INT. J. GEOGRAPHICAL INFORMATION SYSTEMS, f9.92.. V<)L. 6~ NO. 1, 31--45

**Ciência da informação geográfica**

MICHAEL F. GOODCHILD

Centro Nacional de Informação e Análise Geográfica, Universidade da Califórnia, Santa Barbara, CA 93106, EUA!

Abstrato. Artigos de pesquisa. em conferências como o EGIS e os Simpósios Internacionais sobre · Espaciais. O tratamento de dados trata de um conjunto de perguntas que vão bem. além do limitado. capacidades técnicas da tecnologia atual em sistemas de informação geográfica. Este artigo analisa os tópicos. que pode ser incluído em uma ciência da informação geográfica. A pesquisa sobre essas questões fundamentais é uma perspectiva melhor para a sobrevivência e aceitação a longo prazo na academia do que o desenvolvimento de capacidades técnicas. Este artigo analisa o estado atual da pesquisa · em uma série de áreas-chave e especula por que o progresso tem sido tão desigual. A seção final do artigo mostra o futuro e as novas áreas de potencial significativo nessa área de pesquisa.

1. Introdução

A comunidade do sistema de informação geográfica (SIG) tem um longo caminho na última década. Principais programas de pesquisa e treinamento foram estabelecidos em vários países, novas aplicações foram encontradas, novos produtos surgiram da indústria. Uma indústria que continua a se expandir a um ritmo espetacular, melhora dramática nas capacidades das plataformas e novas melhorias significativas. conjuntos de dados estão disponíveis. É tentador dizer que a pesquisa GIS, e as reuniões nas quais essa pesquisa é apresentada, são simplesmente uma parte desse entusiasmo e entusiasmo muito maiores, mas deveria haver mais do que isso. Qual é, afinal, o propósito de toda essa atividade? Expressões como 'manipulação de dados espaciais' podem descrever o que fazemos, mas não dão a entender por que fazemos. Esse foi um dos temas por trás do discurso de Tomlinson no Primeiro Simpósio Internacional sobre Tratamento de Dados Espaciais em Zurique, em 1984 (Tomlinson, 1984). O título da conferência sugere que os dados espaciais são de alguma forma difíceis de lidar, mas será sempre assim? Ele sugere um nível de distanciamento dos dados. Eles mesmos, como se o US Geological Survey enviasse fitas rotuladas com o aviso genérico "manipular com dificuldade". É uma reminiscência do nome da antiga Comissão de Sensoriamento e Processamento de Dados Geográficos da União Geográfica Internacional. Uma rápida revisão dos títulos dos artigos nessas reuniões ou nas reuniões subsequentes deve ser suficiente para garantir a qualquer um que seus autores estejam preocupados com muito mais do que o mero manuseio e processamento de dados - da perspectiva dos EUA. De que a comunidade é mais do que os Estados Unidos Serviço de encomendas de SIG. · Às vezes, os sistemas de informações geográficas são acusados ​​de serem orientados para a tecnologia. Isso parece ser mais verdadeiro em alguns períodos dos 25 anos de história do SIG do que em outros. Por exemplo, é difícil sugerem que Tomlinson e os desenvolvedores do Sistema de Informações Geográficas do Canadá (CG IS) (Tomlinson et al. (1976)) foram guiados pelo assustadoramente primitivo. capacidades de hardware de 1965. Por outro lado, a perspectiva de um SIG de varredura de menu suspenso em cores e totalmente orientado por menus no computador pessoal baseado em 386 em sua mesa vendeu claramente muitos sistemas nos últimos anos ~ O desenvolvimento tecnológico surge rajadas distintas, e o mesmo ocorre com a tecnologia por trás do GIS. Pode ser a motivação por trás do desejo de lidar com dados espaciais, mas falha em explicar muitos dos diversos esforços de pesquisa relatados nas reuniões e na literatura.

Também houve fases em que os aplicativos direcionaram o SIG. O próprio CGIS era uma pesquisa de aplicação da tecnologia, e a unidade foi suficientemente forte para levar ao protótipo do primeiro scanner de mapas e a vários outros desenvolvimentos tecnológicos (Tomlinson et al., 1976). McHarg havia elaborado os princípios da técnica de sobreposição de mapas (McHarg 1969) muito antes de Berry e outros os automatizarem no MA: P e seus derivados (Berry, 1987); o software de roteamento de ônibus escolar existe há muito mais tempo do que a implementação do problema em um SIG padrão. Porém, novamente, grande parte do objeto da pesquisa está muito além de qualquer aplicação razoavelmente previsível. Também houve fases em que os aplicativos direcionaram o SIG. O próprio CGIS era A ampla distribuição das imagens Landsat e SPOT e a disponibilidade de modelos digitais de elevação e arquivos de ruas em muitos países certamente o levaram a isso. aplicativos muito além daqueles usados para justificar a compilação dos dados. TIGER, para por exemplo, parece estar gerando sua própria indústria de atualizadores, reembaladores e desenvolvedores de aplicativos, embora exista apenas em princípio apenas para atender às necessidades do Censo dos EUA de 1990. No entanto, embora o banco do motorista de SIG esteja sem dúvida lotado, eu gostaria de lidar neste artigo em grande parte com o quarto motorista localizado aparentemente irrelevante no banco de trás, a palavra 'S'. Parece-me que há uma necessidade premente de reconhecer e desenvolver o papel da ciência no SIG. Isso significa dois sentidos. O primeiro tem a ver com a extensão em que o SIG como um campo contém um conjunto legítimo de questões científicas, a extensão em que elas podem ser expressas e a extensão em que são genéricas, em vez de específicas para campos de aplicação ou contextos específicos . Até que ponto a comunidade de pesquisa de GIS é motivada pela curiosidade intelectual sobre a natureza da tecnologia GIS · e as questões que ela suscita? E se o GIS pode ser · motivado pela ciência, então quais são seus subcampos, quais são suas perguntas e qual é sua agenda? O segundo sentido · tem a ver com o papel do SIG como uma caixa de ferramentas na ciência em geral - com o SIG para a ciência e não com a ciência do SIG. O que precisamos fazer para garantir esse SIG e espacial. tecnologia de manipulação de dados, desempenham seu papel legítimo no apoio às ciências para as quais a geografia é uma chave significativa ou uma fonte significativa de insights, explicações e compreensão? . . Para fazer isso, precisamos primeiro estabelecer que os dados espaciais, ou melhor, geográficos, são únicos e que seus problemas não podem, portanto, ser incluídos em algum campo maior. Também devemos estabelecer que existem problemas genéricos para todos os dados geográficos ou, pelo menos, estabelecer que é possível distinguir aqueles que são daqueles que não são. Por exemplo, a precisão dos atributos em um mapa detalhado de estatísticas de crimes parece ser muito pouco informada pelo conhecimento da precisão de atributos para dados geográficos em geral, mas requer um nível de entendimento dos problemas específicos das estatísticas de crimes. No entanto, a precisão das estimativas da população para um polígono definido arbitrariamente pode muito bem ser conhecida ou pelo menos informada pelas propriedades gerais da área modificada. problema de unidade (Openshaw, 1977).

2 .. O que é único nos dados espaciais?

Em muitos sistemas de gerenciamento de instalações, o papel do SIG é fornecer uma chave alternativa aos dados, um método de acesso baseado na localização geográfica. Em essência, um banco de dados espacial possui chaves duplas, permitindo que os registros sejam acessados por atributos ou por locais. No entanto, chaves duplas não são incomuns. A chave espacial é distinta, pois permite definir operações que não estão incluídas nas linguagens de consulta padrão. Por exemplo, é possível recuperar todos os registros de pontos dentro de um usuário arbitrário. polígono definido, uma operação que não é definida em linguagens de consulta padrão, como SQL Em essência, a chave espacial é multidimensional, mas novamente as chaves multidimensionais são conhecidas de outras áreas, e análogos de ponto na recuperação de polígonos podem ser definidos para dimensões não espaciais. · O que distingue dados espaciais é · o fato de que a chave espacial é baseada em duas dimensões contínuas. É possível visitar qualquer local (x, y) no mundo geográfico real, definido em princípio com precisão ilimitada e retornar um valor para uma variável, por exemplo, elevação topográfica z. O terreno é assim caracterizado por um. infinito número de tuplas (x, y, z). Nos aplicativos de rede, z é definido apenas para locais na rede, mas o número de tuplas ainda é infinito se a variação for contínua ao longo dessa estrutura unidimensional de links e nós. As séries temporais também possuem chaves contínuas, mas raramente são concebidas, medidas ou representadas como contínuas, e parece haver pouca semelhança de interesse nos problemas de manipulação de dados temporais. Por outro lado, · há ampla evidência de semelhança no tratamento de dados espaciais disciplinas.

Muitos de nossos modelos de dados, principalmente redes poligonais e redes irregulares trianguladas (TINs), refletem uma visão subjacente do espaço como contínua e a necessidade de acomodar o usuário que deseja determinar z de alguma forma arbitrária e precisa (x, Ji). ·. Uma implicação disso é que existe uma multiplicidade de possíveis modelos conceituais de dados para dados espaciais, e que a escolha entre eles para um determinado fenômeno é uma das questões mais fundamentais do tratamento de dados espaciais. Outra característica distintiva dos dados espaciais é o que Anselin (1989) chama de dependência espacial, a propensão para locais próximos influenciarem-se mutuamente e possuírem atributos semelhantes. Sem dependência espacial, não haveria perspectiva razoável de criar visões aproximadas de variação espacial contínua dentro de uma máquina finita e discreta. Não é incomum que tuplas que tenham valores semelhantes de uma chave tenham valores semelhantes de outros atributos, mas a estrutura da dependência espacial é incomum, dependendo das duas dimensões da chave (x, y), com a similaridade determinada por uma métrica. Por fim, os dados geográficos são distribuídos sobre a superfície curva da terra, fato que muitas vezes é esquecido nas áreas limitadas de estudo de muitos projetos de SIG. Durante séculos, preocupamo-nos em como retratar a superfície da Terra em uma folha plana de papel: e desenvolvemos uma extensa tecnologia de projeções de mapas. No entanto, como resultado, temos poucos métodos para analisar dados sobre a esfera ou esferóide e sabemos pouco sobre como modelar processos em sua superfície curva. Além disso, wc tendem a tratar os displays GIS como se fossem folhas virtuais. e insistiu em visualizar dados geográficos como se fossem projetados em uma superfície plana, em vez de explorar o potencial da tela eletrônica para criar vistas do próprio globo. Precisamos desenvolver as técnicas apropriadas para trabalhar com o mundo e usar modelagem sólida em vez de gráficos bidimensionais convencionais, se quisermos entender os processos geográficos em escala global e contribuir efetivamente para a ciência global. Devemos resgatar a projeção ortográfica de sua obscuridade atual.

3. O conteúdo da ciência da informação geográfica

Tendo estabelecido que a informação geográfica possui propriedades e problemas únicos, agora podemos revisar o conjunto de perguntas genéricas que podem constituir uma ciência da informação geográfica. Isso pode ser feito de uma maneira amplamente linear, da coleta de dados à análise, embora alguns temas tendam a atravessar esse arranjo simples. No entanto, parece apropriado iniciar esta revisão com um aviso. O que apresento neste artigo é, sob muitos aspectos, o meu ponto de vista, e eu esperava que fosse. desafiado. Acho que meus próprios preconceitos ficarão claros - a seguir. Por causa do diversidade e dinamismo do campo, é difícil, se não impossível, para qualquer um. tentar uma visão geral. O que se segue é, portanto, quase inevitavelmente incompleto e desigual.

A pesquisa é frequentemente identificada como pura ou aplicada = conduzida por princípios básicos e inocentes curiosidade humana ou pelas necessidades práticas diárias da sociedade humana. Muitos GIS são uma resposta às necessidades humanas de gerenciamento e análise de informações e, nesse sentido. .. pode-se esperar que a pesquisa GIS seja mais aplicada do que pura. No entanto, uma visão de pesquisa pura é que são pesquisas que ainda não encontraram aplicação; pura pesquisa isa. investimento de longo prazo, assim como aplicado. a pesquisa é um investimento de curto prazo. Do ponto de vista acadêmico, a pesquisa pura é frequentemente associada a maior prestígio, mas a pesquisa aplicada com maior financiamento. Eu tentei cobrir toda a gama de puro a aplicado, sentindo que ambos são importantes para o SIG. Ao mesmo tempo, 'pesquisa básica' é o principal objetivo de. o Centro Nacional de Informações Geográficas e Análise dos EUA, e o centro tem muita sorte em ser financiado para fazer pesquisas · cujas aplicações podem ocorrer anos ou mesmo décadas no futuro. Durante a fase de design do CGIS na década de 1960, ficou claro que a única maneira prática de inserir o grande número de mapas necessários seria por alguma forma de dispositivo de digitalização (Tomlinson et al. 1976). Naquela época, não existia scanner para documentos do tamanho de mapas e era necessário inventar um. Um protótipo de scanner de tambor foi construído pela IBM Canadá e testado com sucesso, o que pelos padrões modernos seria considerado uma despesa imensa. Outras partes da equipe de design do CGIS estavam ocupadas inventando outras soluções igualmente fundamentais e agora familiares para problemas técnicos de SIG, como a ordem Morton. Nas quase três décadas de desenvolvimento de SIG que estão agora atrás de nós, pesquisas semelhantes sobre 'como fazê-lo' produziram um grande número de algoritmos, estruturas de dados, esquemas de indexação espacial e outras soluções tecnológicas. Alguns desses. são exclusivos do GIS, mas muitos foram reinventados em várias disciplinas relacionadas. A ordem de Morton, por exemplo, ocorre na literatura de · vários campos de manipulação de dados espaciais com nomes diferentes (Samet 1989) e descrições de algoritmos para encontrar polígonos de Thiessen estão espalhados por uma ampla variedade de periódicos. Ao mesmo tempo, há um crescente senso na pesquisa de GIS de que nossa ênfase mudou, à medida que mais e mais problemas técnicos subjacentes do SIG são resolvidos. A atenção passou de algoritmos primitivos e estruturas de dados para os problemas muito mais complexos do design de bancos de dados e para os problemas que envolvem o uso da tecnologia GIS em aplicativos reais. As seções a seguir identificam alguns desses problemas principais.

3.1 Coleta e medição de dados

Se a realidade espacial é contínua e sujeita a estruturas complexas de dependência espacial, então como deve ser compilada e mensurada? De maneira mais geral, como as pessoas percebem o mundo real da variação geográfica, estruturam-no e aprendem sobre ele? Embora muitas dessas perguntas façam parte das agendas de pesquisa do sensoriamento remoto, fotogrametria, geodésia e psicologia cognitiva, as linhas de demarcação estão longe de serem distintas. O SIG ou o sensoriamento remoto devem preocupar-se com os problemas de transferência de informações de uma tecnologia para outra e, mais importante, com o bom senso? É um SIG ou informações geográficas sensoriais remotas se forem usadas para melhorar a precisão da classificação ou se uma imagem é usado para atualizar uma camada GIS? Por fim, pouco importa a qual dos muitos buracos de pombos atribuímos cada tópico. Sem dúvida, existem questões científicas substanciais aqui, que exigem uma compreensão profunda da natureza da variação espacial, e o sensoriamento remoto de uma pessoa pode muito bem ser outra ciência da informação geográfica. · O processo de discretização, com generalização, abstração e aproximação implícitas, ocorre quando os dados são coletados, interpretados ou compilados, e são feitas escolhas nesse estágio que afetam os usos finais dos dados. Quando esses usuários mudam, como têm feito com o. uso generalizado de SIG, pode ser necessário ou benéfico repensar o processo de dados. coleção. Por exemplo, com gerenciamento digital e entrega de dados do censo, ainda é apropriado realizar um censo a. base decenal? A abordagem tradicional do mapeamento de campos geológicos é a mais apropriada se o objetivo final for uma representação tridimensional digital <? A subsuperfície? Como o mapeamento topográfico mudará agora que é econômico pesquisar novos recursos usando o Sistema de Posicionamento Global ·? Dados geográficos coleção é frequentemente o domínio de especialistas em áreas bem estabelecidas. disciplinas, portanto pode levar .. · muitos anos antes que esse tipo de pergunta seja investigado ou respondido. Para datar o A introdução do SIG parece ter tido muito pouco efeito no processo de coleta de dados.

3.2 Captura de dados

Grandes avanços foram feitos na tecnologia de captura de dados geográficos digitais na última década, e os sistemas atualmente no mercado são capazes de um alto nível de inteligência na interpretação de documentos de mapas digitalizados. O problema continua sendo a baixa qualidade dos documentos e as ambiguidades causadas por eles. aspectos do design do mapa. Como resultado, a digitalização manual continua sendo uma abordagem amplamente usada, apesar de seu alto custo, tédio e falha em mostrar melhorias significativas na ineficiência. Duas tendências podem mudar essa situação substancialmente nos próximos anos. Uma é a crescente evitação do documento do mapa como uma etapa no processo de compilação e entrada de dados. A topografia e a fotogrametria estão se afastando da compilação usando mapas em papel, e os campos mais interpretativos, como uso do solo, vegetação ou mapeamento do solo, provavelmente seguirão o exemplo. A estação total digital provavelmente será seguida pela mesa plana digital e talvez até o caderno de fieldgeologia digital. A outra é a possibilidade há muito reconhecida de que mudanças comparativamente menores no design de um mapa podem tornar muito mais fácil escanear e interpretar (Shiryaev 1987).

3.3 Estatísticas espaciais

Como os dados espaciais são sempre uma aproximação ou generalização da realidade, eles são cheio de incerteza e precisão. Uma mudança de modelo ou escala de dados pode introduzir uma perda de informações, assim como digitalização ou varredura. O processamento em uma máquina finita também insere sua própria forma de incerteza, embora isso geralmente seja insignificante em relação aos erros inerentes aos próprios dados. Muitas construções geográficas humanas são implicitamente incertas, incluindo objetos espaciais ('Oceano Índico', 'Europa') e seus relacionamentos ('in', 'across'). Quer pensemos incerta em termos teóricos definidos através de noções de imprecisão ou em termos estatísticos através do cálculo de probabilidades, o estudo de a incerteza de dados espaciais, sua medição e modelagem, e a análise de sua propagação através dos processos de manipulação de dados espaciais são, sem dúvida, parte da ciência da informação geográfica. Como se deve compilar uma representação precisa da variação geográfica para entrada em um banco de dados? Como se deve representar a incerteza ou imprecisão presente em uma representação digital? Como a incerteza pode ser propagada do banco de dados para os produtos GIS? Dados geográficos trazem seu próprio conjunto especial de problemas para as estatísticas espaciais. Enquanto nas imagens médicas o problema pode ser determinar a verdadeira localização dos objetos a partir de imagens "sujas" (Besag 1986), nas imagens geográficas muitas vezes não existe um conceito claro de verdade, pois os objetos costumam ser produtos de interpretação ou generalização. Precisamos de métodos muito melhores para medir e descrever a incerteza, particularmente nos complexos objetos espaciais comuns no SIG. Precisamos de métodos melhores para lidar com o mundo como um conjunto de contínuos sobrepostos, em vez de forçar o mundo a moldar objetos rigidamente limitados. A maioria das respostas a essas perguntas terá de vir de estatísticas espaciais, mas especialistas em informações geográficas devem fornecer a motivação e os exemplos e definir os objetivos e restrições gerais. . Embora todos os dados geográficos sejam incertos até certo ponto, toda a geração atual de SIG segue a prática comum em cartografia e representa objetos geográficos como se suas posições e atributos fossem perfeitamente conhecidos; a qualidade dos dados pode · ou não pode ser tratada em uma declaração separada · · As consequências da incerteza para produtos GIS nunca são estimadas. Pesquisas recentes seguiram várias linhas diferentes e produtivas na tentativa de abordar o problema da qualidade dos dados. Uma é combinar precisão com exatidão. Em um sentido local, isso significa usar precisão limitada na representação e processamento dos dados com mais freqüência, através do uso de uma varredura cujo tamanho é determinado pela precisão dos dados. Várias formas de estrutura quadtree também foram usadas para ajustar a precisão local a níveis conhecidos de precisão. Existem vários artigos recentes sobre processamento de resolução finita no SIG (por exemplo, Franklin 1984, Dutton 1989) e a geometria de resolução finita é uma área de pesquisa ativa em matemática. Outra abordagem produtiva foi incorporar técnicas de geoestatísticos, notadamente kriging, pois a base estatística dessas técnicas torna explícita a incerteza. Agora, temos vários modelos úteis de erro de digitalização e suas consequências para medidas estimadas, como área (por exemplo, Chrisman e Yandell 1988, Keefer et al. 1988). Finalmente, houve vários esforços bem-sucedidos para modelar conjuntos de dados geográficos como campos aleatórios ou derivadas de campos aleatórios e usar essa abordagem para modelar incertezas em objetos GIS (por exemplo, Goodchild 1989). Entre todos esses métodos, provavelmente agora temos um conjunto adequado de modelos de precisão para criar um G ~ S de rastreamento de erros. No entanto, a estatística espacial não é um campo fácil, e muitas dessas técnicas vão muito além da estatística elementar em sua sofisticação conceitual.

3.4. Modelagem de dados e teorias de dados espaciais

Modelos de dados são as estruturas lógicas que usamos para representar bancos de dados indigitais de variação geográfica. Como cada um deve ser uma aproximação, a escolha entre modelos alternativos restringe não apenas as funções disponíveis, mas também a precisão de produtos. De todos os desenvolvimentos no SIG na década passada, talvez o mais empolgante tenha sido a proliferação de modelos de dados e a crescente literatura sobre seus méritos relativos. O debate sobre varredura e vetor remonta aos primeiros dias, mas agora é acompanhado por debates sobre objetos, camadas, a filosofia da orientação a objetos, modelos hierárquicos de objetos complexos e toda a gama de possibilidades inerentes à dependência do tempo e às três dimensões. Apesar do interesse, ainda não temos um estrutura completa e rigorosa para modelagem de dados geográficos, mesmo no caso estático bidimensional, e sem um é difícil ver como o GIS pode escapar das restrições impostas pelas implementações específicas do sistema. Quanta capacidade está sendo perdida ao forçar aplicativos contemporâneos ao modelo de varredura multicamada usado por muitos sistemas ou ao modelo de cobertura de área de linha de ponto usado por muitos outros? Este é um problema de pesquisa pura e aplicada. Por um lado, devemos desenvolver uma estrutura abrangente para a distribuição geográfica. modelagem de dados, com uma terminologia associada, · para fornecer a base para padrões e um ideal. contra os quais sistemas específicos podem ser medidos. Por outro lado, uma estrutura abstrata é de pouco valor se não influenciar a prática, através da implementação nos produtos dos fornecedores. Aqui, a verdadeira questão é se é possível ampliar ou "modernizar" o modelo de dados. subjacente a um produto existente ou se qualquer · tentativa de fazê-lo está condenado a causar inconsistência e incoerência. Essas questões estão precipitando animada discussão sobre toda a questão do grau para o qual · nós vemos. analisar, representar e modelar o mundo como discreto ou contínuo, como uma coleção de objetos ou um conjunto de campos. Pensamos em termos de variáveis ​​com valores definidos em todo lugar no espaço, ou de um espaço · cheio de objetos possivelmente sobrepostos? Em essência, essas questões trouxeram o debate GIS da obscuridade comparativa das estruturas internas de dados para as questões muito mais gerais de como entendemos a variação geográfica. Experiência humana cotidiana. vê · um mundo de objetos, mas · a ciência dos processos naturais · lida mais com a variação contínua (frank e Mark 1991). Assim, o debate orientado a objetos ameaça colocar os New Agers contra. os restos em conflito do Iluminismo, e o que poderia ser mais estimulante do que isso?

3.5 Estruturas, algoritmos e processos de dados

Muitos dos resultados de pesquisas básicas acumuladas nos últimos 25 anos nesse campo de pesquisa dizem respeito a representações internas de dados e aos algoritmos que operam neles. O quadtree (Samet 1989), os algoritmos de varredura de banda para sobreposição (White 1977), a análise da complexidade computacional (Preparata e Shamos 1988) e a estrutura de dados dos nós do arco (Peucker e Chrisman 1975) são todos avanços intelectuais de significância duradoura. Muitos problemas desafiadores permanecem, por exemplo, no design de algoritmos eficientes para minimizar a sobreposição e em outras áreas do design cartográfico, ou no desenvolvimento de melhores métodos para a conversão entre vários modelos de dados do terreno. Muitos sistemas agora lidam com dados através de sistemas de gerenciamento de banco de dados, e os problemas de estrutura de dados mudaram mais. e mais no campo da ciência da computação. Parece, no entanto, que chegamos a um ponto em que todos os problemas mais simples e genéricos foram resolvidos e o que resta é um conjunto de problemas difíceis e específicos ao contexto. Parece claro, por exemplo: nple. que avanços adicionais na versão de modelos de dados do terreno (por exemplo, do contorno ao TIN) exigirão a. entendimento muito melhor da natureza do terreno (Mark 1979) e talvez tenha que ser específico ao tipo de terreno (por exemplo, fluvial versus glacial). Também continuará sendo necessário pesquisar sobre métodos eficientes de armazenamento e acesso para lidar com os enormes volumes de dados que provavelmente estarão disponíveis na próxima década.

3.6 Exibição

Os sistemas de informações geográficas geralmente são criticados por não dar atenção adequada aos princípios do design cartográfico (Buttenfìeld e Mackaness 1991), ou por considerar o mapa como um simples armazenamento de informações e não como uma ferramenta para comunicação. Se pensarmos no banco de dados como a verdade, um mapa não passa de uma loja, pois geralmente existe uma correspondência simples entre objetos no banco de dados e objetos no mapa. No entanto, se o banco de dados é visto apenas como uma aproximação da verdade geográfica, o design das exibições de saída é crítico, pois pode afetar a visão do mundo do usuário. Coisas simples, como a escolha da cor de fundo ou o contraste entre polígonos adjacentes · (McGranaghan 1991) podem ter um efeito significativo. Os recursos do display eletrônico vão muito além dos da cartografia convencional. Precisamos de pesquisas sobre o design de telas animadas, telas tridimensionais, o uso de ícones e metáforas nas interfaces do usuário, gradação contínua de cores e tons, zoom e navegação, várias mídias, incluindo dispositivos de voz e apontadores, várias janelas que permitem acesso simultâneo a séries espaciais e temporais de dados multivariados. Precisamos usar o meio eletrônico para pensar muito além das melhorias no design dos mapas de coropletas. Todos esses são problemas fundamentais para uma ciência da informação geográfica.

3.7 Ferramentas analíticas

Um GIS é uma ferramenta para apoiar uma ampla gama de técnicas de análise espacial, incluindo processos para criar novas classes de objetos espaciais, analisar as localizações e atributos de objetos e modelar usando várias classes de objetos e as relações entre eles. eles .. · inclui operações geométricas primitivas, como o cálculo dos centróides de polígonos ou a criação de buffers em torno de linhas, além de operações mais complexas, como a determinação do caminho mais curto através de uma rede. A funcionalidade dos principais produtos continua a crescer, sem um fim óbvio à vista. Apesar do amplo reconhecimento de que a análise é central para o objetivo de um SIG, a falta de integração do SIG e da análise espacial e a simplicidade comparativa da funcionalidade analítica de muitos sistemas continuam sendo uma grande preocupação. Nos primeiros dias do pacote estatístico SAS, houve um aumento muito rápido na gama de testes e técnicas implementadas no. sistema. Infelizmente, o mesmo não é verdade GIS, e notavelmente pouco progresso foi feito na incorporação da gama de técnicas conhecidas de análise espacial. produtos atuais. Há muitas razões para isto. Uma razão óbvia é a forte ênfase no mercado de GIS no gerenciamento de informações, e não na análise. Os mercados lucrativos da tecnologia GIS têm necessidades comparativamente pouco sofisticadas. enfatizando consultas e tabulações simples. Outra é a relativa obscuridade da análise espacial, um conjunto de técnicas desenvolvidas em várias disciplinas, sem nenhum sistema claro de codificação ou forte quadro conceitual ou teórico. Mesmo agora · é difícil identificar mais de um punhado de textos (por exemplo, Haining, 1990, Upton e Fingleton, 1985). Embora se possa esperar que o SIG possa fornecer a base para um sistema de codificação para análise espacial, o baixo nível de entendimento atual dos modelos de dados geográficos é uma grande dificuldade. Tomlin (1990) fez uma das poucas tentativas de adicionar algum tipo de estrutura ou estrutura à proliferação de funções GIS, que no caso do ARC / INFO já estão em torno de 103. Nós realmente. precisam de uma taxonomia de análise espacial, desenvolvida talvez a partir de um conjunto enumerado de modelos de dados, mas indo muito além das operações geométricas primitivas. Nesta fase, a integração do SIG e da análise espacial está ocorrendo lentamente, em pelo menos três modos diferentes. Alguns recursos analíticos estão sendo adicionados diretamente ao GIS, por exemplo na recente expansão da funcionalidade em vários módulos para análise de rede. Algum progresso está sendo feito na análise pouco acoplada, em que um módulo de análise independente depende do aGIS para seus dados de entrada e para funções como exibição. No entanto, ainda está faltando uma forma eficaz de acoplamento rígido, no qual os dados podem ser passados entre um GIS e um módulo de análise espacial sem perda de estruturas superiores, como topologia, identidade de objeto, metadados ou vários tipos de relações. No momento, isso é impossível, em grande parte devido à falta de padrões para modelos de dados. Em vez de. o acoplamento deve ocorrer em um nível inferior e as estruturas superiores precisam ser reconstruídas de forma arbitrária.

A integração entre o SIG e a análise espacial também pode levar de uma · linguagem, cujos elementos primitivos representariam as operações fundamentais de análise espacial. O início de uma linguagem já existe nas macro-linguagens de muitos da geração atual de GIS e em várias tentativas de estender o SQL para operações espaciais. No entanto, todos eles são específicos e dependem fortemente de modelos de dados limitados, e há notavelmente pouca similaridade entre eles no momento. Em Santa Barbara, tentamos definir uma linguagem comum a partir de uma análise das linguagens usadas por uma variedade de GIS atuais, mas uma solução mais satisfatória começaria com a estrutura conceitual fornecida por um modelo de dados abrangente. . Outro problema na integração de SIG e análise espacial é que, na Cornier, a discretização do espaço é explícita, enquanto que em muitas formas de análise espacial, ela é frequentemente implícita ou não especificada. Muitas formas de análise espacial são escritas em campos contínuos e não conseguem lidar com as incertezas introduzidas pelo inevitável processo de discretização. Por exemplo, no GIS, não pode haver medida da inclinação independente da discretização, e da mesma forma o comprimento do limite de um objeto de área depende de sua representação digital. No entanto, inclinação e comprimento geralmente aparecem como parâmetros não qualificados em modelos espaciais. Nesse sentido, a integração do GIS e a análise espacial é um processo bidirecional, no qual as inadequações do GIS e da análise espacial devem ser abordadas. A maior parte da geração atual de GIS fornece algum tipo de recurso de macro ou script, permitindo o usuário definir produtos de sequências complexas de operações, mas invocá-los com uma única instrução. Embora muitas vezes incluam a capacidade de construir ambientes e interfaces customizados, ainda não fornecem ferramentas específicas para as necessidades da análise espacial. Uma exceção limitada é o ATB da Prime / Wild, um conjunto de ferramentas construídas sobre o System / 9 que permite ao usuário trabalhar com análises complexas, visualizar suas seqüências e gerenciar resultados intermediários. Ferramentas como essa serão cada vez mais necessárias para que o GIS mude para uma era de análise e suporte à decisão mais sofisticados, porque não é incomum o GIS relativamente simples produtos para envolver o processamento de dezenas de camadas através de números semelhantes de etapas primitivas. Precisamos pesquisar métodos para acompanhar a linhagem de dados e a propagação de erros, voltar atrás para recuperar resultados intermediários e impedir que o usuário combine operações de maneira incorreta ou sem sentido (Lanter, 1990). Também precisamos de pesquisas sobre maneiras de incorporar esse tipo de análise no processo de aquisição e planejamento de GIS. Essa ênfase na análise complexa de vários estágios e na geração de produtos a partir de um banco de dados multicamada parece muito diferente da pesquisa em sistemas baseados em conhecimento, raciocínio espacial e consulta espacial. Uma das atrações do campo GIS é a sua variedade de aplicações e a variedade extremamente extrema de ambientes para o design de interfaces de usuário. Na modelagem de dados, a questão importante não é se modelos relacionais estendidos ou orientados a objetos são melhores para dados geográficos, mas que tipos de dados geográficos são melhor modelados por cada abordagem. Da mesma forma, a importante questão de pesquisa no design de interfaces de usuário é determinar a melhor ambiente para cada um dos muitos tipos de aplicativos GIS. O que é melhor para um sistema de navegação de veículos pode ser totalmente diferente do que é melhor para um gerente de recursos florestais com um profundo medo de teclados e VDUs, coloridos ou monocromáticos.

3.8 Questões institucionais, gerenciais e éticas

As pesquisas estão começando a aparecer sobre as questões envolvidas na implementação e gerenciamento de SIG, especialmente em grandes instituições. Essa é uma pesquisa difícil e as generalizações não são descobertas facilmente. No entanto, o sucesso de vários grandes projetos nos EUA e as discussões em torno de várias grandes aquisições por agências federais criaram a oportunidade para vários estudos de caso úteis. São necessários muitos mais, especialmente dada a importância de tais pesquisas para melhorar o ambiente institucional no futuro. Precisamos de uma compreensão muito melhor dos processos de adoção da tecnologia GIS e seus efeitos nas organizações; do valor da informação geográfica e dos benefícios do SIG; e de processos para a utilização de informações geográficas na tomada de decisões. As estruturas teóricas para abordar muitas dessas questões já existem nas disciplinas relevantes das ciências sociais, e precisamos · fazer · uso muito mais eficaz delas · lidar com questões específicas do SIG. Apesar dos problemas envolvidos na adoção de qualquer nova tecnologia, os GIS foram amplamente adotados no governo local, nos serviços públicos e nas agências de gerenciamento de recursos. De fato, a introdução do SIG teve um grande efeito no gerenciamento de informações geográficas na sociedade. Ao mesmo tempo, há uma preocupação crescente com o poder do SIG pela vigilância e invasão de privacidade. A comunidade de pesquisa tem a responsabilidade de monitorar e estudar os principais aspectos do fenômeno GIS, incluindo seu significado para a sociedade como um todo. O que o SIG significa para o equilíbrio de poder na sociedade? Serão uma tecnologia disponível apenas para os que têm poder ou de alguma forma servirão para equilibrar a distribuição de energia? Até o momento, existem poucos estudos sobre a ética do SIG.

4. Testes de comunalidade

As seções anteriores examinaram várias áreas candidatas para inclusão em uma ciência da informação geográfica. Em cada caso, há claramente questões científicas desafiadoras a serem colocadas e pesquisadas. Não há razão para acreditar que a lista esteja completa ou que não haja perguntas adicionais e substanciais em outras áreas relacionadas. Em cada caso, o contexto espacial parece ser distinto. embora claramente seja mais em alguns do que em outros. Por exemplo, podemos discutir se o contexto espacial era distinto na área da teoria da decisão, mas a questão parece clara para a modelagem de dados. . . No plano de pesquisa da NCGIA (NCGIA 1989), discutimos isso. a ausência de soluções para questões como essas constituía impedimentos para as aplicações efetivas da tecnologia GIS. Outras discussões da agenda de pesquisa de GIS chegaram a conclusões semelhantes, embora com ênfases diferentes (Craig 1989, Maguire 1990, Masser 1990). Muitos são problemas antigos, reconhecidos muito antes do advento do SIG. em áreas como cartografia, geodésia e geografia. Alguns podem não ser exclusivos do SIG. Por exemplo, não é imediatamente óbvio que a tecnologia GIS se difunda de uma maneira fundamentalmente diferente ou mostre padrões fundamentalmente diferentes de adoção de outras tecnologias. A medição dos benefícios do SIG é um problema único ou um exemplo do problema mais geral de medir os benefícios da tecnologia da informação? É claro que essas questões são questões de pesquisa em si mesmas.

Ao mesmo tempo, é muito importante identificar as áreas em que o GIS criou · problemas novos e únicos que não são comuns a outros campos. Nos primeiros dias do SIG, era possível argumentar que a tecnologia estava preenchendo uma lacuna existente e possibilitando tarefas que haviam sido identificadas anteriormente, mas que não eram fáceis de executar manualmente. O uso do GIS para análise de adequação, sobrepondo camadas (Tomlin, 1990), espelha a técnica manual popularizada por. Por McHarg, embora admitindo adicionar novas capacidades interessantes. O CGIS foi justificado com o argumento de que o computador foi um custo-benefício. alternativa à medição manual. áreas cobertas. Mas GIS faz isso. É possível fazer coisas com dados que os coletores de dados talvez nunca tenham imaginado.

A tecnologia GIS está produzindo mudanças radicais na maneira como os dados geográficos são coletados, manipulados e analisados, e levará muitos anos até que o impacto da tecnologia existente seja sentido, sem falar nos impactos de desenvolvimentos futuros. Aqui estão alguns dos problemas que parecem exclusivos para o GIS: como modelar dados geográficos dependentes do tempo; como capturar, armazenar e processar dados geográficos tridimensionais; como modelar dados para distribuições geográficas cobertas por superfícies incorporadas em três dimensões; como explorar esses dados, por exemplo, quais metáforas exploratórias são úteis e como avaliar a perspectiva geográfica da informação. e processos relativos a perspectivas mais convencionais? Essas são questões importantes para o GIS e as necessidades da comunidade GIS. um forte compromisso com a pesquisa, se isso for significativo. progresso neles. Como questões que surgem no contexto do SIG, elas não são motivo de grande preocupação em outras disciplinas.

No entanto, ao mesmo tempo, a comunidade GIS pode se beneficiar enormemente da pesquisa interdisciplinar. Os estatísticos podem dar uma contribuição muito valiosa para resolver o problema do erro no SIG, e a pesquisa em psicologia cognitiva pode ser útil para projetar os aspectos cognitivos das interfaces de usuário no SIG. Esse argumento leva naturalmente a uma definição proposta de pesquisa GIS: pesquisa sobre questões genéricas que envolvem o uso da tecnologia GIS, impede sua implementação bem-sucedida ou surge do entendimento de suas potencialidades. Isso é 'pesquisa sobre SIG' ou 'pesquisa com SIG'? Em certo sentido, são os dois, porque são questões fundamentais para a tecnologia do SIG e também questões que devem ser resolvidas antes que a tecnologia possa ser aplicada com sucesso. Se os problemas de pesquisa com SIG forem genéricos, eles serão melhor abordados como parte da agenda de pesquisa de SIG. No entanto, problemas específicos à aplicação do SIG em um campo específico precisam claramente ser abordados no contexto desse campo e com o benefício de sua experiência. Questões de precisão fornecem um exemplo útil. Existem aspectos do problema de precisão que abrangem uma ampla gama de tipos de dados geográficos e precisam ser resolvidos usando modelos genéricos de incerteza, análogos ao papel desempenhado pela distribuição gaussiana na teoria do erro de medição. No entanto, como observado anteriormente, uma análise de dados de crimes usando um SIG também levantará problemas de precisão específicos para esse aplicativo em particular e necessitarão de um entendimento dos processos que operam em criminologia e na coleta de dados de crimes, se quiserem ser usados. entendido completamente. No entanto, a mera existência de questões científicas está longe de ser uma base adequada para uma ciência. Existe um ponto em comum de interesse aqui? Esses subcampos podem encontrar bases suficientes para a interação para desenvolver os últimos aportes de uma ciência, como periódicos, sociedades, livros e filósofos? Os pesquisadores desses subcampos serão considerados um grupo de estudiosos? Existe uma analogia válida entre os sistemas e a ciência da informação geográfica, por um lado (ferramentas de apoio aos pesquisadores) e pacotes estatísticos e estatísticas, por outro? A estatística é uma disciplina altamente formalizada, mas grupos mais tecnologicamente orientados podem ser encontrados em áreas como dados exploratórios análise, visualização estatística e estatística aplicada. Certamente, às vezes, a relação entre ciência e ferramentas é tempestuosa, mas, no entanto, vital para o sucesso de ambas. O debate em andamento sobre o valor do software estatístico no ensino de estatística tem implicações interessantes para o mesmo problema no SIG. · Pode ser útil examinar brevemente os argumentos de interesse comum em ciência da informação geográfica, primeiro em princípio e depois na prática. O campo é uma pequena retórica sobre o crescimento do setor, à parte, ninguém sugeriria que o campo do GIS é uma disciplina importante. II é distinto, com seu próprio conjunto de perguntas razoavelmente único. E é certamente desafiador e inatamente atraente. Do lado negativo, é multidisciplinar, competindo com clivagens e rivalidades de longa data. Falta uma disciplina central, diferentemente da analogia estatística, onde houve um crescimento constante do número e tamanho dos departamentos acadêmicos nas últimas décadas. Um dos principais requerentes, a geografia, tem sido tradicionalmente um campo não técnico e, em algumas áreas da geografia social, existe uma forte e fundamental antipatia às abordagens tecnológicas. ·. Na prática, o interesse comum é evidente na proliferação de reuniões GIS,. · E estamos começando a ver a. fornecimento de livros e revistas. Contudo, o conhecimento científico O rastreamento nas reuniões GIS geralmente é pequeno. As pessoas que participam das reuniões de GIS precisam de um suprimento constante de novidades, seja em pesquisa científica ou em produtos de fornecedores, e logo desaparecerão se o suprimento secar.

5. Opções para o futuro

Olhando para trás ao longo de quase três décadas de pesquisa em GIS, fica claro que o maior progresso foi alcançado nos problemas mais bem definidos e fáceis, onde as soluções estão nos avanços da própria tecnologia. Houve um rápido progresso em algoritmos e estruturas de dados nas décadas de 1970 e 1980, mas muitos dos problemas difíceis da modelagem de dados. permanece a modelagem de erros, a integração da análise espacial e as questões institucionais e gerenciais. Alguns deles podem ser insolúveis: por exemplo, pode não haver simplesmente generalidades a serem descobertas no processo de · adoção de SIG por agências governamentais, por mais fácil que seja colocar a questão da pesquisa. Outras questões já foram resolvidas no sentido de pesquisa pura, mas a implementação continua sendo uma questão importante de pesquisa aplicada. Em termos de precisão, por exemplo, um conjunto substancial de técnicas foi definido, mas o problema ou a sua transferência para a aplicação real permanece. O ambiente de pesquisa acadêmica é criado para buscar áreas significativas de pesquisa, mas geralmente é pobre em fornecer os meios de implementação. Para isso, precisamos de uma indústria de software fortemente acoplada à comunidade de pesquisa, mas capaz de encontrar os recursos para motivar o desenvolvimento. Mais importante, precisamos de um sistema educacional que responde rapidamente a novas pesquisas e é. capaz de criar novos conceitos rapidamente em seus programas. Infelizmente, o setor de ensino superior é muitas vezes caracterizado pelo conservadorismo. Pode levar muitos anos para que novas idéias se inseram no currículo. A pesquisa em SIG é como dados geográficos, quanto mais se olha, mais questões interessantes aparecem. A pesquisa GIS apenas começou a abordar questões importantes no. a agenda de pesquisa. Estamos em uma posição invejável, trabalhando em um campo com uma motivação tão forte e uma indústria subjacente tão forte, e com um conjunto de problemas tão interessantes que abrangem tantas disciplinas e campos. Espero ter demonstrado neste artigo que o manuseio de informações espaciais com a tecnologia GIS apresenta uma variedade intelectual e científico. desafios muito maiores do que a expressão 'manipulação de dados espaciais' implica, de fato, uma ciência da informação geográfica. O termo "geográfico" parece essencial muito do que se refere à pesquisa de SIG, concernente ao mundo geográfico e às nossas relações. com ele, e o termo é muito mais rico que "espacial". A mudança no significado da palavra 'S' de sistemas para ciência parece estar indo bem, como evidenciado pelo sucesso da série de conferências sobre manipulação de dados espaciais, a mudança da série AutoCarto para artigos com referências completas, novos textos, assinaturas de Revista Internacional de Sistemas de Informação Geográfica e submissões de GIS documentos para · periódicos estabelecidos como .Geoqraphical Analysis, Computers .e Geociências, visão computacional, gráficos e processamento de imagens e publicações da Regional Science Association e do IEEE.

   Espero também ter demonstrado que um forte programa científico serve não apenas a si mesmo, mas também às necessidades da indústria e dos usuários de GIS. Os SIG precisam de um forte componente científico e intelectual para que sejam mais do que um fenômeno comercial, a. flash de curta duração no panorama tecnológico. É muito fácil selecionar o GIS atual como uma tecnologia de hardware e softwate em busca de aplicativos e ver o campo do GIS definido de acordo com os limites funcionais de seus principais produtos de fornecedores. Precisamos mudar do sistema para a ciência, para estabelecer o SIG como a interseção entre um grupo de disciplinas com interesses comuns, apoiado por uma caixa de ferramentas tecnológica e, por sua vez, apoiando a tecnologia por meio de sua pesquisa básica. Como atualmente percebido, o GIS às vezes parece tão próximo de uma ciência quanto o FORTRAN é da álgebra.

 Nos últimos anos, vimos uma crescente clivagem no SIG entre duas tradições: a de informação espacial, por um lado, e a de análise espacial, por outro. A tradição de informações espaciais enfatiza grandes bancos de dados de inventário e atribui à geografia o papel de um mecanismo de acesso. A tradição da análise espacial enfatiza uma rica funcionalidade e uma variedade de modelos de dados, e atribui à geografia um papel fundamental na análise e modelagem. As duas tradições compartilham estruturas e algoritmos comuns de dados e contam com as mesmas fontes de dados e hardware. No entanto, isso não é suficiente para convencer a academia da existência de um campo científico. Para reivindicar isso, precisamos ter uma visão mais ampla e incluir precisão na modelagem de dados, cognição, raciocínio, interfaces homem-computador (HCI) e visualização, e mostrar como essas são partes integrais de ambas as tradições.

Sem esses argumentos, o campo GIS se fragmentará e a tempestade GIS se explodirá. Associações tão fundamentalmente desassociadas quanto a Association of American Geographers e AM / FM acharão impossível justificar o patrocínio conjunto de conferências. Os fornecedores se especializarão em estações de trabalho de entrada de dados, estações de análise espacial ou sistemas de gerenciamento de instalações, com pouco potencial para interação ou integração. Isso seria trágico.

Como podemos garantir um futuro duradouro para os sistemas de informação geográfica e para a ciência? As disciplinas são como tribos, com seus próprios totens, símbolos e membros regras, idiomas e redes sociais, a tribo GIS é atualmente muito coesa; é bem financiado, o campo é empolgante e muita pesquisa útil está sendo feita. No entanto, a longo prazo, o campo não se saiu bem em beber como ciência, e a academia ainda é. duvidoso sobre a necessidade de ser levado a sério. A ciência é difícil e coloca obrigações para com seus praticantes. Estamos muito ocupados e a tecnologia está se movendo muito rapidamente. Muita da nossa literatura está em anais de conferências, que trazem exposição rápida, mas apenas para audiências limitadas, e carecem de controle de qualidade suficiente. Poucas pessoas tiveram tempo para escrever os livros, identificar o núcleo intelectual ou publicar bons exemplos.

Acredito que garantimos o futuro do SIG pensando mais na ciência do que em sistemas e identificando as principais questões científicas do campo e percebendo sua amplitude intelectual. Os sistemas de informação geográfica são uma ferramenta para a ciência da informação geográfica, a qual, por sua vez, levará à sua eventual melhoria. Precisamos falar com a academia, diretamente e através de artigos e textos importantes, sobre a filosofia, metodologia e fundamentos do campo, e colocando artigos de GIS em periódicos fortes. Todos os três usuários, fornecedores e pesquisadores da comunidade têm papéis vitais e simbióticos a desempenhar, e serviremos aos três melhores ao desempenhar o nosso · no sentido mais amplo possível. ·

Agradecimento .

O Centro Nacional de Informação e Análise Geográfica é suportado pela National Science Foundation através da concessão SES 88-10917.

