City Data Science and Technology

Prof. Fabio Kon
Department of Computer Science
IME-USP, Brazil

INCT Future Internet for Smart Cities
interscity.org

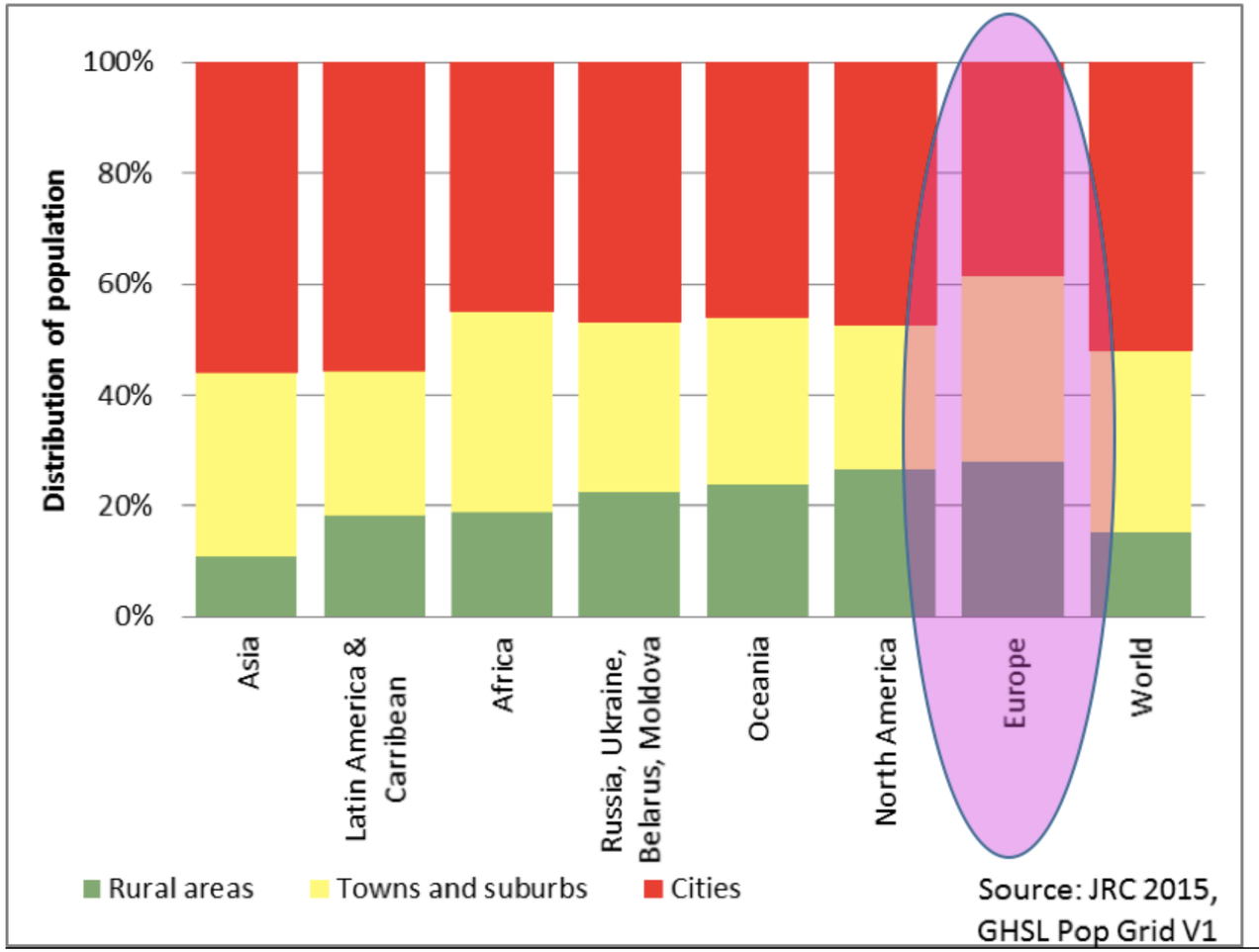
We live in cities

- old methodology: ~55% of the people (UN)

- **newer methodologies: >80% (European Commission)**
  - Satellite images
  - Image Processing / Machine Learning
  - Population databases

Unveiling new perspectives of urbanization

applications



if the **density of population** it is used as objective criteria for distinction of urban vs. rural areas, Europe is the region showing the *highest share of rural population in 2015*
(and the World it is > 80% urban)

The State of European Cities 2016, European Commission, DG REGIO, 2016

GEO GROUP ON EARTH OBSERVATIONS

- Source: European Commission Joint Research Center

Urban public policies

- Most of the times:
  - Designed with no scientific basis at all

- A few times:
  - Based on methods and technologies from the XX century

Evidence-based public policies

1. Create and collect rigorous scientific evidence of what works, including costs and benefits

2. Monitor the execution of programs and measure the impact

3. Use scientific evidences to improve the programs, increase their scale and cancel the programs that don’t work.

4. Promote innovation and test new approaches.

INCT InterSCity Collaborations

- 35 CS professors +
  - Architects, Urban Planners, Economists, Health Professionals, Transportation Engineers
- City governments (unfortunately, very weak collaboration)

- FAPESP PhD and Post-doc fellowships available

InterSCity lab in Brazil

~60 people working:

- USP, PUC-Rio, Scipopulis, UFABC, UFG, UFMA, UFMS, UFRJ, Unicamp, FGV, Unifesp
- FAPESP / CNPq / CAPES (2017 to 2023)
- **InterSCity.org**
- Open Source software
- Open Datasets

Our view

Smart City =
"a city in which its social, business, and technological aspects are supported by ICT to improve the quality of life of its citizens in an integrated, affordable, and sustainable way."

we're interested in developing a
Software platform for Smart Cities

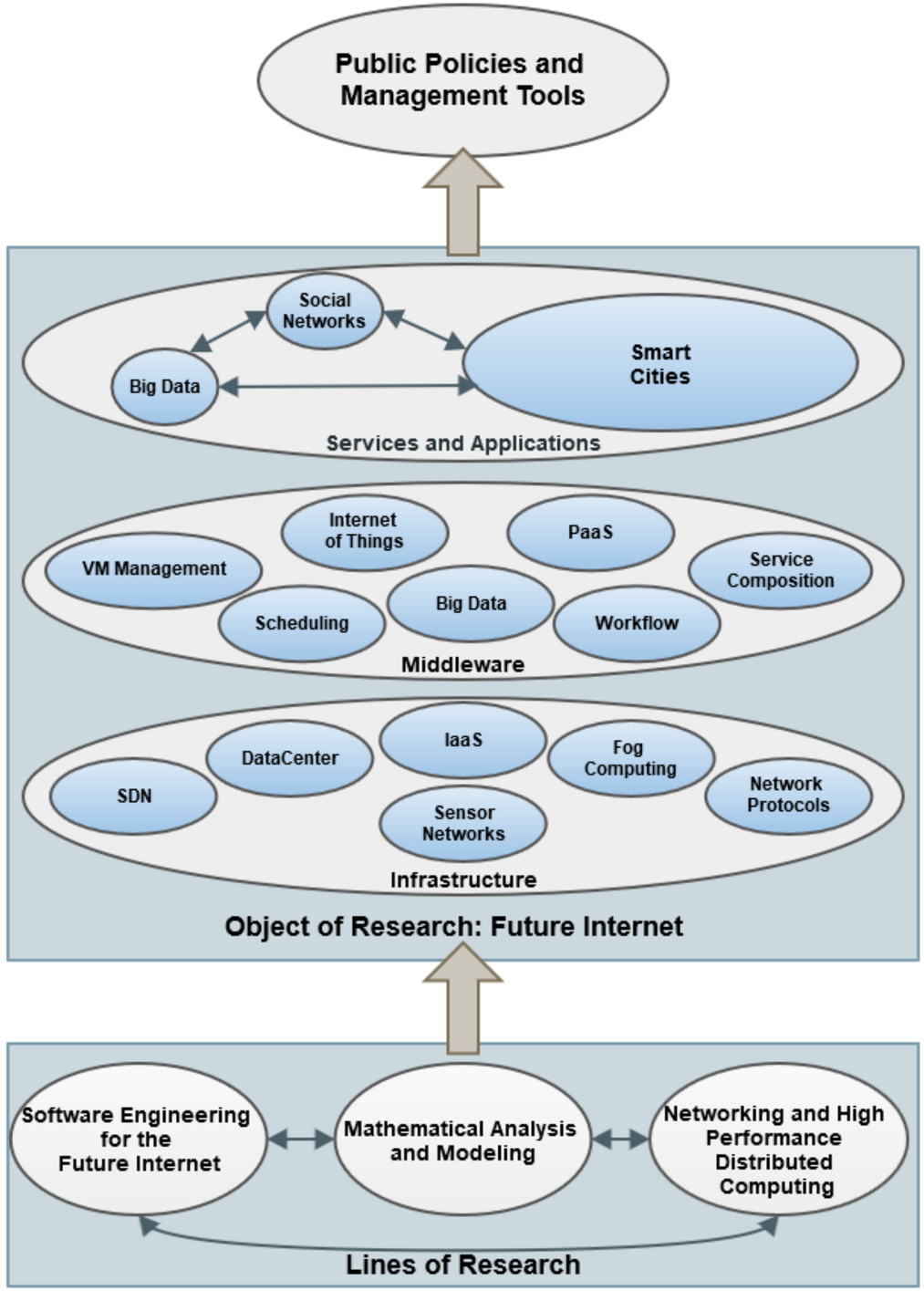
Our view on Smart Cities

Although we don’t ignore high-tech solutions for the elite, we prefer to focus on:

- people (technology is a means not an end)
- low-income populations
- developing countries
- underprivileged neighborhoods

The InterSCity Project

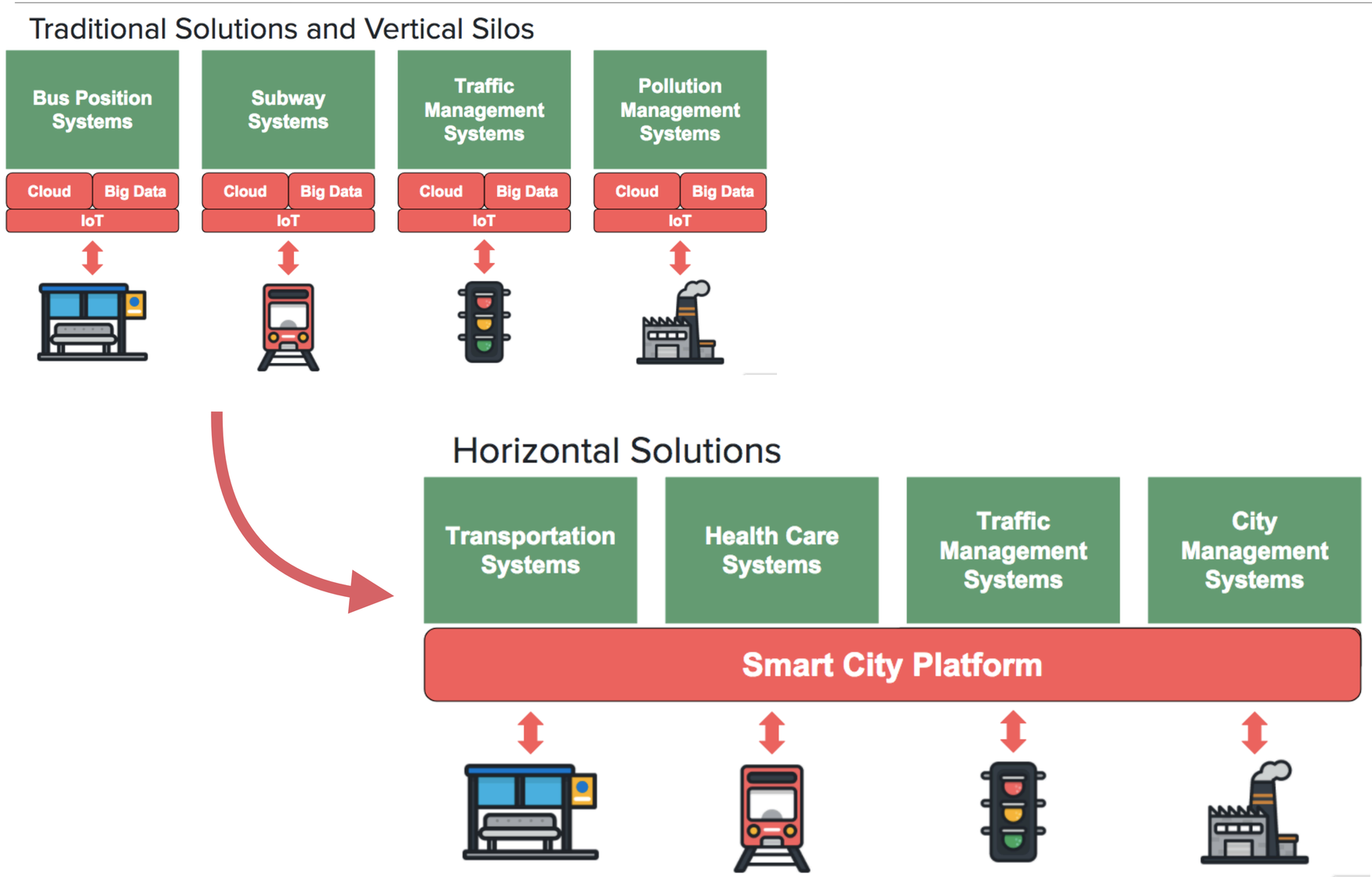
- 3 lines of research
- 3 levels



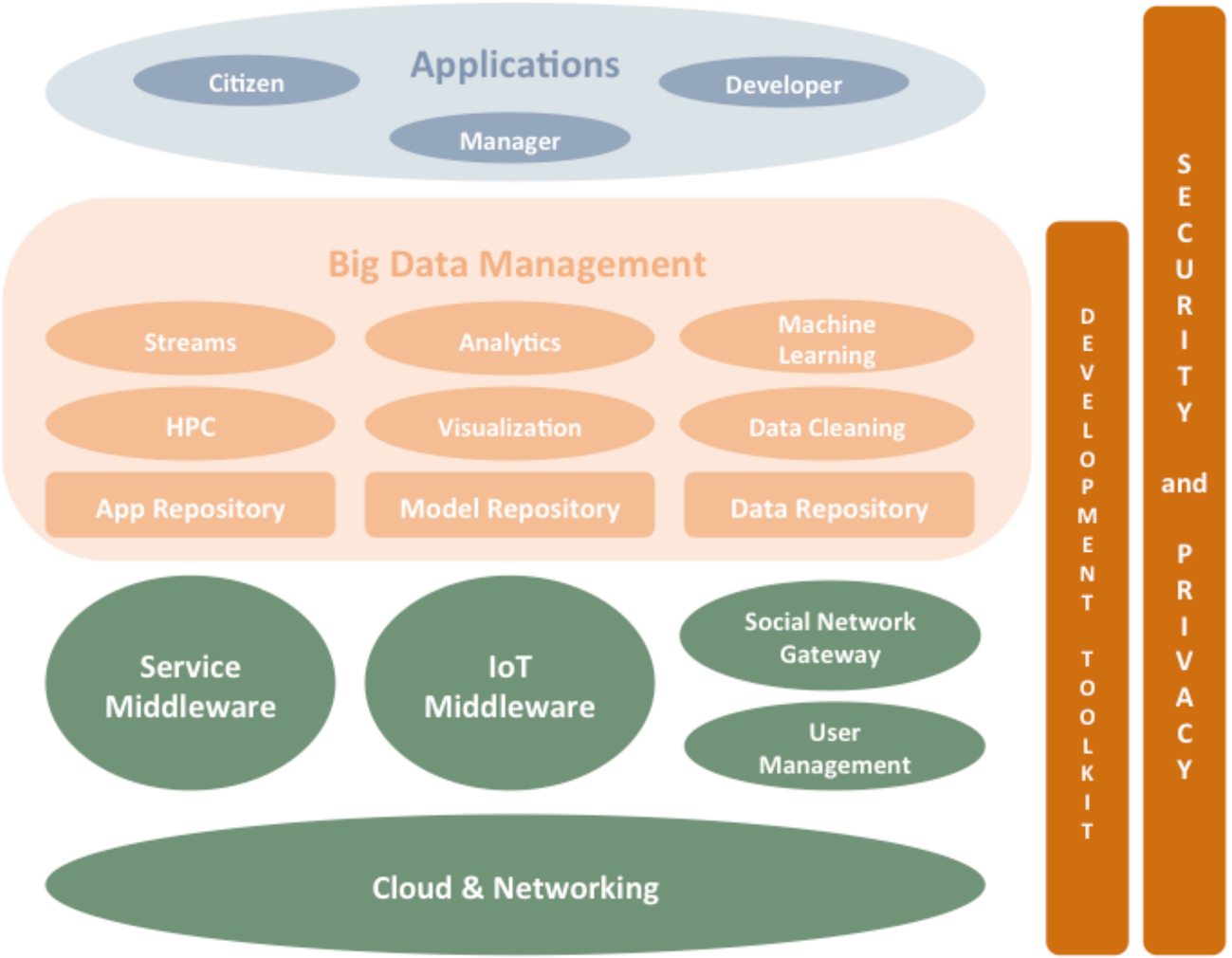
Projects

1. Smart City Software Platform
2. City Simulator
3. Health Dashboard
4. Accessibility Ranking
5. Scipopulis Startup
6. BikeSCience @ MIT Senseable City Lab

1 - A generic Software Platform for Smart Cities



Survey and proposed reference architecture for Smart City Software Platforms

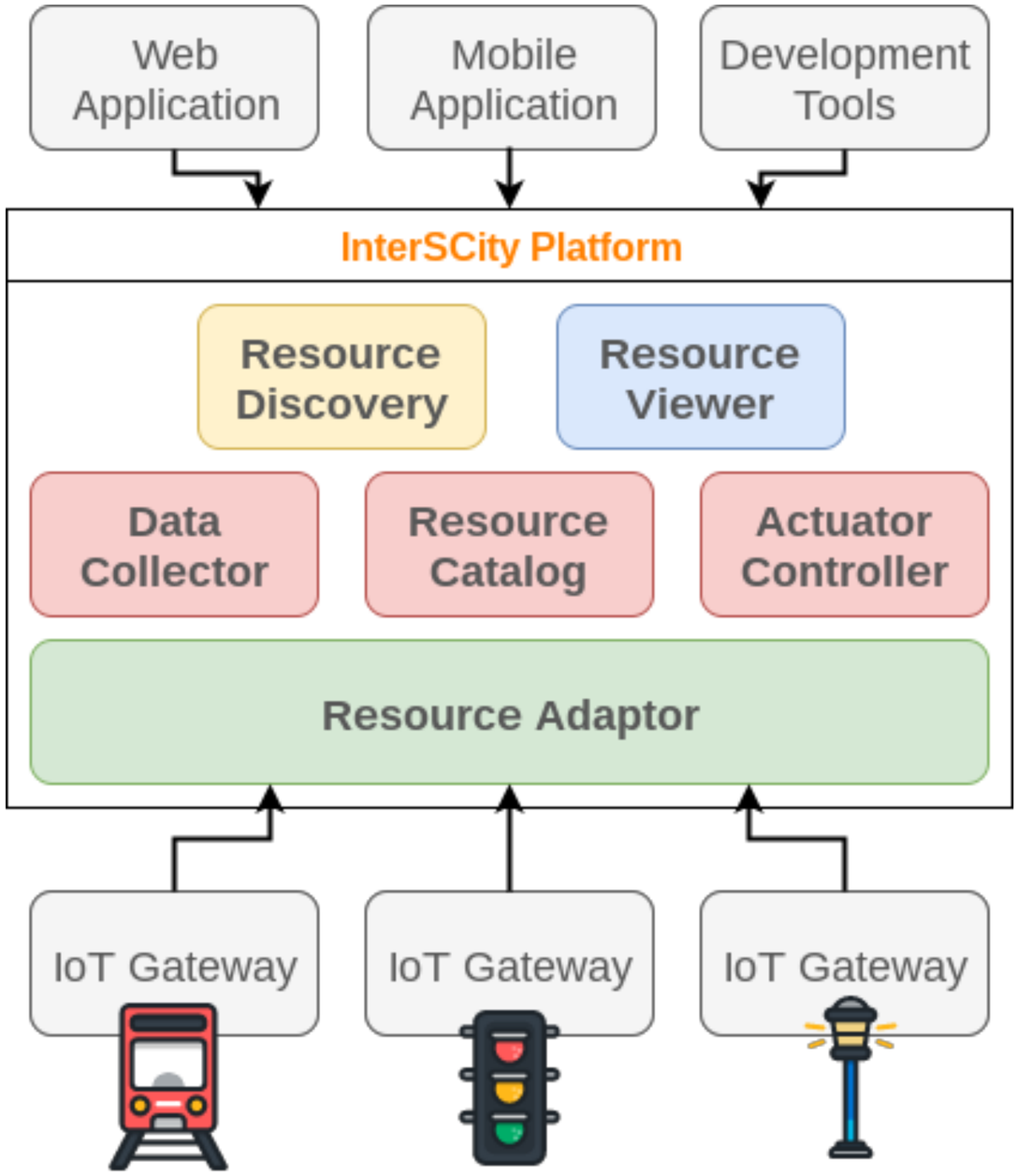


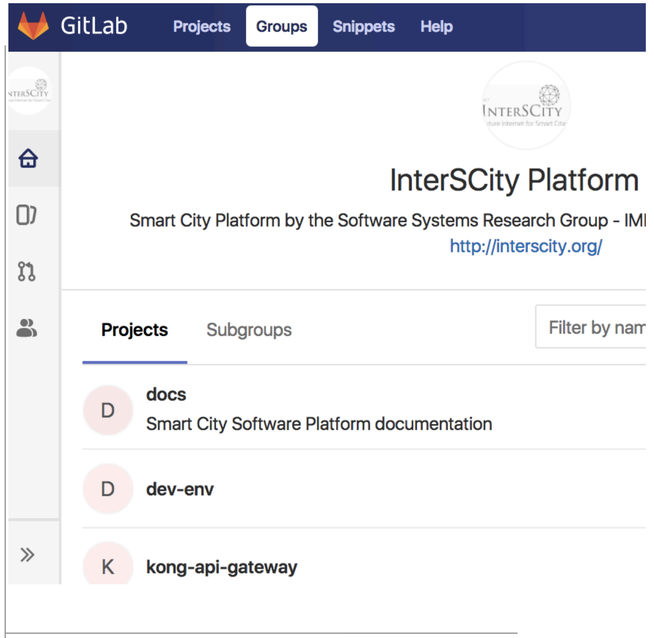
ACM Computing Surveys

Software Platforms for Smart Cities: Concepts, Requirements, Challenges, and a Unified Reference Architecture

Eduardo Felipe Zambom Santana, University of São Paulo
Ana Paula Chaves, Federal Technological University of Paraná
Marco Aurelio Gerosa, University of São Paulo
Fabio Kon, University of São Paulo
Dejan S. Milojicic, Hewlett Packard Labs Palo Alto

Making cities smarter help improve city services and increase citizens' quality of life. Information and communication technologies (ICT) are fundamental for progressing towards smarter city environments. Smart City software platforms potentially support the development and integration of Smart City applications. However, the ICT community must overcome current significant technological and scientific challenges before these platforms can be widely used. This paper surveys the state-of-the-art in software platforms for Smart Cities. We analyzed 23 projects with respect to the most used enabling technologies, as well as functional and non-functional requirements, classifying them into four categories: Cyber-Physical Systems, Internet of Things, Big Data, and Cloud Computing. Based on these results, we derived a reference architec-





InterSCity: A Scalable Microservice-based Open Source Platform for Smart Cities

Arthur de M. Del Esposte[1], Fabio Kon[1], Fabio M. Costa[2] and Nelson Lago[1]
[1]*Department of Computer Science, University of São Paulo, R. do Matão, 1010 - Cidade Universitária, 05508-090, São Paulo, São Paulo, Brazil*
[2]*Institute of Informatics, Federal University of Goiás, Alameda Palmeiras, Quadra D, Câmpus Samambaia, 74690-900, Goiânia, Goiás, Brazil*
{esposte, kon, lago}@ime.usp.br, fmc@inf.ufg.br

eywords: Smart Cities, Software Platform, Microservices, Scalability, Open Source Software

bstract: Smart City technologies emerge as a potential solution to tackle common problems in large urban centers by using city resources efficiently and providing quality services for citizens. Despite the various advances in middleware technologies to support future smart cities, there are no universally accepted platforms yet. Most of the existing solutions do not provide the required flexibility to be shared across cities. Moreover, the extensive use and development of non-open-source software leads to interoperability issues and limits the collaboration among R&D groups. In this paper, we explore the use of a microservices architecture to address key practical challenges in smart city platforms. We present InterSCity, a microservice-based open source smart city platform that aims at supporting collaborative, novel smart city research, development, and deployment initiatives. We discuss how the microservice approach enables a flexible, extensible, and loosely coupled architecture and present experimental results demonstrating the scalability of the proposed platform.

INTRODUCTION

he rapid growth of cities around the world has cre-
ed large, densely populated urban centers charac-
rized by complex interconnected structural, social
d economic organizations. This urbanization phe-

et al., 2014). The Internet of Things (IoT), Big Data, and Cloud Computing are key enabling technologies of smart cities that offer a wide range of opportunities and challenges, both in the academy and industry. To fully exploit the potential of these enablers, future smart cities will demand a unified ICT infrastructure

Underlying technologies (open source tools)

- Relational Database

- NoSQL Database

- Asynchronous messaging

- Dynamic Programming Language

- Web Framework

Underlying technologies (open source tools)

- Container

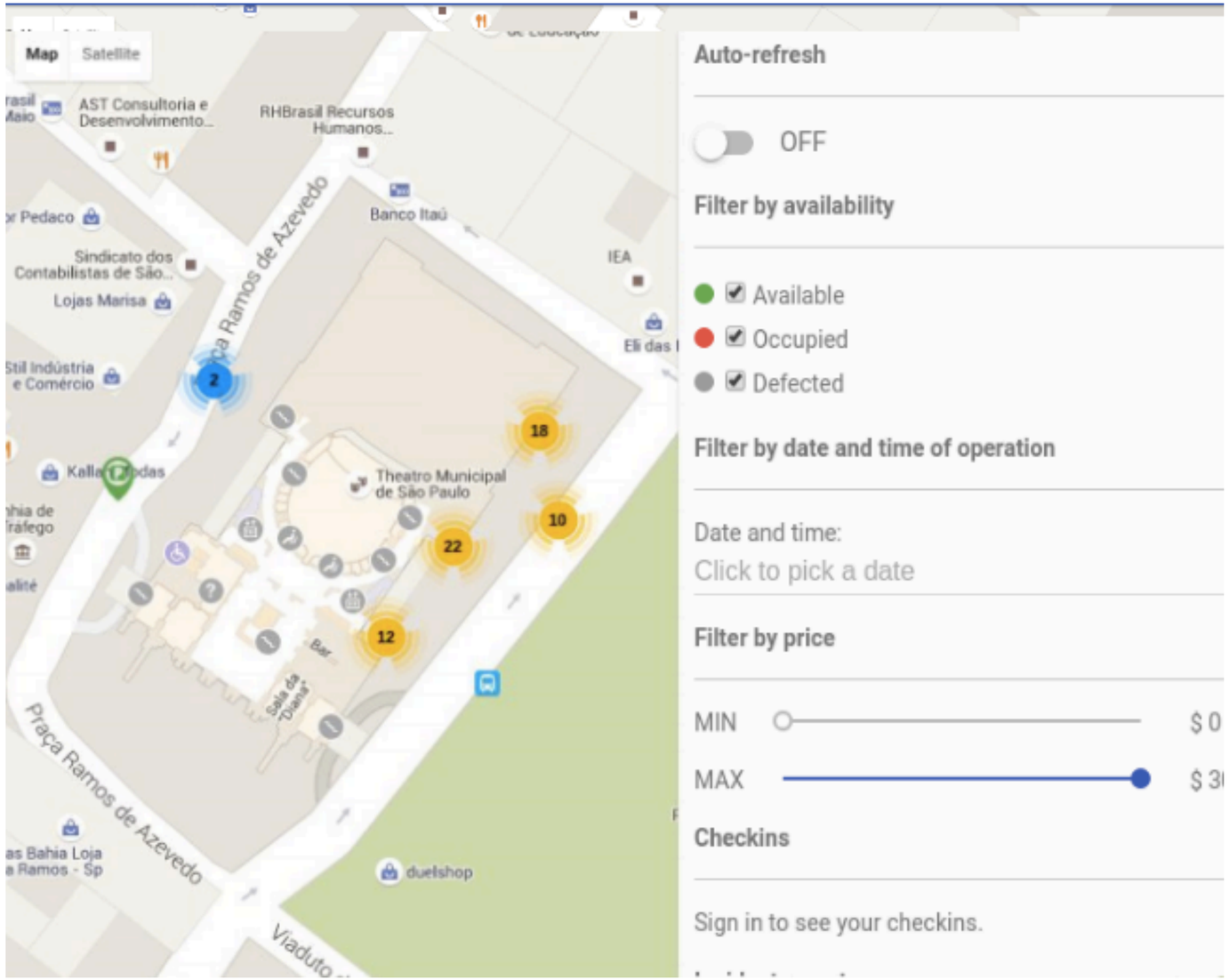
- Microservice API Gateway

Exemplos de uso da plataforma

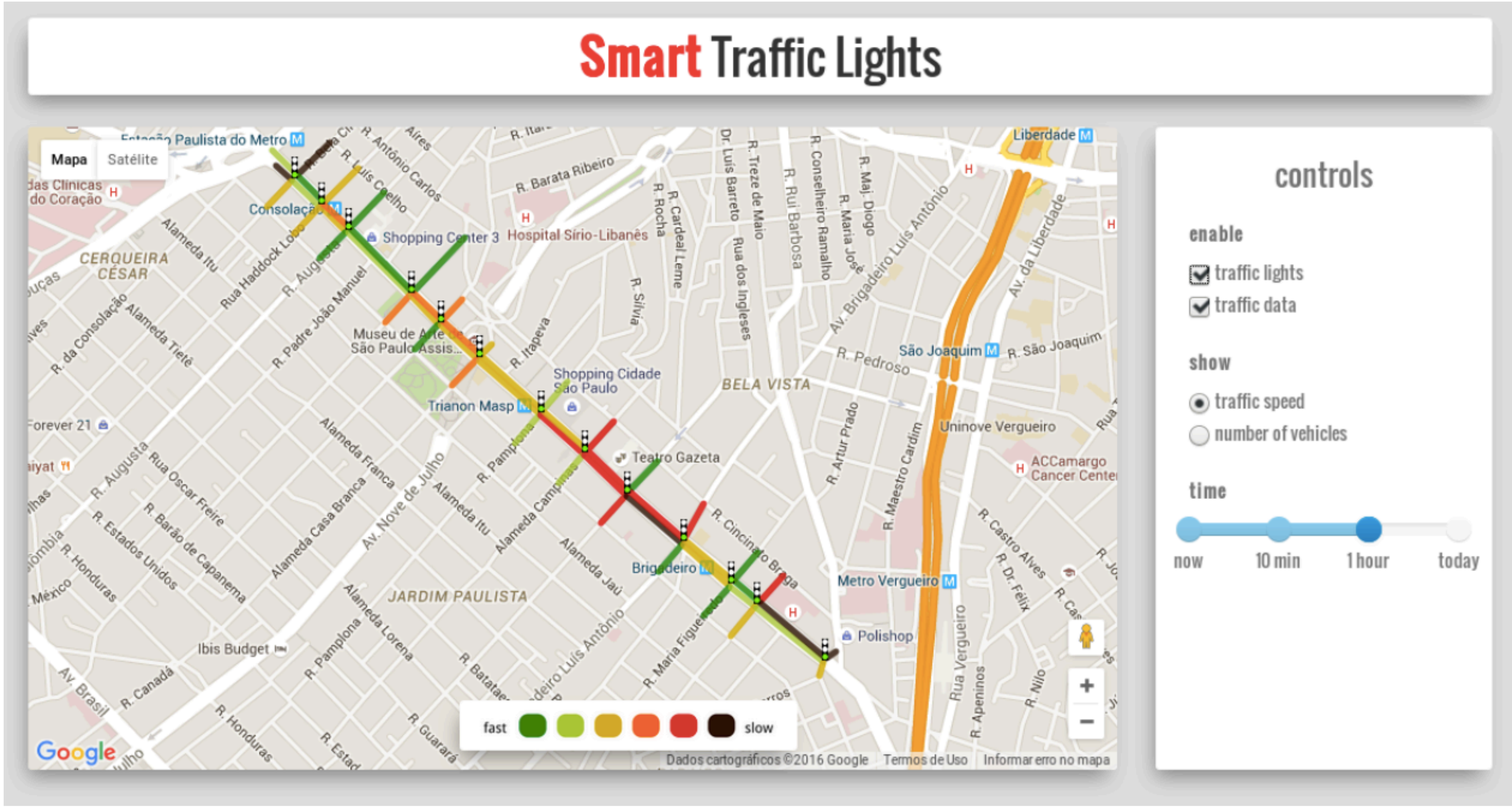
Smart Parking APP

single-page app that helps users in the hard task of finding available parking spots around the city.

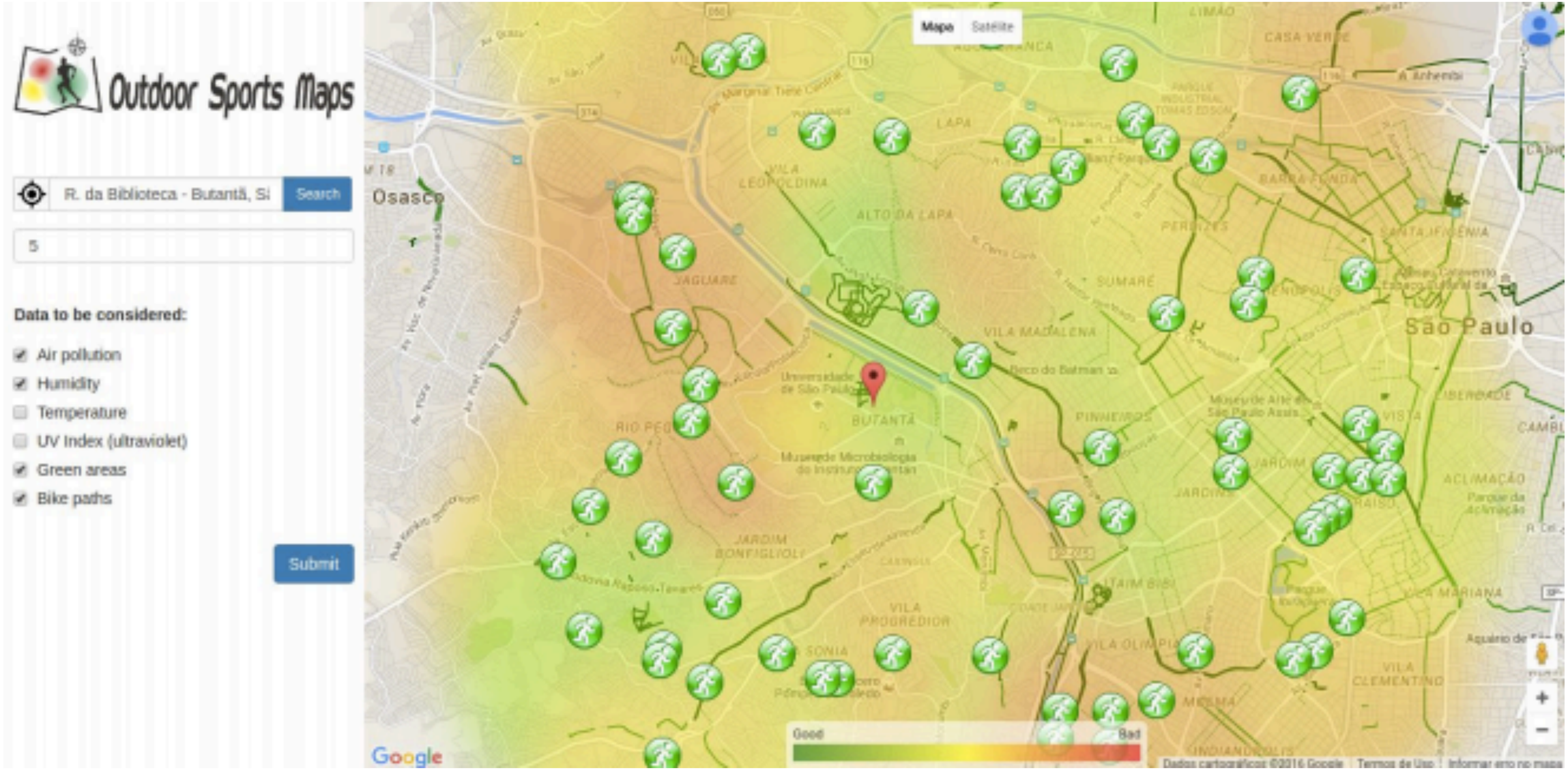
It used the platform services to access simulated data.



Exemplos de uso da plataforma



Outdoor Sports Map



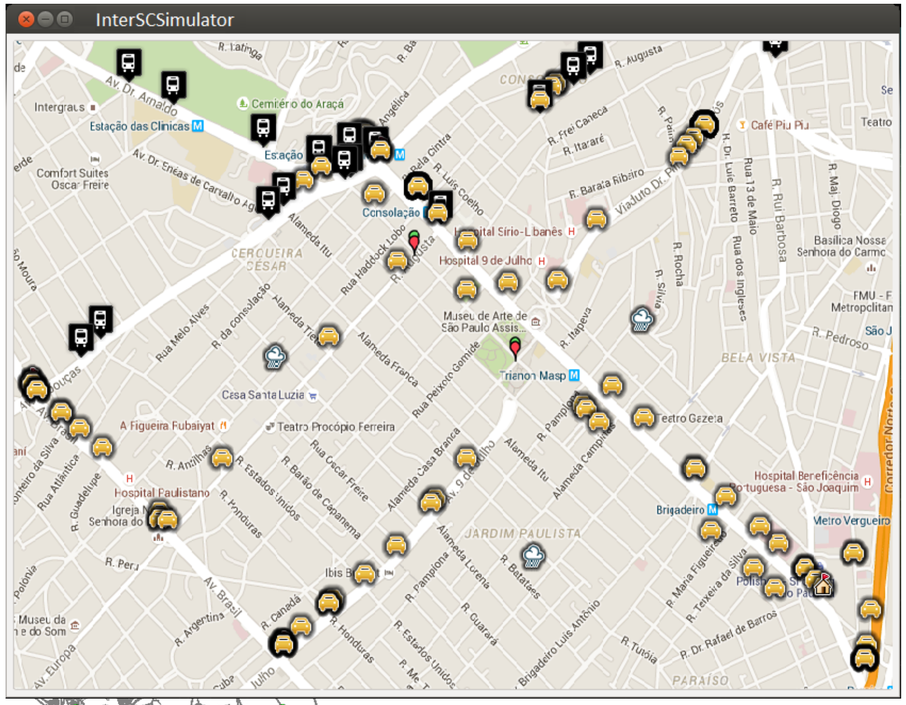
Other uses

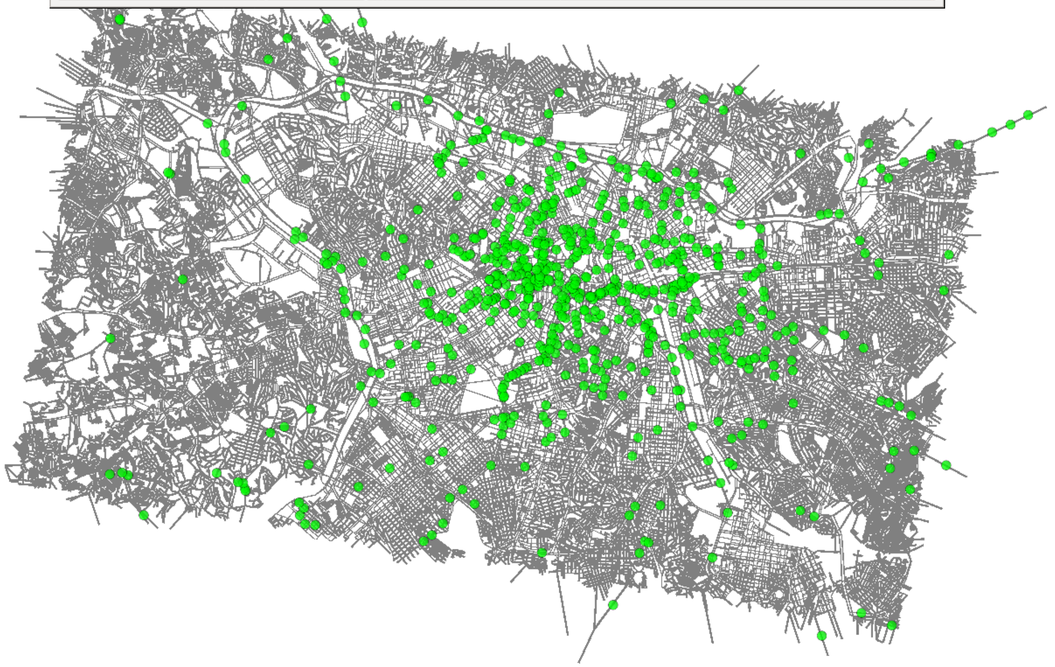
- Smart City Hackathons
- Graduate Course at USP
- Undergraduate Course at UFMA
- Research (MSc, PhD, and post-doc @ USP/Puc-Rio/UFMA)

- In the future: real cities?

2 - InterSCimulator

- Erlang-based large-scale simulator for Smart Cities
- Simulations with 17 million agents in super-real-time
- Multimodal transportation
  - cars, pedestrians, buses, subway, (bicycles).
  - Impact analysis of changes in the transportation infrastructure and associated costs.
  - Population from Paraisópolis favela (slum) in SP.





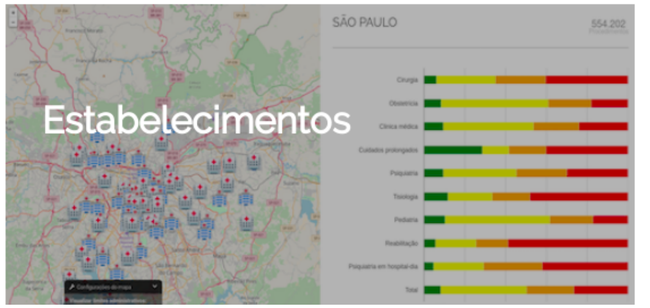
3 -

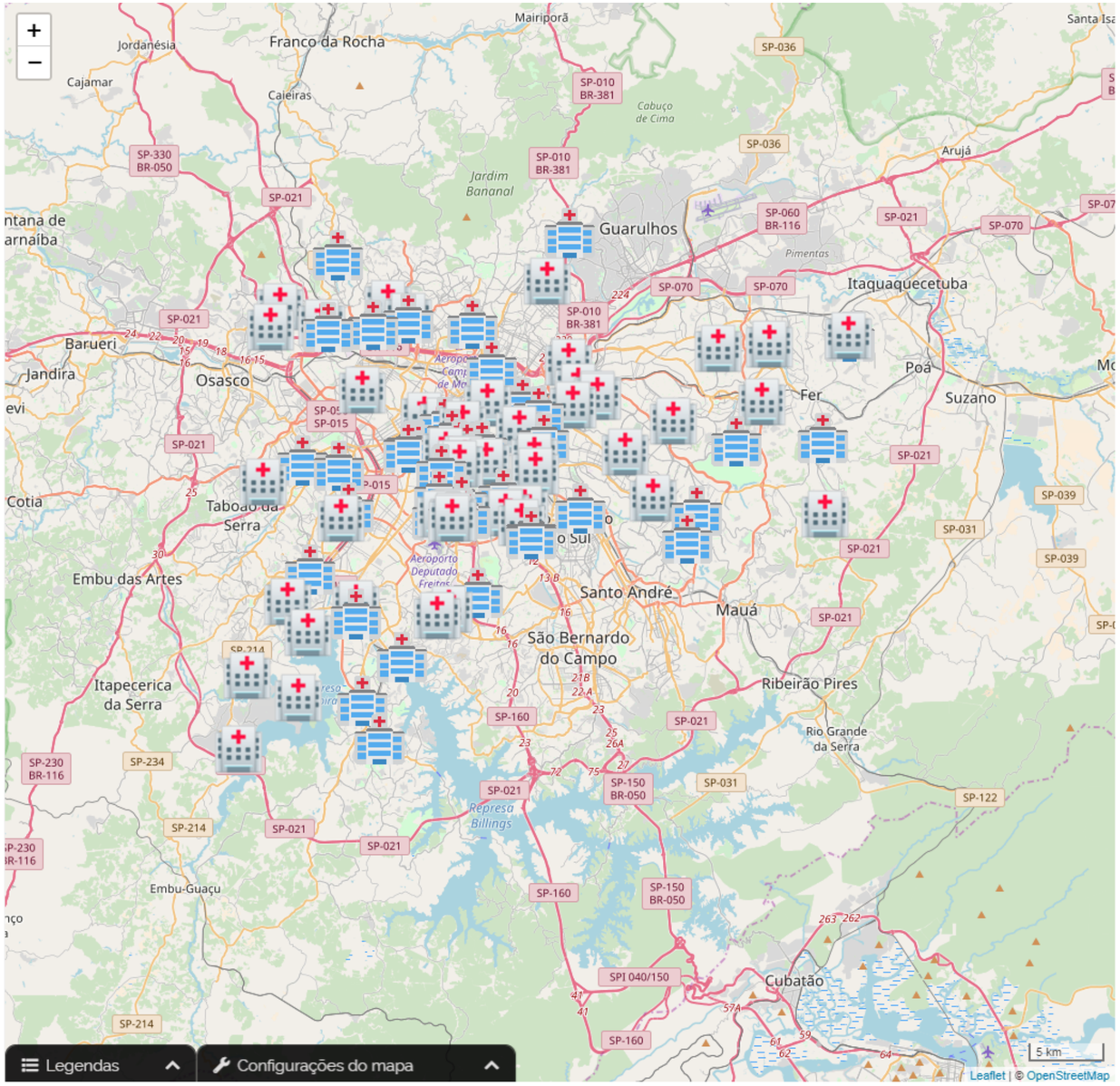
PREFEITURA DE SÃO PAULO
SAÚDE

Health Dashboard

554.202	97	9	8,43 km
Internações Hospitalares	Estabelecimentos de Saúde	Especialidades	Média de Deslocamento

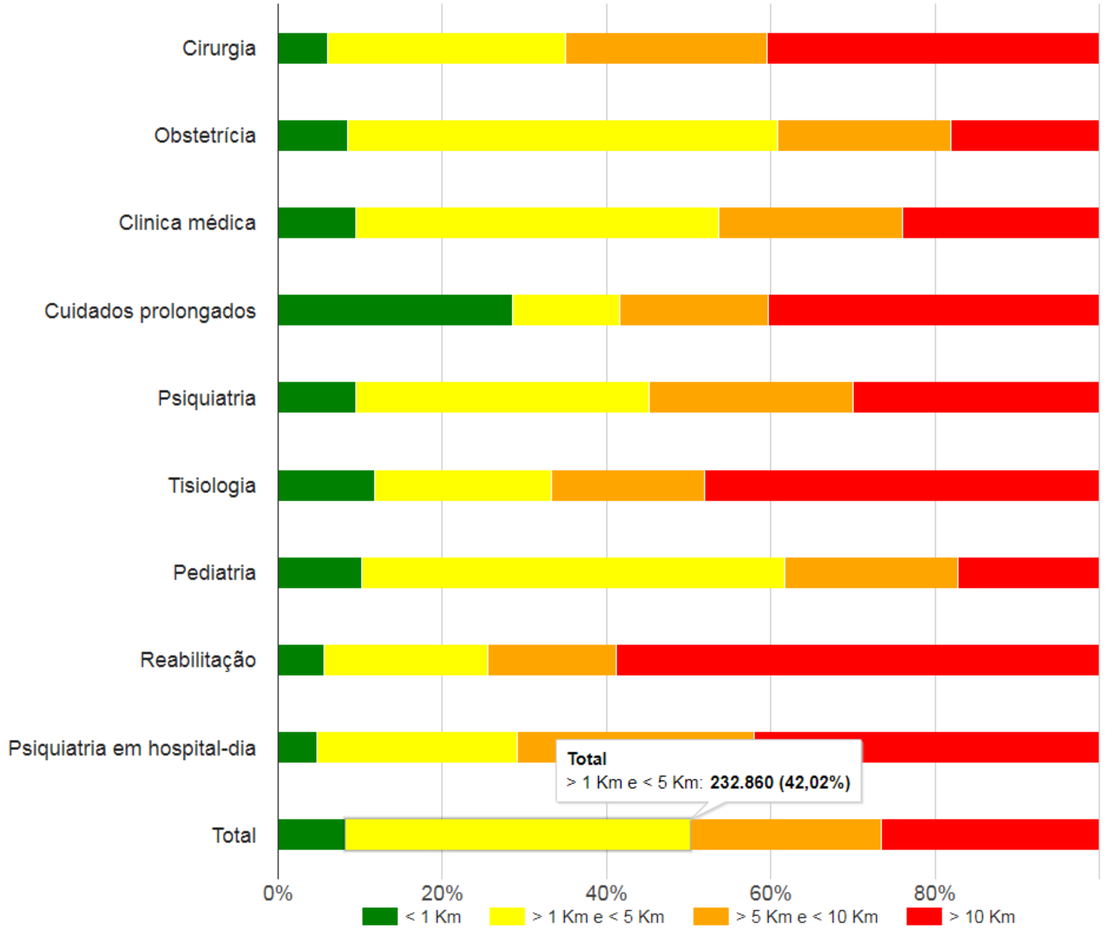
SERVIÇOS



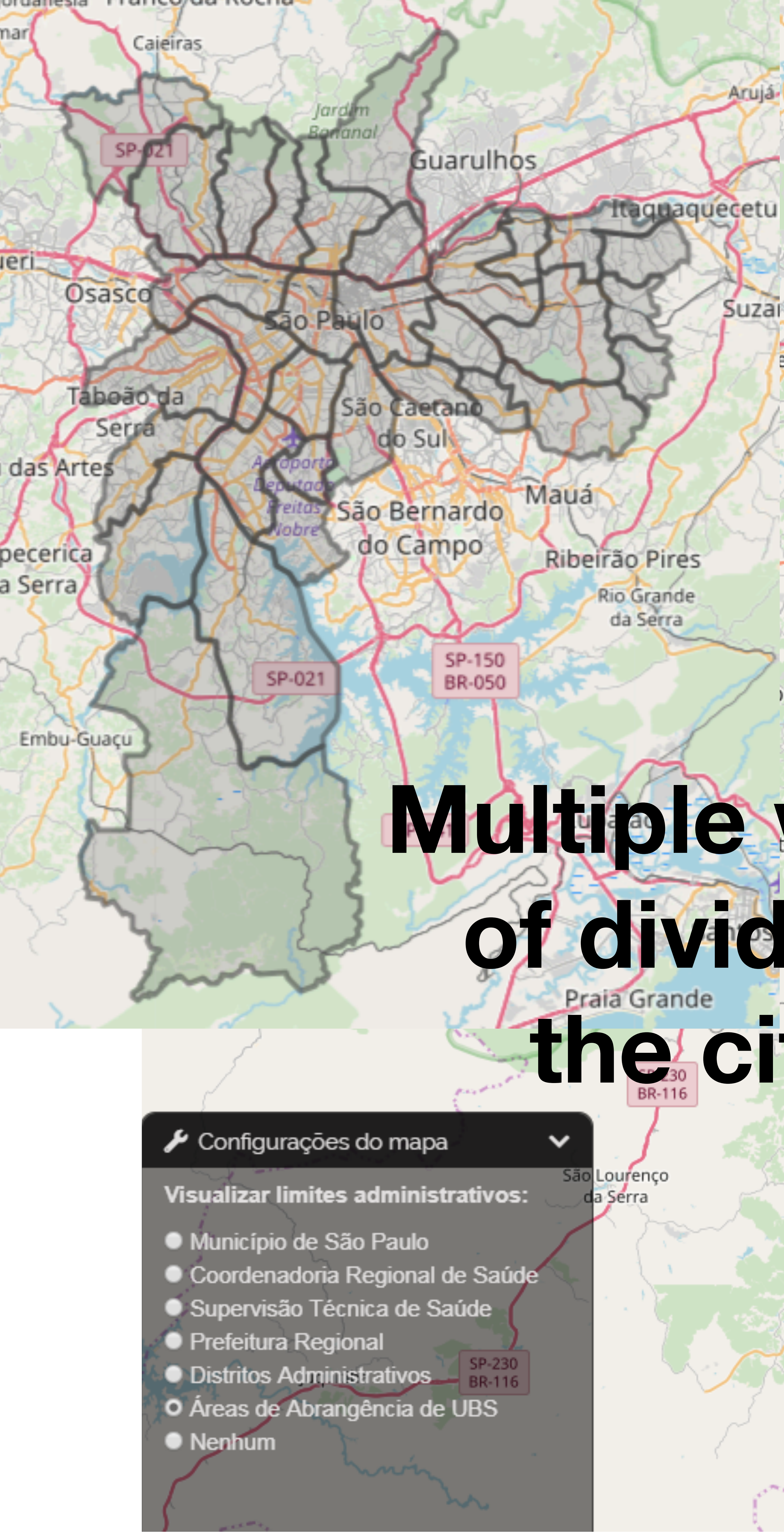


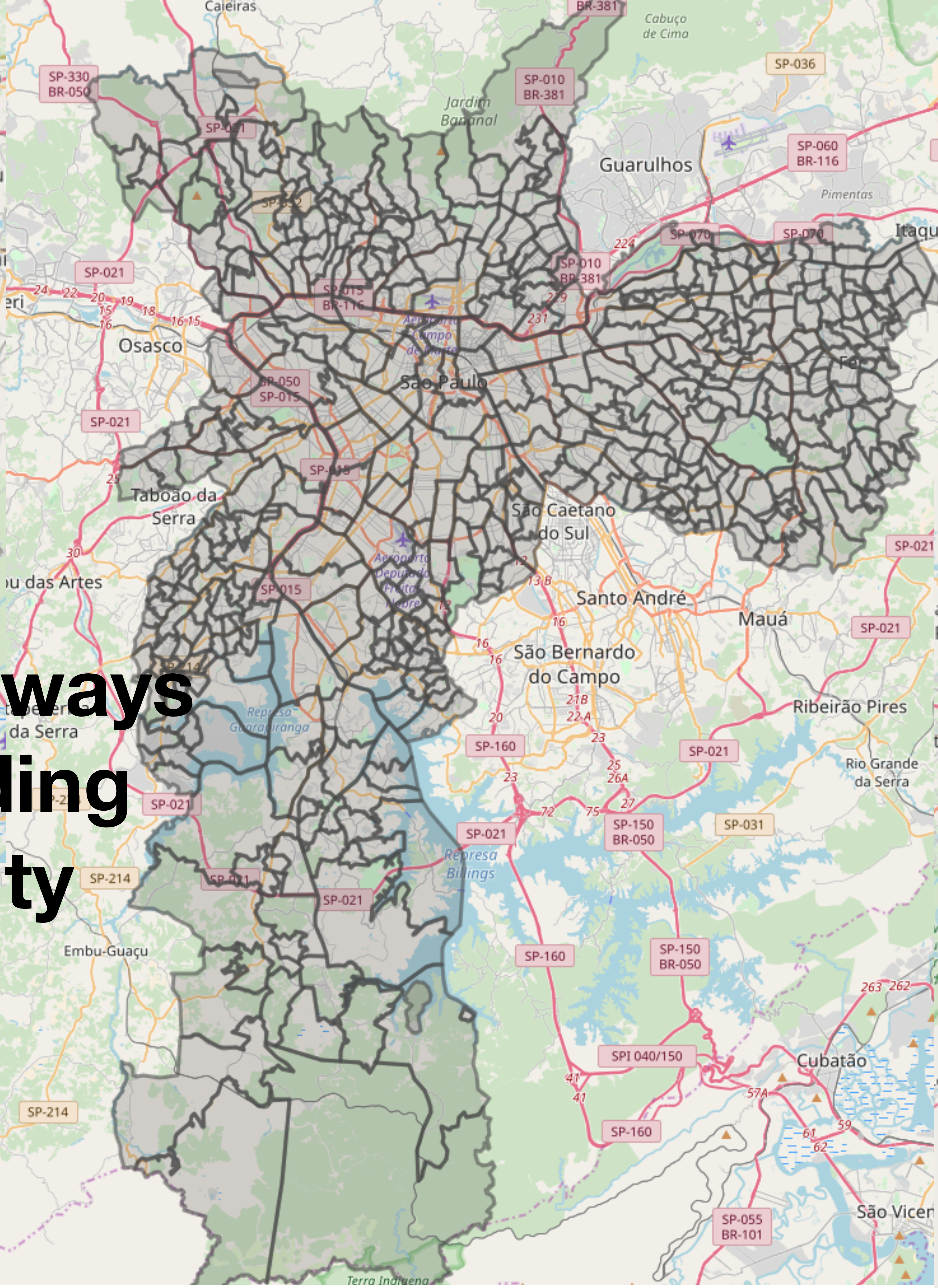
SÃO PAULO

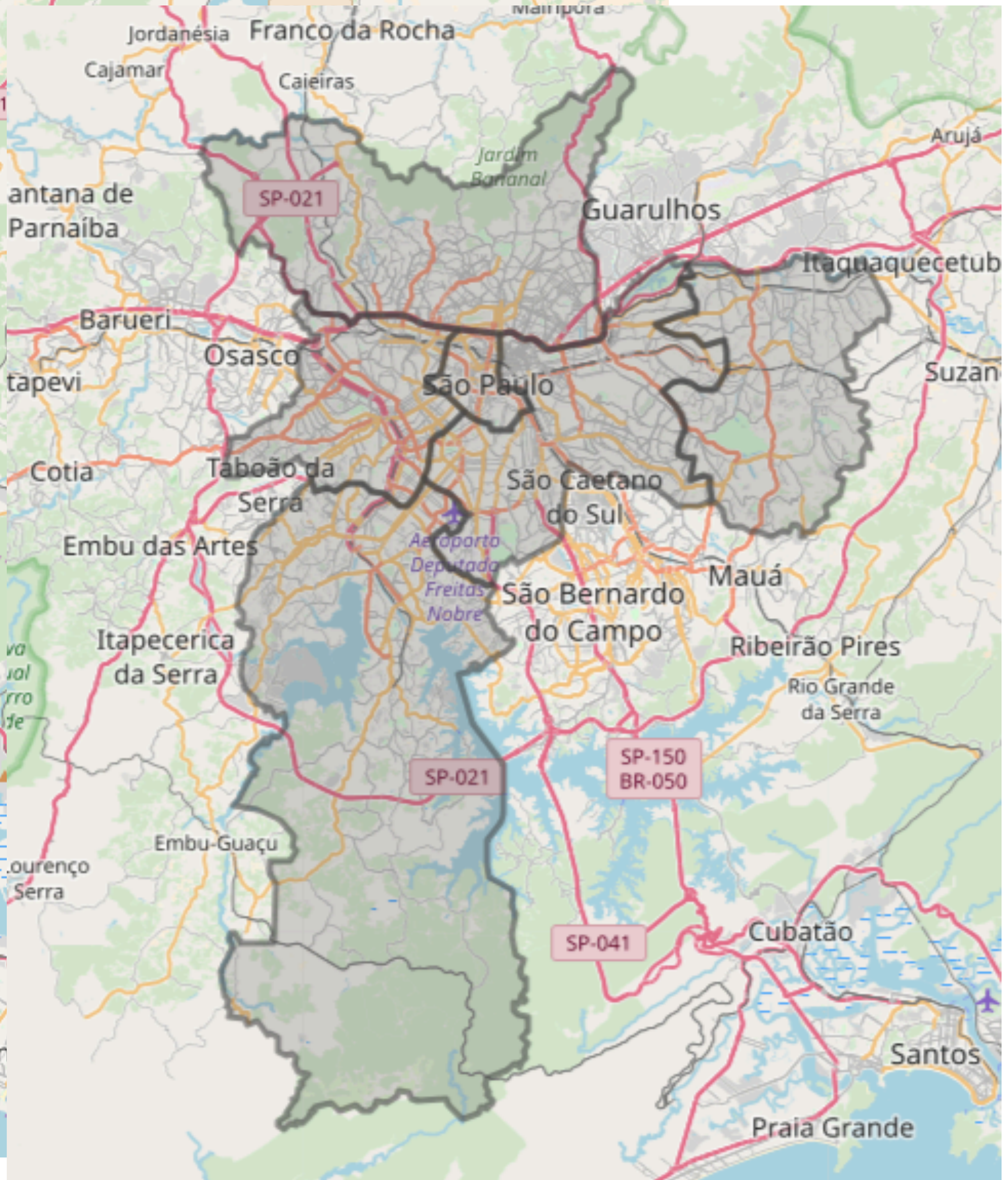
554.202
Procedimentos



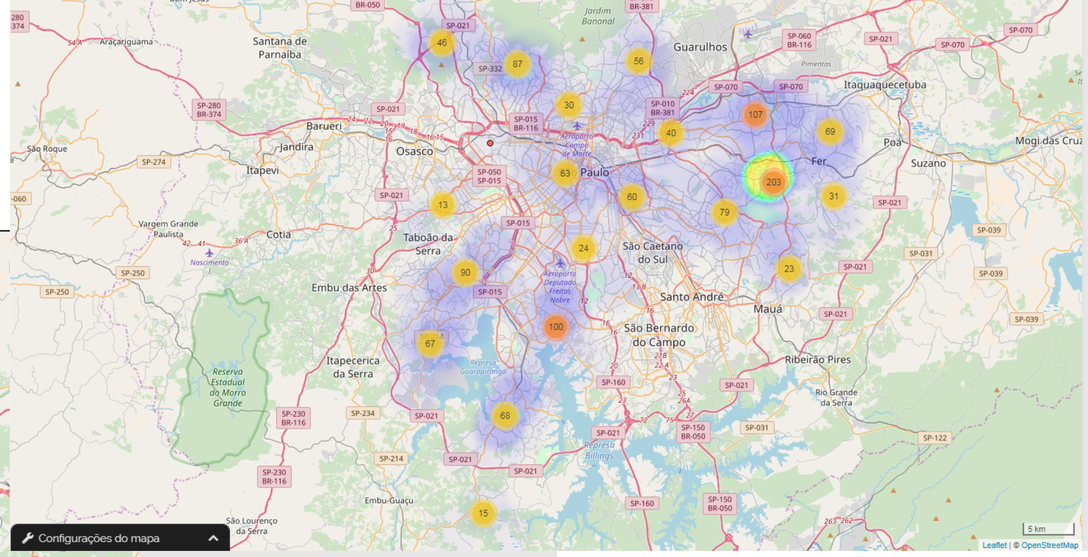
Multiple ways of dividing the city







Leukemia



Competência (aaaamm)
Todos

Grupo do procedimento autorizado
Todos

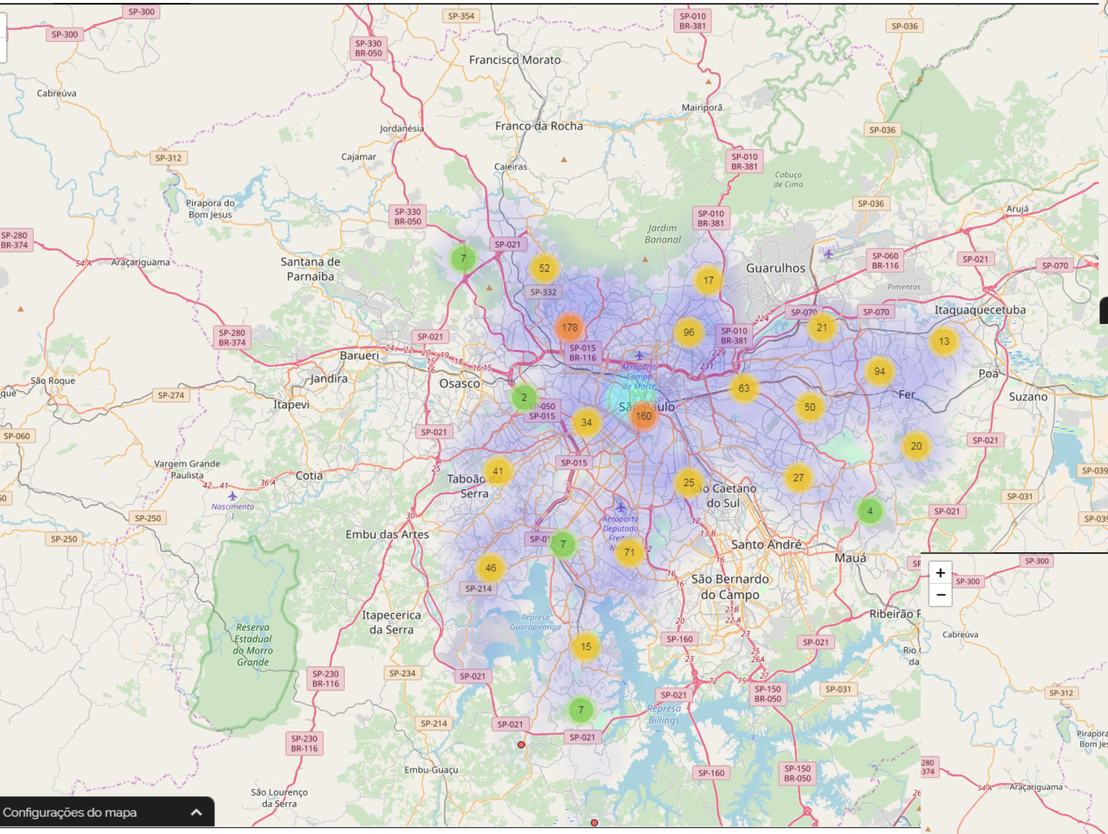
Especialidade do leito
Todos

Caráter do atendimento
Todos

Diagnóstico principal (CID-10)
× C91 - Leucemia Linfóide

Diagnóstico secundário (CID-10)
Todos

Diagnóstico secundário 2 (CID-10)



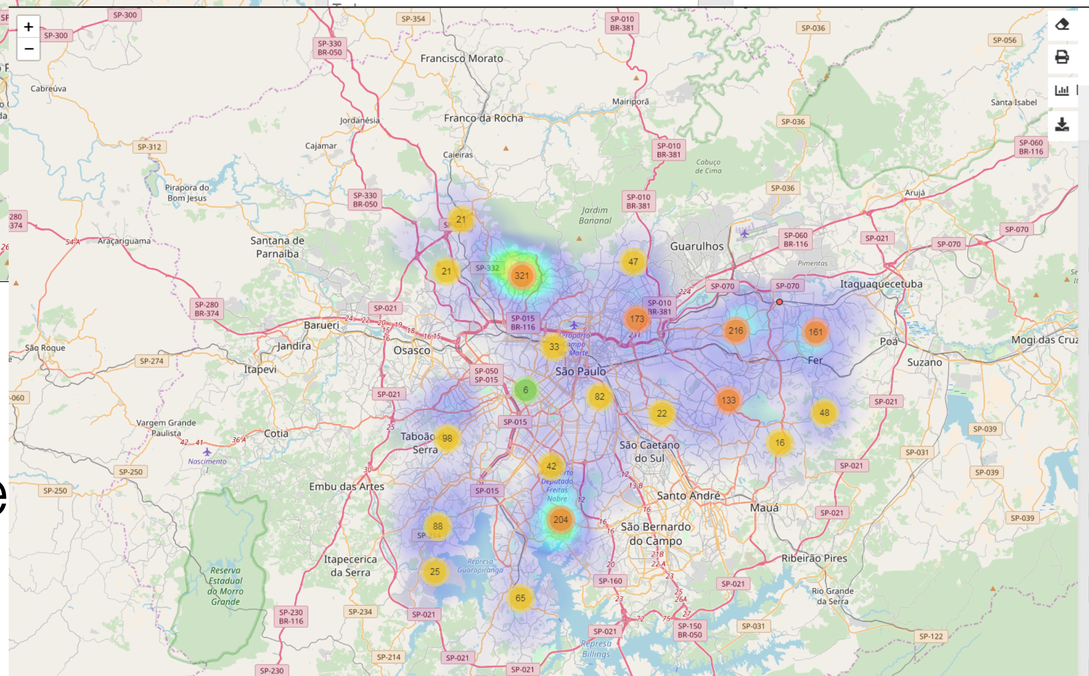
Grupo do procedimento autorizado
Todos

Especialidade do leito
Todos

Caráter do atendimento

HIV

Dengue Fever



Busca automática:

Diagnóstico

Estabelecimento de ocorrência
Todos

Competência (aaaamm)
Todos

Grupo do procedimento autorizado
Todos

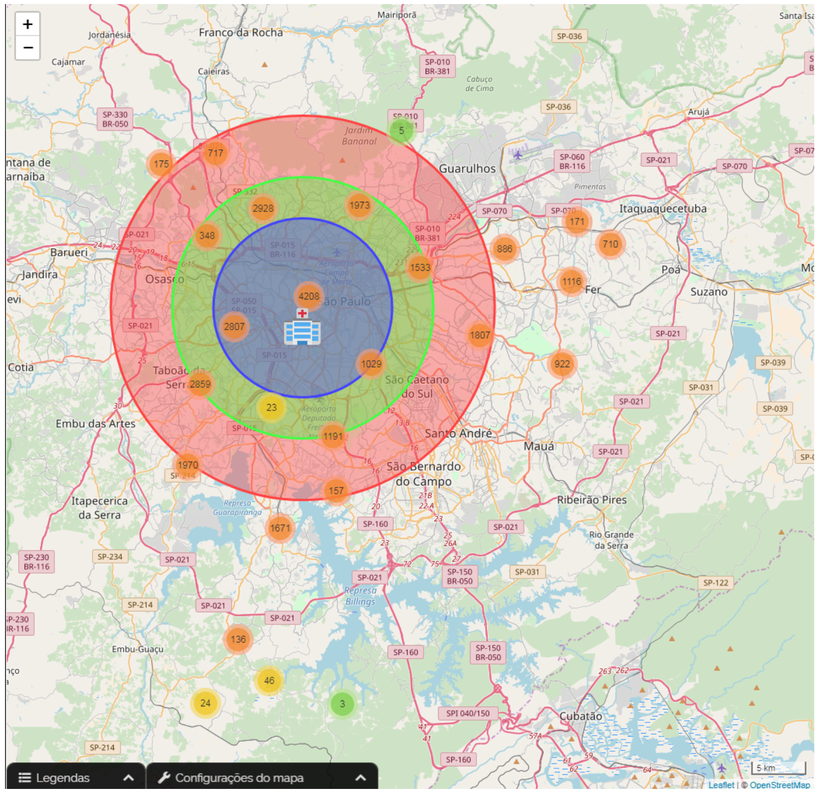
Especialidade do leito
Todos

Caráter do atendimento
Todos

Diagnóstico principal (CID-10)
× A90 - Dengue [dengue Clássico]

Diagnóstico secundário (CID-10)
Todos

Metropolitan Hospital

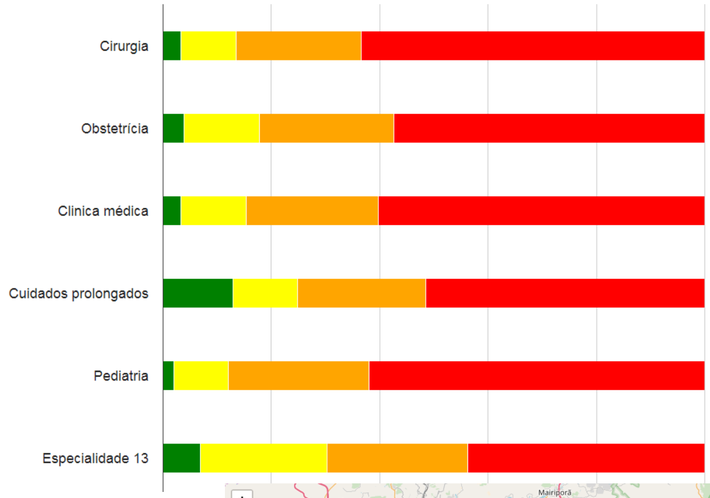


HC DA FMUSP HOSPITAL DAS CLINICAS SAO PAULO

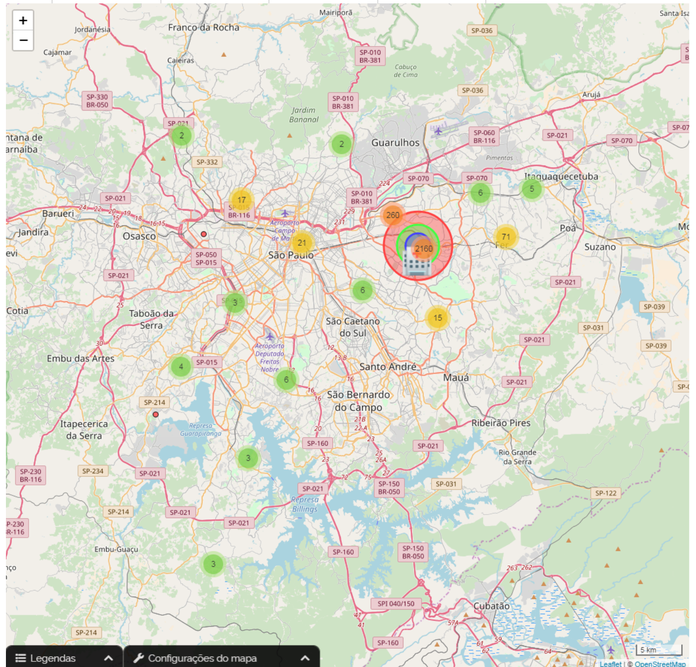
29.415
Procedimentos

Telefone: (11)3087-5456
Leitos: 1506
Distrito Administrativo: JARDIM PAULISTA
Prefeitura Regional: PINHEIROS
Supervisão Técnica de Saúde: LAPA / PINHEIROS
Coordenadoria Regional de Saúde: OESTE

Análise | Esconder Detalhes | Esconder info



Regional Hospital

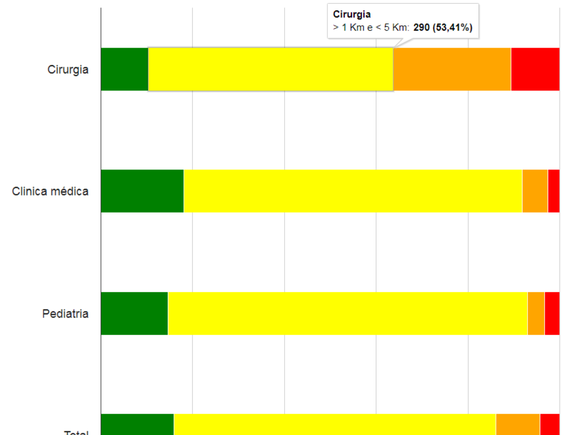


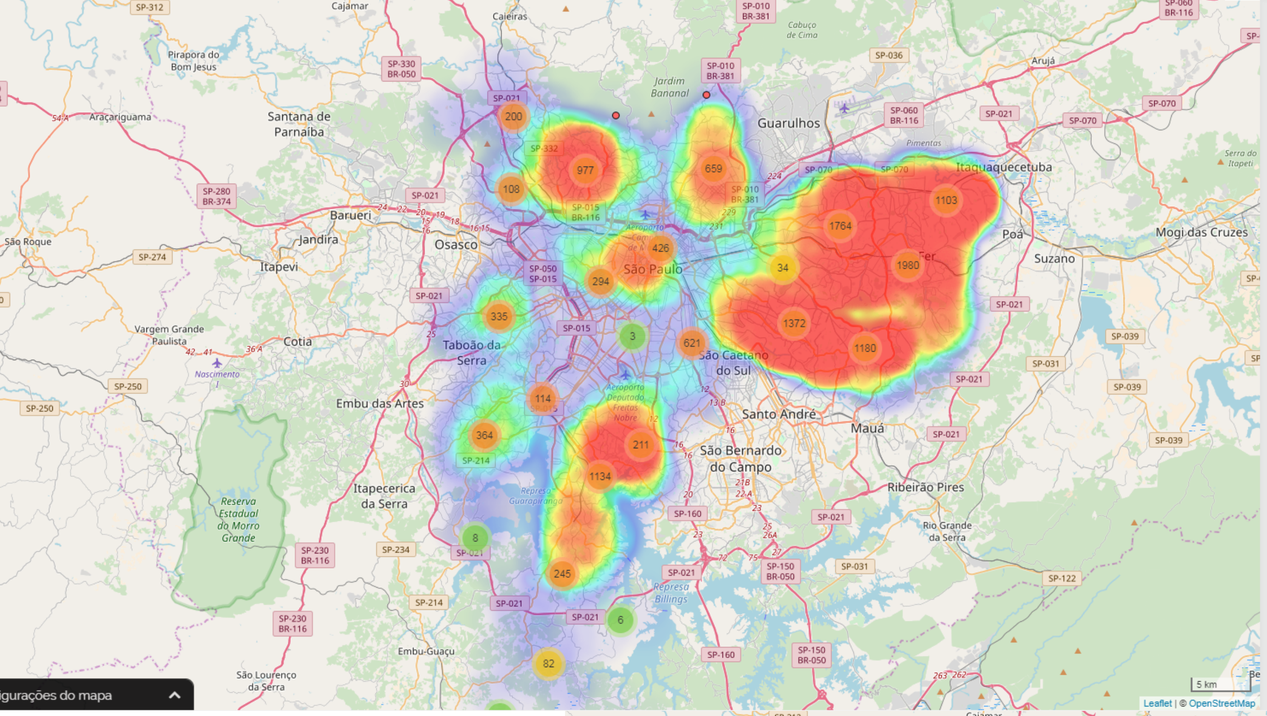
HOSP MUN DOUTOR ALEXANDRE ZAIO

2.586
Procedimentos

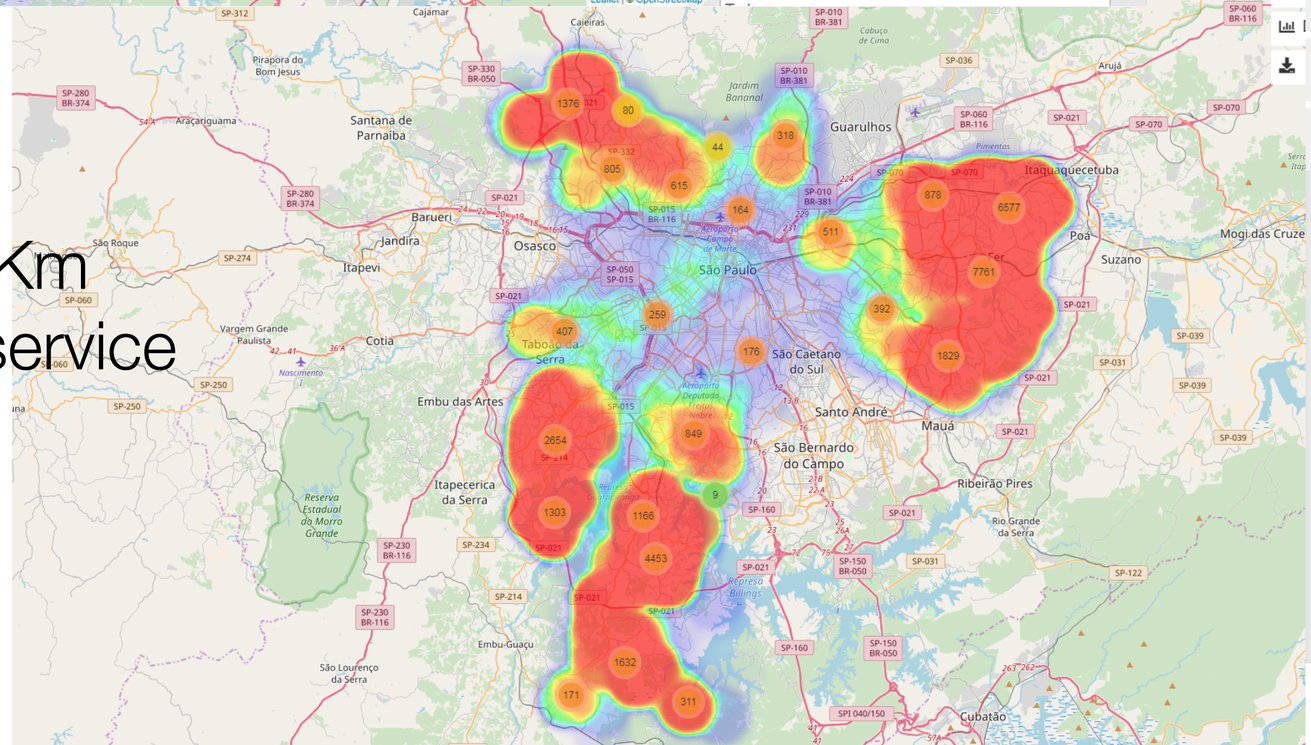
Telefone: (11)3394-9210
Leitos: 42
Distrito Administrativo: VILA MATILDE
Prefeitura Regional: PENHA
Supervisão Técnica de Saúde: PENHA
Coordenadoria Regional de Saúde: SUDESTE

Análise | Esconder Detalhes | Esconder info





Hypertension
(most frequent)

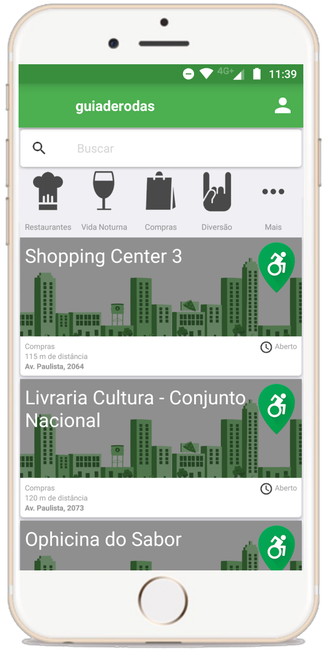


>20Km
to get service

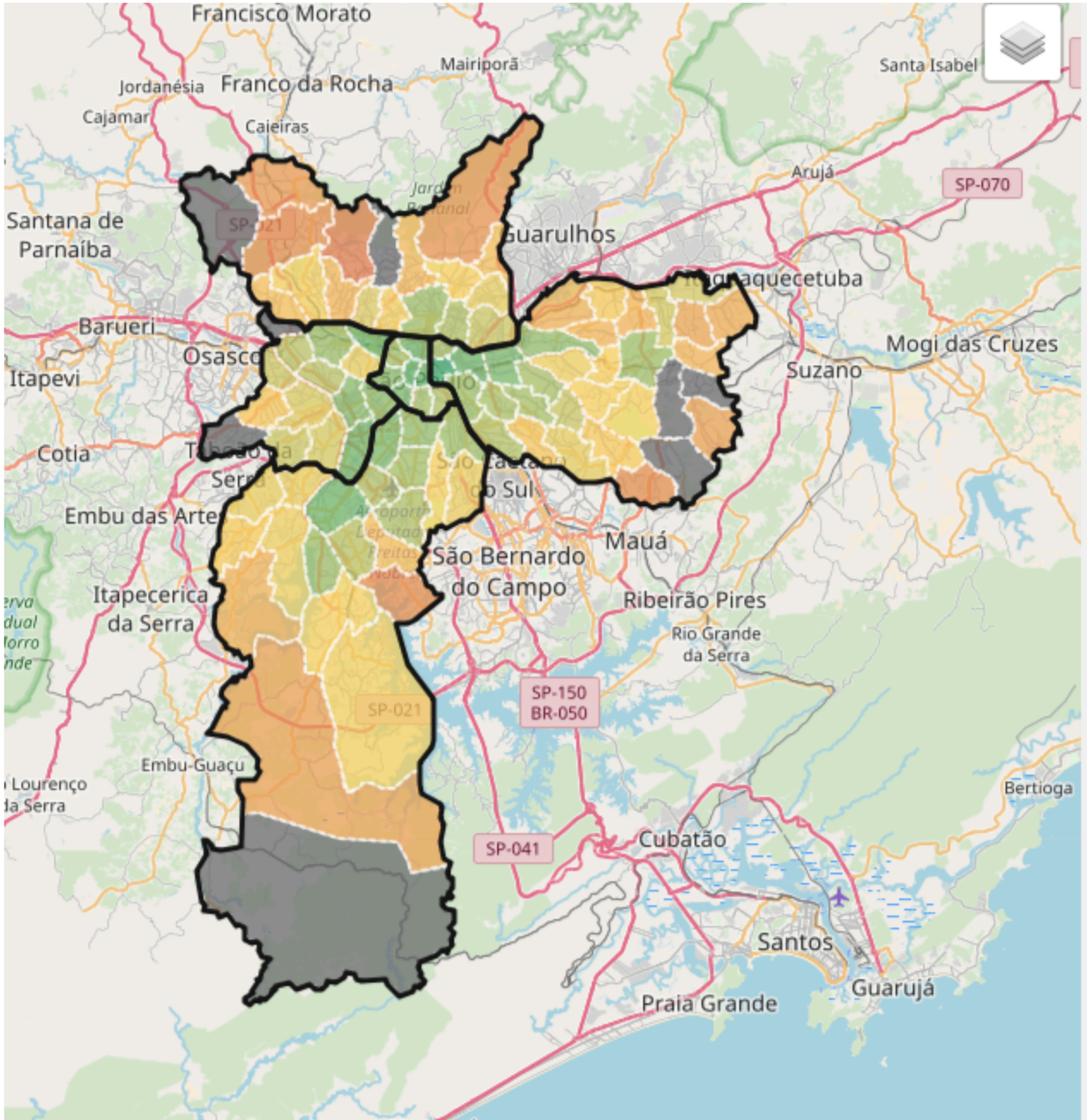
Health Dashboard Challenges

- At the moment it’s a useful interactive tool for the public health professional
- But can we automate part of the work?
- Using ML to detect different patterns for different kinds of diseases?
- Using AI to trigger warnings to health officials?
- Develop models to support long-term planning?

4 -
Crowdsourcing startup App: *guiaderodas*

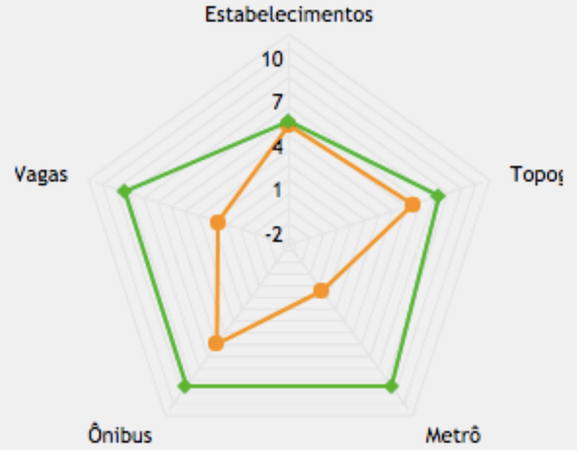


Accessibility Ranking



Sé

8.44 3º

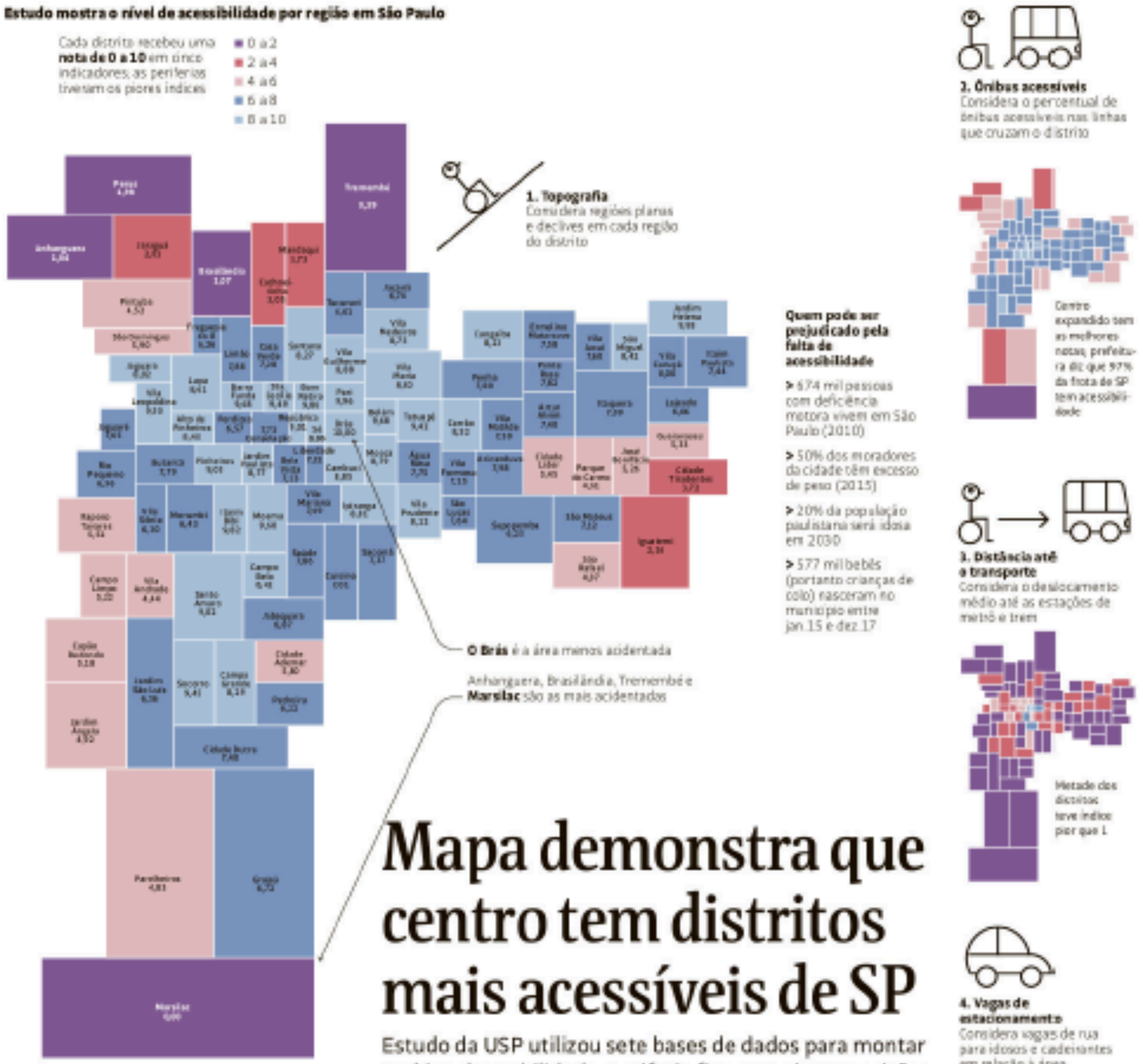


Estabelecimentos	Topografia
6.50 50º	8.86 17º

Metrô	Ônibus
10.0 1º	10.0 1º

Vagas
9.88 4º

Estudo mostra o nível de acessibilidade por região em São Paulo



Mapa demonstra que centro tem distritos mais acessíveis de SP

Estudo da USP utilizou sete bases de dados para montar ranking de mobilidade; periferia fica com piores posições

Jairo Marques e Fábio Takahashi

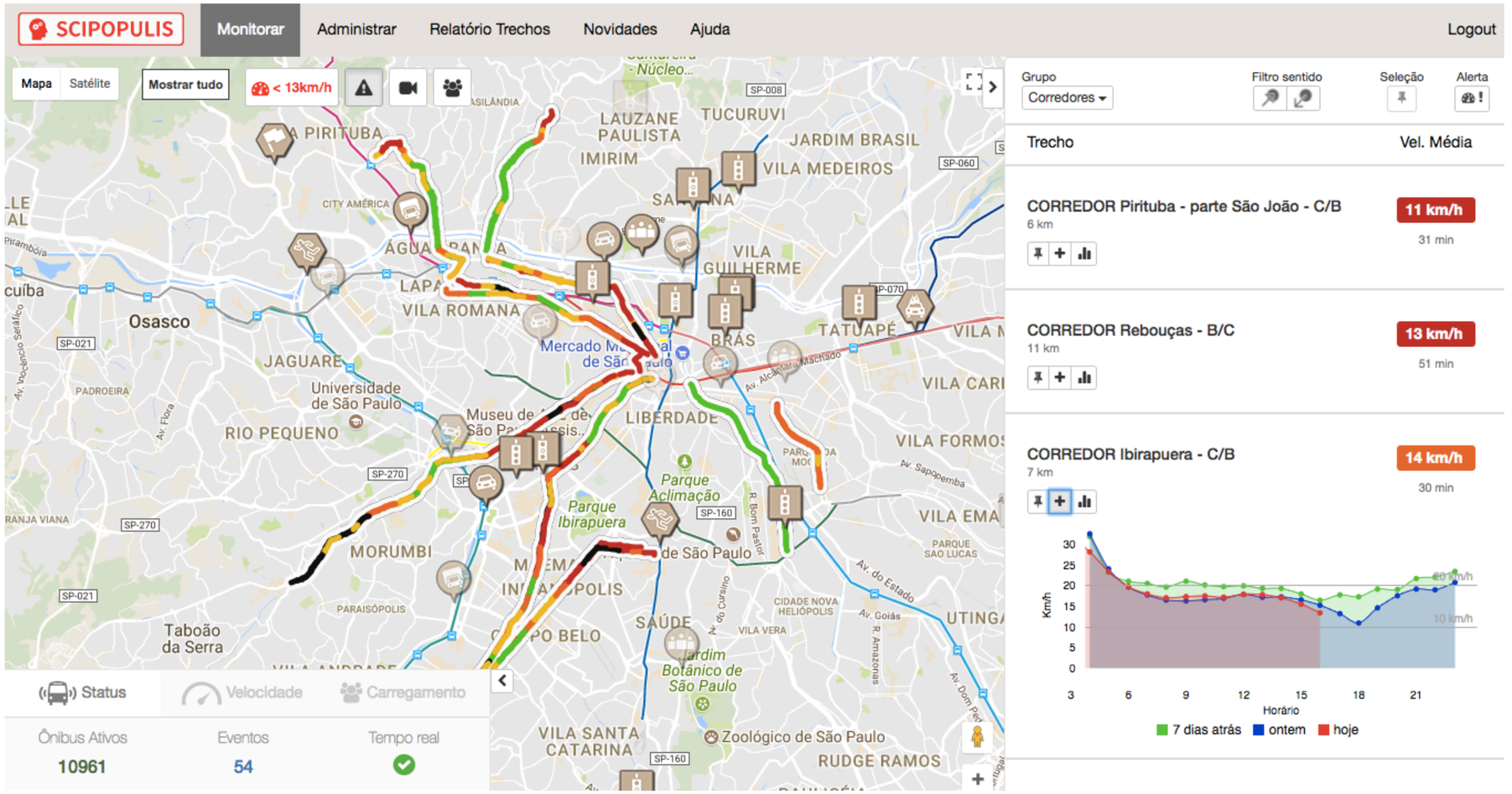
Scipopulis' COLETIVO APP

(for citizens)

SCIPOPULIS

REAL TIME DASHBOARD

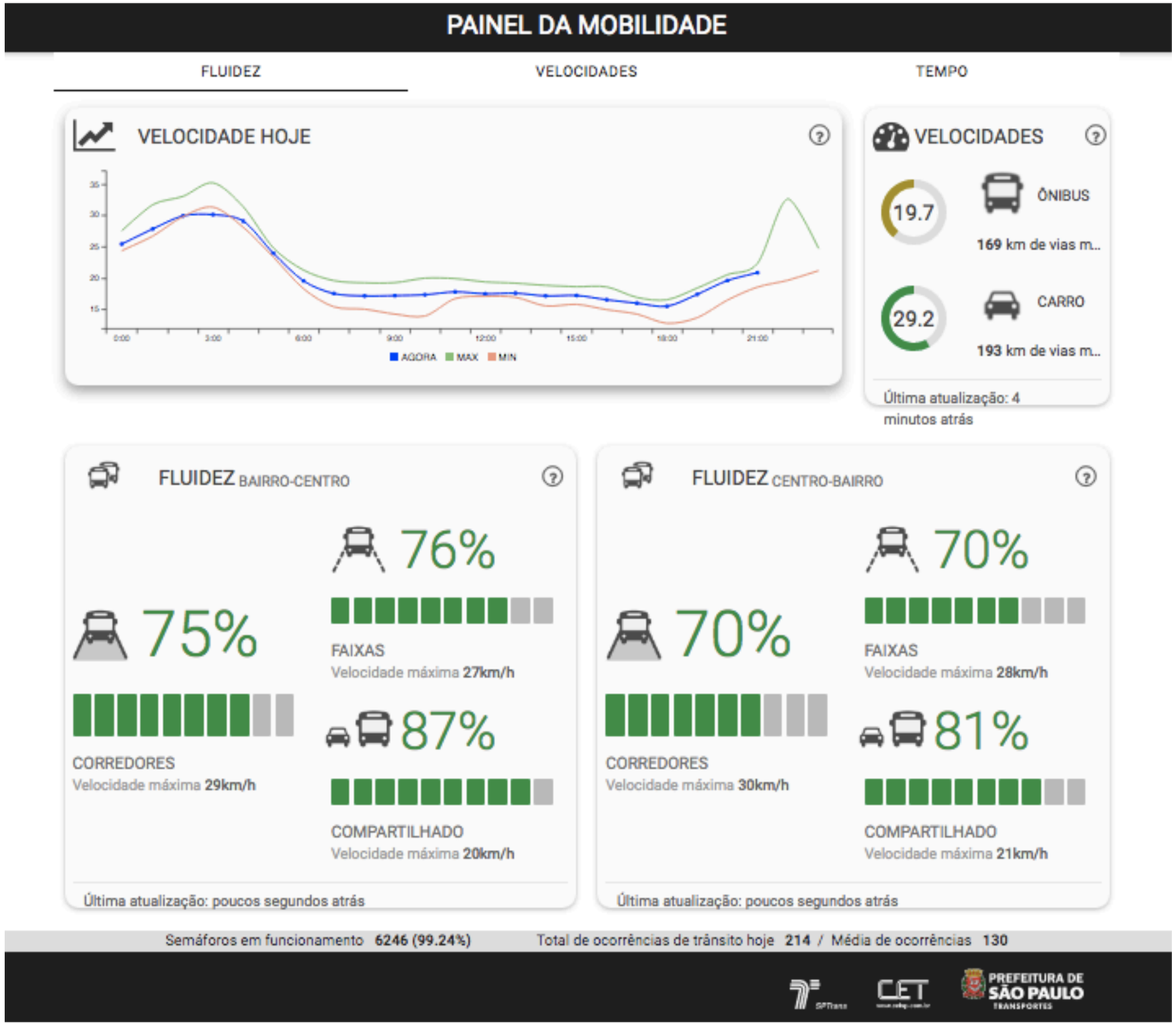
(for system operators)



- In use by the São Paulo secretary of transportation
- in test at: Rio de Janeiro, Curitiba, Santiago (Chile), Brasilia, etc.

MOBILITY PANEL

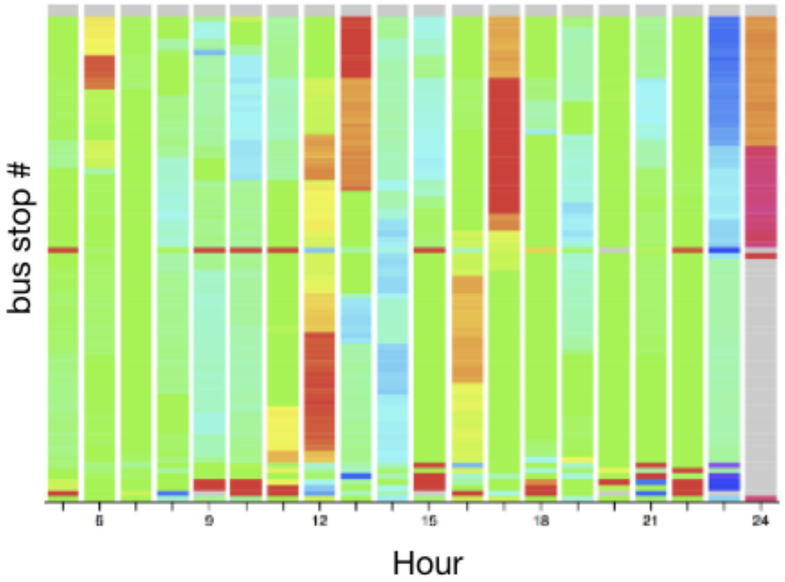
(CONSOLITADED BUS SPEEDS for citizens)



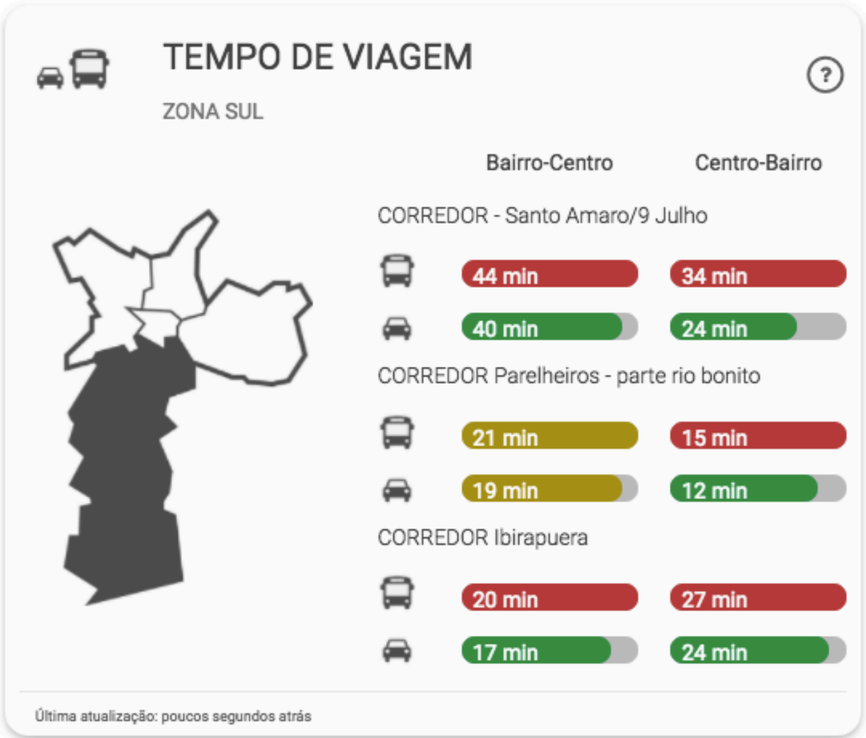
DATA ANALYSIS and visualization

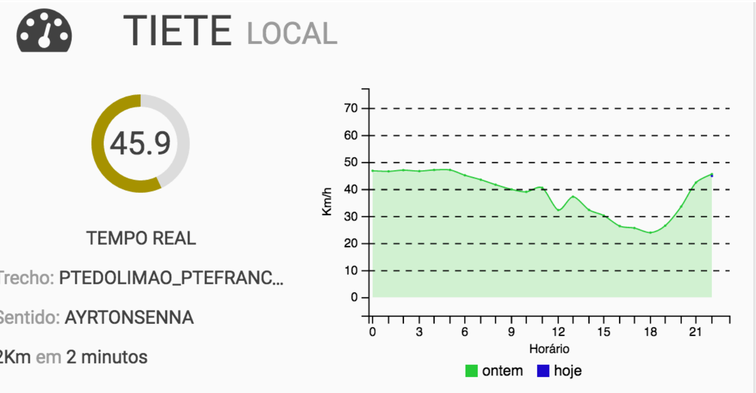
Historical data

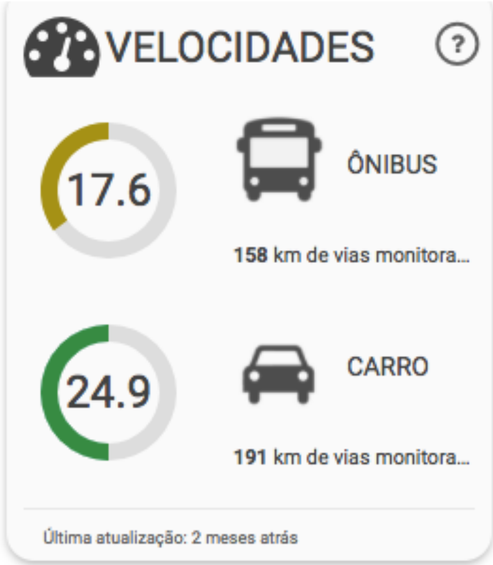
Headway discrepancy per bus stop



Comparing bus x auto



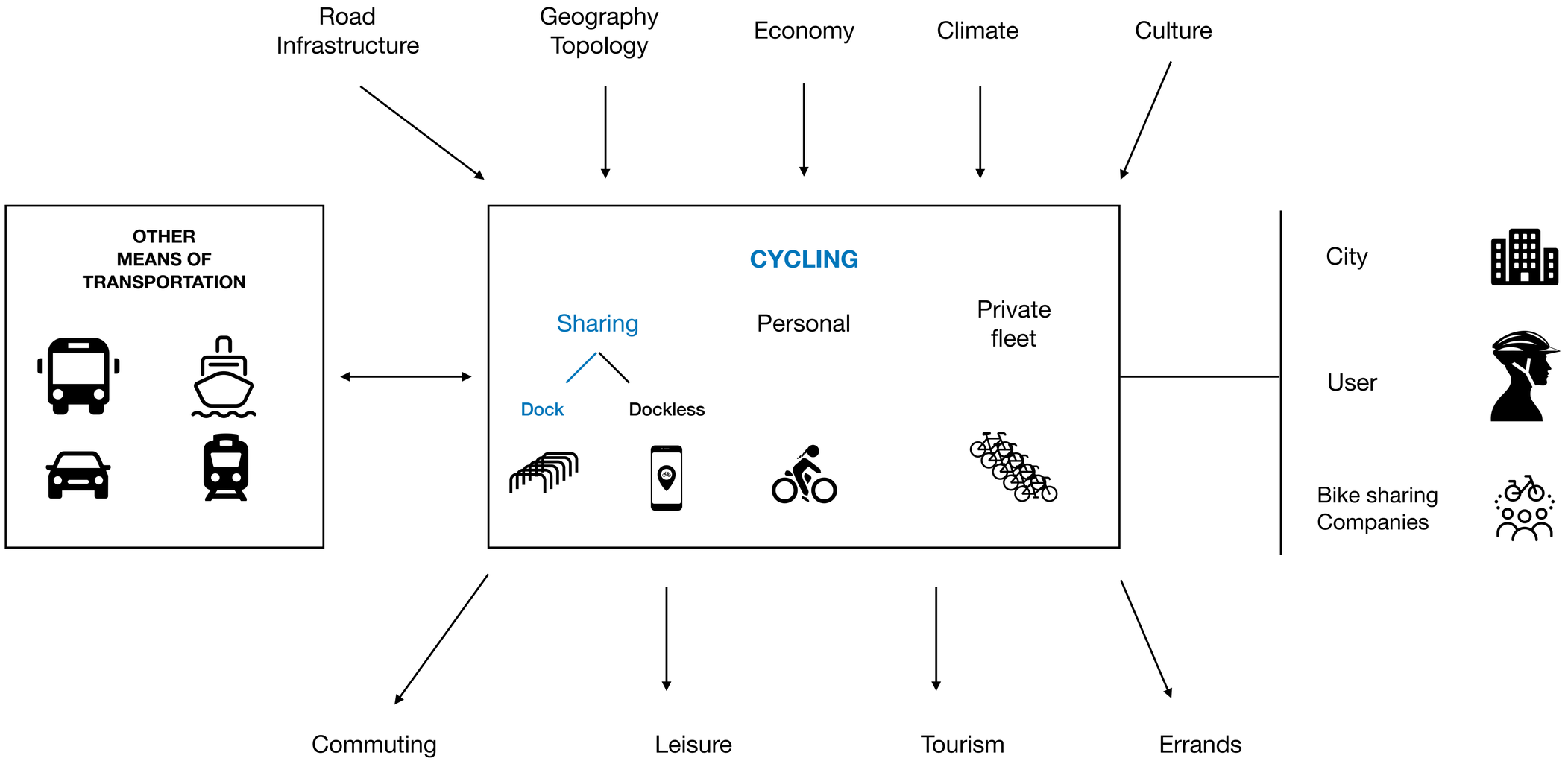




BikeScience w/ MIT Senseable City Lab

- Use of bikes for urban transportation is increasing
  - 18+ million shared bikes, increasing rapidly
- Bike transportation has numerous advantages:
  - for the city
  - for the planet
  - for the user
- But it is highly under-utilized

How can we foster cycling as a serious means of urban transportation

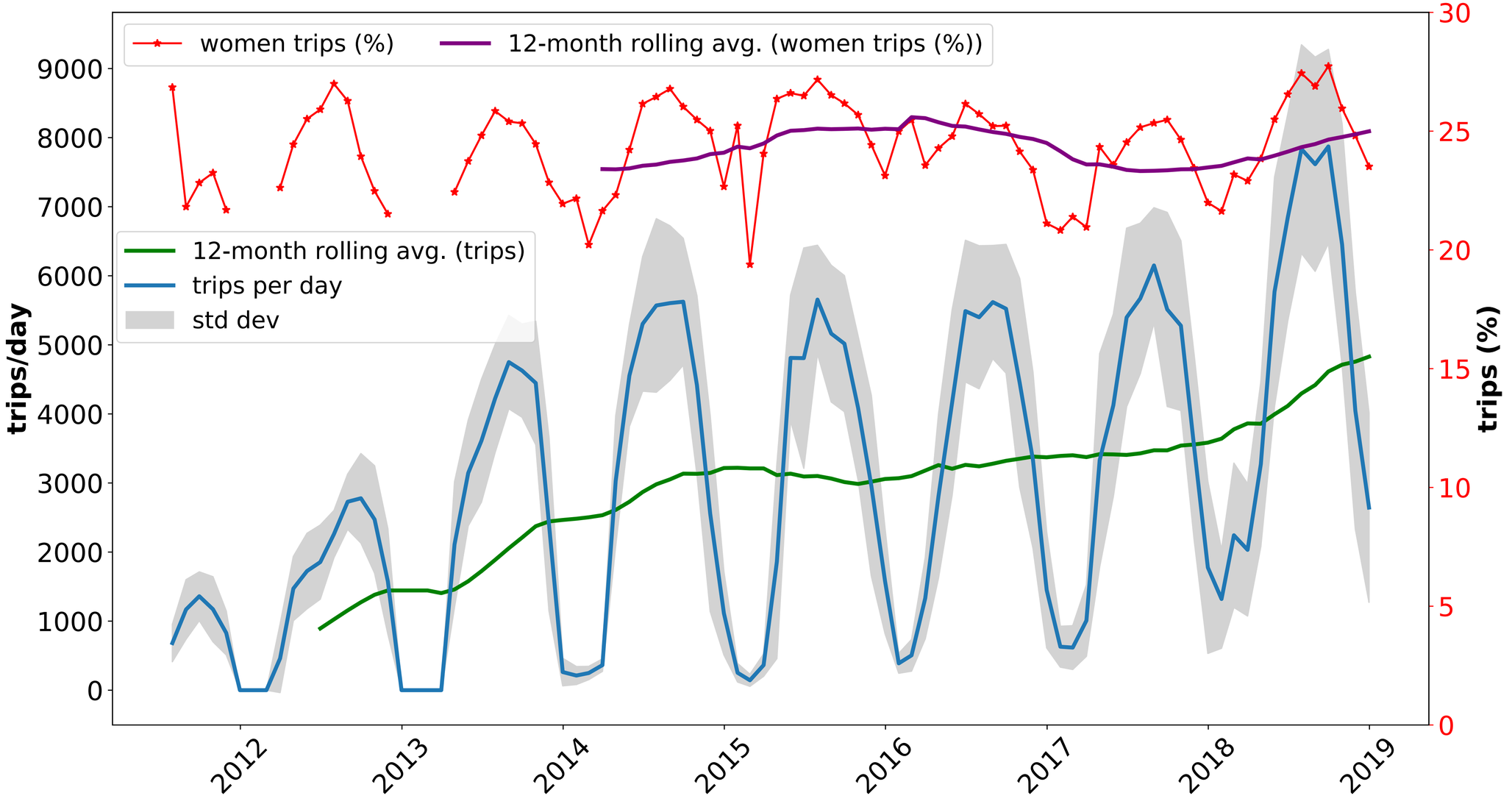


Bike Data Science

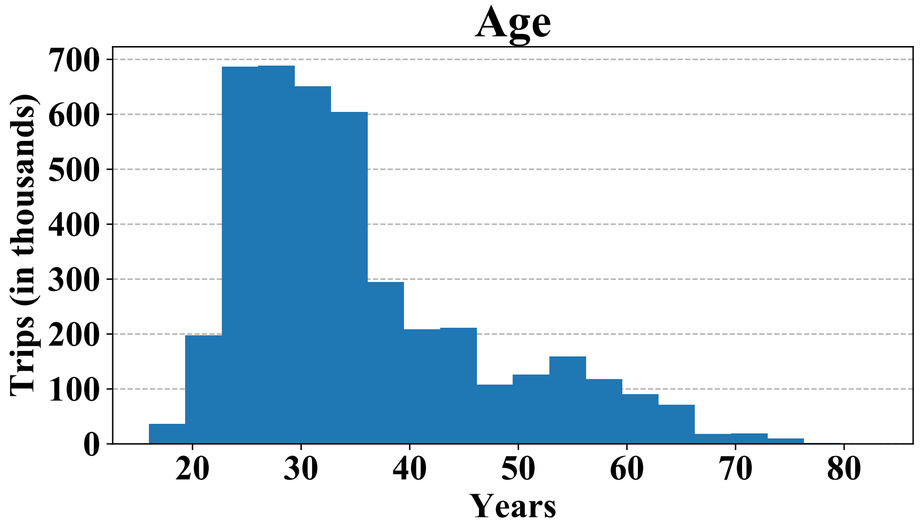
- Analyzing millions of bike trips from 20 cities
  - Starting with Greater Boston (and São Paulo)
  - Dock-based vs. Dockless

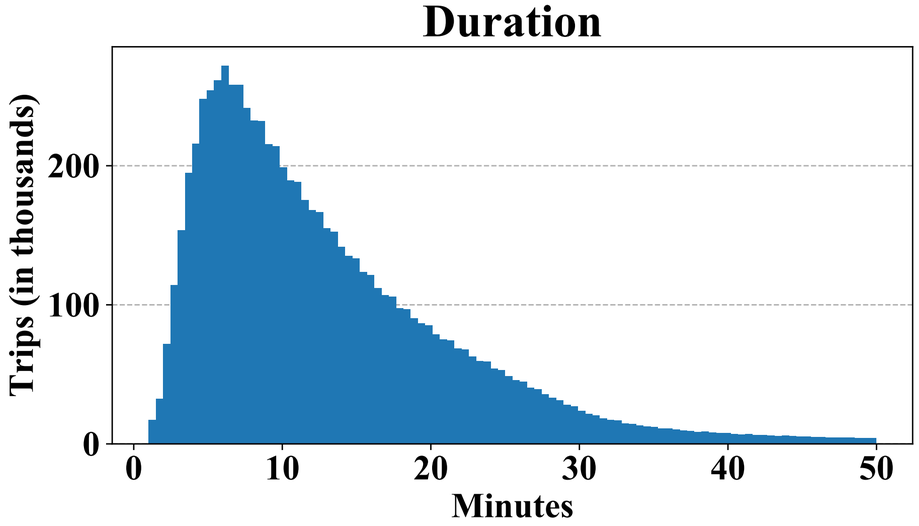
- Greater Boston:
  - 8 million trips since 2011

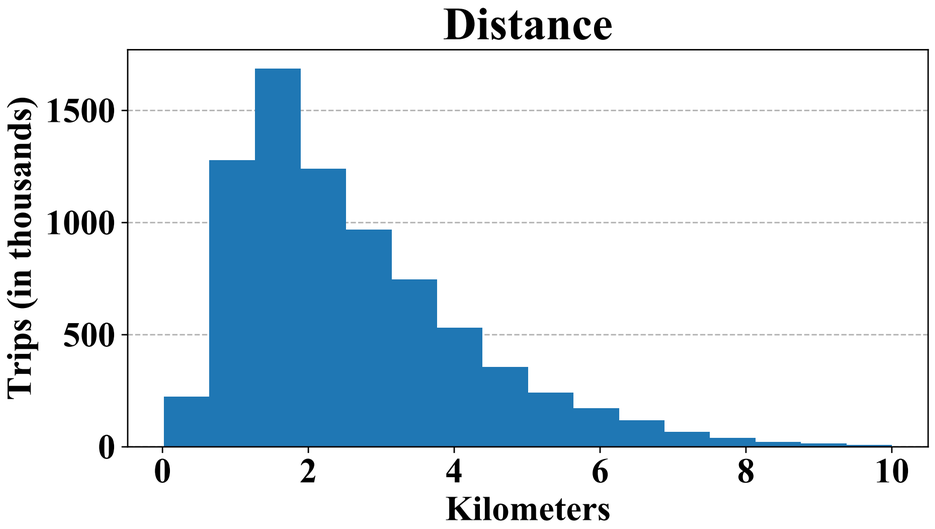
Boston bike-sharing evolution

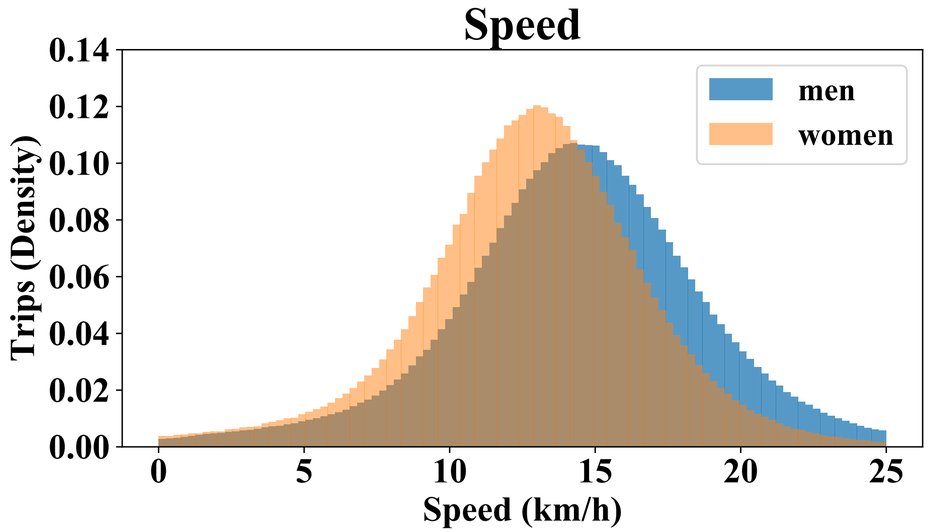


Descriptive Statistics

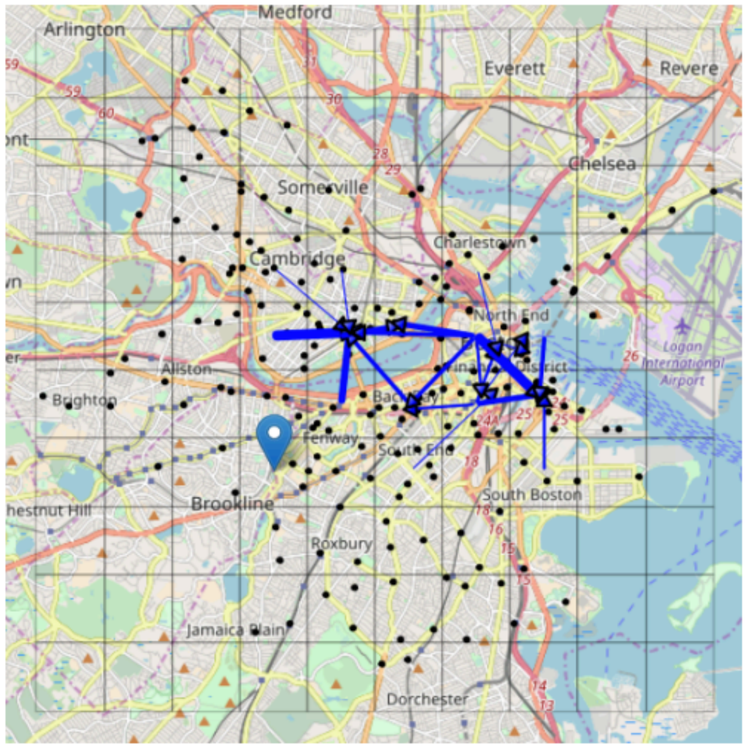




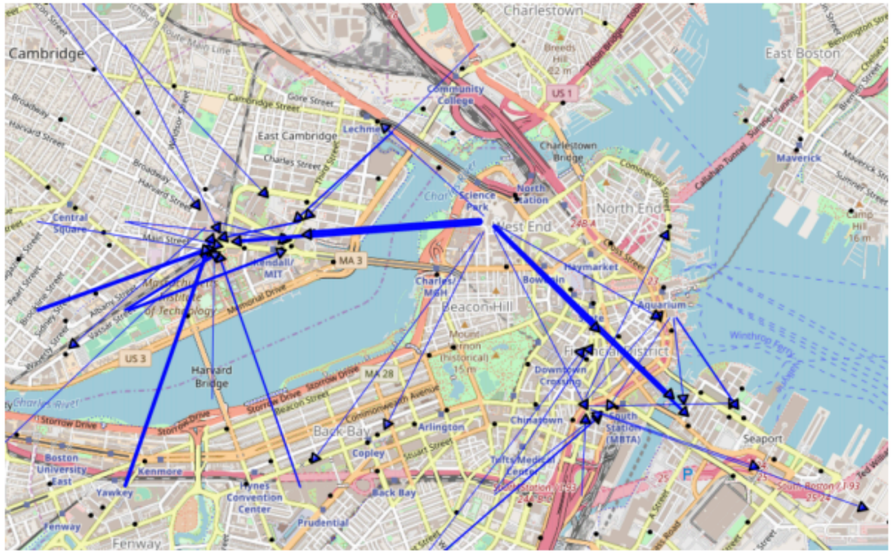




Bike Mobility Flows



(a) 10x10 grid - across cities

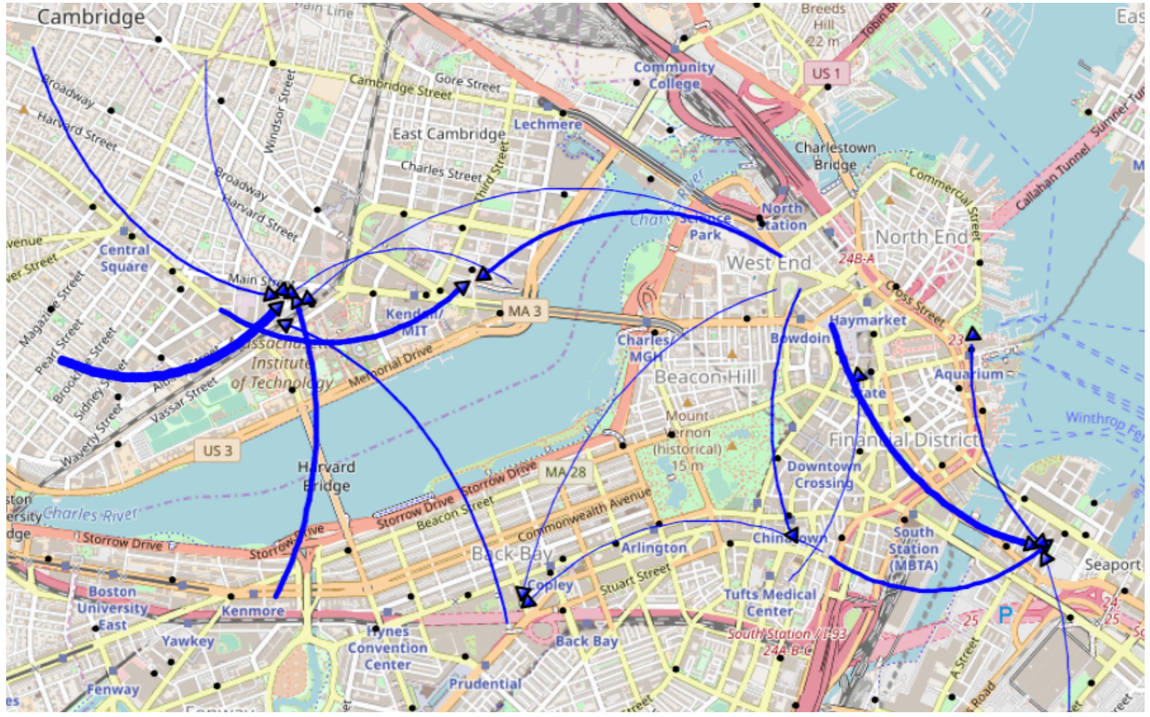


(b) 20x20 grid

(c) 30x30 - flows within a neighborhood

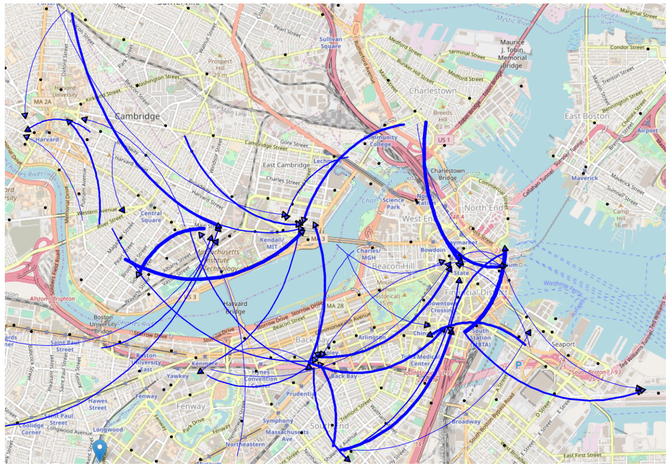
Supporting Public Policy:
Flow popularity and infrastructure investments

Tier 4: 18 flows → 1% of flows

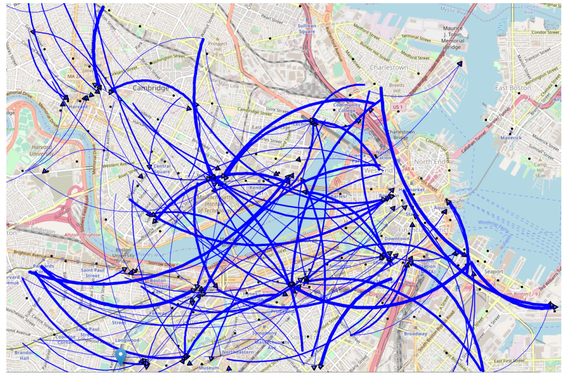


Total: 1629 different flows

Tier 3: 46 flows → 3% of flows



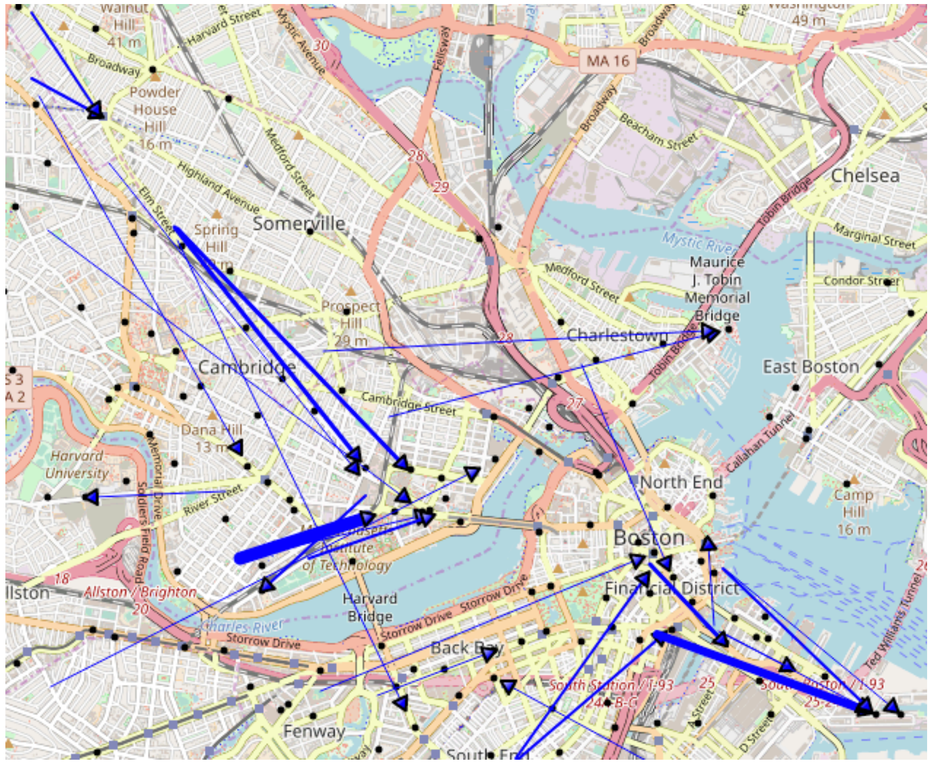
Tier 2: 119 flows → 7% of flows



Tier 1: 1446 flows - 89% of flows

Profile of Speeders (>15Km/h - euclidean distance)

- 4.5% of trips
- 90% are men / 10% are women
- 50% of them are between 18 and 30 years old
- They are present in all age ranges under 52…
  - but higher tendency to drive dangerously fast: 25 to 30
- Speedy trips length is 20% longer
  - (they might speed because they need to go farther away)

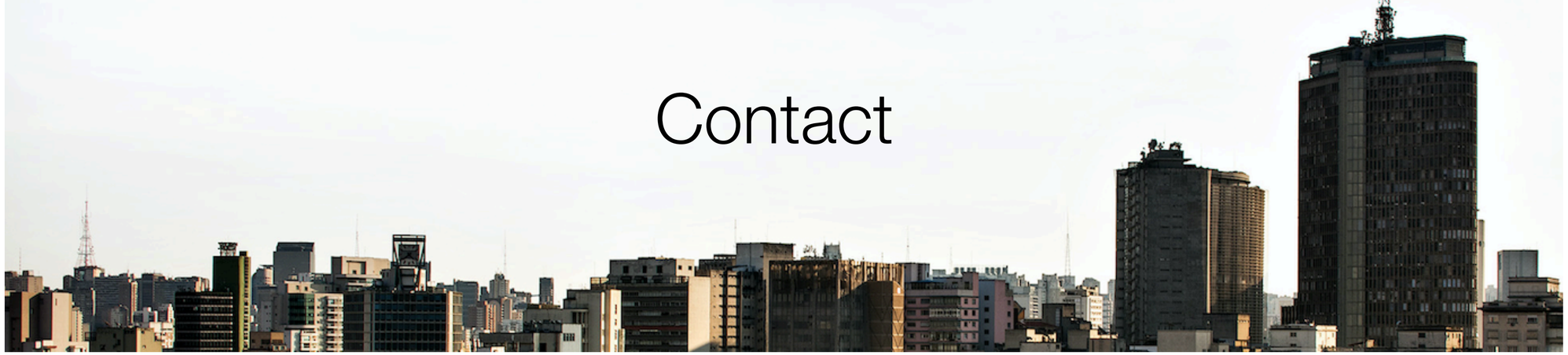


- Speedy trips duration is half of the average (they want to get there quickly)

- A subscriber (normally a resident) is 5 times more likely to be a speeder than an isolated customer (normally a tourist)

The Future

- Advanced collaborative research among InterSCity partners
  - Middleware implementation: scalability, performance, usability by developers
  - Big Data processing, analysis, and visualization
  - Machine Learning to improve city services
- Establish and strengthen collaborations



Contact

kon@ime.usp.br

interscity.org

(PhD and post-doc fellowships available)

Bikes are underutilized

- London - bikes are faster
  - than public transport for most trips < 8 miles [Properly 2013]
  - 1/3 of current car trips [City of London 2017]
- USA [Dept. of Transportation 2017]
  - 35% of car trips are < 2 miles / 46% < 3 miles
  - 1% of trips are on a bike

Next steps

- Analyze flows in 20 cities identifying
  - common patterns and different classes of cities
- Analyze relations with socioeconomic and topographic data from city districts → develop ML model
- Analyze data from dockless systems
-