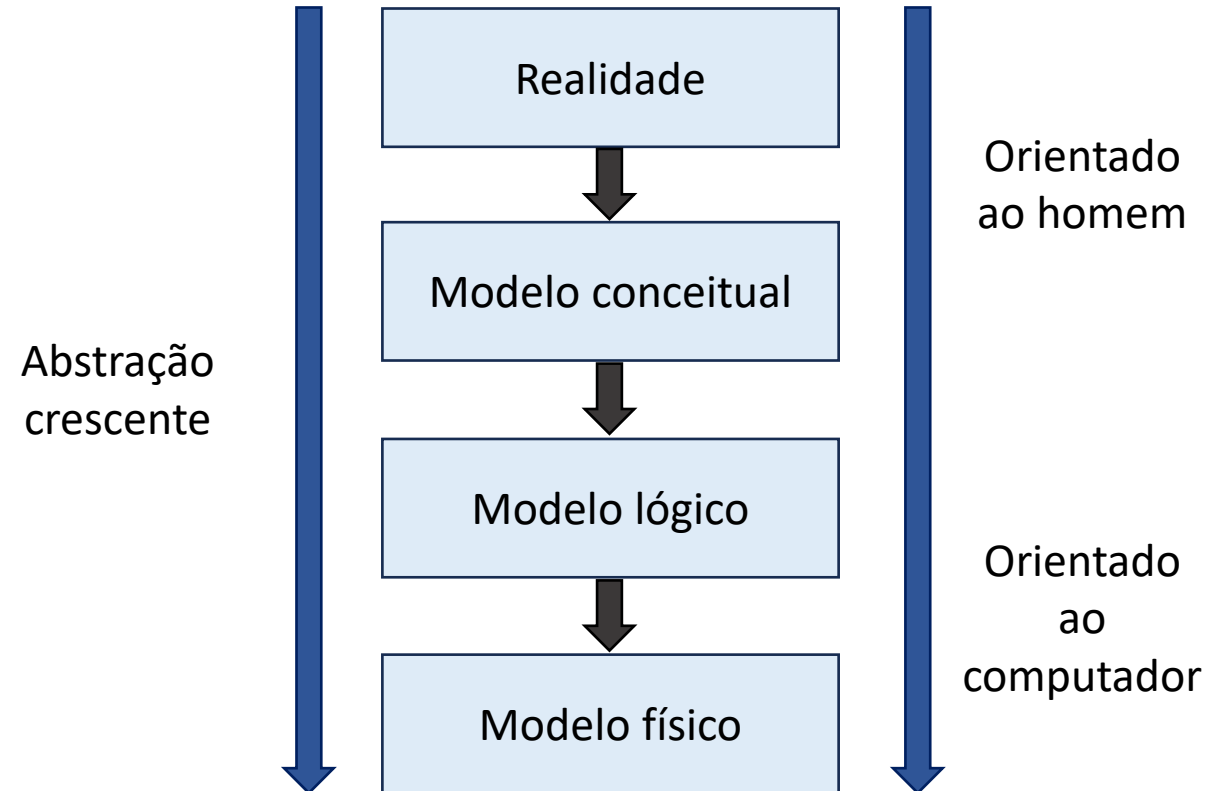
Estabelecer como as entidades se relacionam no espaço: adjacência, contenção, interseção, proximidade e conectividade.



Níveis de Modelagem

- Modelo Conceitual;
- Modelo Lógico;
- Modelo Físico.

Níveis de Modelagem

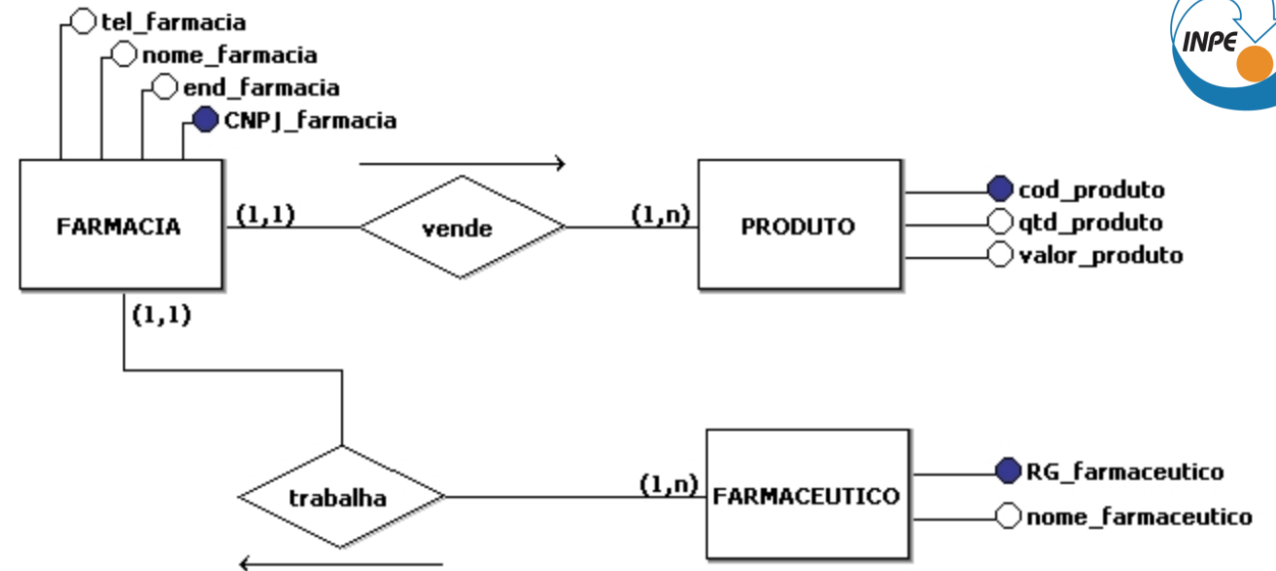


Níveis de abstração de um modelo de dados em SIG. Fonte: Longley et al. (2015)



Modelagem Conceitual

- A modelagem conceitual é uma técnica para descrever abstratamente os conceitos, entidades, relacionamentos e processos de um sistema;



https://sae.unb.br/cae/conteudo/unbfga/sbd/new_bancomodelagem.html

- Ela é independente da tecnologia utilizada para implementação;
- A modelagem conceitual é uma etapa importante do processo de desenvolvimento de software e banco de dados, bem como na seleção das variáveis envolvidas;

Modelagem Conceitual

- A modelagem conceitual é composta por diagramas e documentações textuais que descrevem os conceitos e relações envolvidos no sistema;
- Ela ajuda a evitar problemas de ambiguidade e inconsistência nos requisitos;
- A modelagem conceitual permite que os desenvolvedores e *stakeholders* discutam e validem as necessidades e funcionalidades do sistema de forma mais clara e objetiva.
- Ela serve como base para as etapas seguintes do desenvolvimento, como a modelagem lógica e física.

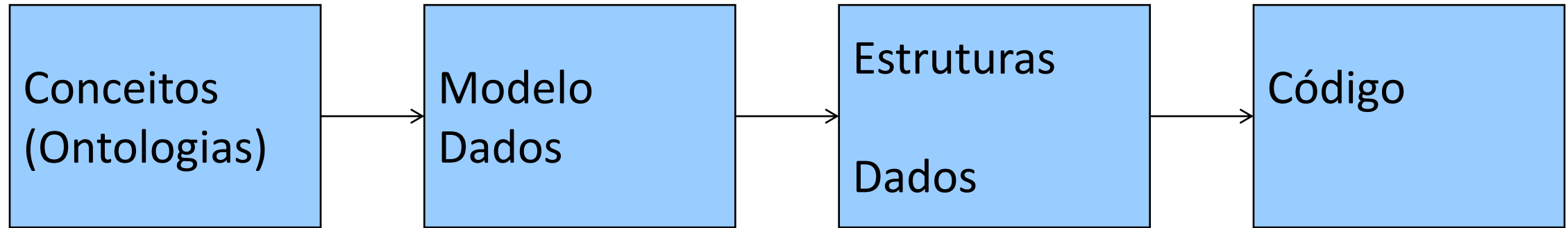
Modelagem Conceitual

- A modelagem conceitual é uma representação das ideias e conceitos subjacentes a um sistema, e não de sua implementação física.
- Os diagramas são uma das principais ferramentas para auxiliar na modelagem conceitual de sistemas e auxiliam na identificação de requisitos, variáveis, funcionalidades e problemas de design do sistema.
- Eles fornecem uma visão geral do sistema, tornando mais fácil para os desenvolvedores e *stakeholders* compreenderem e discutirem as funcionalidades do sistema.

Modelagem de dados geográficos

- especificar entidades espaciais e seus relacionamentos
- conceitos gerais: geo-campos e geo-objetos
- problemas reais: instanciar classes abstratas
- notação para descrição dos dados

Processo de Representação Computacional



Níveis de abstração

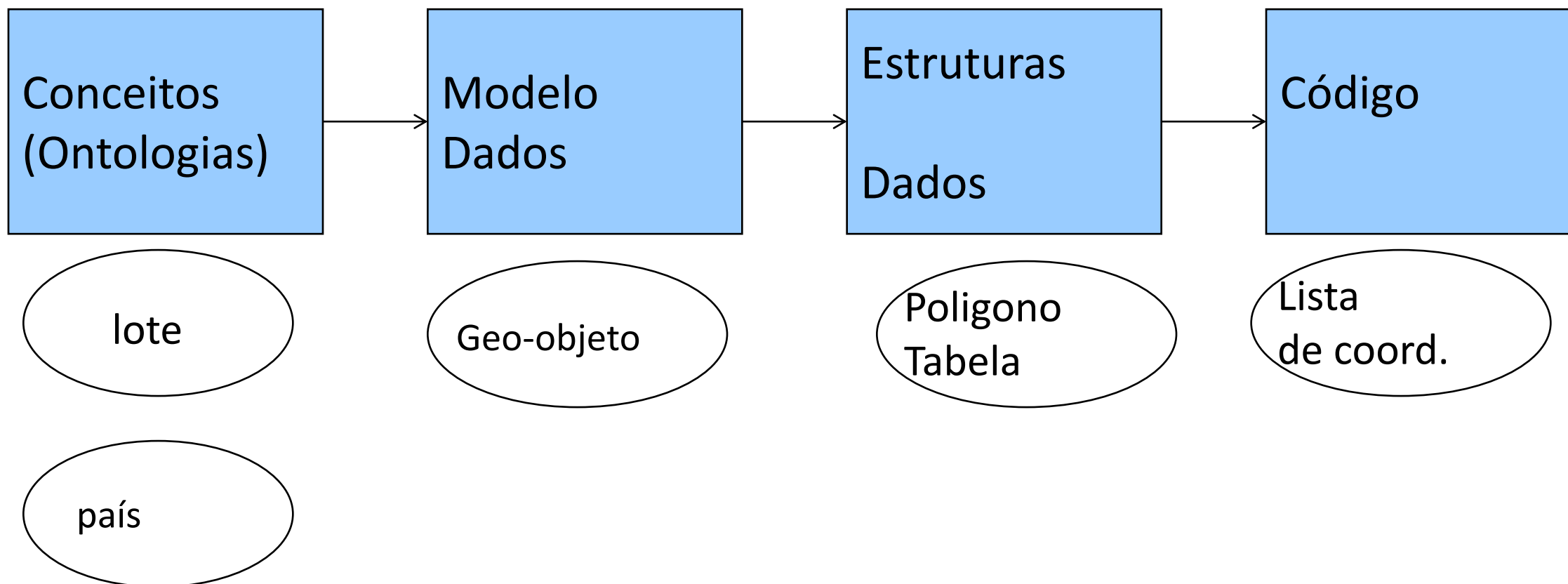
Ontologias (conceitos do mundo real): lote, classes de solos

Modelo de Dados: entidades (objetos) x processos (campos)

Estruturas de dados: matrizes, vetores, model. Dinâmicos

Implementação: R-trees, quad-trees

Processo de Representação Computacional





Campos e Objetos: Abstrações da Realidade

Objetos

Aproximações da realidade

ex. Linha da costa brasileira

Convenções socialmente aceitas

ex. Fronteiras entre municípios

Campos

Discretização de fenômenos contínuos

ex. Altimetria, temperatura

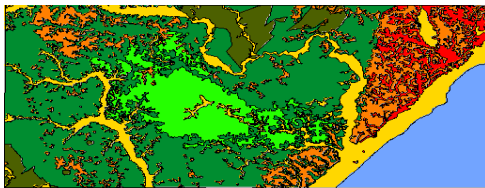
Campos e Objetos: Abstrações da Realidade

Campos (superfícies, distribuições)

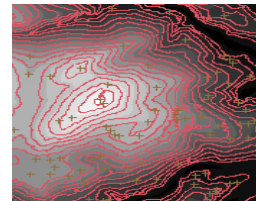
Definidos por suas propriedades

Representam um fenômeno contínuo

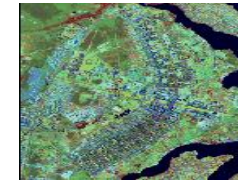
Para cada ponto da região, temos um valor distinto



Temático



Numérico



Imagem

Objetos (entidades, feições)

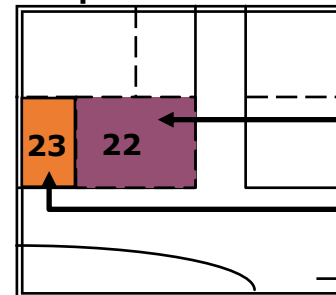
Definidos por seus limites (fronteiras)

Homogeneidade interna

Descrito por seus atributos

Só descreve o que está “dentro” dele

Mapa Cadastral de Lotes



Objeto :Lotes

id	dono	endereço	cadastro IPTU
22	Guimarães	Caetés 768	250186
23	Bevilacqua	São João 456	110427
...

Modelagem Conceitual

Há diversas técnicas de modelagem conceitual, a maioria voltada para bancos de dados convencionais, em que as características espaciais dos objetos não são incluídas.

Ex.:

- Visio
- draw.io
- BrModelo
- etc

O modelo de dados OMT-G

- *Object Modeling Technique* – OMT (Técnica de Modelagem de Objetos)
 - OMT-G - *Object Modeling Technique for Geographic Applications*
- É baseado em três conceitos principais:
 - *classes,*
 - *relacionamentos e*
 - *restrições de integridade espaciais.*

Diagrama de Classe



- O *diagrama de classe* é usado para descrever a estrutura e o conteúdo de um banco de dados geográfico.

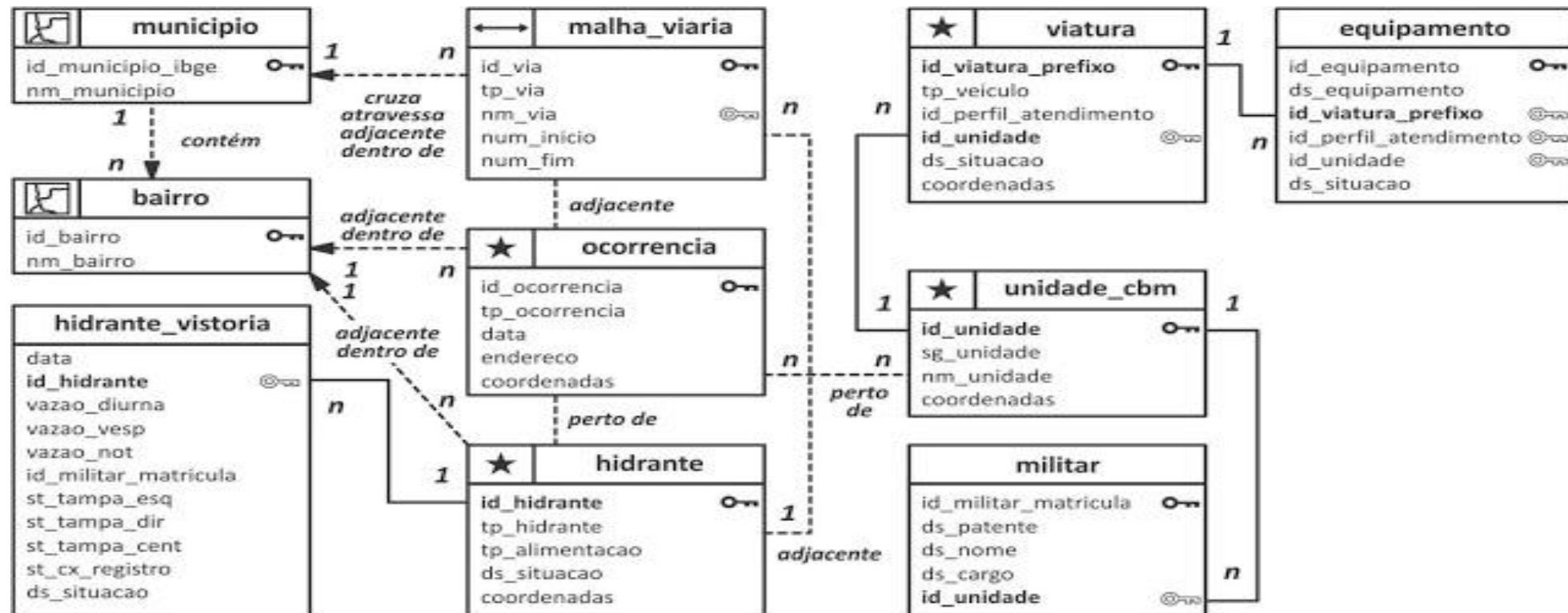


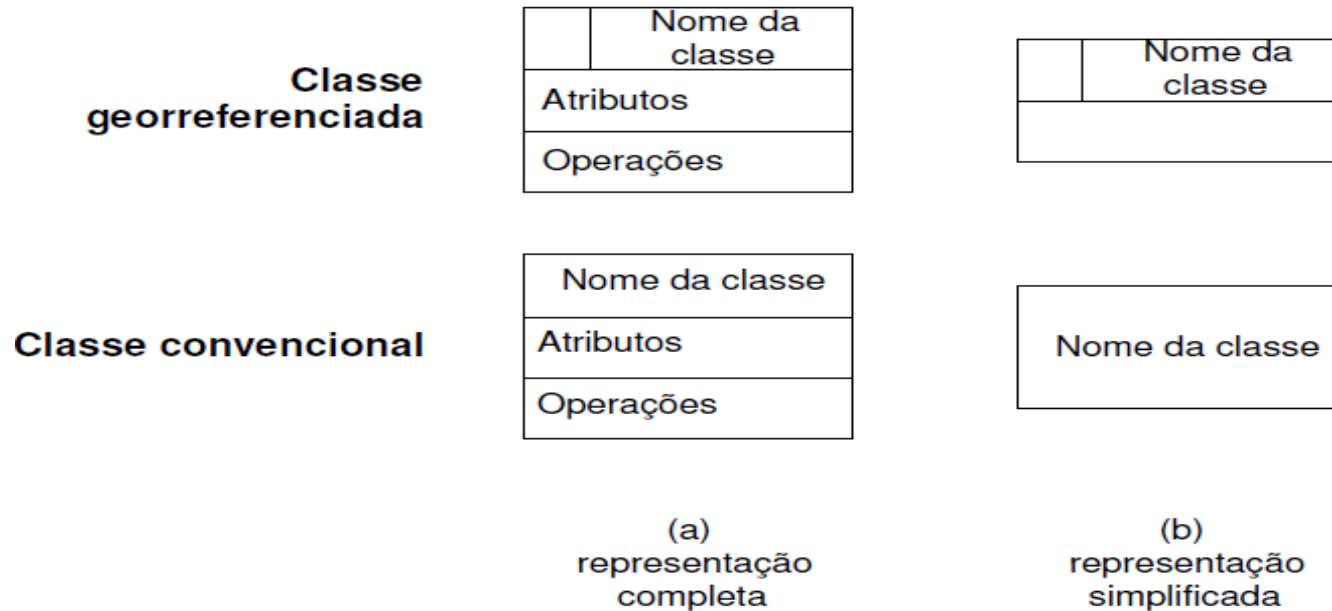
Diagrama OMT-G de Sistema de Controle de Ocorrência

Fonte: <http://geopara.blogspot.com.br/>

Diagrama de Classe



- Atua nos níveis de representação conceitual e apresentação



Notação gráfica para as classes do modelo OMT-G

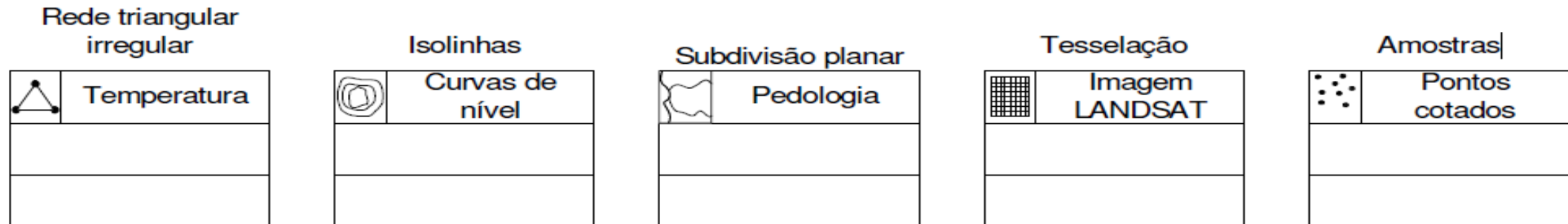
A **Unified Modeling Language (UML)** é uma linguagem de modelagem não proprietária de terceira geração.

A UML permite que desenvolvedores visualizem os produtos de seus trabalhos em diagramas padronizados

Diagrama de Classe

- O modelo OMT-G define cinco classes descendentes de geo-campo:

isolinhas, subdivisão planar, tesselação, amostragem e malha triangular.



Classes de Geo-campos no OMT-G

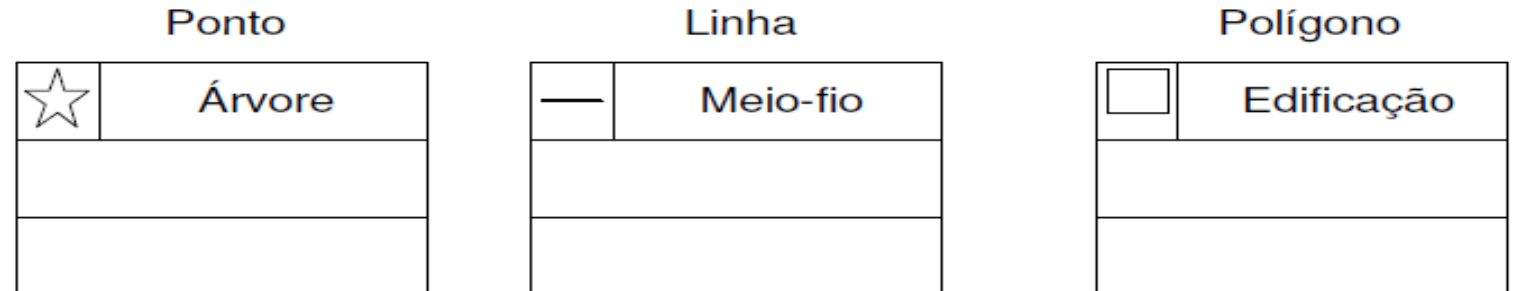
O modelo de geo-campos enxerga o espaço geográfico como uma superfície contínua, sobre a qual variam os fenômenos a serem observados.

Diagrama de Classe



O geo-objeto representa o espaço geográfico como uma coleção de entidades distintas e identificáveis, onde cada entidade é definida por uma fronteira fechada.

Geo-objetos com geometria



Geo-objetos com geometria e topologia

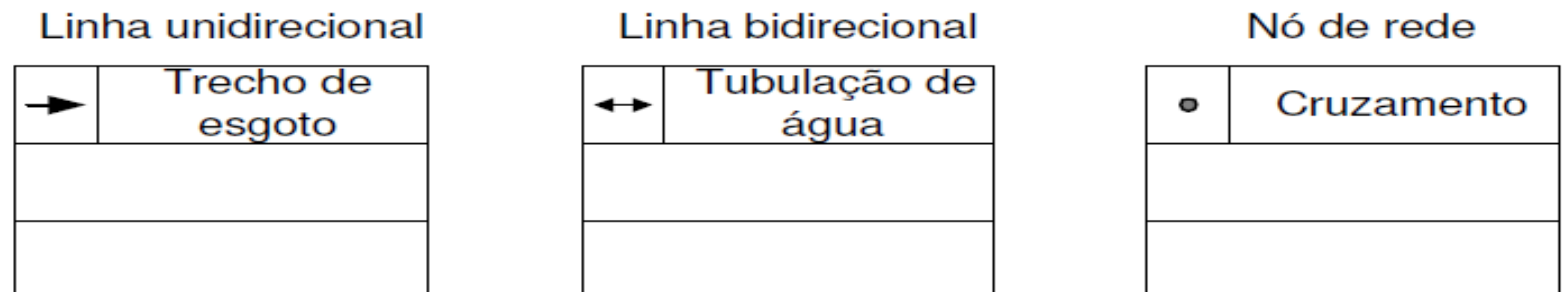
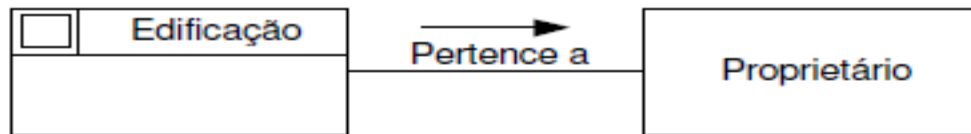


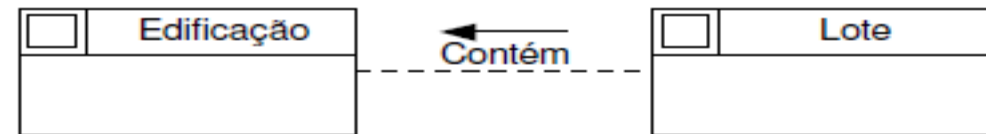
Figura 5 – Geo-objetos.

Relacionamentos

O modelo OMT-G representa três tipos de relacionamentos entre suas classes: *associações simples*, *relacionamentos topológicos em rede* e *relacionamentos espaciais*.



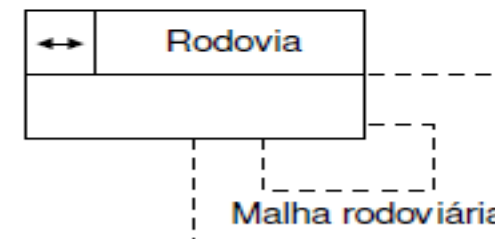
(a) Associação simples



(b) Relacionamento espacial



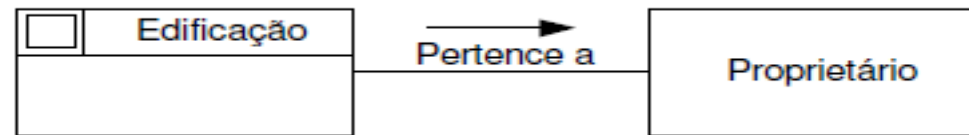
(c) Relacionamento de rede arco-nó



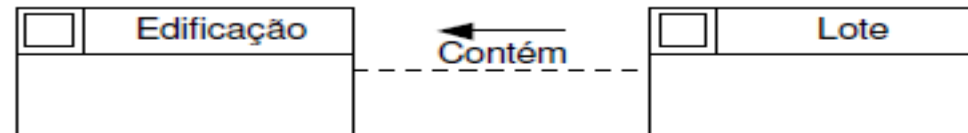
(d) Relacionamento de rede arco-arco

Relacionamentos

- *Associações simples* representam relacionamentos estruturais entre objetos de classes diferentes, convencionais ou georreferenciadas.



- *Relacionamentos espaciais* representam relações topológicas, métricas, de ordem e fuzzy.



(b) Relacionamento espacial

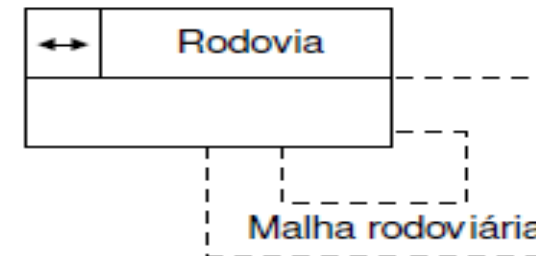
associações simples são indicadas por linhas contínuas, enquanto relacionamentos espaciais são indicados por linhas pontilhadas

Relacionamentos

- *Relacionamentos topológicos em rede* são relacionamentos entre objetos que estão conectados uns com os outros.



(c) Relacionamento de rede arco-nó

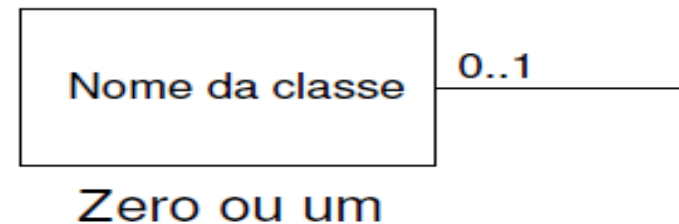
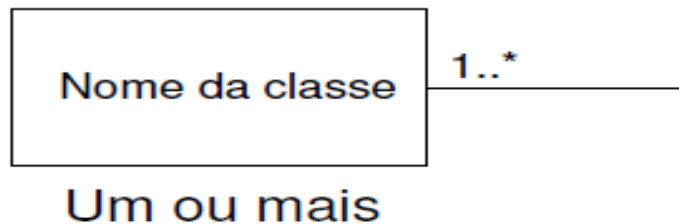
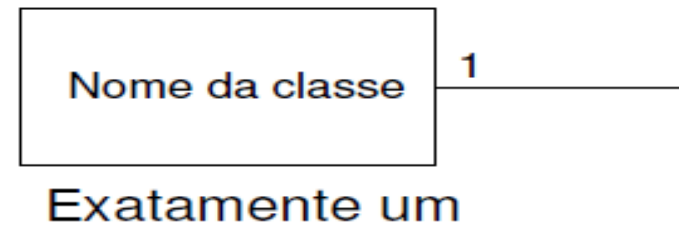
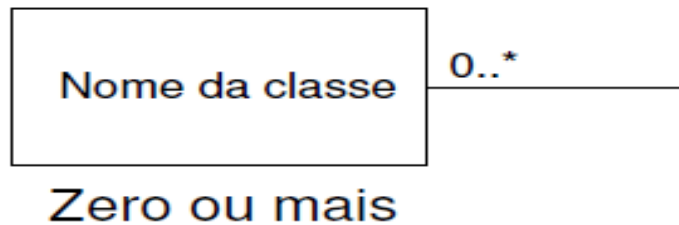


(d) Relacionamento de rede arco-arco

Relacionamentos de rede são indicados por duas linhas pontilhadas paralelas, entre as quais o nome do relacionamento é anotado

Relacionamentos

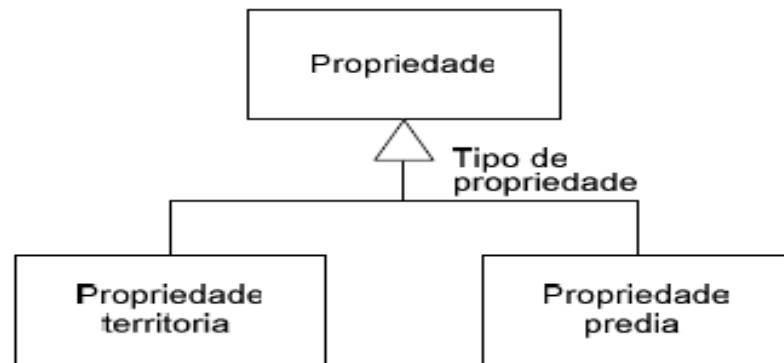
- *A cardinalidade* representa o número de instâncias de uma classe que podem estar associadas a instâncias de outra classe.



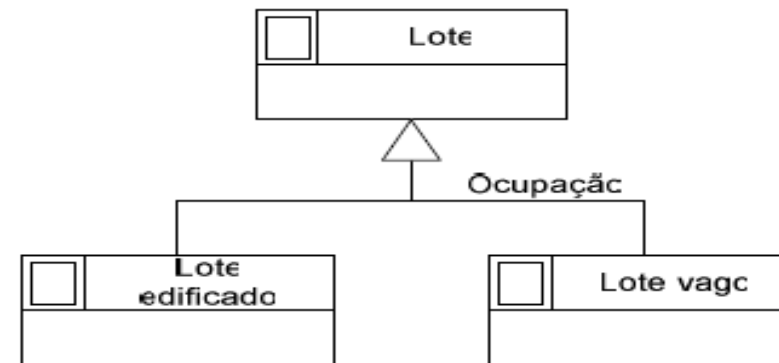
Cardinalidade.

Generalização e especialização

- ▶ *Generalização* é o processo de definição de classes mais genéricas (*superclasses*) a partir de classes com características semelhantes (*subclasses*).



(a) Notação UML



(b) Generalização espacial

A especialização é o processo inverso, no qual classes mais específicas são detalhadas a partir de classes genéricas, adicionando novas propriedades na forma de atributos.

Tipos Geométricos

<i>Representação OMT-G</i>		<i>Representação OpenGIS (Simple Features Specification)</i>
Geo-objeto	Ponto	Point
Geo-objeto	Linha	LineString
Geo-objeto	Polígono	Polygon
Geo-objeto	Nó de rede	Point
Geo-objeto	Arco unidirecionado	LineString
Geo-objeto	Arco bidirecionado	LineString
Geo-campo	Amostras	Point
Geo-campo	Isolinhas	LineString e/ou Polygon
Geo-campo	Subdivisão planar	Polygon
Geo-campo	Triangulação	Point (vértices) e Polygon (triângulos)
Geo-campo	Tesselação	GeoRaster, campo binário longo

Material de apoio

Usando o ChatGPT do OMT-G ao PostGIS:

https://docs.google.com/document/d/1_GpG6mpUGUaW7nXADWvZeGDFwWVi2sX/edit?usp=sharing&oid=118166016499607858652&rtpof=true&sd=true

- <https://homepages.dcc.ufmg.br/~clodoveu/bdg2005/cap3.pdf>
- <http://www.dpi.inpe.br/gilberto/livro/introd/cap4-modelos.pdf>
- http://www.dpi.inpe.br/DPI/livros/pdfs/tutorialbdgeo_geobrasil2006.pdf