

MODELAGEM CONCEITUAL E GEOGRÁFICA: O ESTADO DA ARTE NACIONAL E INTERNACIONAL NO NOVO MILÊNIO

Conceptual and Geographic Modeling: The National and International State of the Art in the New Millennium

Vivian da Silva Celestino Reginato
Universidade Federal de Santa Catarina
PPGTG – Departamento de Engenharia Civil
vivian.celestino@ufsc.br

Francisco Eduardo de Mello Franco Sá
Universidade Federal de Santa Catarina
PPGTG
franciscoeduardo.mfsa@gmail.com

Arthur Campagnoli Morossino
Universidade Federal de Santa Catarina
PPGTG
arthurcampagnolimorossino@gmail.com

Ricardo Augusto Eger
Universidade Federal de Santa Catarina
PPGTG
ricardo.a.eger@gmail.com

André Felipe Bozio
Universidade Federal de Santa Catarina
PPGTG
andrefbozio@gmail.com

Resumo:

Processos que envolvem a modelagem ganharam relevância nos últimos anos nas áreas relacionadas à gestão territorial, pois utilizam modelos de dados baseados no formalismo orientado à objetos, para abstrair os diferentes fenômenos do mundo real e gerar esquemas conceituais que representam as linguagens que deverão ser captadas pelos Sistemas de Informações Geográficas (SIG). Esses sistemas quando desenvolvidos de forma organizada facilitam a análise e integração de dados e auxiliam na tomada de decisão. Muita informação tem sido produzida nos últimos anos, mas em algumas áreas, é verificada uma carência em relação aos trabalhos que envolvem modelagem de dados. Neste sentido esta pesquisa teve por objetivo analisar o estado da arte da produção científica nacional e internacional acerca da modelagem conceitual e geográfica no novo milênio, realizando uma revisão sistemática que retrata o comportamento e o desenvolvimento do tema através de métricas. Como resultados da revisão, após processo de filtragem, foram analisados 37 artigos científicos publicados entre os anos de 2001 e 2022 advindos de quatro bases científicas, na área de modelagem, através da busca de quatro palavras-chave. Foram analisados os resultados e apontadas as lacunas teóricas e empíricas existentes, identificando caminhos para futuras ações, onde o tema do cadastro foi diagnosticado como o mais urgente de ser tratado em pesquisas, tanto no Brasil quanto no restante do mundo.

Palavras-chave: Revisão sistemática; Gestão Territorial; Diagrama de classes, Dicionário de dados.

Abstract

Processes involving modeling have gained relevance in recent years in areas related to territorial management, as they use data models based on object-oriented formalism to abstract the different phenomena from the real world and generate conceptual schemes that represent the languages that should be captured by Geographic Information Systems (GIS). These systems, when developed in an organized way, facilitate the analysis and integration of data and help in decision making. Much information has been produced in recent years, but in some areas, there is a lack of work that involves data modeling. In this sense, this research aimed to analyze the state of the art of national and international scientific production on conceptual and geographic modeling in the new millennium, carrying out a systematic review that portrays the behavior and development of the theme through metrics. As a result of the review, after a filtering process, 37 scientific articles published between the years 2001 and 2022 were analyzed from four scientific bases, in the area of modeling, through the search of four keywords. The results were analyzed and the existing theoretical and empirical gaps were pointed out, pointing out ways for future actions, where the issue of registration was identified as the most urgent to be addressed in research, both in Brazil and in the rest of the world.

Keywords: Systematic review; Territorial Management; Class diagram, Data dictionary.

1 INTRODUÇÃO

A modelagem ganhou relevância nos últimos anos, fruto dos desenvolvimentos metodológico e tecnológico. Por um lado, as áreas ligadas à gestão territorial têm desenvolvido ferramentas de análise e, por outro, as componentes de *hardware* e *software* também evoluíram substancialmente. Em qualquer processo de modelagem são utilizados modelos de dados com base em algum formalismo conceitual, como o Modelo Entidade Relacional (ER), ou o modelo Orientado ao Objeto (OO) (LISBOA FILHO; IOCHPE, 1999). O resultado do processo de modelagem conceitual é o esquema conceitual, representado através de uma linguagem formal de descrição, que possui uma sintaxe e/ou uma notação gráfica (LISBOA FILHO; IOCHPE, 2001).

Já a modelagem geográfica, utilizada desde a década de 1980 em Sistemas de Informações Geográficas (SIG) (BÉDARD e PAQUETE, 1989), era trabalhada dentro das estruturas do sistema, forçando o usuário a adequar a realidade geográfica que estava sendo modelada às estruturas disponíveis no SIG, ou seja, o processo de modelagem não oferecia mecanismos para que o usuário pudesse modelar a realidade de uma forma mais próxima ao seu modelo mental (BORGES; DAVIS JR.; LAENDER, 2005).

De lá para cá, juntamente com o desenvolvimento dos SIG, a modelagem tem facilitado os trabalhos de captura, análise, manipulação, integração e gestão de grandes quantidades de dados relativos à Terra, estruturando e organizando de uma forma que estes possam subsidiar problemas complexos de planejamento e gerenciamento, podendo ser utilizados por públicos diversos e, não somente, àqueles ligados à gestão do território.

Com este cenário exposto, o seguinte questionamento é realizado: como estão sendo realizadas e aplicadas as pesquisas relativas à modelagem conceitual e geográfica no Brasil e no mundo? Para responder a esse questionamento este trabalho tem por objetivo analisar o estado da arte da produção científica nacional e internacional acerca da modelagem conceitual e geográfica no novo milênio, realizando uma revisão sistemática que retrate o comportamento

e o desenvolvimento da temática, apontando métricas e lacunas teóricas e empíricas para futuras ações.

2 MODELAGEM CONCEITUAL E GEOGRÁFICA: O ESTADO DA ARTE

A modelagem é o processo de geração de elementos adicionais a partir de conjuntos de dados existentes. Em SIG diversas metodologias consideram a etapa da modelagem importante, (BERTINI; CESAR NETO, 2004) por utilizarem entidades geográficas, ao invés das convencionais, com relacionamentos diversos, como os topológicos, aos utilizados comumente nos dados relacionais advindos de tabelas (entidades semânticas).

Já os modelos de dados são conjuntos de conceitos que podem ser usados para descrever a estrutura e as operações de um Banco de Dados (BD), buscando sistematizar o entendimento que é desenvolvido a respeito das feições, elementos, objetos e fenômenos que serão representados em um sistema informatizado. Os modelos são de vital importância para um SIG porque controlam a forma como os dados são armazenados e tem um impacto significativo nos tipos de operações analíticas que podem ser realizadas (LONGLEY *et al.*, 2013).

Segundo Lisboa Filho e Iochpe (2001), o projeto de um BD geográfico deverá ser pensado com base em um alto nível de abstração da realidade geográfica, já que independe de *hardware* e *software*. Desta forma, pelos SIG englobarem dados de diferentes fontes, é necessário acordar métodos e padrões comuns de modelagem de dados geográficos, a fim de estabelecer o modelo conceitual de dados que será adotado.

Existem vários tipos de modelos de dados, podendo ser divididos, basicamente em conceituais, lógicos, físicos quando atrelados a parte física e computacional e, conceituais e geográficos, quanto utilizados para aplicações de dados geográficos.

A modelagem conceitual trata da descrição da semântica de aplicativos de *software* em um alto nível de abstração, ou seja, em termos de entidades, relacionamentos e restrições, comportamentos, funções, transições, interações e interfaces entre usuário e informações trocadas. Tem o objetivo de descrever os possíveis conteúdos dos dados, além das estruturas e regras a eles aplicáveis, fornecendo uma base formal, notacional ou semântica, para as ferramentas e técnicas utilizadas no suporte da modelagem de dados (LISBOA FILHO; IOCHPE, 1999). O esquema conceitual é composto por um conjunto de diagramas, que expressam esquematicamente como os dados devem ser coletados, organizados e relacionados entre si, para atender a uma aplicação específica (QUEIROZ; FERREIRA, 2006).

Na modelagem lógica o esquema conceitual se transforma em um esquema de dados compatíveis com o tipo de modelo de Sistema Gerenciador de Banco de Dados (SGBD) utilizado. Nesta fase o mapeamento do esquema conceitual é transformado em representação de dados do modelo escolhido (LAGO, 2006).

Na modelagem física, também chamada de projeto físico, são definidos aspectos de implementação do BD (tipos de dados, estruturas de armazenamento, caminhos de acesso, particionamento, agrupamento, etc.), levando em consideração o SGBD que está sendo utilizado, permitindo ao projetista, desta forma, planejar aspectos ligados à eficiência do mesmo (LISBOA FILHO; IOCHPE, 2001).

Com os avanços nas técnicas de modelagem conceitual de dados geográficos, inúmeros modelos e técnicas foram desenvolvidos. São encontrados na literatura diversos modelos

conceituais de dados para aplicações geográficas. Um dos modelos mais conhecidos no que tange à OO é o modelo OMT (*Object Modeling Technique*) desenvolvido por Rumbaugh *et al.* (1991). A partir de então, este modelo conceitual passou por algumas melhorias, principalmente a partir do trabalho desenvolvido por Borges e Laender (1997), o qual trouxe primitivas geográficas para a abstração. É o modelo utilizado pela Infraestrutura Nacional de Dados Espaciais (INDE), segundo a Comissão Nacional de Cartografia (CONCAR, 2010). O modelo OMT-G abarca termos e primitivas para modelar geometricamente e topologicamente os dados espaciais, ou seja, compreende diferentes estruturas topológicas, várias visualizações de objetos e diversas relações espaciais (BORGES, DAVIS JR.; LAENDER, 2005). É usado normalmente para modelar dados relativos aos cadastros urbanos, por permitir os relacionamentos topológicos.

O OMT-G tem sua origem nos princípios da *Unified Modeling Language* (UML), criada para padronizar a notação gráfica da técnica OMT, que engloba todos os requisitos de modelagem para qualquer aplicação em todas as etapas de implementação do *software*. Portanto, a UML passou a ser usada por toda a comunidade de projetistas e desenvolvedores de sistemas que fazem uso do formalismo OO (GONÇALVES, 2018).

3 FERRAMENTAS DE BUSCA SISTEMÁTICA

A busca sistemática é uma ferramenta de investigação científica que tem como principal objetivo a eliminação de vieses provocados por pesquisas muito exaustivas na literatura por meio do planejamento de buscas em base de dados científicos e a organização dos principais resultados em um portfólio bibliográfico. O método *Systematic Search Flow* (SSF) classifica a revisão de literatura em três tipos: narrativa, sistemática e integrativa (FERENHOF e FERNANDES, 2016), sendo que a primeira delas pode ser denominada também de busca exploratória e é um método tradicional que não segue uma sistematização do processo de busca e não se preocupa com o volume e natureza das fontes de informação.

A segunda, revisão sistemática, adota um método replicável, respeita critérios científicos e transparentes de busca e seleção de fontes bibliográficas (COOK *et al.*, 1997) e permite que a análise seja realizada por mais de um pesquisador, que discutem e definem critérios de inclusão e exclusão dos resultados de busca (ERCOLE *et al.*, 2014).

O terceiro tipo de revisão de literatura, revisão integrativa, é um método de revisão bibliográfica utilizado para organizar e sintetizar os principais resultados de uma pesquisa científica a partir de um tema delimitado, de forma ordenada sistemática e abrangente, que tem por função promover um aprofundamento maior do tema investigado (ROMAN e FRIEDLANDER, 1998; ERCOLE *et al.*, 2014), podendo ser realizada por apenas um pesquisador.

4 PROCEDIMENTOS METODOLÓGICOS DA PESQUISA

Como procedimentos metodológicos para realizar a pesquisa exploratória foi adotada a revisão sistemática e integrativa de literatura. Para definir as etapas e fases da revisão foram utilizados os princípios de estratégia de pesquisa propostos por Ferenhof e Fernandes (2016) no método SSF, estruturados em 4 etapas, conforme exposição do Quadro 1.

Quadro 1- Princípios de estratégia de pesquisa do método SSF.

ETAPA	DEFINIÇÕES
1 - Protocolo de pesquisa	Fase 1 - Definição da estratégia de busca: identificação da necessidade de revisão sistemática, avaliação e delimitação do campo da temática, atualidade, relevância, extensão temporal da investigação científica, limitações e motivações
	Fase 2 - Consulta em base de dados
	Fase 3 - Gestão de documentos
	Fase 4 - Padronização e seleção dos documentos
	Fase 5 - Composição do portfólio de documentos
2 - Análise de documentos	Caracterização da natureza dos resultados da busca sistemática, onde são descritas informações sobre o número de publicações encontradas, distribuição geográfica das mesmas, maior número de citações, obtenção de dados bibliométricos, a fim de propor sugestões para trabalhos futuros
3 - Síntese	As informações relativas à busca sistemática são condensadas para formar uma síntese e todos os artigos devem passar pelos critérios de inclusão e análise crítica. É importante que nessa etapa seja avaliada as evidências práticas das principais soluções encontradas e seus desdobramentos (pontos fracos e pontos fortes)
4 - Redação de resultados	Redação dos principais resultados obtidos por meio da revisão integrativa

Fonte: Adaptado de Ferenhof e Fernandes (2016).

Como a Etapa 1 “Protocolo da pesquisa” é formada por cinco fases interdependentes, foram estabelecidos critérios para realizar as buscas, sendo que no Quadro 2 estão expostos os critérios definidos na Fase 1 “Definição da estratégia de busca”.

Quadro 2- Critérios definidos na estratégia de busca.

Estratégia de busca	Critério definido
Base de dados: utilização de bases de dados relevantes e internacionais	“Scopus”, “Web of Science”, “SciELO” e “Science Direct”
Palavras-chave da busca: utilização de quatro palavras-chave no escopo da temática	“Modelagem Conceitual”, “Modelagem Geográfica”, “Diagrama de Classes”, “Dicionário de Dados”
Filtros a serem aplicados: idioma, área temática, tipo de publicação, intervalo temporal de busca	“Inglês, Espanhol e Português”, “Engenharia e Ciências Exatas e da Terra”, “Artigos científicos publicados em periódicos com revisão por pares”, “2001 a 2022”
Estratégias de busca: pelas palavras-chave definidas nos artigos buscados nas bases de dados	“Título”, “Palavras-chave”, “Resumo”

A partir da definição exposta no Quadro 2, seguiu-se para a Fase 2 “Consulta em base de dados” da Etapa 1. Desta forma foram então realizadas as buscas na base de dados através de combinações de expressões alfanuméricas que direcionaram a pesquisa através da aplicação dos critérios estabelecidos. Além dos critérios qualitativos, foi também utilizado o operador lógico do tipo “OR”, que ao ser utilizado como critério em uma busca apresenta como dados de saída resultados que contenham, pelo menos um, entre dois ou mais termos definidos.

A Fase 3 “Gestão de documentos” e a Fase 4 “Padronização e seleção dos documentos” da Etapa 1 foram realizadas de forma simultânea através de processo de seleção estabelecido,

primeiramente, pela leitura superficial das seções resumo dos artigos selecionados na Fase 2. Com a finalidade de se obter o conhecimento das principais temáticas e aplicações em modelagem desenvolvidas internacionalmente, todos os artigos que não estavam relacionados à gestão territorial nesta fase foram descartados.

Os artigos resultantes foram selecionados para compor o portfólio de documentos da Fase 5 “Composição do portfólio de documentos”, cuja etapa foi realizada em simultâneo com a “Análise de documentos” da Etapa 2. Nesta fase todos artigos foram lidos de forma integral e foi desenvolvido um “Painel de indicadores”, utilizado como ferramenta de análise dos dados. Para tanto os dados foram organizados e exportados para uma planilha eletrônica do tipo *Excel*, de forma a organizar os resultados em: tipo de modelagem, “conceitual, lógica, física, geográfica, outros”; país de publicação; palavras-chave; ano e idioma da publicação.

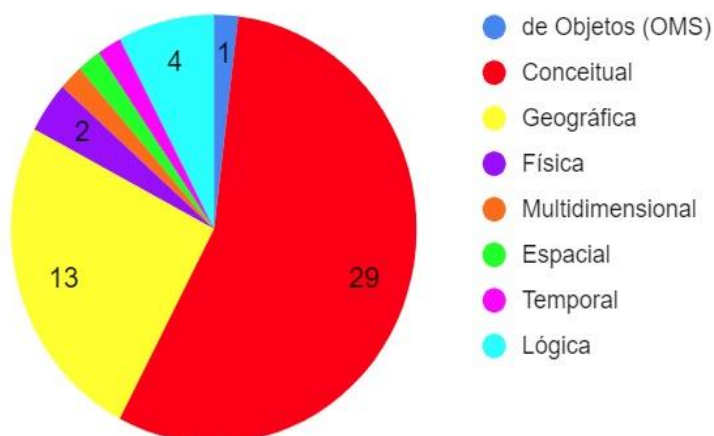
Após aplicação das sub etapas, como forma de sintetizar os resultados da Etapa 3 “Síntese” foi utilizado o *software Excel* para produzir os gráficos de resultados por meio de diferentes cruzamentos de colunas. Assim foi possível perceber os pontos fracos e fortes dos trabalhos pesquisados, bem como tecer as análises e escrita de resultados da Etapa 4 “Redação de resultados”.

5 RESULTADOS E ANÁLISES

A partir da Etapa 1, do processo de revisão descrito no capítulo 4, foram sendo gerados os resultados da pesquisa onde foram encontrados 141 artigos científicos nas bases de dados que atendiam aos critérios especificados. Através da filtragem descrita foram descartados 104 artigos. Assim, a seleção final resultou em 37 artigos que abrangeram a modelagem conceitual e geográfica, sendo 26 no idioma inglês, oito em espanhol e somente três em português.

Na Etapa 2 de “Análise de documentos” foi verificado que os 37 artigos filtrados apresentaram estudos de caso envolvendo oito aplicações distintas na área de modelagem. Ver Figura 1, onde pode ser percebido que os tipos de modelagens mais utilizadas foram as que abrangeram o universo conceitual, com 29, seguida da geográfica, com 13. Os demais tipos catalogados: lógica, física, entre outros, tiveram participação incipiente. Dentre essas abrangências da modelagem, as temáticas mais citadas foram meio ambiente e recursos hídricos, com 13 trabalhos, água/energia e gestão da informação, com cinco cada, agricultura e saúde com três, cada uma, banco de dados e mapeamento com duas cada e, diversas outras aplicações como: engenharia reversa, BIM, 3D, sustentabilidade, monitoramento, cadastro, com uma cada uma.

Figura 1 – Aplicações na área de modelagem.



Se por um lado, os tipos de modelagens e as temáticas não se evidenciaram, para melhor exposição dos resultados foi gerada nuvem de palavras com os termos mais presentes nas palavras-chaves dos artigos pesquisados, ou seja, os termos “modelagem” e “diagrama” aparecem em destaque, seguidos de outros termos corriqueiros como: modelos, representação, procedimentos, objetos, dados, saúde, sistema, uml. Ver Figura 2. Termos como “diagrama de classes”, embora não apareça na nuvem, foi citado três vezes nas palavras chave dos artigos, assim como o termo “dicionário de dados”.

Há que se destacar que 25 artigos não especificaram áreas de aplicação e desenvolveram trabalhos somente conceituais de aplicação universal, porém o termo “modelagem”, sem o acréscimo “conceitual”, ou até mesmo o termo “conceitual” sozinho, aparece pequeno na Figura 2 porque surgiu somente nas palavras-chave de dez artigos, não corroborando, entretanto, com o grande número de artigos da temática de modelagem conceitual identificados nas aplicações (Figura 1).

Figura 2 – Nuvem de palavras mais presentes nas palavras-chave dos artigos pesquisados.

temática conceitual abrangeram os termos “engenharia de produto, cadeia de produção e ciclo de vida, modelagem de ambiente interno, revisão teórica e conceitual e gestão da informação”. Os artigos que aplicaram a modelagem em áreas de estudo abrangeram os termos “gestão da informação, saúde pública, energias renováveis e meio ambiente”.

Figura 3 – Número de publicações na área de modelagem por ano.

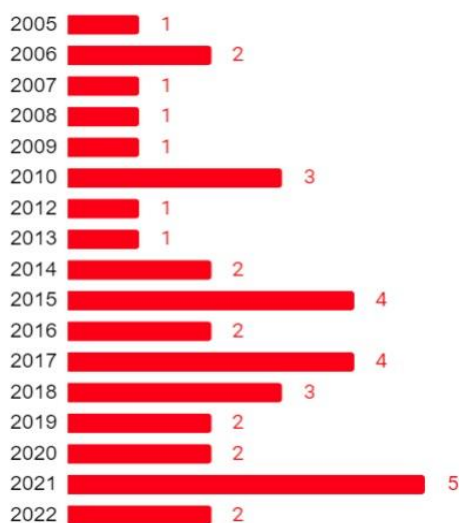


Figura 4 – Distribuição geográfica e quantitativa das publicações.

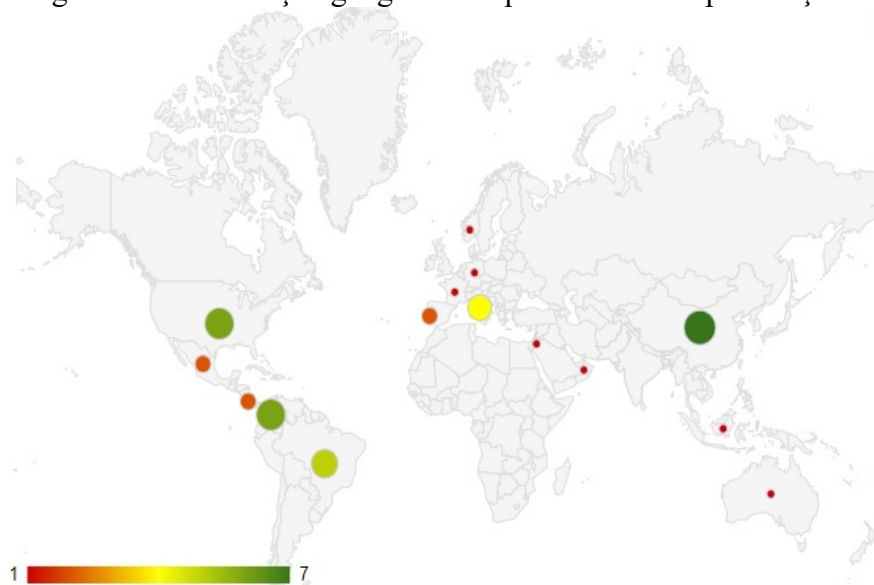


Figura 5 – Gráfico de distribuição quantitativa (valores absolutos) por país de publicação.



6 CONSIDERAÇÕES FINAIS

Através dos resultados expostos foi possível concluir que as produções científicas mais numerosas foram as que tiveram por finalidade os temas que envolvem justamente a modelagem conceitual e a geográfica. Os termos “diagrama de classes” e “dicionário de dados”, comumente utilizados nas publicações que abarcam modelagem quase não aparecem nas buscas, evidenciando uma necessidade de aumento de pesquisas nessas áreas. As temáticas mais citadas nas publicações foram as que envolveram meio ambiente, recursos hídricos, água, energia e saúde, áreas que necessitam de informações especializadas, fundamentais para realizar a gestão de qualquer território. O ano com maior número de publicações foi 2021 com cinco artigos e o país que mais publicou na área foi a China, com sete trabalhos.

A pesquisa ainda apresentou uma visão geral do estado da arte da modelagem através do exposto, donde pode ser concluído que não estão sendo publicados trabalhos com diferentes temáticas, como as cadastrais, desde o urbano, passando pelo rural, até mesmo o Cadastro 3D, sendo este um tema com grande oportunidade de pesquisa de campo para novos estudos. Muitos outros resultados foram produzidos com a pesquisa, mas, por falta de espaço, neste artigo, não puderam ser apresentados.

Referências

BÉDARD, Y.; PAQUETTE, F. Extending entity/relationship formalism for spatial information systems. In: **in Proceedings of AUTOCARTO**. 1989.

BERTINI, G. C.; CESAR NETO, J. Uma Modelagem Orientada a Objeto para o Mapa Urbano Básico de Belo Horizonte (MUB/BH). **Informática Pública**, v. 6, 2004. p. 33-51.

BORGES, K. A. V.; DAVIS JÚNIOR, C. A.; LAENDER, A. H. F. Modelagem conceitual de dados geográficos. In: CASANOVA, *et al.* Banco de Dados Geográfico. **MundoGEO**: Curitiba, 2005. Cap. 1, p. 83-136.

BORGES, K. A. V.; LAENDER, A. H. F. Geo-OMT—uma extensão do modelo OMT para aplicações geográficas. *In: IV Simpósio Brasileiro de Geoprocessamento*. 1997. p. 252-262.

CONCAR - Comissão Nacional de Cartografia. **Plano de Ação para Implantação da Infraestrutura Nacional de Dados Espaciais**. 2010.

COOK, D. J.; MULROW, C. D.; HAYNES, R. B. Systematic Reviews: Synthesis of Best Evidence for Clinical Decisions, *Annals of Internal Medicine*, 126 (5), 1997, p.: 376–380.

ERCOLE, F. F.; MELO, L. S.; ALCOFORADO, C. L. G. C. Revisão integrativa versus revisão sistemática, *Revista Mineira de Enfermagem*, v. 18, n. 1, 2014. p.: 9–12.

FERENHOF, H. A.; FERNANDES, R. F. Desmistificando a revisão de literatura como base para redação científica: Método SSF. *Revista ACB*, v. 21, n. 3, 2016. p. 550-563.

GONÇALVES, R. P. **Modelagem conceitual de bancos de dados geográficos para cadastro técnico multifinalitário em municípios de pequeno e médio porte**. Dissertação (Mestrado em Engenharia Civil) - Programa de Pós-Graduação em Engenharia Civil, Universidade Federal de Viçosa: Viçosa, 2018.

LAGO, D. Modelagem de banco de dados geográfico para subsídio a gestão integrada de recursos hídricos. *Revista de Ciências Exatas e Tecnologia*, v.1, 2006. p. 79-90.

LISBOA FILHO, J.; IOCHPE, C. Modelagem de Bancos de Dados Geográficos. *Anais do XX Congresso Brasileiro de Cartografia*, 2001.

LISBOA FILHO, J.; IOCHPE, C. Um estudo sobre modelos conceituais de dados para projeto de bancos de dados geográficos. *Revista IP-Informática Pública*, v. 1, n. 2, p. 37-90, 1999.

LONGLEY, P.A. *et al. Sistemas e Ciência da Informação Geográfica*. 3. ed. Brasil: Bookman Companhia Editora, 2013. 560 p.

QUEIROZ, G. R.; FERREIRA, K. R. **Tutorial sobre Bancos de Dados Geográficos**. GeoBrasil, 2006. INPE. Disponível em: <www.dpi.inpe.br/TutorialBdGeo_GeoBrasil2006.pdf> Acesso em: 15 de junho de 2022.

ROMAN, A. R.; FRIEDLANDER, M. R. Revisão integrativa de pesquisa aplicada à enfermagem. *Cogitare Enfermagem*, v. 3, n. 2, 1998.

RUMBAUGH, J. *et al. Object-oriented modeling and design*. Englewood Cliffs, NJ: Prentice-hall, 1991.