

Pauliceia 2.0: Uma plataforma espaço-temporal para Humanidades Digitais

Relatório Científico Final - fevereiro de 2019

Pesquisadores Principais:

Dr. Luis Antonio Coelho Ferla (Unifesp)

Dr. Karine Reis Ferreira Gomes (Inpe)

Instituições-sede do projeto:

Universidade Federal de São Paulo

Instituto Nacional de Pesquisas Espaciais

Equipe:

Nome	Qualificação	Instituição	Bolsas
Luis Ferla	Doutor	Unifesp	
Karine Reis Ferreira	Doutora	Inpe	
Cristiane Regina Myiasaka	Doutora	Prefeitura do Município de Campinas	
Daniela Leal Musa	Doutora	Unifesp	
Fernando Atique	Doutor	Unifesp	
Gilberto Ribeiro de Queiroz	Doutor	Inpe	
Jeffrey Lesser	Doutor	Emory University	
Michael Page	Doutor	Emory University	
Nandamudi Vijaykumar	Doutor	Inpe	
Thomas Rogers	Doutor	Emory University	
Andrew Graham Britt	Doutor	Emory University	
Ana Maria Alves Barbour	Mestranda	Unifesp	
Orlando Guarnier Farias	Mestrando	Unifesp	
Rodrigo Monteiro Mariano	Mestrando	Inpe	Mestrado
Gabriel Sansigolo	Mestrando	Inpe	
Marta Cristina Pacoret Rodríguez	Engenheira civil	Consejería de Educación de Andalucía	
Janaina Yamamoto Santos	Geógrafa	Arquivo do Estado	
Priscila Machado Meireles	Arquiteta e urbanista	Universidade São Francisco e Senac	
Wania Mazzarello	Historiadora		
Carlos Alberto Noronha	Graduado em Análise de Sistemas	Fatec	IC
Cintia Rodrigues de Almeida	Graduanda em História	Unifesp	IC
Denis Taveira L. de Lima	Graduando em Ciência e Tecnologia	Unifesp	IC
Ester Dantas Reis Nunes	Graduanda em História	Unifesp	IC
Luanna M. do Nascimento	Graduanda em História	Unifesp	IC
Monaliza Caetano	Graduanda em História	Unifesp	IC (CNPq)
Yasmin Wassef	Graduando em Ciência e Tecnologia	Unifesp	IC
Eduardo de Souza Goiabeira	Graduando em História	Unifesp	IC
Vitória Martina Fontes	Graduanda em História	Unifesp	IC (CNPq)
Júlio Oliveira	Graduando em História	Unifesp	

Número do Processo: 2016/04846-0

Período de vigência: 01/02/2017 a 31/01/2019

Período coberto por esse relatório: 01/02/2018 a 30/01/2019



Luis Antonio Coelho Ferla



Karine Reis Ferreira Gomes

Sumário

1. Resumo do projeto proposto.....	3
2. Realizações no período.....	5
2.1 Atividades gerais.....	5
2.1.1 Reuniões gerais do projeto.....	5
2.1.2 Apresentações da plataforma.....	6
2.1.3 Estudo do meio no centro de São Paulo.....	8
2.1.4 Design gráfico.....	10
2.2 Núcleo “material geo-histórico”.....	11
2.2.1 Capacitação.....	11
2.2.2 Georreferenciamento e vetorização.....	13
2.2.3 Banco de dados para a geolocalização.....	14
2.2.4 Determinação da origem das ruas.....	15
2.2.5 Disciplina de Laboratório de Ensino e Pesquisa em História I.....	16
2.2.6 II Ciclo de Oficinas do Grupo Hímaco.....	16
2.3 Núcleo “plataforma computacional”.....	18
2.3.1. Capacitação e acompanhamento das atividades.....	18
2.3.2. Projeto e desenvolvimento da plataforma computacional.....	18
2.3.3. Arquitetura Docker da plataforma computacional Pauliceia 2.0.....	27
2.3.4. Projeto e desenvolvimento da plataforma Pauliceia 2.0.....	29
3. Apoio institucional recebido no período.....	37
4. Lista de publicações e participações em eventos.....	39
4.1 A Plataforma Pauliceia 2.0.....	39
4.2 Apresentações em conferências internacionais.....	39
4.3 Apresentações em conferências nacionais.....	40
4.4 Trabalhos publicados em anais de conferências nacionais.....	41
4.5 Trabalhos aceitos para publicação em revista.....	41
5. Perspectivas para o futuro.....	42
Anexos.....	43
Anexo 1 – Design gráfico.....	44
Anexo 2 – Tutorial de alimentação do banco de dados.....	49
Anexo 3 – Determinação da origem das ruas.....	57
Anexo 4 – Ementa da disciplina de Lab. 1.....	62
Anexo 5 – Artigo apresentado no GeoInfo 2018.....	66
Anexo 6 – Tutorial de uso da plataforma.....	81
Anexo 7 – Relatórios sintéticos dos bolsistas.....	98

1. Resumo do projeto proposto

As chamadas humanidades digitais vêm se constituindo em realidade crescentemente incontornável para aqueles que trabalham com o conhecimento histórico. Já não se trata mais tanto de questionar suas possibilidades de afirmação junto à comunidade acadêmica em questão, mas de discutir os desdobramentos dela, tais como a necessidade (ou não) de definir alguma identidade epistemológica para a área e as implicações teóricas e práticas no cotidiano do pesquisador e nos sistemas institucionais de avaliação do seu trabalho. O presente projeto objetiva a concepção e o desenvolvimento de uma plataforma computacional para pesquisa histórica colaborativa. A ideia principal é desenvolver ferramentas de software no estado da arte, tais como um portal web e plugins de Sistema de Informações Geográficas (SIG), que permitam a pesquisadores de ciências humanas criar, organizar, armazenar, integrar, processar e publicar conjuntos de dados de história urbana. A plataforma integrará todas essas ferramentas. Os dados de história urbana, enquanto representação da evolução de localizações espaciais urbanas ao longo do tempo, são dados espaço-temporais. Dessa forma, o desenvolvimento da plataforma implica em muitos desafios em Geoinformática:

- (a) como modelar, armazenar, geocodificar, manipular e visualizar conjuntos de dados espaço-temporais;
- (b) como integrar informações de história urbana com conjunto de dados espaço-temporais de outras áreas, tais como saúde pública e mobilidade, que estão disponíveis na rede por meio de diferentes provedores;
- (c) como definir e vincular informação semântica a um conjunto de dados de história urbana; e
- (d) como promover o trabalho colaborativo entre pesquisadores, facilitando o compartilhamento e o acesso ao conhecimento.

Para testar e validar a plataforma proposta será desenvolvido um estudo de caso com uma base cartográfica histórica da cidade de São Paulo referente ao período de sua modernização urbano-industrial (1870-1940). As informações serão organizadas em uma



base de dados espaço-temporal. A plataforma permitirá o acesso a essa base de dados e a interação entre pesquisadores interessados, que por sua vez poderão alimentar a base com eventos passíveis de representação espacial e temporal. Dessa forma, pesquisadores serão capazes de produzir mapas e visualizações de suas próprias pesquisas, ao mesmo tempo em que alimentarão o sistema com suas informações.

Pretende-se, assim, criar as condições para o enriquecimento das abordagens da história de São Paulo daquele período, fazendo-o em conformidade com os mais recentes e interessantes desdobramentos das chamadas humanidades digitais, voltados ao trabalho colaborativo e à livre circulação do conhecimento.

2. Realizações no período

A organização das atividades cotidianas da equipe seguiu a estrutura apresentada no projeto de pesquisa aprovado, dividida em duas áreas de concentração: *material geo-histórico* e *plataforma computacional*. As interações entre os dois grupos assim constituídos foram bastante recorrentes, dadas as necessidades da própria pesquisa. Além disso, algumas atividades tiveram um caráter geral e não se enquadram naquela divisão. Dessa forma, assim como feito no relatório parcial, as realizações do período serão relatadas em três seções: atividades gerais, material geo-histórico e plataforma computacional.

2.1 Atividades gerais

2.1.1 Reuniões gerais do projeto

Para além de diversas reuniões presenciais e por video-conferência realizadas por vários dos membros e sub-grupos do projeto, foram realizadas no período três reuniões gerais, destinadas a discutir de conjunto a concepção e o andamento do projeto (em 2018, foram duas). III Reunião Geral do Projeto foi realizada no dia 24 de maio, no INPE. A IV Reunião foi realizada no dia 24 de agosto, no campus Guarulhos da Unifesp. Nesta oportunidade, o andamento do projeto foi apresentado a uma delegação da Emory University, composta por Jeffrey Lesser, Diretor do Halle Institute for Global Research and Learning e membro da equipe do projeto, Lanny Liebeskind, Vice Provost for Strategic Research Initiatives, Philip Wainwright, Vice Provost for Global Strategy and Initiatives, e Kelly Richmond Yates, Diretora Assistente do Halle Institute for Global Research and Learning (figura 1). Por fim, na manhã do dia 30 de outubro foi realizada a V Reunião Geral, nas dependências do Arquivo Público do Estado, tendo como objetivo

principal as preparações finais do lançamento da versão beta da plataforma, na tarde do mesmo dia.



Figura 1 - IV Reunião Geral do Projeto (agosto de 2018)

2.1.2 Apresentações da plataforma¹

O caráter de ciência aberta do projeto implica na busca sistemática de interlocução com a comunidade-alvo, constituída prioritariamente pelos estudiosos da história da cidade de São Paulo, mas também por aqueles possíveis interessados em

¹ As comunicações em eventos científicos estão indicadas no item 4.

utilizar a tecnologia e a metodologia desenvolvidas pelo projeto em outros recortes espaciais e temporais. Tais interlocuções tiveram dois principais objetivos: divulgar propriamente o projeto e receber sugestões quanto ao seu desenvolvimento. Como afirmado no relatório parcial, quando era justificado o evento de lançamento do projeto, a identidade colaborativa da plataforma implica na sua ampla divulgação junto aos possíveis interessados, de forma a não se incorrer em uma das principais fragilidades de projetos similares que fracassam, qual seja o pequeno número de acessos e a pouca utilização da ferramenta. Muitas vezes, esse resultado indesejado é consequência de problemas da própria estrutura da ferramenta disponibilizada, mas outras vezes é resultado do desconhecimento da sua existência por parte dos usuários potenciais. Além disso, quando estes são instados a opinar e tecer críticas quanto a características da ferramenta a ser desenvolvida, crescem as possibilidades de participação e engajamento.

Com essa perspectiva, foi realizada o lançamento da versão beta da plataforma, no dia 30 de outubro de 2018, no auditório do Arquivo Público do Estado. O cronograma inicial do projeto previa tal evento para agosto, mas dois fatores levaram ao adiamento: por um lado, a legislação eleitoral restringia a realização de eventos públicos nas dependências do Arquivo do Estado, até o término das eleições majoritárias. Por outro, decidiu-se avançar no aprimoramento técnico da plataforma para sua apresentação mais ampla. Compareceram cerca de 75 pessoas, e a transmissão pelas redes sociais atingiu 1150 visualizações.² Se tal público pode ser considerado bastante satisfatório, para uma tarde de dia de semana, é também verdade que está bem abaixo daquele que compareceu ao lançamento do projeto, em abril de 2017 (cerca de 120 pessoas), principalmente em se considerando o apelo presumivelmente maior do segundo evento. A explicação para tal está na grande dificuldade encontrada na divulgação do lançamento, em um período de plena campanha para o segundo turno das eleições presidenciais (o evento teve lugar apenas dois dias depois deste). Além disso, o Arquivo

2 Para ver o vídeo na íntegra, clicar em <https://www.facebook.com/arquivoestado/videos/704166923288544> (acessado em 7 de fevereiro de 2019).

do Estado tampouco pode veicular o evento nas suas plataformas de divulgação, pelas mesmas restrições legais indicadas acima. De todo modo, como se dera no evento de 2017, a qualidade do debate e das intervenções do público presente garantiram o sucesso pretendido.

A apresentação da versão beta em outubro foi precedida de dois pré-lançamentos, realizados como parte das programações do XXIX Encontro Regional de História da ANPUH/SP e do Colóquio Histórias de São Paulo, ambos no Campus Guarulhos da Unifesp, nos dias 5 de setembro³ e 3 de outubro⁴, respectivamente. Esses pré-lançamentos foram extremamente oportunos, pois se deram em eventos com grande concentração de possíveis usuários da plataforma.

Já após o lançamento da versão beta, foram realizadas três reuniões para discutir a plataforma com interlocutores estratégicos: no dia 14 de novembro, com Juliana Marques da Silva, assistente de direção da Escola de Ciências Sociais (CPDOC) da FGV do Rio de Janeiro, quando foram aventadas possibilidades de parceria e foi definida uma apresentação da plataforma para um público mais amplo em maio desse ano, nas dependências do CPDOC; no dia 3 de dezembro, com pesquisadores da Faculdade de Arquitetura e Urbanismo da USP, em reunião organizada pelos professores Ana Lana e Fernando Atique, quando ficou decidido a realização de um mapathon de alimentação da plataforma em março desse ano, na própria FAU; e no dia 31 de janeiro, em Atlanta, quando foram discutidas a plataforma e as suas perspectivas futuras com os membros da equipe do projeto da Emory University.

2.1.3 Estudo do meio no centro de São Paulo

No dia 21 de abril de 2018, foi realizado um estudo do meio no centro histórico de São Paulo (figura 2). O objetivo da atividade, coordenada pelo Prof. Fernando Atique, foi possibilitar à equipe do projeto, principalmente àqueles da área da computação, algo do

3 https://www.encontro2018.sp.anpuh.org/conteudo/view?ID_CONTEUDO=395 (acessado em 7 de fevereiro de 2019).

4 <https://historiasdesaopaulo.wordpress.com/lancamento-pauliceia-2-0-2> (acessado em 7 de fevereiro de 2019).



suporte do conhecimento histórico que está relacionado ao seu escopo. Essa demanda surgiu da própria equipe do INPE participante do projeto, e expressa a ideia de humanidades digitais enquanto trabalho comum, ainda que interdisciplinar, e não como um conjunto de atividades isoladas e desconectadas entre si. Esse é o mesmo significado, em sentido inverso, do curso de banco de dados geográficos que as bolsistas do projeto do curso de História fizeram no INPE em janeiro desse ano.



Figura 2 - Estudo do meio no centro de São Paulo, em 21 de abril de 2018

2.1.4 Design gráfico⁵

Dando continuidade às criações iniciadas em 2017, os elementos de Identidade Visual foram concluídos em meados de março de 2018, contemplando a marca do Projeto “Pauliceia 2.0”, o papel timbrado e demais ícones, visando principalmente a legibilidade para a plataforma web. Tais elementos de branding foram fundamentais para estabelecer uma unidade de apresentação da plataforma, como a paleta de cores, as tipografias e os ícones gráficos que compuseram todo o conjunto digital.

Por sua vez, a interface digital desenvolveu-se em duas etapas concomitantes: a elaboração de um Portal de apresentação do projeto, contendo as informações gerais, dados da equipe envolvida no processo, assim como subsídios para contato e suporte aos interessados; e a criação da Plataforma de interação do mapa, com seus meios de acesso através de barras laterais e pictogramas que pudessem tornar a atividade amigável e didática aos futuros usuários e pesquisadores. Estes layouts gráficos foram produzidos dentro do contexto de HTML/CSS, compatível com a maioria dos navegadores e dispositivos de acesso a internet atualmente, adotando como base inicial a resolução de tela com 1600 x 900 pixels.

Todos os elementos visuais, tanto do Portal quanto de interação da Plataforma, foram concluídos no final do mês agosto de 2018 objetivando a apresentação da versão Beta de testes do projeto, que ocorreu ao final de outubro do mesmo ano. Nesta etapa de criação definiram-se as quatro páginas de acesso e respectivos conteúdos, que foram afixados no cabeçalho disponível na borda superior da tela, intitulados “Home”, “Mapa”, “Sobre” e “Contato”, assim como foram desenvolvidos os pictogramas e botões de interação com o mapa, concentrados nas barras flutuantes disponíveis nas laterais da tela de navegação.

O processo de refinamento dos objetos gráficos ocorreu após o lançamento da versão beta (testes da plataforma), concentrando-se entre os meses de novembro de

5 Essa atividade ficou sob a responsabilidade de Priscila Meireles. O relatório completo referente ao design gráfico, com as figuras respectivas, se encontra no Anexo 1.

2018 a janeiro de 2019, na finalização dos elementos correspondentes ao Painel de Usuário.

Após a finalização dos produtos gráficos, a etapa de transição da versão HTML foi desenvolvida inteiramente pela equipe do INPE, uma vez que a página on line está hospedada em um dos seus servidores. Todos os elementos de programação visual elaborados podem ser consultados através do endereço eletrônico: www.pauliceia.dpi.inpe.br.

Como qualquer website, há atualizações e ajustes pontuais que possivelmente serão detectados e corrigidos a longo prazo, visando adaptar-se à dinâmica de acessos e necessidades dos usuários e pesquisadores da plataforma “Pauliceia 2.0”. Portanto, os meios de comunicação e contato com os usuários foram tratados de maneira essenciais neste processo de construção coletiva da plataforma digital.

2.2 Núcleo “material geo-histórico”

2.2.1 Capacitação

Ao longo do período em questão, foram organizados grupos de estudos para a discussão de bibliografia atinente ao projeto. Os temas abordados se concentraram ao redor de três eixos principais: as humanidades digitais e o ofício do historiador; as relações entre história e geografia; e a história da cidade de São Paulo. Muitos dos textos estudados se situavam em mais de um desses eixos. A programação completa desses grupos de estudos, referente ao período desse relatório, se encontra na tabela abaixo.

Data	Título da obra	Autor(es)
07/03/2018	Crowdsourcing, Citizen Science or Volunteered Geographic Information? The Current State of Crowdsourced Geographic Information.	Linda See ¹ , Peter Mooney, Giles Foody, Lucy Bastin, Alexis Comber, Jacinto Estima, Steffen Fritz, Norman Kerle, Bin Jiang, Mari Laakso, Hai-Ying Liu, Grega Milčinski, Matej Nikšić, Marco Painho, Andrea Podör, Ana Maria Olteanu-Raimond e Martin Rutzinger.
04/04/2018	Activist-driven Innovation: uma história interpretativa do software livre	CARAMEZ, Maria e ORTELLADO, Pablo
02/05/2018	Introduction to Historical GIS and the Study of Urban History.	DEBATS, Donald A. e GREGORY, Ian N
20/06/2018	Mapas, saber e poder	HARLEY, Brian
01/08/2018	Relatório parcial de projeto de Iniciação Científica	REIS, Ester
29/08/2018	Relatório parcial de projeto de Iniciação Científica	CAETANO, Monaliza
26/09/2018	Ciência Aberta em Questão	ALBAGLI, Sarita
17/10/2018	Ciência cidadã: modos de participação e ativismo informacional.	PARRA, Henrique
07/11/2018	Relatório parcial de Monografia de final de curso	SILVA, Carlos Alberto.
05/12/2018	Orfeu Extático na Metrópole: São Paulo, sociedade e cultura nos frementes dos anos 20	SEVCENKO, Nicolau

2.2.2 Georreferenciamento e vetorização⁶

Ao longo da pesquisa, o subgrupo de vetorização foi responsável pelo desenvolvimento de uma metodologia de georreferenciamento, vetorização e revisão que contemplassem as cartas SARA, seus vetores e demais mapas selecionados para o projeto. Nos meses iniciais, os pesquisadores se preocuparam em georreferenciar o conjunto de cartas do levantamento SARA Brasil, utilizando um gradeamento para este fim. Posteriormente os mesmos mapas foram vetorizados, utilizando enquanto base os vetores do Mapa Digital da Cidade (MDC), levantamento oficial da Prefeitura de São Paulo produzido em 2006. Vetores que não existiam ou que se alteraram ao longo do tempo foram apagados ou redesenhados, a fim de que os vetores presentes no MDC convergissem com as ruas indicadas no SARA Brasil. Este mesmo processo, intitulado pelo grupo como “vetorização negativa”, foi utilizado na vetorização de todas os mapas utilizados no projeto. Regressivamente, cada mapa teve seus vetores viários atualizados conforme o desenho da carta respectiva.

Após as etapas descritas acima, o material produzido foi carregado na plataforma de edição, também utilizada pela equipe de numeração. A partir de então, as demais equipes, que lidavam diretamente com os vetores para cumprir suas próprias tarefas, passaram a revisar o material, apontando possíveis melhorias e problemas de uso. Esta etapa foi essencial para que os vetores atendessem às necessidades daqueles que alimentaram a o banco de dados de numeração. Neste sentido, os vetores foram corrigidos e alimentados novamente na plataforma. Alguns vetores foram realinhados, aumentados ou renomeados através da tabela de referência.

Outra importante frente de trabalho foi a extensão do gradeamento produzido para o georreferenciamento das cartas SARA Brasil. Essa demanda foi motivada pela própria extensão do recorte espacial inicial do projeto: com a agregação da região do

⁶ Para essa atividade, foi constituído um sub-grupo da equipe do projeto, constituído por Monaliza Caetano e Orlando Guarnier, estudantes de História da Unifesp e membros do Grupo Hímaco.

Pari, novas folhas SARA foram selecionadas e adicionadas. A reformulação do gradeamento foi pesquisada e posteriormente feita. Esta etapa foi importante marco no desenvolvimento da metodologia de georreferenciamento da pesquisa e do próprio grupo, principalmente para a anexação futura de outras regiões da cidade, conforme os interesses de pesquisadores.

2.2.3 Banco de dados para a geolocalização⁷

Como discutido no projeto inicial e no relatório parcial, o banco de dados de números de imóveis das ruas da área piloto constitui o suporte histórico-técnico essencial para o funcionamento do geolocalizador de endereços da plataforma. Aqui, houve uma mudança de metodologia significativa em relação ao que havia sido definido e encaminhado no primeiro ano do projeto. A alimentação de dados via um web-editor, como explicada no relatório parcial, foi substituída pela alimentação direta em uma tabela, em que cada registro corresponde a um imóvel determinado (ver amostra na Figura 3). Um dos campos da tabela se refere à distância em metros do imóvel em relação ao começo da rua, informação que passou a ser utilizada pela Prefeitura após a promulgação da Lei nº 3427, de 19 de novembro 1929 (conhecida como Código de Obras de Arthur Saboya), para a definição do próprio número do imóvel. Dessa forma, o ponto respectivo pode ser localizado com razoável precisão no espaço cartográfico digital, e as numerações que esse ponto teve anteriormente puderam ser recuperadas dos próprios livros de emplacamento, que informam sempre o número novo e antigo de cada imóvel. Assim, cada ponto alimentado ao banco de dados teve sua localização espacial e os números sucessivos que recebeu determinados com precisão muito maior do que possibilitado pela metodologia anterior. Essa modificação metodológica foi resultante das discussões da equipe responsável com o restante da equipe do projeto,

⁷ Da mesma forma como se deu no período anterior, essa atividade ficou a cargo de um sub-grupo da equipe do projeto, constituído pelos bolsistas Cintia Rodrigues, Luanna Nascimento, Eduardo Goiabeira e Vitória Martins Fontes, estudantes de História da Unifesp, sob a coordenação da Dra. Cristiane Myiasaka.

principalmente quando das três reuniões gerais realizadas no período. No Anexo 2 se encontra o tutorial elaborado para orientar o preenchimento do banco de dados.

	A	B	C	D	E	F	G	H	I	J	K
1	Id da rua	Id ponto	metragem	logradouro	número	ero original	data inicial (DD)	data final (DD)	fonte	autor da alime	Data
104	165	165-23	22.40	rua lopes de oliveira	23	23	06/01/1936		livro de emplacamento 20.p	Luanna	12/10/2018
105	165	165-23	22.40	rua lopes de oliveira	3	3	11/10/1927	05/01/1936	livro de emplacamento 20.p	Luanna	12/10/2018
106	165	165-23	22.40	rua lopes de oliveira	1	1	09/04/1908	10/10/1927	livro de emplacamento 11.p	Luanna	12/10/2018
107	165	165-437	436.50	rua lopes de oliveira	437	437	06/01/1936		livro de emplacamento 20.p	Luanna	12/10/2018
108	165	165-437	436.50	rua lopes de oliveira	51	51	11/10/1927	05/01/1936	livro de emplacamento 20.p	Luanna	12/10/2018
109	165	165-597	598.00	rua lopes de oliveira	597	597	06/01/1936		livro de emplacamento 20.p	Luanna	12/10/2018
110	165	165-597	598.00	rua lopes de oliveira	89	89	11/10/1927	05/01/1936	livro de emplacamento 20.p	Luanna	12/10/2018
111	165	165-597	598.00	rua lopes de oliveira	65	65	09/04/1908	10/10/1927	livro de emplacamento 11.p	Luanna	12/10/2018
112	165	165-597	598.00	rua lopes de oliveira	39	39	08/04/1908	08/04/1908	livro de emplacamento 6.p	Luanna	12/10/2018
113	165	165-615	615.90	rua lopes de oliveira	615	615	06/01/1936		livro de emplacamento 20.p	Luanna	12/10/2018
114	165	165-615	615.90	rua lopes de oliveira	93	93	11/10/1927	05/01/1936	livro de emplacamento 20.p	Luanna	12/10/2018
115	165	165-615	615.90	rua lopes de oliveira	69	69	10/10/1927	10/10/1927	livro de emplacamento 11.p	Luanna	12/10/2018

Figura 3 - Amostra da tabela de numeração de imóveis

2.2.4 Determinação da origem das ruas⁸

Como se pode imaginar, a determinação correta de qual extremidade de uma rua é o seu início, no que diz respeito à numeração dos imóveis, é fundamental para o funcionamento adequado do geolocalizador. O referido Código Saboya também definiu critérios para tal, baseados na distância das extremidades das ruas em relação a determinados eixos viários da cidade, em mais uma expressão do seu caráter racionalista. No entanto, verificou-se que esses critérios nem sempre eram atendidos, o que levou à necessidade de um teste controlado e à definição de uma nova metodologia para a determinação do início das ruas. Por um lado, chegou-se à conclusão de que não se poderia fiar no critério legal, dadas as inúmeras exceções encontradas. Por outro, admitiu-se a necessidade de verificação do início de todas as ruas da área-piloto, uma a uma, com base na documentação dos livros de emplacamento. O relatório com a metodologia adotada e os resultados alcançados se encontra no Anexo 3.

⁸ Essa atividade ficou sob a responsabilidade de Monaliza Caetano dos Santos, bolsista Pibic do projeto.

2.2.5 Disciplina de Laboratório de Ensino e Pesquisa em História I

No segundo semestre de 2018, o coordenador do projeto ficou responsável pelas aulas de Laboratório de Ensino e Pesquisa em História I (turma do noturno), disciplina obrigatória da grade curricular do curso de graduação em História da Unifesp e destinada ao conhecimento da metodologia para utilização e análise da documentação histórica. Na ementa do curso (Anexo 4), constava: “Por fim, pretende-se abordar os influxos das novas tecnologias nessas temáticas e as consequentes reconfigurações do ofício do historiador, dentro do contexto mais geral da afirmação das chamadas humanidades digitais”. Para dar conta desse objetivo, os alunos foram instados a trabalhar com geotecnologias, em geral, e com a plataforma Pauliceia 2.0, em particular. Por um lado, isso permitiu uma aplicação prática das discussões envolvendo humanidades digitais mobilizadas no curso; por outro, permitiu a execução de testes e a formulação de críticas e sugestões que colaboraram no aprimoramento da plataforma. Bolsistas do projeto atuaram como monitores do curso, e as aulas tiveram lugar em um dos laboratórios de informática do Campus de Guarulhos da Unifesp, com o suporte do Laboratório de Humanidades Digitais, do mesmo campus, para atendimento dos alunos.

2.2.6 II Ciclo de Oficinas do Grupo Hímaco

De 19 a 27 de abril de 2018, foi realizado o II Ciclo de Oficinas do Grupo Hímaco⁹, nas dependências do Laboratório de Humanidade Digitais da Unifesp (Figura 4). Cadastrado como atividade de extensão da universidade, o ciclo teve como objetivo “capacitar os interessados para colaborar com o Projeto Pauliceia 2.0 (Mapeamento colaborativo da história de São Paulo). As atividades práticas do curso estiveram voltadas a auxiliar no desenvolvimento de um geolocalizador de endereços da cidade de São

9 O grupo Hímaco (História, Mapas e Computadores) é coordenado por Luis Ferla e foi criado em 2010 na Unifesp com o objetivo de explorar as possibilidades das tecnologias digitais no trabalho do historiador. Grande parte da equipe do projeto pertence ao grupo (mais informações em www.unifesp.br/himaco; acessado em 20 de fevereiro de 2019).

Paulo do período que vai de 1870 a 1940, o que foi feito pela alimentação de um banco de dados a partir de fontes históricas. A experiência pretendeu facilitar o contato com as humanidades digitais e com o trabalho colaborativo”.



Figura 4 - II Ciclo de Oficinas do Grupo Himaco, em abril de 2018

2.3 Núcleo “plataforma computacional”

2.3.1. Capacitação e acompanhamento das atividades

Durante o período deste relatório, os bolsistas de computação participaram do Worcap 2018: Workshop de Computação Aplicada, promovido pelo INPE de 21 a 23 de agosto de 2018 (<http://www.inpe.br/worcap/2018/>). Nesse evento, os bolsistas tiveram oportunidades de apresentar e discutir os seus trabalhos, fazer cursos de curta duração e conhecer as pesquisas em computação aplicada desenvolvidas pelo INPE e institutos da região.

O bolsista de Mestrado, Rodrigo M. Mariano, escreveu e apresentou o seu artigo científico (Anexo 5), relacionado a sua dissertação, no evento GEOINFO (Simpósio Brasileiro de Geoinformática), que foi realizado de 5 a 7 de dezembro de 2018, em Campina Grande, na Paraíba. O GEOINFO é uma conferência anual para explorar pesquisas, desenvolvimento e aplicações inovadoras em ciência da informação geográfica e áreas afins. Ele também apresentou uma demonstração do projeto Pauliceia 2.0 ao vivo para os participantes do evento. Para a ida a este evento científico, o aluno usou a sua Reserva Técnica provida pela FAPESP.

Ao longo desse período, foram realizadas reuniões semanais no INPE todas as sexta-feiras e algumas outras durante a semana, para o acompanhamento das atividades dos bolsistas e discussões técnicas relacionadas ao desenvolvimento da plataforma Pauliceia 2.0.

2.3.2. Projeto e desenvolvimento da plataforma computacional

O sistema Pauliceia 2.0 é código aberto, ou seja, o código está disponível online no repositório do Github (<https://github.com/pauliceia>). A plataforma é online e orientado a serviços. A arquitetura que foi desenvolvida é mostrada na Figura 5.



Arquiteturas orientadas a serviços são adequadas para fornecer uma melhor interoperabilidade entre os sistemas.

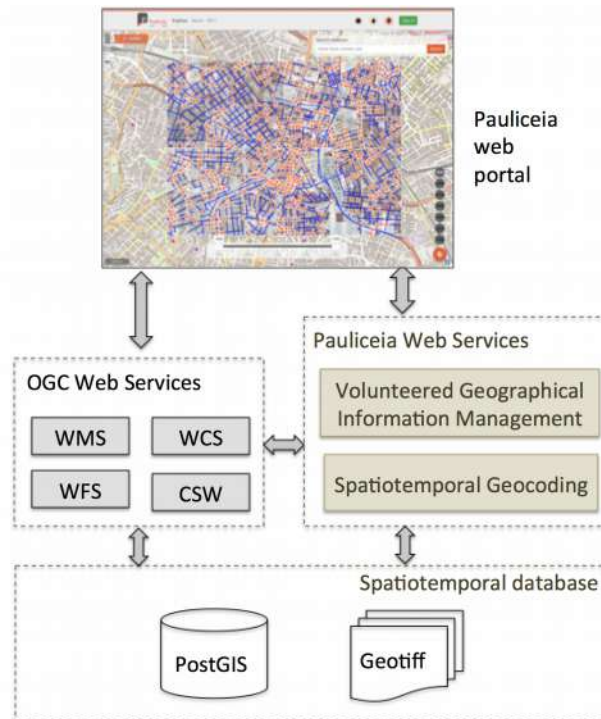


Figura 5: Arquitetura da plataforma computacional Pauliceia 2.0

A plataforma Pauliceia 2.0 foi implementada utilizando a biblioteca de SIG TerraLib e o framework para construção de portais web para dados geográficos TerraBrasilis, ambos desenvolvidos pelo INPE¹⁰.

A arquitetura tem dois grupos de serviços web. O primeiro é composto pelos seguintes serviços web geográficos definidos pela OGC¹¹ (*Open Geospatial Consortium*): *Web Map Service* (WMS), *Web Feature Service* (WFS), *Web Coverage Service* (WCS) e *Catalogue Service Web* (CSW). O grupo OGC tem desempenhado um papel crucial na interoperabilidade de dados geográficos, propondo padrões de serviços web para

10 Câmara, G., Vinhas, L., Ferreira, K. R., Queiroz, G. R. D., Souza, R. C. M. D., Monteiro, A. M. V., Carvalho, M. T. D., Casanova, M. A., and Freitas, U. M. D. (2008). TerraLib: An open source gis library for large-scale environmental and socio-economic applications. Open source approaches in spatial data handling, pages 247–270.

11 <http://www.opengeospatial.org/>.

visualização, disseminação e processamento de dados geográficos. Os serviços da OGC são providos por um GeoServer.

A disseminação dos dados do projeto Pauliceia 2.0 através de serviços web OGC é importante para a interoperabilidade, integração com outras aplicações e compartilhamento de dados. As especificações da Infraestrutura Nacional de Dados Espaciais Brasileira (INDE) são baseadas em serviços web OGC. Portanto, a criação desses serviços na plataforma do projeto permite que os dados históricos do projeto Pauliceia 2.0 sejam disseminados de acordo com as especificações da INDE.

O segundo grupo é composto por dois serviços web que foram projetados e desenvolvidos para estender as funcionalidades dos serviços web OGC, atendendo às demandas específicas do projeto Pauliceia 2.0. Esses serviços são o VGIMWS e o geocoding, que serão descritos nas próximas seções.

Os dados espaço-temporais vetoriais do projeto Pauliceia 2.0 são armazenados em um banco de dados PostgreSQL com extensão espacial PostGIS, enquanto os dados matriciais (i.e. rasters) são salvos em arquivos no formato GeoTIFF.

2.3.2.1 Serviço web VGI de dados históricos

Originalmente o termo VGI foi definido por Goodchild¹² como um conjunto de ferramentas usadas para criar, reunir e disseminar dados geográficos fornecidos voluntariamente pela comunidade. Como o projeto Pauliceia 2.0 é uma plataforma colaborativa, ferramentas VGI são imprescindíveis para a coleta e validação dos dados que são utilizados, visto que eles são fornecidos pela comunidade.

A abordagem colaborativa da coleta dos dados, bem como o próprio VGI, apresentam alguns desafios, entre eles a validação, controle de qualidade e uniformização dos dados. Pensando nesses desafios, foram estudadas e desenvolvidas ferramentas e métodos para o projeto Pauliceia 2.0, que farão a validação dos dados da plataforma através da própria comunidade que utiliza o portal.

12 Goodchild, M. F. (2007). Citizens as sensors: the world of volunteered geography. *Geo-Journal*, 69(4):211–221.

Durante o período ao qual esse relatório se refere, foi finalizado o protocolo VGI para dados históricos e o Serviço Web de Gerenciamento de Informações Geográficas Voluntárias (*Volunteered Geographic Information Management Web Service*, VGIMWS) para dados históricos seguindo esse protocolo, ambos no contexto do projeto Pauliceia 2.0. A explicação em detalhes sobre ambos podem ser encontrados no artigo contido no Anexo 5.¹³

A definição desse protocolo VGI determinou técnicas de coleta de dados e metadados, um modelo de dados para o projeto, e possíveis métricas e formas de controle de qualidade para esses dados obtidos pela comunidade. Métodos para validação e coleta uniformizada dos dados também foram definidos, bem como a organização geral do portal e sua interação com usuário.

Este protocolo VGI tem como principais etapas: (1) a plataforma fornece um tutorial de como usar o sistema, para que os usuários possam lê-lo para resolver as principais questões. Este tutorial pode ser encontrado no Anexo 6; (2) os voluntários podem criar camadas sobre seus dados de pesquisa, manualmente ou por importação em massa; (3) os colaboradores podem adicionar pessoas de confiança em suas camadas para auxiliá-los na coleta de dados, gerando uma qualidade local; (4) usuários fora das camadas, que encontram problemas nos dados, podem fazer comentários e observações através das notificações; (5) historiadores podem promover HistMapathons para grande coleta de dados de vetorização de mapas históricos; (6) o usuário pode receber notificações relevantes do sistema; (7) se os voluntários descobrirem dados inadequados na plataforma, podem informá-los por meio de denúncias; e (8) os voluntários podem fornecer feedback de sua experiência nos canais disponíveis do projeto. O protocolo VGI para dados históricos aplicados ao projeto Pauliceia 2.0, desenvolvido aqui, pode ser aplicado a outros projetos históricos VGI existentes que não possuem dados autoritativos para comparação. Portanto, é possível usar as técnicas de indicadores de qualidade e abordagens descritas aqui e aplicá-las a outros projetos VGI.

13 Mariano, R. M.; Ferreira, K. R.; Ferla, L. A. C. VGI Protocol and Web Service for Historical Data Management. Proceedings XIX GEOINFO, December 05-07, 2018, Campina Grande, PB, Brazil. p 103-115.

O VGIMWS é um serviço Web RESTful para lidar com dados espaço-temporais. Esta API foi desenvolvida com base no protocolo VGI para dados históricos. O formato padrão de troca de dados do serviço web é o GeoJSON ou JSON, para trabalhar com dados com informação geográfica ou sem, respectivamente. Ele foi desenvolvido em linguagem Python, usando o framework web chamado Tornado. Cada recurso do VGIMWS pode ser encontrado através de uma URL única.

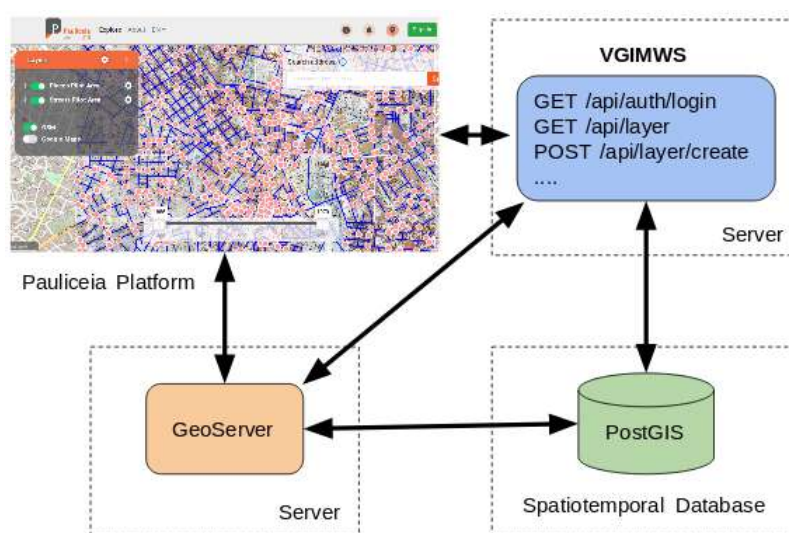


Figura 6: Arquitetura simplista do projeto Pauliceia 2.0, relacionado ao VGIMWS

A Figura 6 mostra o processo de uso da API. Um usuário acessa o portal Pauliceia 2.0 através de um navegador. Assim, um componente front-end dentro da plataforma chamará o VGIMWS para gerenciar os dados, através de funções bem definidas. Finalmente, o VGIMWS armazenará essas informações em um banco de dados espaço-temporal. Por exemplo: quando um usuário cria uma nova camada, uma função específica é chamada para armazenar as informações dentro do banco de dados e o VGIMWS publicará essa camada no GeoServer. Em seguida, o portal Pauliceia 2.0 pode traçar os recursos no mapa usando o GeoServer, através do OpenLayers.

A documentação de como executar o serviço e de como usar as suas funcionalidades pode ser encontrada no Github do projeto: <https://github.com/Pauliceia/vgiws/blob/master/doc/README.md>.

2.3.2.2 Serviço web para geocodificação de dados históricos

Uma característica crucial da plataforma Pauliceia 2.0 é fornecer funcionalidades que permitam aos historiadores compartilhar dados geográficos históricos resultantes de suas pesquisas. Nesse caso, a maioria dos conjuntos de dados históricos possui endereços textuais para indicar localizações espaciais no passado. Assim, é necessário fornecer um algoritmo de geocodificação capaz de transformar endereços textuais históricos em coordenadas geográficas.

Geocodificação é o processo de transformar dados textuais em informação geográfica. A obtenção de coordenadas a partir de endereços textuais é um dos mais importantes métodos de geocodificação¹⁴. Martins et. al. propõem um método de geocodificação para endereços urbanos cuja produção inclui um indicador de certeza geográfica, que informa a qualidade esperada dos resultados.

Na literatura, existem muitas propostas de geocodificadores eficientes para endereços atuais, mas eles não lidam com dados históricos. Um geocodificador para informações históricas deve operar em conjuntos de dados espaço-temporais, isto é, entidades espaciais cujas geometrias e atributos variam ao longo do tempo. Os desafios da criação de um sistema de geocodificação de endereços para dados históricos estão relacionados principalmente à variação de nomes, geometrias e numerações de ruas e edifícios ao longo do tempo. No banco de dados do Pauliceia 2.0, cada entidade espacial, como um segmento de rua e um local com um endereço, possui um período associado que indica quando é válido. Assim, o método de geocodificação para esse

¹⁴ Martins, D., Davis Jr, C. A., and Fonseca, F. T. Geocodificação de endereços urbanos com indicação de qualidade. Proceedings of XIII Brazilian Symposium on Geoinformatics (GeoInfo 2012), 2012.

banco de dados deve levar em consideração todos os períodos válidos associados a entidades espaciais.

O serviço web de geocodificação que foi projetado no projeto Pauliceia 2.0 considera todas essas particularidades dos conjuntos de dados históricos. Usando esse serviço, os historiadores podem geocodificar um único endereço ou um conjunto de endereços por meio de arquivos CSV. Cada endereço deve conter o nome da rua, o número e o ano. O serviço calcula as coordenadas geográficas associadas aos endereços usando os locais históricos e os segmentos de rua armazenados no banco de dados do projeto. Além das coordenadas geográficas, o serviço retorna um grau de certeza associado a cada coordenada. Esse grau indica a confiança de uma coordenada geográfica, com base no número de entidades históricas disponíveis que foram usadas no processo de geocodificação.

Para preencher o banco de dados do projeto, primeiramente foi criado um portal web para a edição de endereços históricos, onde neste portal fornecia funcionalidades para inserir, excluir, editar e pesquisar endereços históricos. Neste período optou-se pelo uso de uma tabela formatada, conforme já indicado acima. Esta tabela formatada é uma tabela no formato CSV que contém os endereços históricos distribuído nas linhas, enquanto as colunas descrevem os seus atributos. Nela os membros do projeto coletam e inserem endereços históricos da cidade de São Paulo de 1870 a 1940, o que posteriormente é importado no banco de dados do projeto. Cada endereço tem um nome de rua, um número de local e um período em que é válido.

O serviço de geocodificação depende dos endereços históricos e segmentos de rua armazenados no banco de dados do projeto. Assim, a construção de um banco de dados de boa qualidade é crucial para o sucesso do processo de geocodificação. Quanto maior o número de endereços identificados, maior a precisão da base resultante. Os membros do projeto estão usando diferentes tipos de fontes históricas, como documentos legislativos, jornais, livros de matrículas e folhetos publicitários, para coletar esses endereços. A documentação mais importante é constituída pelos livros de emplacamento do município, como indicado mais acima.

O processo de geocodificação de endereços é mostrado na Figura 7. No primeiro quadro (i.e. A), as ruas são mostradas como feições lineares, com um nome e um intervalo de validade. Os lugares conhecidos são representados por características pontuais descritas por vários atributos, como nome, localização e intervalo de validade. A relação entre cada local conhecido e sua rua também se materializa no banco de dados. Portanto, para geocodificar um endereço como "Rua Augusta, 50, 1937", uma rua no período chamado "Augusta" é pesquisada primeiro. Então os lugares associados a esta rua, cujo intervalo de tempo contém o ano de "1937" são recuperados (i.e. quadro B). Finalmente, uma interpolação linear do endereço pesquisado é calculada tomando a geometria linear da rua encontrada (i.e. quadro C). Embora ilustrada como um procedimento simples, a geocodificação tem que lidar com alguns casos especiais, como endereços cujas ruas digitalizadas não possuem locais conhecidos associados ao intervalo de tempo solicitado ou cuja numeração está fora do intervalo definido pelos locais conhecidos.

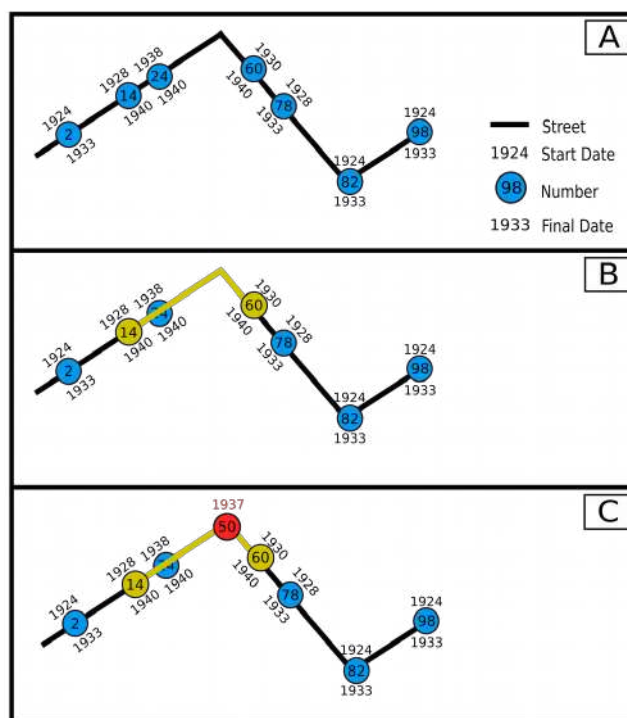


Figura 7: Geocoding de endereços históricos.

2.3.3. Arquitetura Docker da plataforma computacional Pauliceia 2.0

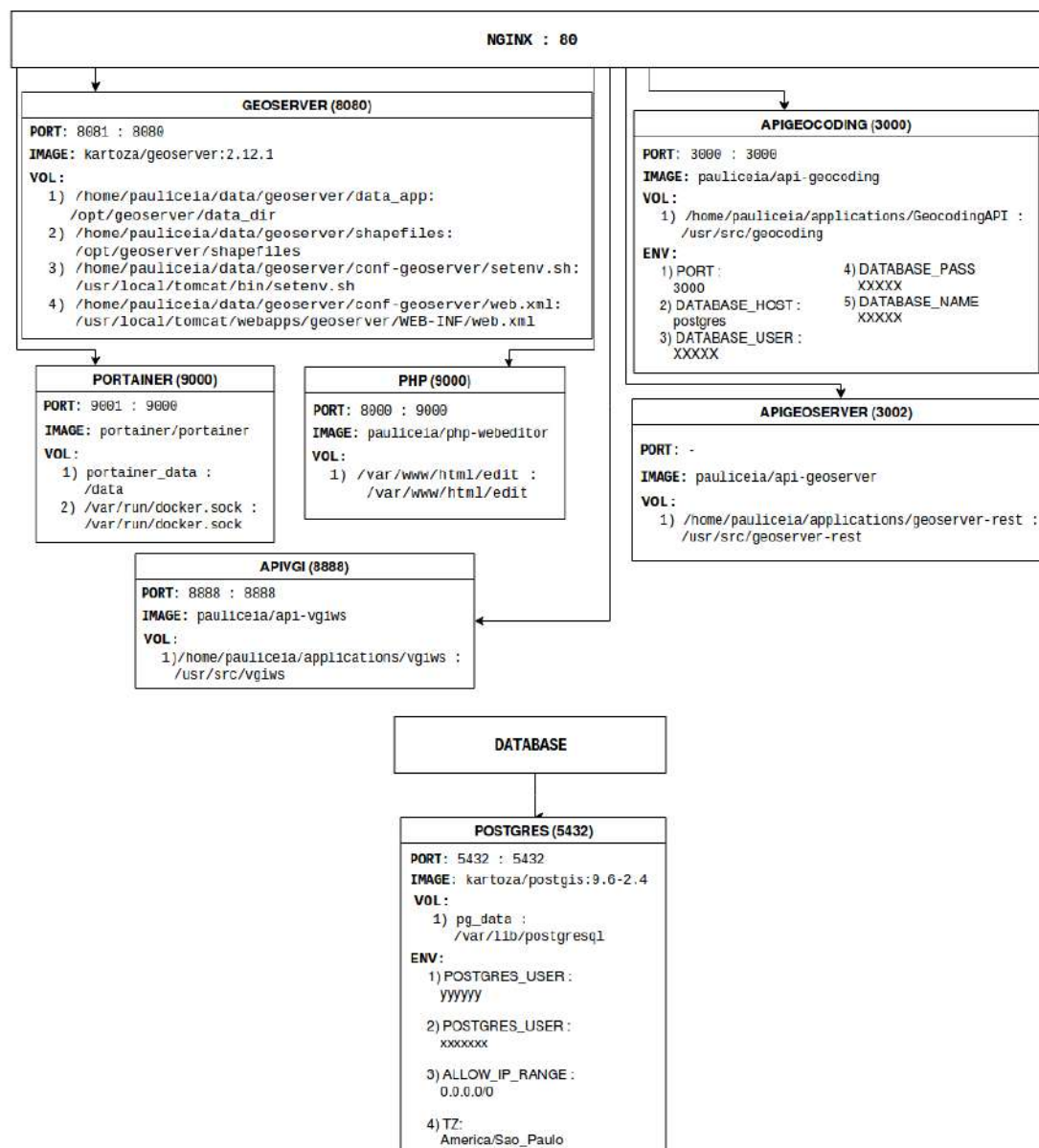


Figura 8: Arquitetura Docker da plataforma computacional Pauliceia 2.0

Toda a plataforma Pauliceia 2.0 é dividida em contêineres Docker, onde cada contêiner executa um serviço específico, como mostrado pela Figura 5. Todos os

arquivos do projeto, como portal web, serviços web e dados espaço-temporais, estão salvos em um servidor hospedado no INPE.

A Figura 8 ilustra a arquitetura interna do servidor Pauliceia 2.0 usando contêineres Docker. Há um servidor de proxy reverso usando o Nginx, que escuta a porta 80 e não está dentro de um contêiner. O Nginx guia os pedidos feitos pelo cliente para os serviços de back-end adequados, como VGIMWS, GeoServer, API de Geocoding e assim por diante. Por exemplo: um cliente lista as camadas disponíveis através do portal e, em seguida, a plataforma faz uma solicitação para a seguinte URL: <http://www.pauliceia.dpi.inpe.br/api/vgi/api/layer>. O Nginx obtém esse pedido e envia para o serviço apropriado, que é o VGIMWS, que está sendo executado na porta 8888. Cada contêiner Docker na Figura pode ter um nome com a porta do contêiner, uma propriedade “PORT” que define as portas ligadas (host:contêiner), um atributo “IMAGE” que é a imagem usada, um campo “VOL” que são os volumes (host:contêiner) e “ENV” que é um metadado com uma lista de variáveis de ambiente.

APIVGI é o nome do contêiner do Docker que executa o VGIMWS. A porta do host exposta é o 8888 e a porta do contêiner é a mesma (i.e. “8888:8888”, ambas as portas estão vinculadas), isso significa que o Nginx envia as solicitações para a porta 8888 no host e redireciona para a porta 8888 no contêiner. A imagem do Docker utilizada é a “pauliceia/api-vgiws”. Há um volume (i.e. uma pasta dentro do servidor que armazena dados produzidos por contêineres Docker) que usa a pasta host “/home/pauliceia/applications/vgiws” vinculada à pasta do contêiner “/usr/src/vgiws”. As imagens do Docker do Pauliceia 2.0 são armazenadas no Docker Hub: <https://hub.docker.com/u/pauliceia>, enquanto os Dockerfiles estão salvos no Github: <https://github.com/pauliceia/dockers-imagens>.

O contêiner GEOSERVER contém um GeoServer. O GeoServer é um servidor baseado em Java de código aberto para gerenciar dados geográficos. Ele usa os padrões definidos pela OGC, como WMS, WFS e entre outros. O GeoServer administra os dados espaciais do banco de dados do Pauliceia 2.0, como plotar os mapas históricos na plataforma.



O PORTAINER é um contêiner que consiste em um aplicativo do Portainer. O Portainer é um site da Web para monitorar contêineres, volumes, imagens e redes do Docker. O Portainer lida com os contêineres do servidor Pauliceia 2.0, em uma interface de usuário amigável.

O container PHP possui a primeira versão do portal Pauliceia 2.0, chamada Edit ou Web Editor. APIGEOCODING é um contêiner que inclui o serviço web de geocodificação. Este serviço faz a geocodificação dos endereços históricos. O contêiner APIGEOSERVER contém um middleware entre o GeoServer e o VGIMWS. O VGIMWS se comunica com o GeoServer através deste serviço, a fim de publicar e cancelar a publicação de uma tabela de recursos (i.e. camada) do banco de dados do Pauliceia 2.0. O POSTGRES é um contêiner que armazena os dados vetoriais da plataforma Pauliceia 2.0 usando um banco de dados PostgreSQL com extensão PostGIS.

2.3.4. Projeto e desenvolvimento da plataforma Pauliceia 2.0

Durante o período ao qual esse relatório se refere, foi projetado e desenvolvido a plataforma Pauliceia 2.0, que disponibiliza um conjunto de funcionalidades para inserção, edição, exclusão e consulta de dados históricos providos pelos colaboradores. A interface de mapa da plataforma Pauliceia 2.0 é mostrada na Figura 9 e ele está disponível no endereço <http://www.pauliceia.dpi.inpe.br>.

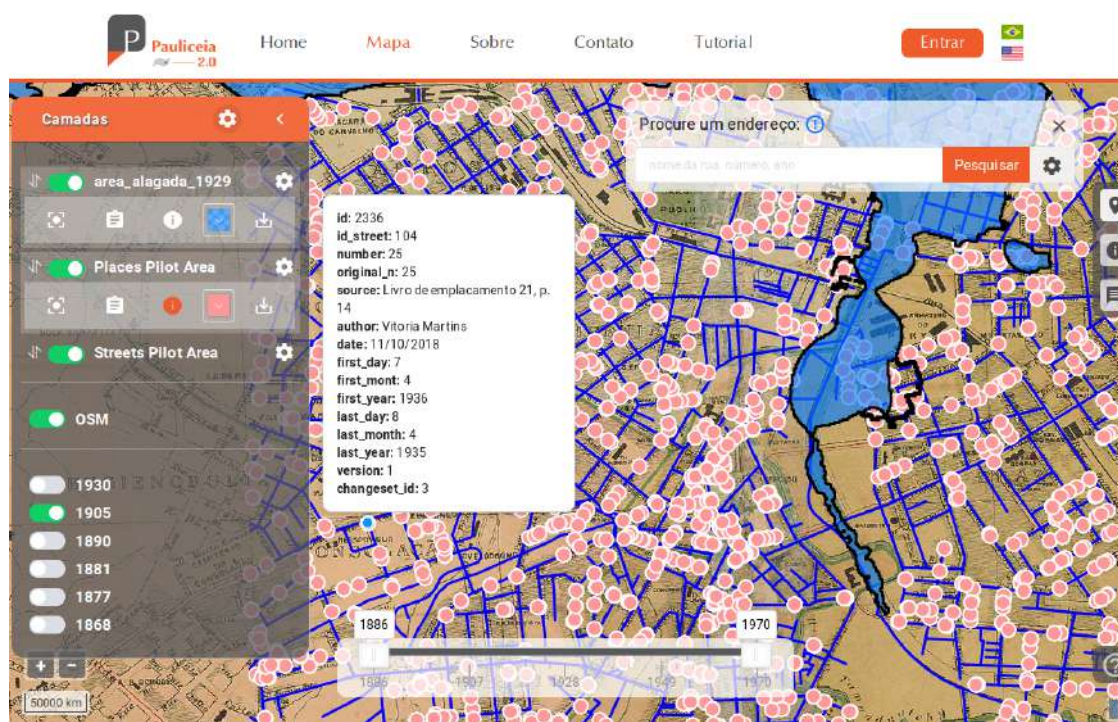


Figura 9: Página do mapa da plataforma computacional Pauliceia 2.0

Através desse portal, os pesquisadores do projeto ou qualquer interessado podem inserir no banco de dados históricos dados espaço-temporais fornecidos por suas pesquisas históricas sobre a cidade de São Paulo no período de 1870 a 1940, tais como casas, hotéis, igrejas, fábricas, praças e entre outros.

A plataforma possui uma interface amigável e intuitiva. Sua manutenção e aprimoramento são feitos de acordo com as necessidades dos pesquisadores envolvidos com o projeto Pauliceia 2.0. Para uma melhor organização e controle dos dados inseridos, o usuário necessita de uma autenticação ao iniciar suas atividades no sistema. Para isso, foram estabelecidos, até o presente momento, três níveis de usuários: normal, curador e administrador. Os normais podem adicionar ou editar os dados históricos, produzindo camadas relacionadas à sua pesquisa histórica associada a palavras-chave ou referências. Isso pode ser feito manualmente ou por importação em massa. Os curadores são aqueles que promovem a articulação de certas palavras-chave e espaço geográfico dentro da plataforma. Eles podem atualizar as informações da camada de qualquer

usuário e organizá-las, criando novas palavras-chave para a camada. Os administradores podem inserir, atualizar e excluir todas as entidades da plataforma Pauliceia 2.0, gerenciando também as denúncias feitas pelos usuários.

A Figura 9 mostra a página do mapa da plataforma computacional Pauliceia 2.0. Ao fundo aparece ativado o mapa da área central de São Paulo de 1905. Quando logado, o usuário pode inserir camadas de dados vetoriais, sendo que a Figura mostra três camadas que foram alimentadas por algum usuário, onde: (1) 'area_alagada_1929' é uma camada de polígonos, que indica a área alagada de São Paulo em 1929 (i.e. polígono azul); (2) 'Places Pilot Area' é uma camada de pontos, que ilustra os endereços históricos da área central de São Paulo em 1930 (i.e. pontos rosas); e (3) 'Streets Pilot Area' é uma camada de linhas, que mostra as ruas da área central de São Paulo em 1930 (i.e. linhas azuis).

Quando ativada uma camada no sistema, há um ícone de uma engrenagem ao lado do nome da camada, com o qual o usuário pode: (1) dar zoom geral da camada, visualizando-a na tela inteira; (2) visualizar as informações da camada; (3) obter a informação de uma feature específica; (4) editar a visualização da camada (e.g. mudando suas cores); e (5) baixar o Shapefile daquela camada, após o que o usuário poderá abri-la novamente em algum SIG (e.g. QGIS, gvSIG ou TerraView).

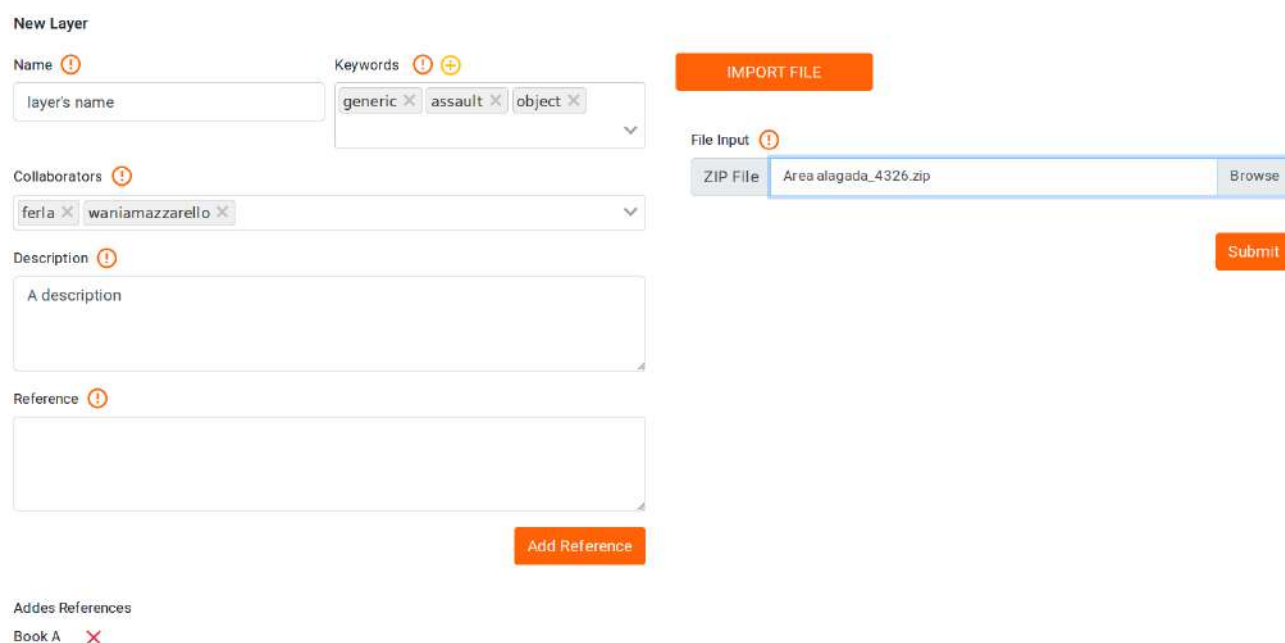
No menu lateral à direita, o usuário pode pesquisar por algum endereço histórico, selecionar informações de uma feature ou um conjunto dela, ou o voluntário pode visualizar as notificações relacionadas aquela camada. O usuário pode pesquisar por algum endereço histórico usando o padrão "rua, número, ano" na barra de pesquisa, através do serviço web de geocoding. O voluntário também pode clicar na engrenagem ao lado da barra de pesquisa e pesquisar uma lista de endereços históricos adicionados dentro de um arquivo CSV.

Na parte inferior do mapa, há o componente *Slider*, que é um visualizador de dados espaço-temporais. Ele é utilizado junto ao mapa, tendo a função de filtrar o intervalo de tempo desejado pelo usuário, através de um tempo inicial e final definido no portal. Dessa forma, o usuário poderá escolher o intervalo de tempo que ele deseja



trabalhar, não ficando limitado a um único ano. Além disso, o *Slider* permite fazer uma animação dos dados históricos ao longo do tempo.

Um colaborador pode inserir uma nova camada no sistema, se ele quiser. Primeiramente ele precisa se autenticar na plataforma, depois ele irá em dashboard e por fim em criar nova camada. A Figura 10 ilustra a página de criação de uma nova camada, onde o usuário irá preencher com as informações daquela camada, tais como nome, lista de colaboradores, lista de palavras-chave, descrição, lista de referências e um arquivo zip contendo um Shapefile. O Shapefile é um formato padrão bem-definido para armazenamento de dados vetoriais geográficos.



The screenshot shows the 'New Layer' form in the Pauliceia 2.0 platform. The form is divided into several sections:

- Name:** A text input field with the placeholder 'layer's name'.
- Keywords:** A list of tags including 'generic', 'assault', and 'object'.
- Collaborators:** A list of names including 'ferla' and 'waniamazzeello'.
- Description:** A large text area with the placeholder 'A description'.
- Reference:** A large text area for adding references.
- Import File:** A section with an 'IMPORT FILE' button and a 'File Input' field containing 'Area alagada_4326.zip' and a 'Browse' button.
- Submit:** An orange button at the bottom right.
- Add Reference:** An orange button at the bottom center.
- Addes References:** A section at the bottom left with a 'Book A' entry and a red 'X' icon.

Figura 10: Página de criação de uma nova camada na plataforma Pauliceia 2.0

Após inserir uma nova camada, será solicitado para o usuário preencher os dados temporais daquela camada específica, como a caixa delimitadora temporal (i.e. o período que compõe aquela camada), qual campo do Shapefile contém o a data inicial e a data final, e a máscara usada naquele campo. Essas informações temporais são

importantes para que o *Slider* espaço-temporal funcione corretamente. A Figura 11 exibe esta tela.

Temporal Columns







Start date 	Start date column 	Start date mask 
<input type="text" value="01 / 01 / 1900"/>	<input type="text" value="result"/>	<input type="text" value="YYYY-MM-DD"/>
End date 	End date column 	End date mask 
<input type="text" value="01 / 01 / 1930"/>	<input type="text"/>	<input type="text"/>
<input type="button" value="Submit"/>		

Figura 11: Página de inserção de colunas temporais na plataforma Pauliceia 2.0

Depois que um usuário insere uma nova camada, ela estará disponível na página principal para a ativação dela no mapa e para atualização/remoção dela no dashboard , como mostrado pelas Figuras 12 e 13, respectivamente.



Figura 12: Listando camadas disponíveis na plataforma Pauliceia 2.0

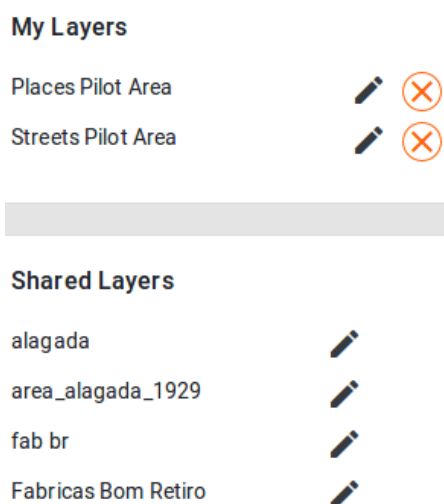


Figura 13: Página de atualização/remoção de uma camada na plataforma Pauliceia 2.0

Os voluntários podem escrever comentários e observações nas camadas disponíveis na plataforma, através das notificações. As notificações são comentários que um usuário recebe relacionado a uma camada, como feedbacks. Elas podem ser visualizadas na página de dashboard, como mostrado na Figura 14, ou na página do

mapa. Um usuário pode ainda enviar uma notificação geral para todos os membros da comunidade Pauliceia 2.0, como o anúncio de um evento científico.


Notificações

GENERAL

PERSONAL

FOLLOWING

Submit

**Pauliceia Admin**
2019-02-10 18:17:32

Welcome to the Pauliceia 2.0 platform.

Comments




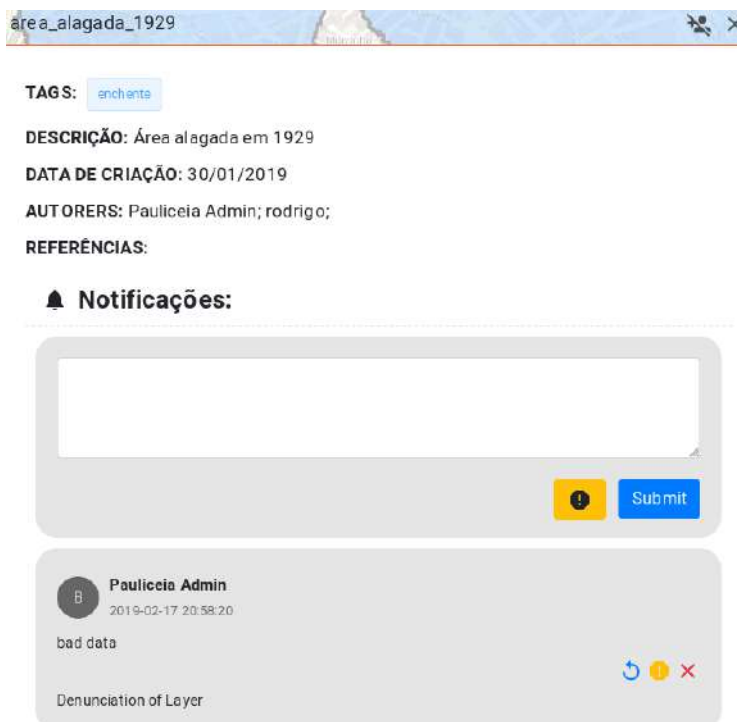
  

Figura 14: Página de gerenciamento de notificações da plataforma Pauliceia 2.0



area_alagada_1929

TAGS: enchente

DESCRIÇÃO: Área alagada em 1929

DATA DE CRIAÇÃO: 30/01/2019

AUTORES: Pauliceia Admin; rodrigo;

REFERÊNCIAS:

Notificações:

bad data

Denunciation of Layer

Figura 15: Página de denúncia de uma camada da plataforma Pauliceia 2.0

Se um usuário encontrar algum dado com conteúdo inapropriado (e.g. dados com direitos autorais), ele deve reportar isto através das denúncias. As denúncias são um tipo de notificação em que o colaborador do portal Pauliceia 2.0 pode reportar problemas ou situações inadequadas que ele encontrar. Isto é mostrado pela Figura 15. O usuário descreve o problema na caixa de texto e clica no botão amarelo para enviar sua denúncia.

3. Apoio institucional recebido no período¹⁵

O projeto foi desenvolvido nas dependências das quatro instituições parceiras: a UNIFESP (campi Guarulhos e São José dos Campos), o INPE, a Emory University e o Arquivo do Estado de São Paulo.

No campus de Guarulhos da Unifesp está localizada a Escola de Filosofia, Letras e Ciências Humanas. No primeiro semestre de 2016 foram inauguradas suas novas instalações, e um ano depois terminadas as reformas das antigas. O novo edifício acadêmico conta com “cerca de 20 mil metros quadrados de área útil, com 38 salas de aula, 14 laboratórios de pesquisas e estudos, biblioteca com mais de 100 mil exemplares, anfiteatro com capacidade para 150 lugares, restaurante universitário e sistemas inteligentes de automação e de captação de águas pluviais”.¹⁶ Em uma de suas salas encontra-se o Laboratório de Humanidades Digitais, espaço compartilhado por cinco grupos de pesquisas da EFLCH, o Hímaco dentre eles. Ali existem doze estações de trabalho, sendo seis em uso pelo Hímaco, com equipamentos adquiridos por financiamentos de projetos anteriores do grupo. Grande parte das atividades do projeto em questão vem sendo desenvolvida nesse laboratório. Além disso, o projeto contou com o apoio de duas secretárias do Departamento de História e dos técnicos do Setor de Informática para a resolução de problemas atinentes ao laboratório. O Setor de Transporte, por sua vez, deu todo o suporte necessário para viabilizar o deslocamento de membros da equipe entre as instituições parceiras. Houve uso intenso da biblioteca da unidade pelos participantes do projeto. Ainda com relação à EFLCH, deve ser registrada a concessão de uma bolsa PIBIC/CNPq a uma aluna com um projeto relacionado ao Pauliceia 2.0.

O INPE disponibiliza para o projeto salas de trabalho com computadores, internet e telefones para os bolsistas que executam atividades relacionadas ao desenvolvimento

¹⁵ Item escrito com base no que foi informado no relatório parcial, apenas com as atualizações necessárias.

¹⁶ Um documento descritivo da infraestrutura de pesquisa da EFLCH/Unifesp pode ser acessado em: www.humanas.unifesp.br/pgpq_elfch/a-elfch-unifesp-e-sua-infraestrutura-para-pesquisa (acessado em 13 de janeiro de 2019).

da plataforma computacional. Esses bolsistas podem utilizar todas as facilidades do INPE como bibliotecas e salas de reuniões. Além disso, o INPE forneceu o transporte da equipe para reuniões do projeto que aconteceram em São Paulo e Guarulhos. O INPE também disponibiliza sua infraestrutura computacional, como por exemplo servidores web, para hospedar o banco de dados do projeto Pauliceia e os sistemas web que estão sendo desenvolvidos.

A Emory University participou da equipe do projeto com três dos seus professores (Jeffrey Lesser, Michael Page e Thomas Rogers) e um doutorando (Andrew Britt). A instituição recebeu Luis Ferla em janeiro de 2019, quando houve a oportunidade de discutir o andamento do projeto e as possibilidades da continuidade da parceria.

O Arquivo Público do Estado, que tem sua participação no projeto disciplinada por um termo de convênio com a Unifesp para esse fim¹⁷, disponibilizou para o uso da equipe um laboratório com duas estações de trabalho e uma mesa de reuniões, nas dependências do Núcleo de Acervo Cartográfico. O espaço, além de abrigar várias das atividades da equipe e muitas de suas reuniões, também serviu para facilitar o contato com pesquisadores interessados em colaborar com o projeto. Além disso, o Apesp sediou e ajudou a divulgar e organizar o evento de apresentação da versão beta da plataforma, realizado em 30 de outubro de 2018 (ver item 2.1.2). Por outro lado, a instituição tem a guarda de um acervo documental, principalmente cartográfico, de interesse do projeto.

17 O termo de convênio foi assinado em abril de 2016 e renovado em abril de 2018 (publicação do despacho no Diário Oficial do Estado disponível em https://www.imprensaoficial.com.br/DO/BuscaDO2001Documento_11_4.aspx?link=%2f2018%2fexecutivo%2520secao%2520i%2fabril%2f04%2fpag_0001_9OUOK1RRG4GQAe2QJ4PMVKOG9RA.pdf&pagina=1&data=04/04/2018&caderno=Executivo%20I&paginaordenacao=100001; acessado em 13 de janeiro de 2019).

4. Lista de publicações e participações em eventos

4.1 A Plataforma Pauliceia 2.0

De uma certa forma, a publicação mais importante do projeto é a própria plataforma Pauliceia 2.0, no ar desde outubro de 2018. Desde então, como explicado acima, ela se encontra em fase de testes e vem sendo sistemática e produtivamente aprimorada. O acesso é feito pelo link: www.pauliceia.dpi.inpe.br. A explicação das suas principais funcionalidades está no tutorial respectivo, acessado pelo mesmo portal (ou pelo link: <https://drive.google.com/file/d/1KT9NTIFYM4u-siL5F88NIg255VjSgl2I/view>).

4.2 Apresentações em conferências internacionais

Ferla, Luis; Ferreira, Karine Reis; Vijaykumar, Nandamudi; Miyasaka, Cristiane; Musa, Daniela; Atique, Fernando; Queiroz, Gilberto Ribeiro de; Lesser, Jeffrey; Page, Michael; Rogers, Thomas; Britt, Andrew; Meireles, Priscila; Noronha, Carlos Alberto; Almeida, Cintia; Taveira, Denis; Reis, Ester; Sansigolo, Gabriel; Dardi, Ivan; Diniz, Jaíne; Nascimento, Luanna; Caetano, Monaliza; Mariano, Rodrigo; Yamamoto, Janaina; Guarnieri, Orlando. Pauliceia 2.0: Mapeamento Colaborativo da História de São Paulo (1870-1940). I Congresso Internacional em Humanidades Digitais. Rio de Janeiro, Brasil, abril de 2018.

Miyasaka, Cristiane, Ferla, Luis; Ferreira, Karine Reis; Queiroz, Gilberto Ribeiro de; Almeida, Cintia Rodrigues; Nascimento, Luanna Mendes; Diniz, Jaíne; Caetano, Monaliza; Guarnieri, Orlando; Noronha, Carlos Alberto; Sansigolo, Gabriel. Desafios para a construção de um geolocalizador para dados históricos: o caso de São Paulo. I Congresso Internacional em Humanidades Digitais. Rio de Janeiro, Brasil, abril de 2018.

Nascimento, Luanna; Reis, Ester; Rodrigues, Cintia; Ferla, Luis. História e o uso da tecnologia: as experiências de um grupo de pesquisa nas Humanidades Digitais. I Congresso Internacional em Humanidades Digitais. Rio de Janeiro, Brasil, abril de 2018 (pôster).

Caetano, Monaliza. Tutorial gvSIG Aplicado aos Estudos Históricos: possibilidades e avanços. I Congresso Internacional em Humanidades Digitais. Rio de Janeiro, Brasil, abril de 2018 (pôster).

Ferla, Luis; Ferreira, Karine Reis. Pauliceia 2.0: uma plataforma espaço temporal para a história de São Paulo. 14º Congresso Internacional da Associação de Estudos Brasileiros (BRASA). Rio de Janeiro, Brasil, julho de 2018.

4.3 Apresentações em conferências nacionais

Mariano, R. M. ; Ferreira, K. R. ; Ferla, Luis . VGI Protocol and Web Service for Historical Data Management. In: GeoInfo 2018 - XIX Brazilian Symposium on Geoinformatics, 2018, Campina Grande. Proceedings XIX GEOINFO, p. 103-115, 2018.

Mariano, Rodrigo. Demonstration of Pauliceia 2.0 Computational Platform. Brazilian Symposium on Geoinformatics - GEOINFO, 2018.

Mariano, Rodrigo. VGI Management Web Service for Historical Data. WORCAP, 2018.

Rodrigues, Cintia; Nascimento, Luanna. Pauliceia 2.0: Mapeamento colaborativo da história de São Paulo (1870-1940). IV Congresso acadêmico da Unifesp. 2018.

Ferla, Luis. Pauliceia 2.0: mapeamento colaborativo da história de São Paulo (1870-1940). Colóquio Histórias de São Paulo: construções e desconstruções, Guarulhos, agosto de 2018.

Ferla, Luis. Pauliceia 2.0: Mapeamento colaborativo da história de São Paulo (1870-1940). 24 Encontro Estadual de História, Guarulhos, setembro de 2018.

4.4 Trabalhos publicados em anais de conferências nacionais

Mariano, R. M. ; Ferreira, K. R. ; Ferla, Luis . VGI Protocol and Web Service for Historical Data Management. In: GeoInfo 2018 - XIX Brazilian Symposium on GeoInformatics, 2018, Campina Grande. Proceedings XIX GEOINFO, p. 103-115, 2018.

4.5 Trabalhos aceitos para publicação em revista

Ferreira, Karine; Ferla, Luis; Queiroz, Gilberto de; Vijaykumar, Nandamudi; Noronha, Carlos; Mariano, Rodrigo; Wassef, Yasmin; Taveira, Denis; Dardi, Ivan; Sansigolo, Gabriel; Guarnieri, Orlando; Musa, Daniela; Rogers, Thomas; Lesser, Jeffrey; Page, Michael; Britt, Andrew; Atique, Fernando; Yamamoto, Janaína; Moraes, Diego; Miyasaka, Cristiane; Almeida, Cintia; Nascimento, Luanna; Diniz, Jaíne; Caetano, Monaliza. *A Platform for Collaborative Historical Research based on Volunteered Geographical Information*. (aceito para publicação em 2019 na Journal of Information and Data Management).

5. Perspectivas para o futuro

O escopo espacial do projeto a que se refere esse relatório foi limitado a uma área-piloto, correspondente à cidade de São Paulo indicada na carta de 1868 (de Carlos Frederico Rath), dados os desafios técnicos e metodológicos envolvidos.¹⁸ Para que não haja solução de continuidade da plataforma, a equipe do projeto está preparando uma nova submissão ao programa eScience, da Fapesp, com o objetivo de expandir o recorte espacial para toda a mancha urbana da cidade, respeitando o mesmo recorte temporal (1870-1940). As instituições parceiras e grande parte da equipe atual já manifestaram interesse em continuar no projeto.

Além da expansão espacial do escopo da pesquisa, o projeto da fase 2 terá mais dois objetivos que são definidos pelo que a primeira fase desenvolveu: o amadurecimento tecnológico da infraestrutura computacional da plataforma; e o aprofundamento do engajamento da comunidade-alvo da plataforma, os pesquisadores da história da cidade de São Paulo. Além disso, da etapa anterior permanecem como objetivos a capacitação da equipe do projeto e a difusão dos resultados alcançados.

18 Ver item 2.1 do Projeto.

Anexos

Anexo 1 – Design gráfico



Relatório de Atividades - 2018/2019 - Design Gráfico

Dando continuidade às criações iniciadas em 2017, os elementos de Identidade Visual foram concluídos em meados de março de 2018, contemplando a marca do Projeto “Pauliceia 2.0”, o papel timbrado e demais ícones, visando principalmente a legibilidade para a plataforma web. Tais elementos de *branding* foram fundamentais para estabelecer uma unidade de apresentação da plataforma, como a paleta de cores, as tipografias e os ícones gráficos que compuseram todo o conjunto digital.



Marca:



Imagem 1: símbolos de construção da Marca do projeto “Pauliceia 2.0”



Por sua vez, a interface digital desenvolveu-se em duas etapas concomitantes: a elaboração de um Portal de apresentação do projeto, contendo as informações gerais, dados da equipe envolvida no processo, assim como subsídios para contato e suporte aos interessados; e a criação da Plataforma de interação do mapa, com seus meios de acesso através de barras laterais e pictogramas que pudessem tornar a atividade amigável e didática aos futuros usuários e pesquisadores. Estes layouts gráficos foram produzidos dentro do contexto de HTML/CSS, compatível com a maioria dos navegadores e dispositivos de acesso a internet atualmente, adotando como base inicial a resolução de tela com 1600 x 900 pixels.

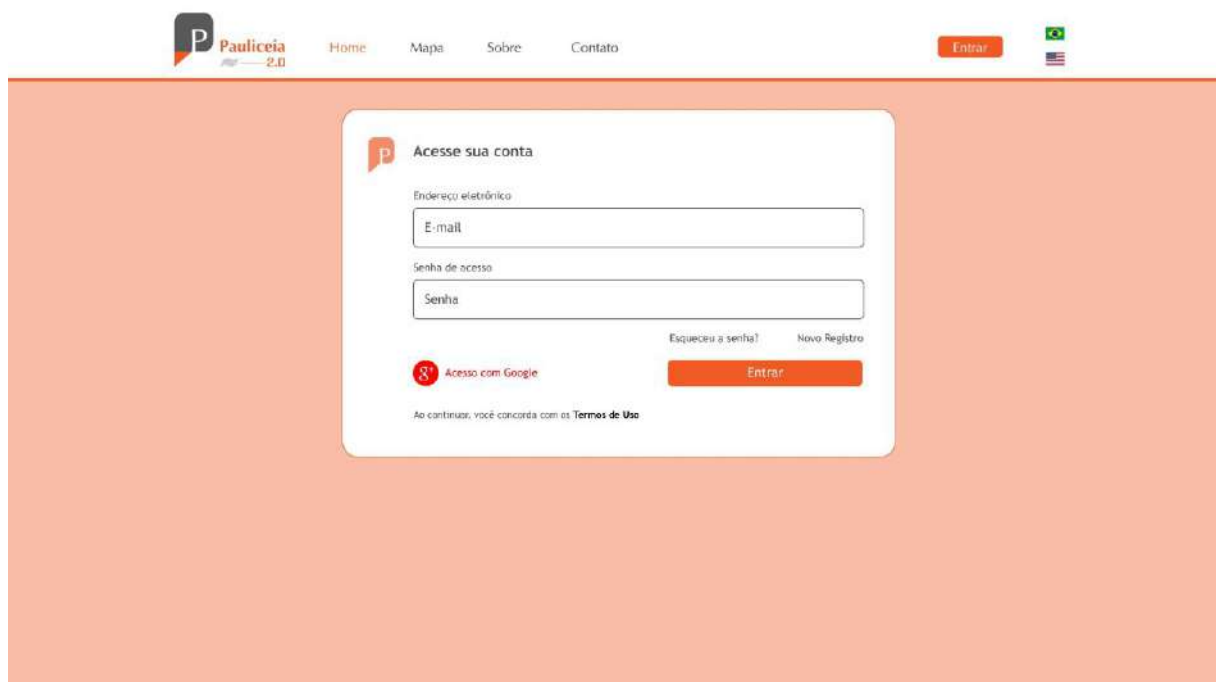
Todos os elementos visuais, tanto do Portal quanto de interação da Plataforma, foram concluídos no final do mês agosto de 2018 objetivando a apresentação da versão Beta de testes do projeto, que ocorreu ao final de outubro do mesmo ano. Nesta etapa de criação definiram-se as quatro páginas de acesso e respectivos conteúdos, que foram afixados no cabeçalho disponível na borda superior da tela, intitulados “Home”, “Mapa”, “Sobre” e “Contato”, assim como foram desenvolvidos os pictogramas e botões de interação com o mapa, concentrados nas barras flutuantes disponíveis nas laterais da tela de navegação.



Imagem 2: visão geral da página de acesso ao Portal “Pauliceia 2.0”



O processo de refinamento dos objetos gráficos ocorreu após o lançamento da versão Beta (testes da plataforma), concentrando-se entre os meses de novembro de 2018 a janeiro de 2019, na finalização dos elementos correspondentes ao Painel de Usuário - *dashboard* - que contempla as informações e tarefas para os usuários já cadastrados na plataforma e que possuem acesso às páginas relativas ao Painel de Controle. Nesta última etapa de produtos gráficos foram desenvolvidas as telas de “Login” e “Novo Cadastro” dos usuários, incluindo a página “Meu perfil”, e o menu que compreende o “Painel de Controle”, as “Camadas” e “Palavras-chave”.



The image shows the login interface of the Pauliceia 2.0 platform. At the top, there is a navigation bar with the Pauliceia 2.0 logo, links for Home, Mapa, Sobre, and Contato, and an Entrar button next to flags for Brazil and the United States. The main content area has an orange background and features a white login box. Inside the box, the title 'Acesse sua conta' is followed by two input fields: 'Endereço eletrônico' (containing 'E-mail') and 'Senha de acesso' (containing 'Senha'). Below these fields are links for 'Esqueceu a senha?' and 'Novo Registro'. A 'Google' logo is next to the text 'Acesso com Google'. An orange 'Entrar' button is at the bottom right of the box. At the very bottom of the box, a small line of text states: 'Ao continuar, você concorda com os Termos de Uso'.

Imagem 3: login de acesso ao usuário cadastrado na plataforma

Como qualquer website, há atualizações e ajustes pontuais que possivelmente serão detectados e corrigidos a longo prazo, visando adaptar-se à dinâmica de acessos e necessidades dos usuários e pesquisadores da plataforma “Pauliceia 2.0”. Portanto, os meios de comunicação e contato com os usuários foram tratados de maneira essenciais neste processo de construção coletiva da plataforma digital.

As atividades foram desenvolvidas dentro dos prazos e cronograma previstos para o período do projeto:

Atividades	meses												
	2018												19
	02	03	04	05	06	07	08	09	10	11	12	01	
Atividades gerais													
- design e programação visual													

- 1) Finalização da Identidade Visual - março/2018;
- 2) Proposta preliminar do website (portal e plataforma) - agosto/2018;
- 3) Lançamento da Versão Beta para testes (versão HTML) - outubro/2018;
- 4) Refinamento dos elementos gráficos e Painel de Usuário - janeiro/2019.

Após a finalização dos produtos gráficos, a etapa de transição da versão HTML foi desenvolvida inteiramente pela equipe do INPE, uma vez que a página *on line* está hospedada no servidor do Instituto Nacional de Pesquisas Espaciais. Todos os elementos de programação visual elaborados podem ser consultados através do endereço eletrônico: www.pauliceia.dpi.inpe.br.

Atenciosamente,

Priscila Machado Meireles

fevereiro/2019



Anexo 2 – Tutorial de alimentação do banco de dados

Instruções básicas para o preenchimento do banco de dados do geolocalizador da plataforma Pauliceia 2.0 - “TABELÃO”

- As ruas serão divididas entre os integrantes para o preenchimento da tabela, cada um ficará responsável por um conjunto de ruas, isso será registrado e controlado na planilha CONTROLE_TABELÃO.¹ É só colocar seu nome como responsável por um conjunto de ruas que esteja com o campo “responsável” em branco.
- A geolocalização será feita a partir da metragem, informada nos livros de emplacamento pós saboya, e associando os endereços anteriores, através de um ID (campo ID_PONTO) por isso é essencial que o preenchimento desses campos seja feito com muito cuidado e atenção.
- Cada integrante preencherá uma tabela com suas ruas, ao final todas serão reunidas.
- Seguiremos a lógica que usávamos para o webeditor, **um ponto no início, um no fim, um no meio da rua**, para cada lado da rua (ímpar - par), ou seja **6 pontos**. Para ruas muito extensas colocar dois pontos intermediários. (8 pontos no total). Para as muito pequenas o começo e o fim serão suficientes. (4 pontos)

1. <https://docs.google.com/spreadsheets/d/1Edf-fxSytpXhHNOLjO-cPvYJrhVnrkZwsAyNfVIGELM/edit?usp=sharing>
2. <https://www.dropbox.com/sh/3wxiz6fyjeb8sdg/AAD-hKhG1csEZ5lE09cMVieUa?dl=0>

Preenchimento

1. Consultar na planilha 'CONTROLE_TABELÃO' as ruas ainda sem responsável e colcar seu nome nas que for fazer.

2. Procurar a rua na pasta 'Emplacamentos'² no dropbox, verificar se todos os registros informados no campo 'Livros de Emplacamento' da planilha de controle estão na pasta.
(caso falte algum, não comece o preenchimento da rua, registre na aba PENDÊNCIAS na planilha de controle, e registre isso também na aba de controle, no campo observação colocar "Pendência".)

3. Copiar o **ID_rua** da planilha de controle para a sua planilha;
Criar um **ID_ponto**, constituído do ID_rua + traço (-) + o número do imóvel de acordo com o emplacamento pós saboya. (ver no exemplo a seguir) E esse ID_ponto será o **mesmo** para todas as informações daquele mesmo ponto, em todas as décadas.
Portanto é necessário iniciar pelo emplacamento mais novo (pós 30) e ir regredindo. Fazendo o histórico daquele ponto em todos emplacamentos.

Por exemplo: Uma rua tem emplacamentos em 1930, 20 e 10, coloco todas as informações sobre aquele ponto disponíveis nos três livros.

Depois eu volto ao emplacamento de 1930 para começar um novo ponto. Faremos o histórico por ponto, não por década.

* Caso na rua não exista emplacamento pós saboya (pós 1930), colocar na planilha de pendências e pular para a próxima rua. É possível verificar isso já na planilha de controle.*

1. <https://docs.google.com/spreadsheets/d/1Edf-fxSytPXhHNOLjO-cPvYJrhVnrkZwsAyNfVIGELM/edit?usp=sharing>
2. <https://www.dropbox.com/sh/3wxiz6fyjeb8sdg/AAD-hKhG1csEZ5IE09cMVieUa?dl=0>

Id da rua como informado na planilha de controle.

Id_rua-numero do imóvel no “pós-saboya”

A	B	C	D	E	F	G	H	I	J	K
Id da rua	Id_ponto	metragem	logradouro	número	n. original	data_inicial (DD/MM/AAAA)	data_final (DD/MM/AAAA)	fonte	autor_da_alime	Data
156	156-40	40.00	rua da alegria	40	40	18/04/1936		livro de emplacamento 21, p. 118	Luanna	11/10/2018
156	156-40	40.00	rua da alegria	2	2	12/08/1930	17/04/1936	livro de emplacamento 18, p. 399	Luanna	11/10/2018
156	156-40	40.00	rua da alegria	6	6	31/07/1908	11/08/1930	livro de emplacamento 18, p. 399	Luanna	11/10/2018
156	156-40	40.00	rua da alegria	2	2 tinta	30/07/1908	30/07/1908	livro de emplacamento 1, p. 118	Luanna	11/10/2018
156	156-178	177.50	rua da alegria	178	178	18/04/1936		livro de emplacamento 21, p. 118	Luanna	11/10/2018
156	156-178	177.50	rua da alegria	40	40	31/07/1908	17/04/1936	livro de emplacamento 21, p. 118	Luanna	11/10/2018
156	156-178	177.50	rua da alegria	28	28	30/07/1908	30/07/1908	livro de emplacamento 18, p. 399	Luanna	11/10/2018
156	156-336	336.20	rua da alegria	336	336	18/04/1936		livro de emplacamento 21, p. 118	Luanna	11/10/2018
156	156-288	288.00	rua da alegria	288	288	18/04/1936		livro de emplacamento 21, p. 118	Luanna	11/10/2018
156	156-288	288.00	rua da alegria	78	78	31/07/1908	17/04/1936	livro de emplacamento 21, p. 118	Luanna	11/10/2018
156	156-288	288.00	rua da alegria	56	56	30/07/1908	30/07/1908	livro de emplacamento 1, p. 118	Luanna	11/10/2018

Metragem informada pelo livro de emplacamento pós saboya. Este mesmo número se repete pros mesmos pontos anteriores que não possuem a metragem. Com decimais, separados por ponto (.)

Nome do logradouro conforme tabela de controle

logradouro	numero	numero original	data_inicial (DD/MM/AAAA)	data_final (DD/MM/AAAA)	fonte	autor_da_alimenta	Data
viaducto santa ephigenia	1.1	1a		13/03/1941	livro de emplacamento 28, p. 118	Luanna	08/10/2018
ladeira da constituicao	23	23	13/9/1937		livro de emplacamento 22, p. 399	Luanna	08/10/2018
ladeira da constituicao	1	1		12/09/1937	livro de emplacamento 22, p. 399	Luanna	08/10/2018
ladeira da constituicao	69	69	13/9/1937		livro de emplacamento 22, p. 399	Luanna	08/10/2018
ladeira da constituicao	28	28	13/9/1937		livro de emplacamento 22, p. 399	Luanna	08/10/2018
ladeira da constituicao	2	2		12/09/1937	livro de emplacamento 22, p. 399	Luanna	08/10/2018
ladeira da constituicao	74	74	13/9/1937		livro de emplacamento 22, p. 399	Luanna	08/10/2018
ladeira da constituicao	8	8		12/09/1937	livro de emplacamento 22, p. 399	Luanna	08/10/2018

Número: o numero que o sistema vai calcular.

Nos casos dos alfanuméricos adicionar decimais. Ex. 1a, nesse campo é 1.1. (usar **ponto**, não vírgula)

Número original: Como o número aparece no emplacamento.(1a, 1s, 1t), se for um número comum, é só repetir. 1s, 1t no número é só 1.

Data inicial daquela numeração, indicada pelo “realizado em” de cada emplacamento. Em formato DD/MM/AAAA

Data final daquela numeração, um dia anterior ao emplacamento seguinte


Livro(s) de emplacamento usados para aquela informação.

Caso não haja mudança da numeração em dois emplacamentos diferentes, juntar as informações na mesma linha, aumentando o período que aquele número ficou vigente. Informar na coluna 'fonte' todos os emplacamentos que foram usados para obter essa informação.

Quando chegar no emplacamento mais antigo, não há como saber desde quando o “número antigo” informado ali estava vigente, usaremos então a data desde emplacamento como data inicial. Se a numeração nova e a antiga coincidirem, adicionar uma linha para aquele ponto com uma “data discreta”, usaremos um dia anterior à data anterior daquele emplacamento nos campos de data inicial e final. O exemplo a seguir mostra um caso como esse, no ID 78-42.

Exemplo: Rua Araújo

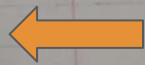
No emplacamento de 1936 (pós-saboya) a rua Araújo tem como primeiro número o 42, também é possível ver neste mesmo emplacamento que o antigo número deste imóvel era o 55. No emplacamento de 1930, o número 55 aparece nas colunas de “antigo” e “novo”, isto quer dizer que em 1930 a numeração deste imóvel permaneceu a mesma. Já em 1908, o número 55 aparece na coluna “novo”, entende-se que foi a partir desse emplacamento que aquele número passou a ser válido e que o anterior a ele era o 61.



Rua da República			9,70	9	340	339,30		
55	42	41,00		7	350	350,40	N	31.10.58
53	54	54,60		5	362	361,70	N	-
51	62	61,70	N 14.2.64	3	372	372,60	N	24.5.54
49	84	84,60		1	392	391,50	405,90	
47	86	85,80	93,00		Rua da Revolução			
0	0							

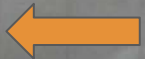
Emplacamento de 1936

23	24			21	23	47	47
24	26			23	25	51	51
26	28			25	27	53	53
28	30			27	29	55	55



Emplacamento de 1930

1	45
53	47
55	49
57	51
59	53
61	55



Emplacamento de 1908

Rua Araújo no Tabela

Os números 42, 55, e 61 correspondem ao mesmo imóvel e por isso possuem o mesmo Id_Ponto, que é atribuído sempre em referência a numeração pós saboya, são as colunas de “data inicial” e “data final” que informam o tempo em que determinada numeração ficou vigente.

Referente ao mesmo ponto, possuem o mesmo ID, a mesma metragem, usando uma linha para cada mudança da numeração. Informando a data em que o número ficou vigente

	A	B	C	D	E	F	G	H	I	J	K
1	id da rua	Id ponto	metragem	sistema métrico	logradouro	numero	numero original	data inicial	data final	fonte	autor da alimenta
2	78	78-42	41	s	Rua Araújo	42	42	15/05/1936		Livro de Emplacamento, n.21, p.122	Cintia
3	78	78-42	41	n	Rua Araújo	55	55	19/09/1908	14/05/1936	Livro de Emplacamento, n.21, p.122; Livro	Cintia
4	78	78-42	41	n	Rua Araújo	61	61	18/09/1908	18/09/1908	Livro de Emplacamento, n.1, p.2	Cintia
5	78	78-276	276.4	s	Rua Araújo	276	276	15/05/1936		Livro de Emplacamento, n.21, p.122	Cintia
6	78	78-276	276.4	n	Rua Araújo	21	21	08/08/1930	14/05/1936	Livro de Emplacamento, n.21, p.122; Livro	Cintia
7	78	78-276	276.4	n	Rua Araújo	19	19	19/09/1908	7/8/1930	Livro de Emplacamento, n.21, p.122; Livro	Cintia
8	78	78-392	391.5	s	Rua Araújo	392	392	15/05/1936		Livro de Emplacamento, n.21, p.122	Cintia
9	78	78-392	391.5	n	Rua Araújo	1	1	19/09/1908	14/05/1936	Livro de Emplacamento, n.21, p.122; Livro	Cintia
10	78	78-73	74.1	s	Rua Araújo	73	73	15/05/1936		Livro de Emplacamento, n.21, p.121	Cintia
11	78	78-73	74.1	n	Rua Araújo	42	42	08/08/1930	14/05/1936	Livro de Emplacamento, n.21, p.121; Livro	Cintia
12	78	78-73	74.1	n	Rua Araújo	44	44	19/09/1908	08/08/1930	Livro de Emplacamento, n.21, p.121; Livro	Cintia
13	78	78-73	74.1	n	Rua Araújo	60	60	18/09/1908	18/09/1908	Livro de Emplacamento, n.21, p.121; Livro	Cintia
14	78	78-281	281.8	s	Rua Araújo	281	281	15/05/1936		Livro de Emplacamento, n.21, p.121	Cintia
15	78	78-281	281.8	n	Rua Araújo	10	10	19/09/1908	14/05/1936	Livro de Emplacamento, n.21, p.122; Livro	Cintia
16	78	78-399	399.6	s	Rua Araújo	399	399	15/05/1936		Livro de Emplacamento, n.21, p.121	Cintia

Em casos de dúvidas e/ou inconsistências nas fontes adicionar a rua na aba “pendências” na planilha de controle para que seja analisada em equipe.

Anexo 3 – Determinação da origem das ruas

Teste e Checagem de Direção das Ruas – Pauliceia 2.0

Execução: Monaliza Caetano dos Santos

Supervisão: Cristiane Miyasaka e Luis Ferla

O presente texto tem como principal objetivo a apresentação dos resultados do teste de direção de ruas do projeto Pauliceia 2.0, apresentaremos também algumas das problemáticas e conclusões provenientes do mesmo. O teste de direção das ruas presentes no recorte proposto para o projeto foi desenvolvido e discutido ao longo do primeiro semestre de 2018. A problemática girava em torno da necessidade de designação do sentido e início de cada uma das ruas, para que a distribuição da numeração e pontos pudesse ser feita através de um script computacional. A partir desta demanda, a bibliografia e fontes documentais a respeito da temática foram levantadas, principalmente o Código de Obras de Arthur Saboya, datado de 1929.

O código apontava uma sequência de ruas que, quando traçadas, formavam eixos norte/sul e leste/oeste ao longo de toda a cidade (vetor destacado em vermelho, consultar imagem 01 e 02). Estes vetores seriam os principais responsáveis por definir o início das ruas ao seu redor. Ainda segundo o código, cada início de rua seria determinado a partir de sua proximidade com um dos dois eixos, priorizando-se o eixo norte/sul e, no caso de ruas paralelas ao mesmo, utilizando-se o eixo leste/oeste. Ou seja, a extremidade de determinada rua que estivesse mais próxima do eixo principal seria consequentemente o início da mesma. Caso a rua estivesse distante do eixo norte/sul, a mesma regra poderia ser aplicada ao eixo leste/oeste.

No entanto, discussões produzidas pelo grupo levantaram a hipótese de que, na prática cotidiana de lida com os problemas estruturais e legais da metrópole, o que definia o sentido das ruas seria a intersecção de eixos norte-sul e leste-oeste. Além disso, ambos eixos cruzam-se próximo ao ponto zero da cidade, na Catedral da Sé (figura 01, círculo em amarelo). Este último fator atribuiu ainda mais pertinência à hipótese levantada.

A fim de comprovar a positividade ou não da hipótese levantada, foi elaborado um teste. Com o auxílio das ferramentas disponíveis no software livre gvSIG, reunimos os mapas que compõem o recorte inicial do projeto. Os mesmos foram divididos em 4 quadrantes, respeitando-se a divisão proveniente dos próprios eixos norte-sul e leste-oeste (figura 02). Em cada quadrante foram selecionadas dez ruas para o teste. Esta escolha buscou variar as formas das ruas, distâncias da intersecção e eixo, proximidade ou não com os vetores delimitados pelo código Saboya e sua posição. Após a seleção, separamos os livros de emplacamento correspondentes, posteriormente consultando o início e fim de cada rua apontada pelos mesmos. Novamente com o auxílio do software gvSIG, passamos a apontar a extensão, formato e início (segundo o livro de

emplacamentos), de cada rua selecionada. Os vetores em amarelo indicam as ruas escolhidas, e as formas em círculo indicam o início das mesmas (figura 02).

O teste foi capaz de levantar algumas questões bastante importantes, tanto para a discussão historiográfica do tema, quanto para o projeto Pauliceia 2.0: a maioria das ruas tem seu início direcionado para a intersecção entre os eixos norte-sul e leste-oeste, ou seja, ainda que determinada rua esteja próxima ao eixo leste-oeste, seu início poderá estar associado a intersecção, contrariando as disposições presentes no Código Saboya. Por outro lado, algumas ruas seguem lógicas que não se aplicam à hipótese levantada, muitas das vezes mantendo sua direção anterior à implementação do código. Parece de importância bastante relevante que pesquisas a esse respeito sejam produzidas e aprofundadas, para que possamos compreender quais foram os critérios utilizados em casos onde a intersecção não foi adotada enquanto método. Em última instância, o teste também foi capaz de demonstrar que algumas das diretrizes presentes no código Saboya foram simplificadas ou não se aplicaram na prática da cidade. Há de se ressaltar que o teste e seus resultados foram importantes primeiramente para o próprio escopo e questões do projeto e, em segundo lugar, como contribuição para as discussões sobre leis e seus reais alcances e implementações. O teste ainda sugere a necessidade de maiores análises sobre o tema, apontando perguntas em aberto.

Após a constatação de que uma parcela das ruas não responderia a lógica da intersecção de eixos, o grupo percebeu a necessidade de que cada uma das ruas da base do projeto tivesse seu sentido apontado e revisado. Dois principais fatores contribuíram para que essa alimentação fosse feita manualmente, consultando-se cada uma das 523 ruas e livros de emplacamento, ao invés da criação de um script computacional que “puxasse” os pontos de início manualmente. Em primeiro lugar, os sistemas de medição e definição computacionais não foram capazes de assimilar algumas das subjetividades presentes em ruas que, apesar de estarem mais próximas aos vetores ou intersecção, ainda assim tinham início na extremidade oposta. Por outro lado e em decorrência do primeiro fator, havia indiscutível necessidade de que cada um dos pontos zeros estivessem corretos, pois qualquer irregularidade nesta etapa inicial acarretaria em imprecisões nas futuras alimentações de dados provenientes de livros de emplacamentos e outros documentos.

Nesta checagem, feita a partir da própria lista de ruas existentes no interior da área piloto, cada rua foi conferida e, a partir de webeditor para alimentação de dados na plataforma, seu início foi apontado. Esta etapa foi fundamental para uma delimitação da metodologia utilizada na criação da plataforma, colaborando para que os demais dados pudessem ser inseridos corretamente.

IMAGEM 01

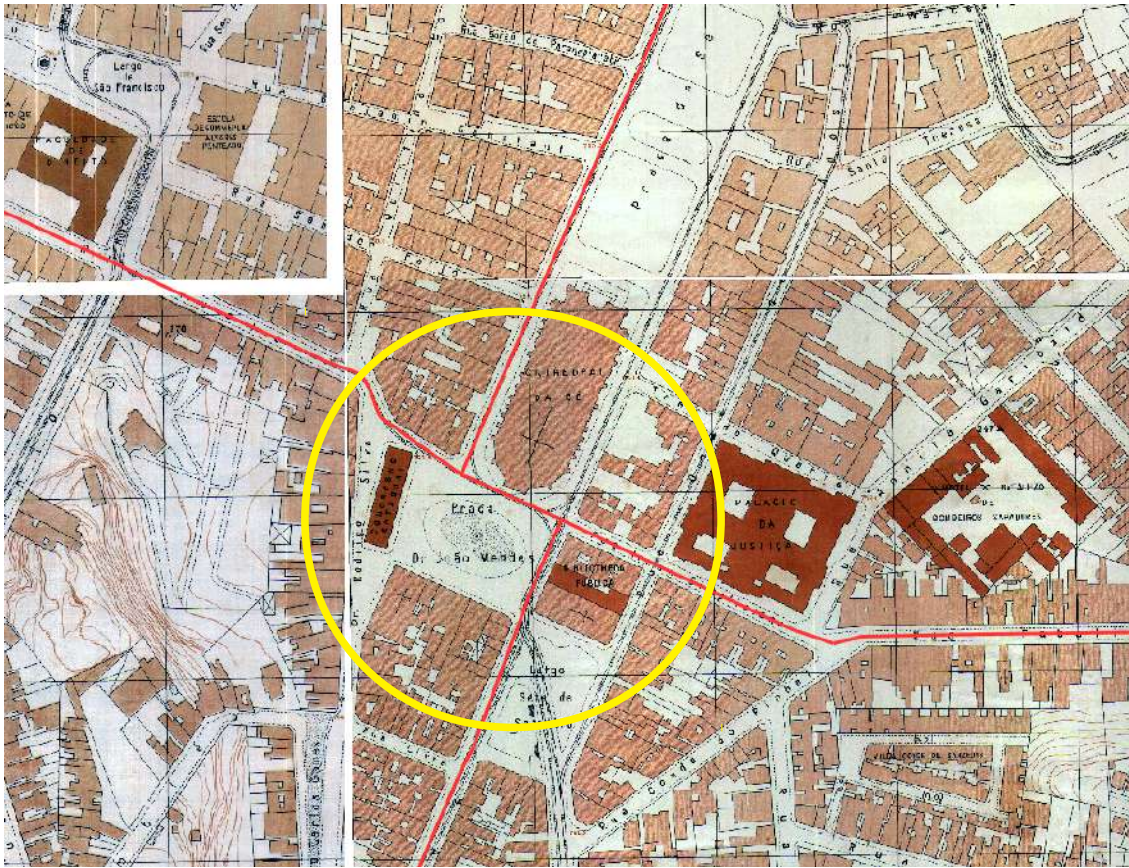


IMAGEM 02



Anexo 4 – Ementa da disciplina de Lab. 1

UNIDADE CURRICULAR (UC): Laboratório de Ensino e Pesquisa I	
Professor responsável: Profº Drº Luis Ferla	Contato: 11-25978170
Ano Letivo: 2018	Semestre: 2
Departamentos/Disciplinas participantes: História	
Carga horária total: 150h	
Carga horária p/ prática: 60h	Carga horária p/ teoria: 90h
OBJETIVOS Gerais: Com base no pressuposto de que pesquisa e ensino são atividades complementares, esta Unidade Curricular buscará propiciar uma reflexão sobre as fontes históricas, suas instituições de guarda documental e as atuais alternativas tecnológicas para sua disponibilização. Ao longo do curso, os alunos terão a oportunidade de experimentar atividades próprias do ofício do professor/pesquisador por meio da reflexão acerca das principais metodologias de leitura crítica de fontes escritas, bem como do processo de seleção e análise das mesmas. Por fim, pretende-se abordar os influxos das novas tecnologias nessas temáticas e as consequentes reconfigurações do ofício do historiador, dentro do contexto mais geral da afirmação das chamadas humanidades digitais.	
OBJETIVOS Específicos: <ul style="list-style-type: none"> - Conhecer a diversidade de instituições de custódia documental e suas potencialidades para o trabalho do historiador. - Conhecer as diferentes tipologias das fontes textuais. - Introduzir noções de leitura paleográfica. - Desenvolver a capacidade de definir um corpus documental, de caracterizá-lo e de analisá-lo em seu contexto histórico. - Capacitar o aluno a fazer leitura e análise de fontes históricas, cruzando diversas variáveis sociais e espaciais. - Capacitar o aluno ao uso de algumas das ferramentas de investigação disponibilizadas pelo mundo digital na análise das fontes; - Refletir criticamente acerca do impacto das novas tecnológicas no ofício do historiador. 	
EMENTA O objeto da Unidade Curricular é proporcionar aos discentes experiências práticas características do ofício do pesquisador/professor de História no trabalho com as fontes. Assim, é objeto do curso o trabalho crítico com as fontes e a percepção de sua historicidade, tanto na pesquisa, quanto na sala de aula e em materiais didáticos. O influxo das novas tecnologias digitais nas práticas do historiador e seu impacto nos temas do curso também serão abordados.	
CONTEÚDO PROGRAMÁTICO <ul style="list-style-type: none"> • A formação do campo do historiador professor e pesquisador: perspectivas profissionais. • Instituições de custódia documental. • Pesquisa, formas de disponibilização e tipologia de fontes. • Paleografia. • Fontes documentais: usos e possibilidades para ensino e pesquisa. • Noções básicas de cartografia. • A afirmação das humanidades digitais. • As novas tecnologias e o ofício do historiador. 	
METODOLOGIA DE ENSINO Aulas expositivas-dialogadas Orientação de leitura e discussão de textos Aulas práticas de seleção e análise de fontes documentais Visita e pesquisa em instituições de guarda Estudos de meio dirigidos Uso de ferramentas e tecnologias digitais	

<p>RECUROS INSTRUCCIONAIS</p> <p>Datashow, livros, computadores para os alunos, acesso Internet</p>
<p>AVALIAÇÃO</p> <p>A avaliação consistirá em uma composição dos seguintes itens:</p> <ul style="list-style-type: none"> - Transcrição de fonte documental - Relatório de visita à instituição de custódia documental - Relatório de trabalho em campo - Elaboração de um trabalho final a partir de uma fonte, levando em conta o cruzamento de variáveis sociais e espaciais e a visualização digital de dados pesquisados para sua análise, fazendo uso da tecnologia de geoprocessamento.
<p>BIBLIOGRAFIA BÁSICA</p> <p>ARQUIVO NACIONAL. Normas técnicas para transcrição e edição de documentos manuscritos. s.d. Disponível em www.portal.arquivonacional.gov.br/Media/Transcreve.pdf. Acesso 12 jun. 2010.</p> <p>BACELLAR, Carlos. Uso e mau uso dos arquivos. In: Pinsky, Carla Bassanezi (org.). Fontes históricas. 2 ed. São Paulo: Contexto, 2006.</p> <p>BELLOTTO, Heloísa Liberalli. Arquivos permanentes: tratamento documental. 4ª ed. Rio de Janeiro: FGV, 2006. Capítulo 2 – Documento, informação e meios institucionais de custódia e disseminação, pp. 35-43.</p> <p>Oficina de Paleografia. Disponível em: http://www.ufrj.br/graduacao/PETHistoria/arquivos_PET/atividades/paleografia/Apostila_Oficina-Paleografia-I.pdf. Acesso: 10/08/2011.</p> <p>FREITAS, Frederico. "Spatial History e Humanidades Digitais: uma entrevista com o brasilianista Zephyr Frank". Revista História e Cultura, Franca-SP, v.3, n.1, p.375-393, 2014.</p> <p>NOIRET, Serge. "História Pública Digital". Liinc em Revista, Rio de Janeiro, v.11, n.1, p. 28-51, maio 2015.</p> <p>Raths, Louis E et al. Ensinar a pensar: teoria e aplicação. São Paulo: Herder, 1972. (371.305 / E59).</p>
<p>BIBLIOGRAFIA COMPLEMENTAR</p> <p>ARÓSTEGUI, Julio. A pesquisa histórica: teoria e método. [La investigación histórica: teoría y método]. Bauru: Edusc, 2006. 591 p. (História). ISBN 8574603007. (907.2 / A769p)</p> <p>BERWANGER, Ana Regina; LEAL, João Eurípedes. Noções de paleografia e de diplomática. Santa Maria, EditorUFSM, 2008.</p> <p>BITTENCOURT, Circe (org). O saber histórico na sala de aula. São Paulo: Contexto, 2006.</p> <p>BITTENCOURT, Circe. Livro didático e saber escolar (1810-1910). Belo Horizonte: Autêntica, 2008.</p> <p>CADIU, François et al. Como se faz a história: historiografia, método e pesquisa. Petrópolis: Vozes, 2007.</p> <p>CARDOSO, Ciro Flamarion; BRIGNOLI, Héctor Pérez. Os métodos da História: introdução aos problemas, métodos e técnicas da história demográfica, econômica e social. 6 ed. Rio de Janeiro: Graal, 2002.</p> <p>FACHIN, Phablo Roberto. Descaminhos e dificuldades de leitura de manuscritos do século XVIII. Goiânia, Trilhars Urbanas/ FAPESP, 2008.</p> <p>FERNANDES, Paula Porta (org.). Guia dos Documentos Históricos na Cidade de São Paulo, 1554/1954. São Paulo: Editora Hucitec/NEPS, 1998.</p> <p>FLEXOR, Maria Helena Ochi. Abreviaturas manuscritos dos séculos XVI ao XIX. 3ed. rev. e aum. Rio de Janeiro: Arquivo Nacional, 2008. 600p.</p> <p>FLORES, Daniel; SANTOS, Henrique M. dos. "Preservação de documentos arquivísticos digitais: reflexões sobre as estratégias de encapsulamento". Liinc em Revista, Rio de Janeiro, v.11, n.1, p. 167-180, maio 2015.</p> <p>Gregory, Ian; Ell, Paul. Historical GIS: Technologies, methodologies and scholarship. Cambridge: Cambridge University Press, 2007.</p> <p>KNOWLES, Anne (ed.). Placing history: how maps, spatial data, and GIS are changing historical scholarship. Redlands: Esri Press, 2008.</p> <p>LÉTOURNEAU, Jocelyn (org.). Ferramentas para o pesquisador iniciante. São Paulo: Martins Fontes, 2011.</p>

MENDES, Ubirajara Dolácio. Noções de paleografia. São Paulo: Arquivo Público do Estado, 2008.

RATHS, Louis E et al. Ensinar a pensar: teoria e aplicação. São Paulo: Herder, 1972. (371.305 / E59)

REBELATTO, Marta; FREITAS, Frederico. “Desafios e possibilidades ao uso de Sistemas de Informação Geográfica na história”. In: 5º Encontro Escravidão e Liberdade no Brasil Meridional, de 11 a 13 de maio de 2012, UFRGS, Porto Alegre; texto disponível em: <http://www.labhstc.ufsc.br/vencontro/pdf/REBELATTO,%20Martha%20e%20FREITAS,%20Frederico.pdf> Salomon

Délcio Vieira. Como fazer uma monografia. 11. ed. São Paulo: Martins Fontes, 2004. (001.81 / S174c)

SEVERINO, Antônio Joaquim. Metodologia do trabalho científico. 12.ed. rev. e ampl. São Paulo: Cortez, 1985. 237 p. (001.42 / S498m)

SEVCENKO, Nicolau. A Cidade Vertigem e o Urbanismo Inflacionário. Incursões na Entropia Paulista In: CAMARGO, Ana Maria de Almeida (coord.) São Paulo. Uma Viagem no Tempo. São Paulo: CIEE, 2005, p.p. 185-202.

DOCENTE PARTICIPANTES				
Nome	Origem (Curso)	Titulação	Regime de Trabalho	Carga horária
Luis Antonio Coelho Ferla	História	Doutor	DE	

Anexo 5 – Artigo apresentado no GeoInfo 2018

VGI Protocol and Web Service for Historical Data Management

Rodrigo M. Mariano¹, Karine R. Ferreira¹, Luis A. C. Ferla²

¹ National Institute for Space Research (INPE)
São José dos Campos – SP – Brazil

² Federal University of São Paulo (UNIFESP)
Guarulhos – SP – Brazil

{rodrigo.mariano,karine.ferreira}@inpe.br, ferla@unifesp.br

Abstract. *Volunteered Geographic Information (VGI) is a phenomenon that uses the web to produce, assemble and disseminate geographic information provided by volunteers. VGI techniques generate detailed geographical data with low cost, taking advantage of citizens local knowledge. The definition of a VGI protocol is crucial to improve the quality of citizen-derived geographical data sets collected by a project. Protocols are also important to facilitate the reuse of VGI data for other projects and applications different from what was originally collected. This paper presents a VGI protocol that was defined for the Pauliceia 2.0 project and a web service that was built based on this protocol. Pauliceia 2.0 project aims to use VGI and crowdsourcing techniques to produce historical geographical data sets of São Paulo city from 1870 to 1940.*

1. Introduction

VGI, citizen science, crowdsourcing and collaborative mapping are examples of different terms used to refer to the general subject of collaborative work and citizen-derived geographical information. See et al. [See et al. 2016] present a good review of these terms and categorize them according to three main aspects: (1) information or process; (2) active or passive contributions; and (3) spatial or non-spatial user-generated information.

The term VGI was first defined by Goodchild [Goodchild 2007] as “*the harnessing of tools to create, assemble, and disseminate geographic data provided voluntarily by individuals*”. Goodchild and Li [Goodchild and Li 2012] define VGI as a version of crowdsourcing, focused on manipulating geographical information. Estellés-Arolas and Guevara [Estellés-Arolas and González-Ladrón-de Guevara 2012] define crowdsourcing as “*a type of participative online activity in which an individual, an institution, a non-profit organization, or company proposes to a group of individuals of varying knowledge, heterogeneity, and number, via a flexible open call, the voluntary undertaking of a task. The undertaking of the task, of variable complexity and modularity, and in which the crowd should participate bringing their work, money, knowledge and/or experience, always entails mutual benefit.*”.

There are many projects that use VGI techniques to collect data, such as OpenStreetMap, Wikimapia and Flickr. OpenStreetMap is the most well-known general platform that implements VGI successfully [Goodchild and Li 2012]. It is an editable map

of the world, provided by volunteers, being possible to handle free geographic data [OpenStreetMap 2017a]. It adopts the local expertise of the users to make updated maps.

Mapping agencies use robust protocols that drive the geographic data collection, while VGI projects regularly contain lack of standards or just supply vague instructions. The definition of a VGI protocol is crucial to improve the quality of citizen-derived geographical data sets collected by a project and to facilitate the reuse of these data sets for other projects and applications [Mooney et al. 2016].

Mooney et al. [Mooney et al. 2016] propose a generic protocol to drive VGI projects. This protocol establishes crucial issues that must be defined in the context of a VGI project in order to improve the understanding of volunteers and so the quality of the data produced by them. These issues include vector geographical data collection and management, user control, self-assessment and quality metrics and feedback to the community.

Pauliceia 2.0 project aims to build a computational platform for collaborative historical research based on VGI and crowdsourcing techniques [Ferreira et al. 2017]. Through this platform, citizens can contribute to produce historical geographical information of São Paulo city from 1870 to 1940. These contributions can be done in different ways, for example, by doing the vectorization of streets and buildings from historical maps or by uploading photos and information about historical places. Besides that, this platform allows historians to share data sets resulting from their researches.

This paper presents a VGI protocol for historical data that was defined for the Pauliceia 2.0 project and a web service based on this protocol, called VGI Management Web Service (VGIMWS), that was built in the Pauliceia platform. Ferreira et al. [Ferreira et al. 2017] introduce the Pauliceia project and its platform generally; while this paper presents a detailed description of the VGI protocol for historical data and VGIMWS service.

2. Related Work

This section presents projects that also use VGI and crowdsourcing techniques to produce historical geographical information, similar to Pauliceia 2.0 project.

OpenHistoricalMap [OpenStreetMap 2018] and HistOSM¹ projects are built on the OpenStreetMap (OSM) platform. OpenHistoricalMap is an effort to use the OSM infrastructure to produce a universal and historical map of the world. HistOSM is a web application to explore the historical objects of OSM, such as castles, ruins, monuments and memorials.

Building Inspector² is a web-based platform that allows citizens to produce, correct and analyze data from historical maps of New York city from 1853 to 1930. In this project, Budig et al. [Budig et al. 2016] propose a consensus polygon algorithm to extract a single polygon to represent each building from all polygons provided voluntarily.

ATLMaps³ is a web portal of the Atlanta Explorer project that handles historical information of Atlanta city for post Civil War to 1940 [Page et al. 2013]. This portal

¹<http://histosm.org/>

²<http://buildinginspector.nycl.org/>

³<https://atlmmaps.org/>

allows user to visualize and explore historical maps, events and places. Users can produce their own projects using the layers that are ready to use in the portal and contribute with audios, annotations or images related to them.

Many VGI projects use the OSM infrastructure and its API (Application Programming Interface) to build their web platforms. In the beginning of the Pauliceia 2.0 project, our team evaluated if it was possible to build the Pauliceia 2.0 platform using the OSM infrastructure and its API. However, after some studies, we concluded that the OSM data model and operation are not suitable for the Pauliceia project due to the following reasons:

1. In the Pauliceia platform, historians can share data sets resulting from their researches and these data sets can not be edited by anyone. In OSM, data can be updated by anyone.
2. The historical features in the Pauliceia database are spatiotemporal, that is, they have a period to indicate when they existed. There are features that do not exist today anymore. In OSM database, the entities are not spatiotemporal. OSM considers that all entities stored in its database exist today.
3. In the Pauliceia platform, the historical data sets are organized in layers, while OSM data sets are not.
4. The community and domain of the Pauliceia project are very specific and structured, while the OSM is very general. The Pauliceia project has a specific domain with a particular spatial and temporal scope, generating a structured community.

Therefore, we decided not to use the OSM infrastructure and API. We defined a specific VGI protocol for the Pauliceia project and built a web service for VGI data management based on this protocol. These protocol and web service are crucial parts of the Pauliceia platform and they are described in the next sections.

3. VGI Protocol for Historical Data

Mooney et al. [Mooney et al. 2016] propose a generic protocol that organizes the issues related to citizen-derived geographical data management in five main stages, that are shown in Figure 1: (1) Initialisation; (2) Vector data collection; (3) Self-assessment and quality control; (4) Data submission; and (5) Feedback to the community.

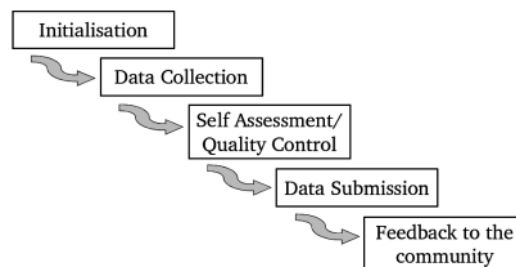


Figure 1. Main stages of VGI Protocol [Mooney et al. 2016]

A VGI protocol defines crucial issues that improve the understanding of volunteers about the project and all its mechanisms and methods to collect, manage and assess the quality of the citizen-derived geographical data. Thus, this helps to improve the quality of the data sets collected by volunteers in a project.

In this work, we define a VGI protocol specific for the Pauliceia 2.0 project following the guide proposed by Mooney et al. [Mooney et al. 2016]. The Pauliceia 2.0 VGI protocol is described in next sections.

3.1. Data types and initialization

In the Pauliceia 2.0 project, volunteers can upload and edit vector geographical data. The Pauliceia platform provides tools that allow citizens to include and edit geometries, such as points, lines and polygons, as well as textual and numerical values associated to geographical entities or features. The platform does not provide tools to edit and create raster data types.

One of the project goals is to use VGI and crowdsourcing techniques for the vectorization of features, such as streets and buildings, from historical raster maps. In this case, the data set gathered by volunteers can have a set of distinct geometries to represent the same feature. To extract the most accurate geometry to represent a single feature from this data set, we intend to employ methods that compute a single geometry that represents the majority opinion, as proposed by Budig et al. [Budig et al. 2016].

The users access the Pauliceia 2.0 platform through an online browser. Before starting the contribution, the collaborator needs to register himself/herself to the platform and accept a "Use Term" that describes mainly that the portal is not responsible for the collected data sets and that these data sets are public. The registration can be done by creating a new user or by using a social login through Google and Facebook accounts. Everybody can access the platform and visualise its data sets freely, however just registered volunteers can edit or add new historical data sets.

The Pauliceia 2.0 database is made available under the Creative Commons Attribution-ShareAlike 4.0 license (CC BY-SA)⁴. In a nutshell, this license authorises the people freely copy, share, modify and use the data for any propose, since the users cite Pauliceia 2.0 and its contributors. If the user reproduces the data, he or she must use the same license for the results.

To motivate volunteers to vectorize streets and buildings as well as historians to share their historical data sets, we intend to organize events oriented to this purpose, following the same idea of the mapathon events promoted by Google Maps [Tech2 2014] and OpenStreetMap [OpenStreetMap 2017b]. These events, called HistMapathon (Historic Mapping Marathon), will be organized in universities with historians and their students to promote the mass contribution of geographical data in the Pauliceia 2.0 platform.

3.2. Data model

Figure 2 shows the concepts of the Pauliceia 2.0 VGI protocol and their relationships, using an entity-relationship diagram. Its main concepts are: user, layer, reference, keyword and notification.

In the platform, the data sets are organized in layers as in GIS (Geographic Information System). A layer groups geographical features related to a subject that are described by the same set of properties. Each feature has spatial and non-spatial properties [Herring 2006]. The spatial properties are represented by geometric data types, such as

⁴<https://creativecommons.org/licenses/by-sa/4.0/>

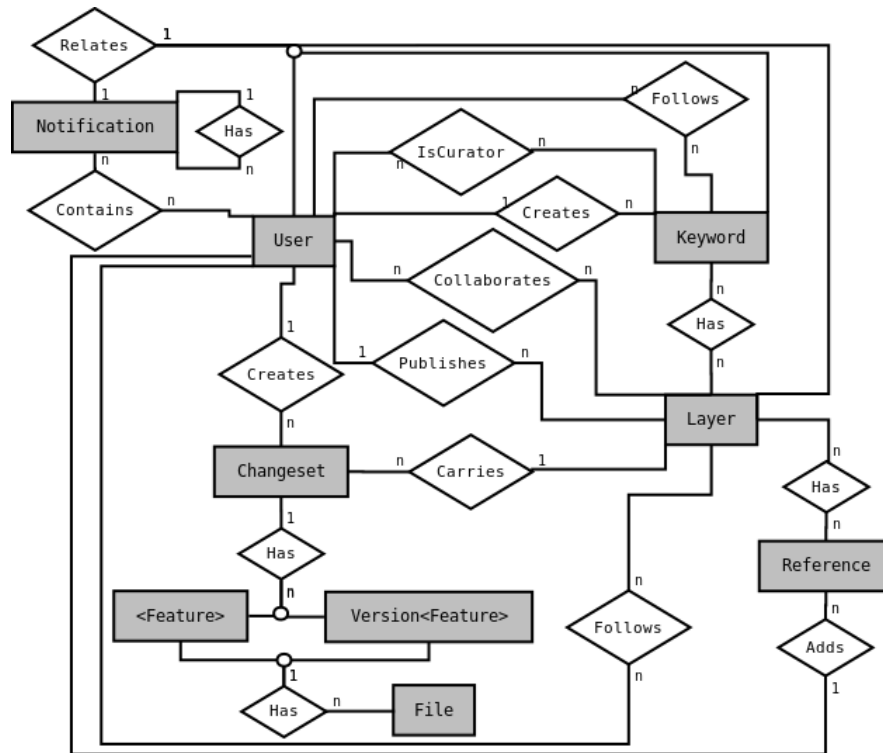


Figure 2. The Pauliceia 2.0 project data model

points, lines and geometries. The non-spatial properties are represented by alphanumeric data types, such as texts and numbers. A non-spatial property of a feature can contain links to media or documents, e.g photos and videos, that are stored in other repositories, like Google Drive, YouTube or Dropbox.

A layer contains a set of features and their versions along time. The changeset entity controls the features of a layer and their versions along time, keeping the history about when and what user realized each change. A changeset is a group of changes related to the features of a layer made by users in a period.

A layer can be associated to one or more keywords (e.g. crimes or factories) and to one or more bibliographical references, such as book, thesis or article. In the platform, the keywords are used in the search mechanism to select layers associated to specific themes. Each layer has an owner user who creates it in the platform and a set of collaborators. Collaborators are users that have permission to edit, delete and include new features into a layer. The collaborators of a layer are defined by its owner. Users can only edit data sets in layers where they are collaborators.

The communication among the platform users, called Pauliceia community, is done through notifications. Notifications can be reviews of data, comments or denuncia-tion. Users can write notifications about a specific layer or about an another notification. Besides that, they can write general notifications for all Pauliceia community.

In the platform, there are two types of users: logged and unlogged ones. The un-

logged users can visualize, search and download the platform data sets as well as read all its notifications. The logged users can be of three types: normal, curator and administrator. The normal users can edit and contribute with new data sets, creating new layers and associating them with keywords and references. The curator users can edit all layers of the platform by adding new keywords to them. The administrator users have permission to create, edit and remove all entities of the platform.

3.3. Data collection methods

There are two ways of data collection methods in the Pauliceia 2.0 platform: manual contribution and bulk import.

In the manual edition, users manually create and edit the spatial locations or boundaries of features by clicking on the historical maps presented in the web portal. Besides that, users can edit manually all attribute values associated to the features.

In the bulk importing, users can upload a group of features stored in well-known file formats of vector geographical data, such as shapefile or geojson. In the manual edition, users have to inform all metadata associated to the features. In the bulk importing, some types of metadata can be extracted automatically by the platform from the file content.

The shapefiles generated by the spatiotemporal geocoding tool of the Pauliceia platform can be directly used in the bulk import. Using the spatiotemporal geocoding tool, a user can upload a CSV (Comma Separated Values) file that contains a set of textual historical addresses and get a shapefile with all spatial locations of these addresses produced by the geocoder. So, this shapefile can be imported in the platform, creating a new layer in the Pauliceia database.

3.4. Quality control

In the literature, there are several proposals to evaluate VGI data quality. In a nutshell, these methods are described as quality measures, quality indicators and quality approaches. Quality measures verify the accuracy of VGI data in relation to the authoritative data provided by mapping agencies. Quality indicators measure the quality of VGI data in an abstract way when there is not authoritative data for comparison [Senaratne et al. 2017]. Quality approaches determine the degree of a fact, if it is possible to be true, and it can be automated or used of human intervention [Goodchild and Li 2012]. One proposal of VGI data quality introduced by Goodchild and Li [Goodchild and Li 2012] is the crowdsourcing approach. One interpretation of this approach is the use of the Linus Law⁵. Linus Law is the ability of using the people to verify the contributed data of VGI to converge to the truth.

To use quality measures, it is necessary to use authoritative data for comparison [Senaratne et al. 2017]. In the Pauliceia 2.0 project context, there is no authoritative data and so we can not use quality measures. Thus, the Pauliceia 2.0 team evaluated different types of quality indicators and approaches to evaluate the citizen-derived geographical data of the project, such as gamification (e.g. ranking) [Senaratne et al. 2017], trustworthiness and user reputation [D'Antonio et al. 2014]. Nevertheless, a consensus

⁵“Given enough eyeballs, all bugs are shallow” [Raymond 2017]

was reached that the best thing would be to adopt a crowdsourcing approach, using notifications and denunciations provided by the Pauliceia 2.0 community.

In the Pauliceia 2.0 platform, a notification is a comment from a user related to a layer or to another comment. A user can write a notification describing the positives or negatives points of a layer, warnings or suggestions to improve it, such as suggestions of new bibliographic references related to the layer. A denunciation is a special kind of notification made to alert administrators that a layer contains inappropriate data (e.g. copyright data or owned by another researcher). The administrators of the platform receive these reports, evaluate the layers associated to denunciations and can remove them from the platform as well as its owner user.

3.5. Feedback to the Community

A collaborative project become better as more users assist in it. Hence, it is important that users supply feedbacks about their experience with the project [Mooney et al. 2016]. In the Pauliceia 2.0 platform, volunteers are encouraged to make comments, give opinions and observations about their experiences on the platform, indicating the positive aspects and suggestions to improve it. This feedback can be done by the available mailing list and social networks that are managed by the Pauliceia 2.0 team. The feedback is important to improve the platform.

Using the Pauliceia 2.0 platform, researchers can disseminate and share their historical data sets as well as receive feedback from other researchers about them through notifications. Such data sets will be freely available on the portal and thus will achieve a greater visibility in the scientific community and dissemination.

Historians can write notifications on their layers or on the layers of other users, providing feedback on their status, such as remarks, praise or hints (e.g. indicate a new reference for that layer). Users can also write general notifications for all members of the Pauliceia 2.0 community, such as research event announcements. If a user finds unappropriated data in the platform, he or she can report it through denunciations. Besides that, users may follow layers of interest from other authors and receive notifications about them by e-mail.

4. VGI Management Web Service for Historical Data

Figure 3 shows the Pauliceia 2.0 platform architecture [Ferreira et al. 2017]. The platform is open source, online and service oriented. Service-oriented systems are well appropriate to supply a better interoperability among applications. The spatiotemporal data sets of the project are stored in a PostgreSQL database with spatial extension PostGIS (vector data) and in GeoTIFF files (raster data).

The platform architecture contains two groups of web services. The first group contains standards of geographic web services specified by the Open Geospatial Consortium (OGC), such as Web Map Service (WMS), Web Feature Service (WFS), Web Coverage Service (WCS) and Catalogue Service Web (CSW). The second group is composed of two web services designed and implemented to augment the functionalities of the OGC standard services, attending to specific and crucial demands of the Pauliceia 2.0 project.

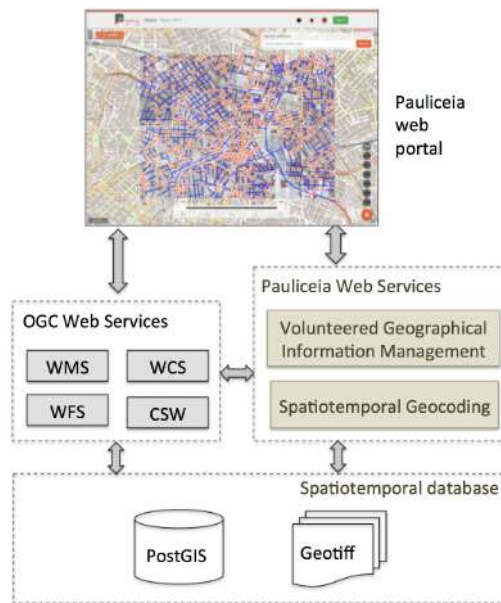


Figure 3. Pauliceia 2.0 platform architecture. [Ferreira et al. 2017]

This section describes the Volunteered Geographical Information Management Web Service (VGIMWS) that was designed and built based on the Pauliceia VGI protocol described in section 3. It provides all necessary functionalities for dealing with historical citizen-derived geographical information, such as user control, spatiotemporal features management as well as user edition of notifications and denunciations.

Figure 4 shows the architecture of VGIMWS. It is a RESTful web service developed in Python language. The chosen standard for data exchange is GeoJSON and JSON, that handle data with geographic information or without, respectively. It provides function to create, edit and remove all concepts described in section 3, such as user, layer, features and notifications.

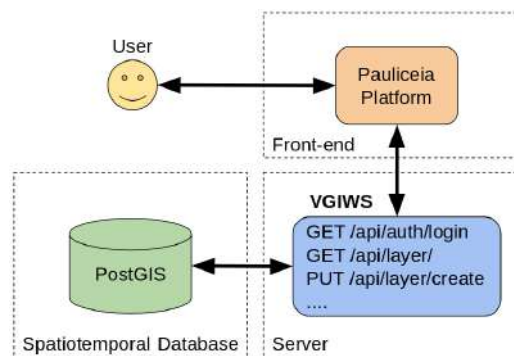


Figure 4. VGIMWS architecture.

All software codes of Pauliceia 2.0 project are free and open source and can

be found at the Github of the project ⁶. This Github contains the source code of the VGIMWS⁷, its documentation⁸ and instruction about how to run the web service and to install its dependencies.

Figure 5 shows a sequence diagram of one function of the VGIMWS that creates a new layer. First, the user tries to log in the platform using one URL and the VGIMWS returns a HTTP status, success or error. If the user is able to log in the platform, he or she can enter on the dashboard, create a new layer and associate keywords and references to it. After that, the volunteer can import a shapefile using the bulk import or create an empty layer. Lastly, the user must inform other metadata about the layer, such as its temporal columns.

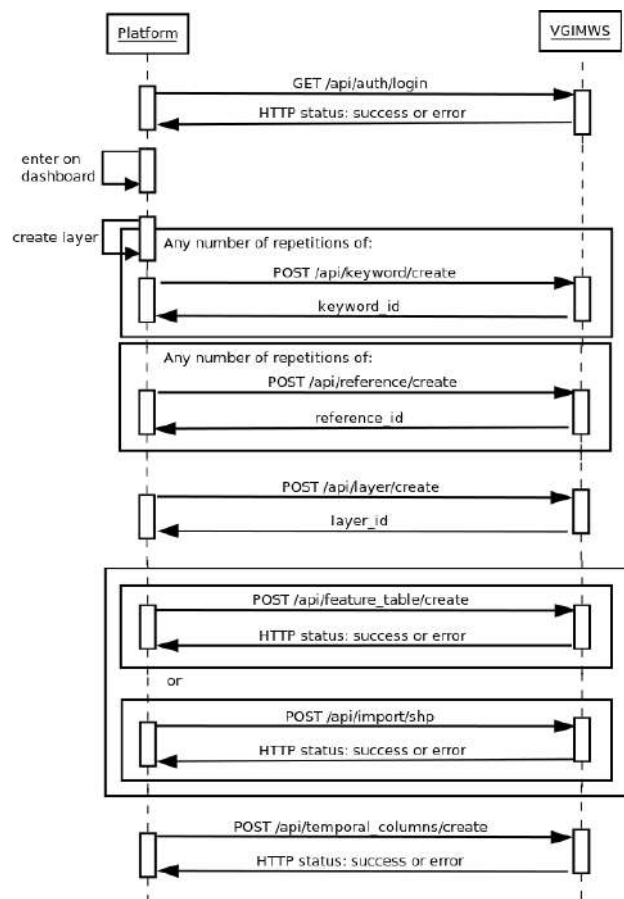


Figure 5. Add a new layer.

A set of tables is proposed to store metadata related to the temporal information and media attributes of the layers, shown in Figure 6. It is an extension of the Simple Feature Access Model proposed by OGC. The GEOMETRY_COLUMNS and the

⁶<https://github.com/Pauliceia>

⁷<https://github.com/Pauliceia/vgiws>

⁸<https://github.com/Pauliceia/vgiws/blob/master/doc/README.md>

FEATURE TABLE are tables defined by OGC, while the MEDIA_COLUMNS, TEMPORAL_COLUMNS and MASK are proposed in this work.

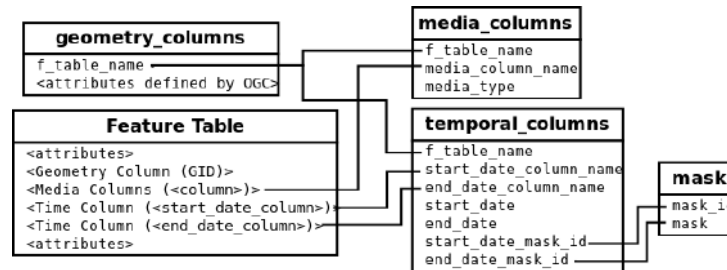


Figure 6. Feature metadata tables.

FEATURE TABLE is a table that stores features, where the columns are the attributes and the rows are the features. GEOMETRY_COLUMNS is a table that contains metadata about the geometry properties of a feature table [Herring 2006]. MEDIA_COLUMNS is a table that contains the metadata about the properties of a feature table that are links to media, such as photos or videos, that are stored in other repositories, as Google Drive, YouTube or Dropbox. TEMPORAL_COLUMNS is a table that defines the temporal attributes of the feature table. Its attributes are the start date, end date and the temporal bounding box of a feature. MASK is a table that saves the possible masks for the start date and end date, such as "YYYY-MM-DD". Both MEDIA_COLUMNS and TEMPORAL_COLUMNS contain a reference to a feature table that is registered in the GEOMETRY_COLUMNS.

Figure 7 presents the complete spatiotemporal database model of Pauliceia 2.0 project. This model express the conceptual model described in section 3.2. It contains tables to store the concepts of the Pauliceia 2.0 project, such as user, layer, reference, keyword, notification, changeset, feature table, media, temporal information and followers. This database was built using PostgreSQL with the spatial extension PostGIS.

5. Conclusion

VGI has emerged with the purpose of collecting geographical data sets fast and with low cost. However, to improve the quality and the reuse of these data sets, it is necessary to define protocols to guide VGI projects.

This paper presents a VGI protocol for historical data that was defined in the context of Pauliceia 2.0 project. This project aims to develop an online platform for collaborative research of historical data, using VGI and crowdsourcing techniques. Besides that, this paper describes a RESTful web service, called VGIMWS, that was built in the Pauliceia platform based on the VGI protocol. VGIMWS manipulates all the protocol concepts through specific URLs.

The VGI protocol helps to increase the quality of the historical citizen-derived geographical data of the Pauliceia 2.0 platform. It defines crucial issues that improve the understanding of volunteers about the Pauliceia project and all its mechanisms and methods to collect, manage and assess the quality of the citizen-derived geographical data. The proposed VGI protocol and the VGI management web service are generic, so both can be used to other collaborative historical project.

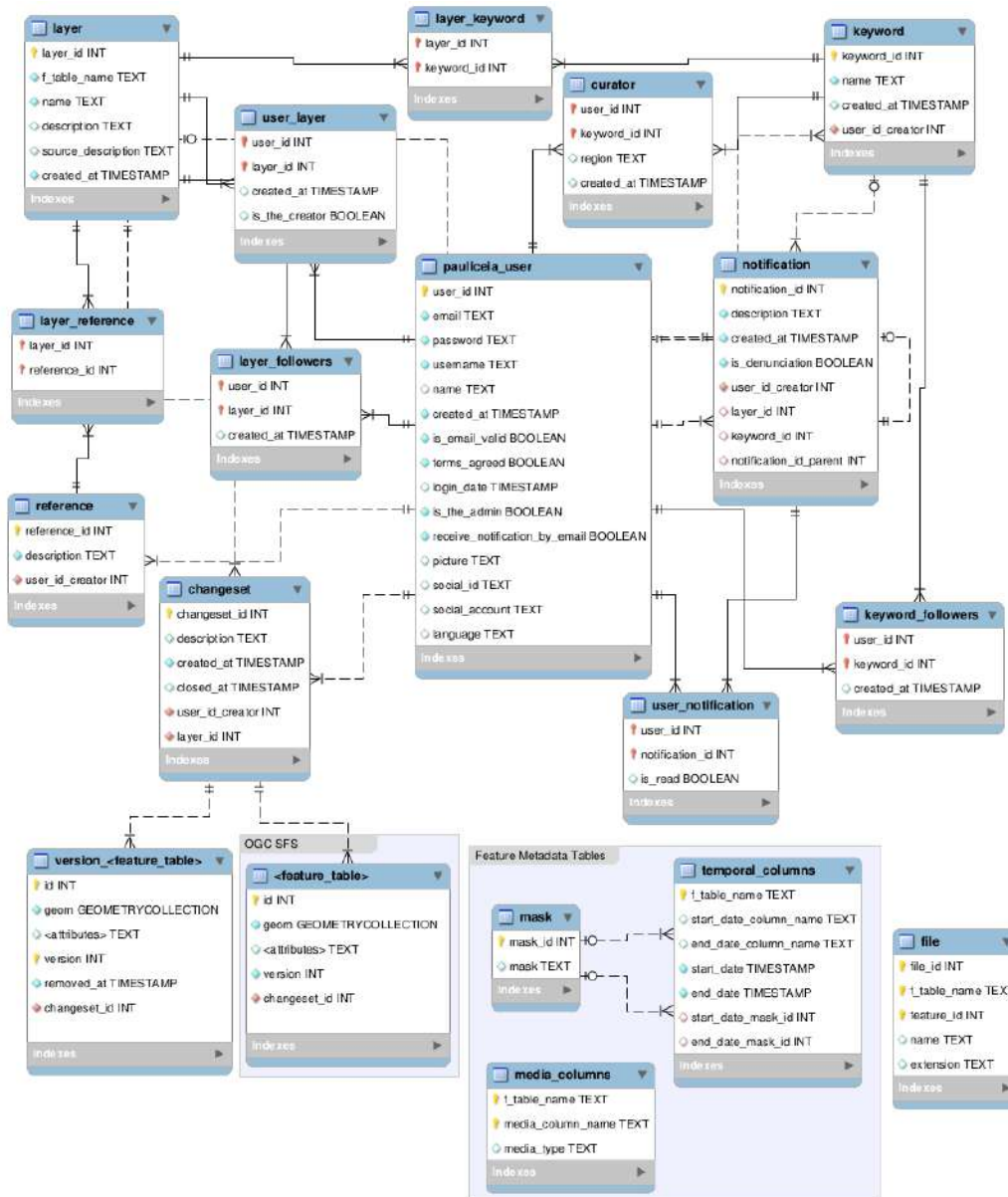


Figure 7. Database Model of the Pauliceia 2.0 project.

6. Acknowledgment

Pauliceia 2.0 project is funded by São Paulo Research Foundation (FAPESP) eScience Program, grant #2016/04846-0. The authors thank FAPESP for granting student scholarship #2017/03852-9.

References

Budig, B., van Dijk, T. C., Feitsch, F., and Arteaga, M. G. (2016). Polygon consensus: smart crowdsourcing for extracting building footprints from historical maps. In

- Proceedings of the 24th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, page 66. ACM.
- D'Antonio, F., Fogliaroni, P., and Kauppinen, T. (2014). Vgi edit history reveals data trustworthiness and user reputation.
- Estellés-Arolas, E. and González-Ladrón-de Guevara, F. (2012). Towards an integrated crowdsourcing definition. *Journal of Information science*, 38(2):189–200.
- Ferreira, K. R., Ferla, L., de 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., de Almeida, C. R., do Nascimento, L. G. M., Diniz, J. A., and dos Santos, M. C. (2017). Pauliceia 2.0: A computational platform for collaborative historical research. *Proceedings XVIII GEOINFO, December 04th to 06th, 2017*, pages 28–39.
- Goodchild, M. F. (2007). Citizens as sensors: the world of volunteered geography. *Geo-Journal*, 69(4):211–221.
- Goodchild, M. F. and Li, L. (2012). Assuring the quality of volunteered geographic information. *Spatial statistics*, 1:110–120.
- Herring, J. (2006). Opengis implementation specification for geographic information-simple feature access-part 2: Sql option. *Open Geospatial Consortium Inc.*
- Mooney, P., Minghini, M., Laakso, M., Antoniou, V., Olteanu-Raimond, A.-M., and Skopeliti, A. (2016). Towards a protocol for the collection of vgi vector data. *ISPRS International Journal of Geo-Information*, 5(11):217.
- OpenStreetMap (2017a). About. <https://www.openstreetmap.org/about>. Accessed on 02/08/2018.
- OpenStreetMap (2017b). Mapathon. <http://wiki.openstreetmap.org/wiki/Mapathon>. Accessed on 18/08/2018.
- OpenStreetMap (2018). Open historical map. https://wiki.openstreetmap.org/wiki/Open_Historical_Map. Accessed on 05/08/2018.
- Page, M. C., Durante, K., and Gue, R. (2013). Modeling the history of the city. *Journal of Map & Geography Libraries*, 9(1-2):128–139.
- Raymond, E. S. (2017). Release early, release often. <http://www.catb.org/~esr/writings/cathedral-bazaar/cathedral-bazaar/ar01s04.html>. Accessed on 10/08/2018.
- See, L., Mooney, P., Foody, G., Bastin, L., Comber, A., Estima, J., Fritz, S., Kerle, N., Jiang, B., Laakso, M., et al. (2016). Crowdsourcing, citizen science or volunteered geographic information? the current state of crowdsourced geographic information. *ISPRS International Journal of Geo-Information*, 5(5):55.
- Senaratne, H., Mobasher, A., Ali, A. L., Capineri, C., and Haklay, M. (2017). A review of volunteered geographic information quality assessment methods. *International Journal of Geographical Information Science*, 31(1):139–167.

Tech2 (2014). Why is google's mapathon in hot waters in india? all you need to know. <http://www.firstpost.com/tech/news-analysis/why-is-googles-mapathon-in-hot-waters-in-india-all-you-need-to-know-3655197.html>. Accessed on 18/08/2018.

Anexo 6 – Tutorial de uso da plataforma

Tutorial Plataforma Pauliceia 2.0



Realização:



Apoio:



Financiamento:



Software:



A Plataforma e Suas Possibilidades

A plataforma Pauliceia 2.0 busca desenvolver e disponibilizar na rede mundial de computadores uma base cartográfica digital histórica da cidade de São Paulo, abordando o período de sua modernização e crescimento (1870-1940).

A interface permite a interação entre pesquisadores interessados no tema e recorte, através da alimentação de dados e eventos espacializáveis de suas próprias investigações. Dessa forma, há a possibilidade de produção de mapas e visualizações de pesquisas próprias ou de terceiros, a partir da base fornecida. Essa alimentação tende a enriquecer os dados disponibilizados, acrescentando aos mesmos novas informações que, por sua vez, poderão ser utilizadas para cruzamento de informações e novas análises.

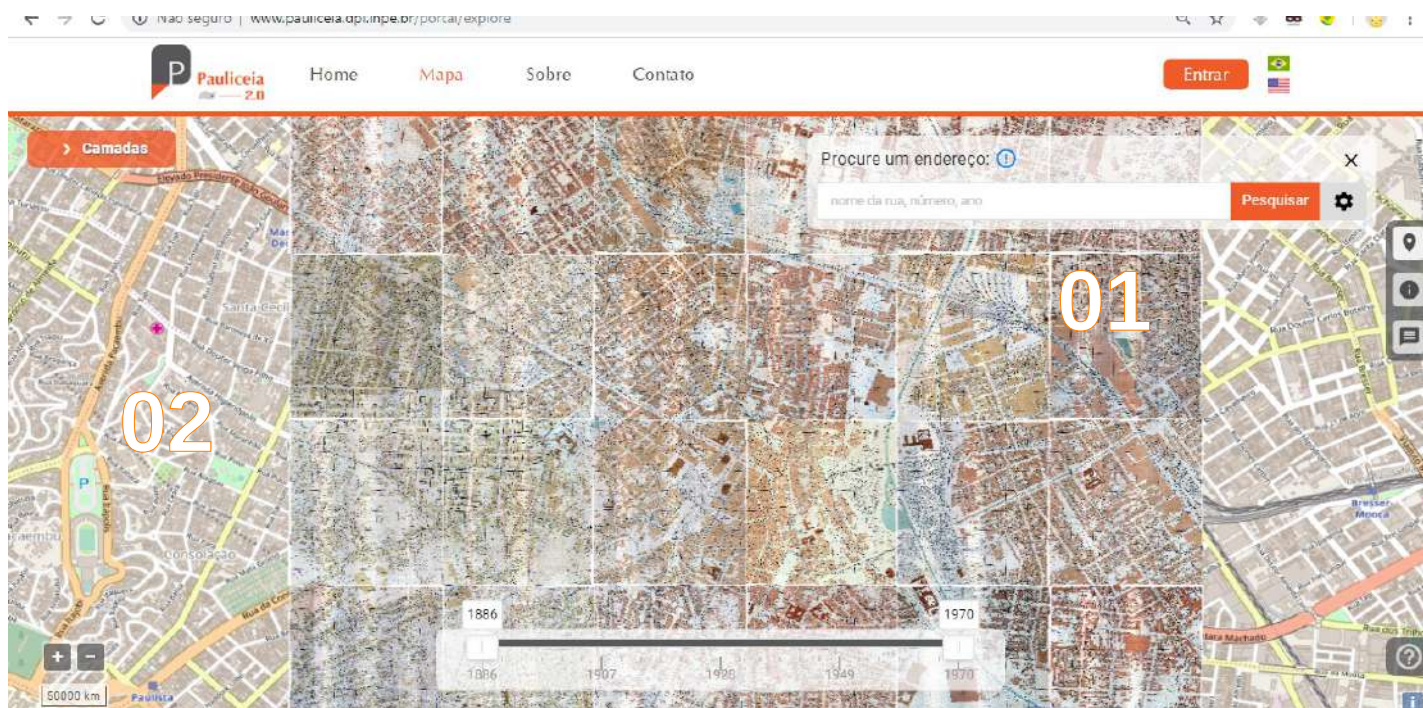
A principal função deste tutorial é apontar as possibilidades de uso e principais ferramentas da plataforma, para que sua manipulação seja ainda mais proveitosa e útil.

Primeiros passos: acesso ao portal e cadastro

O acesso ao portal pode ser feito através deste link:

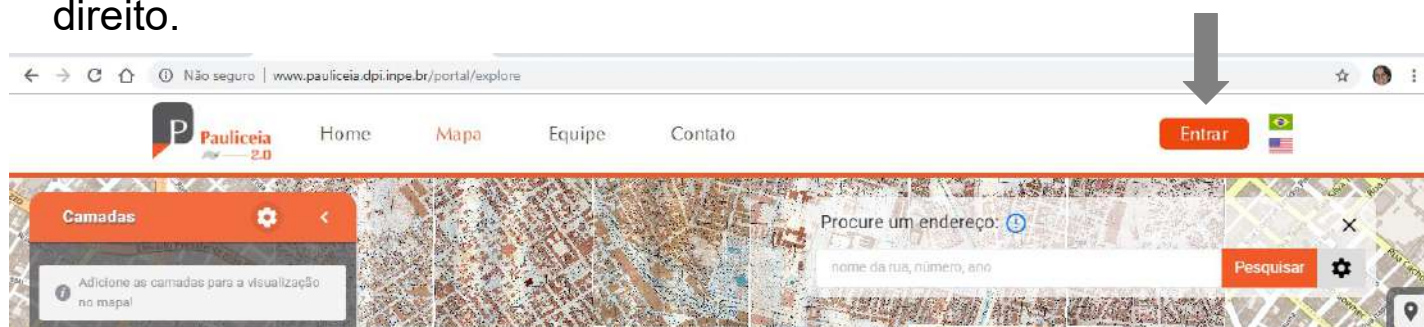
<http://www.pauliceia.dpi.inpe.br/portal/>

Nesta vista inicial, o portal nos apresenta uma parcela do mapa da cidade de São Paulo de 1930 (levantamento SARA Brasil), por de trás do mesmo está o o mapa atual da Cidade de São Paulo, apresentado via OpenStreetMap (seção 01 e 02, respectivamente). Também podemos visualizar uma série de botões e ferramentas que, nas próximas páginas, o tutorial explorará.



O portal permite a pesquisa de endereços e sobreposição de camadas/mapas sem login ou criação de usuário. No entanto, para ter acesso a todas as funcionalidades presentes na plataforma, é necessário que se faça o cadastro (criando um usuário ou utilizando um login social como Facebook ou Google).

Para o cadastro clique no botão **entrar**, localizado no canto superior direito.



re Mapa Equipe Contato

Para novo cadastro basta clicar em **registre-se agora**, digitar seus dados (nome, e-mail, nome de usuário e senha) e clicar em cadastrar.

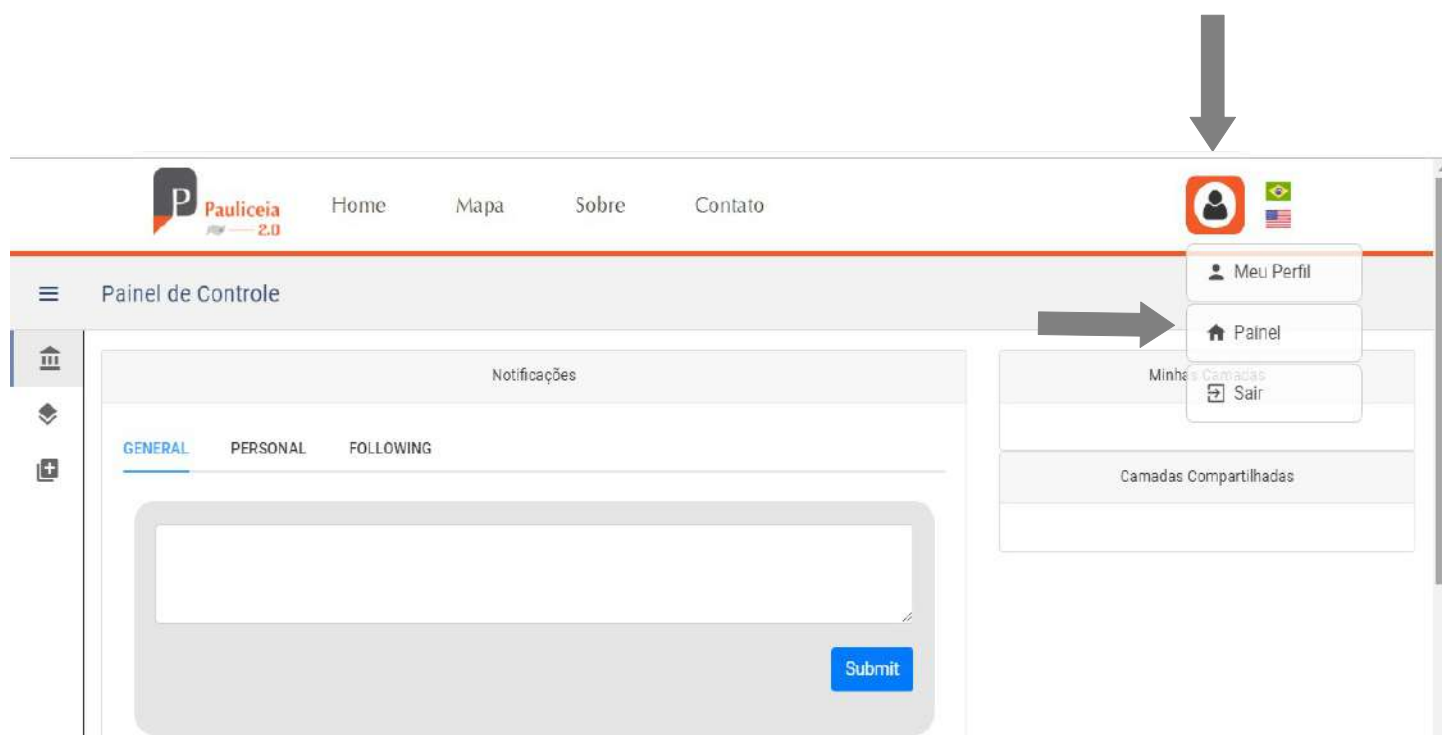
Para cadastro através do login social, basta clicar nos botões de Facebook/Google+ e inserir seus dados de acesso (caso não esteja logado na respectiva rede).

A imagem mostra o formulário de login e registro do portal Pauliceia 2.0. O formulário tem o título 'Login' e contém campos para 'E-mail' e 'Senha'. Abaixo dos campos, há um link 'Registre-se agora' e um botão verde com o texto 'Entrar'. Uma seta cinza aponta para o link 'Registre-se agora'. No rodapé do formulário, há uma mensagem: 'Se você fizer login via rede social, estará automaticamente concordando com os termos do projeto.' e um link '(leia os termos aqui)'. Abaixo disso, há dois botões: um azul com o ícone do Facebook e o texto 'Facebook', e um vermelho com o ícone do Google+ e o texto 'Google+'.

Atenção: não se esqueça de ler os termos de uso.

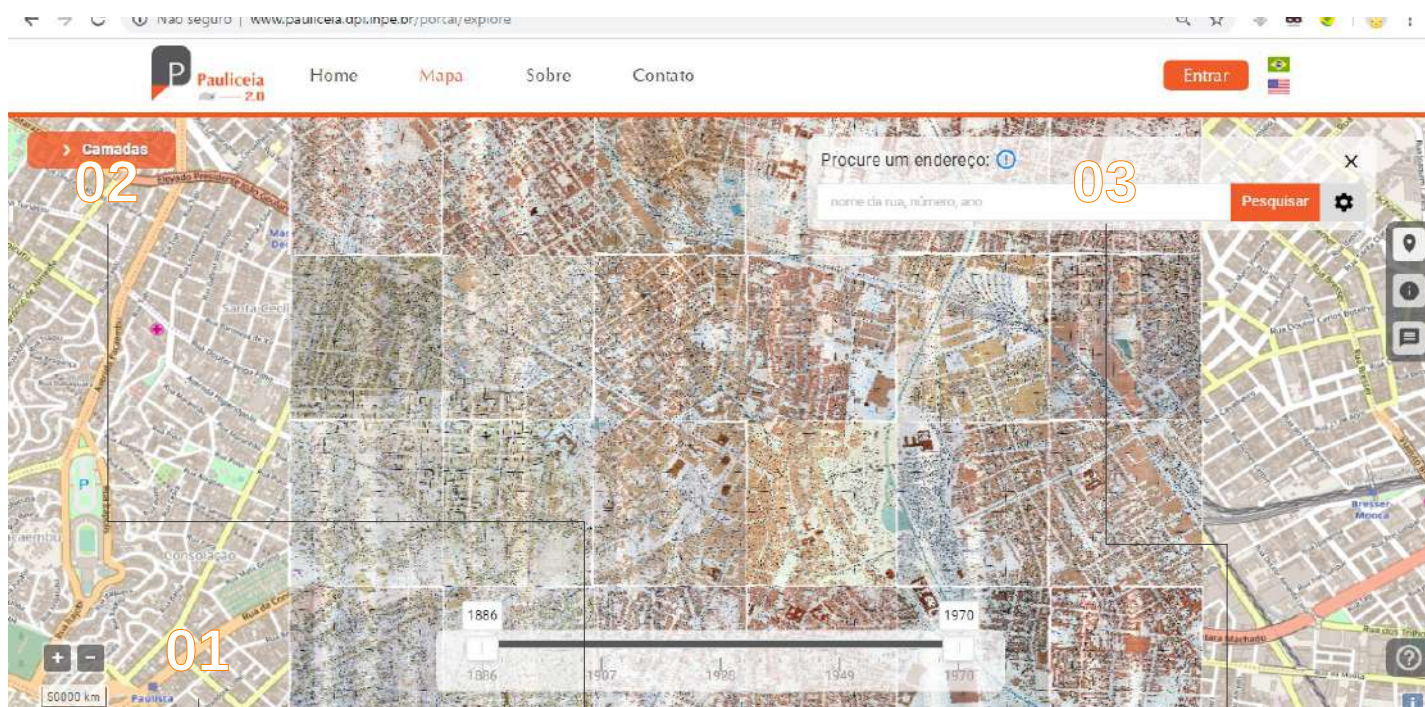
Após cadastro e login, no espaço onde antes estava o botão “entrar” agora estará um **ícone com um perfil**. Ao clicar no mesmo, você terá acesso ao seu perfil (onde poderá indicar suas informações pessoais e carregar sua fotografia), painel e logout.

Na aba **painel**, você poderá acessar informações sobre as camadas ou usuários que segue na plataforma. Essa é uma boa maneira de acompanhar as atualizações de dados sobre camadas ou temas de seu interesse. Neste espaço você também acompanhará as camadas que criou.



Principais botões da vista inicial

Na imagem abaixo destacamos algumas ferramentas que o usuário poderá utilizar para a pesquisa e manipulação dos dados presentes na plataforma. São eles: zoom (01), camadas (02) e caixa de busca (03).



O botão camada abre a aba de mapas/camadas disponíveis em toda a plataforma.

A Barra de pesquisa é a ferramenta utilizada para inserir endereço e ano que se pretende visualizar.

Aproxima e distancia o mapa ou camada que esteja sendo visualizado.

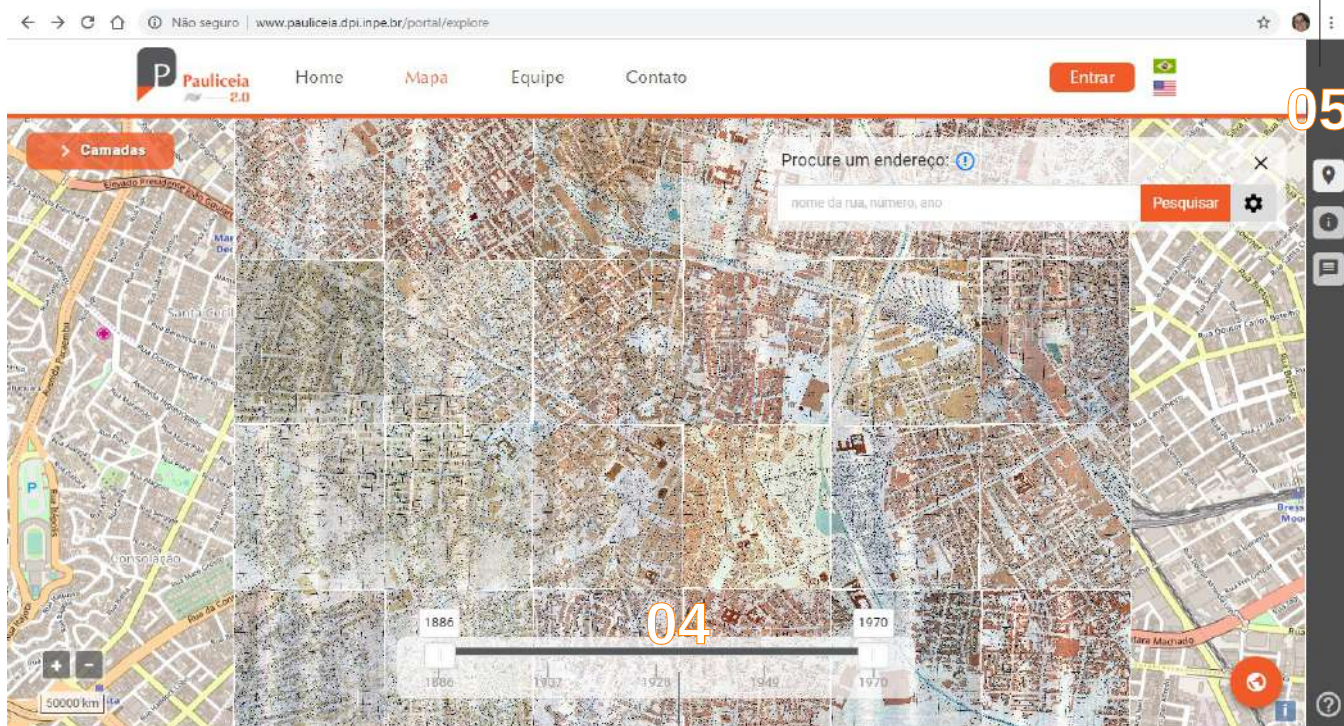
Temos ainda o *slider* espaço temporal (04) e a caixa com botões para busca, visualização, informação e notificações (05).

Abre caixa de texto com últimas notificações feitas por usuários.



Botão para que a barra de busca de endereços seja oculta ou não.

Habilita busca por localização e por região, informando ainda dados sobre os pontos selecionados (consultar próximas páginas do tutorial).



Para que essa ferramenta funcione o usuário deve ativar ao menos uma camada.

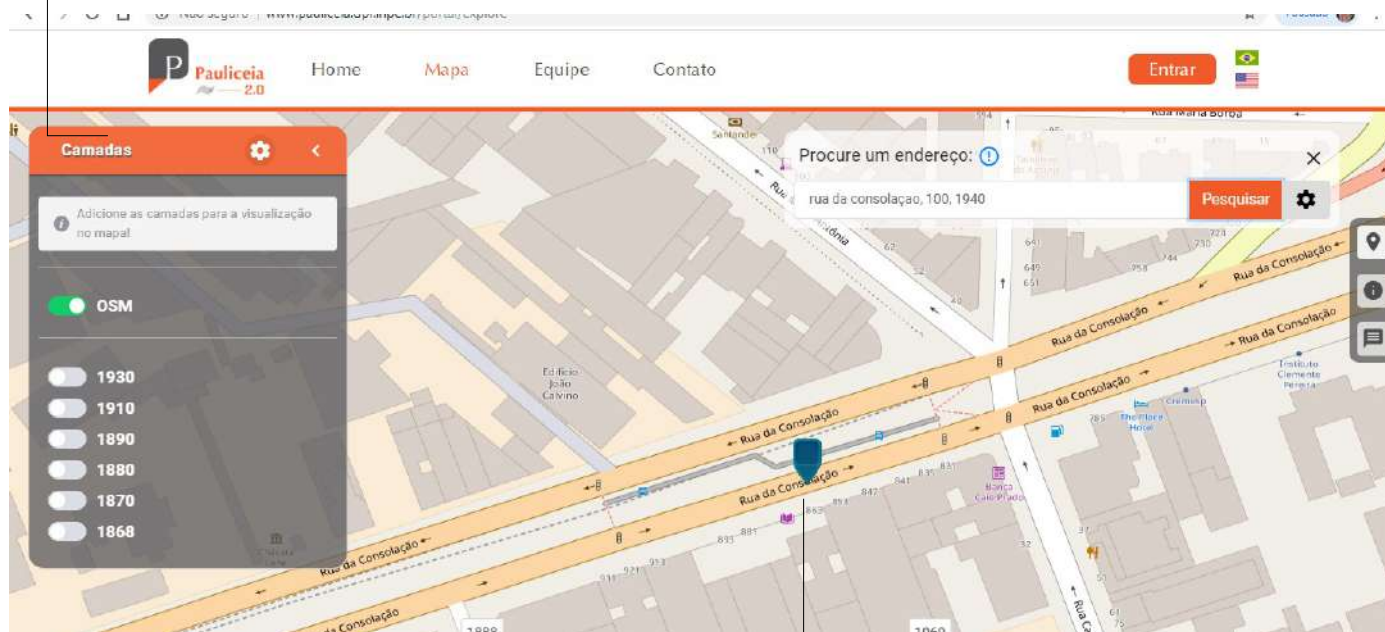
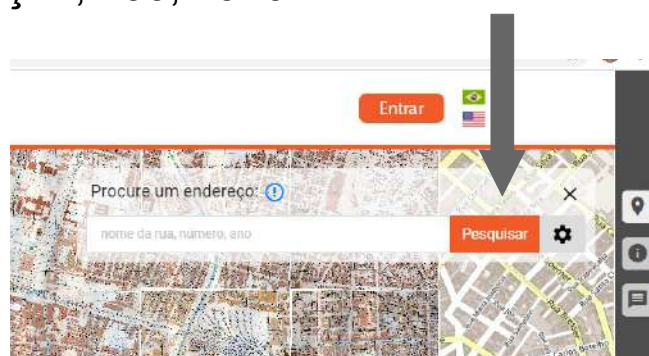
Ao arrastar a barra, você poderá visualizar as alterações sofridas em determinada região ou camada de pontos.

Buscando e Visualizando Endereços

Através da plataforma o usuário pode pesquisar um endereço específico ou uma lista de endereços. Para procurar um único endereço, devemos digitar na caixa de pesquisa a seguinte ordem de informações e depois clicar em pesquisar:

nome da rua, numero, ano
EX: rua da consolação, 100, 1940.

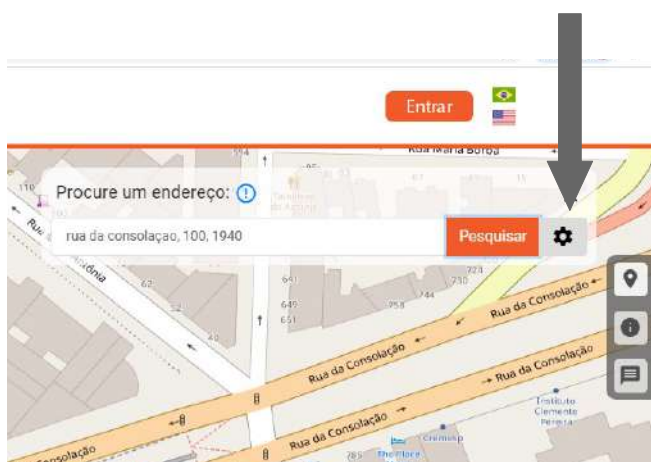
Para visualizar o mesmo ponto em algum ano específico, basta escolher a década através dos botões da janela de **camadas**. Ou utilizar o slider temporal.



Um ponto aparecerá no mapa.

Para pesquisar um grupo de endereços o usuário deverá carregar um arquivo no formato .CSV (arquivo produzido geralmente em tabela excel e salvo neste formato). Para que possa ser lido corretamente pelo portal, o arquivo deve possuir o seguinte formato:

Primeira linha: Cabeçalho (nome da rua, número, ano).
Demais linhas: Endereços escritos por extenso (cada linha com um endereço).



	A	B	C	D	E
1	rua	nro	ano		
2	rua anita	12	1945		
3	rua das araras	21	1910		
4	avenida dos est	58	1930		
5					
6					
7					

Exemplo de tabela

Com a tabela preenchida e salva em formato .CSV, clique no **símbolo de engrenagem** presente na caixa de busca de endereços, para adicionar seu novo arquivo.

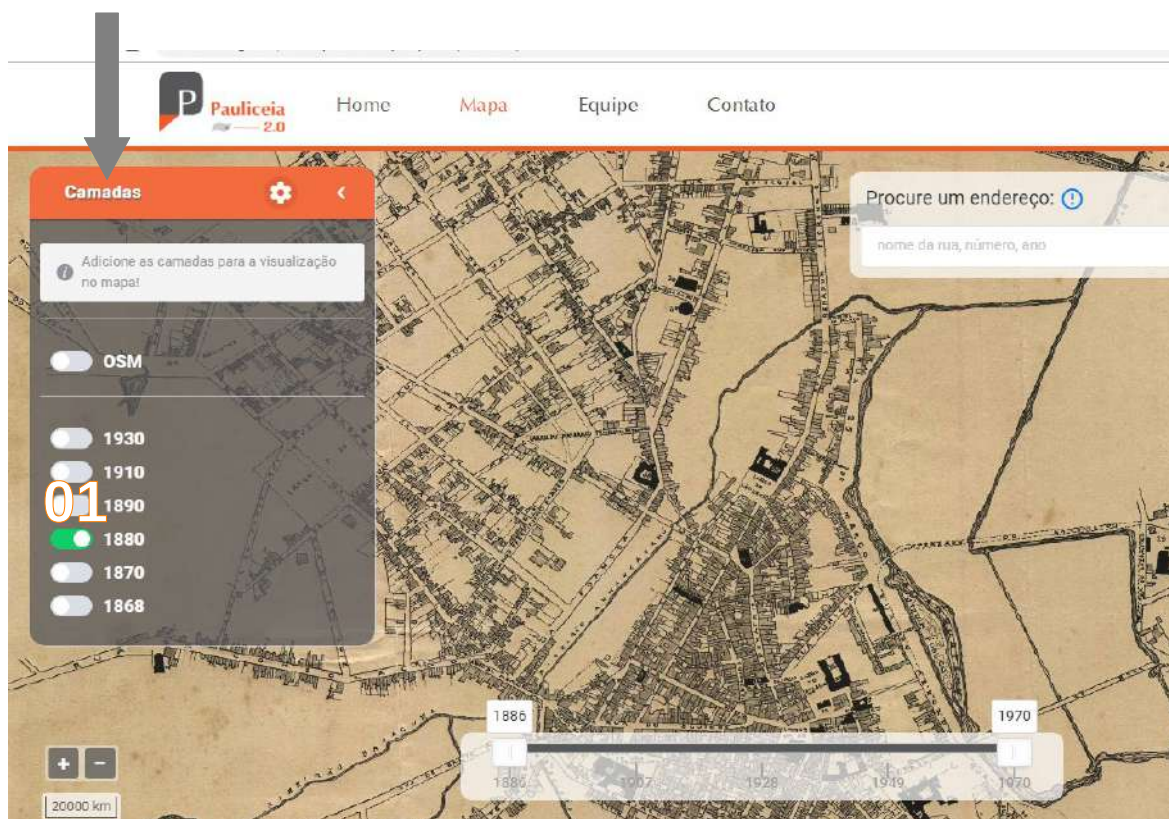


Busque a tabela em seu computador e carregue-a através do botão **escolher arquivo**. Após reconhecê-la, a plataforma solicitará que você indique qual o nome foi atribuído para as colunas rua, número e ano no arquivo.

Você poderá optar entre visualizar os endereços na própria plataforma e/ou realizar download destes dados para visualização em um software de informação geográfica de sua escolha.

Manipulando Camadas

Para acessar as camadas disponíveis clique em **camadas**. Para habilitar ou ocultar determinado mapa basta clicar nos botões assinalados (01). Os itens em verde são aqueles acionados no momento. Na figura selecionada, o mapa de 1880 foi selecionado para visualização.

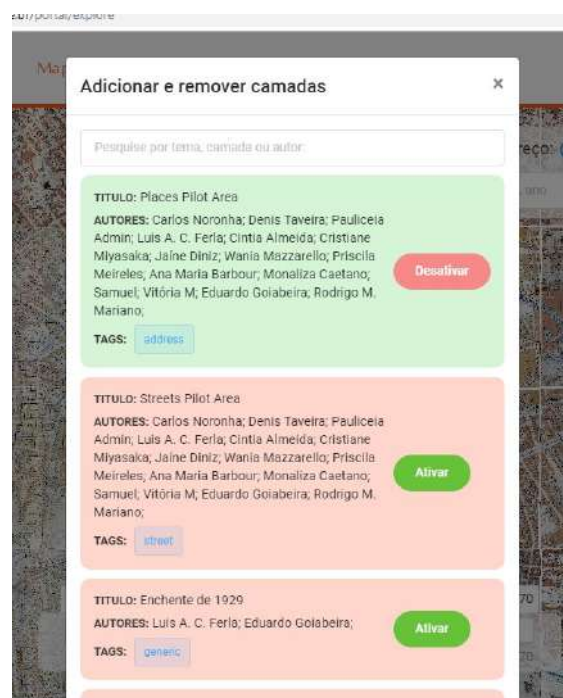


ATENÇÃO: duas ou mais camadas podem ser habilitadas para uma mesma visualização, afim de comparação e cruzamento de dados.

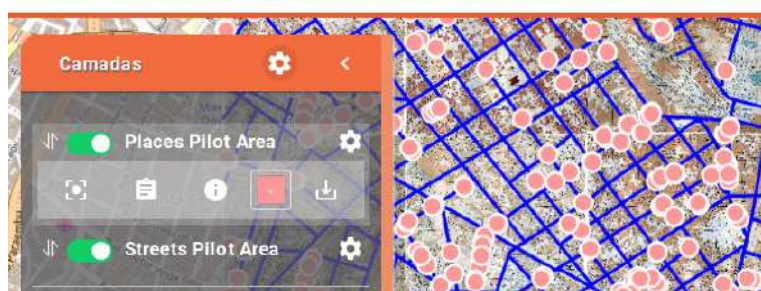
O usuário cadastrado poderá também ativar ou desativar camadas, que ficarão disponíveis para visualização na plataforma. Para ativar uma camada, clique no **botão de engrenagem** presente na caixa de camadas.



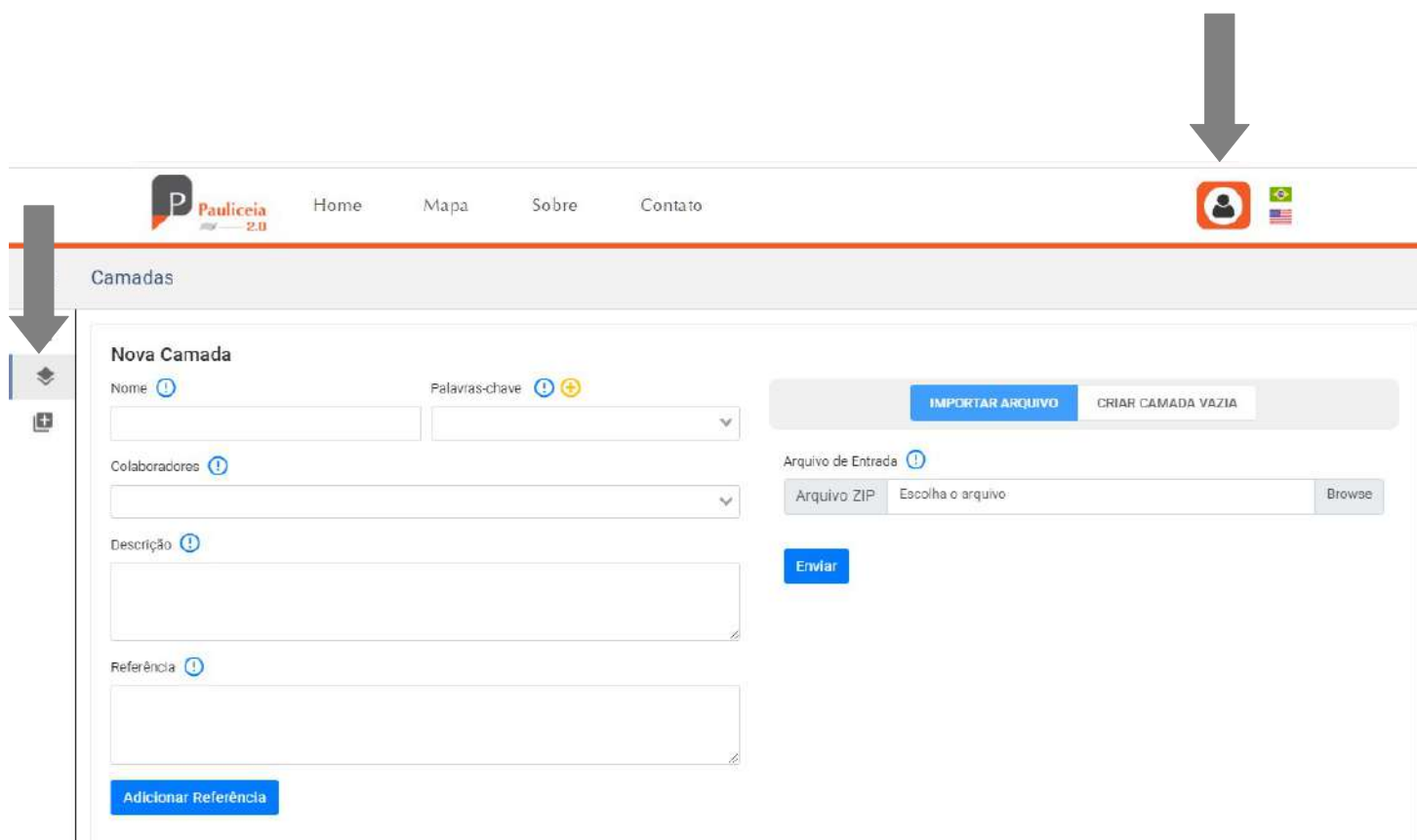
A opção de ativar (para que apareça no mapa selecionado enquanto fundo) ou desativar camadas presentes na plataforma aparecerão em uma nova janela.



Após ser ativada, essa camada passa a integrar a lista de camadas disponíveis para a visualização. No exemplo abaixo, a camada de ruas da área piloto do projeto e pontos de geolocalização de endereços estão ativados.



Há ainda a opção de criação de nova camada com dados a respeito de determinado tema e período. Acesse **sua conta** na plataforma e clique no **ícone cujo símbolo são camadas sobrepostas**.



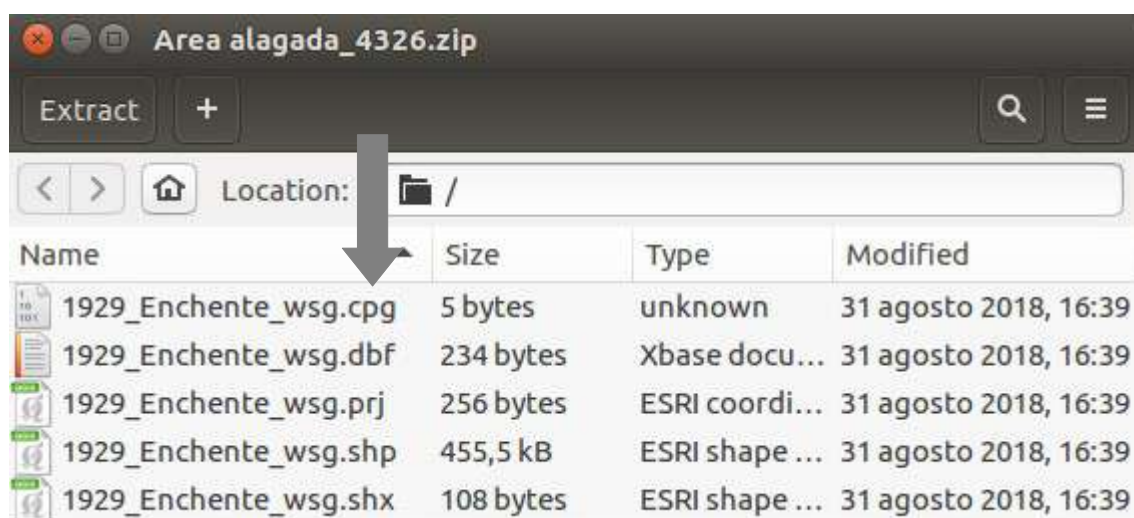
The screenshot displays the 'Pauliceia 2.0' web application. The top navigation bar contains links for 'Home', 'Mapa', 'Sobre', and 'Contato'. On the right side of the bar, there is a user profile icon and a flag icon. A sidebar on the left shows a 'Camadas' (Layers) icon, which is highlighted by a grey arrow. The main content area is titled 'Nova Camada' (New Layer) and contains several input fields: 'Nome' (Name), 'Palavras-chave' (Keywords), 'Colaboradores' (Collaborators), 'Descrição' (Description), and 'Referência' (Reference). Each field has a blue exclamation mark icon next to it. To the right of these fields, there is a section for 'Arquivo de Entrada' (Input File) with a 'Browse' button and an 'Enviar' (Send) button. At the bottom left of the form, there is a blue button labeled 'Adicionar Referência' (Add Reference).

Você deverá inserir os dados solicitados pelo portal, em cada ponto de exclamação em azul, há uma breve explicação sobre o que se espera que seja preenchido no espaço em branco. O arquivo da camada (shapefile), feito anteriormente, deverá ser colocado em pasta de formato ZIP.

O zip deve conter **somente** o shapefile que será importado na plataforma. Não devem existir subpastas, pois o portal não encontrará o arquivo. O .zip baixado pelo geolocalizador, derivado do arquivo .csv, já vem em perfeitas condições de ser importado.

Atenção: para que um Shapefile seja válido na plataforma Pauliceia 2.0, é obrigatório que ele esteja em uma pasta .zip. Essa pasta deve conter, no mínimo, arquivos com as seguintes extensões: .shp, .prj, .dbf e .shx.

Portanto, é interessante que seja verificado se o .zip contém tais arquivos, para que o shapefile responda corretamente.



















Se você gerar um Shapefile a partir de um CSV utilizando o geocoding, o Shapefile resultante já vem com os arquivos necessários.


Não se esqueça de preencher os metadados temporais da camada adicionada após o cadastro inicial .

Através destes dados, você aponta ao sistema em qual ano as informações presentes na tabela iniciam e quando terminam. Ou seja, você indicará em que período a camada existiu temporalmente (no exemplo abaixo, entre 1900 e 1930).

Além de gerar dados localizáveis em tempo e espaço, estas informações também são necessárias para que o slider temporal funcione.

Dados Temporais

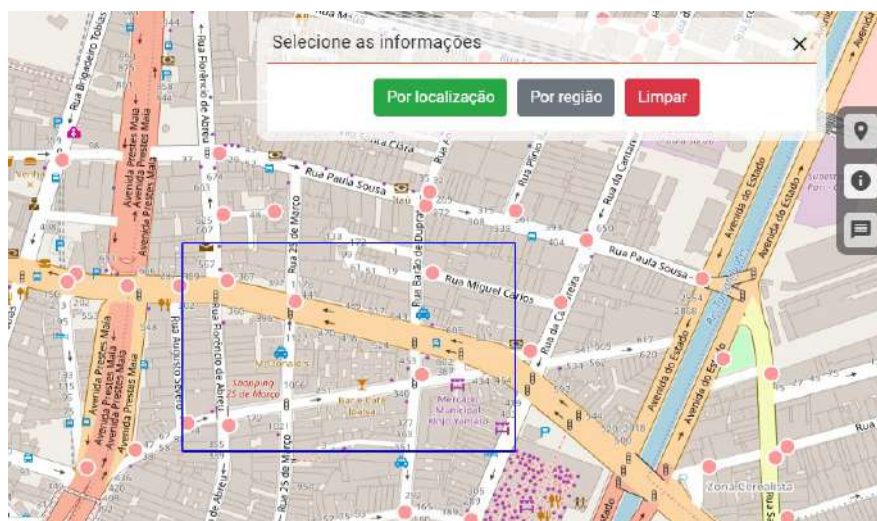
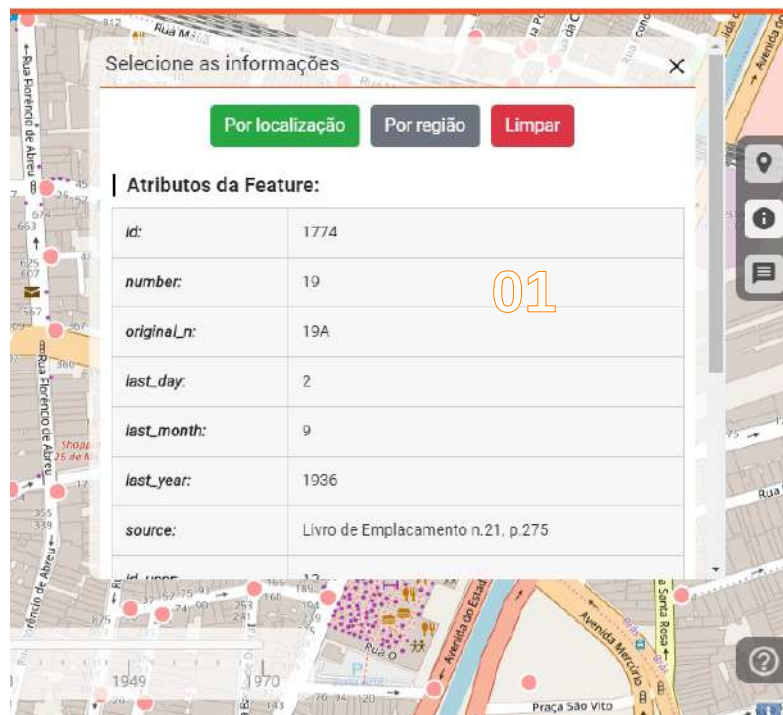
Data inicial 	Coluna da data inicial 	Mascara da data inicial 
<input type="text" value="01 / 01 / 1900"/> 	<input type="text" value="start_date"/>  	<input type="text" value="YYYY"/>  
Data final 	Coluna da data final 	Mascara da data final 
<input type="text" value="01 / 01 / 1930"/> 	<input type="text" value="end_date"/>  	<input type="text" value="YYYY"/>  

 Enviar

Visualizar Informações Sobre Determinado Ponto

Para acessar a ferramenta de seleção de informações sobre um ponto, você deve clicar no **ícone de informações** presente na lateral direita da tela.

Uma janela aparecerá, solicitando que você opte pela seleção por localização ou região. Ao clicar em determinado ponto, as informações sobre ele são descritas (01).



A seleção por área de interesse também é possível. Ao invés de um único ponto o usuário deve clicar no botão por região e arrastar o mouse no mapa até cobrir a área desejada, formando um retângulo. Informações sobre aquela região aparecerão também em uma janela.

Ficou com várias dúvidas? Não sabe exatamente como colaborar? Gostaria de enviar alguma sugestão? Entre em contato conosco:



pauliceia_support@googlegroups.com



[/grupohimaco](https://www.facebook.com/grupohimaco)

Anexo 7 – Relatórios sintéticos dos bolsistas

1. Título do projeto

Histórico da numeração dos imóveis da cidade de São Paulo e banco de dados respectivo (1870-1940)

2. Nome da Bolsista

Cintia Rodrigues de Almeida

3. Informação sobre o nível e período de usufruto da Bolsa

Bolsa de Iniciação Científica pelo Programa em Inovação Tecnológica eScience

Vigência: 01/04/2018 a 31/01/2019

4. Descrição das atividades do bolsista no projeto de pesquisa

Fiz parte juntamente com a bolsista Luanna Mendes, Eduardo Goiabeira e Vitória Martins do Projeto Pauliceia 2.0, mais precisamente do subgrupo chamado “Banco de Endereços Históricos” coordenado pela Dra. Cristiane Miyasaka. Este subgrupo teve como objetivo a alimentação de um banco de dados de endereço histórico que foi utilizado para a criação de um geolocalizador que transforma dados textuais em coordenadas geográficas.

Para tanto, foi acordado que o Inpe ficaria com a tarefa de criar uma interface para padronizar as informações inseridas no banco de dados, este portal recebeu o nome de “Web Editor de Endereços” e permitia que a equipe de Endereços Históricos editasse e organizasse os dados das numerações dos imóveis. No decorrer do ano de 2018 percebeu-se a limitação desta interface e uma nova metodologia foi adotada, sendo isto melhor explicado nas páginas seguintes.

Durante a vigência da bolsa o grupo realizou tarefas diversas, como por exemplo: a divulgação e apresentação dos resultados parciais do projeto no meio da comunidade científica; realização de um ciclo de oficinas; reuniões gerais com o subgrupo do Instituto Nacional de Pesquisas Espaciais e é claro, a alimentação do banco de dados.

Nossa primeira apresentação dos resultados parciais do projeto ocorreu no I Congresso de Humanidades Digitais, na cidade do Rio de Janeiro em abril de 2018. Ali foi possível trocar informações com os pares da comunidade científica, pegar sugestões, ouvir suas críticas e elogios. No geral, o projeto foi recebido com muito entusiasmo e também nos deu valiosos contatos de estudiosos interessados em testar a plataforma em sua versão beta.

No mesmo mês, ocorreu na Universidade Federal de São Paulo, um Ciclo de Oficinas que tinha como intuito fomentar a alimentação do banco de dados através de colaboradores. Na primeira parte do Ciclo foi apresentada aos participantes a plataforma e um tutorial de como utilizar o Web Editor, nos dias que se seguiram os participantes ficaram responsáveis de alimentar dados na interface, sendo sempre auxiliados pelas monitoras do projeto. O Ciclo de Oficinas foi de grande importância, não só por ser uma maneira de alimentar o banco de dados de forma rápida e eficiente, mas também por colocar alunos em contato com uma área da História que é pouca difundida, isto é, as Humanidades Digitais.

Dando continuação nos trabalhos de divulgação do projeto, o mesmo foi apresentado no IV Congresso Acadêmico da Unifesp, sendo recebido pelo público como uma inovação na área. Os questionamentos feitos neste evento foram diferentes daqueles do Congresso de Humanidades Digitais, já que os estudiosos que participaram do Congresso da Unifesp não tinham muito contato com o campo da História que o projeto faz parte.

Foram nas Reuniões Gerais, que juntavam as equipes da Unifesp e Inpe, que as principais decisões acerca da Pauliceia 2.0 foram tomadas. Na III Reunião Geral, por exemplo, foi decidido que as pesquisas alimentadas deveriam ter *tags*, palavras-chaves que estivessem intimamente ligadas ao conteúdo alimentado e que iriam existir três tipos de usuário na plataforma, sendo eles: normal, curador e administrador. O usuário normal seria o estudioso que alimenta a plataforma com a sua pesquisa, criando camadas; o curador ficaria responsável por determinado tema, como “História Ambiental”, assim, caberia a ele estimular pesquisadores desta temática a participar da

plataforma, fomentar debates, organizar eventos, etc; o administrador seria o único capaz de excluir camadas e *tags*.

Foi na IV Reunião Geral do Projeto que ocorreu a mudança de metodologia de inserção dados no Banco de Endereços Históricos. Até aquele momento, os bolsistas utilizavam o Web Editor para editar e organizar os dados provenientes dos Livros de Emplacamento, que era a nossa principal fonte de informações. Nesta interface, cada rua deveria ter três pontos, um no início (primeiro imóvel da rua), no meio (imóvel na metade da rua) e no final (último imóvel da rua), sendo que cada ponto continha a numeração par e ímpar, e os números antigos daquele imóvel.

Com o decorrer do projeto, percebemos que esta era uma interface demasiada simples quando comparada com a complexidade das fontes. Isto se deu, porque o Web Editor não foi criado levando em consideração a Lei nº 3427, de 19/11/1929, do engenheiro Arthur Saboya, que dizia que a numeração se dava entre a distância do meio da soleira do imóvel e o início da rua. Como as numerações ímpares e pares tinham que ficar no mesmo ponto, dava-se a falsa impressão de que os imóveis estavam próximos. Imaginemos que a primeira casa ímpar de uma rua é a de número 5, e que a par é a 212, pela lógica da utilização do Web Editor, ambas estariam no mesmo ponto, mas sabemos que devido ao Código de Arthur Saboya, na realidade elas estão há mais de 200 metros de distância. Também foi normal a demolição ou construção de imóveis, o que acabava por alterar o imóvel inicial e final dos emplacamentos de uma mesma rua.

Desta forma, se fez necessário pensar em uma maneira de inserir as informações no banco de dados levando em consideração a metragem, a solução foi alimentá-la através de um planilha do excel em que cada imóvel ganhou um "ID", isto é, um número de identificação. Assim, um imóvel que teve, por exemplo, as numerações 212, 32, 12, 2 ao longo do tempo, seria inserido na tabela com o mesmo ID, apenas se indicando o período em que cada uma das numerações foi vigente. Esta solução trouxe pontos positivos e negativos. Como positivos temos o fato de que ao utilizar a

metragem, o geolocalizador ficou muito mais preciso, entretanto, foi necessário recomeçar do zero a alimentação do banco de dados.

Em outubro, ocorreu nas dependências da Unifesp, o XXIV Encontro Regional Estadual da Anpuh. Com o intuito de divulgação, o projeto Pauliceia 2.0 foi apresentado tendo já boa parte do seu banco de dados preenchido. Assim, neste evento, já foi possível mostrar o carregamento de camadas e pesquisas de endereços históricos.

O lançamento da versão beta da plataforma ocorreu no auditório do Arquivo Público de São Paulo, no mês de outubro. Esta versão da interface permite pesquisar endereços; carregar/sobrepôr camadas criadas em programas de SIG e planilhas no formato .csv; além de promover a interação entre os usuários, sendo possível “seguir” determinado tema e comentar as alimentações feitas por outros pesquisadores.

Como a plataforma se encontra no momento na versão de testes, foi criado um e-mail de suporte para que os usuários consigam tirar dúvidas e reportar os erros que encontraram, para que assim a equipe possa otimizar o funcionamento do Pauliceia 2.0.

5. Informar e justificar caso tenham ocorrido mudanças e, eventualmente, os ajustes realizados nas atividades de pesquisa do bolsista, em relação ao Plano de Atividades

Segundo o plano de atividades, a alimentação do banco de dados da plataforma deveria ter sido finalizada antes do tempo em que realmente ocorreu. Entretanto, como já explicitado, houve uma mudança na metodologia de inserção de dados, o que fez com que a equipe tivesse que reiniciar a alimentação.

Apesar disso, creio que a troca de metodologia foi essencial para que o geolocalizador apresentasse um bom funcionamento, não há dúvidas de que se ele tivesse sido criado a partir do Web Editor apresentaria resultados muito inferiores ao que apresenta atualmente.

6. Avaliação do impacto das atividades do bolsista sobre o andamento do projeto.

A criação do Banco de Dados de Endereços Históricos foi parte essencial do Projeto Pauliceia 2.0, pois sem este material, o geolocalizador não conseguiria extrapolar os endereços pesquisados pelos usuários, o que consequentemente faria com que o principal objetivo do projeto não fosse alcançado, isto é, o de promover o mapeamento colaborativo da cidade de São Paulo.

1. Título do projeto

Histórico da numeração dos imóveis da cidade de São Paulo e banco de dados respectivo (1870-1940)

2. Nome da bolsista

Luanna Gabriely Mendes do Nascimento

3. Informação sobre o nível e período de usufruto da Bolsa.

Bolsa de Iniciação Científica pelo Programa em Inovação Tecnológica eScience

Vigência: 01/04/2017 a 31/01/2019

4. Descrição das atividades do bolsista no projeto de pesquisa

As minhas atividades dentro do projeto *“Pauliceia 2.0: uma plataforma espaço-temporal para Humanidades Digitais”* são, sobretudo, relacionadas à pesquisa sobre o histórico de mudanças na numeração dos imóveis em São Paulo, para constituir um banco de dados, essencial na formulação do geolocalizador da plataforma. Além de mim, os bolsistas Cintia Rodrigues e Eduardo Goiabeira também trabalham com esse mesmo objetivo, todos nós coordenados pela Dr. Cristiane Miyasaka.

No período anterior foi realizado um levantamento das ruas abarcadas pela área piloto para que fosse possível investigar quais fontes seriam produtivas para o nosso recorte espacial e temporal. Dentre fontes textuais como jornais, guias de ruas, processos de patrimonialização de imóveis, os livros de emplantamento foi o tipo de documento que nos disponibilizou dados mais precisos, o que possibilitou um uso mais sistemático para a construção do banco de dados.

Em Abril, realizamos um ciclo de oficinas que buscava a colaboração no preenchimento do banco de dados fora da equipe do projeto. O evento foi organizado e ministrado por mim, Cintia Rodrigues e Monaliza Caetano, sob a supervisão do professor Luis Ferla. Aconteceu entre os dias 19 a 27 de abril de 2018, foram sete encontros com duração de 1h30 cada. Sete alunos de graduação e um servidor técnico administrativo do campus participaram ativamente da construção do banco de dados, sob nossa supervisão tiveram contato com a

documentação histórica e a interface web destinada para o preenchimento dos dados.

A metodologia adotada no período anterior para o preenchimento do banco de dados, prevista no plano de atividades para esse período, se mostrou insuficiente para os dados que utilizamos.¹ A nova metodologia foi formulada entre reuniões da equipe responsável pelo histórico da numeração, em reuniões de planejamento entre os pesquisadores da EFLCH-UNIFESP e APESP e principalmente nas reuniões gerais com toda a equipe do projeto, onde pudemos discutir as necessidades computacionais e históricas do banco de dados na nova metodologia.

Geramos então um modelo de tabela com os dados necessários, e desde Agosto de 2018 trabalhamos sistematicamente com o objetivo de integrar dados de toda área piloto. As informações da numeração de imóveis de referência em cada rua eram atreladas a uma localização, com base na distância métrica do início da rua informada pela fonte, e a uma data de vigência de cada numeração indicada em cada um dos livros de emplacamento.

Como resultado tivemos um banco de dados com 4.155 entradas de informações temporais relacionadas a um espaço, que possibilita a plataforma localizar, por extrapolação, numerações na maior parte da área piloto, a partir da primeira década do século XX, que foi o que as fontes puderam viabilizar.

Além da construção do banco de dados, minhas atividades relacionadas ao projeto também se relacionaram a leitura e discussão de bibliografia pertinente em grupos de estudos que aconteceram com frequência mensal; reuniões de planejamento sobre todo o projeto também com frequência mensal.

Entre os dias 12 e 14 de dezembro eu e a bolsista Cintia Rodrigues realizamos um curso sobre banco de dados geográficos realizado no INPE- São José dos Campos para que pudéssemos entender melhor a formulação e funcionamento do banco de dados na parte computacional e assim melhor inserir os dados históricos que apresentavam menos substancialidade.

Durante o percurso de pesquisa divulgamos nossos objetivos, avanços e desafios no I Congresso Internacional em Humanidades Digitais no Rio de Janeiro,

¹ Ver mais no item 5. Ajustes realizados nas atividades de pesquisa do bolsista, em relação ao Plano de Atividades

em Abril de 2018; no IV Congresso Acadêmico da Unifesp e V Colóquio de Humanidades em Julho, no 24º Encontro Estadual de História - ANPUH-SP em Setembro e no Colóquio Histórias de São Paulo realizado na Unifesp em Outubro. Todas foram oportunidades de tornar o trabalho ainda mais colaborativo e próximo do conceito de ciência aberta.

5. Ajustes realizados nas atividades de pesquisa do bolsista, em relação ao Plano de Atividades

O ajuste mais significativo nas atividades planejadas foi a mudança de metodologia. O banco de dados que seria preenchido através de um sistema web programado para esse fim acabou sendo executado através de uma tabela, com os atributos necessários para relacionar a informação textual - as numerações vigentes em determinados períodos, fontes da informação, observações - a uma localização geográfica. A localização foi dada através das fontes que informam a distância métrica de cada imóvel em relação ao início da rua, devido ao Código Saboya de 1929, que determina que a numeração do imóvel será relativa a essa metragem. Para os períodos anteriores a essa regulamentação relacionamos os números, regressivamente, aos dos emplacements que constam a metragem.

Com a tabela conseguimos usar essas informações que garantiram uma precisão muito maior para o geolocalizador, mas como todos dados precisaram ser reinseridos no novo padrão, o cronograma foi reajustado para que pudéssemos finalizar essa tarefa e não foi possível buscar novas fontes para as duas primeiras décadas do recorte, período que os livros de emplacements não cobrem.

6. Avaliação do impacto das atividades do bolsista sobre o andamento do projeto.

Considero que as atividades aqui descritas foram essenciais para a realização do projeto. Pois além de participar da discussão sobre todo o projeto, as minhas ações nessa pesquisa proporcionaram a construção do banco de dados que subsidia a geolocalização de endereços históricos, uma função primordial na plataforma.

Todas as apresentações que participei também tiveram um impacto positivo na pesquisa, já que durante essas apresentações recebemos críticas, sugestões para os desafios enfrentados e conhecemos as demandas do principal público-alvo do projeto.

1. Título do projeto

Histórico da numeração dos imóveis da cidade de São Paulo e banco de dados respectivo (1870-1940)

2. Nome do bolsista

Eduardo de Souza Goiabeira

3. Informação sobre o nível e período de usufruto da Bolsa.

Bolsa de Iniciação Científica pelo Programa em Inovação Tecnológica eScience
Vigência: 01/08/2018 a 01/01/2019

4. Descrição das atividades do bolsista no projeto de pesquisa

Na execução do projeto “*Pauliceia 2.0: uma plataforma espaço-temporal para Humanidades Digitais*”, em parceria com as bolsistas: Luanna Mendes, Cintia Rodrigues e Vitória Martins, sob a coordenação da Dr^a Cristiane Miyassaka, nos encarregamos da construção e alimentação de um banco de dados de endereços históricos que serviu de base informacional para a nossa ferramenta digital localizar coordenadas geográficas de lugares de São Paulo dentro do recorte estipulado no projeto (1870-1940).

Para a execução desta atividade foi necessário padronizarmos a forma como alimentamos os dados para o banco de dados por meio do desenvolvimento de metodologias que buscassem a maior obtenção de dados históricos de possível espacialização. Neste processo, o INPE (Instituto Nacional de Pesquisas Espaciais) desenvolveu um website que possibilita a inserção, edição e organização instantânea do banco de dados. No decorrer da utilização desta ferramenta percebemos a limitação deste método e, após reuniões com todo o grupo envolvido no projeto, nova metodologia foi criada, mais detalhes sobre está serão fornecidos com o desenvolver deste relatório.

Para a utilização do Webeditor era necessário o preenchimento de três pontos: um inicial, um intermediário e um final, tanto para o lado ímpar e para o lado par de cada rua. Estes pontos eram preenchidos com base nas informações disponíveis nos Livros de Emplacamento, principal fonte primária para a construção do banco de numeração.

O webeditor não levava em consideração a metragem dos imóveis em relação à soleira, em muitos casos podemos perceber que a metragem dos imóveis pares e ímpares eram distantes demais para colocarmos pontos pares e ímpares em um mesmo lugar na ferramenta, além de não considerar também possíveis demolições e crescimento de ruas. Estas descobertas fornecem dados de extrema relevância para a construção do banco de dados e precisávamos de uma metodologia que acompanhasse nossas novas necessidades. Esta foi elaborada na 4ª Reunião geral do projeto.

As reuniões quinzenais foram essenciais para a criação do banco de dados. Reuníamos-nos para discutir os obstáculos encontrados na construção do banco de dados, como por exemplo, a obtenção de fontes com informações de possível espacialização dentro de nossa área piloto dentro outras atividades. O grupo Hímaco também realizou mensalmente reuniões com o objetivo de estudar obras que possibilitassem reflexões sobre o trabalho colaborativo, a livre circulação do conhecimento propostos pela Ciência Aberta, a utilização de geotecnologias na produção historiográfica, além do crescimento e urbanização da cidade de São Paulo. Estes encontros revelam a preocupação do grupo Hímaco em não automatizar o conhecimento historiográfico e sua construção. E são abertos para todos os estudantes ou pesquisadores interessados.

Além das reuniões do grupo de estudo realizamos reuniões para apresentar os resultados parciais do desenvolvimento do projeto em questão e discutir o aprimoramento de nossa metodologia em conjunto com as demais integrantes. Destes eventos, um que merece destaque especial foi a 4ª Reunião do Projeto realizada no final de agosto no campus Guarulhos da UNIFESP.

Neste dia foi apresentado um novo layout da plataforma, além da proposta, discussão e aprovação da metodologia que nos permitiu avançar com a composição do banco por meio da utilização de uma tabela que incorporasse a utilização do sistema métrico dos Livros de Emplacamento. Nesta “tabela o campo “ID do ponto” e ID da rua” possibilitaram posicionar com maior precisão os endereços nos mapas de São Paulo.

Podemos destacar também a 3ª Reunião geral do projeto como um evento que possibilitou avanços significativos. Realizada no INPE no final de maio, a reunião contou com a apresentação de resultados parciais do banco de numeração, apresentação de um novo layout da plataforma e a definição do perfil dos usuários

da plataforma, sendo eles: usuário comum, curador e administrador. O usuário comum pesquisa e alimentam dados formando camadas. Os curadores organizariam as camadas em temas e fomentariam debates e eventos sendo responsável por estimular os pesquisadores a utilizarem a plataforma. Os administradores seriam capazes de excluir camadas e palavras-chaves de pesquisa, além de prestar suporte aos demais usuários.

Além das atividades relacionadas à construção do banco de dados, participei também da organização e apresentação de eventos de divulgação da plataforma. Pode-se destacar o XXIV Encontro Regional Estadual da ANPUH realizado no Campus Guarulhos da UNIFESP. Neste evento já foi possível apresentar a utilização de ferramentas da plataforma como a pesquisa de endereços e o carregamento de camadas referentes a pesquisas.

Um marco da divulgação foi o evento de lançamento da versão beta da plataforma ocorrido no Arquivo Público do Estado de São Paulo no dia 30/10/2018. A versão beta apresentada possibilita os usuários a pesquisarem endereços, criar camadas com o auxílio de programas que lidam com SIG e planilhas com grandes quantidades de endereços, seguir e comentar “tags” de pesquisa, em um esquema semelhante a redes sociais. A plataforma atualmente esta aberta para teste públicos, devido a este fato criamos um e-mail de suporte para os usuários que esperamos que seja utilizado para reportar erros e tirar dúvidas consequentemente possibilitando a melhoria da plataforma.

5. Informar e justificar caso tenham ocorrido mudanças e, eventualmente, os ajustes realizados nas atividades de pesquisa do bolsista, em relação ao plano de atividades.

O ajuste realizado em relação ao plano de atividades inicial mais significativo foi a mudança na metodologia de alimentação do banco de dados da numeração dos imóveis. Anteriormente utilizávamos o Web-Editor, site criado pelo INPE no qual inseríamos manualmente pontos iniciais, intermediários e finais para cada rua, em um lado e duplicávamos para o outro lado da rua sem considerar a metragem dos imóveis. Como descrito acima a não consideração do sistema métrico na

metodologia impossibilitava a inserção de informações em geográficas em seus respectivos locais e devido a esse fato desenvolvemos uma nova metodologia.

A tabela proposta e discutida na 4ª Reunião Geral do projeto possibilitava a incorporação da metragem como informação crucial e conseqüentemente gerava maior precisão dos algoritmos da plataforma. Ao encontrarmos uma metodologia mais adequada para as nossas necessidades reiniciamos as atividades de alimentação do banco de dados. Acreditamos que a utilização do webeditor não traria os resultados satisfatórios como foi com a utilização das tabelas que permitem a inserção de mais dados com mais facilidade no banco de numeração dos imóveis.

6. Avaliação do Impacto das atividades do bolsista sobre o andamento do projeto.

As atividades acima descritas foram fundamentais para a execução do projeto em questão. A equipe de numeração foi a grande encarregada por montar um banco de dados sólido com informações fundamentadas na análise de documentação histórica. Sem este banco de dados a plataforma não seria capaz de geolocalizar, aproximar e exibir endereços históricos, característica fundamental desta ferramenta digital. Também é necessário destacar o engajamento na divulgação do projeto entre colegas que possibilitou o contato com diferentes opiniões sobre a proposta que foi discutida em reuniões como, por exemplo, o layout.

1. Título do Projeto

O uso do Sistema de Informações Geográficas em pesquisas históricas

2. Nome da bolsista

Ester Dantas Reis Nunes

3. Período de usufruto da bolsa

Bolsa de Iniciação Científica pelo Programa em Inovação Tecnológica eScience.

Vigência: 01/04/2017 a 31/01/2019

4. Descrição das atividades do bolsista no projeto de pesquisa

As atividades do bolsista em questão no projeto de pesquisa Pauliceia 2.0: uma plataforma espaço-temporal para humanidades digitais, consistem em pesquisar o uso do Sistema de Informações Geográficas (SIG) em pesquisas históricas, levantando em conta a origem e trajetória das geotecnologias em investigações históricas no Brasil e fora dele.

No período anterior, a pesquisa concentrou-se em uma capacitação teórica voltada para o uso do SIG histórico e a elaboração de um inventário com SIGs encontrados na Internet. O inventário foi dividido em dois principais tópicos: SIGs nacionais e SIGs urbanos. No atual período, a pesquisa concentrou-se nos seguintes nichos: pesquisa de novos SIGs urbanos na Internet que tivessem a presença de tecnologias livres; SIGs Históricos elaborados por instituições brasileiras; participação do I Congresso de Humanidades Digitais – HDRIO, realizado em abril de 2018 na cidade do Rio de Janeiro; elaboração de um texto para ser discutido no Grupo de Estudos; elaboração de um SIG com os SIGs do inventário.

Sobre os novos SIGs, foi adicionado ao inventário o SIG urbano The NYC Space/Time Directory, um SIG elaborado pela The New York Public Library (disponível em <<http://spacetime.nypl.org>> acesso em 29 jan. 2019), com o recorte temporal de

1860 a 1939. O projeto utiliza o *software* livre Quantum GIS (QGIS) e foi uma inspiração para a construção do projeto Paulicéia 2.0.

A respeito de SIGs elaborados por instituições brasileiras, foi encontrado o Atlas Digital da América Lusa, elaborado pela Universidade de Brasília. O Atlas, assim como o Paulicéia 2.0, é uma proposta colaborativa, no qual há ênfase no trabalho de equipe e troca livre de informações. O projeto utiliza os *softwares* livres QGIS e i3GEO, *software* brasileiro livre e público, integrado com o GVSIG. O recorte temporal do projeto compreende o período entre 1500 a 1800 e o espacial, a América portuguesa (atual território do Brasil).

A respeito da participação do I Congresso de Humanidades Digitais – HDRIO, o Grupo Hímaco apresentou o projeto Paulicéia 2.0, fato que além de trazer visibilidade para o projeto, permitiu aos componentes terem uma experiência e contato com outros âmbitos das humanidades digitais. Visto que no congresso estavam presentes professores que são referência no Brasil e no exterior no âmbito de SIGs históricos, sendo um deles organizador de um dos SIGs que se encontram no inventário - Tiago Gil (UNB), organizador do Atlas Digital da América Lusa -, foi feita uma entrevista com cada um visando uma melhor compreensão entre o SIG histórico e a universidade. Os professores entrevistados são Tiago Gil (UNB), Carlos Valencia (UFF) e Daniel Alves (Universidade Nova de Lisboa), respectivamente.

Em conjunto com o professor orientador, foi proposto que fosse elaborado um SIG dos SIGs pesquisados que constam no inventário, com o objetivo de promover a visualização da espacialização dos SIGs, utilizando o *software* gvSIG. Visto que o inventário possui três recortes principais – SIGs nacionais, SIGs urbanos, SIGs brasileiros -, o SIG foi elaborado da seguinte maneira: foram baixados quatro tipos de shapes, sendo estes um shape do mapa mundi (disponível em <<http://murilocardoso.com/materiais/>> acesso em 30 jan. 2019); um shape dos estados dos EUA (disponível em <<https://tapiquen-sig.jimdo.com/descargas-gratuitas/estados-unidos/>> acesso em 01 fev. 2019); um shape das cidades brasileiras (disponível em

<<http://murilocardoso.com/materiais/>> acesso em 01 fev. 2019); um shape das cidades estadunidenses (disponível em <<https://tapiquen-sig.jimdo.com/descargas-gratuitas/estados-unidos/>> acesso em 01 fev. 2019).

A partir disso, foram criadas quatro camadas, sendo essas: 1 – Mundo, para os SIGs nacionais SIG Histórico da China (CHGIS), SIG Histórico da Alemanha (HGIS Germany), SIG Histórico da Grã Bretanha, História das Fronteiras Administrativas da Rússia, SIG Histórico Nacional dos EUA (NHGIS), SIG Histórico Belga, SIG Histórico dos Países Baixos (NLGIS), The Down Survey of Ireland e também o Atlas Digital da América Lusa que compreende o atual território brasileiro; 2- Oeste EUA, para o SIG “Mapeando correios do oeste dos EUA no século XIX”; 3- Rio de Janeiro, para os quatro SIGs com o recorte espacial na cidade do Rio de Janeiro: “O mercado de escravos no Rio de Janeiro: Movimento, Contexto e Experiência Social”, “Localizador de endereços históricos do Rio de Janeiro”, “Epidemia de febre amarela de 1850: escala, tempo e espaço”, “ImagineRio”; 4- Cidades EUA, onde estão representados os demais SIGs urbanos, “Mapeamento do declínio: São Luis e a Cidade Americana”, “The Digital Harlem”, “SIG Histórico de Chico”, “Mapas Históricos de Cleveland”, “A Evolução do Espaço de Abatedouro, 1852-1870 em São Francisco”, “The NYC Space/Time Directory”.

Visto que o enfoque da pesquisa foi voltado a maior parte do tempo para a área teórica, trabalhar com o gvSIG foi um novo aprendizado, pois foi possível compreender melhor o *software* que utilizamos em nosso trabalho no que diz respeito ao seu funcionamento e suas ferramentas.

À vista de todas essas experiências obtidas durante os dois anos de pesquisa, foi elaborado um texto que foi discutido no grupo de estudos.

5. Avaliação do impacto das atividades do bolsista sobre o andamento do projeto.

A partir das atividades que foram desenvolvidas nesse período e o impacto delas sobre o andamento do projeto Pauliceia 2.0, a pesquisa em questão contribui sobre uma parte teórica do uso do SIG, trazendo a contextualização necessária e sua trajetória nas pesquisas históricas. Como foi possível observar ao longo desse período, as humanidades digitais têm potencializado a interdisciplinaridade, o trabalho colaborativo e trazem uma ciência digital mais colaborativa. Entretanto, embora a tecnologia esteja muito presente em nossas vidas e as humanidades digitais venham crescendo, esta passa por desafios no âmbito do SIG em pesquisas históricas, em que há uma certa resistência por parte dos receptores dessa tecnologia e até mesmo um desconhecimento acerca desse campo.

Entretanto, foi possível estabelecer os diferenciais do projeto Pauliceia 2.0 - defesa do trabalho colaborativo, uso de tecnologia livre, etc - e seus desafios, que consistem em utilizar dessas metodologias para que despertem o interesse dos estudantes em reconhecer a importância desse âmbito de pesquisa e, posteriormente, sintam-se familiarizados para poderem participar como pesquisadores.

VGI Protocol and Web Service for Historical Data Management

1. Nome do Bolsista

Rodrigo Monteiro Mariano

2. Informação sobre o nível e período de usufruto da Bolsa

Bolsas Concedidas como Itens Orçamentários em Auxílios / BCO - Mestrado - Fluxo Contínuo. Período de 30/01/2018 à 31/01/2019.

3. Descrição das atividades do bolsista no projeto de pesquisa

3.1. Introdução

O projeto Pauliceia 2.0 tem como objetivo desenvolver uma plataforma computacional online para o gerenciamento colaborativo de dados históricos. Os pesquisadores poderão produzir mapas e visualizações de suas próprias pesquisas, contribuindo para os dados do sistema. Uma atividade online com participação da comunidade em um propósito, é conhecido como crowdsourcing e quando esses voluntários contribuem com informações geográficas, é conhecido como Volunteered Geographic Information (VGI). O projeto Pauliceia 2.0 enriquece a compreensão da história de São Paulo durante o período de 1870 a 1940 usando técnicas de VGI, criando assim um modelo inovador de pesquisa na área de Humanidades Digitais.

Uma questão importante relacionada a VGI é a qualidade dos dados coletados pela comunidade, pois eles podem não garantir qualidade. A imprecisão dos dados VGI é explicado pelo fato de que os seres humanos expressam as regiões geográficas e suas relações através de conceitos vagos. Por conta disso, fornecer serviços confiáveis ou retirar informações úteis desses dados, requer das contribuições pelo menos um padrão de qualidade. Para que isso ocorra e os possíveis problemas possam ser minimizados, são utilizadas medidas, indicadores ou abordagens de qualidade, e também a definição de um protocolo VGI. A falta de especificação, padrões e protocolos para projetos VGI, são fatores que podem causar a má qualidade nos dados, por isso eles são importantes para facilitar e ampliar a reutilização dos dados VGI [Senaratne et al. 2017] [Goodchild and Li 2012] [Mooney et al. 2016].

Para ajudar a melhorar a qualidade dos dados colaborativos coletados pela plataforma Pauliceia 2.0, foi desenvolvido primeiramente um protocolo VGI para lidar com os dados históricos do projeto. Depois disso, foi projetado e construído um Serviço Web de Gerenciamento de VGI (VGI Management Web Service, VGIMWS) para gerenciar os dados espaço-temporais (i.e. dados históricos), coletados colaborativamente pela plataforma Pauliceia 2.0. Este serviço web, junto com os demais desenvolvidos pela equipe Pauliceia 2.0, estão disponíveis no Github do projeto¹ e estão atualmente sendo executados em containers Docker em um servidor no INPE.

3.2. Atividades desenvolvidas

Durante o período deste relatório, o estudante escreveu um artigo científico (i.e. "VGI Protocol and Web Service for Historical Data Management") e ajudou a escrever outro

¹<https://github.com/pauliceia>

(i.e. "A Platform for Collaborative Historical Research based on Volunteered Geographical Information"). Além disso, o estudante também está finalizando a sua dissertação de Mestrado.

O estudante participou do evento científico GEOINFO em Dezembro de 2018, em Campina Grande na Paraíba, onde ele apresentou o seu artigo científico, citado anteriormente, e fez a apresentação de um demo do projeto Pauliceia 2.0, ambos oralmente. Para ir a esse evento e apresentar seu artigo, o aluno foi subsidiado com recursos da FAPESP. Ele também participou do evento Worcap em Agosto de 2018, que foi realizado no INPE, onde ele apresentou uma prévia da sua dissertação de Mestrado. Ele ajudou também a organizar o evento de lançamento da plataforma Pauliceia 2.0 no Arquivo Público do Estado de São Paulo em Outubro de 2018.

Neste período o estudante finalizou o Protocolo VGI para dados históricos e o VGIMWS baseado neste protocolo, no contexto do projeto Pauliceia 2.0, que são relacionados a sua pesquisa de Mestrado. Além de que, foi finalizado a documentação do VGIMWS, que está disponível no Github do serviço².

O estudante também ajudou a gerenciar e organizar servidor do projeto Pauliceia 2.0, onde estão os containers Docker do projeto, que contêm as aplicações sendo executadas, inclusive o próprio VGIMWS. Ele ajudou a dar manutenção no servidor de proxy reverso do Pauliceia 2.0 que é o Nginx, configurando seus arquivos para as necessidades do projeto.

Durante este período, o estudante participou de reuniões semanais com a equipe do projeto Pauliceia 2.0, principalmente com o time do INPE, enquanto foram feitas reuniões quinzenais ou mensais com o time da UNIFESP Guarulhos.

Além disso, o estudante escreveu dois relatórios para a FAPESP, sendo um completo e outro sintético, contendo suas principais atividades desenvolvidas durante o período e outras informações importantes, tais como artigos escritos e uso da Reserva Técnica.

4. Informar e justificar caso tenham ocorrido mudanças e, eventualmente, os ajustes realizados nas atividades de pesquisa do bolsista, em relação ao Plano de Atividades

Durante o período deste relatório, o aluno finalizou o que tinha planejado e a única mudança significativa que ocorreu nas atividades de pesquisa dele, foi que ele ficou responsável por gerenciar, organizar e cuidar do servidor do projeto Pauliceia 2.0.

5. Avaliação do impacto das atividades do bolsista sobre o andamento do projeto.

O principal objetivo do projeto Pauliceia 2.0 é a construção de uma plataforma online para gerenciar os dados históricos espaço-temporais providos de pesquisas de historiadores sobre a cidade de São Paulo entre 1870 à 1940. Estes dados espaço-temporais são compartilhados pelos pesquisadores de história e seus alunos, criando assim um modelo inovador de pesquisa em Humanidades Digitais.

²<https://github.com/Pauliceia/vgiws/blob/master/doc/README.md>

Quando voluntários contribuem com informação geográfica em sistemas, isto é conhecido como VGI. Foi usada técnicas de VGI para a coleta e compartilhamento colaborativo dos dados espaço-temporais no projeto Pauliceia 2.0. A plataforma fornece maneiras para a comunidade contribuir com informações históricas (i.e. importação de um arquivo Shapefile) e para controle de qualidade provido pelos voluntários (i.e. notificações e denúncias). Para que isto fosse possível, foi projetado e construído um Serviço Web de Gerenciamento de VGI (VGIMWS) para lidar principalmente com os dados espaço-temporais do projeto Pauliceia 2.0.

O desenvolvimento do VGIMWS foi importante para o intercâmbio dos dados históricos do portal online com o lado do servidor. O VGIMWS atende às necessidades descritas no Protocolo VGI para dados históricos, fornecendo todas as funcionalidades essenciais para lidar com informações geográficas históricas derivadas de cidadãos. O VGIMWS trabalha principalmente com: (1) controle do usuário, gerenciando suas funções apropriadas (e.g. usuários fora de uma camada não podem editar seus dados); (2) gerenciamento de dados espaço-temporais, manipulando conjuntos de dados históricos com seus atributos, como informações geográficas e temporais, e a anexação de documentos históricos; e (3) fornecer feedback de/para a comunidade, através das notificações e denúncias. VGIMWS faz a manipulação desses dados colaborativos, seu gerenciamento e por fim o armazenamento deles no banco de dados PostgreSQL com extensão espacial PostGIS.

6. Uso da Reserva Técnica

Durante o período deste relatório, o bolsista fez uso da Reserva Técnica concedida pela FAPESP, para que ele fosse apresentar o seu artigo científico (i.e. "VGI Protocol and Web Service for Historical Data Management") no evento científico GEOINFO em Dezembro de 2018, em Campina Grande na Paraíba. Neste evento, ele apresentou o seu artigo científico, citado anteriormente, e fez a apresentação de um demo do projeto Pauliceia 2.0, ambos oralmente.

O valor total dos gastos foi R\$ 2061,96, sendo que: R\$ 1340,09 de passagem aérea, R\$ 180,00 de inscrição do evento e R\$ 541,87 de diárias. A prestação de contas foi feita no sistema Agilis, como solicitado pela FAPESP, em Dezembro de 2018.

Referências

- [Goodchild and Li 2012] Goodchild, M. F. and Li, L. (2012). Assuring the quality of volunteered geographic information. *Spatial statistics*, 1:110–120.
- [Mooney et al. 2016] Mooney, P., Minghini, M., Laakso, M., Antoniou, V., Olteanu-Raimond, A.-M., and Skopeliti, A. (2016). Towards a protocol for the collection of vgi vector data. *ISPRS International Journal of Geo-Information*, 5(11):217.
- [Senaratne et al. 2017] Senaratne, H., Mobasher, A., Ali, A. L., Capineri, C., and Haklay, M. (2017). A review of volunteered geographic information quality assessment methods. *International Journal of Geographical Information Science*, 31(1):139–167.

Portal Web e Serviços OGC Web

1. Nome do Bolsista

Carlos Alberto Ferreira de Noronha

2. Informação sobre o projeto

Processo: 2018/08044-0

Vigência: 01/07/2018 à 31/10/2018

Período coberto pelo relatório em questão: 01/07/2018 à 31/10/2018

3. Atividades Desenvolvidas

Esse trabalho iniciado em julho de 2018 e com duração de 4 meses, possibilitou ao estudante participar de reuniões gerais com a equipe do projeto Paulicéia 2.0, dar continuidade ao desenvolvimento da plataforma web, criando o setor de visualização dos mapas e dados históricos coletados pela equipe, realizar a organização das aplicações desenvolvidas em um servidor web, orquestrados por containers em Docker e a participação do Workshop de Computação Aplicada (WORCAP 2018), realizado no Instituto Nacional de Pesquisas Espaciais.

3.1 Reunião geral

A Reunião Geral do projeto Paulicéia, realizada em 24 de agosto de 2018, no campus de estudos da UNIFESP de Guarulhos, possibilitou à equipe do projeto: discutir alguns pontos sobre a alocação dos endereços coletados no Web Editor de Endereços Históricos, aplicação web criada inicialmente para a inserção de dados espaciais da área piloto do projeto no banco de dados, modelado anteriormente pelos integrantes do Paulicéia 2.0, e visualizar o visual da plataforma criados pelos responsáveis pelo design.

3.2 Transição do novo visual para a Plataforma Web

Após a reunião do dia 24 de agosto, no qual foi apresentado o novo visual da plataforma, o estudante desse trabalho iniciou a modificação do layout criado anteriormente para o novo, desenvolvido pela equipe de design. Essa transição aconteceu de forma gradativa, para que não comprometesse o

cronograma do projeto. Inicialmente, foi modificado o setor de visualização de mapas e dados históricos na plataforma, tendo em vista, que essa parte era prioritária e estava com o protótipo finalizado. Posterior a essa implementação, outros bolsistas de computação realizaram as modificações nos painéis de administradores e demais setores do site.

3.3 Visualização de dados geográficos do projeto Paulicéia 2.0

Para o desenvolvimento do setor de visualização de informações geográficas coletadas pela equipe do projeto Paulicéia, foram realizados estudos e testes em outros sistemas similares, que disponibilizam funções para a visualização de dados espaciais, através de sistemas computacionais. Evidenciando TerraView e TerraBrasilis, criados pelo INPE, e OpenStreetMap (OSM).

Após os estudos, houve um alinhamento de ideias entre os integrantes da equipe do projeto Paulicéia, para que o setor de design pudesse criar o protótipo aplicado pelo estudante desse trabalho.

O setor de visualização de dados espaciais, disponibilizado através do link: <http://www.pauliceia.dpi.inpe.br/portal>, possui funcionalidades que possibilitam a navegação entre os mapas da cidade de São Paulo, no período de 1868 a 1940. Dentre as funcionalidades, tem-se a visualização: dos dados cadastrados na plataforma de forma colorida, dinâmica e filtrada por um período desejado, das informações e comentários dos mapas e dos atributos de cada vetor. Além disso, possibilita a edição de um vetor do mapa e o download de uma informação previamente cadastrada no sistema.

3.4 Organização do servidor para alocação das aplicações desenvolvidas

Outra tarefa executada pelo estudante foi a organização inicial de um servidor web, para a alocação das aplicações desenvolvidas pela equipe do projeto Paulicéia 2.0. Nessa perspectiva, foi iniciada uma máquina virtual, hospedada em um dos servidores do INPE, que roda um sistema operacional Linux. Essa instância está sendo configurada de acordo com as necessidades das aplicações do Pauliceia em containers, usando a tecnologia Docker.

A máquina virtual foi apontada para o DNS: www.pauliceia.dpi.inpe.br, assim, foi possível rotear as aplicações e serviços de maneira intuitiva, ou seja, para que um usuário acesse o Web editor criado, deverá entrar no link citado acima com o acréscimo de um '/edit' em sua URL, caso queira acessar a documentação da API de geolocalização, deverá adicionar '/docs/geocoding' e, assim, sucessivamente.

Esse servidor teve apenas duas portas liberadas, por motivos de segurança e organização. Essas portas são relacionadas ao banco de dados postgresql e ao servidor HTTP, chamado NGINX, o qual tem a função de orquestrador de rotas.

3.5 WORCAP 2018

O Workshop de Computação Aplicada de 2018 (WORCAP), realizado no Instituto Nacional de Pesquisas Espaciais, durante dos dias 21 a 23 de agosto de 2018, possibilitou ao estudante desse projeto apresentar um Pôster de demonstração de um dos trabalhos desenvolvidos no INPE e que estava sendo utilizado pelos integrantes do projeto Pauliceia 2.0. O trabalho conteve a apresentação do Web Editor de Endereços Históricos, um portal web que possibilita a edição de dados espaciais e que estão alocados em um banco de dados, baseado no modelo espaço-temporal, e da plataforma que estava sendo desenvolvida para o projeto Pauliceia 2.0, no qual o estudante estava vinculado.

4. Avaliação do Apoio Institucional

Para a execução das pesquisas e tarefas realizadas referentes à esse trabalho, o aluno contou com todo o apoio institucional da UNIFESP, campus de Guarulhos e do Instituto Nacional de Pesquisa Espaciais (INPE). Sendo o INPE, o maior fornecedor dos equipamentos utilizados na confecção do projeto. Além disso, foi fornecido transporte, sempre que necessário, para locomoção até a UNIFESP ou ao Arquivo do Estado de São Paulo, para participação de reuniões ou realização de pesquisa no local designado.

Outro ponto a se destacar que auxiliou no desenvolvimento e na conclusão desse trabalho, foi o acompanhamento e a colaboração dos Pesquisadores de Projeto Paulicéia 2.0, principalmente, a Dra. Karine Reis Ferreira, o Dr. Gilberto Riberio Queiroz, o Dr. Nandamudi Lankalapalli Vijaykumar e o Dr. Luis Antonio Coelho Ferla, que estavam sempre dispostos a ajudar e sanar possíveis dúvidas do aluno.

1. Título do projeto

Visualização Espaço-Temporal de Dados Históricos

2. Nome do bolsista

Denis Taveira Laurentino de Lima

3. Informação sobre o nível e período de usufruto da Bolsa

Bolsa de Iniciação Científica pelo Programa em Inovação Tecnológica eScience

Vigência: 01/06/2017 a 31/01/2019

4. Descrição das atividades do bolsista no projeto de pesquisa

As atividades desenvolvidas pelo bolsista foram:

4.1 Revisão bibliográfica:

A primeira parte do projeto foi fazer uma revisão bibliográfica sobre técnicas de visualização espaço-temporal para dados históricos. Nesta etapa, ao mesmo tempo em que o bolsista pesquisou sobre as técnicas já existentes de visualização espaço-temporal, o mesmo aprendeu sobre dados espaciais, visualização de dados e desenvolvimento de programas em linguagem Web.

Os principais artigos científicos que me ajudaram na revisão literária, foram: José (2016), Shrestha (2014) e Oliveira e Baptista (2013).

4.2 Projetar componentes Web:

Após o estudo das técnicas existentes de visualização espaço-temporal, foi possível projetar componentes voltados para o projeto, visando auxiliar ao máximo a experiência do usuário com o portal, com a visualização e interpretação dos dados históricos.

O *Slider* será um dos componentes utilizados no mapa, que terá a função de filtrar o intervalo de tempo desejado para o usuário, o mesmo possuirá um tempo inicial e um tempo final.

Outro componente será o *Seletor*, onde o usuário poderá selecionar de um conjunto de dados para ser exibidos no mapa, a fim de compará-los, e cada conjunto possuirá uma cor. Outra funcionalidade do componente *Seletor* será a de escolher quais mapas serão visualizados. O projeto consiste na visualização da cidade de São Paulo em décadas passadas, em que as ruas e construções eram diferentes das atuais, logo é necessário mais de um mapa da cidade.

E finalmente, o mapa terá um componente que mostrará, sempre que o usuário clicar em um ponto, o histórico daquela coordenada, exibindo se lá já esteve outra construção ou se possuía outro endereço.

4.3 Entrevistas:

A entrevista com Jean Souza, pesquisador do INPE, ocorreu no dia 01 de Novembro de 2017. Nela, o pesquisador mostrou o seu projeto atual, a plataforma computacional TerraMA² ¹, e com base nela e em outros projetos já desenvolvido por ele, apresentou as tecnologias utilizadas e deu dicas sobre o desenvolvimento de sistemas de informações geográficas (GIS) para Web.

4.4 Desenvolvimento dos componentes Web:

Para a visualização dos dados no espaço e no tempo, foram desenvolvidos certos componentes já citados anteriormente. Um dos componentes é o slider, que possui a função de filtrar todos os dados ativos no mapa que estão dentro de período selecionado com o mesmo. Mais especificamente, se o dado tiver data inicial menor do que o primeiro seletor ou entre eles, ou tiver data final maior que o segundo seletor ou entre eles, o mesmo será mostrado no mapa, pois significa que o mesmo esteve presente no período especificado pelo usuário.

Outro componente desenvolvido foi o seletor de camadas, este possui a função de ativar e desativar camadas, especificando quais serão mostradas no mapa, além disso é possível centralizar a camada desejada, ver informações da camada, tais como comentários, realizar um comentário na camada, ativar o modo de informações dos dados, onde se ativada aparecerá as informações dos dados que a pessoa clicar, alterar a cor dos dados da camada, e por fim, é possível realizar o download da mesma.

O seletor de camadas possui três divisões, sendo a primeira responsável por mostrar as camadas cadastradas por usuários escolhidas para serem mostradas no mapa, a segunda são as camadas base do mapa mundial, e a terceira responsável por mostrar as camadas de mapas antigos da região selecionada pelo projeto e é possível a visualização de um mapa por vez. É possível ainda escolher quais camadas cadastradas por usuários a pessoa deseja visualizar na primeira divisão do seletor, mostrado na Figura 6, permitindo a pessoa trabalhar apenas com as camadas desejadas.

Por fim, também foi desenvolvido o componente de visualizar as informações dos dados, sendo possível visualizar informações de um dado específico ou de vários dados dentro de uma região. Se a localização selecionada tiver mais de um dado em períodos diferentes, as duas informações serão mostradas ao usuário.

¹ <http://www.terrama2.dpi.inpe.br/> acesso em 10 de Janeiro de 2018.

A Figura 1 mostra o mapa do portal desenvolvido, nela é mostrado a região de São Paulo onde se tem o mapa do Sara Brasil de 1930, as linhas em azul são as vetorizações das ruas deste mapa, estes vetores possuem a informações, como o nome da rua e o intervalo de tempo que ela existiu. Os pontos em verde são os lugares, já integrado ao banco de dados pela equipe do projeto, que existiam em São Paulo nesta época. As ferramentas utilizadas, por enquanto, no projeto base é o Geoserver, Html, Css, JavaScript, npm, JQuery, Bootstrap para o visual de alguns componentes de JavaScript, e Openlayers.

Todos os códigos desenvolvidos no projeto são open source e estão disponíveis no Github.

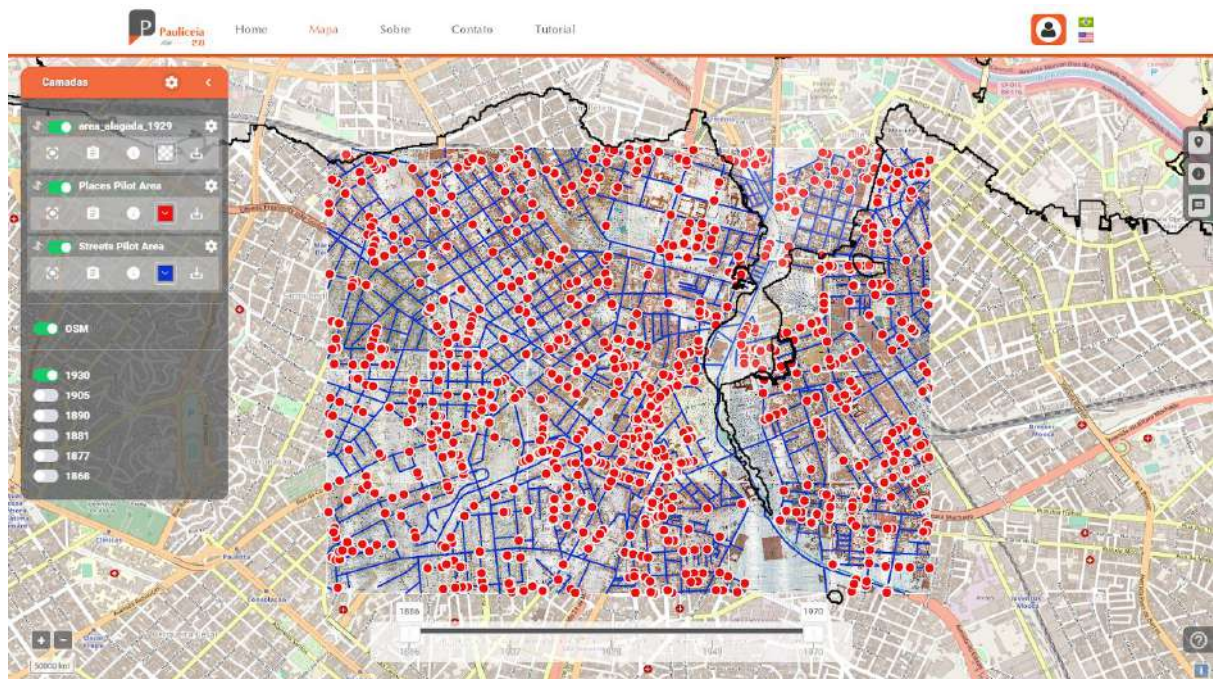


Figura 1: Mapa do projeto, autoria própria (2019).

Os componentes frontend desenvolvidos foram baseados nas funções da api de VGI. Estes possuem a função fundamental de fazer a ligação do usuário com a API desenvolvida, de maneira que fique fácil para um usuário, que não tenha domínio sobre computação, manipular seus dados históricos e trocar informações com outros usuários. A seguir são listados os componentes desenvolvidos:

O componente de notificação possui abas que mostram as notificações globais, as notificações pessoais e as das camadas que o usuário segue, além disso é possível por meio dele, inserir uma notificações global e responder ou denunciar alguma notificação. As notificações que o usuário inseriu podem ser apagadas pelo mesmo ou pelos administradores da página. Esse componente também pode ser visualizado no mapa.

Os componentes de visualização de camadas possui o objetivo de mostrar as camadas que o usuário pode alterar ou excluir a mesma, este último apenas no caso do componente minhas camadas.

O componente de inserir uma nova camada, que possui o objetivo de criar uma nova camada dentro do portal com os dados do usuário e com outras informações, como palavras chaves, colaboradores, descrição e referência. Para criação da camada é necessária algumas etapas, primeiro se cria uma nova camada, segundo se cria um changeset da camada, onde a versão dos dados é controlada, após ocorre a importação dos dados do usuário, e por último é obtido do usuário as informações temporais da layer.

As informações temporais dos dados podem ser obtidas de duas formas, seja o usuário informando a data inicial única e a data final única de todos os dados da camada, ou seja o mesmo especificando qual coluna do seu shapefile possui a data inicial e qual coluna possui a data final dos dados.

Desenvolveu-se ainda, o componente de inserir uma nova palavra-chave e a de visualizar as que o usuário já criou anteriormente, podendo também apagar alguma se desejar.

E por último, o componente de gerenciar as denúncias feitas no portal, seja ela de uma camada ou de um comentário. Apenas um administrador do portal pode acessar essa página, e o mesmo pode visualizar as mensagens da denúncia e as informações da layer ou mensagem denunciada, e caso julgue necessário administrador pode excluir a mesmas.

4.5 Reuniões semanais:

Durante o período do projeto, foram realizadas reuniões semanais com o grupo de pesquisa do INPE em Geoinformática, para se discutir sobre o projeto. Nessas reuniões, eram realizadas diversas atividades: (1) Apresentação das atividades realizadas na semana passada; (2) Explicações sobre dúvidas técnicas do bolsista, por parte dos pesquisadores do INPE; (3) Discussões sobre conceitos importantes relacionado ao projeto; e (4) Definições dos próximos passos a serem realizados na semana.

4.6 Levantamento de requisitos:

Os requisitos necessários para a elaboração do componente espaço-temporal foram definidos nas reuniões semanais e na entrevistas com parceiros do projeto.

4.7 Capacitação realizada:

No período do projeto, os bolsistas realizaram um curso, voltado para Mestrado e Doutorado, de Banco de Dado Geográficos no INPE, administrado por

Gilberto Ribeiro de Queiroz², Karine Reis Ferreira³ e Lúbia Vinhas⁴. Este serviu como base para o desenvolvimento do projeto, pois apresentou ao bolsistas conhecimentos específicos e importantes sobre Dados Geográficos e como se trabalhar com eles. O curso teve duração de três meses e era realizado, às segundas e quartas feiras, com duração de duas horas.

5. Informar e justificar caso tenham ocorrido mudanças e, eventualmente, os ajustes realizados nas atividades de pesquisa do bolsista, em relação ao Plano de Atividades (v. item 8, a seguir)

O plano de atividades do relatório parcial descrevia como continuação do projeto, o desenvolvimentos de novos métodos de visualização espaço-temporal além do slider, no entanto, com a saída de membros do projeto, o bolsista ficou responsável pelo desenvolvimento do componentes frontend de VGI do portal e pela criação das páginas de informações do projeto.

Esta alteração foi necessária pois os componente VGI do portal são extremamente importante para o funcionamento do mesmo, pois eles que permitem a interação dos dados dos usuários com o portal e com os outros usuários.

6. Avaliação do impacto das atividades do bolsista sobre o andamento do projeto

A visualização espaço-temporal para dados históricos é uma parte importante do projeto Pauliceia. Componentes de visualização espaço-temporal permitirá aos usuários do portal do projeto fazer análises exploratórias dos dados históricos do projeto de maneira intuitiva e simples. Isso motivará a utilização do portal por historiadores, pesquisadores e alunos.

Além disso, os componentes de VGI desenvolvido pelo bolsista são de extrema importância para o projeto, pois esses componentes permitem o usuário utilizar as funções da API de VGI, tais como, criar camadas, editar, comentar, entre outras.

8. Bibliografia

José, Juan Franklin Rodriguez Vila. Clusterização e visualização espaço-temporal de dados georreferenciados adaptando o algoritmo marker clusterer: um caso de uso em Curitiba. 2016. 100 p. Dissertação (Mestrado em Computação Aplicada) - Universidade Tecnológica Federal Do Paraná, Curitiba PR, 2016. Disponível em: <http://repositorio.utfpr.edu.br/jspui/bitstream/1/2832/1/CT_DAINF_M_Rodriguez%2

² <http://lattes.cnpq.br/9981634193003068> acesso em 15 de Janeiro de 2018.

³ <http://lattes.cnpq.br/1014575950191035> acesso em 15 de Janeiro de 2018.

⁴ <http://lattes.cnpq.br/6187040703676041> acesso em 15 de Janeiro de 2018.

0Vila%2C%20Juan%20Jos%C3%A9%20Franklin_2016.pdf>. Acesso em: 15 novembro 2017.

Shrestha, Ayush. "Visualizing Spatio-Temporal data." Dissertação - Georgia State University, 2014. Disponível em: <https://scholarworks.gsu.edu/cs_diss/92> Acesso em: 12 Novembro 2017.

OLIVEIRA, Maxwell Guimarães de ; BAPTISTA, Cláudio de Souza . An Approach to Visualization and Clustering-based Analysis on Spatiotemporal Data. 2013. 12 p. Dissertação (Departamento de Sistemas e Computação)- Universidade Federal de Campina Grande (UFCG), Brasil. Disponível em: <<https://seer.ufmg.br/index.php/jidm/article/view/232>>. Acesso em: 09 jan. 2018.

1. Título do projeto

Banco de dados espaço-temporal para o protocolo VGI

2. Nome do bolsista

Yasmin Wassef

3. Informação sobre o nível e período de usufruto da Bolsa.

Bolsa de Iniciação Científica pelo Programa em Inovação Tecnológica eScience

Vigência: 01/06/2017 a 31/05/2018

4. Descrição das atividades do bolsista no projeto de pesquisa

4.1 Introdução

O projeto Pauliceia 2.0 tem como objetivo desenvolver uma plataforma computacional online para gerenciamento colaborativo de dados históricos espaço-temporais, oferecendo um modelo inovador de pesquisa para as Humanidades Digitais utilizando o conceito de VGI (Volunteered Geographic Information).

Originalmente o termo VGI é definido por Goodchild [1] como *“the harnessing of tools to create, assemble, and disseminate geographic data provided voluntarily by individuals”*, ou seja, é o conjunto de ferramentas usadas para criar, reunir e disseminar dados geográficos fornecidos voluntariamente pela comunidade. Como o projeto Pauliceia 2.0 se trata de uma plataforma colaborativa de dados geográficos, ferramentas VGI são imprescindíveis para a coleta e validação dos dados que serão utilizados, visto que eles serão fornecidos pela comunidade.

A abordagem colaborativa da coleta dos dados geográficos, bem como o próprio VGI, apresentam alguns desafios, entre eles a validação, controle de qualidade e uniformização dos dados. Pensando nesses desafios, foram estudadas e desenvolvidas ferramentas e métodos para o projeto que farão a validação dos dados da plataforma através da própria comunidade que utiliza o portal.

4.2 Atividades desenvolvidas

As atividades desenvolvidas pelo bolsista foram:

- Revisão e estudo de literatura sobre VGI, crowdsourcing e mapeamento colaborativo, com o objetivo de obter melhor compreensão sobre os conceitos relacionados ao aspecto colaborativo do projeto, entendendo seus desafios, ferramentas e técnicas. Essa revisão de literatura também visou buscar exemplos sobre como VGI tem sido usado para o desenvolvimento de outros projetos e o que é possível aprender com eles

para se aplicar ao Pauliceia 2.0, analisando as modificações necessárias para que as ferramentas e técnicas utilizadas se adequem ao nosso portal. As principais referências que estão orientando esse trabalho são: [1], [2], [3], [4], [5] e [6].

- Definição de um protocolo VGI para o projeto em conjunto com o mestrando Rodrigo Mariano. Esse protocolo estabelece como as ferramentas e técnicas estudadas anteriormente serão aplicadas ao projeto, e que modificações serão necessárias. A definição desse protocolo determinou técnicas de coleta de dados e metadados, um modelo de dado para o projeto, e possíveis métricas e formas de controle de qualidade para esses dados obtidos pela comunidade. Métodos para validação e coleta uniformizada dos dados também foram definidos, bem como a organização geral do portal e sua interação com usuário. A partir dessa organização geral, alguns conceitos intrínsecos e necessários ao portal foram definidos, como: usuário normal, curador e administrador; feature; camadas (conjunto de features), entre outros.
- Definição da arquitetura do banco de dados do projeto, em conjunto com o mestrando Rodrigo Mariano. A definição dessa arquitetura correlaciona os conceitos definidos no protocolo VGI do projeto de forma a permitir sua modelagem em um banco de dados espaço-temporal para o projeto.

5. Informar e justificar caso tenham ocorrido mudanças e, eventualmente, os ajustes realizados nas atividades de pesquisa do bolsista, em relação ao Plano de Atividades (v. item 8, a seguir).

O Plano de Atividades original descrevia como principal atividade a criação de uma base de dados histórica urbana, utilizando scripts SQL para criar automaticamente os bancos de dados históricos urbanos em um Sistema de Bancos de Dados Espaciais, e preenchê-los com os conjuntos de dados urbanos históricos de 1870 a 1940. Como a coleta desses dados acontecerá de forma colaborativa com a comunidade, inevitavelmente seria necessário um estudo mais aprofundado de VGI e suas ferramentas e técnicas, para que fosse possível modelar o banco e organizá-lo de forma a otimizar o armazenamento dos dados colaborativos.

Conforme os estudos e revisão de literatura avançavam, e as reuniões e discussões delimitavam de forma mais definida as funcionalidades e aspectos do portal, percebeu-se que a modelagem do banco de dados histórico urbano dependia completamente das ferramentas VGI e da forma como os usuários colaborarão com o portal do projeto. Um modelo eficiente e otimizado de banco de dados histórico urbano está diretamente ligado aos tipos de dados que serão armazenados, bem como a forma que serão coletados. Por conseguinte, priorizou-se a definição e o aperfeiçoamento dessas ferramentas VGI, de forma que a criação de uma base de dados histórica urbana adequada ao projeto se tornou uma consequência da definição do protocolo VGI que será utilizado no portal.

Dessa forma, o Plano de Atividades passou a incluir a definição e desenvolvimento do protocolo VGI do projeto, para que possa auxiliar na criação de um banco de dados espaço-temporal otimizado e adequado ao escopo do projeto.

6. Avaliação do impacto das atividades do bolsista sobre o andamento do projeto.

O projeto Pauliceia 2.0 propõe um modelo inovador de pesquisa para Humanidades Digitais, utilizando a própria comunidade como fonte de dados para o portal, criando uma plataforma colaborativa. Para que essa coleta e distribuição de dados fornecidos pela comunidade seja feita de forma otimizada e permitindo a validação e controle de qualidade pela própria comunidade, é de suma importância a estruturação de um banco de dados espaço-temporal otimizado que atendam às necessidades do portal de forma adequada.

A primeira fase do projeto se constituiu principalmente de estudos e definições das funcionalidades que o portal deve ter, e quais ferramentas e técnicas VGI serão necessárias ao desenvolvimento do projeto. A abrangência e diversas possibilidades de abordagem desses conceitos fez com que a etapa de discussões e definições se prolongasse, e por fim resultou na escolha adequada de funcionalidades para o portal, o que traz benefícios e agiliza o andamento do projeto.

A etapa de desenvolvimento dos componentes VGI ainda está em andamento, porém já mostra resultados positivos com a integração da API de VGI desenvolvida pelo aluno mestrando que também está no projeto, o que traz resultados positivos para o andamento do projeto ao iniciar o processo de integração entre as partes do portal.

O desenvolvimento do banco de dados geográficos do projeto, conjuntamente com o mestrando que está no projeto, permitiu que o portal fosse desenvolvido de forma adequada e atendesse aos requisitos estabelecidos na etapa de de discussão e estudo do projeto.

8. Bibliografia

[1] GOODCHILD, M. F. *Citizens as sensors: the world of volunteered geography*. GeoJournal, v. 69, n. 4, p. 211-221, agosto de 2007.

[2] MOONEY, P. et al. *Towards a Protocol for the Collection of VGI Vector Data*. ISPRS International Journal of Geo-Information, v. 5, n. 11, p. 217, novembro de 2016.

[3] SEE, L. et al. *Crowdsourcing, citizen science or volunteered geographic information? The current state of crowdsourced geographic information*. ISPRS International Journal of Geo-Information, v. 5, n. 5, p. 55, abril de 2016.

[4] DAVIS JR, C. A.; VELLOZO, H. S.; PINHEIRO, M. B. *A Framework for Web and Mobile Volunteered Geographic Information Applications*. In: GEOINFO, XIV, 2013, Campos do Jordão, 2013. p. 147-157. Disponível em: <http://www.geoinfo.info/proceedings_geoinfo2013.split/paper16.pdf>. Acesso em: 25 jan. 2018

[5] CURA, R. *Historical collaborative geocoding*. Artigo em andamento submetido a arXiv: 1703.07138, março de 2017. Disponível em: <<https://arxiv.org/abs/1703.07138>> Acesso em: 25 jan. 2018

[6] JAMAL, J. A. et al. *OpenStreetMap in GIScience*. 2015.