

# VGI e Humanidades Digitais

Matheus Monteiro Mariano, Rodrigo Monteiro Mariano

# 1 Resumo

A informação geográfica voluntária (Volunteered geographic information, VGI) é um fenômeno recente que oferece um mecanismo alternativo para a aquisição e compilação de informações geográficas (GOODCHILD, 2007). VGI é uma versão do crowdsourcing, na qual o usuário contribui gerando e manipulando dados geoespaciais em websites, sendo estes dados inseridos em um bancos de dados. Estes dados contribuídos especificam localizações geográficas, com suas descrições e propriedades (GOODCHILD, 2012). Humanidades Digitais (HD) nasceu do encontro entre humanidades tradicionais e métodos computacionais. Com a migração de materiais culturais para ambientes de rede, as questões relativas à produção, disponibilidade, validade e administração destes materiais apresentam novos desafios e oportunidades para os humanistas (BURDICK et al., 2012). HD é um campo que abrange uma ampla gama de métodos e práticas computacionais, como visualizações de grandes conjuntos de imagens, modelagem 3D de artefatos históricos e sua análise, jogos de realidade alternados e entre outros (GOLD, KLEIN, 2017).

Pauliceia é um projeto FAPESP que tem como objetivo desenvolver uma plataforma computacional para criação, organização, armazenamento, processamento e publicação de dados históricos urbanos de maneira colaborativa. Para validar essa plataforma, será implementado um estudo de caso com os dados históricos da cidade de São Paulo no período de 1870 a 1940. No contexto desse projeto, os objetivos desse trabalho são:

1. Proposta computacional de como integrar dados do OpenStreetMap na Plataforma do projeto Pauliceia.
2. Mostrar um exemplo desta arquitetura de integração funcionando.

## 2 Geoprocessamento do mapa de São Paulo

Está sendo feita a digitalização, georreferenciamento e vetorização do mapa da cidade de São Paulo, de 1868, correspondente à área espacial estabelecida para o projeto.

Com base em critérios científicos, definidos pela equipe, um conjunto de mapas, extraídos do período entre 1870 e 1940, serão selecionados para rastrear com maior precisão as transformações da área definida pelo mapa de 1868.

Camadas serão vetorizadas para indicar as mudanças nessa área durante o período de tempo do projeto.

O mapa base é o do Sara Brasil, que foi um mapeamento de São Paulo feito com aerofotogrametria entre os anos de 1928 a 1933 (Gestão Urbana SP, 2017) e será utilizado neste projeto o de 1930.

## 2.1 Diagrama OMT-G

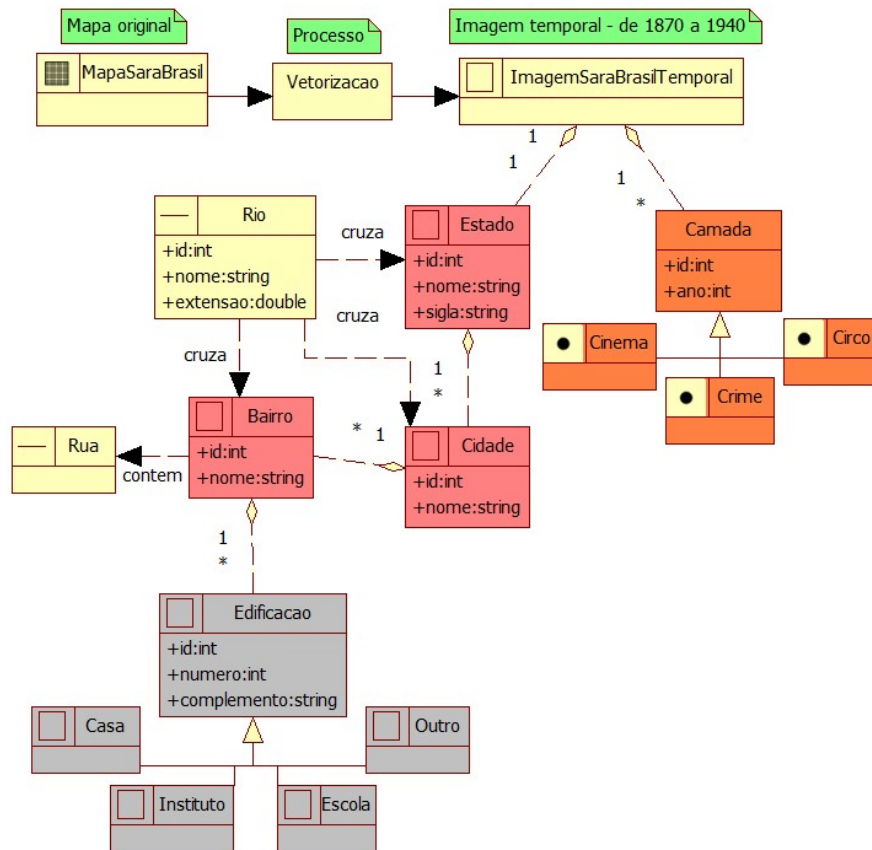


Figura 1: Diagrama OMT-G

A figura 1 mostra o Diagrama OMT-G do projeto. Temos o mapa original do Sara Brasil, que após vetorizado gerou-se um mapa vetorial que é temporal. Nele há georreferenciado os rios, ruas e edifícios. Este mapa terá inúmeras camadas, correspondente a um tipo de informação que os pesquisadores irão inserir na base de dados, como informação de crimes, cinema, circo e etc.

A resolução espacial é dada pela escala 1/1000, a resolução temporal é um mapa histórico por década, no período entre 1870 e 1940, e os formatos das imagens são TIFF e SHP.

### 3 VGI

Volunteered Geographic Information (VGI) é uma abordagem na área geográfica dado por Goodchild (2007) que define uma forma colaborativa entre o uso e a disponibilização das informações geográficas através da Internet de forma aberta. Portanto, qualquer pessoa com acesso à algum hardware que colete dados geográficos (como GPS, por exemplo) e tenha acesso à Internet poderia colaborar com projetos geográficos, disponibilizando seus dados e fazendo uso das informações já disponíveis sem a necessidade de ser um profissional da área, tornando o acesso à informação livre.

Desta forma, a informação se torna mais democrática, pois não dependeria de agências governamentais ou privadas para ter a informação. Isto também torna os sistemas geoespaciais de instituições melhores, pois como as informações são geradas em tempo real daria confiabilidade a um sistema ao testar diversas fontes de dados, cada vez mais aprimorando seus sistemas. Como a informação é disponível de forma pública, os próprios usuários poderiam averiguar aquele dado ou editar o sistema, tornando a plataforma mais democrática. Este conceito é parecido com outras abordagens Open Source, como Wikipédia.

Apesar do VGI ser uma plataforma colaborativa que tenha diversos prós, ainda persistem alguns desafios. Um desses desafios é verificar a precisão dos dados disponibilizados. Por exemplo, Haklay (2010) analisou os dados disponibilizados entre o OpenStreetMap e de agências governamentais britânicas, e percebeu que os dados no OpenStreetMap haviam um deslocamento posicional médio de cerca de 6 metros. Este problema normalmente acontece porque quem disponibiliza não necessariamente tem um conhecimento sobre cartografia ou geoprocessamento, podendo afetar na qualidade da informação. Mesmo assim, isso não invalida o uso do VGI, pois mesmo que os dados possuam deslocamentos, isto pode ser corrigido através de outros dados disponibilizados por outras pessoas ou até atualizado por quem conhece sobre o assunto, entretanto instituições que fazem uso de sensores para a coleta dos dados geralmente possuem um tempo de resposta mais longo para receber os dados, tornando o processo de atualização mais complicado.

Desta forma, o VGI ainda é um conceito relevante para a área geográfica, sendo de grande importância para tornar os dados cada vez mais públicos e de uso cotidiano pela sociedade.

O principal modelo VGI hoje é o OpenStreetMap, uma ferramenta colaborativa de mapeamento geográfico construída a partir da colaboração de diversas pessoas no mundo. Um usuário não precisa ser, necessariamente, um cartógrafo para gerar os dados e disponibilizá-lo no portal, tornando a ferramenta de domínio público para uso.

### 4 Humanidades Digitais

Humanidades Digitais é uma área de pesquisa e ensinamento que visa a intersecção das tecnologias digitais com as disciplinas da humanidade (Burdick, A.

et al; 2012), criando uma abordagem multidisciplinar que une a tecnologia da computação com as diversas disciplinas baseadas na cultura humana, como artes, literatura, música, dança, filosofia, entre outros. É uma metodologia que está preocupado em criar mecanismos de análise e síntese que apresente esta informação da humanidade de uma forma digital.

O termo "humanidade" foi primeiro usado na época da Renascença em escolas italianas que trabalhavam com em estudos da antiguidade clássica. De acordo com Johanna Drucker (et al, 2014), o termo "humanidade" enfatiza a mudança da visão centrada do homem medieval para "o homem é uma medida de todas as coisas". Desta forma, humanidade se refere a ensinamentos voltados ao homem pela cultura, como a música, artes, literatura, dança, que enfatizam a maneira como o homem vê o mundo. Humanidade também pode ser traduzido como Humanas, que são disciplinas focadas no homem e suas ações.

Com isso, pode-se deduzir que Humanidades Digitais são comportamentos da vida humana através da sua cultura em sociedade que são alteradas pelo ambiente digital do século XXI. Em outras palavras, é o reposicionamento ideológico da sociedade através da tecnologia computacional (Inacio, C.S. et al; 2016). Dentro do contexto geoespacial, é o modo como as pessoas interagem com as informações geográficas, em que a todo momento geram dados sobre localização ou aspectos do meio ambiente no geral, e acabam disponibilizando para a comunidade ou ferramentas privadas que possam tirar proveito desta informação. No caso deste trabalho, a criação de um mapa histórico de São Paulo é uma forma de Humanidade Digital, pois traz para a população informações digitalizadas sobre seu passado que podem ser utilizadas a uso próprio, de forma livre e democrática. Além disto, também entra no contexto multidisciplinar, pois este projeto é uma colaboração entre pesquisadores de História interessantes em restaurar o passado da cidade, com profissionais da Tecnologia da Informação.

Então, o projeto Pauliceia é um exemplo de como a Humanidade Digital pode afetar a sociedade positivamente, realizando trabalho colaborativo entre as frentes humanas e exatas.

## 5 O Projeto Pauliceia

As humanidades digitais estão se tornando uma realidade crescente e inevitável para quem trabalha com conhecimento histórico. Eles ganharam suporte intelectual e institucional entre os estudiosos em todas as disciplinas, que passaram a argumentar por seu potencial para discutir suas práticas e implicações. Tais discussões incluem a questão de definir uma identidade epistemológica para o campo, juntamente com as implicações teóricas e práticas para a erudição e dentro das estruturas institucionais que avaliam o trabalho de um erudito.

Este projeto visa projetar e construir uma plataforma computacional para pesquisa histórica colaborativa. O principal objetivo é desenvolver ferramentas de software, como um portal web, que permitem aos pesquisadores de humanidades criar, organizar, armazenar, integrar, processar e publicar conjuntos de

dados da história urbana.

Os conjuntos de dados urbanos históricos representam a evolução dos locais espaciais urbanos ao longo do tempo e são um conjuntos de dados espaço-temporais.

A plataforma proporcionará acesso a este banco de dados e permitirá a interação entre pesquisadores, que poderão contribuir para os eventos do banco de dados que podem ser representados espacialmente e temporalmente. Os estudiosos poderão produzir mapas e visualizações de suas próprias pesquisas, e ao mesmo tempo, contribuir para os dados dentro do sistema.

Este projeto irá enriquecer a compreensão da história de São Paulo durante o período de 1870 a 1940, além de oferecer um modelo inovador de pesquisa para as Humanidades Digitais que promova o trabalho colaborativo e o fluxo de conhecimento gratuito.

## 6 GeoServer

GeoServer é um servidor multiplataforma de mapas que tem como objetivo fazer a informação geográfica o mais aberto possível para a sociedade. Para isto, o GeoServer que segue padrões abertos da comunidade geoespacial como o Consórcio Geoespacial Aberto (OGC), Web Map Service (WMS), Web Coverage Service (WCS) e Web Feature Service (WFS-T). Tem como foco facilitar o uso e suporte aos padrões abertos, tornando possível qualquer pessoa, seja ela da área geoespacial ou não, compartilhar informações de forma interoperável, garantindo consistência nos dados.

Desenvolvido em Java, permite a integração com diversos Banco de Dados Geográficos como o Oracle Spatial e PostGIS, além dos relacionais tradicionais como SQL Server, MySQL, DB2 entre outros, para os formatos vetoriais. Já para o formato raster, possui suporte para diversos GeoTIFF, ArcGrid, Gtopo30, ImageMosaic, MrSID, ECW e WorldImage. Possui licença GNU/GPL, desta forma a distribuição do software é livre.

## 7 OpenStreetMap

OpenStreetMap (OSM) é um projeto de banco de dados de mapa colaborativo que visa a difusão de geodados por seus usuários (Neis, P. et al; 2011), transmitindo informações geoespaciais através um portal web que pode ser acessado por qualquer pessoa. O projeto é uma das formas de VGI mais conhecidas, tendo conseguido uma comunidade de voluntários de cerca de 1 milhão de usuários, que a mantém constantemente atualizada.

OpenStreetMap foi idealizado por Steve Coast, tendo sido revelado na 40<sup>a</sup> Escola de Verão Anual de Cambridge, em 2005. Por ser gerenciado pela própria comunidade, OSM muitas vezes é comparado com o Wikipédia.

Os usuários que colaboram com o OSM podem inserir os dados espaciais de diversas fontes e formatos, como de GPS, fotografias aéreas ou de conheci-

mentos da região local (Mooney, P., Corcoran, P.; 2012). No caso do projeto Pauliceia, os dados do OSM serão integrados aos dados do portal do projeto para visualização histórica da cidade de São Paulo.

## 7.1 Upload de dados no OpenStreetMap

Há basicamente duas maneiras de se fazer o upload de dados no OSM, manualmente no Portal do OSM e o envio de um grande volume de dados.

O envio manual dos dados é feito no site deles (<https://www.openstreetmap.org>). Após acessar com suas credenciais, clica-se no botão "Edit" do menu superior, depois em "Edit with iD (in-browser editor)", aparecerá uma tela de edição. Haverá opções de "Point", "Line" e "Area", escolhendo uma delas, pode-se adicionar manualmente algum objeto no mapa.

No caso do projeto Pauliceia, temos mais de 40 Gb de dados, que estão sendo produzidos pela equipe de São Paulo, que está fazendo a vetorização dos mapas antigos. Seria inviável adicionar manualmente no OSM todos esses dados, por conta disso, gostaríamos de fazer o upload deles no OpenStreetMap.

Para fazer o upload de uma grande quantidade de dados é necessário seguir umas diretrizes de importação. Essas diretrizes, juntamente com o código de conduta das Edições Automatizadas (OpenStreetMap, 2017), devem ser seguidas ao importar dados no banco de dados do OpenStreetMap. As importações devem ser planejadas e executadas com mais cuidado e sensibilidade do que outras edições, uma vez que as importações pobres podem ter impactos significativos tanto na comunidade de cartografia de dados quanto na comunidade local. O Grupo de Trabalho de Dados é encarregado pelo OpenStreetMap Foundation (OSMF) de detectar e parar as importações que não cumprem as diretrizes. Portanto, não seguir estas diretrizes pode colocar sua conta em risco de ser bloqueada.

As importações não devem ser vistas como uma alternativa para construir a comunidade de mapeamento, executar partes de mapeamento e gerar publicidade para se envolver com mais contribuidores, contudo isso está aberto a discussão, como na lista de envio de importações e na página de discussão (OpenStreetMap, 2017).

## 7.2 Diretrizes de importação

Algumas das diretrizes, que necessitam ser seguidas, são as seguintes:

- Ter familiaridade com os conceitos básicos do OpenStreetMap. Analise o que pode dar errado nas importações;
- Identificar os dados que deseja importar. Isso pode ser linhas centrais de rua, construção de contornos, canais, endereços, etc. e os requisitos da licença de dados;
- Entrar em contato com a comunidade do OSM para ver se há interesse em importar os dados:



- Existem áreas geográficas de interesse, com diferentes níveis de aceitação para as importações. O mesmo conjunto de dados pode ser aceito em uma área e ser rejeitado em outra;
  - Enviar um e-mail para a comunidade OSM para notificá-los dos seus planos, incluindo um link para sua página wiki. Pode-se fazer isso com um e-mail para [imports@openstreetmap.org](mailto:imports@openstreetmap.org), [talk-\(seu país\)@openstreetmap.org](mailto:talk-(seu país)@openstreetmap.org) e o grupo OSM específico para a área pela importação. Precisa fazer parte das listas em <http://lists.openstreetmap.org/> antes de se enviar um e-mail;
  - Importações mais complexas e em larga escala devem ser revistas com a ajuda de voluntários mais tecnicamente orientados e experientes do OSM;
  - Não se pode importar dados sem o acordo da comunidade.
- Deve-se ter permissões e licenças adequadas, do proprietário dos dados, para usar os dados no OSM. Basicamente poder compartilhar, criar e modificar os dados, os dados se tornam públicos no OSM;
  - Criar uma página do wiki do OSM que descreva os detalhes do projeto e dos dados;
  - Siga o planejamento, fornecendo atualizações para a comunidade. Avisar na lista quando tiver terminado (OpenStreetMap, 2017).

Atualmente estamos esperando a comunidade do OpenStreetMap adicionar-nos nas listas de discussões, para podermos mandar um e-mail explicando sobre os dados do projeto Pauliceia.

## 8 Resultados

Como estamos esperando a comunidade do OpenStreetMap adicionar-nos nas listas de discussões, foi-se criado um protótipo do Portal do projeto Pauliceia, em que os dados são salvos em um computador local.

Neste Portal é possível criar tipos de mapas (de estradas, crimes, etc), manipular mapas existentes e publicar o mapa em um Geoserver local, salvando os registros no PostGIS.

Tem-se a possibilidade de criar um grupo de trabalho para os estudiosos, que utilizarão o portal do projeto.

## 9 Conclusão

Através do VGI, criou-se um protótipo do projeto com a possibilidade de utilizar as informações geográficas do projeto de maneira colaborativa pela internet. Esse protótipo utilizou das Humanidades Digitais para criar uma interseção

entre as áreas de tecnologias digitais e de história, para ser utilizado pelos profissionais de História em suas pesquisas.

O mapa base do protótipo vem do OpenStreetMap podendo-se inserir dados espaciais que serão salvos no PostGIS, enquanto que sua publicação é dada no GeoServer, por enquanto ambos em máquinas locais.

O propósito original é utilizar o OpenStreetMap para a persistência dos dados, porém como há um grande volume desses dados, será necessário esperar a adição da equipe do projeto na lista de importação para poder entrar em contato com a comunidade do OSM, o que a equipe do projeto está esperando até o momento.

Para trabalhos futuros, espera-se que:

- possamos utilizar o OpenStreetMap como repositório para os dados históricos. Caso não seja possível após conversarmos com a comunidade criaremos um servidor com um Geoserver oficial;
- criar uma nova versão do protótipo, com novos recursos, como a categorização dos dados inseridos.

## 10 Referências

- BURDICK, A., DRUCKER, J., LUNENFELD, P., PRESNER, T., SCHNAPP, J. DIGITAL\_HUMANITIES. Massachusetts Institute of Technology, 2012.
- Gestão Urbana SP. Prefeitura disponibiliza mapa histórico de 1930 no Geosampa. Disponível em <http://gestaourbana.prefeitura.sp.gov.br/noticias/prefeitura-disponibiliza-mapa-historico-de-1930-no-geosampa/> Acesso em: 11/03/2017.
- GOODCHILD, M. F. Citizens as sensors: the world of volunteered geography. *GeoJournal*, Vol. 69, p. 211–221, 2007.
- GOODCHILD, M. F., LI, L. Assuring the quality of volunteered geographic information. *Elsevier - Spatial Statistics*, Vol. 1, p. 110–120, 2012.
- GOLD, M. K., KLEIN, L. F. Digital Humanities: The Expanded Field. Disponível em <http://dhdebates.gc.cuny.edu/debates/2> Acesso em: 25/03/2017.
- HAKLAY, M., 2010. How Good is volunteered geographical information? a comparative study of OpenStreetMap and ordnance survey datasets. *Environment and Planning B: Planning and Design* 37 (4), 682–703
- MOONEY, P., CORCORAN P. Characteristics of Heavily Edited Objects in OpenStreetMap. 2012. Disponível em <http://www.mdpi.com/1999-5903/4/1/285/htm>
- NEIS, P., ZIELSTRA, D., ZIPF, A. The Street Network Evolution of Crowd-sourced Maps: OpenStreetMap in Germany 2007–2011. 2011. Disponível em: <http://www.mdpi.com/1999-5903/4/1/1/htm>
- OpenStreetMap. Import/Guidelines. Disponível em <http://wiki.openstreetmap.org/wiki/Import/Guidelines> Acesso em: 03/06/2017.
- OpenStreetMap. Automated Edits code of conduct. Disponível em [http://wiki.openstreetmap.org/wiki/Automated\\_Edits\\_code\\_of\\_conduct](http://wiki.openstreetmap.org/wiki/Automated_Edits_code_of_conduct) Acesso em: 03/06/2017.
- Pauliceia 2.0: A Spatiotemporal Platform for Digital Humanities. Disponível em [http://www.dpi.inpe.br/pauliceia/lib/exe/fetch.php?media=wiki:final\\_project.pdf](http://www.dpi.inpe.br/pauliceia/lib/exe/fetch.php?media=wiki:final_project.pdf) Acesso em: 25/03/2017.