

# Arrumação de dados de um Map-Based Survey com R: Desafios e Lições Aprendidas

Dra. Beatriz Milz



Trans-Atlantic Platform for Social Sciences  
and Humanities (T-AP)

Partner institutions:



LAPLAN LABORATÓRIO DE  
PLANEJAMENTO TERRITORIAL



PROGRAMA DE  
PÓS-GRADUAÇÃO  
EM PLANEJAMENTO E  
GESTÃO DO TERRITÓRIO



technische universität  
dortmund



International  
Planning Studies

IRPUD



UNIVERSITY of the  
WESTERN CAPE



Bundesministerium  
für Bildung  
und Forschung

## 2018: Comecei a estudar programação

Acadêmico



2012-2015  
Graduação em  
Gestão Ambiental

2016-2018  
Mestrado em Análise  
Ambiental Integrada



2019-2024  
Doutorado em  
Ciências Ambientais

2024 - 2025  
Pós-doc no Laboratório de  
Planejamento Territorial



Profissional

2014-2015  
Estágio em Gestão  
Ambiental  
NTGA/SMSP/PMSP



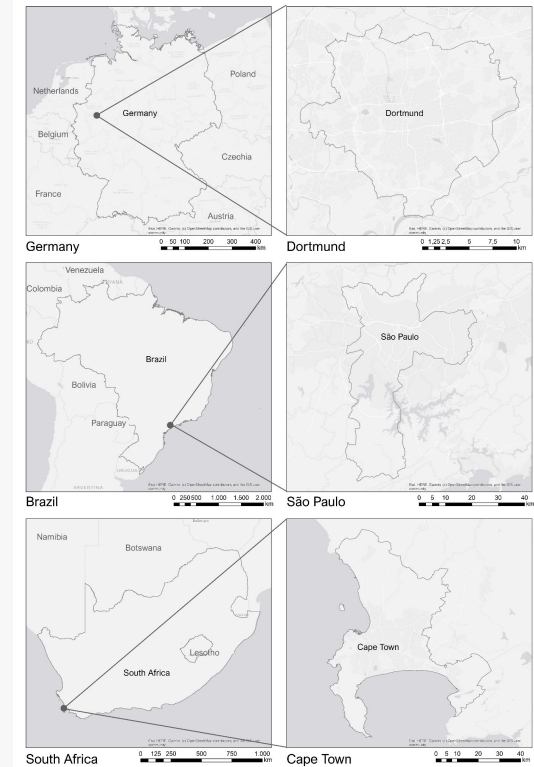
2022-2024  
R6 Consultoria e  
Treinamentos



2012 —————→ 2025

## Projeto ICOLMA

- Tem como objetivo compreender como a pandemia de COVID-19 afetou a mobilidade de pessoas de grupos urbanos marginalizados.
- Locais do estudo: São Paulo (Brasil), Cidade do Cabo (África do Sul), Dortmund (Alemanha).



Fonte da imagem: [Schnittfinke et. al \(2024\)](#)

Partner institutions:



LAPLAN LABORATÓRIO DE  
PLANEJAMENTO TERRITORIAL



PROGRAMA DE  
PÓS-GRADUAÇÃO  
EM PLANEJAMENTO E  
GESTÃO DO TERRITÓRIO



technische universität  
dortmund



International  
Planning Studies

IRPUD



UNIVERSITY OF CAPE TOWN  
IFUNVESTHINI YASEKAPA - UNIVERSITEIT VAN KAAPSTAD



UNIVERSITY of the  
WESTERN CAPE



Bundesministerium  
für Bildung  
und Forschung

# Projeto ICOLMA - 2022-2025 (em andamento)



UNIVERSITY of the  
WESTERN CAPE



Partner institutions:



LAPLAN LABORATÓRIO DE  
PLANEJAMENTO TERRITORIAL



PROGRAMA DE  
PÓS-GRADUAÇÃO  
EM PLANEJAMENTO E  
GESTÃO DO TERRITÓRIO



technische universität  
dortmund



International  
Planning Studies

IRPUD



UNIVERSITY OF CAPE TOWN  
iYUNIVESITHI YASEKAPA - UNIVERSITEIT VAN KAAPSTAD



UNIVERSITY of the  
WESTERN CAPE



Bundesministerium  
für Bildung  
und Forschung

# Sobre a apresentação...

- Projeto foi iniciado em 2022
- Comecei a fazer parte em julho de 2024, após o período de coleta de dados.
- Focada na minha experiência, e alguns desafios e lições aprendidas.

Partner institutions:



**LAPL** LABORATÓRIO DE  
PLANEJAMENTO TERRITORIAL



PROGRAMA DE  
PÓS-GRADUAÇÃO  
EM PLANEJAMENTO E  
GESTÃO DO TERRITÓRIO



technische universität  
dortmund



International  
Planning Studies

**IRPUD**



UNIVERSITY of the  
WESTERN CAPE



Bundesministerium  
für Bildung  
und Forschung

## O que é um map-based survey?

- Os respondentes podem interagir com mapas e fornecer respostas baseadas em localização.
- Tornou-se mais popular nos últimos anos, especialmente em pesquisas participativas.



Source of image: [Geospatial World](#)

Partner institutions:



**LAPLAN** LABORATÓRIO DE  
PLANEJAMENTO TERRITORIAL



PROGRAMA DE  
PÓS-GRADUAÇÃO  
EM PLANEJAMENTO E  
GESTÃO DO TERRITÓRIO



technische universität  
dortmund



International  
Planning Studies



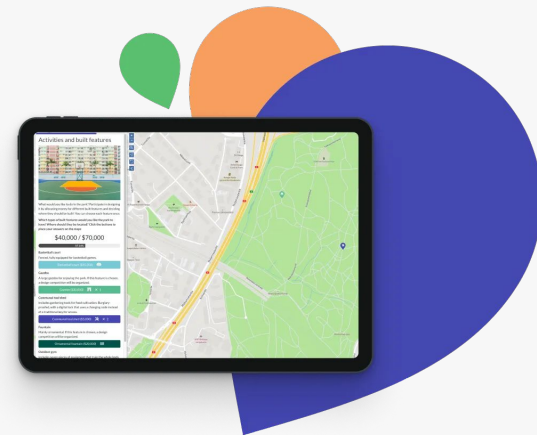
UNIVERSITY of the  
WESTERN CAPE



Bundesministerium  
für Bildung  
und Forschung

## Ferramenta utilizada: Maptionnaire

- Ferramenta que permite fazer a coleta de dados com base em mapas
- Resultados: mais de 300 respostas coletadas presencialmente em três países
- Motivo da escolha: equipe da Alemanha já tinha utilizado em projeto anterior



Source of image:  
<https://www.maptionnaire.com/>

# Desafios e lições aprendidas

Partner institutions:



**LAPLAN** LABORATÓRIO DE  
PLANEJAMENTO TERRITORIAL



PROGRAMA DE  
PÓS-GRADUAÇÃO  
EM PLANEJAMENTO E  
GESTÃO DO TERRITÓRIO



technische universität  
dortmund



International  
Planning Studies

**IRPUD**



UNIVERSITY of the  
WESTERN CAPE



Bundesministerium  
für Bildung  
und Forschung



## Desafio #1: Instabilidade da Internet

- Problemas de conectividade em São Paulo (Cidade Tiradentes)
- O software não permitia o salvamento parcial, apenas após a conclusão da resposta.
- Impacto na completude dos dados
- Reflexão: grupos marginalizados também enfrentam barreiras digitais à participação em pesquisas



## Reflexões

- Cuidado com a extensão do questionário.
- Na etapa de pré-teste, verificar o uso do software também em campo.
- Não depender totalmente do software para a coleta de dados.

Partner institutions:



**LAPL** LABORATÓRIO DE  
PLANEJAMENTO TERRITORIAL



PROGRAMA DE  
PÓS-GRADUAÇÃO  
EM PLANEJAMENTO E  
GESTÃO DO TERRITÓRIO



technische universität  
dortmund



International  
Planning Studies

**IRPUD**



UNIVERSITY of the  
WESTERN CAPE



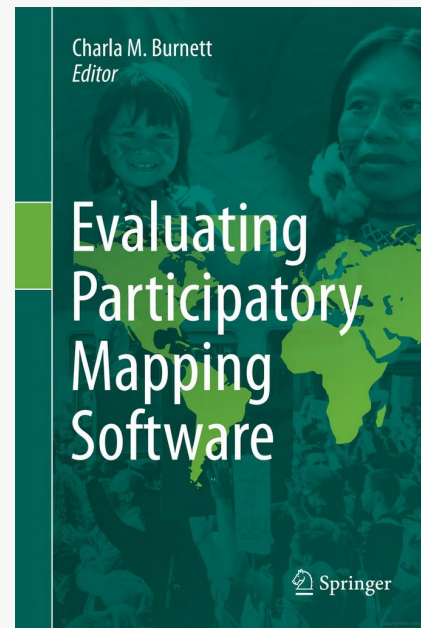
Bundesministerium  
für Bildung  
und Forschung

## Desafio#2: Custos

- Maptionnaire:
  - Software proprietário com altos custos (\$\$\$)
    - ~ 1.000 Euros por mês ([do livro: Evaluating Participatory Mapping Software](#))
  - O acesso é limitado ao período de assinatura ativo
  - Grupos de pesquisa no Sul Global frequentemente enfrentam restrições orçamentárias

## Lições aprendidas

- Explorar alternativas de código aberto!



Partner institutions:



**LAPLAN** LABORATÓRIO DE  
PLANEJAMENTO TERRITORIAL



PROGRAMA DE  
PÓS-GRADUAÇÃO  
EM PLANEJAMENTO E  
GESTÃO DO TERRITÓRIO



technische universität  
dortmund



International  
Planning Studies

**IRPUD**



**UNIVERSITY OF CAPE TOWN**  
IFUNIVESITHI YASEKAPA - UNIVERSITEIT VAN KAAPSTAD



**UNIVERSITY of the  
WESTERN CAPE**



Bundesministerium  
für Bildung  
und Forschung

# Desafio #3: Limpeza e análise de dados

- O Maptionnaire exporta planilhas do Excel com diversas folhas de dados.
- Requer software externo para análise.

S	T	U	V
WKT 26. Where did you mainly go to work or for income g	GeoJSON 26. Where did you mainly go to work or for income	26.1. Whose workplace is this? checkbox-list-8	Respondent checkbox-list-8: option option-identifier-igZhg2u
wkt	geojson		
POINT (-46.673831 -23.537906)	{ "type": "Feature", "geometry": { "type": "Point", "coordinates": [-46.673831, -23.537906] }, "properties": {} }		FALSO
POINT (-46.660782 -23.572405)	{ "type": "Feature", "geometry": { "type": "Point", "coordinates": [-46.660782, -23.572405] }, "properties": {} }		FALSO
POINT (-46.650211 -23.559421)	{ "type": "Feature", "geometry": { "type": "Point", "coordinates": [-46.650211, -23.559421] }, "properties": {} }		VERDADEIRO
POINT (-46.552166 -23.691498)	{ "type": "Feature", "geometry": { "type": "Point", "coordinates": [-46.552166, -23.691498] }, "properties": {} }		VERDADEIRO
POINT (-46.637197 -23.570695)	{ "type": "Feature", "geometry": { "type": "Point", "coordinates": [-46.637197, -23.570695] }, "properties": {} }		FALSO
POINT (-46.569922 -23.598209)	{ "type": "Feature", "geometry": { "type": "Point", "coordinates": [-46.569922, -23.598209] }, "properties": {} }		VERDADEIRO
POINT (-46.464646 -23.654272)	{ "type": "Feature", "geometry": { "type": "Point", "coordinates": [-46.464646, -23.654272] }, "properties": {} }		VERDADEIRO
POINT (-46.628667 -23.575073)	{ "type": "Feature", "geometry": { "type": "Point", "coordinates": [-46.628667, -23.575073] }, "properties": {} }		FALSO
POINT (-46.398281 -23.594065)	{ "type": "Feature", "geometry": { "type": "Point", "coordinates": [-46.398281, -23.594065] }, "properties": {} }		FALSO
POINT (-46.397054 -23.59289)	{ "type": "Feature", "geometry": { "type": "Point", "coordinates": [-46.397054, -23.59289] }, "properties": {} }		FALSO
POINT (-46.722924 -23.094781)	{ "type": "Feature", "geometry": { "type": "Point", "coordinates": [-46.722924, -23.094781] }, "properties": {} }		FALSO
POINT (-46.640073 -23.54371)	{ "type": "Feature", "geometry": { "type": "Point", "coordinates": [-46.640073, -23.54371] }, "properties": {} }		VERDADEIRO
POINT (-46.639932 -23.54371)	{ "type": "Feature", "geometry": { "type": "Point", "coordinates": [-46.639932, -23.54371] }, "properties": {} }		VERDADEIRO
POINT (-46.640422 -23.543727)	{ "type": "Feature", "geometry": { "type": "Point", "coordinates": [-46.640422, -23.543727] }, "properties": {} }		VERDADEIRO
POINT (-46.640053 -23.543968)	{ "type": "Feature", "geometry": { "type": "Point", "coordinates": [-46.640053, -23.543968] }, "properties": {} }		FALSO
POINT (-46.639769 -23.543719)	{ "type": "Feature", "geometry": { "type": "Point", "coordinates": [-46.639769, -23.543719] }, "properties": {} }		VERDADEIRO
POINT (-46.639995 -23.543571)	{ "type": "Feature", "geometry": { "type": "Point", "coordinates": [-46.639995, -23.543571] }, "properties": {} }		VERDADEIRO
POINT (-46.639951 -23.54368)	{ "type": "Feature", "geometry": { "type": "Point", "coordinates": [-46.639951, -23.54368] }, "properties": {} }		VERDADEIRO
POINT (-46.699561 -23.385394)	{ "type": "Feature", "geometry": { "type": "Point", "coordinates": [-46.699561, -23.385394] }, "properties": {} }		FALSO
POINT (-46.639904 -23.543796)	{ "type": "Feature", "geometry": { "type": "Point", "coordinates": [-46.639904, -23.543796] }, "properties": {} }		FALSO
POINT (-46.639768 -23.543704)	{ "type": "Feature", "geometry": { "type": "Point", "coordinates": [-46.639768, -23.543704] }, "properties": {} }		VERDADEIRO
POINT (-46.398544 -23.590165)	{ "type": "Feature", "geometry": { "type": "Point", "coordinates": [-46.398544, -23.590165] }, "properties": {} }		VERDADEIRO
POINT (-46.400103 -23.588259)	{ "type": "Feature", "geometry": { "type": "Point", "coordinates": [-46.400103, -23.588259] }, "properties": {} }		FALSO
POINT (-46.403919 -23.589707)	{ "type": "Feature", "geometry": { "type": "Point", "coordinates": [-46.403919, -23.589707] }, "properties": {} }		FALSO
POINT (-46.402182 -23.584076)	{ "type": "Feature", "geometry": { "type": "Point", "coordinates": [-46.402182, -23.584076] }, "properties": {} }		VERDADEIRO
POINT (-47.059249 -22.905415)	{ "type": "Feature", "geometry": { "type": "Point", "coordinates": [-47.059249, -22.905415] }, "properties": {} }		FALSO
POINT (-46.399663 -23.508683)	{ "type": "Feature", "geometry": { "type": "Point", "coordinates": [-46.399663, -23.508683] }, "properties": {} }		VERDADEIRO
POINT (-46.310218 -23.505828)	{ "type": "Feature", "geometry": { "type": "Point", "coordinates": [-46.310218, -23.505828] }, "properties": {} }		FALSO
POINT (-46.309815 -23.50152)	{ "type": "Feature", "geometry": { "type": "Point", "coordinates": [-46.309815, -23.50152] }, "properties": {} }		FALSO
POINT (-46.404469 -23.600954)	{ "type": "Feature", "geometry": { "type": "Point", "coordinates": [-46.404469, -23.600954] }, "properties": {} }		VERDADEIRO
POINT (-46.481532 -23.597858)	{ "type": "Feature", "geometry": { "type": "Point", "coordinates": [-46.481532, -23.597858] }, "properties": {} }		FALSO

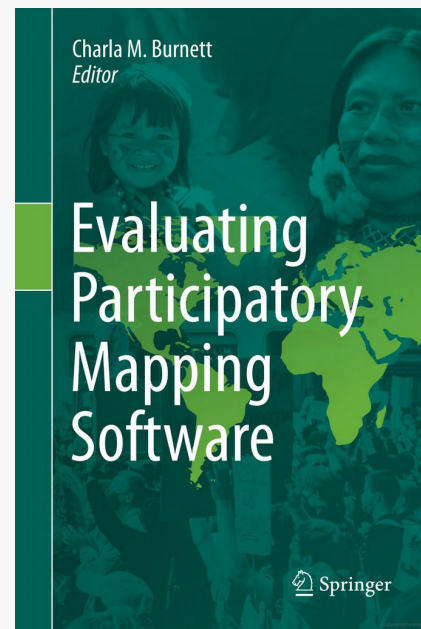
## R para limpeza e análise de dados

- R é uma linguagem de programação de código aberto.
- R permite análises avançadas sem custo financeiro.
- Treinamento técnico é essencial.



## Lições aprendidas

- O Maptionnaire foca na coleta de dados, mas não na validação, limpeza e análise de dados.
- Explorar outras alternativas de software.



Partner institutions:



LAPLAN LABORATÓRIO DE  
PLANEJAMENTO TERRITORIAL



PROGRAMA DE  
PÓS-GRADUAÇÃO  
EM PLANEJAMENTO E  
GESTÃO DO TERRITÓRIO



technische universität  
dortmund



International  
Planning Studies

IRPUD



UNIVERSITY of the  
WESTERN CAPE



Bundesministerium  
für Bildung  
und Forschung

## Desafio #4: Coordenadas

- Software registrou algumas coordenadas fora da área da pesquisa.

Conversamos com a equipe que coletou os dados e descobrimos problemas!

Pesquisadores solicitavam o endereço, e o software registrava como coordenadas.

- Ex: "avenida brigadeiro"



## Lições aprendidas

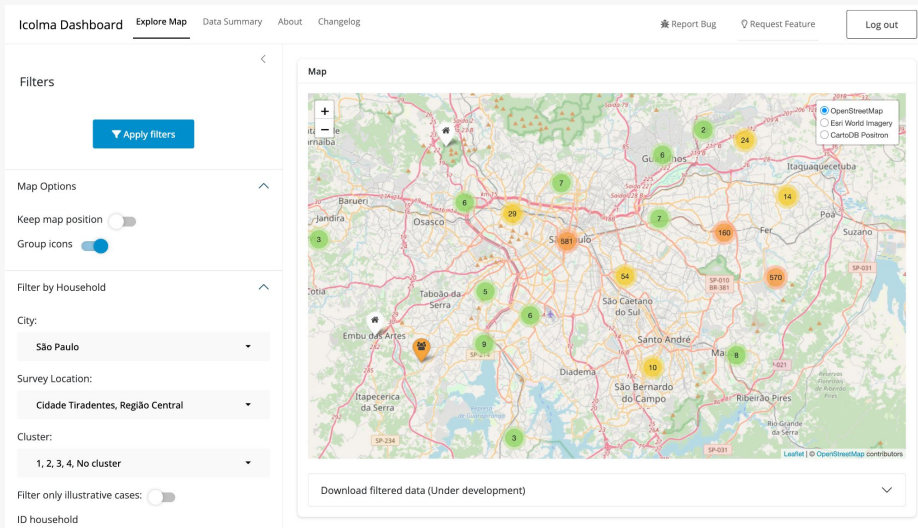
- Seria melhor registrar o endereço -> Com geocoding poderíamos extrair as coordenadas.
- Acompanhamento dos dados coletados ao longo do processo de coleta de dados, para alinhamento com a equipe de pesquisa.

## Validação das coordenadas

- Agrupamento por pessoa/período - Matriz de distâncias entre cada atividade.
  - Distâncias grandes -> indicam que pode haver um erro
- Algumas coordenadas tiveram que ser verificadas manualmente para serem corrigidas, outras tiveram que ser removidas.

# Dashboard para exploração de dados

- **WIP:** Painel interno para ajudar a detectar possíveis erros na coleta de dados e ajudar a explorar os dados (WIP)





▼ Apply filters



## Map



### Response identification

ID: 77662k6g8la4

Person: 1

N. Location: 1

Activity: work

Period: t2

Survey location: Cidade Tiradentes

Cluster: 3

### Responses Geo Questions

#### - Mobility

Mode of transport: Bus; Subway / Tram

Time spent on transport: 2 horas e meia

Time stayed at location: 7 horas

Frequency: 4-3 times a week

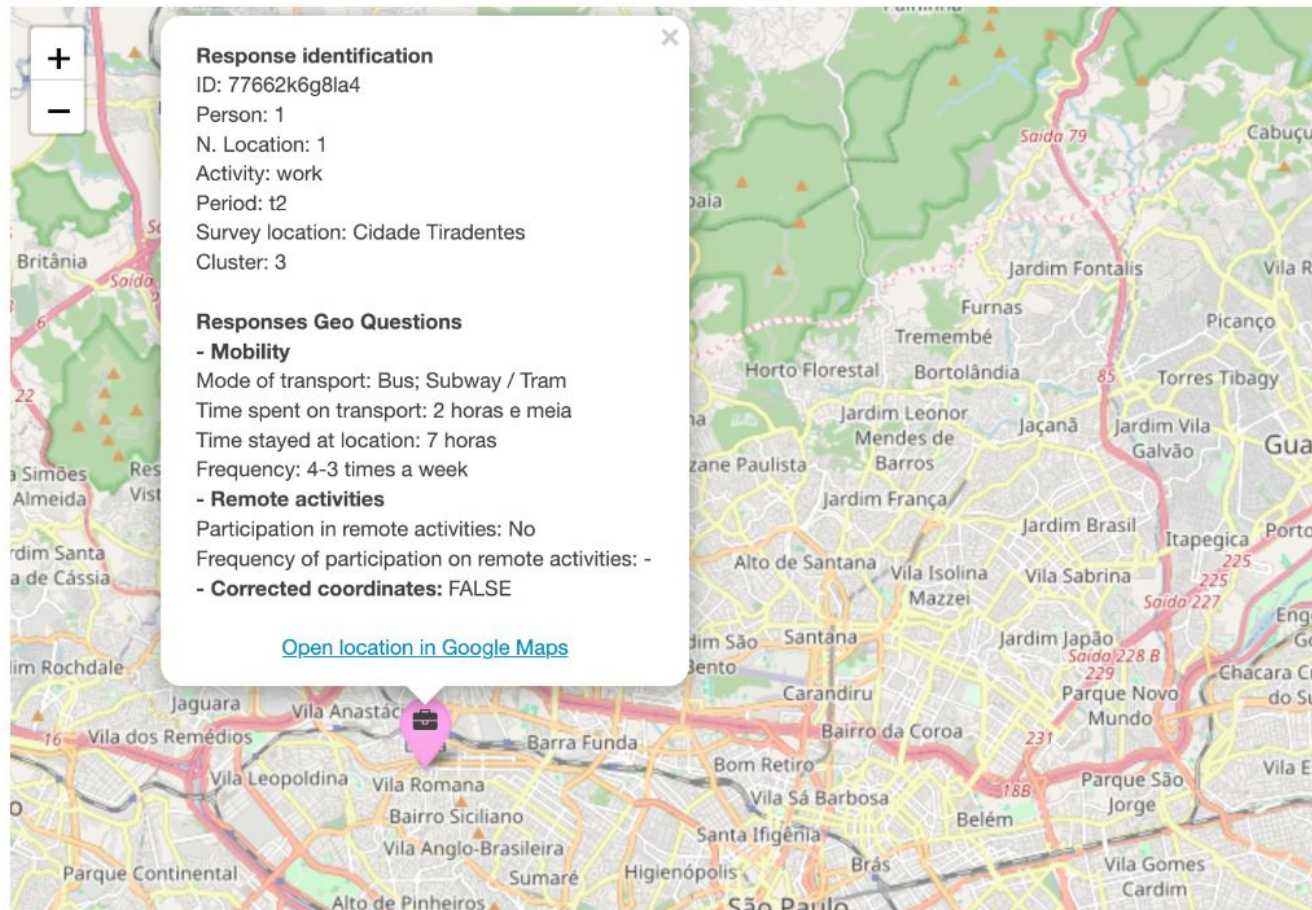
#### - Remote activities

Participation in remote activities: No

Frequency of participation on remote activities: -

- **Corrected coordinates:** FALSE

[Open location in Google Maps](#)



# Principal recomendação

- Considere o planejamento da análise de dados desde o início da pesquisa.

Partner institutions:



**LAPLAN** LABORATÓRIO DE  
PLANEJAMENTO TERRITORIAL



PROGRAMA DE  
PÓS-GRADUAÇÃO  
EM PLANEJAMENTO E  
GESTÃO DO TERRITÓRIO



technische universität  
dortmund



International  
Planning Studies

**IRPUD**



UNIVERSITY of the  
WESTERN CAPE



Bundesministerium  
für Bildung  
und Forschung



# Obrigada

Beatriz Milz

Universidade Federal do ABC (UFABC)

[milz.bea@gmail.com](mailto:milz.bea@gmail.com) - <https://beamilz.com>

Funding: Fapesp process 2021/07554-8 and 2024/05779-0.

Partner institutions:



LAPLAN LABORATÓRIO DE  
PLANEJAMENTO TERRITORIAL



PROGRAMA DE  
PÓS-GRADUAÇÃO  
EM PLANEJAMENTO E  
GESTÃO DO TERRITÓRIO



technische universität  
dortmund



International  
Planning Studies

IRPUD



UNIVERSITY of the  
WESTERN CAPE



Bundesministerium  
für Bildung  
und Forschung