

FLORESTA 2030

*Mapeando processos de crimes
ambientais*

Arthur Clini de Souza – IA, físico, desenvolvedor.

Rodrigo Vieira da Silva – IA e desenvolvedor.

Fabio Akahoshi Collado – IA em Direito.

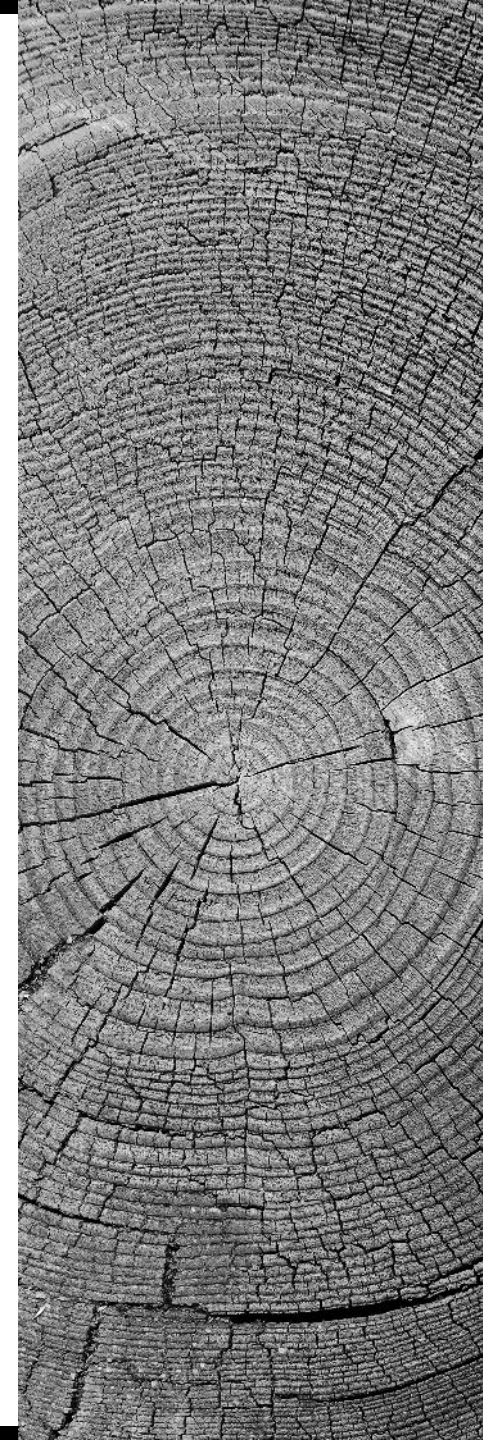


VOCÊ SABE QUAIS CRIMES AMBIENTAIS ESTÃO SENDO INVESTIGADOS E PROCESSADOS PERTO DE VOCÊ?

É essa questão que nosso grupo se propôs a resolver no desafio HACKLIODSCNJ!

Problema

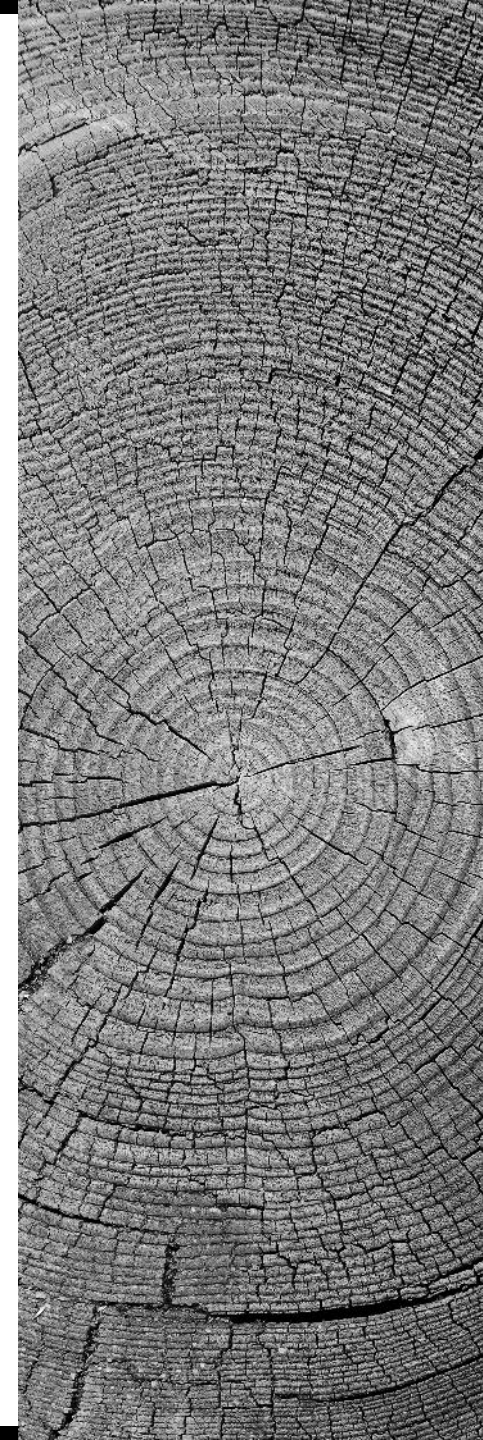
- Situações de interesse social e ambiental: incêndio, desmatamento, degradação, mineração, existência de gado, plantação, trabalho infantil e trabalho análogo ao de escravo.
- Os crimes ambientais relacionados são investigados, processados e julgados pelos órgãos públicos.
- Embora muitos desses processos sejam públicos, permitindo a qualquer cidadão consultá-los e fiscalizá-los, é muito difícil encontrar um processo cujo objeto seja próximo a um local de interesse, como uma reserva indígena ou uma floresta protegida.



QUAL A ESTRUTURA DOS DADOS QUE DEVEM SER FORNECIDOS PELO JUDICIÁRIO?

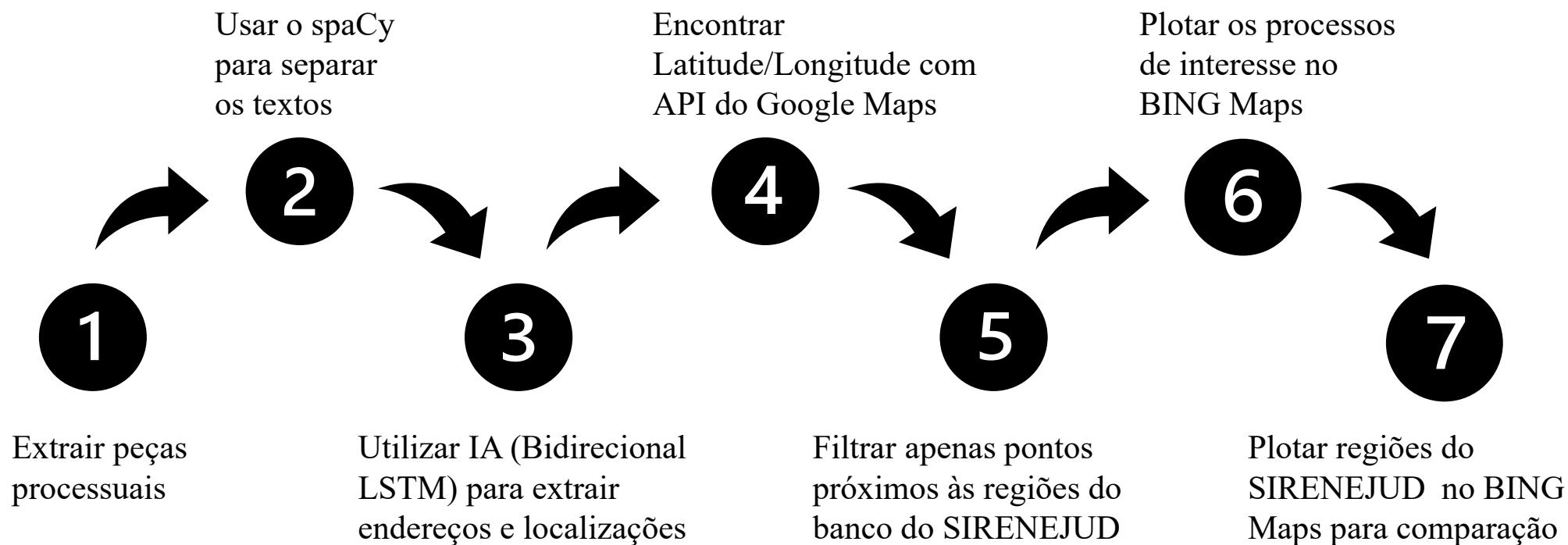
DESAFIO 2

- Se os dados não estiverem estruturados, para correlacionar a região atingida ao número do processo, todos os dados podem ser necessários (muitos dos dados necessários estão em peças que normalmente não são disponibilizadas). Como essa quantidade de dados é muito grande, certamente existirá entre eles dados pessoais e dados sensíveis.
- Com a LGPD, a disponibilização de dados precisa ser cautelosa. Por isso, a estruturação dos dados é imprescindível.
- É necessária e viável a estruturação dos dados de processos cujo objeto são áreas dentro ou próximas daquelas do SIRENE-JUD.
- O principal dado a ser extraído de um processo para resolver o problema em questão é a geolocalização.
- É necessário e viável (tal como implementamos) um algoritmo de IA que extrai informações de localização das peças processuais, transforme-as em latitude/longitude e compare-as com a base do SIRENE-JUD.
- Esses dados podem ser entregues ao jurisdicionado, adicionadas nos sites de pesquisa dos Tribunais ou utilizados em plataformas de buscas mais intuitivas, como a que implementamos.



COMO ESTRUTURAR OS DADOS DAS PEÇAS PROCESSUAIS?

Nossa solução



NOSSA SOLUÇÃO PERMITE TRÊS DIFERENTES POSSIBILIDADES:

1. DISPONIBILIZAÇÃO

- Disponibilização dos dados de maneira estruturada em formato de planilha.
- Como os dados estão estruturados, apenas aquilo que realmente for necessário será disponibilizado, permitindo um maior controle e respeito à LGPD.

2. MELHORAR A BUSCA EXISTENTE

- Melhorar a busca já existente nos sites dos Tribunais, permitindo que se busque processos próximos a certa localização.
- Adiciona-se um campo de busca no site que faz uma comparação com os dados extraídos pelo nosso algoritmo.

3. Fazer uma busca completamente nova, através do Bing Maps

- Ao invés de explicar, veja você mesmo(a): <https://floresta2030.github.io/?semcache>



TERRA INDIGENA
PARAKANA

0003570-73

TERRA INDIGENA PARAKANA

PAE ALDEIA

Quilômetro 13

Nova IPIXUNA

Terra Indígena Mãe Maria

PAE SÃO RAIMUNDO

TERRA INDIGENA MAE MARIA

Itupiranga

MARABÁ

Marabá

Rebojinho

Papagaio

Cigana

São Domingos
do Araguaia

Pedra de
Amolar

São João do
Araguaia

Bacurizinho

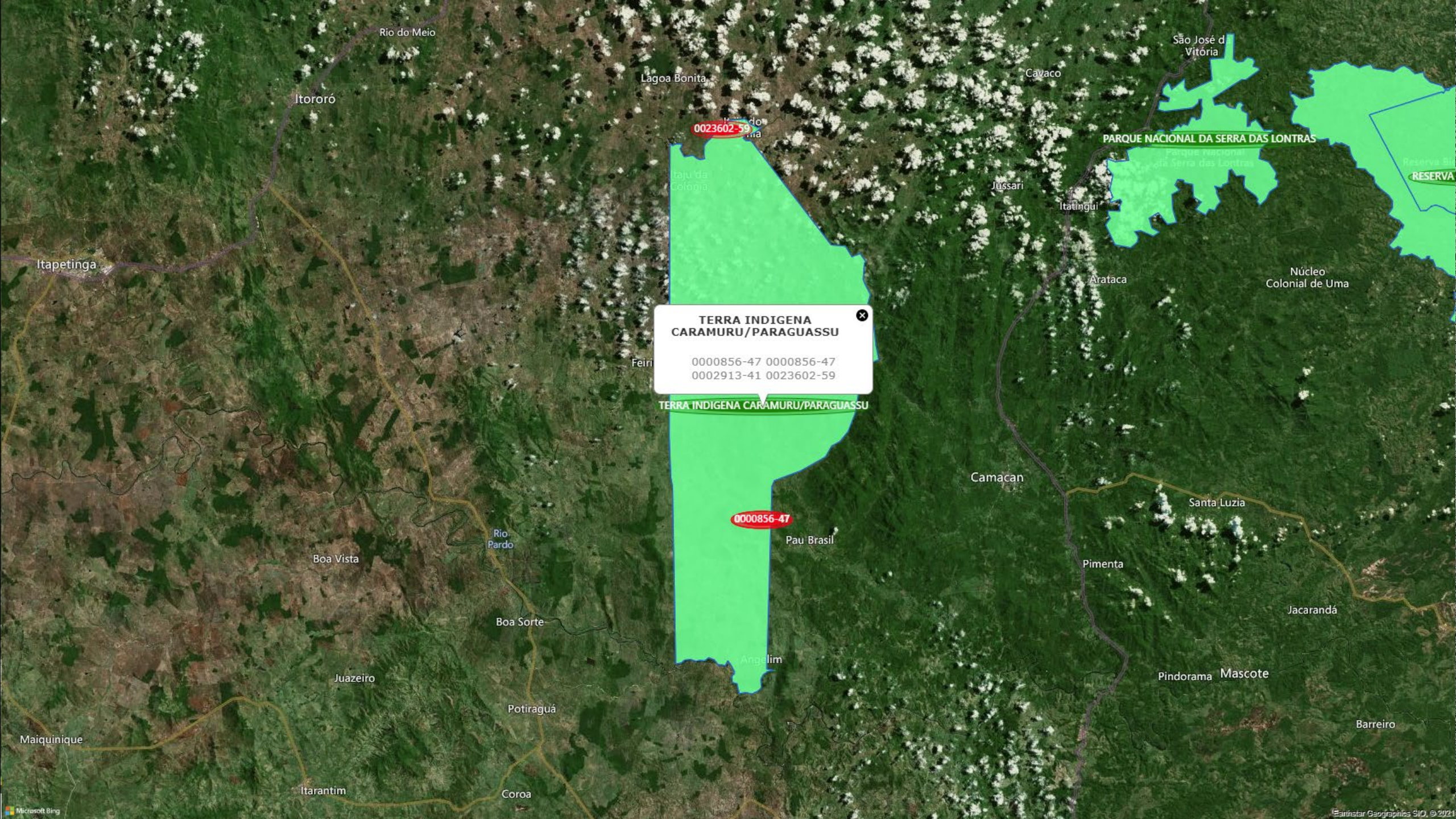
Reserva Biológica do Tapirapé
RESERVA BIOLOGICA DO TAPIRAPE

ITACAIUNAS

SORORÓ

FLORESTA NACIONAL DO TAPIRAPE-AQUIRI

5 milhas 19 km



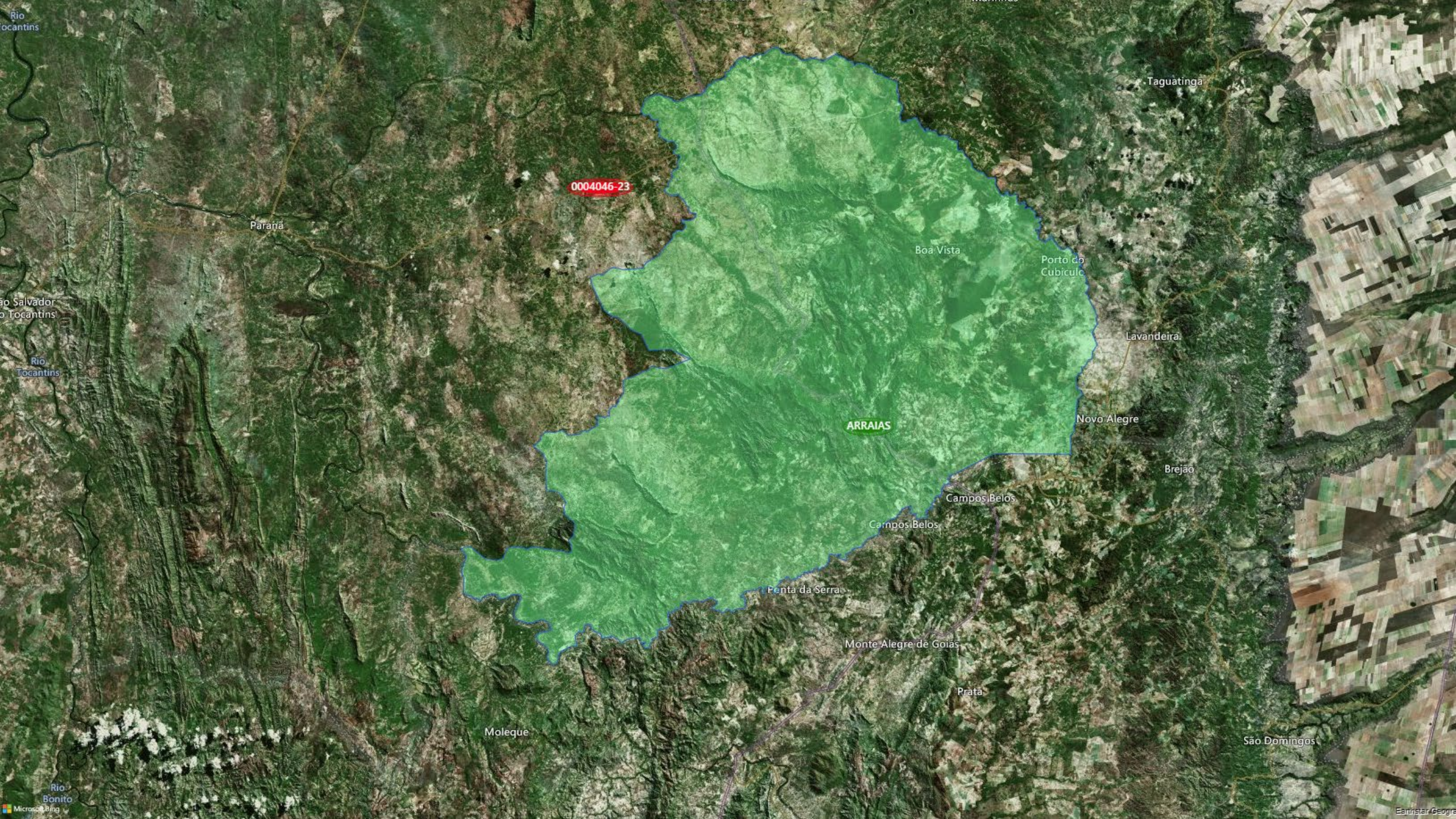
**TERRA INDIGENA
CARAMURU/PARAGUASSU**

0000856-47 0000856-47
0002913-41 0023602-59

TERRA INDIGENA CARAMURU/PARAGUASSU

0023602-59

0000856-47



Rio Tocantins

Rio Salvador
do Tocantins

Rio Bonito
Microsoft Bing

Paraná

0004046-23

ARRAIAS

Boa Vista

Porto do
Cubículo

Taguatinga

Lavandeira

Novo Alegre

Brejão

Campos Belos

Campos Belos

Ponta da Serra

Monte Alegre de Goiás

Prata

Moleque

São Domingos

Earthstar Geospatial

PRINCIPAIS DESAFIOS ENCONTRADOS

- O principal desafio foi o tempo. Foram utilizadas quatro Inteligências Artificiais, duas para estruturação dos dados e duas (k-means) para plotar os dados de forma amigável.
- O segundo desafio foi o cansaço. Trabalhar durante a pandemia até de madrugada no final de semana após uma semana de trabalho foi exaustivo.
- O terceiro desafio foi adaptar técnicas que deveriam ser desempenhadas durante semanas por uma equipe completa em três dias por três pessoas. A principal adaptação foram as dezenas de milhares de dados sintéticos, utilizados para o treinamento da IA principal, que tiveram que ser ajustados várias vezes para aumentar a generalização.
- O quarto desafio foi encontrar e entender os dados geográficos necessários.

MAS, VALEU A PENA!!!

- A solução está implementada e funcionando, do começo ao fim.
- Apesar de usar complexas estratégias computacionais, a experiência para o usuário final é simples e intuitiva, não requerendo qualquer conhecimento específico ou treinamento. É só acessar o site <https://floresta2030.github.io/?semcache> e conferir!
- A solução implementada auxiliará na transparência do combate aos crimes ambientais e o fornecimento de informações para a estruturação de políticas públicas que colaborem com as ODS 13 e 15.



OBSERVAÇÕES FINAIS E AGRADECIMENTOS

Observações:

- Outros dados como área, características da área e informações processuais podem ser extraídos com a mesma estratégia.
- No nosso exemplo estamos mostrando o número do processo porque no desafio descrito no regulamento dá importância a ele. Mas poderíamos mostrar, além do número, a ementa do acórdão ou um resumo da causa.
- Para executar o programa, está tudo no github (<https://github.com/floresta2030/floresta2030.github.io>), exceto os embeddings, que podem ser baixados do site: <http://nilc.icmc.usp.br/nilc/index.php/repositorio-de-word-embeddings-do-nilc>
- Vídeo do pitch: https://www.youtube.com/watch?v=o5YJWKJuH_E
- Vídeo da demonstração: <https://www.youtube.com/watch?v=bFGBH2zYvDE>

AGRADECEMOS PELA OPORTUNIDADE E ATÉ O PRÓXIMO HACKATON!

