

Indicadores Econômicos a partir de Informações de Utilities

Eduardo de Rezende Francisco
FGV-EAESP – *Technology and Data Science*
eduardo.francisco@fgv.br





Professor do Departamento de Informática e Métodos Quantitativos da **FGV-EAESP**. Bacharel em Ciência da Computação pelo IME-USP e Mestre e Doutor em Administração de Empresas pela FGV-EAESP. É pesquisador visitante do *Spatial Information Research Centre* da University of Otago, na Nova Zelândia. Ministra disciplinas e desenvolve pesquisas em *Big Data Analytics*, Geointeligência, Estatística Espacial e Mobilidade Urbana.

É CKO (*Chief Knowledge Officer*) da startup **Meia Bandeirada** e sócio-fundador do **GisBI**, grupo de estudos e fomento da integração entre Geotecnologias e Big Data.

Atua no mercado de GIS, Geomarketing e *Business Intelligence* desde 1994. Foi gerente de tecnologia do projeto GIS e atuou em Planejamento de Mercado, Estratégias de Marketing e Planejamento de Pesquisas de Satisfação de Clientes na **AES Eletropaulo** durante 13 anos. É colunista das revistas InfoGEO e GV Executivo, e consultor em Estatística Espacial e Modelos Preditivos para Crédito e *Real Estate*. É membro do Conselho Curador da **Fundação SEADE**.

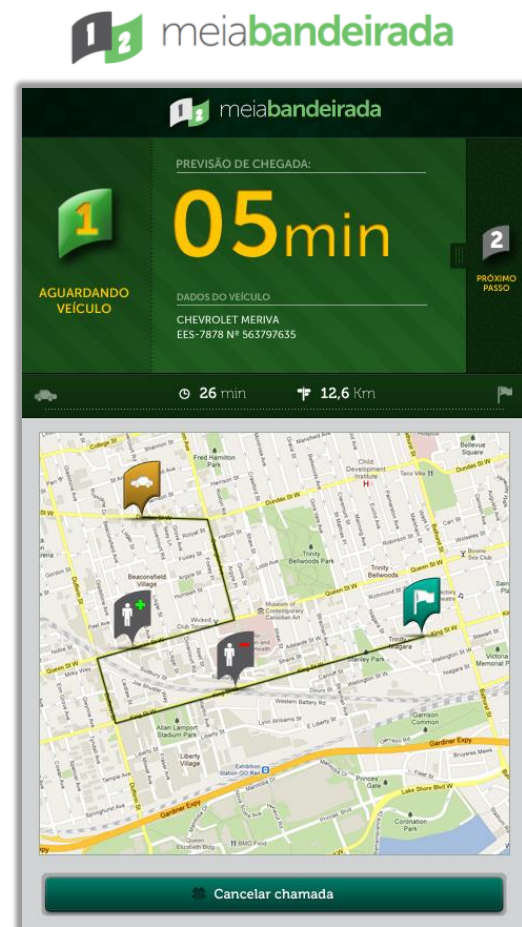


Mobilidade Urbana

Meia Bandeirada

- ❑ Primeira solução automatizada de compartilhamento de táxi (ou multi-modal) baseada em Inteligência Geográfica
- ❑ Economia para passageiros
- ❑ Mobilidade Urbana
- ❑ Mapeamento das "Intenções Geográficas"

APLICATIVO
MOBILE PARA
SMARTPHONES



Mobilidade Urbana

The screenshot displays a web browser window with the URL www.meiabandeirada.com.br/Home. The page features the MB logo and a navigation menu with links: Home, Quem Somos, Soluções, Clientes, and Contato. There are input fields for 'Login' and 'Senha' (Password) with an 'Entrar' (Enter) button. The main content area has a dark green background with a man in a suit using a smartphone. Text on the page reads: 'O aplicativo de mobilidade da MB agora se chama: **MB mobi**'. Below this is a 'Quero Conhecer' (I want to know) button. The footer contains the copyright notice '© 2017 MB - Todos os direitos reservados.com' and social media icons for Facebook and LinkedIn.

MB - MB

Não seguro | www.meiabandeirada.com.br/Home

Apps ★ Bookmarks | Eu Esri 2012 - Submi | 8 negócios promissos | www.aneel.gov.br/ar | Applied Spatial Stati | Independent and De | 5.4 Spatial autoregre

Outros favoritos

MB Home Quem Somos Soluções Clientes Contato

Login Senha Entrar

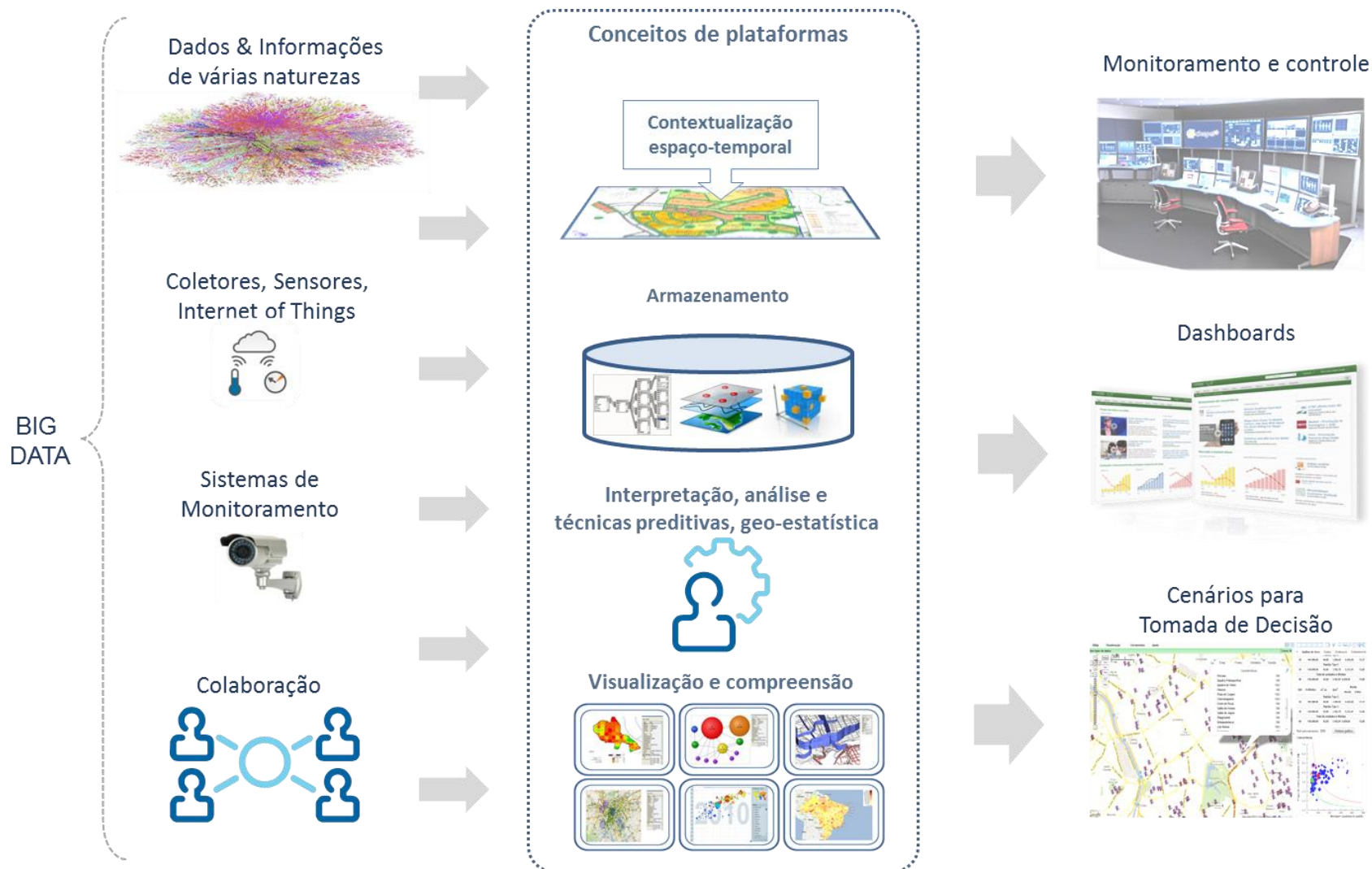
O aplicativo de mobilidade da MB agora se chama:

MB mobi

Quero Conhecer

© 2017 MB - Todos os direitos reservados.com

f in





The screenshot shows a web browser window with the URL www.gisbi.com.br. The page features the gisbi logo at the top left, followed by navigation links: Princípios, Mercados, Produtos, Maptrix, and Mais. A search bar with the text "login | cadastro" and a "Pesquisar" button is located at the top right. The main content area has a large red banner with the text "GEOFirst" in white. Below the banner, the text "O CONCEITO gisbi DE COMO UTILIZAR MAPAS EM EMPRESAS E INSTITUIÇÕES PÚBLICAS E PRIVADAS" is displayed in red. A section titled "Maptrix" by gisbi includes a link to "Conheça o software de georeferenciamento mais fácil de operar". The "Bigdata" section is titled "Bigdata em empresas e instituições públicas" and describes the combination of geotechnologies with Business Intelligence and access to exclusive data banks.

login | cadastro

gisbi

Princípios Mercados Produtos Maptrix Mais

GEOFirst

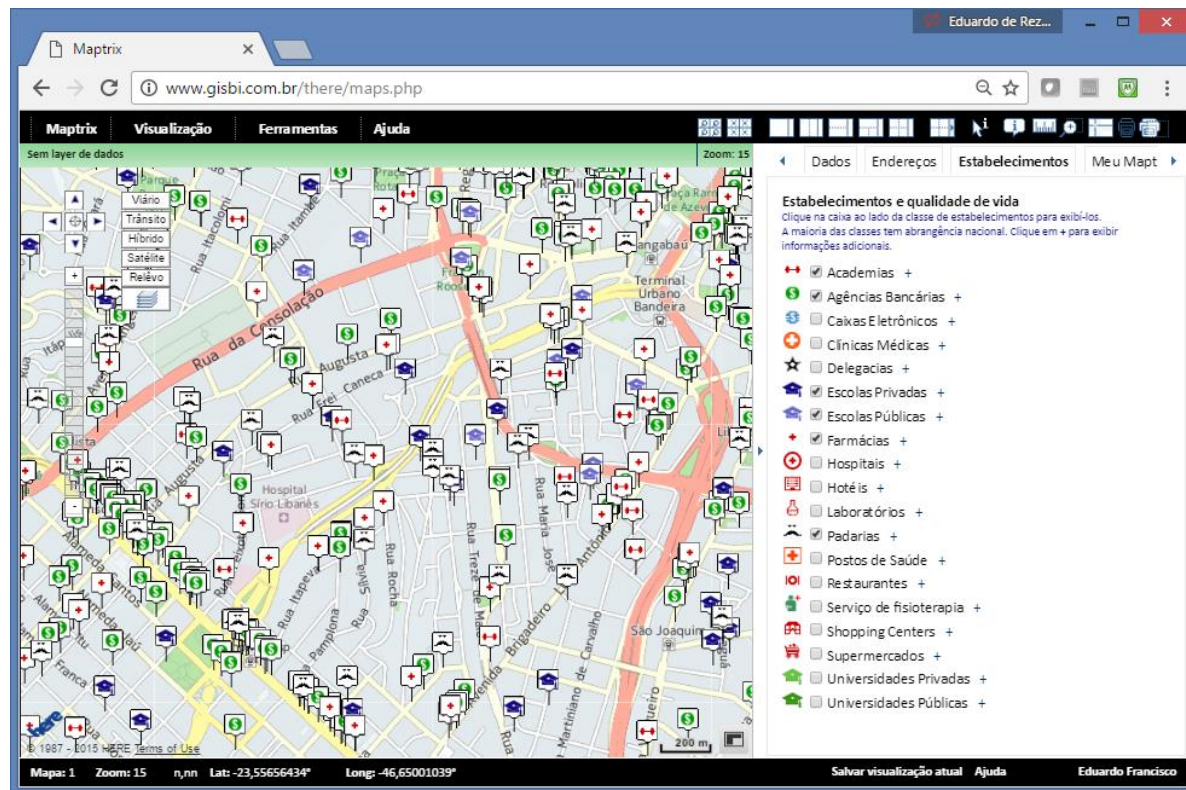
O CONCEITO gisbi
DE COMO UTILIZAR MAPAS
EM EMPRESAS E INSTITUIÇÕES PÚBLICAS E PRIVADAS

Maptrix by gisbi [Conheça o software de georeferenciamento mais fácil de operar](#)

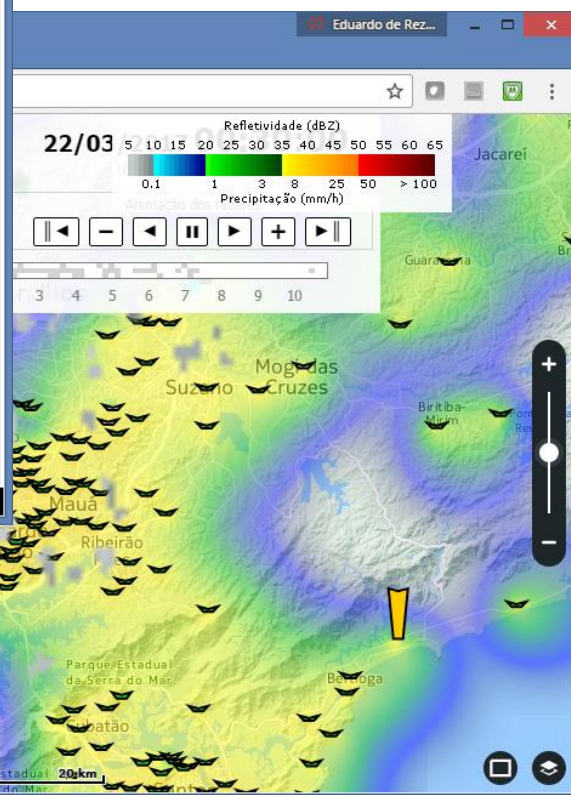
Bigdata

 em empresas e instituições públicas

A combinação das geotecnologias com os conceitos de *Business Intelligence* e o acesso a bancos de dados exclusivos, através do software [Maptrix](#), cria novas condições para melhorar a gestão dos negócios e a administração pública, com o aproveitamento de informações reais da operação empresarial ou das cidades e



Chuvview



Maptrix

Acquisition

Marshalling

Analysis

Action

Data Acquisition



VLDW and BI Appliances



Analytics



BPM & Action



Including Complex Event Processing (CEP) tools

Data Providers



And all your own data
And your partners data

No SQL



Content Management



Data Virtualization COMPOSITE SOFTWARE



BI Tools



Capgemini - Capping IT off
Manuel Sevilla - 2012

Data Governance



BIG DATA & AI LANDSCAPE 2018

INFRASTRUCTURE

HADOOP ON-PREMISE

- cloudera
- hortonworks
- MAPR
- Pivotal
- IBM InfoSphere
- bluedata
- jethro

HADOOP IN THE CLOUD

- aws
- Microsoft Azure
- Google Cloud
- IBM InfoSphere BigInsights
- TREASURE DATA
- Databricks
- Alscale
- CAZENA
- CenturyLink

STREAMING / IN-MEMORY

- aws
- databricks
- strimzi
- confluent
- GridGain
- DATA TORRENT
- dataArtisans
- C4i
- hazelcast
- TERRACOTTA

The collage displays logos for various database technologies, organized into five main categories:

- NoSQL DATABASES:** Includes Google Cloud, AWS, Oracle, Microsoft Azure, mongoDB, MarkLogic, HIEROSPHERE, DATASITE, AmazonDB, Couchbase, redislabs, and SCYLLA.
- NewSQL DATABASES:** Includes SAP HANA, Clustrix, nuodb, Cockroach Labs, MEMSQL, influxdata, Aerospike, TIME SCALE, VOLTE, catusdata, paradigm4, and Trifolium.
- GRAPH DBs:** Includes Neo4j, Amazon Neptune, IBM, and Oracle.
- MPP DBs:** Includes TERADATA, VERTICA, IBM Data Warehouse Systems, Cdata, Kognitio, EXASOL, and dremio.
- CLOUD EDW:** Includes AWS, Google Cloud, Microsoft Azure, Pivotal, snowflake, and Infoworks.

DATA TRANSFORMATION

- talend
- pentaho
- alteryx
- TRIFACTA
- tmr
- Paxata
- StreamSets
- UNIFI

DATA INTEGRATION

- Informatica
- MuleSoft
- inapLogic
- TEALiUM
- Segment
- enigma
- alooma
- ZALONI
- Redwood Data
- Alation
- Waterline
- import.io
- Stitch

DATA GOVERNANCE

- Informatica
- IBM
- ScalePoint
- Nicolas Singh Security Cloud
- collibra
- Alation
- Waterline
- OKERA

MGMT / MONITORING

- AWS
- New Relic
- octoio
- rubrik
- APPROXYNOMICS
- WAVEFORM
- dynatrace
- splunk
- Signalfx
- druid
- Moogsoft
- univention
- pagerduty
- Numery

The collage is organized into six main categories, each with a header and a collection of logos:

- STORAGE**: Includes Google Cloud, Microsoft Azure, Pure Storage, Alluxio, rimmastorage, Ceph, and Cohesity.
- CLUSTER SVCS**: Includes AWS, Kubernetes, Mesos, Docker, and CoreOS.
- APP DEV**: Includes Lightbend, Keen IO, and Rainforest CSK.
- CROWD - SOURCING**: Includes Amazon Mechanical Turk, uWork, figure eight, and HIVE.
- HARDWARE**: Includes Google TPU, ARM, Intel AI, Graphcore, MYTHIC, NVIDIA, Movidius, WAVE, and HALO.
- GPU DBs**: Includes Kinetica, H2O, SAP HANA, Databricks, BLAZINGDB, and bryllyt PG-Stream.

CROSS-INFRASTRUCTURE/ANALYTICS

aws Google Cloud Microsoft IBM SAP Hewlett Packard Enterprise sas 1010DATA vmware TIBCO TERADATA ORACLE NetApp syncsort

ANALYTICS

DATA ANALYST PLATFORMS

- Microsoft
- pentaho
- alteryx
- Digital Reasoning
- guavus
- AYASDI
- ATTIVO
- Datameer
- Quid
- incorta.
- interana
- ClearStory
- Origami logic
- ASSEMBLY
- ENDO3
- MODE
- Bottlenose

DATA SCIENCE PLATFORMS

- IBM
- KNIME
- dataiku
- DOMINO
- rapidminer
- CONTINUUM ANALYTICS
- ALGORITHMIA
- DATAWATCH ANALYTICS

The collage is organized into three vertical sections, each with a header and a collection of logos:

- BI PLATFORMS:** Includes logos for Microsoft, AWS, Domo, Wave Analytics, Looker, Alteryx, ArcGIS Data Analyst, Google Data Studio, and Birst.
- VISUALIZATION:** Includes logos for Tableau, SAP, Google Cloud, Celonis, Qlik, Perficient Data, ZEP, XOOMDATA, CARTIO, and Plotly.
- MACHINE LEARNING:** Includes logos for AWS, Google Cloud, H2O, DataRobot, gamaloon, ELEMENT, VERVISE, VIZENSE, OneDrops, and bonsai.

COMPUTER VISION

- Microsoft Azure
- Amazon Rekognition
- Clarifai
- Cloud Vision API
- EVER AI
- deeppomatic
- twintytwo

HORIZONTAL AI

- Watson
- Cortana
- Face
- Google Assistant
- sentient
- Voyager
- Microsoft
- Affective
- Numenta
- PETuum
- AI
- Curious AI
- OSARO
- Blue Vision

SPEECH & NLP

- Google Cloud
- twilio
- amazon alexa
- narrative
- semantic
- Moovio
- SoundHeard Inc.
- voicebox
- connected to
- snip
- transcribe
- vsync

The collage displays logos for various SaaS companies, organized into four columns:

- SEARCH:** elasticsearch, EXASOL, Lucidworks, swifttype, alphasense, omni-u, SINEQA.
- LOG ANALYTICS:** ORACLE, splunk, sumologic, LOGGLY, ELASTIC, kibana, logz.io.
- SOCIAL ANALYTICS:** Hootsuite, sprinklr, NETBASE, synthesio, track, simplereach, bitly, predata, SimilarWeb.
- WEB / MOBILE / COMMERCE ANALYTICS:** Google Analytics, mixpanel, AMPITUDE, sumal, Airtable, RESCI, SIGOPT, granify, custora.

OPEN SOURCE

The diagram illustrates a comprehensive ecosystem of data science and machine learning tools, organized into 12 functional categories:

- FRAMEWORK**: Includes Hadoop, MapReduce, YARN, Flink, Mesos, Spark, and CDAP.
- QUERY / DATA FLOW**: Includes Spark SQL, Presto, SLAM DATA, and Google Cloud Dataflow.
- DATA ACCESS**: Includes Cassandra, CouchDB, HBase, Nifi, MongoDB, SciDB, and others.
- COORDINATION**: Includes Talend, Apache ZooKeeper, and Apache Ambari.
- STREAMING**: Includes Spark, Flink, Beam, Kafka, and Storm.
- STAT TOOLS**: Includes Python, R, ScalaLab, and SciPy.
- AI / MACHINE LEARNING / DEEP LEARNING**: Includes TensorFlow, Theano, Caffe, Microsoft Cognitive Toolkit, OpenAI, DM TK, Keras, and others.
- SEARCH**: Includes Elasticsearch and Solr.
- LOGGING & MONITORING**: Includes Elasticsearch, Kibana, Sentry, Logstash, and Prometheus.
- VISUALIZATION**: Includes BeakerX, Rodeo, and others.
- COLLABORATION**: Includes Jupyter, Zeppelin, and Anaconda.
- SECURITY**: Includes Apache Ranger, KNOX, and Sentry.

DATA SOURCES & APIs

HEALTH

Apple, VALIDIC, practicefusion, fitbit, GARMIN, HUMAN API, kinsa

IOT

GE Digital, UPTAKE, thingworx, helium, scensara, RASSEMBLY, estimate

FINANCIAL & ECONOMIC DATA

Bloomberg, THOMSON REUTERS, DOW JONES, S&P CAPITAL IQ, CB Insights, xignite, Quandl, INVESTMENT MOOLES, PREMISE, estimate, SECOND MEASURE, Eagle Alpha, StockTwits, PLAID, Thinknum

AIR / SPACE / SEA

Critical Insight, planet, DAYCATCH, Airware, AIRBOTICS, SPIRE, PRECISIONHUB, UNDERSTAY, Decipher, BASIS, telluslabs, DroneDeploy, Manifold

PEOPLE / ENTITIES

ackl0m, experian, EPSILON, InsideView, Crimson Hexagon, Quantcast, MANTAS, SAFE GRAPH

LOCATION INTELLIGENCE

FOURSQUARE, Mapquest, sense360, PlaceIQ, esri, factual, CARTO, Mapillary, Streetline, cuebiq

OTHER

qualtrics, DATA.GOV, data.world, enigma, mobilewalla

DATA RESOURCES

DATA SERVICES

- Palantir
- OPERA
- DATA SCIENCE
- fractal6
- kaggle
- EXL
- DataKind

INCUBATORS & SCHOOLS

- PLURALISIGHT
- GA
- galvanize
- DataCamp
- DataElite
- INSIGHT
- The Data Incubator
- METIS

RESEARCH

- facebook research
- OpenAI
- MIRA
- MIRA MIRI
- VECTOR INSTITUTE
- CSAI
- Qi
- AIZ
- AI2
- AI2 INSTITUTE

“Big Data” são dados e processos cuja escala, distribuição, diversidade e/ou velocidade de criação requer o uso de novas tecnologias de armazenamento e análise para permitir a captura do valor inserido nos mesmos (adaptado de EMC, 2013)

1. Volume de dados

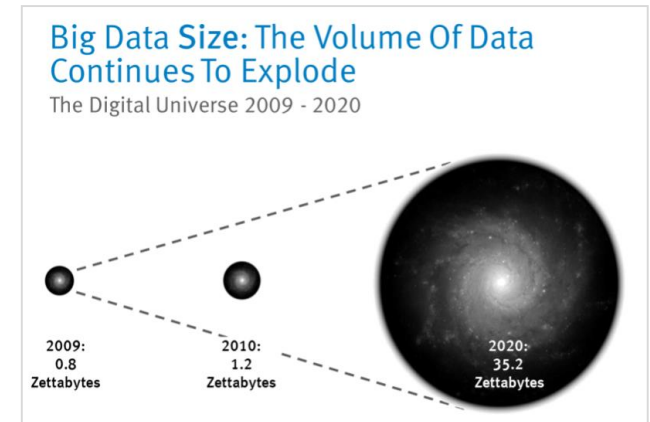
- ▶ Bilhões de linhas x bilhões de colunas
- ▶ Aumento de 44x de 2009 a 2020 (0,9ZB a 35ZB)

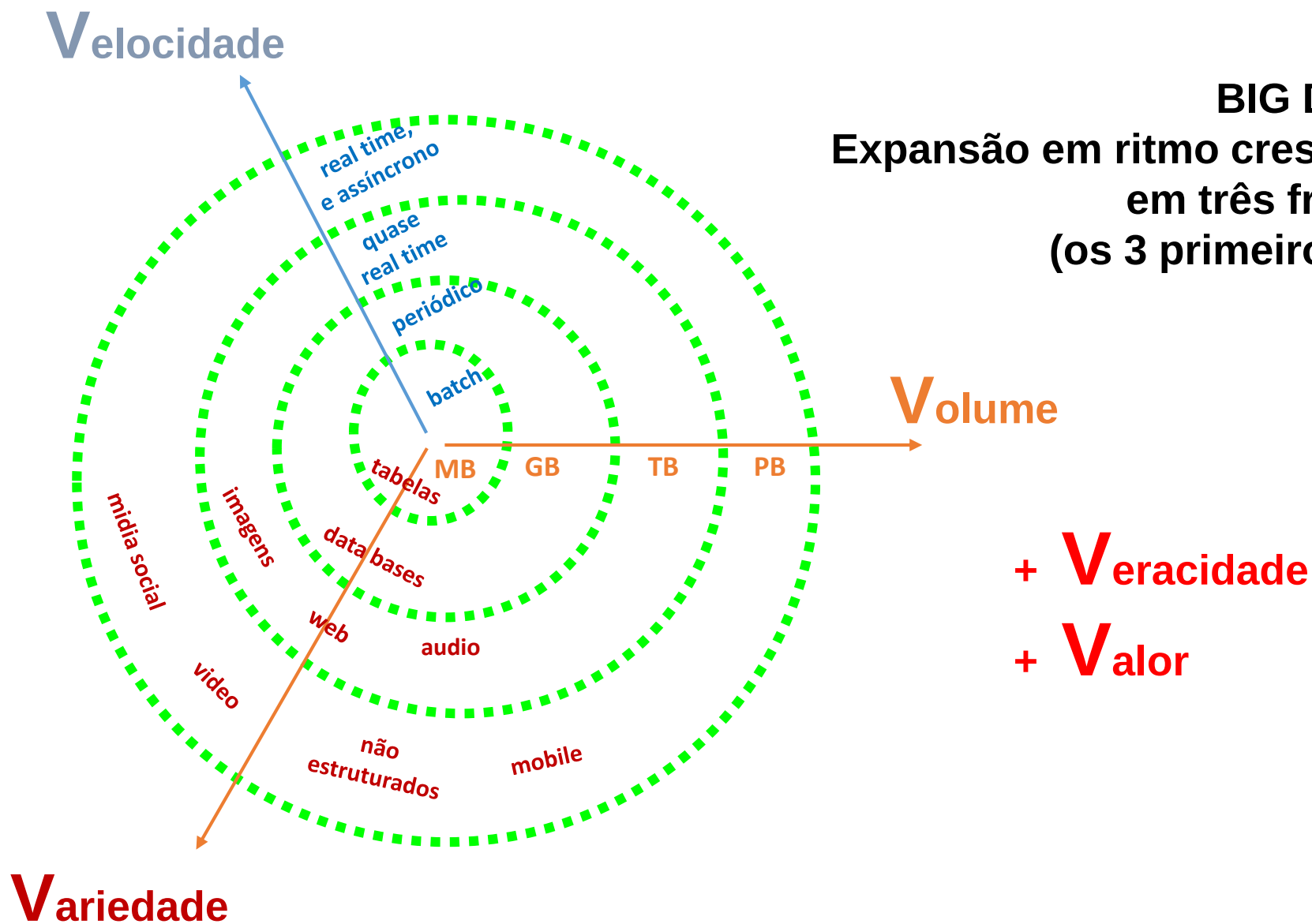
2. Complexidade de Processamento

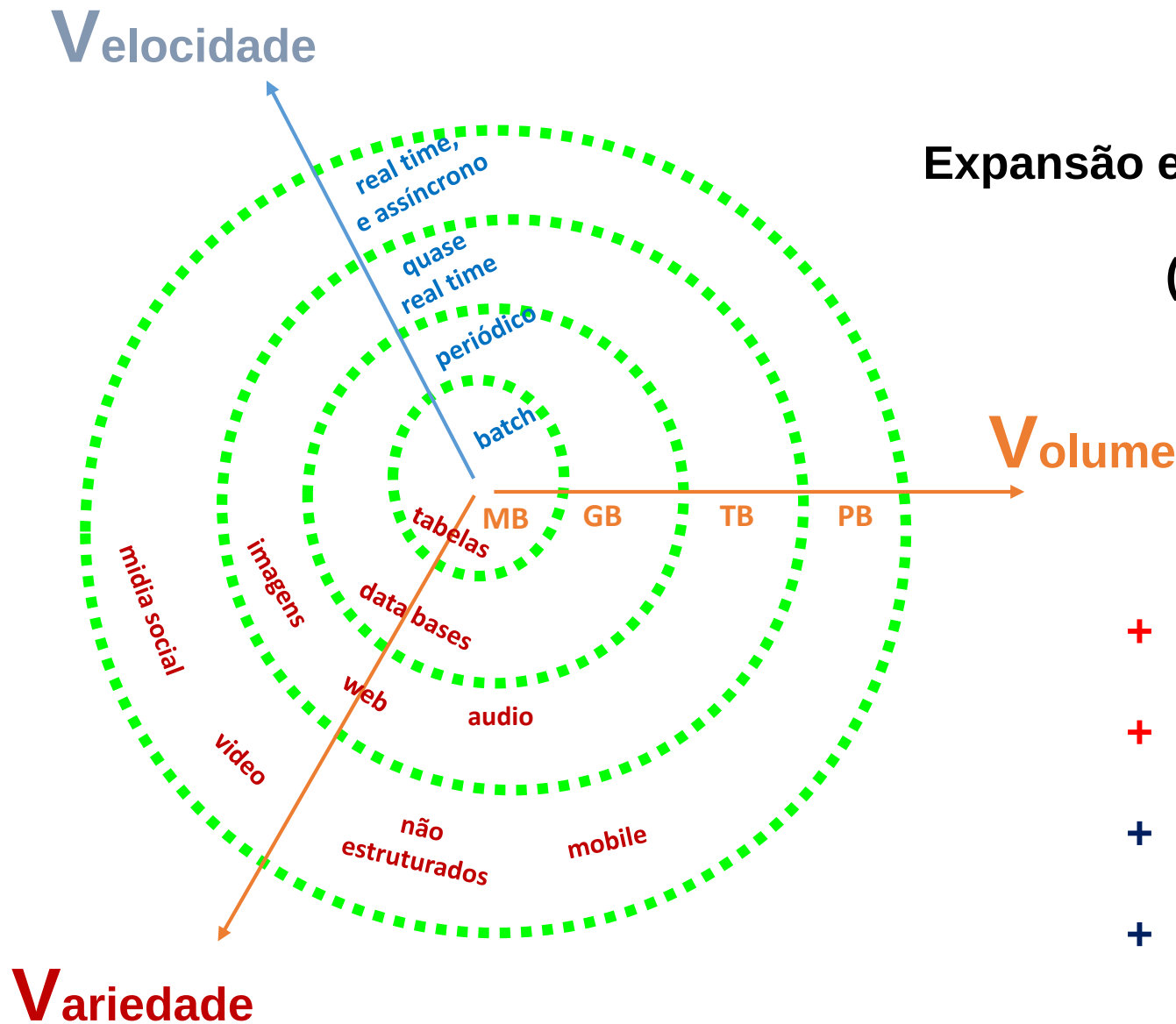
- ▶ Estruturas de dados em constante mudança
- ▶ Necessidade de analisar tais dados em tempo real

3. Estruturas de dados

- ▶ Grande variedade (80-90% não estruturada) a ser analisada
- ▶ Estas características tornam necessário o uso de sistemas de computação paralela e paralela massiva (MPP, massively parallel processing)







BIG DATA:
Expansão em ritmo crescente
em três frentes
(os 3 primeiros Vs)

- + **V**eracidade
- + **V**alor
- + **V**ariabilidade
- + **V**isualização

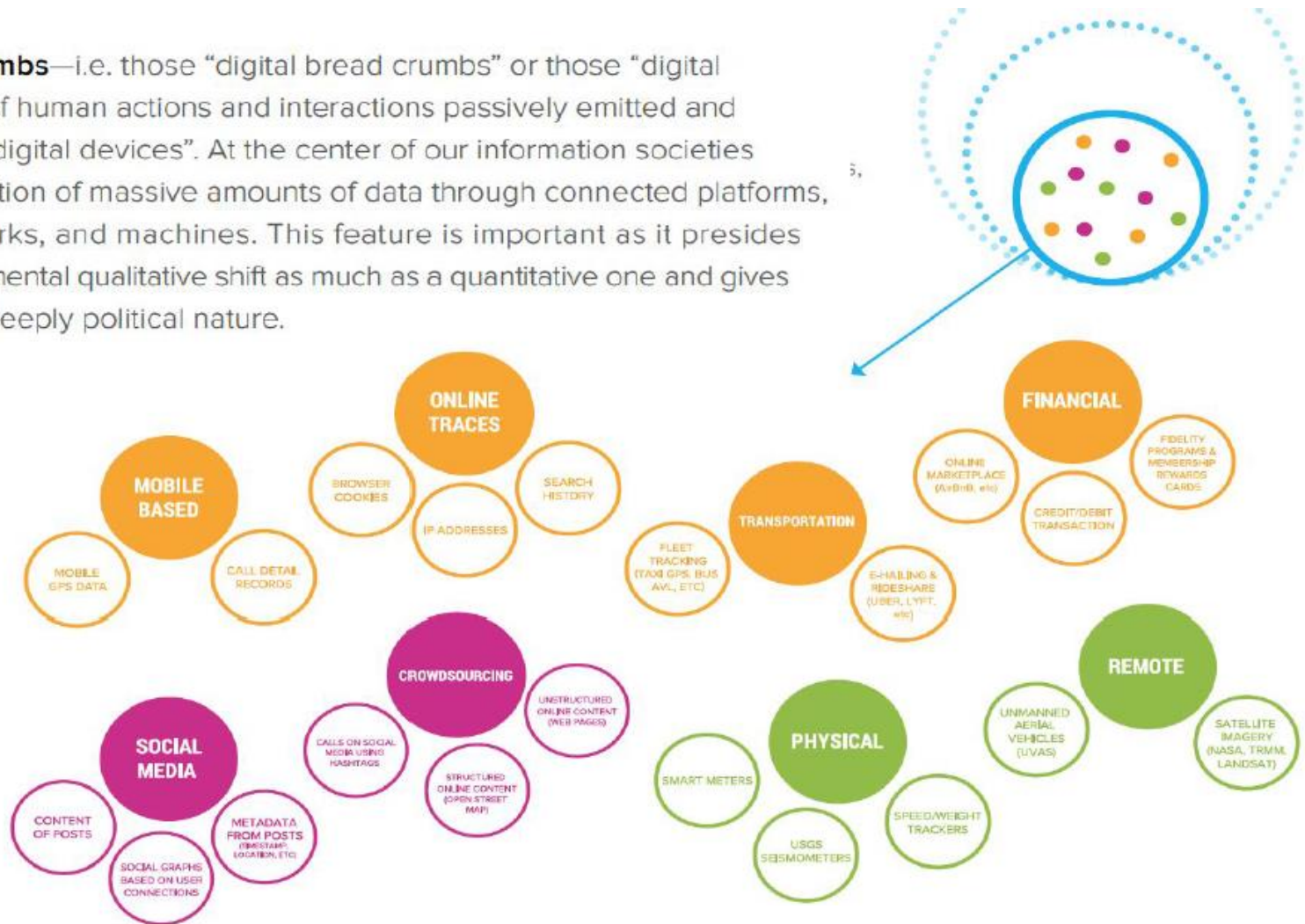
Uma expansão/revisão no posicionamento



Fonte: LETOUZÉ, 2017

Uma expansão/revisão no posicionamento

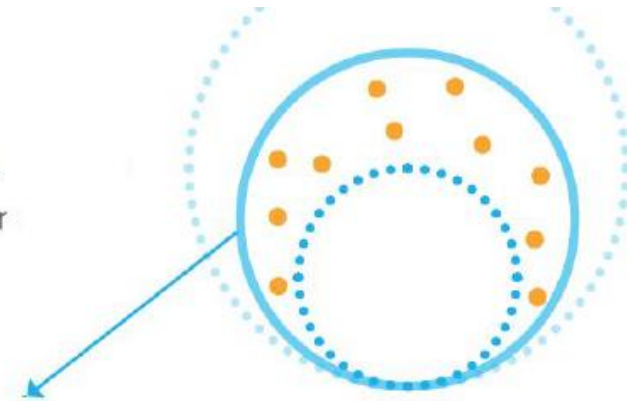
The **C of crumbs**—i.e. those “digital bread crumbs” or those “digital translations of human actions and interactions passively emitted and captured by digital devices”. At the center of our information societies is the production of massive amounts of data through connected platforms, social networks, and machines. This feature is important as it presides over a fundamental qualitative shift as much as a quantitative one and gives Big Data its deeply political nature.



Fonte: LETOUZÉ, 2017

Uma expansão/revisão no posicionamento

The **C of capacities**—i.e. tools and methods to collect, aggregate and analyze data. Algorithms—to be defined and discussed below—fall squarely under capacities, and stand firmly at the center of this ecosystem, as both products and drivers of its expansion. Parallel computing is another key aspect without which Big Data would not exist as a techno-social phenomenon as it allows making computations in a fraction of the time—sometimes years—it would take to run them on one machine.



NETWORK
ANALYSIS
AND SOCIAL
PHYSICS

MACHINE
LEARNING
TECHNIQUES
AND ARTIFICIAL
INTELLIGENCE

TEXT MINING
TOOLS
(SENTIMENT
ANALYSIS, TOPIC
MODELS, ETC)

SPATIAL
ANALYSIS AND
GEOGRAPHIC
INFORMATION
SYSTEMS (GIS)

CROWDSOURCING

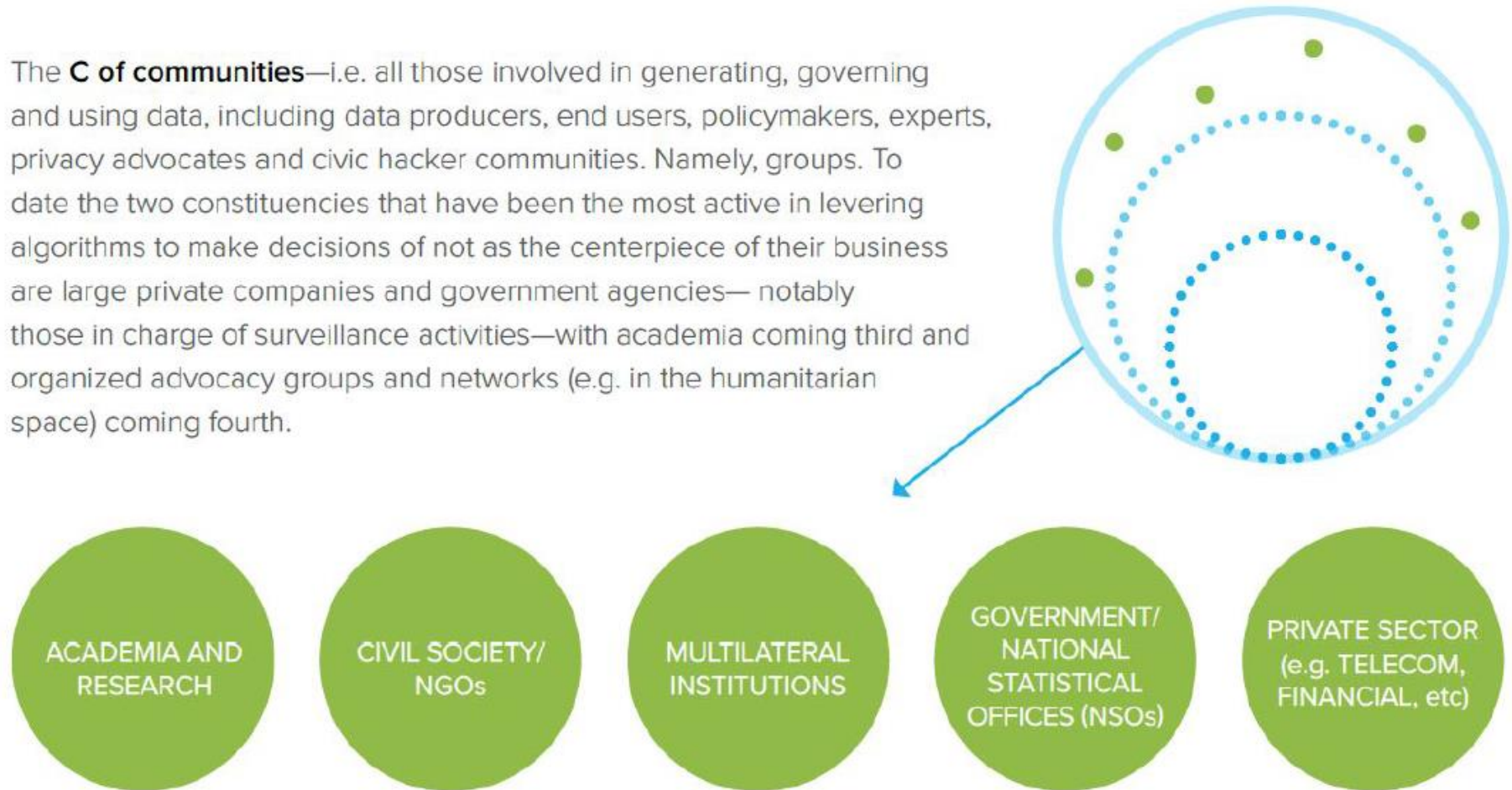
DATASET
DISAGGREGATION
AND UNIFICATION

DISTRIBUTED
NETWORKS OF
DEVICES

CLOUD
COMPUTING
AND STORAGE

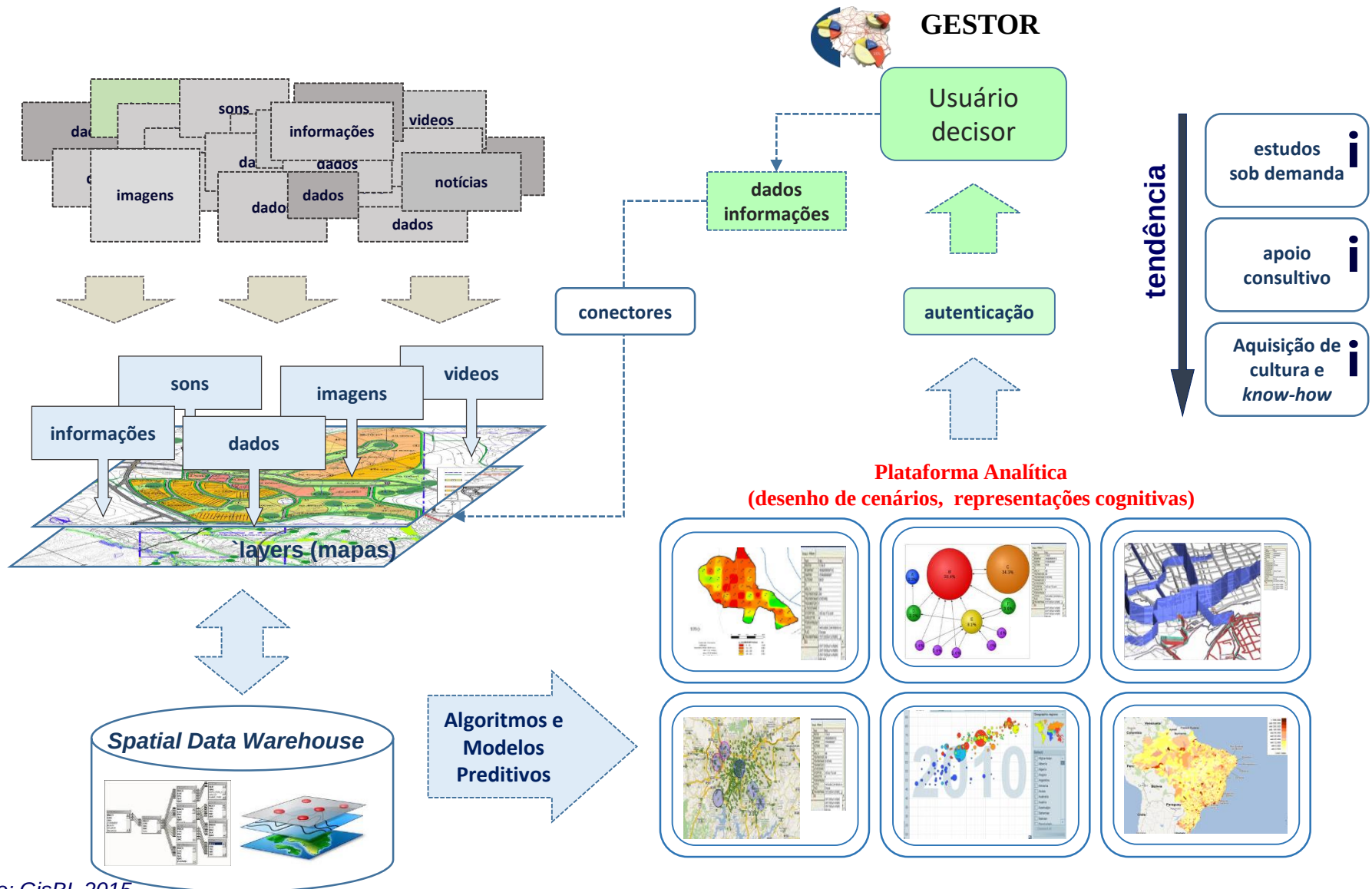
Uma expansão/revisão no posicionamento

The **C of communities**—i.e. all those involved in generating, governing and using data, including data producers, end users, policymakers, experts, privacy advocates and civic hacker communities. Namely, groups. To date the two constituencies that have been the most active in leveraging algorithms to make decisions of not as the centerpiece of their business are large private companies and government agencies— notably those in charge of surveillance activities—with academia coming third and organized advocacy groups and networks (e.g. in the humanitarian space) coming fourth.



Fonte: LETOUZÉ, 2017

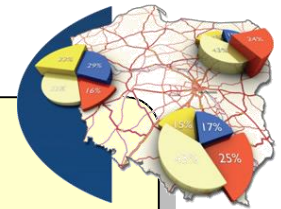
Framework: Business Intelligence Adaptativo



Fonte: GisBI, 2015

◆ Inteligência Geográfica está por aí?

Aproximadamente 70 a 80% das informações relevantes nos processos decisórios têm caracterização espacial (Gartner Group, 2004)



- A geografia é o que possibilita a integração de dados, informações, processos, inclusive “Big Data”
- Essa integração induz uma visão sistêmica (integrada, ampla, abrangente, holística) da maioria das questões necessárias às tomadas de decisão



Geomarketing e GIS: Suporte à Operação

- Descrição, Expansão, Segmentação e Otimização do Território de Atuação



Estatística Espacial: Geoinformação em Modelos Preditivos

- Incorporação da Influência Geográfica nos Modelos de Predição de Renda e de Risco de Crédito
- Aprimoramento do poder de explicação dos Modelos Preditivos

Indicadores Econômicos baseados em Consumo de Energia Elétrica



Renda e Caracterização Socioeconômica

• Renda

- **Indicador tradicional adotado em estudos sobre Condições de Vida, Pobreza e Mercado**
- **Dificuldade em conseguir informações precisas sobre essa variável** (BUSSAB; FERREIRA, 1999)
 - Subdeclaração, superdeclaração, esquecimento, sazonalidade, recusa, etc.
- **Mensurar a Pobreza é desafio extremamente relevante nos dias de hoje** (PRAHALAD, 2005; PÉREZ; PÉREZ, 2008)



• Caracterização (Socio)econômica ou Poder de Consumo através de indicadores

- **Posse de Bens duráveis e Instrução do Chefe de Família**
- **Estoque de bens possuídos indica o nível de conforto alcançado pela família ao longo do tempo**

➤ ***Critério de Classificação Econômica Brasil***

♦ Brasil

- ♦ Critério ABA (1970), ABA-ABIPEME (1982), Proposta de Almeida e Wickerhauser (1991)

♦ CCEB – Critério de Classificação Econômica Brasil

- ♦ Criado pela ANEP em 1996 e mantido pela ABEP a partir de 2004
- ♦ Estima o poder de compra de pessoas e famílias urbanas
- ♦ Classes econômicas a partir de sistema de pontuação

Posse de Itens (Bens duráveis) e Empregada mensalista

Itens	Quantidade de itens				
	0	1	2	3	4 ou +
Televisão em cores	0	1	2	3	4
Rádio	0	1	2	3	4
Banheiro	0	4	5	6	7
Automóvel	0	4	7	9	9
Empregada mensalista	0	2	4	4	4
Aspirador de pó	0	0	0	0	0
Máquina de lavar	0	2	2	2	2
Videocassete e/ou DVD	0	2	2	2	2
Geladeira	0	4	4	4	4
Freezer (aparelho independente ou parte da geladeira duplex)	0	2	2	2	2



Classe	Pontos
A1	42-46
A2	35-41
B1	29-34
B2	23-28
C1	18-22
C2	14-17
D	8-13
E	0-7

Grau de Instrução do chefe de família

Analfabeto / Primário incompleto	Analfabeto / Até 3ª. Série Fundamental	0
Primário completo / Ginásial incompleto	Até 4ª. Série Fundamental	1
Ginásial completo / Colegial incompleto	Fundamental completo	2
Colegial completo / Superior incompleto	Médio completo	4
Superior completo	Superior completo	8



INTRODUÇÃO

MÉTODOS

RESULTADOS

CONCLUSÃO

- **Brasil**

- Critério ABA (1970), ABA-ABIPEME (1982), Proposta de Almeida e Wickerhauser (1991)

- **CCEB – Critério de Classificação Econômica Brasil**

- Criado pela ANEP em 1996 e mantido pela ABEP a partir de 2004
- Estima o poder de compra de pessoas e famílias urbanas
- Classes econômicas a partir de sistema de pontuação

Utilização de variáveis e indicadores que não têm estabilidade ao longo do tempo e que são pouco discriminadores dos estratos da população (PEREIRA, 2004)

Não adequado para a caracterização de famílias posicionadas nos extremos da distribuição de renda – em particular, para o consumidor de Baixa Renda (MATTAR, 1996; SILVA, 2004)

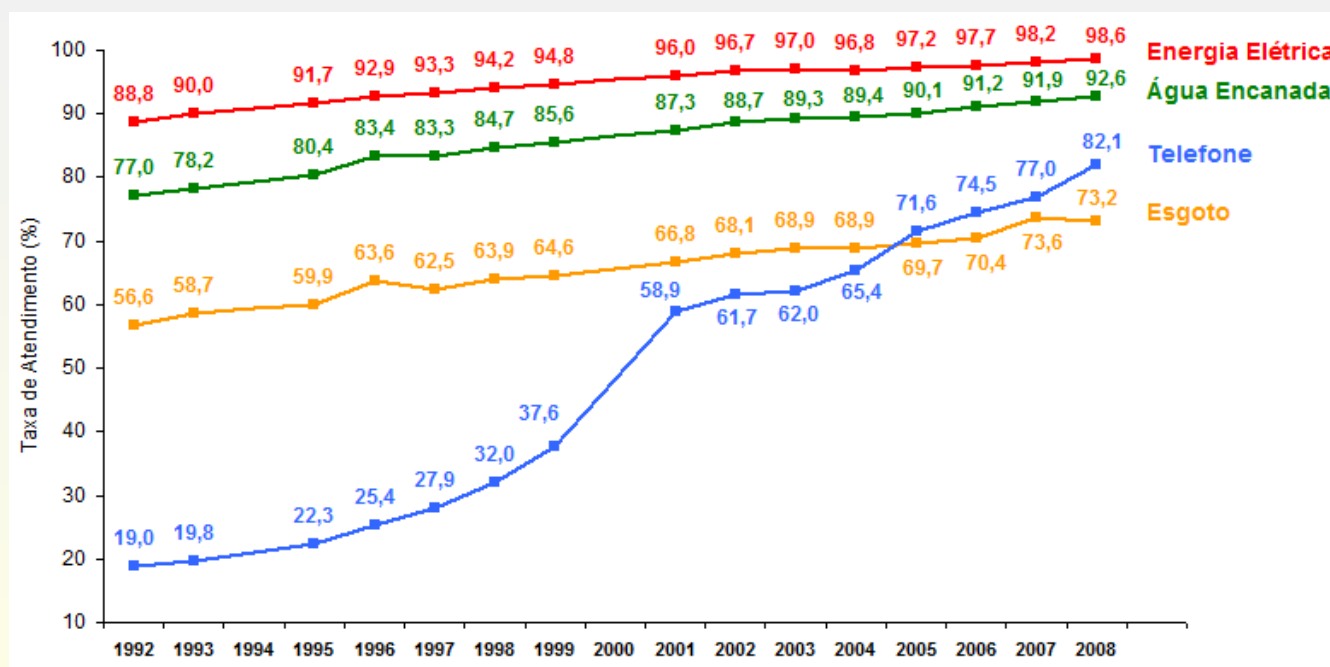
Estudos aprofundados necessitam de especializações e adaptações do Critério Brasil: Indicadores de consumo de grande abrangência, capilaridade e que não necessitem de constante atualização podem ser úteis

Viés: Famílias mais Velhas têm mais condições de ascender economicamente do que as Famílias recém-configuradas

Consumo de Energia Elétrica

O Consumo de Energia Elétrica pode ser um bom indicador para auxiliar processo de caracterização do consumidor

- Serviço Essencial
- Abrangência e Cobertura
 - **98,6%** dos domicílios brasileiros (**99,82%** na área urbana, **99,97%** na Região SE)



Consumo de Energia Elétrica

O Consumo de Energia Elétrica pode ser um bom indicador para auxiliar processo de caracterização do consumidor

- ♦ **Serviço Essencial**
- ♦ **Abrangência e Cobertura**
 - ♦ **98,6%** dos domicílios brasileiros (**99,82%** na área urbana, **99,97%** na Região SE)
- ♦ **Grande Capilaridade**
 - ♦ Maior que a de outras utilidades (água, telefone, gás)
 - ♦ Clientes de A a E
- ♦ **Precisão e Temporalidade**
 - ♦ Endereço, localização geográfica dos consumidores
 - ♦ Coletado mensalmente
 - ♦ Histórico de faturamento e arrecadação
- ♦ **Exerce papel fundamental no dia-a-dia dos domicílios residenciais**

Melhor caracterização das famílias-alvo (em termos socioeconômicos e de potencial de consumo)

INTRODUÇÃO

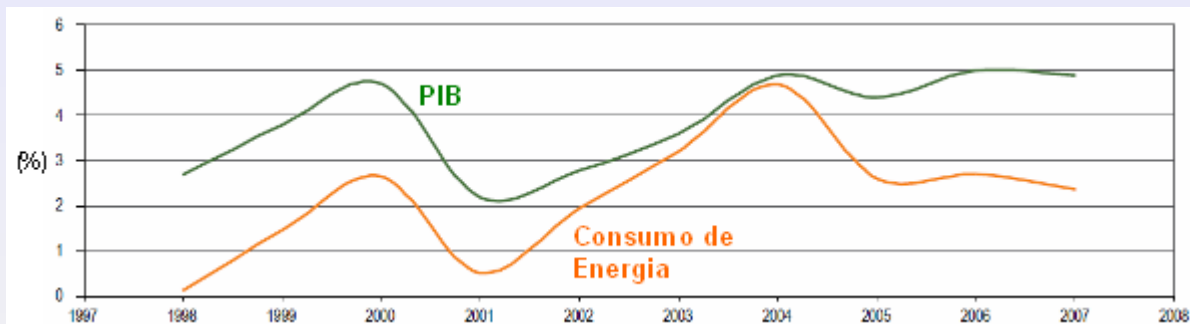
MÉTODOS

RESULTADOS

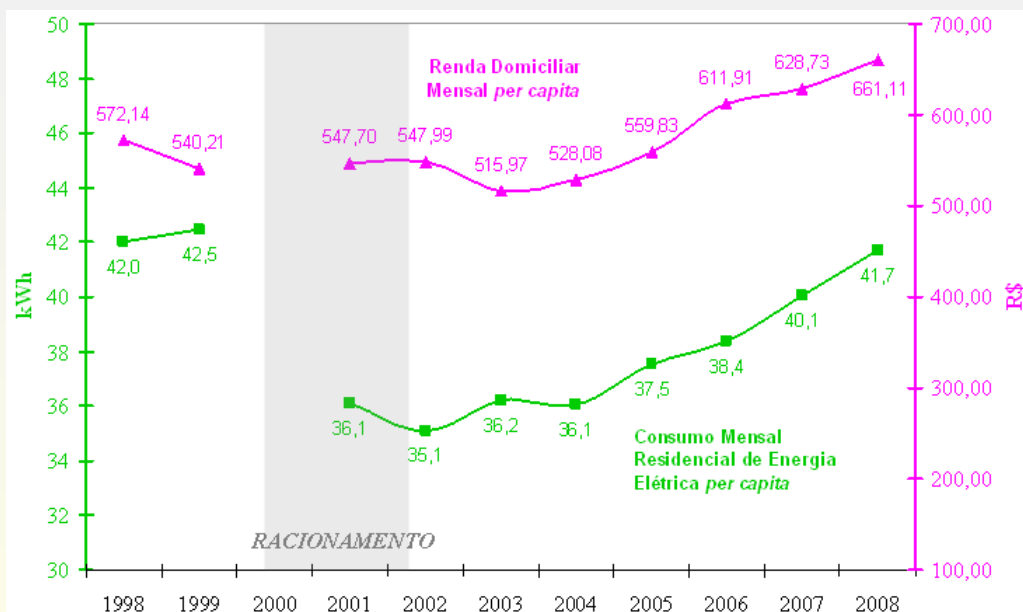
CONCLUSÃO

Consumo de Energia Elétrica

Mundo: Variação do PIB × Variação Consumo de Energia



Brasil: Renda Domiciliar *per capita* × Consumo Residencial de Energia Elétrica *per capita*



Muitos estudos mostram alta aderência do Consumo Total ou do Consumo Industrial com o PIB, com Renda e com Qualidade de Vida

INTRODUÇÃO

MÉTODOS

RESULTADOS

CONCLUSÃO

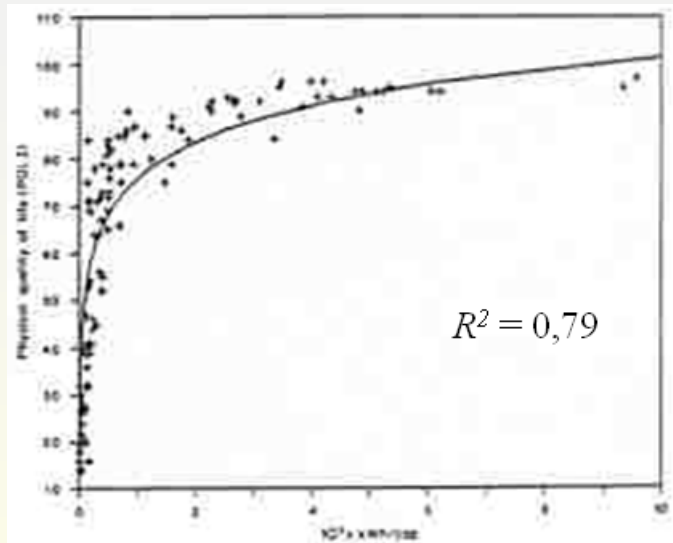
Consumo de Energia Elétrica

♦ SPI – Sinalizador de Produção Industrial

- ♦ Indicador econômico mensal (AES Eletropaulo/FGV) que projeta a evolução da produção física mensal da Indústria Paulista com um mês de antecedência

♦ *Physical Quality of Life Indicator (PQLI) (MARTIN, 1980)*

- ♦ 3 Dimensões: Expectativa média de Vida, Mortalidade Infantil e Taxa de Alfabetização (precursor do IDH)
- ♦ Associação de 79% (R^2) com log do Consumo de Energia *per capita* – análise de dados de 112 países (ALAM et al., 1991)



Consumo de Energia Elétrica

♦ Índice Paulista de Responsabilidade Social (IPRS)

- ♦ Indicador municipal desenvolvido pela Fundação SEADE para o Estado de SP
- ♦ 3 Dimensões: Riqueza, Educação e Longevidade
- ♦ Tipologia (segmentação) de 5 grupos

Dimensão	Componentes (Variáveis)	Contribuição para o Indicador Sintético
Riqueza	Riqueza do Município	Consumo de energia elétrica na agricultura, no comércio e nos serviços
		23,71%
		Valor adicionado fiscal <i>per capita</i>
		13,90%
Longevidade	Renda Familiar	Consumo residencial de energia elétrica
		43,51%
		Remuneração média dos empregados com carteira assinada e do setor público
		19,42%
Escolaridade		Mortalidade perinatal
		30,00%
		Mortalidade infantil
		30,00%
		Mortalidade de adultos de 15 a 39 anos
		20,00%
		Mortalidade de adultos de 60 anos ou mais
		20,00%
		% de jovens de 15 a 17 anos que concluíram o ensino fundamental
		36,08%
		% de jovens de 15 a 17 anos com pelo menos quatro anos de escolaridade
		8,38%
		% de jovens de 18 a 19 anos que concluíram o ensino médio
		35,37%
		% de crianças de 5 e 6 anos que freqüentam a pré-escola
		20,17%

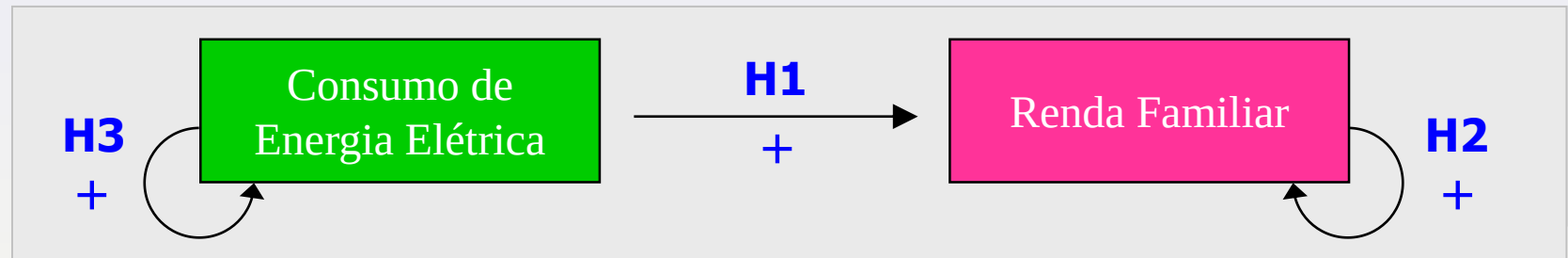
the information gap on financial risk for millions of Americans. More concretely, considering that many of these millions outside the credit mainstream are poorer, less advantaged Americans, the information can direct markets toward a faster alleviation of poverty in this country.



Objetivos e Objeto de Estudo

- Examinar a Relação entre Consumo de Energia Elétrica e Renda Domiciliar
- Propor a construção de um indicador de Classificação Econômica dos domicílios baseado em Consumo de Energia Elétrica, de forma a estender ou substituir o Critério Brasil

Objeto: Município de São Paulo

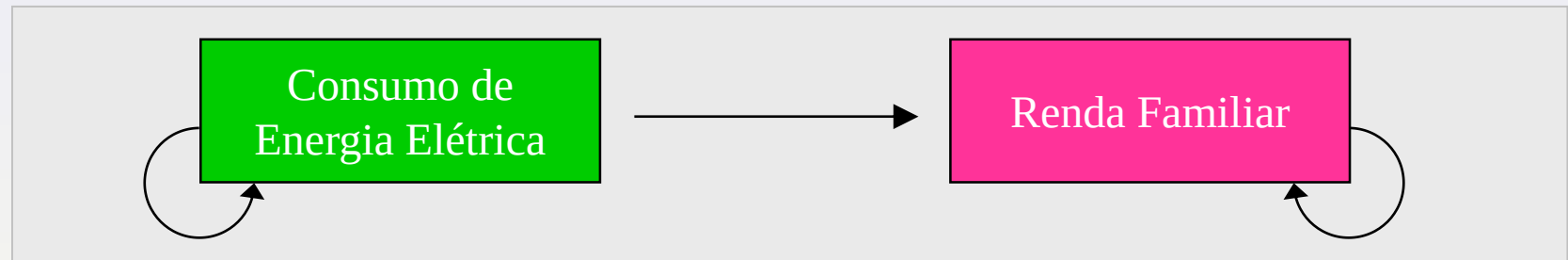


- **H1:** Quanto maior o Consumo de Energia Elétrica residencial, maior a Renda Domiciliar (ou Familiar)
- **H2:** Existe padrão de dependência espacial da variável Renda Domiciliar entre regiões de São Paulo, com renda decrescente no sentido Centro-Periferia
- **H3:** Existe padrão de dependência espacial da variável Consumo de Energia Elétrica entre regiões de São Paulo, com consumo de energia decrescente no sentido Centro-Periferia

Objetivos e Objeto de Estudo

- Examinar a Relação entre Consumo de Energia Elétrica e Renda Domiciliar
- Propor a construção de um indicador de Classificação Econômica dos domicílios baseado em Consumo de Energia Elétrica, de forma a estender ou substituir o Critério Brasil

Objeto: Município de São Paulo



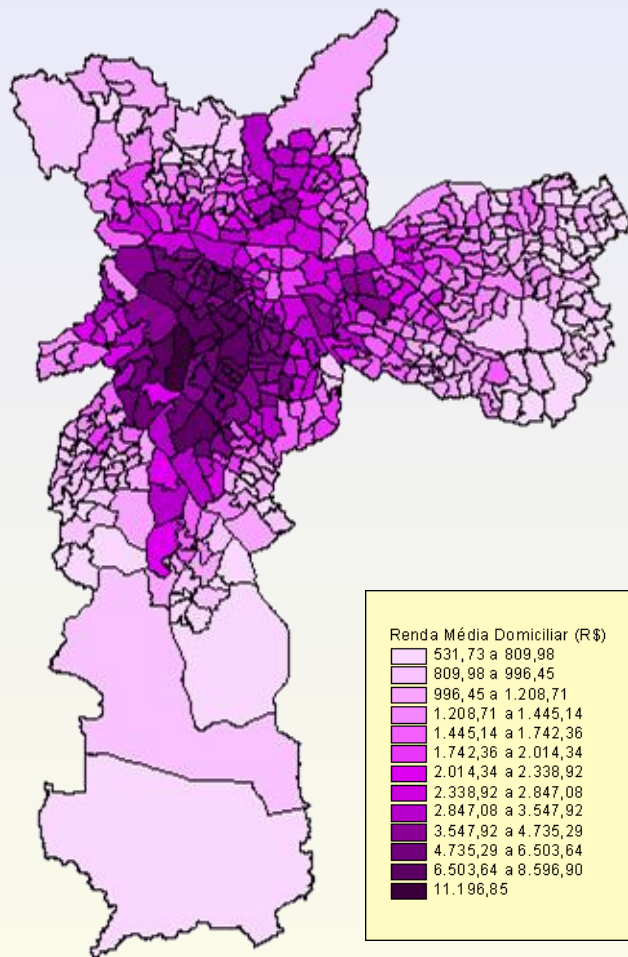
2 Níveis de Investigação :



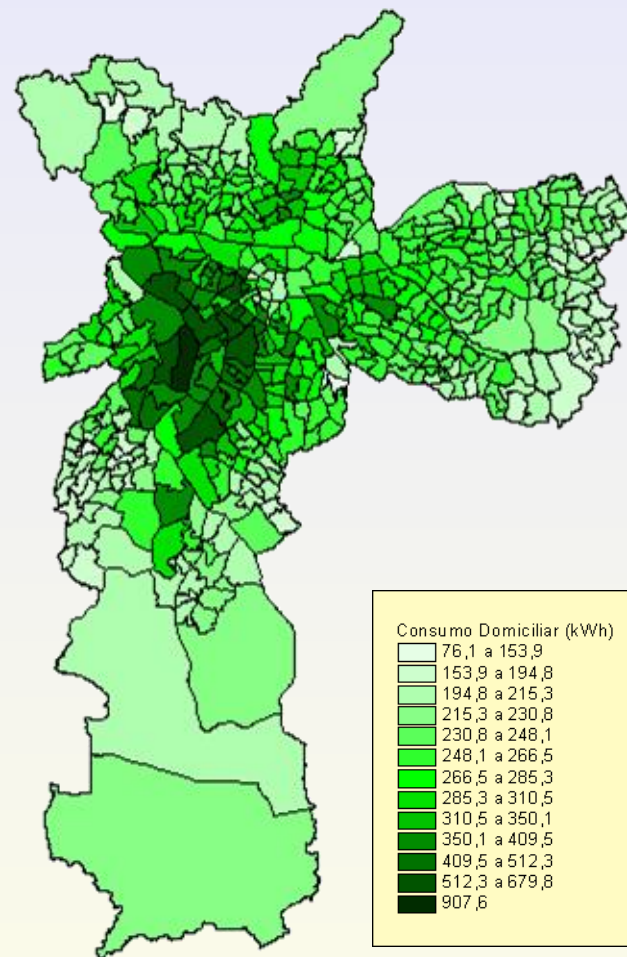
Contribuições Metodológicas:

- Aprimoramento da Avaliação da Relação através do uso de Modelos Espaciais
- Refinamento dos Modelos Espaciais para dados pontuais através de técnicas de espalhamento de pontos no interior dos polígonos

- Mapas das 456 Áreas de Ponderação do município de São Paulo



Renda Domiciliar



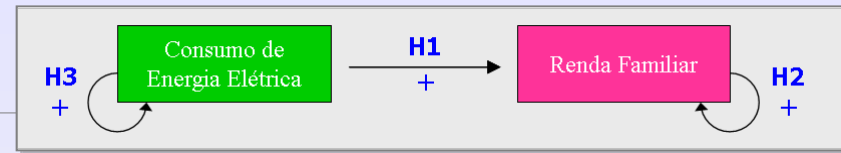
Consumo de
Energia Elétrica

INTRODUÇÃO

MÉTODOS

RESULTADOS

CONCLUSÃO



INVESTIGAÇÃO TERRITORIAL

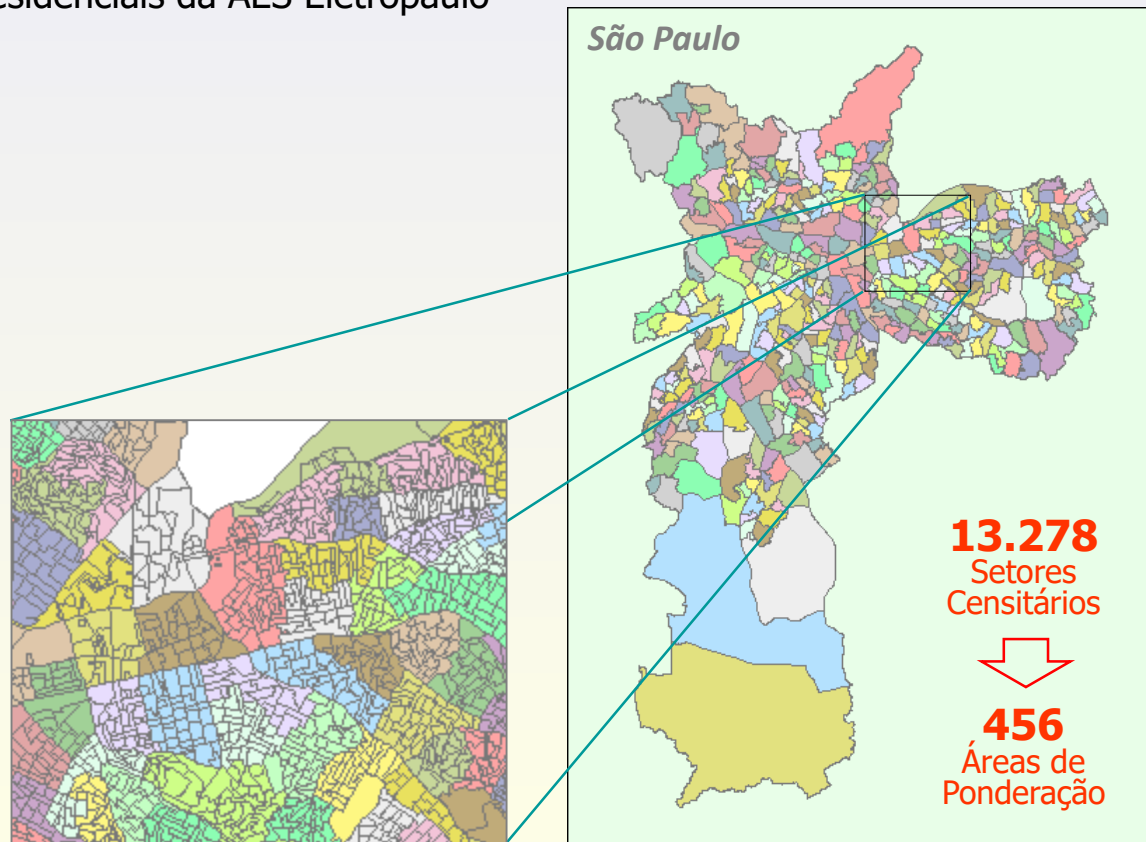
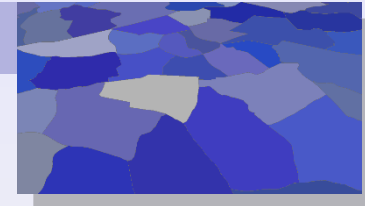
Censo Demográfico 2000 + Consumo de Energia

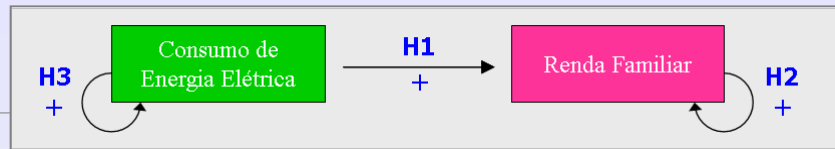
(IBGE)

(AES Eletropaulo)

- Unidade de Análise: **Área de Ponderação**

- 303.669 domicílios amostrados (representando **3.032.095**)
- **3.037.992** clientes residenciais da AES Eletropaulo





INVESTIGAÇÃO TERRITORIAL

Censo Demográfico 2000 + Consumo de Energia

(IBGE)

(AES Eletropaulo)

• Unidade de Análise: **Área de Ponderação**

- 303.669 domicílios amostrados (representando **3.032.095**)
- **3.037.992** clientes residenciais da AES Eletropaulo

• Construção do Critério Brasil Adaptado ao Censo Demográfico 2000 (CBA)

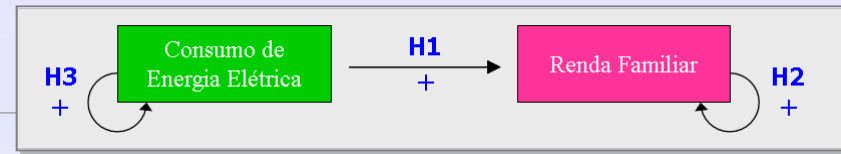
Critério Brasil *

Itens	Quantidade de itens				
	0	1	2	3	4 ou +
Televisão em cores	0	2	3	4	5
Rádio	0	1	2	3	4
Banheiro	0	2	3	4	4
Automóvel	0	2	4	5	5
Empregada mensalista	0	2	4	4	4
Aspirador de pó	0	1	1	1	1
Máquina de lavar	0	1	1	1	1
Videocassete e/ou DVD	0	2	2	2	2
Geladeira	0	2	2	2	2
Freezer (aparelho independente ou parte da geladeira duplex)	0	1	1	1	1

Critério Brasil Adaptado

Itens	Quantidade de itens				
	0	1	2	3	4 ou +
Televisão em cores	0	2	3	4	5
Rádio	0	1	1	1	1
Banheiro + Sanitário	0	2	3	4	4
Automóvel	0	2	4	5	5
Empregada doméstica	0	2	4	4	4
Aspirador de pó	0	0	0	0	0
Máquina de lavar	0	1	1	1	1
Videocassete e/ou DVD	0	2	2	2	2
Geladeira ou Freezer	0	2	2	2	2

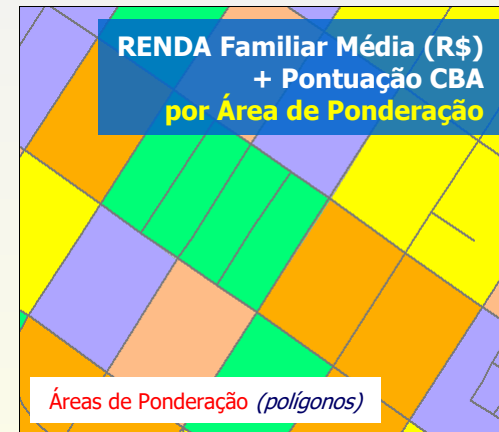
*: Critério Brasil em vigor em 2000

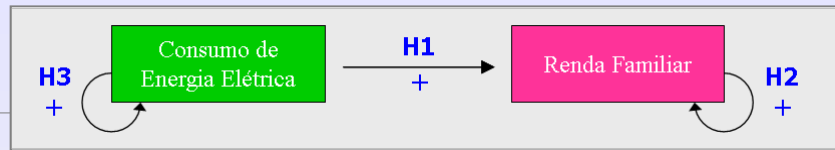


INVESTIGAÇÃO TERRITORIAL

Censo Demográfico 2000 + Consumo de Energia
(IBGE) (AES Eletropaulo)

- Unidade de Análise: **Área de Ponderação**
 - 303.669 domicílios amostrados (representando **3.032.095**)
 - **3.037.992** clientes residenciais da AES Eletropaulo
- Construção do Critério Brasil Adaptado ao Censo Demográfico 2000 (CBA)
- Sobreposição geográfica e Junção Espacial



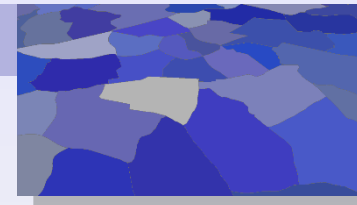


INVESTIGAÇÃO TERRITORIAL

Censo Demográfico 2000 + Consumo de Energia

(IBGE)

(AES Eletropaulo)



• Unidade de Análise: **Área de Ponderação**

- 303.669 domicílios amostrados (representando **3.032.095**)
- **3.037.992** clientes residenciais da AES Eletropaulo
- Construção do Critério Brasil Adaptado ao Censo Demográfico 2000 (CBA)
- Sobreposição geográfica e Junção Espacial
- Técnicas de Análise

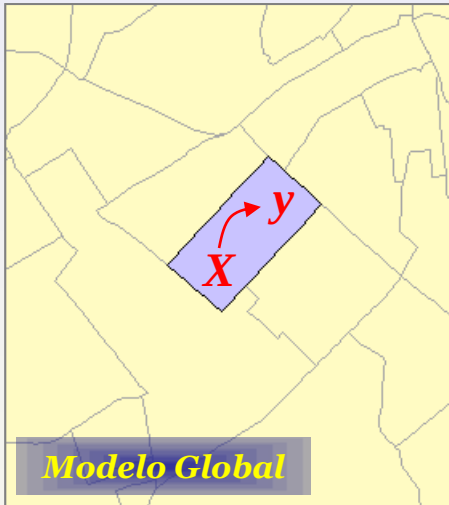
Proposição	Técnica de Análise
H1	Predição da Renda através de OLS, SAR, GWR e GWR+SAR (16 regressões)
H2	Auto-correlação espacial (<i>I</i> de Moran) e LISA maps da Renda Familiar a partir de diversas matrizes de vizinhança
H3	Auto-correlação espacial (<i>I</i> de Moran) e LISA maps do Consumo de Energia Elétrica a partir de diversas matrizes de vizinhança

Estatística Espacial



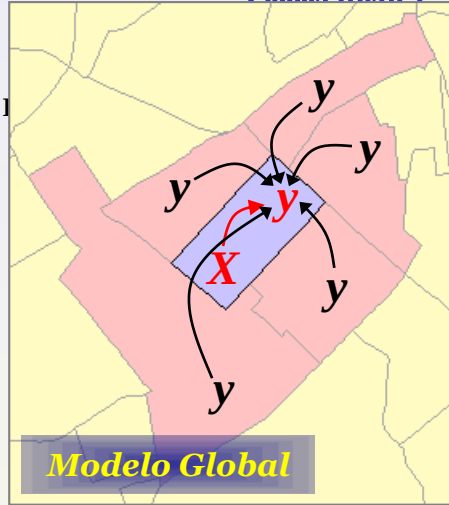
Modelo OLS
(Regressão Linear Tradicional)

$$y = X\beta + \varepsilon$$



Modelo SAR
(Spatial Auto-Regressive)

$$y = \rho W y + X\beta + \varepsilon$$

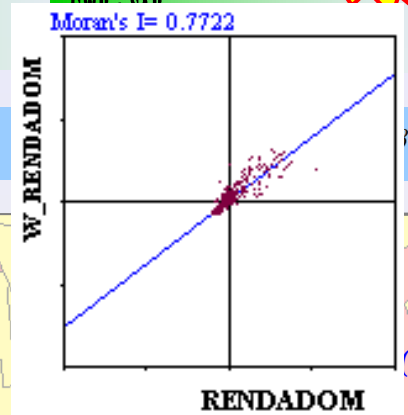
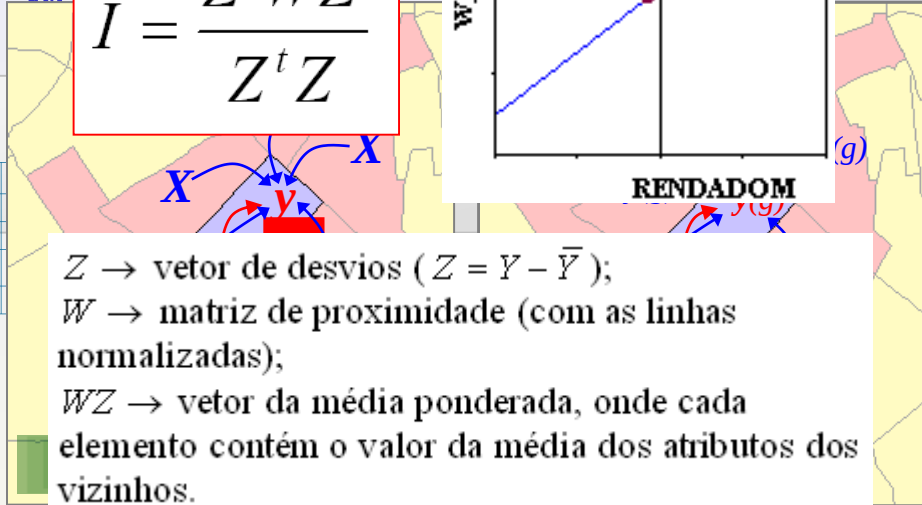


Não capturam a dependência espacial do fenômeno

Modelo GWR
(Geographic Weighted Regression)

I de Moran

$$I = \frac{Z^t W Z}{Z^t Z}$$



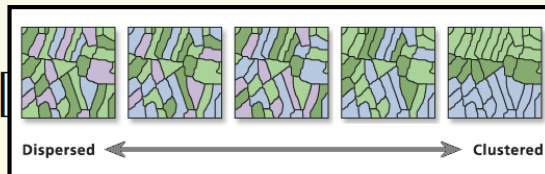
$Z \rightarrow$ vetor de desvios ($Z = Y - \bar{Y}$);
 $W \rightarrow$ matriz de proximidade (com as linhas normalizadas);
 $WZ \rightarrow$ vetor da média ponderada, onde cada elemento contém o valor da média dos atributos dos vizinhos.

- Indicador de Auto-Correlação Espacial

Para pesos diferentes nos elementos da amostra: **LISA Maps**: Mapas de Auto-Correlação Local

Modelo WLM (Weighted Linear Model)

$$y = qX\beta + \varepsilon, \text{ com } q =$$



Minimização por Akaike

Minimização por Cross-Validation (CV)

AIC ou CV para a determinação da Amostra Local (próprio de GWR)

+

Termo Espacial Auto-regressivo (SAR) apropriado nas amostras locais

Indicadores de Renda baseados em Consumo de Energia Elétrica

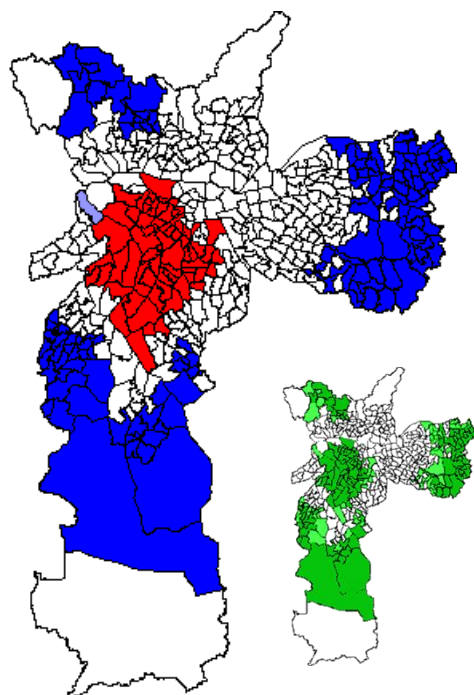
RESULTADOS – INVESTIGAÇÃO TERRITORIAL

- Auto-Correlação Espacial dos Construtos

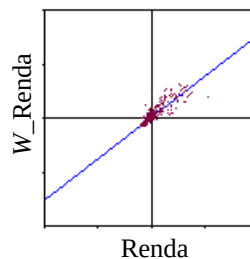
- I de Moran da Renda $> 0,77$ e I de Moran do Consumo de Energia $> 0,58$
na maioria das matrizes de vizinhança adotadas
- LISA maps : alta renda no centro, áreas de transição no entorno e heterogeneidade das baixas rendas nas regiões periféricas
- Quanto maior o número médio de vizinhos por área, menor a magnitude do I de Moran, maior a significância das áreas Baixo-Baixo (principalmente) e Alto-Alto e menores as áreas de transição de regimes

W: Contiguidade de 1ª ordem do tipo Queen

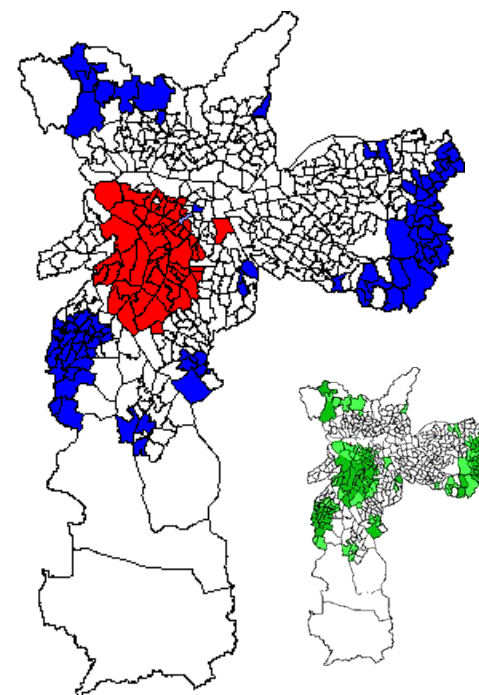
Renda Domiciliar



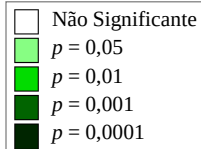
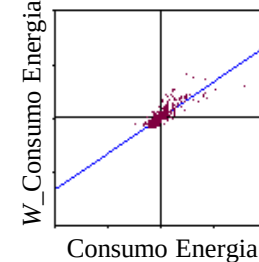
$I = 0,7722$



Consumo de Energia Elétrica



$I = 0,6661$



Indicadores de Renda baseados em Consumo de Energia Elétrica

RESULTADOS – INVESTIGAÇÃO TERRITORIAL



OLS

SAR

Adjacência de 1ª ordem

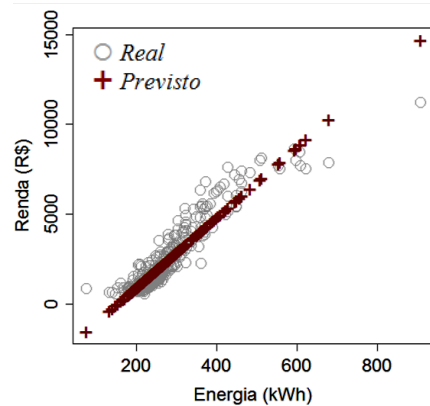
GWR

Kernel adaptável gaussiano – $k=9$

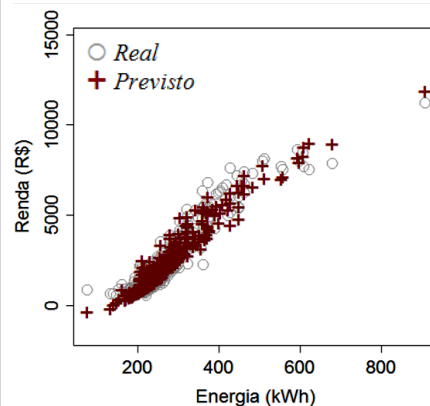
GWR+SAR

Kernel adaptável gaussiano – $k=9$

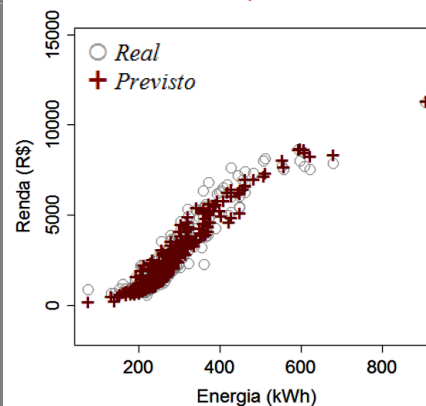
$R^2 = 86,80\%$



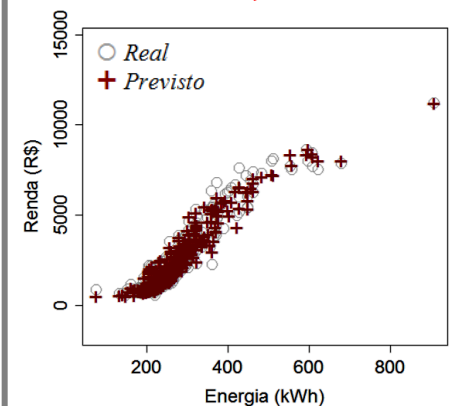
$R^2 = 94,44\%$



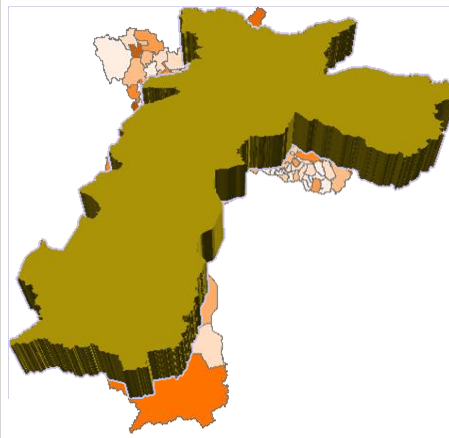
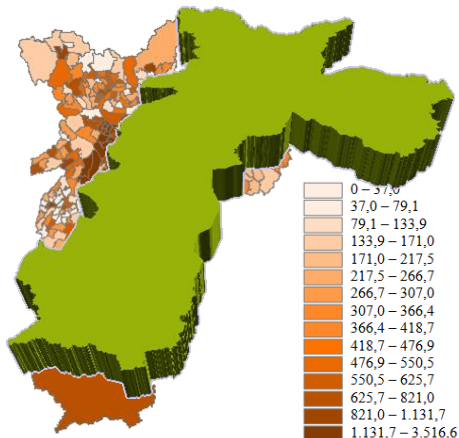
$R^2 = 96,80\%$



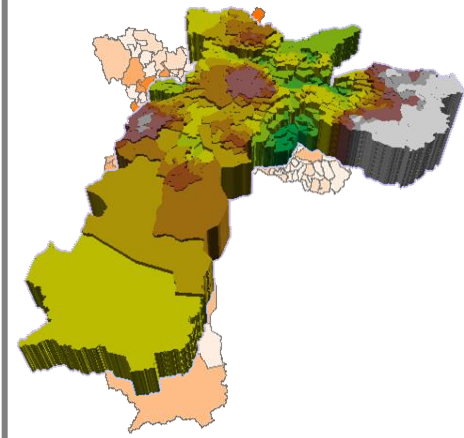
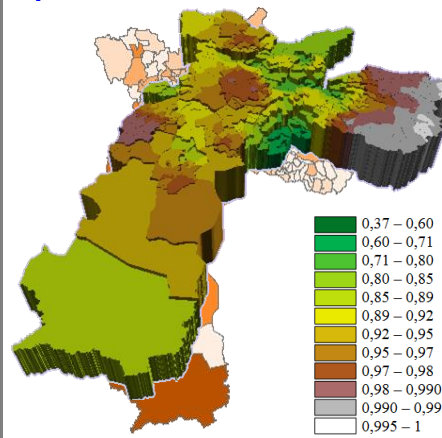
$R^2 = 98,01\%$



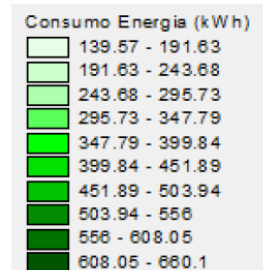
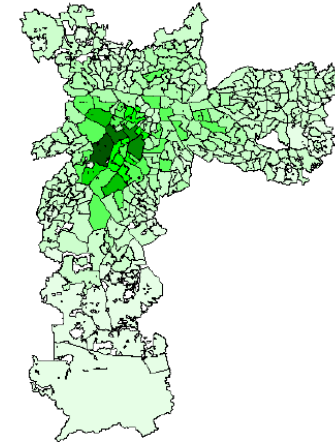
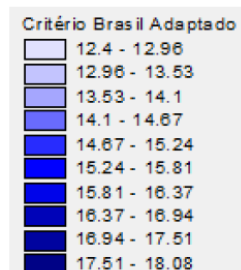
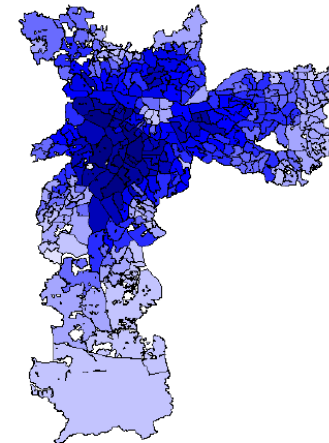
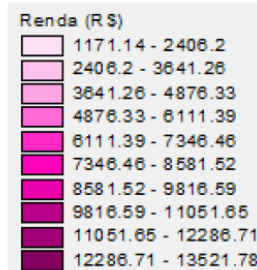
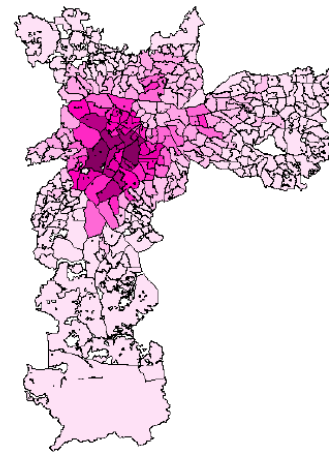
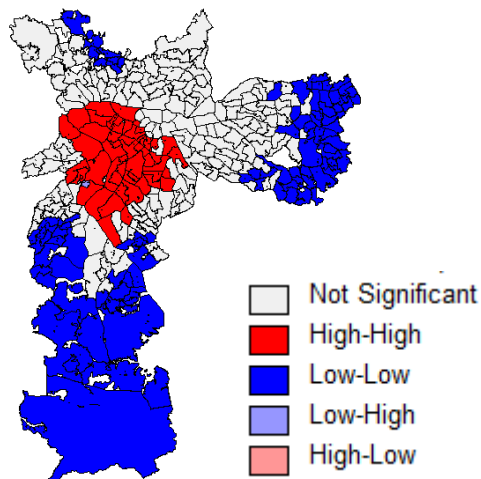
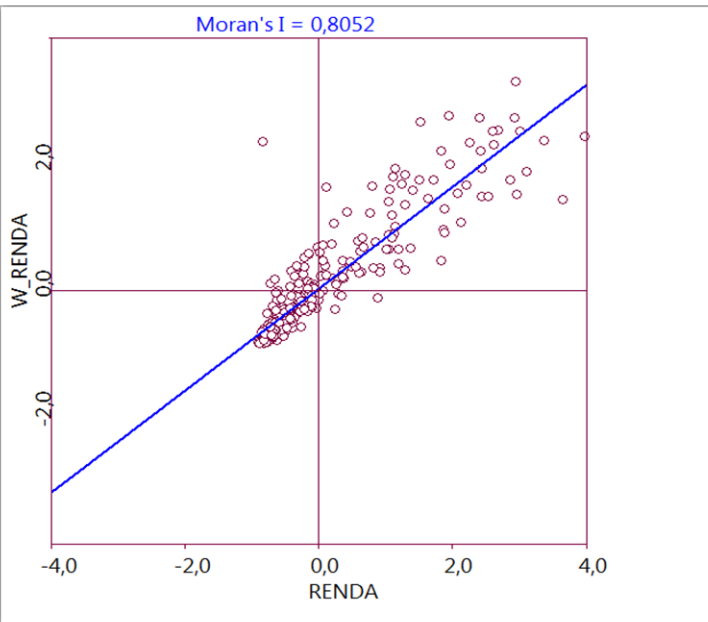
Mapa Resíduos Absolutos



Mapa R^2 Local (3D)



Relação entre Consumo de Energia Elétrica e Renda Domiciliar: Um Estudo Censitário no Município de São Paulo com Dados de 2010



- S.A.R. $\rightarrow R^2 = 0,929$
- R.L.S. $\rightarrow R^2 = 0,812$
- I de Moran = 0,8052

Potenciais Implicações Gerenciais

Exemplos de Agregação Territorial dos Indicadores de Consumo de Energia Elétrica

• Domicílios



Medida Alternativa de Renda: **Atual, abrangente, flexível, e atualizável mensalmente**

Indicadores de Renda baseados em Consumo de Energia Elétrica

Setores Censitários

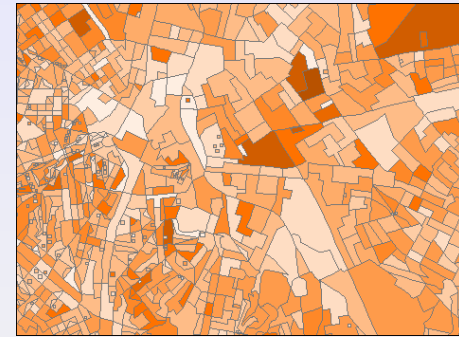
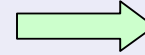
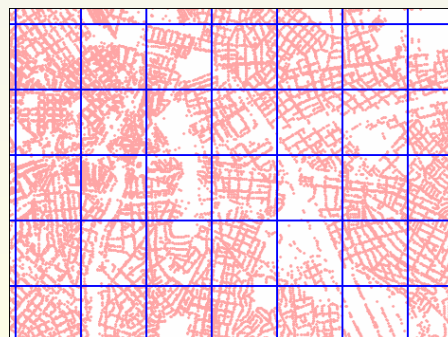


Círculos Concêntricos

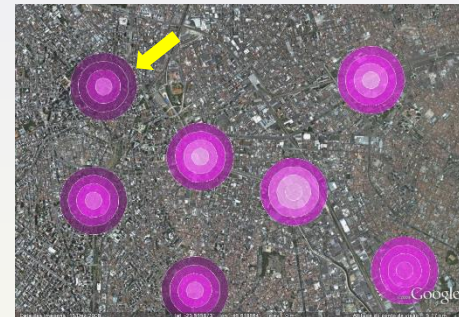
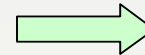
(raios progressivos de 125 m)



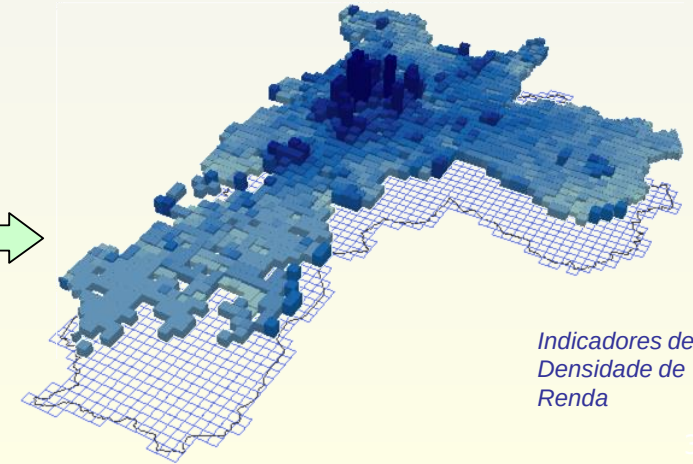
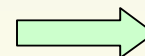
Quadrículas de 1 km²



Indicadores sistemáticos de Renda por Setor Censitário, granularidade convencional do Mercado

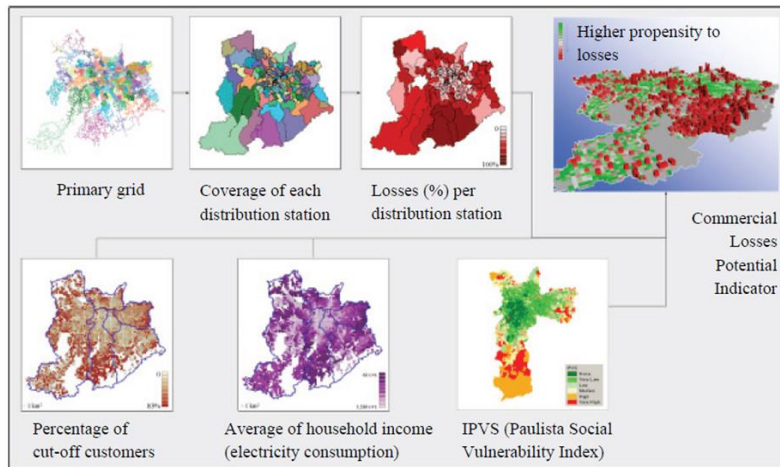


Estudos Ad-hoc (p.e., escolha do melhor ponto comercial para expansão ou reposicionamento no varejo)

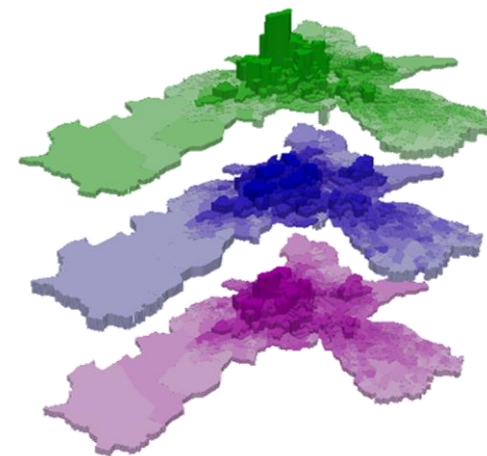


Indicadores de Densidade de Renda

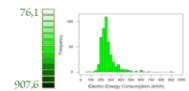
Development of an Indicator of Propensity to Energy Commercial Losses using Geospatial Statistical Techniques and Socio-Economic Data: the Case of AES Eletropaulo



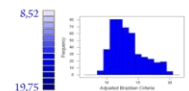
Microcredit Score e Indicadores Sócio-Econômicos baseados em Energia Elétrica



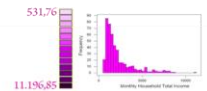
Consumo Residencial de Energia Elétrica (kWh)



Microcredit Score



Renda Domiciliar (R\$)



Muito Obrigado!!!

Eduardo de Rezende Francisco
FGV-EAESP – TDS

eduardo.francisco@fgv.br

