



II Simpósio Regional de Agrimensura e Cartografia

“Ampliando os horizontes e discutindo o futuro da geoinformação e do cadastro territorial do Brasil”

Universidade Federal de Uberlândia – UFU / Campus Monte Carmelo
22 a 24 de novembro de 2021



CONVERSÃO DE DADOS EM FORMATO CAD PARA BANCO DE DADOS GEOGRÁFICOS CONFORME ESPECIFICAÇÃO TÉCNICA PARA ESTRUTURAÇÃO DE DADOS GEOESPACIAIS VETORIAIS DA SECRETARIA DO PATRIMÔNIO DA UNIÃO – ET-EDGV/SPU: COMPARAÇÃO DE PROCESSOS

Ismael Fiuza Ramos¹; Filipe Soares Da Silva²; Elmo Bispo De Oliveira³;
Fabiola Andrade Souza⁴; Fabiano Peixoto Freiman⁵;
Jéssica Carvalho Vianna Có⁶

RESUMO

Este trabalho compara dois processos utilizados na definição da metodologia de conversão de dados vetoriais a partir de formato CAD para um banco de dados geográficos conforme a Especificação Técnica para Estruturação de Dados Geoespaciais Vetoriais da Secretaria do Patrimônio da União (ET-EDGV/SPU). O primeiro processo considerado “manual” utiliza diferentes ferramentas de softwares CAD e SIG no processo de conversão, sendo dependente das ações executadas pelo técnico responsável. O segundo, utiliza uma ferramenta de *Extract-Transform-Load* (ETL) no intuito de automatizar a conversão e tratamento dos dados. A partir da aplicação de ambos na conversão completa da carta denominada Trecho Forte São Diogo/Farol da Barra, localizada no município de Salvador/Bahia, conclui-se que a utilização de uma ferramenta ETL é bastante útil ao processo, especialmente no controle do conjunto de dados e na redução do tempo, embora ainda haja necessidade de execução de tarefas “manuais” por parte do técnico.

Palavras-chave: ET-EDGV, CAD, Banco de dados Geográficos, ETL, Conversão vetorial.

1 INTRODUÇÃO

Atualmente, o portfólio cartográfico da Secretaria de Coordenação e Governança do Patrimônio da União (SPU) é armazenado em meios analógico (plantas, mapas e cartas cadastrais impressos) e digital (cartas em formato *Computer Aided Design* - CAD) de forma descentralizada em suas diversas superintendências estaduais, fator que limita a disseminação dos dados cartográficos à sociedade. Para solucionar esta limitação, a SPU criou a Infraestrutura de Dados Espaciais da SPU (IDE/SPU) de forma a aderir às normas e padrões estabelecidos pela Infraestrutura Nacional de Dados Espaciais (INDE), garantindo compartilhamento, interoperabilidade e disseminação de dados, conforme definido pela

¹ Universidade Federal da Bahia, ismael.ramos@economia.gov.br

² Universidade Federal da Bahia, filipe.soares@economia.gov.br

³ Universidade Federal da Bahia, elmo.oliveira@economia.gov.br

⁴ Universidade Federal da Bahia, fabiola.souza@economia.gov.br

⁵ Universidade Federal da Bahia, fabiano.freiman@economia.gov.br

⁶ Universidade Federal da Bahia, jessica.co@economia.gov.br



II Simpósio Regional de Agrimensura e Cartografia

“Ampliando os horizontes e discutindo o futuro da geoinformação e do cadastro territorial do Brasil”

*Universidade Federal de Uberlândia – UFU / Campus Monte Carmelo
22 a 24 de novembro de 2021*



Comissão Nacional de Cartografia (CONCAR), nas Especificações Técnicas para Estruturação de Dados Geoespaciais Vetoriais 3.0 (CONCAR, 2017).

Para adequação às normas estabelecidas pela INDE, a SPU desenvolveu Especificações Técnicas para Estruturação (ET-EDGV/SPU) e Aquisição (ET-ADGV/SPU) de Dados Geoespaciais Vetoriais, ambas na versão 3.0, que estabelecem o padrão das estruturas de dados vetoriais (pequenas e grandes escalas) e temáticos do patrimônio imobiliário público federal, e almeja futuramente viabilizar o compartilhamento e a interoperabilidade de dados, bem como a racionalização de recursos entre os produtores e usuários de informações cartográficas (SPU, 2020).

Neste contexto, o primeiro desafio consiste na alimentação da IDE/SPU com dados derivados de produtos analógicos e digitais (em formato CAD), os quais não refletem adequadamente as normas e especificações técnicas para serem disponibilizados diretamente no Banco de Dados Geográficos (BDG) desenvolvido pela instituição, fator que insere a necessidade de convertê-los, padronizá-los e adequá-los ao Sistema de Informação Geográfica (SIG).

No que tange os dados digitais em formato CAD, um dos problemas é trabalhar com um sistema de coordenadas métricas, sendo limitado quando analisadas as necessidades de um SIG, como sistemas de referência, sejam eles projetados ou não, assegurando o posicionamento dos objetos geográficos em uma única base evitando problemas como os de diferença de fuso, quando utilizada projeção UTM. Além disso, um aspecto importante é a análise da consistência topológica (vizinhança, proximidade, pertinência) dos dados e integridade espacial das feições que compõem a base cartográfica conforme exigências da ET-EDGV/SPU, em que dados CAD não são prontamente utilizáveis e requer a sua conversão para posterior carregamento de dados no BDG. Sem o armazenamento da topologia possibilitado pelo SIG, não seria possível o desenvolvimento de consultas a um BDG (CÂMARA & ORTIZ, 1998).

Para convertê-los é necessário aplicar um conjunto de ações encadeadas, entretanto, as bases de dados geoespaciais são extensas e complexas e estão associadas a processos de conversão repetitivos e morosos que, quando manuais, são suscetíveis a erros grosseiros e, em consequência, inserem retrabalhos. O método ETL (*Extract-Transform-Load*) foi a solução encontrada para auxiliar na conversão de dados CAD para SIG, de forma automatizada, minimizando tempo de trabalho, custos do processo e erros.

O ETL é um processo que extrai dados de um sistema de Bases de Dados para outro,

depois de serem processados e modificados. Divide-se em três fases - extração, transformações e carga – as quais têm o objetivo de extrair, integrar e transformar os dados oriundos de várias fontes de informação de acordo com as restrições do BDG ao qual os dados serão enfim carregados (FERREIRA *et.al.*, 2010; CORNEJO *et. al.*, 2014).

Diante do exposto, objetiva-se analisar as funcionalidades de uma aplicação ETL no processo de conversão automática de dados CAD diretamente para BDG, comparando com o processo manualmente executado, considerando os aspectos da ET-EDGV/SPU.

2 METODOLOGIA

Para realizar a comparação de dois processos distintos visando determinar uma metodologia única de conversão para todas as bases em formato CAD disponíveis, foi utilizada a carta do Trecho Forte São Diogo/Farol da Barra, localizada no município de Salvador-Bahia, originalmente em formato CAD (.dgn), numa escala de 1:2.000 (Figura 1).

Figura 1 - Carta do Trecho Forte São Diogo/Farol da Barra – CAD sobre ortofoto.



Fonte: Os autores (2021). Ortofoto PMS/SEFAZ 2017.

Em geral, tanto este documento cartográfico quanto inúmeros outros similares possuem erros topológicos sistemáticos e assistemáticos, além de não terem padronização na representação dos objetos. A causa disso pode ser explicada pela ausência de um instrumento de especificação técnica para aquisição de dados geospaciais vetoriais (ADGV)

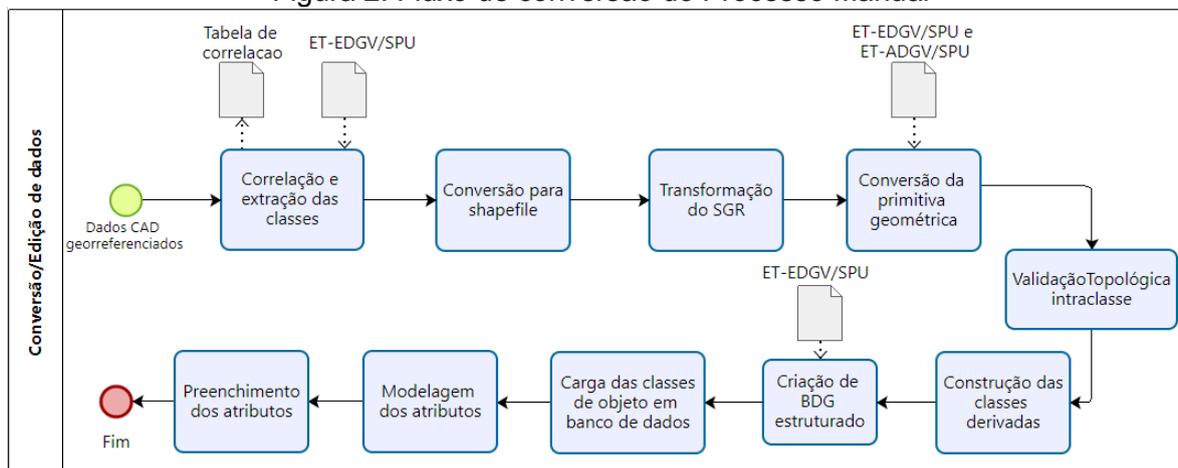
que norteasse a forma de adquiri-los e organizá-los naquela época.

Para estabelecimento de uma metodologia de conversão dos dados em formato CAD para o modelo ET-EDGV/SPU em BDG, dois processos foram testados, um com procedimentos manuais em software SIG (QGIS) e outro com processos mais automatizados, utilizando o software de ETL: *Feature Manipulation Engine* (FME).

2.1 Processo manual de conversão de dados

O processo manual, primeiro procedimento adotado, foi executado seguindo a ordem das etapas do fluxo metodológico ilustrado na Figura 2.

Figura 2: Fluxo de conversão do Processo Manual



Fonte: Os Autores (2020)

A execução das etapas foi realizada em ambiente CAD (AUTOCAD 2020 versão estudantil), SIG (QGIS 3.16) e PostgreSQL 12/PostGIS 3.0. A primeira etapa consiste na “Extração das classes de objetos” do arquivo vetorial em formato CAD por meio da identificação dos objetos e a correlação das camadas às classes da ET-EDGV/SPU.

Os dados vetoriais em formato CAD são constituídos por camadas para organizar e agrupar diferentes tipos de informações da carta por função ou finalidade. Desta forma, é possível ocultar feições que não se deseja visualizar, reduzindo a complexidade visual e facilitando a identificação dos objetos presentes na carta (Autodesk Help, 2018).

Cada camada pode ser denominada por um nome ou número. Quando são nomeados, geralmente, relacionam-se ao tema que foi mapeado, como: poste, edificação, ruas, vegetação. Contudo, quando são numerados, necessita-se utilizar outros elementos da carta ou artifícios (legenda, interpretação visual do objeto, sobreposição da camada a uma



II Simpósio Regional de Agrimensura e Cartografia

“Ampliando os horizontes e discutindo o futuro da geoinformação e do cadastro territorial do Brasil”

Universidade Federal de Uberlândia – UFU / Campus Monte Carmelo
22 a 24 de novembro de 2021



imagem e análise dos relacionamentos topológicos entre as feições que compõem o conjunto de dados) para identificar qual classe de objeto é representada.

Porém, entender a semântica dos objetos na camada analisada não é suficiente. É necessário correlacioná-los à classe de objeto a qual pertence, de acordo com a conceituação apresentada pela ET-EDGV/SPU. Para facilitar a identificação das classes, foi elaborada planilha que liste o número da camada em CAD e o correlacione com a sua respectiva classe de objeto representada, conforme exemplo ilustrado na Tabela 1 (a numeração da camada no CAD e a sua correspondência às classes de objeto será variável de acordo com cada coleção de produtos cartográficos).

Tabela 1 – Exemplo de correlação das camadas CAD às classes de objeto ET-EDGV/SPU

CLASSE DE OBJETO	CAMADA NO CAD
Trecho_rodoviario_L	4
Estrut_apoio_transporte_L	7
Trecho_arruamento_L	8
Edificação_A	10
Trecho_Drenagem	21
Canal_vala_L	22
Massa D'água	23
Terreno_sujeito_inundação	24
Terreno_exposto	26
Linha_costa	27
Curva_nivel	33 (mestra) e 34 (intermediaria)
Ponto cotado altimétrico	35
Trecho_Lpm	60
Trecho Ltm	62
Vegetação	25, 29, 30 e 31

Elaboração: Os autores (2020)

Depois de extraídas, as classes foram convertidas para o formato *shapefile*, transformadas do Sistema Geodésico de Referência (SGR) *South American Datum* (SAD 69) para o SGR para as Américas (SIRGAS 2000), e tiveram sua primitiva geométrica corrigida de acordo com a ET-EDGV/SPU por meio do QGIS - por exemplo, objetos na primitiva geométrica do tipo linha foram convertidos para polígono. É também nessa etapa que inconsistências que necessitam de edição foram corrigidas de forma manual da área dos terrenos cadastrais (originalmente como linhas parciais).

Após as correções, os dados passaram por Validações Topológicas Intraclasse e Entre Classes. Na primeira os dados são submetidos à verificação topológica entre eles mesmos,

respeitando regras básicas como: não deve sobrepor, não deve ter duplicados, não deve ter geometrias inválidas, não devem ter lacunas, entre outros. Já na segunda, a verificação é realizada entre diferentes classes, para que fiquem topologicamente consistidos, respeitando as regras de relacionamento entre os objetos e de ordenamento dos atributos estipuladas na ET-EDGV/SPU.

A ET-EDGV/SPU também prevê a criação de algumas classes temáticas que são derivadas de outras já mapeadas e topologicamente consistidas, como: Faixa_dominio, Limite_mar_territorial, Trecho_Terreno_Acrescido_Marinha, Trecho_terreno_marinha, Mar_Territorial. Esta ação é executada na sequência após as validações topológicas dos dados de referência, para permitir criação das camadas derivadas sem erros.

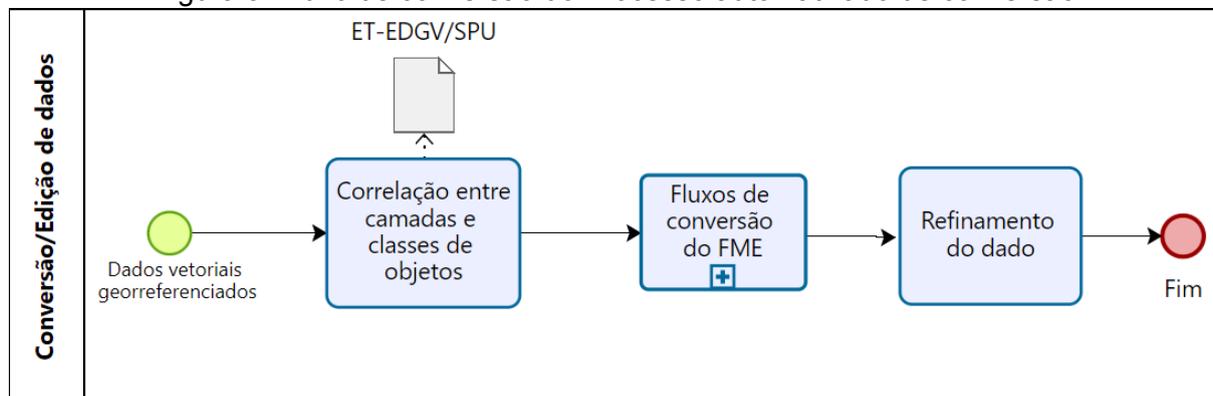
Após conversão, tratamento e criação de classes temáticas, foi criado BDG estruturado conforme ET-EDGV/SPU para carregamento dos dados em banco PostgreSQL/PostGIS, inicialmente em tabelas temporárias, para posterior preenchimento dos atributos através de conexão do banco com o QGIS. Depois de estruturados, os objetos das tabelas temporárias foram inseridos em tabelas definitivas do BDG.

Todas as etapas executadas neste processo de conversão foram consideradas “manuais”, uma vez que o responsável pela execução deveria acionar as ferramentas pertinentes a cada ação no *software*, muitas vezes com análise visual sobre os mesmos.

2.2 Processo automatizado de conversão de dados

O processo automatizado é composto pelas etapas apresentadas na Figura 2, entretanto inseridas no fluxo do FME, conforme Figura 3.

Figura 3: Fluxo de conversão do Processo automatizado de conversão



Elaboração: Os autores (2021)



II Simpósio Regional de Agrimensura e Cartografia

“Ampliando os horizontes e discutindo o futuro da geoinformação e do cadastro territorial do Brasil”

*Universidade Federal de Uberlândia – UFU / Campus Monte Carmelo
22 a 24 de novembro de 2021*



Ao comparar o fluxo dos dois processos de conversão apresentados percebe-se que a principal diferença é a utilização de ferramenta baseada no método ETL - software FME, uma aplicação capaz de se comunicar com diversos formatos e bancos de dados.

Os dados são adquiridos por rotinas de extração, que oferecem informações referentes aos originais, para depois passarem para o processo de transformação, correspondente aos procedimentos de limpeza, filtragem, integração e redefinição. Estas ações não incidem sobre os dados originais, e sim, no que vai ser carregado no BDG. No momento da carga ocorrem os mapeamentos sintáticos e semânticos entre os diferentes esquemas, levando em conta as restrições de integridade com a finalidade de padronizar os dados das diferentes fontes (FERREIRA *et. al.*, 2010).

As ferramentas de transformação desta aplicação automatizaram o processo de conversão através de fluxos que foram criados previamente para o processamento de dados espaciais em cadeia. Nos fluxos elaborados, diversas etapas do processo de conversão manual foram executadas de forma automatizada e encadeada, como: transformador de SGR, conversor de primitiva geométrica, validações topológicas pertinentes e edição dos dados. Outras ferramentas específicas do FME também foram utilizadas, como os filtros por tipo de geometria ou por atributo (para separar e selecionar as geometrias desejadas ou não previstas na ET-EDGV/SPU).

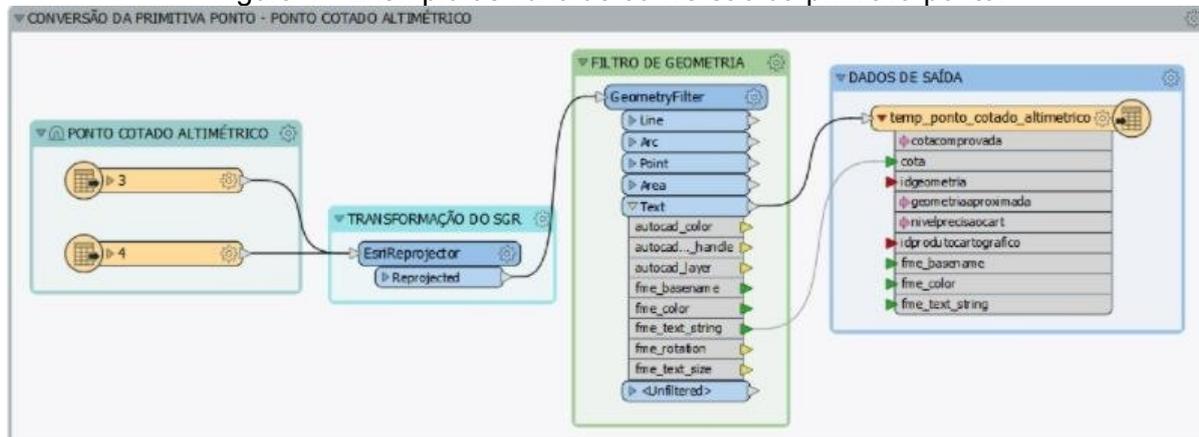
A escolha de quais ferramentas inserir nos fluxos variou conforme a necessidade de ajuste em cada dado, haja vista a ocorrência de erros assistemáticos nas cartas, e em virtude das diferentes primitivas geométricas. Por conta do tratamento diferenciado exigido por cada primitiva, foram criados quatro fluxos para conversão no intuito de recomendar as ferramentas para cada caso – ponto, linha, área e texto (convertido para primitiva do tipo ponto) - como no exemplo da Figura 4.

Como o processo de conversão é realizado por classe de objeto, primeiramente, correlacionou-se as camadas do CAD e as classes da ET-EDGV/SPU, materializada como uma planilha semelhante à da Tabela 1. Posteriormente, as camadas foram inseridas no fluxo de ferramentas por meio das rotinas de extração para serem filtradas, tratadas, editadas e convertidas. O próximo passo foi a carga em BDG já estruturado conforme classe de objetos, porém como uma tabela temporária (`temp_nome_da_classe_de_objeto`).

Com os dados implementados em banco, as próximas etapas consistiram no refinamento dos mesmos, por meio de ferramentas do QGIS, com o objetivo de realizar as etapas que não foram possíveis de serem executadas de forma satisfatória nos fluxos do

FME. Por isso, após o FME, as classes passaram por etapas semelhantes ao processo de conversão manual para, por fim, serem inseridas em tabelas definitivas do BDG.

Figura 4 - Exemplo de fluxo de conversão da primitiva ponto.



Fonte: Os Autores (2021)

3 RESULTADOS E DISCUSSÕES

A fim de possibilitar a comparação entre os dois processos estabelecidos, a carta Trecho Forte São Diogo/Farol da Barra passou por ambos para conversão. Com isso, foi possível perceber que, ao analisar estes processos, a principal característica que os diferencia é a utilização do método ETL para automatização de conversão.

Contudo, esta diferença altera todo o caminho de edição e tratamento dos dados. No processo manual, os dados são convertidos do formato CAD para *shapefile* e, posteriormente, inseridos em BDG. No processo automatizado, os dados são convertidos diretamente do CAD, passando pelo fluxo FME, para o BDG. Desta forma, evita-se perda de informações ao garantir maior controle do dado, por exemplo: quantificação do número de feições no fluxo, minimização de inconsistências, reconhecimento de diversos tipos de primitiva geométrica. Obtém-se, desta maneira, segurança e integridade pós conversão pelo FME ao tratá-lo por meio do QGIS conectado ao BDG ao invés da utilização de arquivo *shapefile*, que é mais vulnerável e limitado, conforme argumentação de Cepicky (2017).

Outra diferença notada foi a otimização do tempo necessário para a execução de toda a atividade ao utilizar o processo automatizado pelo método ETL. No primeiro processo, grande parte das tarefas são realizadas a partir de verificação visual e de forma manual, como: realocação de objetos na camada correspondente, retirada de objetos não previstos pela ET-EDGV/SPU, conversão da primitiva geométrica do tipo arco (reconhecida no QGIS como pequenos segmentos de reta) para linha, que delongam muito tempo, e muitas vezes



II Simpósio Regional de Agrimensura e Cartografia

“Ampliando os horizontes e discutindo o futuro da geoinformação e do cadastro territorial do Brasil”

*Universidade Federal de Uberlândia – UFU / Campus Monte Carmelo
22 a 24 de novembro de 2021*



retrabalho por conta de falhas humanas. Por meio do fluxo automatizado, utilizou-se filtros em atributos do próprio arquivo em CAD que permitem a seleção e execução destes objetos em lote.

Ademais, o processo com uso de uma ferramenta ETL facilita a sistematização das funções a serem utilizadas na conversão. Embora os fluxos elaborados no FME não sejam rígidos e não precisem ser seguidos fidedignamente por conta de erros assistemáticos, ao menos, a sequência das atividades pelas quais o dado passa no fluxo permitem padronizar o máximo possível a conversão das diferentes classes de objetos e suas respectivas primitivas geométricas. A sistematização das etapas pelo processo manual se torna uma tarefa mais difícil em virtude das constantes mudanças de versão do QGIS que alteram os algoritmos de plug-ins e ferramentas, ou por vezes, não habilitam mais a utilização destes.

4 CONCLUSÃO

A atividade de conversão de cartas da SPU mostrou-se desafiadora. A grande quantidade de erros sistemáticos e assistemáticos encontrados e a falta de padronização dos produtos cartográficos trabalhados demonstram a importância da existência de um instrumento norteador que sistematize a produção cartográfica. O desconhecimento quanto à forma como os dados foram adquiridos nos mapeamentos originais (inexistência de metadados ou indisponibilidade de documentos técnicos que detalhassem o processo de produção) e a pouca experiência técnica da equipe em lidar com dados inconsistentes ocasionaram a execução de retrabalhos provocando grande tempo de execução na conversão dos produtos. Contudo, este retrabalho permitiu construir conhecimento empírico necessário para a consolidação de fluxos metodológicos mais assertivos que otimizassem o tempo de execução, especialmente de maneira automatizada.

O processo manual mostrou-se, na fase inicial da atividade, como uma solução viável para converter os dados, principalmente pela realização de testes com inúmeras ferramentas e plug-ins do QGIS, a fim de solucionar as inconsistências de maneira mais célere e evitar os trabalhos manuais. Contudo, embora ainda não solucione todos os problemas e execute todos os processos necessários, a automatização da conversão através do método ETL para o processamento de dados espaciais em cadeia, conforme demonstrado no processo automatizado, mostrou-se eficiente como ferramenta de otimização do tempo de execução e controle da atividade de conversão dos produtos cartográficos em CAD para o BDG padronizado conforme ET-EDGV SPU.



II Simpósio Regional de Agrimensura e Cartografia

“Ampliando os horizontes e discutindo o futuro da geoinformação e do cadastro territorial do Brasil”

*Universidade Federal de Uberlândia – UFU / Campus Monte Carmelo
22 a 24 de novembro de 2021*



Além disso, a automatização das tarefas implicou na redução e acúmulo de erros grosseiros, uma vez que o responsável pela conversão não executou manualmente a maioria das ferramentas, voltando-se apenas à definição das funções a serem utilizadas no fluxo.

Ressalva-se, ainda, que apesar de ter sido utilizado o software FME para execução do processo automatizado, esta não é a única ferramenta de ETL no mercado que permita a automatização de fluxos de trabalho, cabendo análises futuras de ferramentas similares.

REFERÊNCIAS

Autodesk Help; (2018). Sobre Camadas. <<https://knowledge.autodesk.com/pt-br/support/autocad/learn-explore/caas/CloudHelp/cloudhelp/2019/PTB/AutoCAD-Core/files/GUID-6B3E3B5D-3AE2-4162-A5FE-CFE42AB0743B-htm.html>>. Acesso em 20 ago 2021

Câmara, G., & Ortiz, M. J; (1998). Sistemas de informação geográfica para aplicações ambientais e cadastrais: uma visão geral. In Congresso Brasileiro de Engenharia Agrícola (Vol. 27, pp. 59-82). Sn.

CEPICKY, J.; (2018). Switch from Shapefile. Creative Commons Attribution-ShareAlike 4.0 internacional License. <<http://switchfromshapefile.org/>>. Acesso em 20 ago 2021

CONCAR. Comissão Nacional de Cartografia. Especificações Técnicas para Estruturação de dados Geoespaciais Vetoriais (ET-EDGV 3.0). Brasília: CONCAR, 2017.

CORNEJO, R.; NAVARRETE, M.; VALDIVIA, R.; AROCA, P.; & ARACENA, S. (2014). Desarrollo de una base de datos integrada de Censo y encuesta mediante el uso de elementos de inteligencia de negocios y SIG. Ingeniare. Revista chilena de ingeniería, 22(2), 205-217.

FERREIRA, J.; MIRANDA, M.; ABELHA, A.; & MACHADO, J.; (2010). O processo ETL em sistemas data warehouse. In INForum (pp. 757-765).

SPU. Secretaria do Patrimônio da União. Especificação Técnica para Aquisição de Dados Geoespaciais Vetoriais do Patrimônio Imobiliário Público Federal (ET-ADGV/PIPB 3.0). Brasília: SPU, 2021

SPU. Secretaria do Patrimônio da União. Especificação Técnica para Estruturação de dados Geoespaciais Vetoriais do Patrimônio Imobiliário Público Federal (ET-EDGV/PIPB 3.0). Brasília: SPU, 2021