

Pauliceia 2.0: Enriquecendo as Humanidades Digitais com Geocodificação e Informação Geográfica Voluntária

Pauliceia 2.0: Enriching the Digital Humanities with Geocoding and Voluntary Geographic Information

Karla Donato Fook¹, Daniela Leal Musa², Luis Antonio Coelho Ferla³, Nandamudi Vijaykumar^{2,4}, Karine Reis Ferreira⁴, Gilberto Ribeiro de Queiroz⁴, Cristiane Regina Miyasaka³, Fernando Atique³, Jeffrey Lesser⁵, Thomas D. Rogers⁵, Andrew G. Britt⁵, Rafael Laguardia⁵, Rodrigo Monteiro Mariano⁴, Ana Maria Alves Barbour, Orlando Guarnier Farias³, Monaliza Caetano dos Santos³, Gabriel Sansigolo⁴, Janaina Yamamoto Santos⁶, Priscila Machado Meireles, Wania Mazzarello Gonçalves, Cintia Almeida³, Ester Dantas Reis Nunes, Luanna Nascimento³, Vitória Martins Fontes da Silva, Bruno Pricinato, Denis Taveira, Carlos Alberto Ferreira de Noronha⁴

¹Instituto Tecnológico de Aeronáutica (ITA), São José dos Campos, SP, Brasil

²Unifesp/ICT - São José dos Campos, SP, Brasil

³Unifesp/EFLCH - Guarulhos, SP, Brasil

⁴Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos, SP, Brasil

⁵Emory University, Atlanta, GA, EUA

⁶Arquivo Público do Estado de São Paulo, SP, Brasil

Resumo

O Projeto Pauliceia 2.0 objetiva o mapeamento histórico da cidade de São Paulo no período 1870-1940. Para isso foi desenvolvida uma plataforma computacional que permite a inserção de dados espacializáveis de pesquisas sobre a história da cidade de São Paulo por parte de qualquer investigador interessado em fazê-lo. A plataforma se utiliza da área de Ciência da Computação, em particular, VGI (Volunteered Geographic Information) e Geocodificação que vão ao encontro dos anseios da comunidade de Humanidades Digitais. Este texto apresenta como o protocolo VGI foi definido e implementado; além do serviço Web que permite a geolocalização de um estabelecimento contido na plataforma.

Palavras-chave: SIG histórico. Mapeamento Colaborativo. Geocodificação. VGI. Serviços Web.

Abstract

The Pauliceia 2.0 Project has an objective to produce the historical mapping of São Paulo city from 1870 to 1940. A computational platform was developed. This platform allows the insertion of spatializable research data on the history of São Paulo city into its geographical database by any researcher interested in doing so. The platform uses Computer Science, VGI (Volunteered Geographic Information) and Geocoding that meet the needs and specificities of the Digital Humanities community. This paper presents how the VGI

protocol was designed and implemented. It also describes the web service that allows the geolocation of a feature contained in the platform.

Keywords: *historical GIS. Collaborative Mapping. Geocoding. VGI. Web Services.*

1 Introdução

Pauliceia 2.0 é um projeto que objetiva o mapeamento histórico da cidade de São Paulo no período 1870-1940. Este período sofreu transformações radicais, pois a população passou de 31.385 a 1.326.261 habitantes [IBGE, 2011]. Em outras cidades ao redor do mundo também ocorreu algo semelhante, mas a cidade de São Paulo chegou próximo a um milhão de habitantes em poucas décadas.

No projeto foi desenvolvida uma plataforma computacional que permite a inserção de dados espacializáveis de pesquisas sobre a história da cidade de São Paulo por parte de qualquer investigador interessado em fazê-lo. A plataforma se utiliza da área de Ciência da Computação, em particular, VGI (*Volunteered Geographic Information*) e Geocodificação que vão ao encontro dos anseios da comunidade de Humanidades Digitais. O artigo discute como estas tecnologias de estado da arte contribuem para que historiadores possam contar com um meio de divulgar as suas pesquisas e compartilhar num ambiente aberto e de livre acesso sem nenhuma restrição. A abordagem VGI pode ser considerada *crowdsourcing*, onde os cidadãos voluntários são convidados a produzir e disseminar informações geográficas em sites disponibilizados. Estes voluntários são incentivados a contribuir com informações geográficas em sites como OpenStreetMap com base nos seus próprios conhecimentos. A grande vantagem é que esta iniciativa contribui e produz informações com baixo custo. Para garantir a qualidade dos dados foram definidos métricas, indicadores, abordagens e um protocolo VGI para guiar a coleta de dados geográficos a partir dos voluntários. O protocolo desenvolvido permite padronizar essas colaborações. Como a Plataforma é alimentada por dados espacializáveis, é necessário fazer um mapeamento do endereço de um estabelecimento (nome de rua e o número) para a sua geolocalização que consiste em endereço composto por latitude e longitude. Para realizar esta tarefa, foi desenvolvido um serviço web, que é um código computacional

RBHD, Rio de Janeiro, v. 1, n. 1, Dossiê Temático 3, p. 110-133, jan./jun., 2021

para ser executado por meio de chamada na Internet. Este artigo, na forma de uma Comunicação, discute: (i) como o protocolo VGI foi definido e implementado; e (ii) como foi desenvolvido e testado o serviço Web que permite a geolocalização de um estabelecimento alimentado na plataforma. É importante salientar que a Plataforma considera o fator tempo, ou seja, ela permite que a visualização de um endereço seja feita não só no seu espaço, mas também mostra o seu histórico durante um certo período de tempo de acordo com o estabelecido pelo projeto (1870-1940).

O termo georreferência se refere aos dados geográficos associados a uma localização espacial referenciando às coordenadas em termos de latitude e longitude [Carter, 1989]. As características de entidades geográficas são as propriedades específicas e as suas relações espaciais com outras entidades [Medeiros, 1994]. Isso ocorre naturalmente no dia a dia quando alguém precisa localizar um endereço de um certo destino. É natural recorrer ao Google Maps para não só determinar a localização, mas também o caminho até chegar a este destino. Geocodificação é nada mais do que converter atributos textuais como um endereço (Rua ou Avenida e número) para coordenadas geográficas em termos de latitude e longitude. A Plataforma fornece ferramentas para localizar e manipular dados históricos. Algumas destas ferramentas estão implementadas como serviços Web, de forma similar ao protocolo VGI. A ferramenta também permite medir a precisão da geolocalização retornada utilizando o conceito de distância Euclidiana.

O Pauliceia 2.0 conta com uma equipe multidisciplinar de cerca de 30 pessoas, sendo uma parceria entre a Universidade Federal de São Paulo (campi de Guarulhos e de São José dos Campos), o Instituto Nacional de Pesquisas Espaciais, o Arquivo Público do Estado de São Paulo e a Emory University. A primeira fase do projeto contou com financiamento do Programa eScience da Fapesp. A versão beta da Plataforma Pauliceia 2.0 está disponível para testes em www.pauliceia.dpi.inpe.br.

2 Conceitos e Trabalhos Relacionados

As humanidades digitais, por definição, têm um comportamento inerente de trabalhos colaborativos valorizando o compartilhamento, além de incentivar a livre circulação do

RBHD, Rio de Janeiro, v. 1, n. 1, Dossiê Temático 3, p. 110-133, jan./jun., 2021

conhecimento e o emprego de tecnologias [Spiro, 2012] e [Greenspan, 2016]. É comum hoje em dia ver trabalhos científicos sendo desenvolvidos em rede por meio de tecnologias digitais. Projetos como *Wikipedia*, *Open Street Maps* e outros são a prova disso. Um termo que surge, *crowdsourcing*, se refere a este novo cenário que é aberto e colaborativo e é facilitado pela Internet. Outro termo, *Volunteered Geographic Information* (VGI), também surgiu neste contexto. Boas revisões sobre estes termos e definições básicas podem ser encontradas em [See et al., 2016].

Crowdsourcing pode ser considerado um tipo de atividade online e participativa onde um cidadão, uma instituição, uma ONG, uma empresa, dentre outros pode propor a um grupo de indivíduos com diferentes níveis de conhecimento a realizarem uma tarefa. Esta proposta pode ser feita por meio de uma chamada aberta e flexível [Estellés-Arolas & González-Ladrón-De-Guevara, 2012]. O VGI, por sua vez, pode ser definido como um conjunto de ferramentas para criação, organização e disseminação de dados geográficos com a colaboração voluntária de indivíduos [Goodchild, 2007] e [Goodchild & Li, 2012] e considerado como uma versão de *crowdsourcing*. As contribuições para VGI possuem uma localização geográfica e uma descrição de atributos daquela localização [Goodchild & Li, 2012].

Os dados VGI englobam representações espaciais como pontos, linhas e polígonos. Os seus tipos são mapas, imagens e texto. Mapas correspondem às geometrias básicas (pontos, linhas e polígonos). Por exemplo, OpenStreetMap, Google Map Maker, Map Insight e Wikimapia são baseados em mapas. VGI baseadas em imagens se referem às fotos com alguma associação de uma referência geoespacial. Instagram, Flickr e Panoramio são alguns exemplos. No caso de VGI baseada em texto, são criadas implicitamente quando os voluntários fornecem informações em formato de texto com dispositivos como celulares ou notebooks [Senaratne et al., 2017]. A VGI pode não ser precisa. Isto porque os humanos expressam regiões geográficas e as suas relações de forma imprecisa. Este fato também tem a ver com qualidade e limitações de conhecimento espacial (Hollenstein & Purves, 2010). Para evitar estas inconveniências, é importante a adoção de indicadores e métricas de qualidade para as contribuições de VGI [Senaratne et al., 2017].

Em teoria, VGI pode obter padrões de qualidade, mas é necessário adotar protocolos robustos para evitar que erros se propaguem e impactem no seu reuso [Mooney et al., 2016]. Foram definidos alguns estágios para protocolo VGI: inicialização, coleta de dados, auto-avaliação e controle de qualidade, submissão de dados e *feedback* para comunidade.

Os mapas têm sido desenvolvidos por humanos até bem antes de começar a escrita. A cartografia teve a sua evolução desde os desenhos nas cavernas até a era digital. Sistemas de Informações Geográficas (SIGs) se tornaram populares e imprescindíveis para realizar análises de dados espaciais. É possível também customizar mapas online por meio de mapeamento por web. É muito comum nos celulares atuais poder se localizar um certo endereço e até com dicas de restaurantes, teatros, cinemas, etc. [Garrard, 2016]. Os mapas digitais se utilizam de características geográficas descritas como objetos geométricos, tais como, pontos, linhas e polígonos. Estes tipos de dados são conhecidos como dados vetoriais. Por exemplo, um mapa de um estado pode ter suas cidades representadas por pontos, estradas ligando as cidades como linhas, e macro-regiões como polígonos.

Além de geometria, os dados vetoriais podem ter atributos associados a eles. O atributo pode ser o comprimento de uma linha ou a área de um polígono. Mas, a grande vantagem é que os atributos podem ser informações mais detalhadas, como, nome do estado, sua população, etc. [Garrard, 2016]. Podem até associar informações mais complexas como imagens, objetos conhecidos como BLOBs (objetos binários grandes). Dados geográficos são representados por um conjunto de coordenadas de latitude e longitude. Geolocalização é um termo utilizado para se referir às coordenadas geográficas (latitude e longitude) de um dado endereço de um ponto na cidade, por exemplo. Geoprocessamento é uma operação realizada em um conjunto de dados geográficos, ou seja, é uma operação que manipula dados SIG. Geoprocessamento fornece ferramentas e métodos para facilitar a análise de dados espaciais [Zaidan, 2017]. Geocodificação é o processo de transformar dados textuais em informações geográficas, obter coordenadas de endereços textuais é um dos métodos de procedimento de geocodificação mais importantes [Câmara et al.2005]. Neste trabalho, abordamos a

geocodificação histórica usando a estrutura nome da rua, número do endereço e ano pesquisado.

Há projetos históricos com características similares ao projeto Pauliceia.

Digital Harlem é um portal web que faz parte de um projeto de pesquisa colaborativa sobre a vida cotidiana no Harlem entre 1915 e 1930 realizado por historiadores do Departamento de História da Universidade de Sydney, na Austrália. Com ele é possível pesquisar eventos e lugares, gerar mapas interativos, verificar os locais que a população frequentava, e os resultados podem ser mostrados em um mapa com várias camadas [Digital Harlem Blog, 2017]. Os dados gerados no portal são interpretados pelos historiadores.

A plataforma ATLMaps foi desenvolvida com a colaboração entre a Georgia State University e a Emory Universidade. Nessa plataforma é possível combinar mapas, visualizar dados geoespaciais e localizar pontos da cidade de Atlanta. Os usuários podem sobrepor mapas de diferentes períodos de tempo para comparação [White; Gilbert, 2016].

GeoHistoricalData¹ é uma plataforma para digitalização colaborativa de mapas históricos de Paris e da França do século XVIII ao século XX. O principal objetivo do projeto é estudar as evoluções territoriais em diferentes escalas e níveis nessa região e período.

O projeto Atlanta Explorer [Page et al., 2013] consiste em criar Banco de Dados Geográficos Históricos, Geocodificadores e modelos 3D da cidade de Atlanta, nos EUA, para período pós guerra civil até 1940.

Esses projetos serviram de inspiração e referência para o desenvolvimento do Pauliceia.

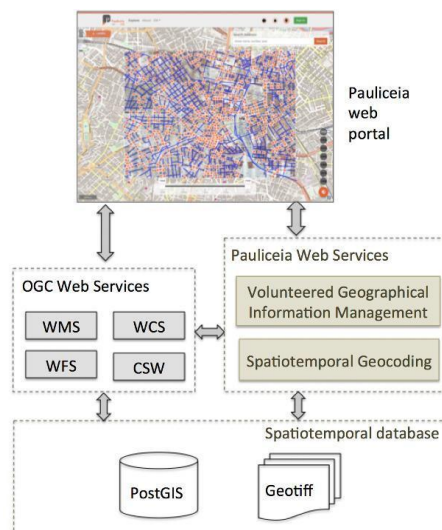
¹ <http://geohistoricaldata.org/>

3 Arquitetura da Plataforma

A plataforma Pauliceia 2.0 foi concebida e desenvolvida na Fase 1 do Projeto Pauliceia 2.0, estando atualmente disponível e plenamente acessível a qualquer interessado. Pode ser usada só para consulta ou para alimentar dados históricos por meio de camadas. Para alimentar dados, é necessário criar um cadastro simples sem exigência de qualquer qualificação ou titulação do usuário. É importante observar que qualquer pessoa pode ter acesso à Plataforma e visualizar livremente os conjuntos disponíveis de dados. No entanto, só aqueles que tiverem um registro, podem editar ou criar camadas com dados históricos. A arquitetura da Plataforma está ilustrada na Figura 1.

O topo da figura ilustra o acesso à Plataforma por meio de um navegador local de um usuário. A Plataforma reside num servidor hospedado em alguma instituição (INPE, Unifesp ou em outro local). A Plataforma é aberta (*open source*), online e orientada a serviço. A parte intermediária da figura mostra serviços Web que retornam alguma resposta quando há uma solicitação de um usuário. Neste sentido, o conceito de sistemas orientados a serviço atende muito bem às aplicações que necessitam de interoperabilidade. Os serviços Web à esquerda se referem aos serviços padrão especificados pelo OGC (Open Geospatial Forum) e os serviços desenvolvidos que seguirem a esta especificação, como é o caso do Projeto Pauliceia, podem ser visualizados por aplicativos como INDE, QGIS, TerraView e outros. Os serviços Web à direita são os desenvolvidos no contexto do projeto e são VGI (Volunteered Geographic Information) e Geocodificação espaço temporal [Ferreira et al., 2017]. Estes serão vistos nas próximas Seções. A última parte da Figura se refere aos repositórios para armazenar os mapas. Mapas vetoriais são armazenados no Banco de Dados PostGIS e os dados *raster* em formato Geotiff.

Figura 1: Arquitetura da Plataforma Pauliceia 2.0

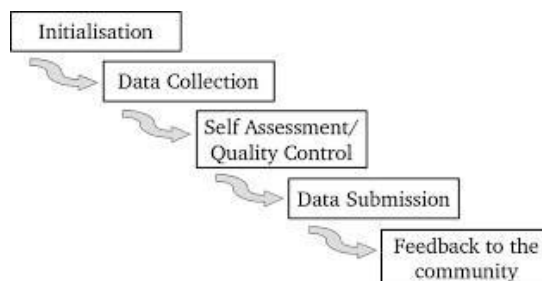


Arquitetura da Plataforma. Fonte: Ferreira et al., 2018

4 Informação Geográfica Voluntária (VGI) no Pauliceia 2.0

Como já mencionado, um protocolo com estágios para gerenciar projetos VGI foi proposto [Mooney et al., 2016]. Seu principal objetivo é melhorar a compreensão dos voluntários para que a qualidade de suas contribuições seja de ótimo nível ao produzir dados por eles. Os estágios se referem à coleta e gerenciamento de dados geográficos, controle de usuário, auto-análise, métricas de qualidade, e retornos (*feedback*) à comunidade. Os estágios serão sucintamente explicados denotando como eles foram adotados para a Plataforma [Mariano et al., 2018].

Figura 2: Protocolo VGI



Estágios Principais do Protocolo de VGI. Fonte: Mooney et al., 2016

4.1 Inicialização

A Plataforma Pauliceia 2.0 consiste de ferramentas para que voluntários possam incluir e editar objetos geométricos (pontos, linhas, polígonos) para captar características como ruas e edifícios a partir de mapas históricos em formato *raster*. Observe que a plataforma não permite editar ou criar objetos *raster*.

Dados produzidos por cidadãos voluntários podem possuir geometrias distintas para representar um mesmo aspecto ou *feature*. Para normalizar esta representação, a Plataforma emprega métodos que extraem uma geometria única que representa a opinião da maioria, já proposto em [Budig et al., 2016]. O acesso online à Plataforma é por meio de um navegador, depois de os voluntários devidamente se registrarem e aceitarem um termo de concordância.

4.2 Modelo de Dados

A Plataforma organiza os conjuntos de dados em camadas onde são agrupados aspectos (*features*) geográficos relacionados a um certo tópico ou tema. Cada aspecto possui suas propriedades espaciais e não-espaciais. No primeiro caso, a representação é por meio de pontos, linhas e outras geometrias. As propriedades não-espaciais se referem aos tipos como textos e números e podem até conter *links* para documentos ou mídia como fotos e vídeos desde que estejam armazenados em outros repositórios como Drives, YouTube ou Dropbox.

Uma camada contém um conjunto de aspectos e suas diversas versões ao longo do tempo. Estas mudanças são registradas junto com informações de quem realizou a modificação e quando. Em geral, as camadas são associadas a palavras-chave, como por exemplo, crimes, fábricas, saúde, eventos. As palavras-chave são importantes no projeto para facilitar a busca no repositório por camadas associadas a certos aspectos ou temas. As camadas são inseridas por usuários que são considerados proprietários desde que estejam devidamente cadastrados na Plataforma. Colaboradores podem ser incluídos por proprietários para auxiliar nas tarefas de inclusão e edição de camadas. A comunicação

entre usuários é permitida por meio de uma funcionalidade implementada na Plataforma, conhecida como notificações. As notificações podem ser um meio para comentar ou até denunciar eventuais problemas na Plataforma.

É importante lembrar que podem ter usuários que não tenham cadastro na Plataforma. Neste caso, esta categoria de usuários pode navegar na Plataforma para fazer consultas e até baixar algum material. Esta categoria não tem permissão para editar camadas. Existe também uma outra categoria de usuário, o curador, cujo papel é incentivar a participação dos pares a darem subsídios à plataforma com suas pesquisas, além de mobilizar o conteúdo alimentado em oficinas e fóruns digitais ou presenciais. Assim um curador tem permissão para editar todas as camadas da Plataforma incluindo adicionar novas palavras-chave às camadas. Há também o perfil de administrador, que tem amplos poderes para criar, editar e remover todas as entidades da Plataforma.

4.3 Coleta de Dados

A Plataforma permite que os dados sejam incluídos ou inseridos por meio de contribuições manuais ou até mesmo uma importação de dados em lote. No modo manual, usuários criam e editam localizações de aspectos (*features*) nos seus mapas de interesse diretamente no portal Web e, também podem associar atributos. É importante, neste caso, que os usuários informem todos os metadados associados aos aspectos (*features*). No caso de importação em lote, os usuários podem transferir (*upload*) um grupo de aspectos desde que estejam armazenados em formatos conhecidos para dados geográficos vetoriais, como por exemplo, *shapefile* ou *geojson*. Ao contrário da contribuição manual, neste, os metadados podem ser extraídos automaticamente a partir do conteúdo dos arquivos. Ao utilizar a ferramenta de geolocalização na Plataforma da Pauliceia, os arquivos em formato *shapefile* são automaticamente gerados e podem ser utilizados no caso de importação em lote.

4.4 Controle de Qualidade de Dados

Dados proprietários são necessários para definir algumas métricas e medir a qualidade [Senaratne et al., 2017]. No contexto do projeto Pauliceia, foram avaliados diferentes tipos de indicadores para qualificar os dados derivados de voluntários como gamificação, por exemplo. Mas, chegou a se um consenso pela equipe que a melhor seria uma abordagem *crowdsourcing* por meio de notificações e denúncias pela própria comunidade [Mariano et al., 2018].

A notificação, na Plataforma, é um comentário de um usuário com respeito a uma camada ou até mesmo a um outro comentário. Podem ser descritas notificações com pontos negativos e positivos, alertas ou sugestões de outras referências bibliográficas. A denúncia é para alertar os administradores que uma certa camada possui material ou dados inadequados (problemas com direitos autorais, por exemplo). A equipe teria que avaliar a denúncia e se tiver mérito, o material ou dados devem ser retirados da Plataforma.

4.5 *Feedback* para Comunidade

Em qualquer projeto colaborativo, em particular, com participação de voluntários, é importante que os participantes recebam algum retorno (*feedback*) a respeito das experiências com a interação com o projeto [Mooney et al., 2016]. Nesta mesma linha, a Plataforma age incentivando os voluntários para compartilharem os pontos positivos e negativos e o que poderia ser feito para melhorar a sua interação com a Plataforma. No caso do Projeto Pauliceia, a comunicação pode ser realizada por meio de e-mails ou redes sociais gerenciadas por Pauliceia e com isso a ideia é aprimorar a Plataforma.

Os pesquisadores, por meio da Plataforma Pauliceia 2.0, podem divulgar e compartilhar os seus conjuntos de dados históricos e podem receber *feedback* de outros pesquisadores por meio de notificações. Estes conjuntos de dados serão disponibilizados livremente no portal Web para que tenham maior visibilidade perante a comunidade científica. Historiadores podem notificar nas suas camadas ou de outros, providenciando um retorno sobre o status, comentários, elogios, dicas. Por exemplo, por meio de notificação pode ser sugerida uma referência para uma certa camada. É possível

também notificar todos os membros da comunidade Pauliceia sobre conferências ou chamadas especiais para um periódico, por exemplo. É importante observar que a notificação também deve ser usada para comunicar algo inapropriado que existe nos dados da Plataforma. Uma vantagem que deve ser comentada sobre a Plataforma é a possibilidade de os usuários seguirem as camadas de interesse criadas por outros autores e até receberem notificações por e-mails.

5 Serviço Web de Geocodificação

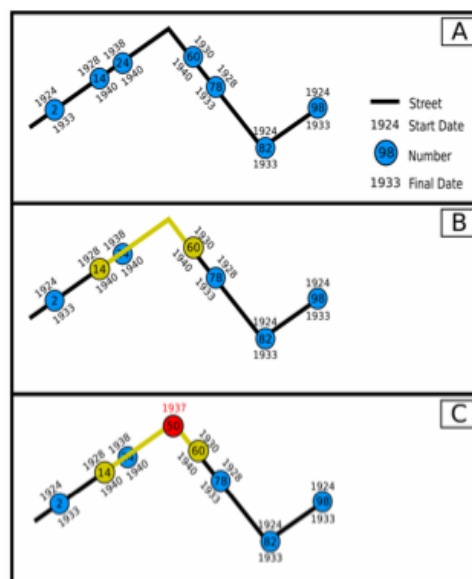
No Pauliceia foi desenvolvida uma solução, um serviço web, para geocodificação dos dados geográficos históricos temporais. Apesar de existir ferramentas de geocodificação de endereços, nenhuma delas lida com dados espaço-temporais, ou seja, não considera a questão temporal dos dados espaciais.

No projeto, toda entidade é georreferenciada como um segmento de rua, um endereço (lugar) e tem um período de tempo associado, que mostra quando era válido. Assim, o nosso método de geocodificação considera todos os períodos de validade associados a entidades espaciais.

A Figura 3 ilustra o processo de Geocodificação implementado por [Sansigolo, 2017] no contexto do projeto. Na parte (A), da figura, há dados da rua e dos locais, pontos de endereço histórico contendo anos inicial e final. Por exemplo, nesta parte (A), o número 98 na rua aparece entre os anos de 1924 e 1933. A partir daí uma busca no banco de dados é realizada. Se o endereço procurado for, por exemplo, 50, a parte representada em (B), ilustra os locais próximos (menor - 14 e maior - 60) ao número 50. Com o uso da distância Euclidiana é localizado o número procurado, conforme mostrado na parte (C) da Figura.

O serviço permite procurar mais de um endereço. Para isso é necessário colocar todos os endereços textuais num arquivo em formato CSV (Valores Separados por Vírgulas) como se fosse importação em lote pela Plataforma. Neste caso, é retornado um shapefile com todas as localizações geográficas, geradas pelo geocodificador, dos endereços fornecidos.

Figura 3: Processo de Geocodificação na Plataforma Pauliceia



Processo de geocodificação para o endereço "Rua Augusta, 50, 1937". Fonte: Ferreira et al., 2018

O serviço web recebe a geometria da rua onde pretende-se fazer a geocodificação de um objeto, o último ponto da rua, o primeiro ponto da rua e o número do endereço pesquisado. Com isso, retorna geometricamente onde está o endereço procurado.

Esse serviço está sendo utilizado no Portal Pauliceia e os resultados obtidos até o momento mostram que o serviço web retorna a localização espacial dentro da faixa temporal buscada. Novos testes ainda estão sendo feitos para aperfeiçoamento do serviço.

6 Estudos de Caso

Como falado anteriormente, a Plataforma do Pauliceia 2.0 disponibiliza uma base cartográfica digital histórica da cidade de São Paulo, abordando o período de sua modernização e crescimento (1870-1940), e encontra-se disponível no link <http://www.pauliceia.dpi.inpe.br/portal/home>.

Atualmente, embora ainda em fase de testes, a Plataforma tem sido utilizada por pesquisadores para a criação de camadas e visualização de estudos Históricos e em

curso e oficinas na UNIFESP e na Universidade da Carolina do Norte (Winston-Salem). Assim, a base tem crescido e novas demandas surgem a partir de seu uso.

Dois estudos de caso que denotam o uso da Plataforma serão apresentados nesta seção. Destaca-se que a utilização do Pauliceia até extrapola a finalidade inicial da Plataforma, permitindo que a Equipe do Projeto vislumbre perspectivas promissoras de aplicação. O primeiro estudo de caso é referente às camadas que retratam enchentes na cidade de São Paulo. O segundo tem como tema a população preta da cidade de São Paulo e visa mapear os locais de presença desta comunidade.

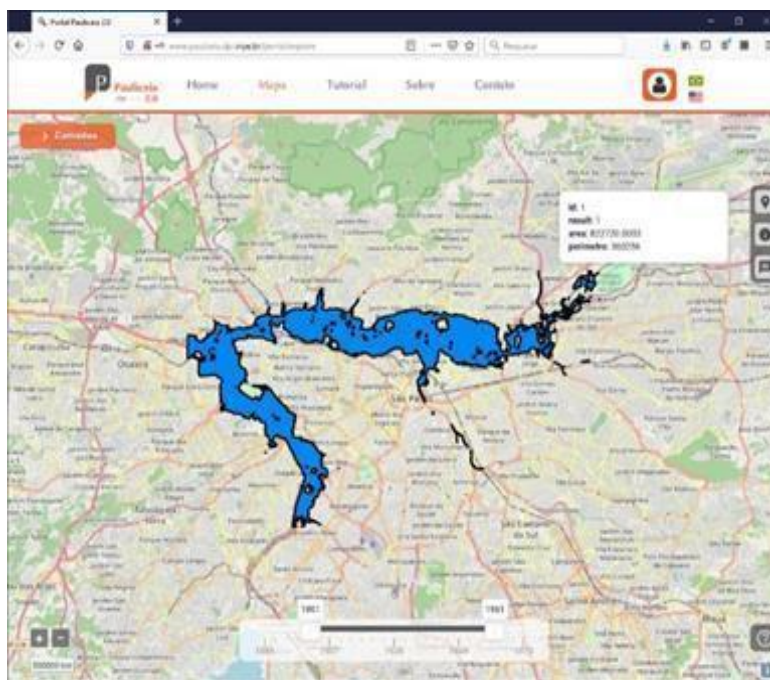
6.1 Camadas com enchentes de São Paulo

Dentre as camadas existentes, a camada com um estudo referente à enchente ocorrida em São Paulo-SP em 2020 foi criada após matéria publicada pela Folha (<https://www1.folha.uol.com.br/cotidiano/2020/02/sao-paulo-revive-mesmas-enchentes-ha-91-anos.shtml>). A referida matéria faz uma comparação com a enchente de 1929, a partir de um mapa produzido pelo grupo Hímaco (já alimentado como camada na plataforma). A contraposição das duas camadas permite fazer análises históricas que identificam continuidades persistentes das áreas alagadas na cidade, como conclui a referida matéria na imprensa.

6.1.1 Visualização da enchente de São Paulo em 1929

O estudo em questão teve como base uma tese de doutoramento na Universidade de São Paulo em 1984 (SEABRA, 1984). A camada foi inserida na Plataforma pelo Grupo Hímaco (História, Mapas e Computadores). A Figura 4 mostra o mapa aproximado da área atingida pela enchente, tendo como mapa de fundo a cidade atual (Open Street Maps).

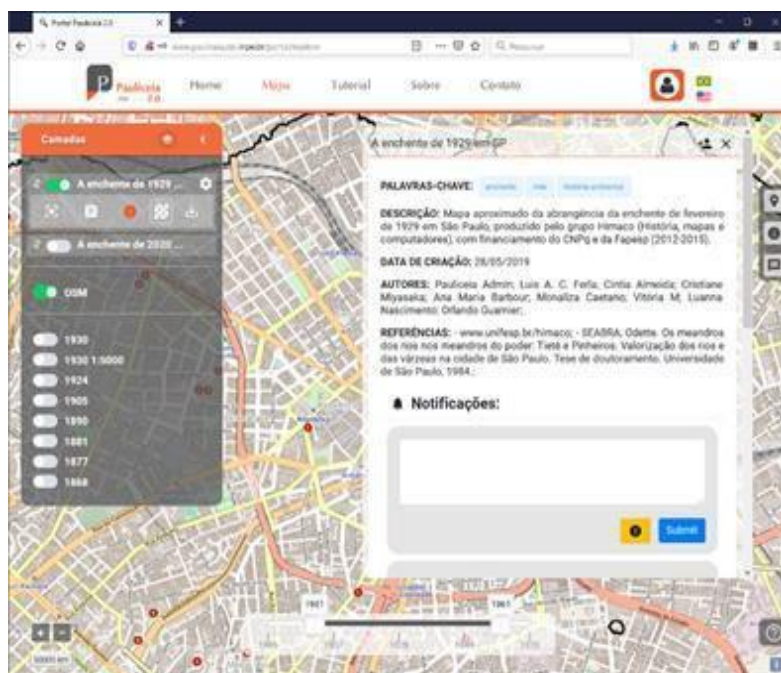
Figura 4: Interface Pauliceia 2.0 para visualização de camada



Mapa aproximado da abrangência da enchente de fevereiro de 1929 em São Paulo-SP. Fonte: Autores

Em complemento ao mapa, informações sobre o estudo, além de metadados, são inseridos na Plataforma. A Figura 5 exhibe os dados para este estudo.

Figura 5: Interface Pauliceia 2.0 para visualização de informações da camada

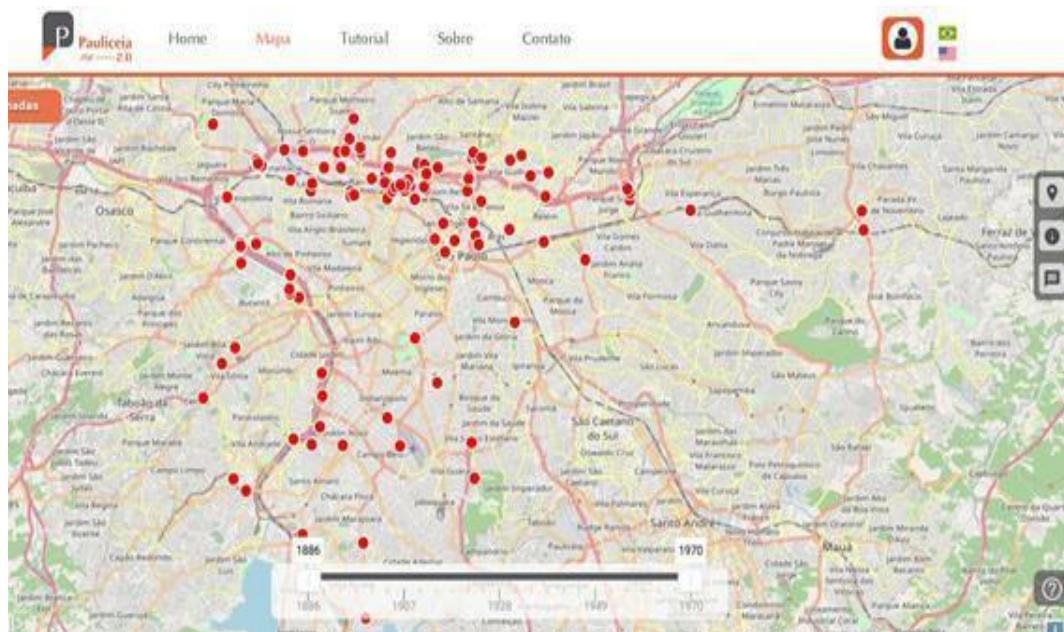


Informações sobre a camada da enchente de fevereiro de 1929 em São Paulo-SP na Plataforma Pauliceia.

6.1.2 Visualização da enchente de São Paulo em 2020

Em fevereiro de 2020, outra enchente de grandes proporções aconteceu na cidade de São Paulo. Foi veiculada pelo jornal Folha de São Paulo uma matéria a partir dos dados do Centro de Gerenciamento de Emergências – GCE (<https://www1.folha.uol.com.br/cotidiano/2020/02/sao-paulo-revive-mesmas-enchentes-ha-91-anos.shtml>). O jornal obteve o mapa de 1929 no site do grupo Hímaco e pode visualizar a área afetada pela enchente de 1929 de forma comparativa à enchente de 2020. Uma nova camada foi inserida na Plataforma Pauliceia pelo Grupo Hímaco (História, Mapas e Computadores) com informações da enchente de 2020. A Figura 6 mostra o mapa com os pontos de alagamento dessa enchente, conforme informações fornecidas pelo Centro de Gerenciamento de Emergências da capital (GCE).

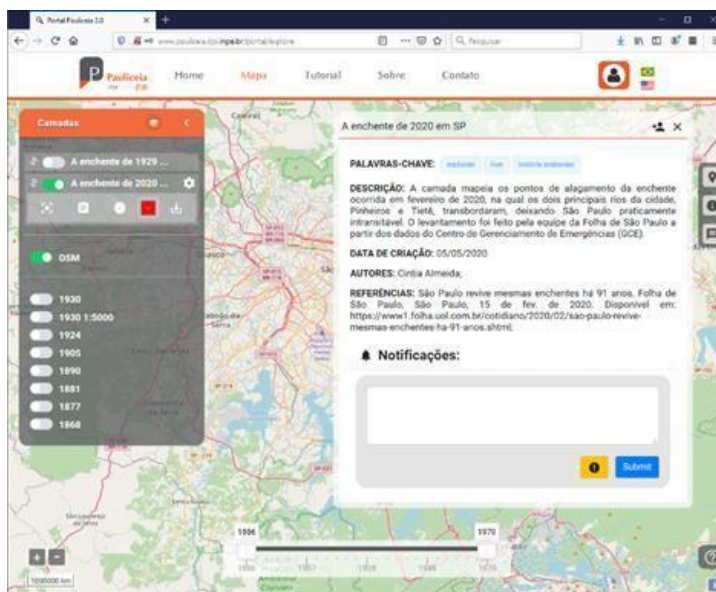
Figura 6: Interface Pauliceia 2.0 para visualização de camada para enchente 2020



Visualização do mapa com os pontos de alagamento da enchente de fevereiro de 2020 em São Paulo-SP.
 Fonte: Autores

As informações sobre o evento recente, além de metadados, podem ser visualizadas na
 Figura 7.

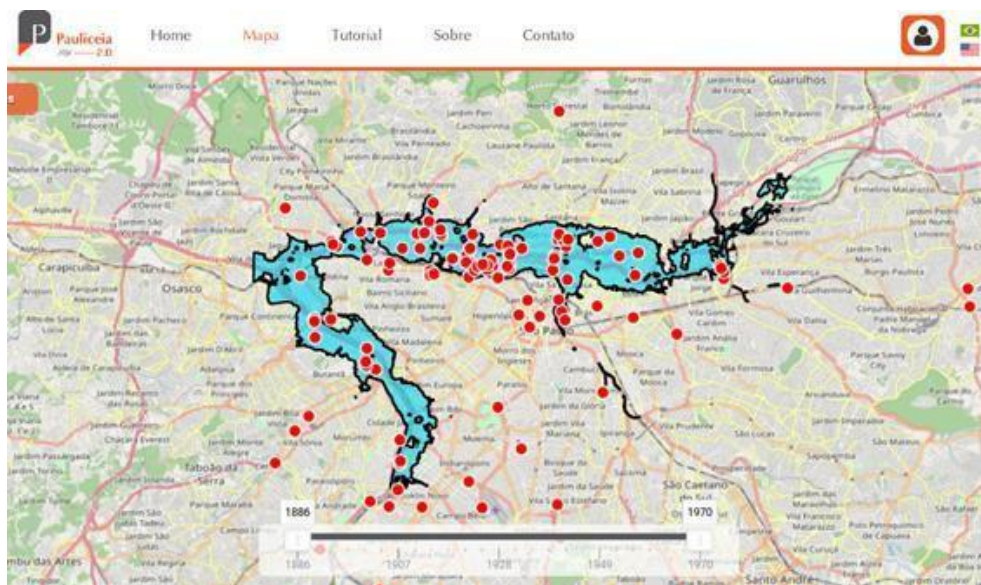
Figura 7: Interface Pauliceia 2.0 para visualização de informações da camada ‘A enchente de 2020 em SP’



Informações sobre a camada da enchente de fevereiro de 1929 em São Paulo-SP na Plataforma Pauliceia. Fonte: Autores.

Destaca-se aqui que, mesmo fora do período de abrangência inicialmente proposto pela Plataforma Pauliceia, a nova camada colabora na compreensão da historicidade de fenômenos atuais da cidade, uma vez que a Plataforma permite visualizar as duas enchentes em um mesmo mapa, conforme apresentado na Figura 8. Esta abordagem é capaz de enriquecer as possibilidades do projeto, ampliando suas perspectivas e indicando possíveis novos caminhos de utilização.

Figura 8: Interface Pauliceia 2.0 para visualização conjunta das enchentes 1029 e 2020

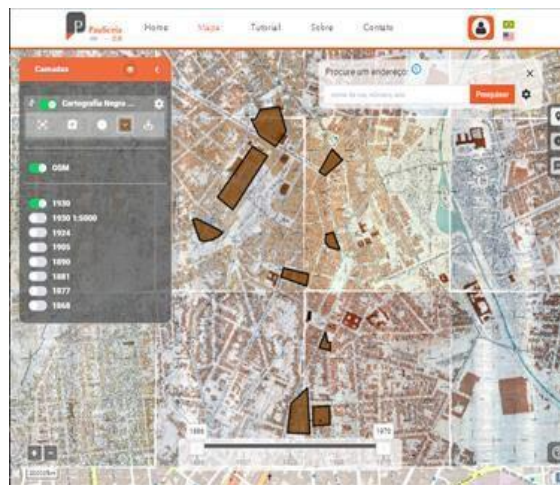


Visualização das enchentes de São Paulo - 1029 e 2020. Fonte: Autores.

6.2 Visualização coletivo Cartografia Negra

A camada produzida pelo coletivo paulistano “Cartografia Negra” é a primeira feita por um grupo de fora da universidade, com uma pesquisa autônoma, sólida e independente, fortalecendo o caráter de ciência aberta que a equipe pretende conferir à plataforma. Abordando a temática da população preta da cidade de São Paulo, a proposta foi a de produzir um material que mapeasse estes locais de presença da comunidade. A figura 9 apresenta a camada.

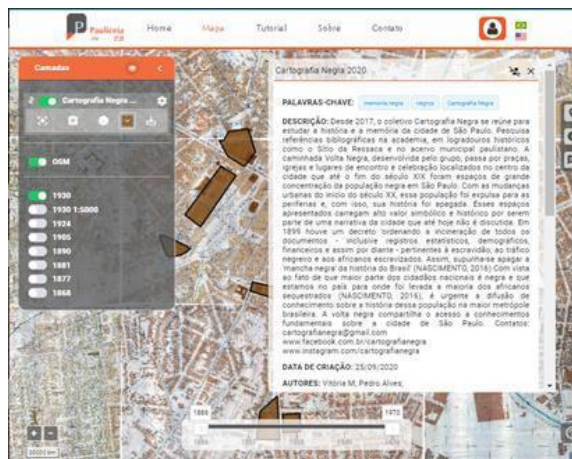
Figura 9: Interface Pauliceia 2.0 para visualização de camada ‘Cartografia Negra’



Camada ‘Cartografia Negra’ aproximada. Fonte: Vitória Maria Fontes – Grupo Himaco

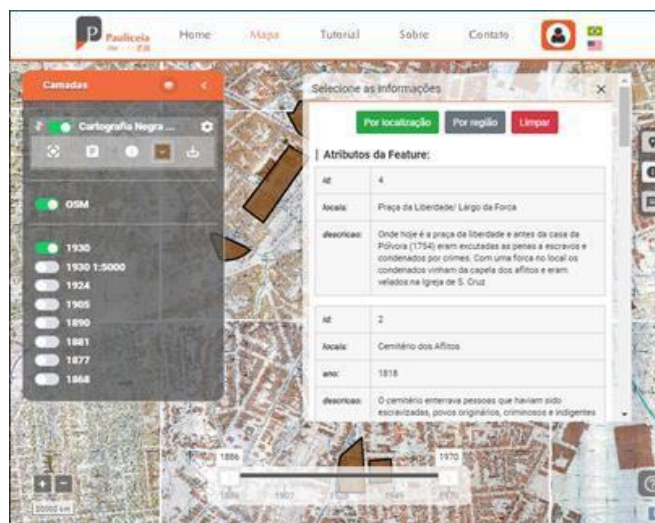
Essa camada contém informações gerais (imagem 10) e um resumo referente a cada polígono da camada (imagem 11). Desta forma é possível compreender a ideia geral da visualização, a bibliografia consultada, os acervos consultados para coleta de informações, e também informações mais específicas, que apresentam mais detalhadamente cada ponto.

Figura 10: Interface Pauliceia 2.0 para visualização de informações gerais



Informações gerais da camada. Fonte: Autores

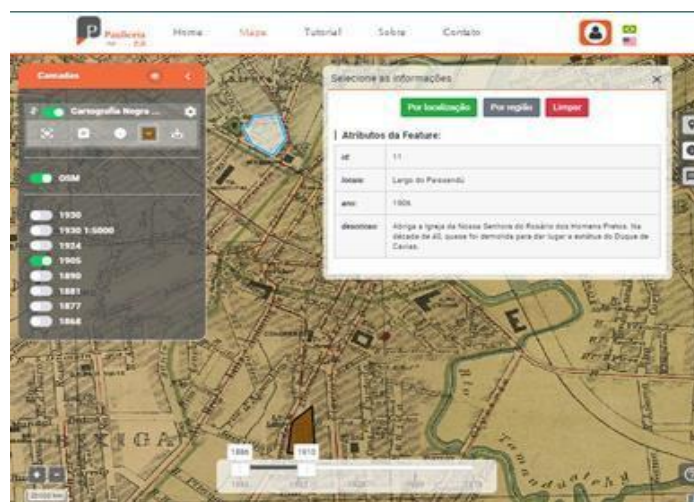
Figura 11: Interface Pauliceia 2.0 para visualização de informações de ‘feature’



Informações específicas de cada polígono que compõem a camada. Fonte: Autores

Assim, a visualização é bastante relevante pois abrange o período completo da plataforma, assim possibilitando o uso de todas as cartografias históricas disponíveis. Ao utilizar o *slider* temporal é possível fazer análises interessantes sobre os apagamentos destes espaços, que estão presentes tanto na cartografia histórica, quanto na realidade. Assim apresenta a Figura 12.

Figura 12: Visualização da camada na cartografia da década de 1910



Camada na cartografia da década de 1910. Com a utilização do *slidder*, na mesma época, possibilitou filtrar o espaço da comunidade preta que existia na cidade. Fonte: Autores.

7 Considerações Finais

A Plataforma Pauliceia 2.0 está em pleno uso não só pela própria equipe de historiadores, mas também por pesquisadores da área. O intenso uso possibilita os ajustes necessários na implementação das funcionalidades do Portal. A equipe está à disposição para compartilhar a Plataforma com outras instituições que queiram fazer um projeto similar para as suas cidades. Neste sentido, na próxima etapa do Projeto já haverá uma colaboração, em caráter de estudo-piloto, com a PUC de Campinas, que desenvolverá o mapeamento histórico para aquela cidade.

O período coberto pela Plataforma é de 1870 a 1940. No entanto, se houver interesse e recursos humanos há planos para incluir outros períodos.

Agradecimentos

O financiamento da Fase 1 do projeto Pauliceia foi fornecido pelo Programa eScience da FAPESP (Bolsa 2016 / 04846-0). Agradecemos à FAPESP também pela concessão de bolsas de estudo: #2017/03852-9, #2017/11637-0, #2017/11625-2 e #2017/11674-3.

Referências

- Budig, B.; van Dijk, T.C.; Feitsch, F.; Artega, M.G. **Polygon Consensus: Smart Crowdsourcing for Extracting Building Footprints from Historical Maps.** *Proceedings of 24th ACM SIGSPATIAL International Conference in Advances in Geographic Information System.* U.S.A. 2016.
- Câmara, et al; **Banco de Dados Geográficos.** Editora Mundogeo. 2005.
- Carter, J. **Fundamentals of Geographic Information Systems: A Compendium Chapter on Defining the Geographic Information Systems.** *American Society for Photogrammetry and Remote Sensing.* 3-8. 1989.
- Digital Harlem Blog. The Project. 2017. Disponível em: <https://digitalharlemblog.wordpress.com/>. Acesso em 28 nov. 2020.
- Estellés-Arolas, E.; González-Ladrón-De-Guevara, F. **Towards an Integrated Crowdsourcing Definition.** *Journal of Information Science.* 38(2). 189-200. 2012.
- Garrard, C. **Georprocessing with Python.** Manning Publications. 2016.
- Goodchild, M.F. **Citizens as Sensors: the World of Volunteered Geography.** *GeoJournal.* 69(4). 211-221. 2007.
- Goodchild, M.F.; Li, L. **Assuring the Quality of Volunteered Geographic Information.** *Spatial Statistics.* Vol. 1. 110-120. 2012.
- Ferreira, K.R.; Ferla, L.; Queiroz, G.R.; Vijaykumar, N.L.; Noronha, C.A.; Mariano, R.M.; Wassef, Y.; Taveira, D.; Dardi, I.B.; Sansigolo, G.; Guarnieri, O.; Musa, D.L.; Rogers, T.; Lesser, J.; Page, M.; Britt, A.G.; Atique, F.; Santos, J.Y.; Morais, D.S.; Miyasaka, C.R.; Almeida, C.R.; Nascimento, L.G.M.; Diniz, J.A.; Santos, M.C.; **A Platform for Collaborative Historical Research based on Volunteered Geographic Information.** *Journal of Information and Data Management.* 8(1). 2018.
- Greenspan, B. **Are Digital Humanities Utopian?** In: Gold, M.; Klein, L. (eds). *Debates in the Digital Humanities.* University of Minnesota Press. Minneapolis. 393-409. 2016.
- Hollenstein, L.; Purves, R. **Exploring Place Through User-Generated Content: Using Flickr Tags to Describe City Cores.** *Journal of Spatial Information Science.* 1(1). 21-48. 2010.
- IBGE. **Tabela 1287-População dos Municípios das Capitais e Percentual da População dos Municípios das Capitais em relação aos das Unidades de Federação nos Censos Demográficos.** 2011. Disponível em: <https://sidra.ibge.gov.br/tabela/1287#/n6/3550308/v/591/p/all/l/v.p.t/resultado>. Acesso em: 09 abr. 2020.
- RBHD, Rio de Janeiro, v. 1, n. 1, Dossiê Temático 3, p. 110-133, jan./jun., 2021

- Mooney, P.; Minghini, M.; Laakso, M.; Antoniou, V.; Olteanu-Raimond, A-M.; Skopeliti, A. **Towards a Protocol for the Collection of VGI Vector Data**. *ISPRS International Journal of Geo-Information*. 5(11). 2016.
- Mariano, R.M; Ferreira, K.R.; Ferla, L.A.C. **VGI Protocol and Web Service for Historical Data Management**. *Proceedings of XIX GeoInfo*. Campina Grande, PB. 103-115. 2018.
- Medeiros, C.B.; Pires, F. **Databases for GIS**. ACM. New York. 1994.
- Page, M.C.; Durante, K.; Gue, R. **Modeling the History of the City**. *Journal of Map and Geography Libraries*. 9(1-2). 128-139. 2013.
- Sansigolo, G; **Web service para geocodificação de endereços em banco de dados espaço-temporais**. Trabalho de Conclusão de curso de Tecnólogo em Análise e Desenvolvimento de Sistemas, FATEC , São José dos Campos, 2017.
- See, L.; Mooney, P.; Foody, G.; Bastin, L.; Comber, A.; Estima, J.; Fritz, S.; Kerle, N.; Jiang, B.; Laakso, M.; **Crowdsourcing, Citizen Science or Volunteered Geographic Information? The Current State of Crowdsourced Geographic Information**. *International Journal of Geo-Information*. 5(5). 2016.
- Senaratne, H.; Mobasheri, A.; Ali, A.L.; Capineri, C.; Haklay, M. **A Review of Volunteered Geographic Information Quality Assessment Methods**. *International Journal of Geographic Information Science*. 31(1). 139-167. 2017.
- Spiro, L.; **This is Why We Fight: Defining the Values of the Digital Humanities**. In: Gold, M.K. (ed). *Debates in the Digital Humanities*. University of Minnesota Press. Minneapolis. 16-35. 2012.
- Zaidan, R.T. **Geoprocessamento Conceitos e Definições**. *Revista de Geografia*. 7(2). 195-201. PPGeo-UFJF. 2017.
- White, J. W.; Gilbert, H. *Laying the foundation*. Purdue University Press, 2016. 3, 15.