

Instituto Nacional de
Pesquisas Espaciais

SEMINÁRIO DE PESQUISA

*Introdução à Geoinformática
(CAP 395-3)*

Autor: Sandro de Sena Machado (CAP)

Orientador: Gilberto Ribeiro Queiroz
Co-orientação: Karine Reis Ferreira

Explorando uma Arquitetura Lakehouse para Análise da Trajetória de Uso e Cobertura da Terra em Áreas de Desmatamento

*Integração dos dados PRODES, DETER e TerraClass
nos Biomas Amazônia e Cerrado (2018–2022)*

26 de junho de 2026



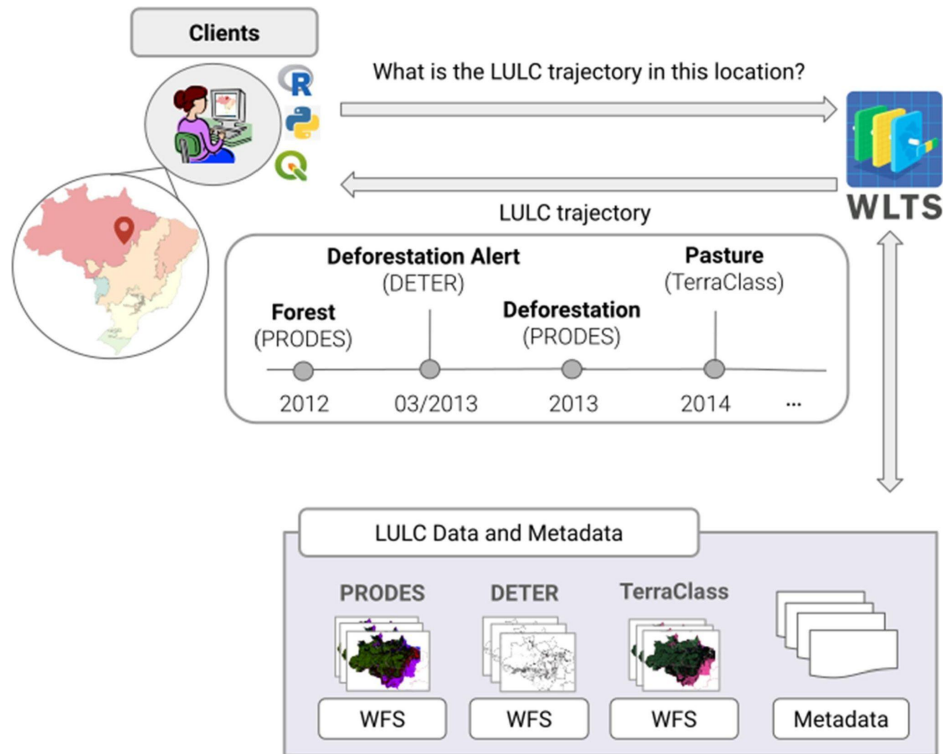
Contexto: Uso do Web Land Trajectory Service (WLTS)

Web Land Trajectory Service (WLTS)

- Solução consolidada do INPE para recuperação de trajetórias de uso e cobertura do solo.
- Integra bases distintas: PRODES, DETER, TerraClass, MapBiomas.

Arquitetura Base Existente

- Persistência em SGBD relacional espacial: PostgreSQL/PostGIS.
- Distribuição via GeoServer e microserviços web.
- Comunicação via padrões OGC: APIs WFS/WMS.
- Paradigma Schema-on-write: modelagem e indexação rígidas prévias à ingestão.



Motivação: Por que explorar o Lakehouse?



Natureza do Experimento

Investigação de viabilidade técnica e desempenho de um novo paradigma de Big Data; não visa substituir sistemas estáveis em produção.



Escalabilidade Distribuída Nativa

SGBDs relacionais operam majoritariamente de forma centralizada (single-node). O ecossistema Lakehouse foi concebido para processamento em cluster.



Desacoplamento de Infraestrutura

Separação entre armazenamento de baixo custo (object storage local) e computação efêmera de alto desempenho (Spark/Sedona).



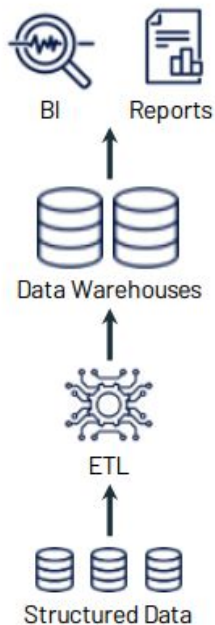
Ingestão Híbrida e Flexível

Manipulação nativa de geometrias vetoriais complexas e matrizes raster de satélite sem necessidade de normalização prévia em tabelas relacionais.

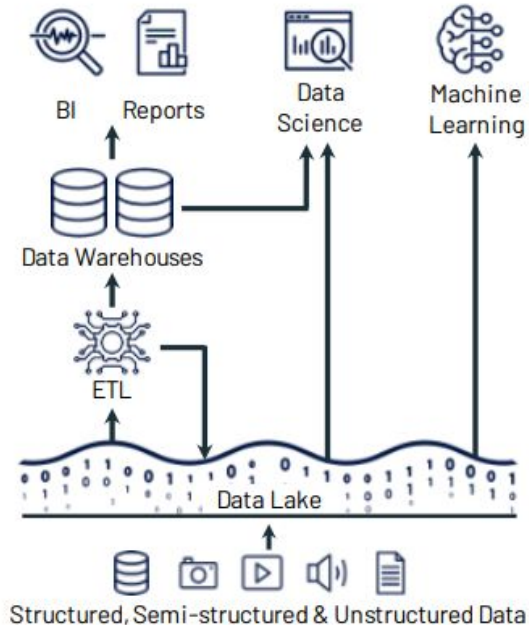
O Panorama Arquitetural das Tecnologias de Dados

Atributo / Paradigma	Data Warehouse / DB Espacial	Data Lake
Tipos de Dados	Estruturados e processados	Estruturados, semi-estruturados e não-estruturados
Formato de Dados	Fechado e proprietário	Aberto e colunar (GeoParquet, Avro)
Abordagem	Schema-on-write	Schema-on-read
Conformidade ACID	Completa — garante integridade máxima	Ausente de suporte nativo
Escalabilidade	Complexa — computação e disco acoplados	Nativa, fácil e de baixo custo

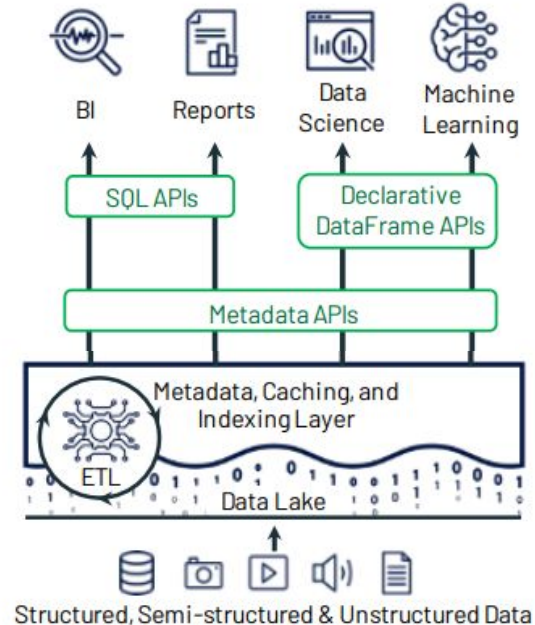
Representação esquemática das arquiteturas



(a) First-generation platforms.



(b) Current two-tier architectures.



(c) Lakehouse platforms.

O Paradigma Lakehouse no Contexto Geoespacial

A Síntese Arquitetural

O lakehouse é apresentado como plataforma unificada que herda do data warehouse a governança e as garantias ACID, e do data lake a flexibilidade, o baixo custo e o suporte a dados heterogêneos.

Desafios para o Domínio Geoespacial

- **Particionamento:** Divisão por dados/espço para manter localidade e tratar polígonos que cruzam partições.
- **Heterogeneidade de Dados:** Superação das barreiras sintáticas, estruturais, semânticas e de modelagem
- **Metadados e Governança:** Alinhamento com a norma ISO 19115 e uso da especificação STAC para catalogação
- **Indexação:** Uso de índices globais/locais e filtragem por retângulos envolventes.

Posicionamento do Paradigma

DATA WAREHOUSE

Schema-on-write · Tabelas · ACID · SQL · BI

DATA LAKEHOUSE

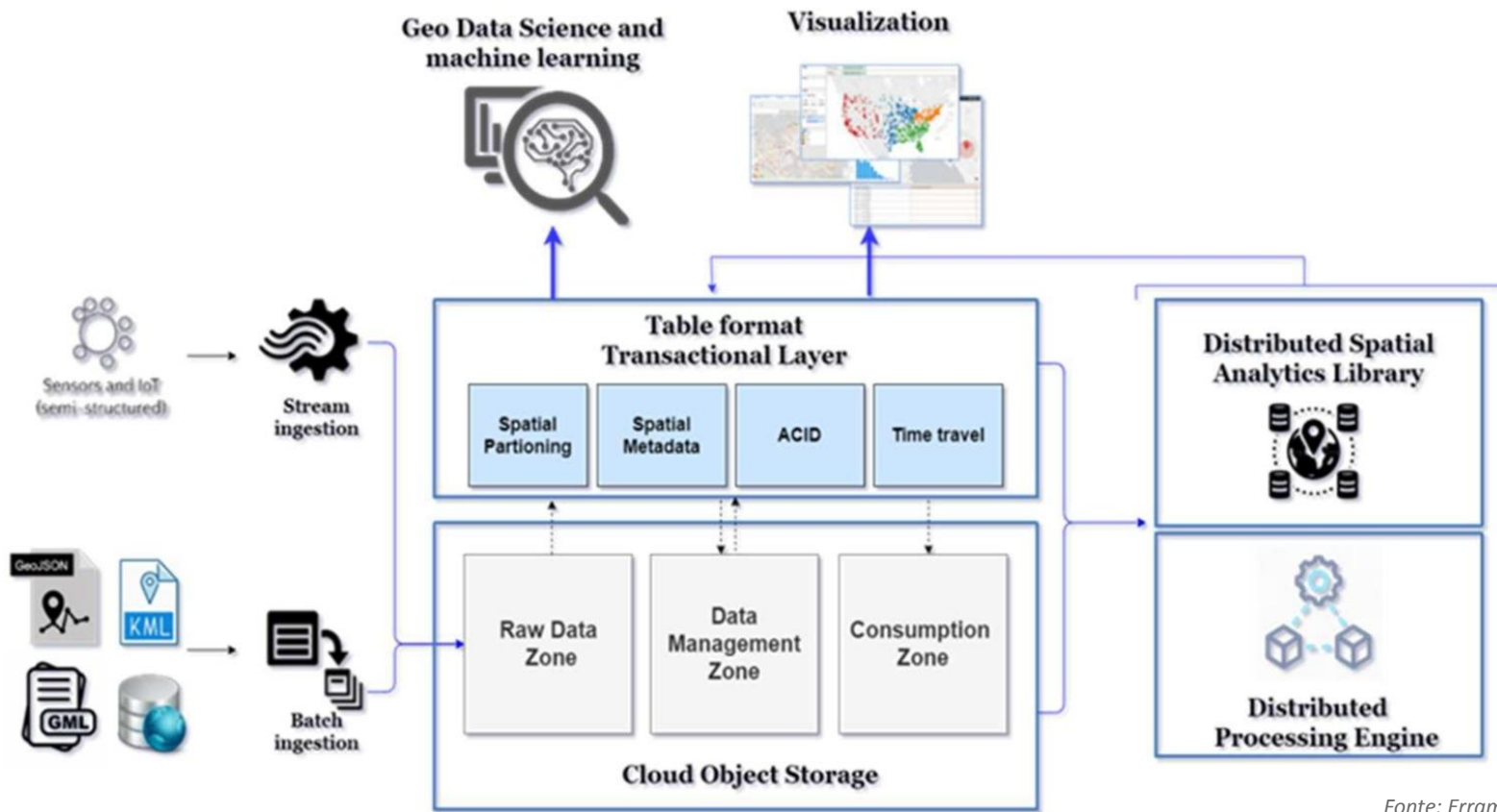
Delta Lake · Spark/Sedona · GeoParquet · COG · BI/ML

DATA LAKE

Schema-on-read · COG · GeoParquet · ML

Lacuna Científica: Tecnologia em estágio inicial de maturação, com escassez de casos de uso formais documentados na literatura para dados geoespaciais.

Esquema de um Lakehouse Geoespacial



Objetivos do Projeto

OBJETIVO GERAL

Desenvolver e avaliar um protótipo local de arquitetura Lakehouse para integração e análise exploratória de dados geoespaciais (PRODES, DETER, TerraClass), validando a viabilidade do paradigma frente aos dados dos biomas Amazônia e Cerrado no recorte temporal de 2018 à 2022.

OBJETIVOS ESPECÍFICOS

- 1 Implementar pipeline distribuído na Arquitetura Medalhão (Bronze, Prata, Ouro) utilizando Apache Sedona e Delta Lake.
- 2 Executar análises exploratórias territoriais agregadas por UFs, Municípios, Terras Indígenas e Unidades de Conservação.
- 3 Extrair trajetórias de mudança de uso da terra em polígonos amostrais do PRODES com os rasters TerraClass (2018, 2020, 2022).
- 4 Avaliar o protótipo à luz dos requisitos formais estabelecidos por Schneider et al. (2024).

Metodologia: Parâmetros Formais de Avaliação do Protótipo

R1

Armazenamento e formato unificados

R5

Garantias de consistência

R2

Suporte a operações CRUD completas sobre arquivos distribuídos

R6

Atomicidade e isolamento

R3

Representação relacional dos dados

R7

Acesso direto ao dado pelo motor de processamento (sem camadas intermediárias)

R4

Suporte a SQL

R8

Integração nativa entre processamento batch e streaming

8 requisitos técnicos independentes de fornecedor

Fonte: Schneider et al., 2024

Metodologia: Ambiente Local e Camada Bronze

Orquestração dos micro serviços — Docker

Container 1

Apache Spark 4.0 + Apache Sedona 1.9
(motor de processamento distribuído)

Container 2

MinIO — Object Storage local
(emula comportamento de bucket S3)

Container 3

Motor de Ingestão Espacial
baseado em GDAL

Estratégia de Ingestão — Camada Bronze

Dados Vetoriais

PRODES, DETER, TIs, UCs, UFs e Municípios

Consumidos via requisições WFS oficiais e persistidos pelo GDAL diretamente em GeoParquet (otimização colunar).

Dados Raster

TerraClass

Download manual e conversão via `gdal_translate` para COG (Cloud Optimized GeoTIFF), viabilizando leituras eficientes por blocos.

WFS / GDAL → Bronze Layer (MinIO / GeoParquet + COG)

Metodologia: Governança e Unificação na Camada Prata

Motor de processamento principal: Delta Lake 4.2 integrado ao Spark 4.0 e Sedona 1.9

Tratamento e Padronização dos dados

- **Suporte Nativo a Geometria:** tipo geometry nativo do Delta Lake 4.2, sem conversão para WKB.
- **Qualidade Topológica:** correção via ST_MakeValid de geometrias inválidas.
- **Padronização:** reprojeção de todos os datasets para EPSG:10857 (SIRGAS 2000 / Albers Equal Area), recálculo do atributo area_ha e modelagem relacional.
- **Indexação:** Z-Order a partir do cálculo do GeoHash.
- **Consolidação:** unificação das camadas e biomas em tabelas únicas (PRODES e DETER), criação de chave surrogada (hashmd5) e particionamento por ano.

Enriquecimento de Metadados

Delta Log

Inclusão de comentários sobre os processos incluídos no log para documentação e governança de tabelas e colunas.

Change Data Feed (CDF)

Ativado para rastreabilidade completa de mutações nos dados (Updates, Inserts, Deletes).

Bronze (GeoParquet / COG) → Silver (Delta Lake · Geometria Nativa · EPSG:10857)

Metodologia: Extração de Trajetórias na Camada Ouro

Desafio Computacional

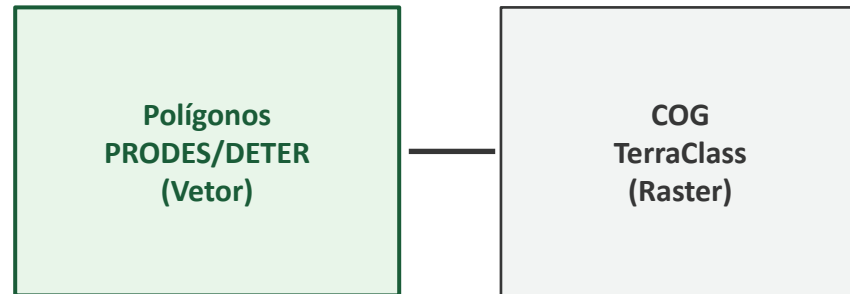
Inviabilidade de cruzar matrizes raster globais completas com geometrias complexas em hardware local com recursos restritos.

Critério de Amostragem

- Filtros espaciais, temporais e de cruzamentos com áreas de interesse.
- 5 maiores polígonos de desmatamento do ano de 2017 por bioma.
- Período anterior à linha de base do TerraClass 2018.

Enriquecimento das tabelas

Join Vetor-Raster — Apache Sedona



RS_Value (Sedona) → Trajetória de Pixels

Anos extraídos: 2018 · 2020 · 2022

Sem vetorização do raster — uso do dado em formato matricial

Silver (Delta Lake) → RS_Value (Sedona) → Gold Layer (Trajetórias 2018·2020·2022)

Metodologia: Contexto Territorial e Regulatório

Enriquecimento com dados territoriais

1

Cruzamento Geométrico

Polígonos cruzados contra camadas oficiais de áreas protegidas.

2

Lógica Condicional

Avaliação de intersecção com Terras Indígenas (TIs) ou Unidades de Conservação (UCs).

3

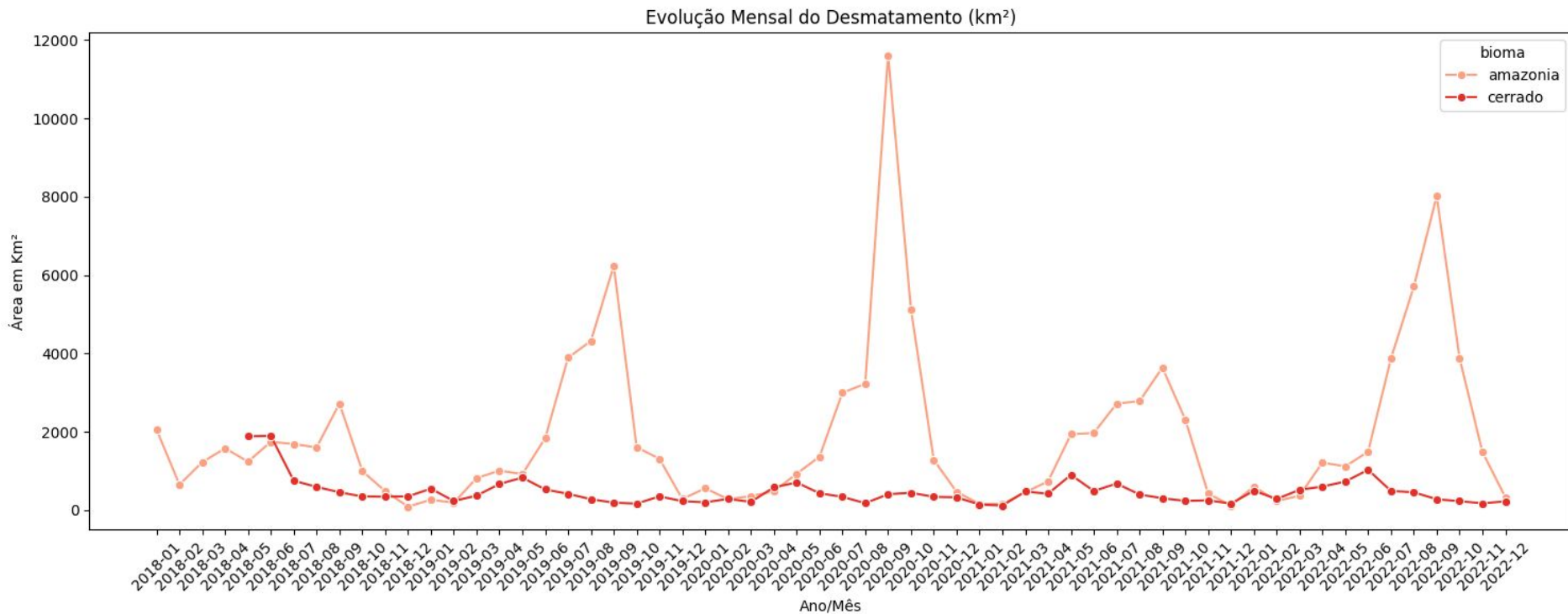
Registro Condicional

Se positivo: nome oficial e área de intersecção em km². Se negativo: atributo persistido como NULL.

Segregação Analítica: Marco Regulatório de Crédito Rural

Segregação da tabela consolidada entre o cenário anterior e posterior à Resolução CMN nº 5.268 de 18/12/2025 (crédito rural).

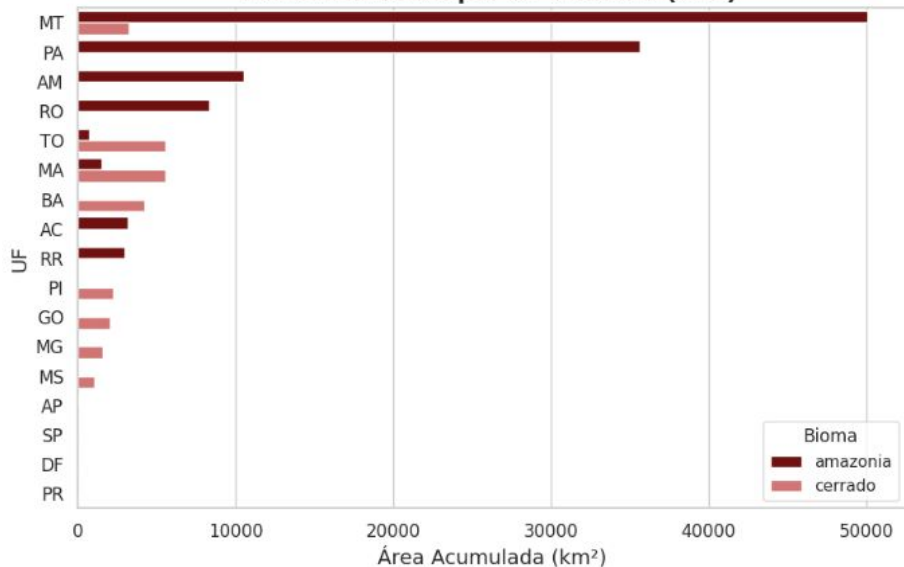
Resultados: Análise Exploratória (DETER - 2018/2022)



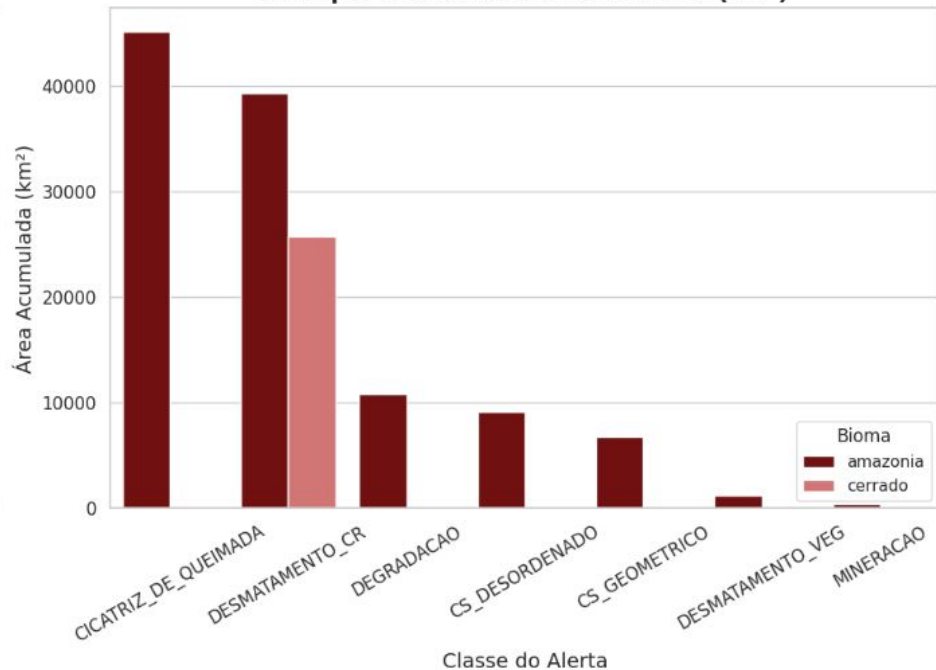
Resultados: Análise Exploratória (DETER - 2018/2022)

DETER - Área por UF e por Classe

Área de Alertas por UF e Bioma (km²)

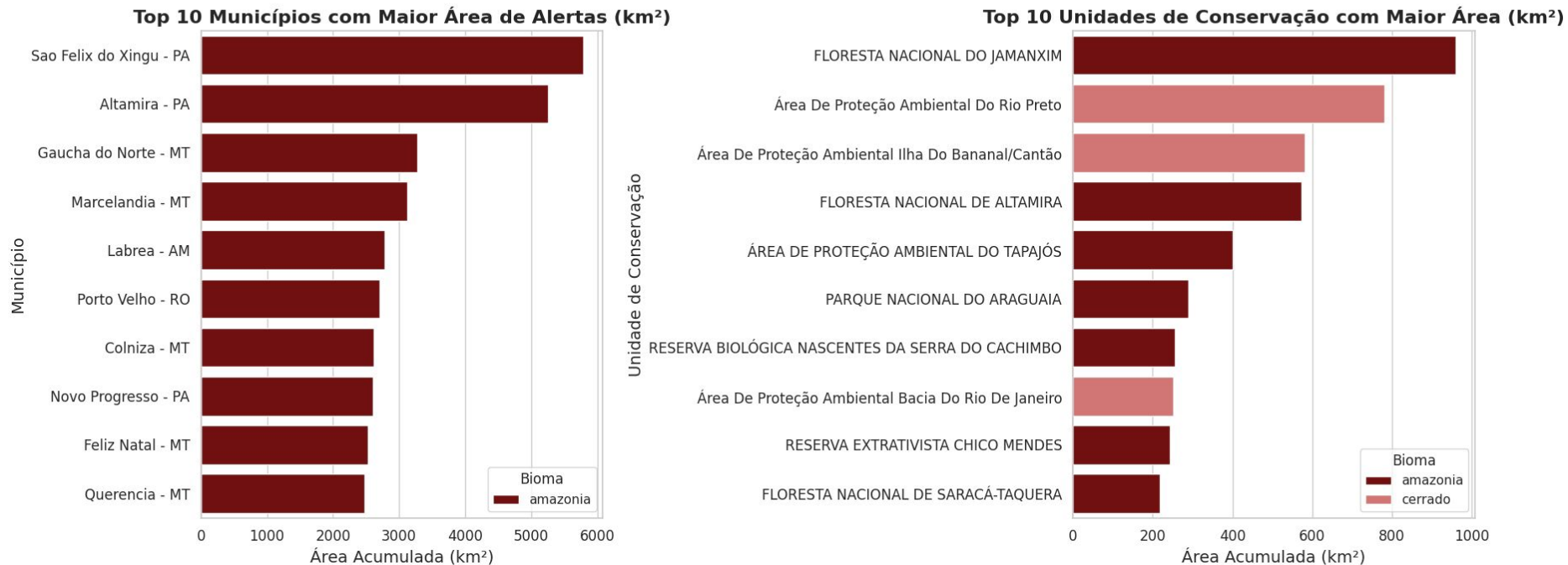


Área por Classe de Alerta e Bioma (km²)



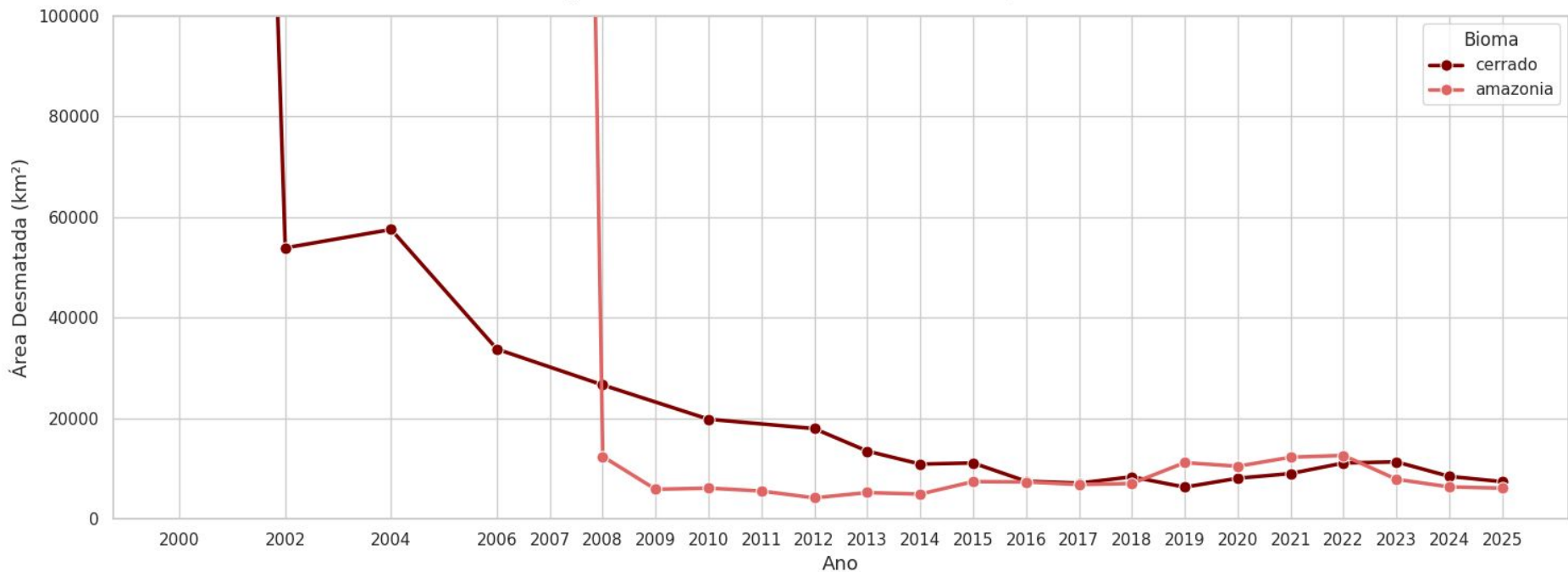
Resultados: Análise Exploratória (DETER - 2018/2022)

DETER - Top 10 Municípios e Unidades de Conservação



Resultados: Análise Exploratória (PRODES - 2018/2022)

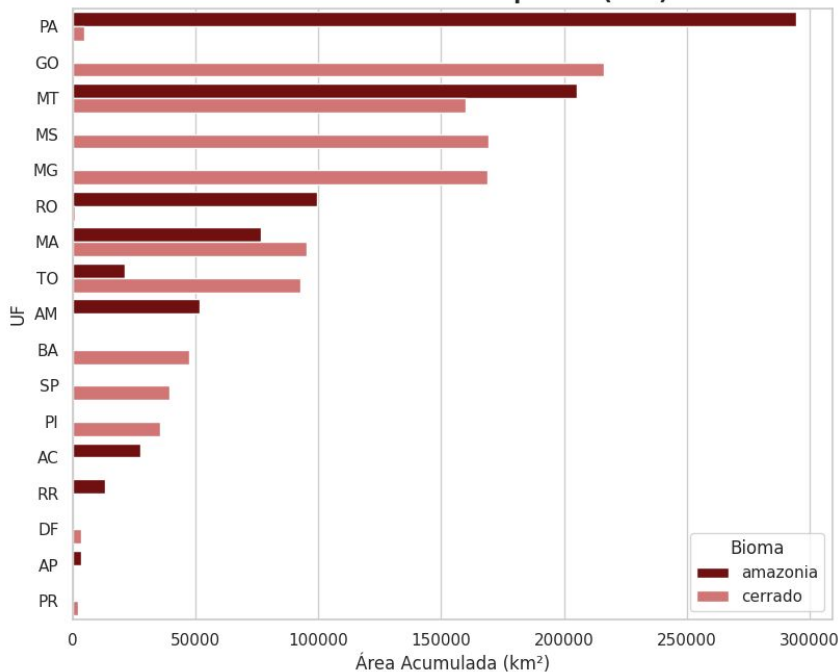
PRODES - Evolução Anual do Desmatamento por Bioma (2018-2022)



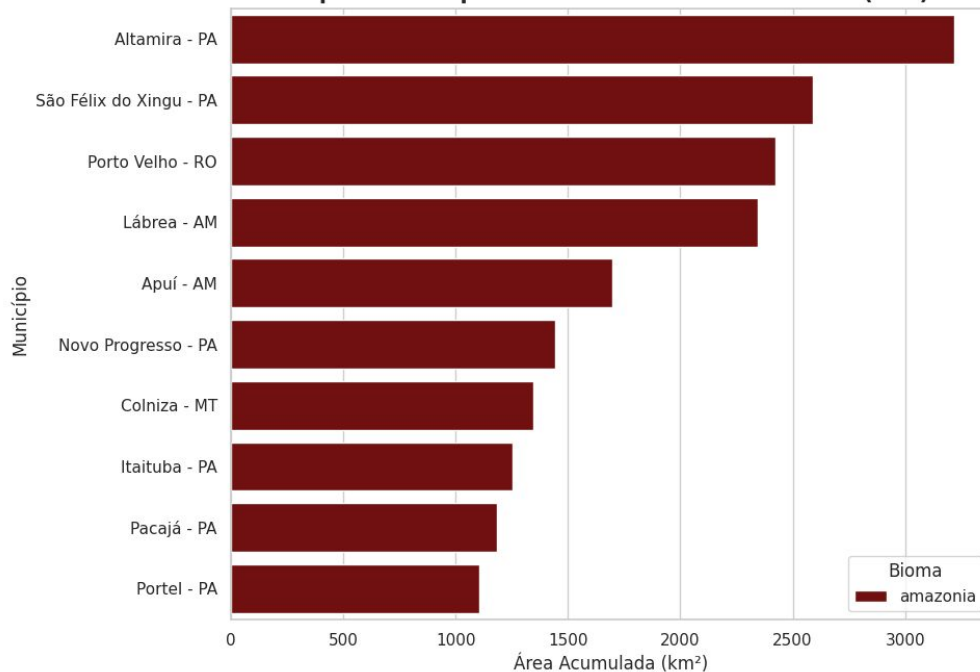
Resultados: Análise Exploratória (PRODES - 2018/2022)

PRODES - Análise Territorial: Estados e Municípios

Área Total Desmatada por UF (km²)

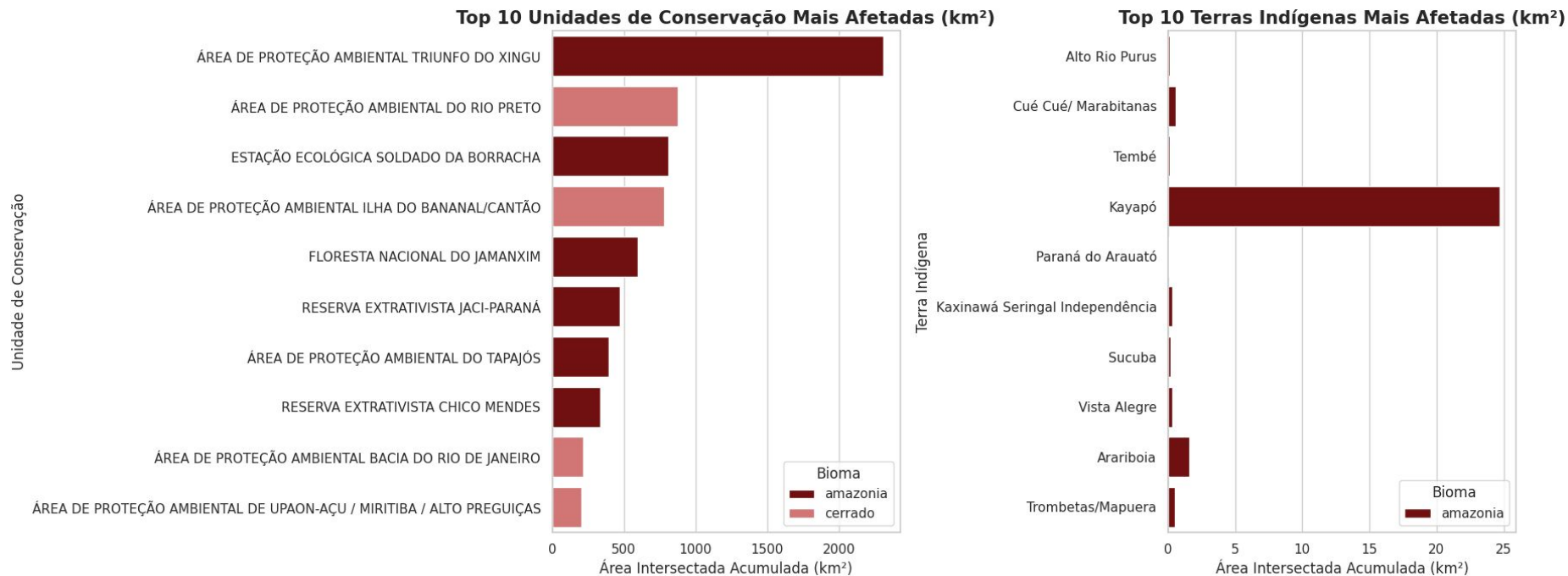


Top 10 Municípios com Maior Área Desmatada (km²)



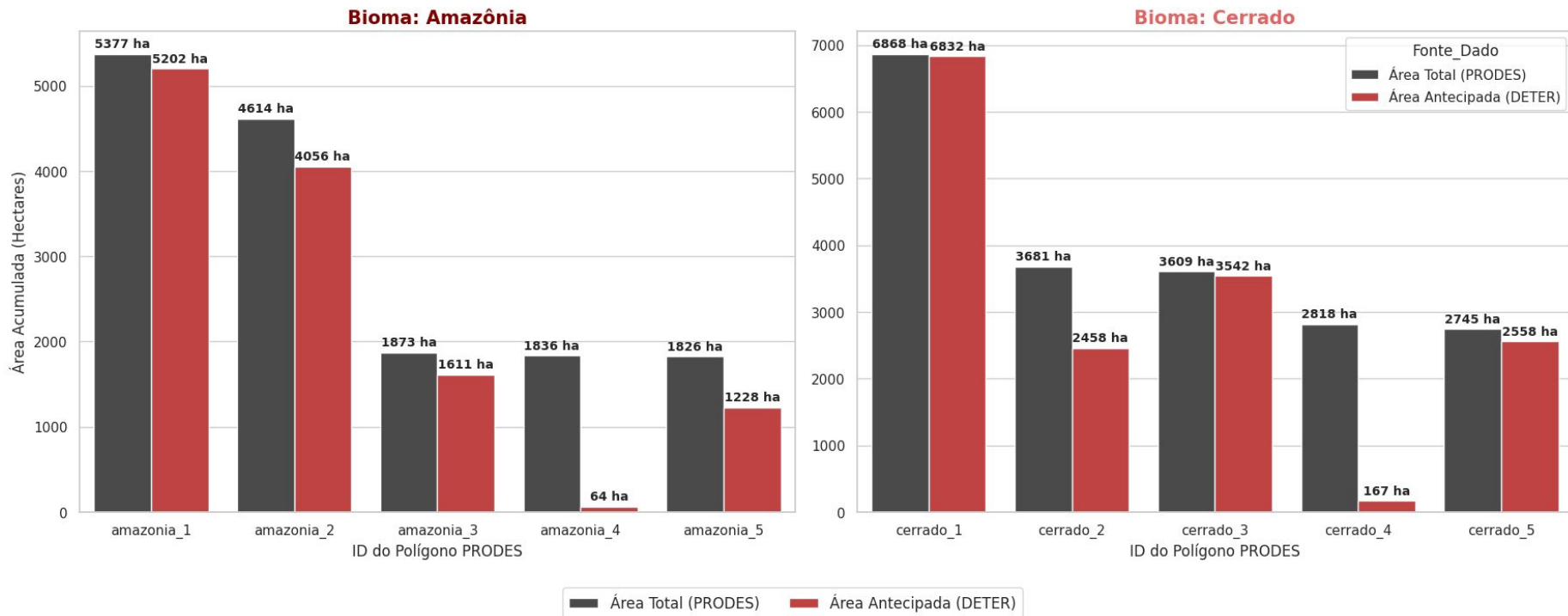
Resultados: Análise Exploratória (PRODES - 2018/2022)

PRODES - Desmatamento em Áreas de Preservação e Demarcação



Resultados: Amostragem para análise de Trajetória de Uso e Ocupação da Terra

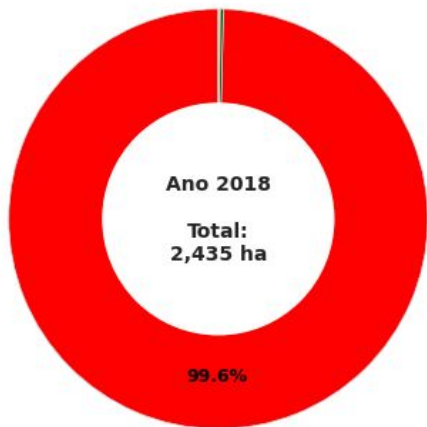
Cruzamento do DETER sobre os Maiores Polígonos do PRODES (2018)



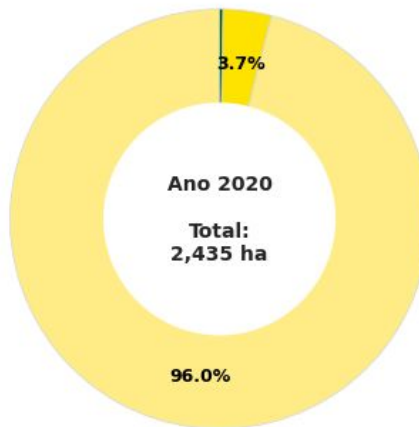
Resultados: Trajetória de Uso e Ocupação da Terra (Cerrado)

Trajatória da Cobertura da Terra - Polígono PRODES: cerrado_1

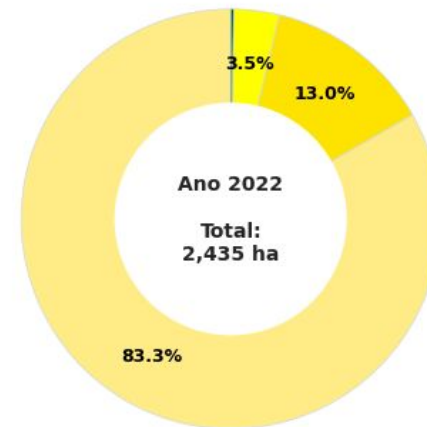
Cenário 2018



Cenário 2020



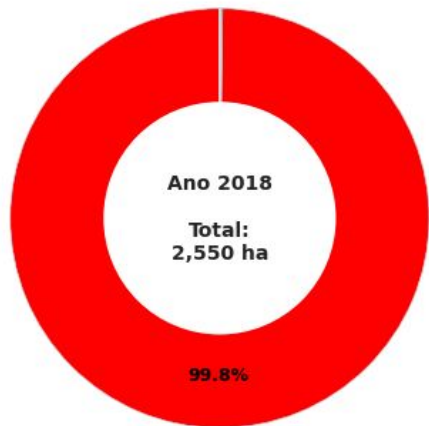
Cenário 2022



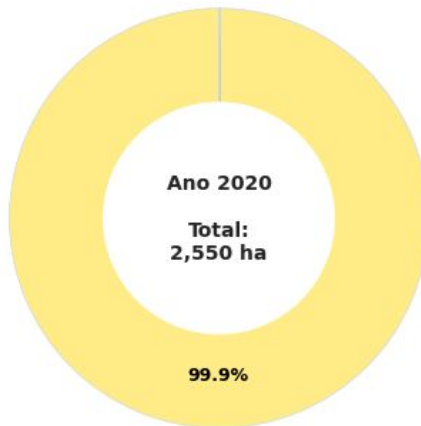
Resultados: Trajetória de Uso e Ocupação da Terra (Cerrado)

Trajetória da Cobertura da Terra - Polígono PRODES: cerrado_2

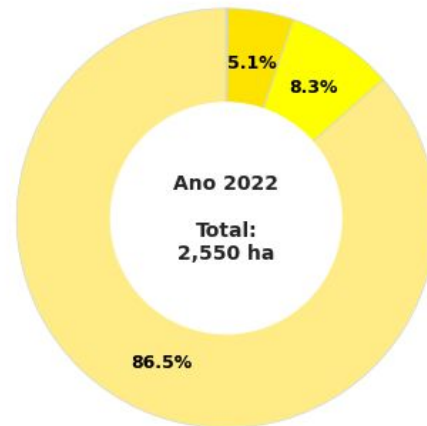
Cenário 2018



Cenário 2020



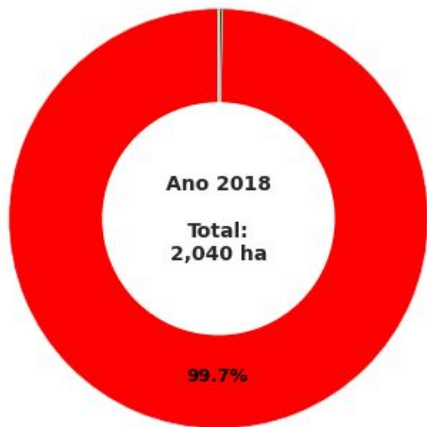
Cenário 2022



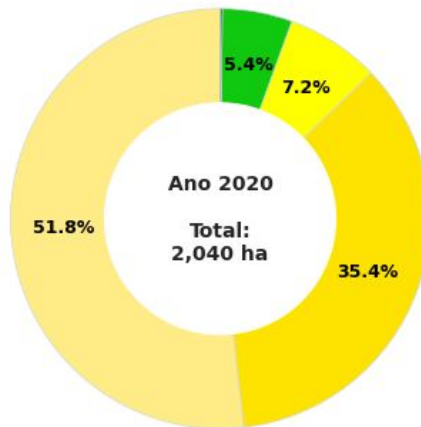
Resultados: Trajetória de Uso e Ocupação da Terra (Cerrado)

Trajetória da Cobertura da Terra - Polígono PRODES: cerrado_3

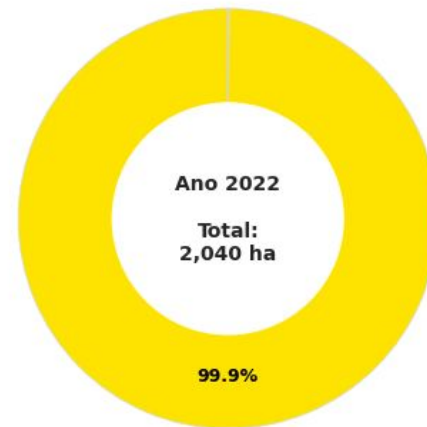
Cenário 2018



Cenário 2020



Cenário 2022



CULTURA AGRICOLA TEMPORARIA DE 1 CICLO
CULTURA AGRICOLA TEMPORARIA DE MAIS DE 1 CICLO

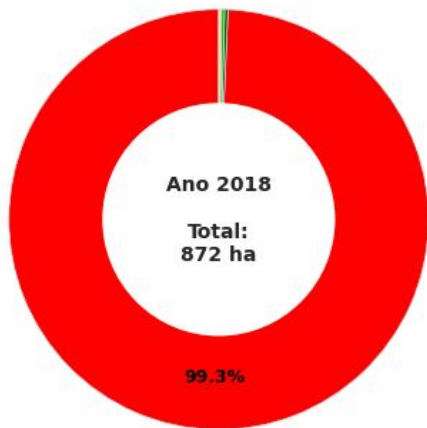
DESFLORESTAMENTO NO ANO
PASTAGEM

VEGETACAO NATURAL PRIMARIA
VEGETACAO NATURAL SECUNDARIA

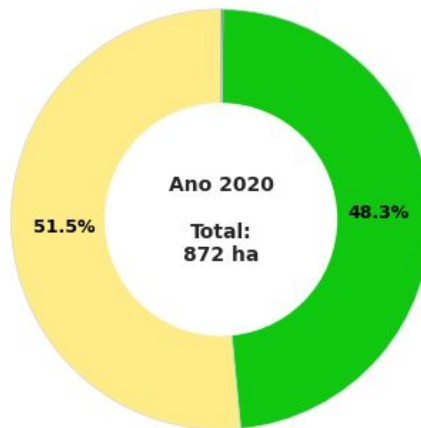
Resultados: Trajetória de Uso e Ocupação da Terra (Cerrado)

Trajetória da Cobertura da Terra - Polígono PRODES: cerrado_5

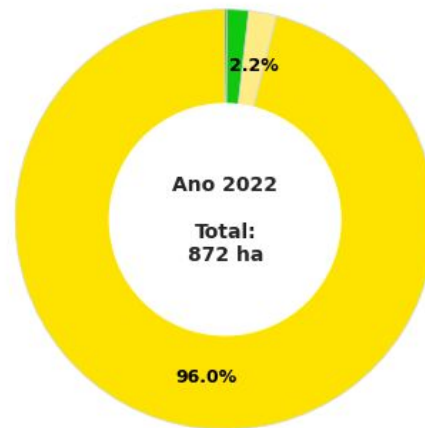
Cenário 2018



Cenário 2020



Cenário 2022



 CULTURA AGRÍCOLA TEMPORÁRIA DE 1 CICLO
 DESFLORESTAMENTO NO ANO

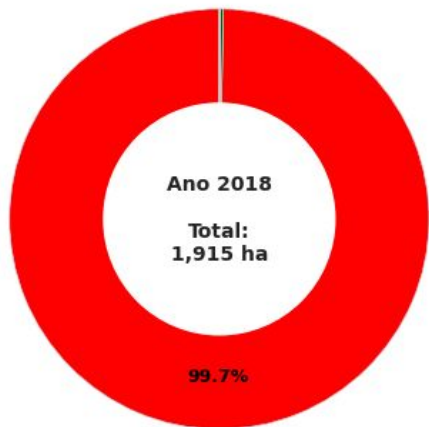
 PASTAGEM
 VEGETAÇÃO NATURAL PRIMÁRIA

 VEGETAÇÃO NATURAL SECUNDÁRIA

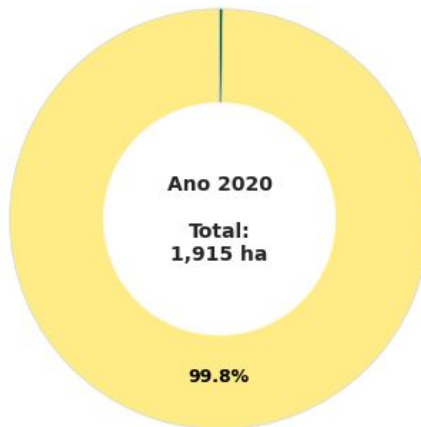
Resultados: Trajetória de Uso e Ocupação da Terra (Cerrado)

Trajetória da Cobertura da Terra - Polígono PRODES: cerrado_4

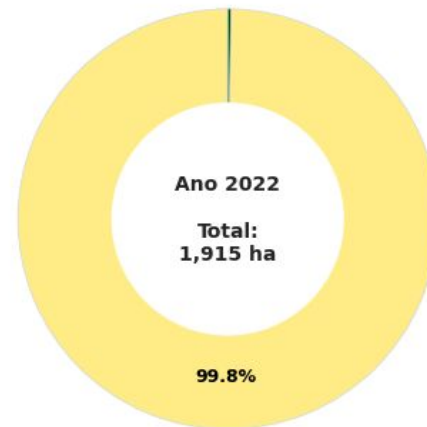
Cenário 2018



Cenário 2020



Cenário 2022

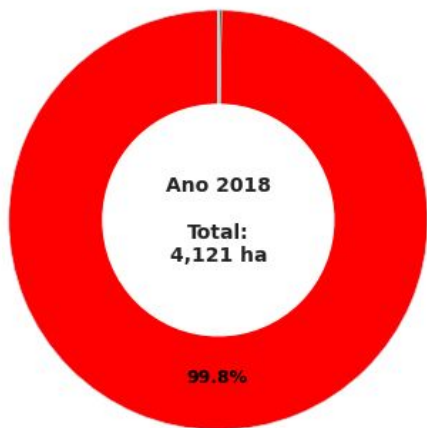


■ DESFLORESTAMENTO NO ANO ■ PASTAGEM ■ VEGETAÇÃO NATURAL PRIMÁRIA

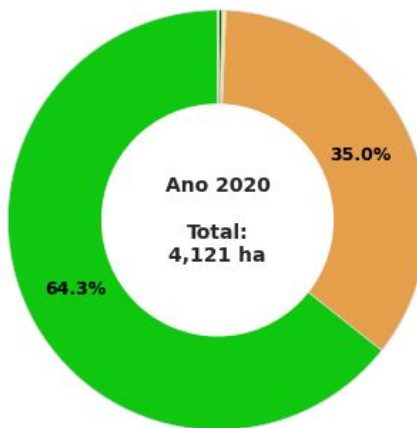
Resultados: Trajetória de Uso e Ocupação da Terra (Amazônia)

Trajetória da Cobertura da Terra - Polígono PRODES: amazonia_1

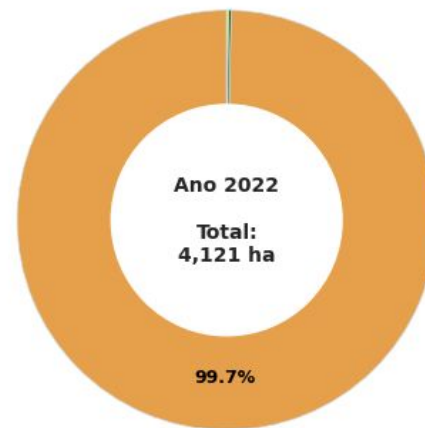
Cenário 2018



Cenário 2020



Cenário 2022



CULTURA AGRICOLA TEMPORARIA DE MAIS DE 1 CICLO
DESFLORESTAMENTO NO ANO

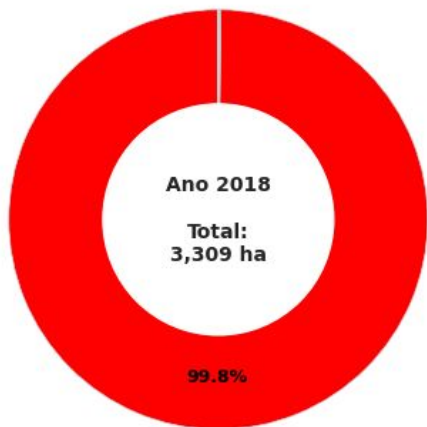
PASTAGEM ARBUSTIVA/ARBOREA
PASTAGEM HERBACEA

VEGETACAO NATURAL FLORESTAL PRIMARIA
VEGETACAO NATURAL FLORESTAL SECUNDARIA

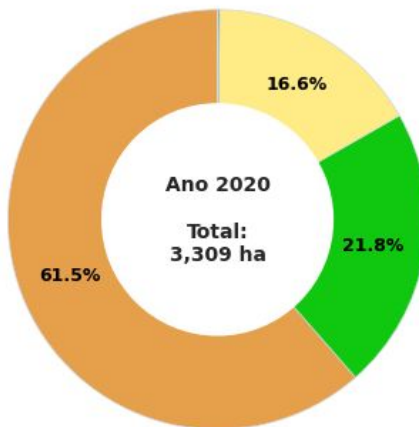
Resultados: Trajetória de Uso e Ocupação da Terra (Amazônia)

Trajetória da Cobertura da Terra - Polígono PRODES: amazonia_2

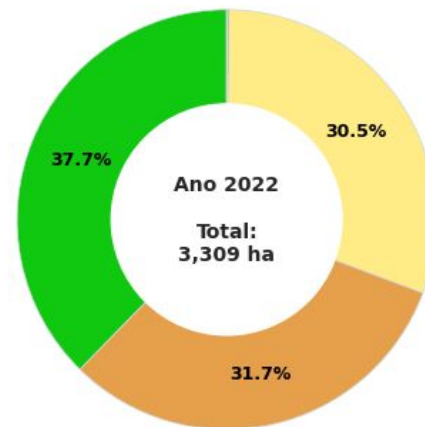
Cenário 2018







Cenário 2020




Cenário 2022



 CORPO DAGUA
 DESFLORESTAMENTO NO ANO
 NATURAL NAO FLORESTAL

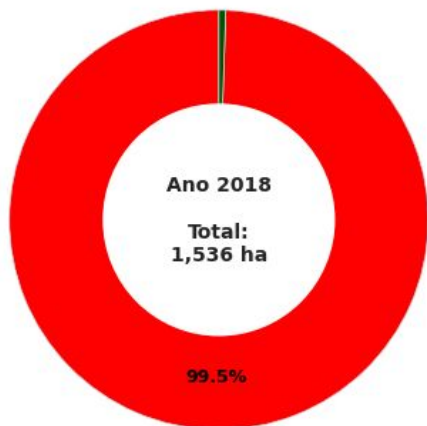
 PASTAGEM ARBUSTIVA/ARBOREA
 PASTAGEM HERBACEA

 VEGETACAO NATURAL FLORESTAL PRIMARIA
 VEGETACAO NATURAL FLORESTAL SECUNDARIA

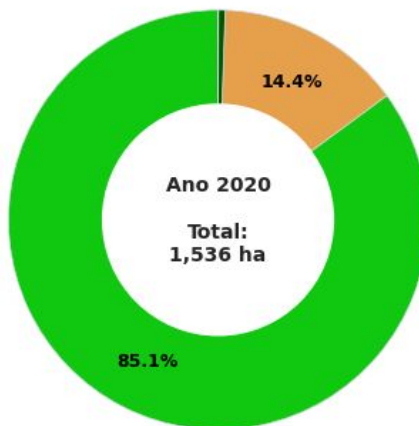
Resultados: Trajetória de Uso e Ocupação da Terra (Amazônia)

Trajetória da Cobertura da Terra - Polígono PRODES: amazonia_3

Cenário 2018



Cenário 2020



Cenário 2022



DESFLORESTAMENTO NO ANO
PASTAGEM ARBUSTIVA/ARBOREA

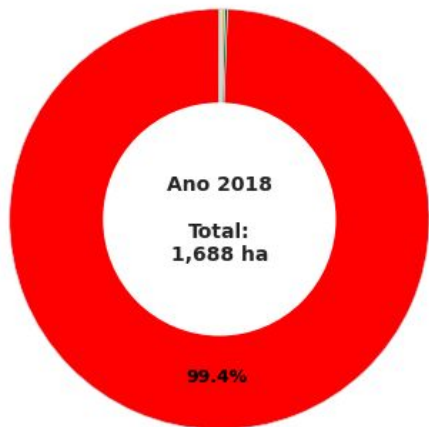
PASTAGEM HERBACEA
VEGETACAO NATURAL FLORESTAL PRIMARIA

VEGETACAO NATURAL FLORESTAL SECUNDARIA

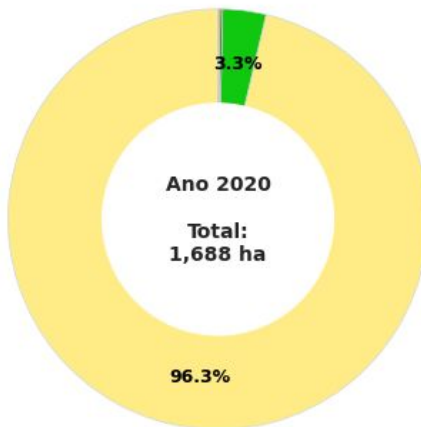
Resultados: Trajetória de Uso e Ocupação da Terra (Amazônia)

Trajetória da Cobertura da Terra - Polígono PRODES: amazonia_4

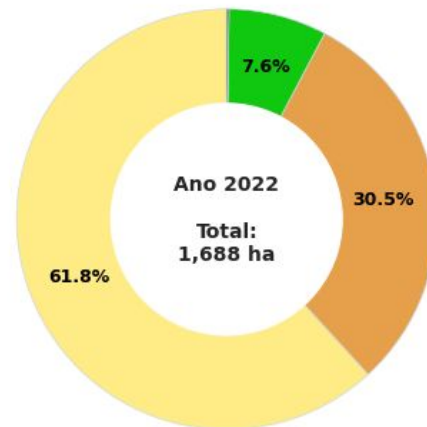
Cenário 2018



Cenário 2020



Cenário 2022



 DESFLORESTAMENTO NO ANO
 PASTAGEM ARBUSTIVA/ARBOREA

 PASTAGEM HERBACEA
 VEGETACAO NATURAL FLORESTAL PRIMARIA

 VEGETACAO NATURAL FLORESTAL SECUNDARIA

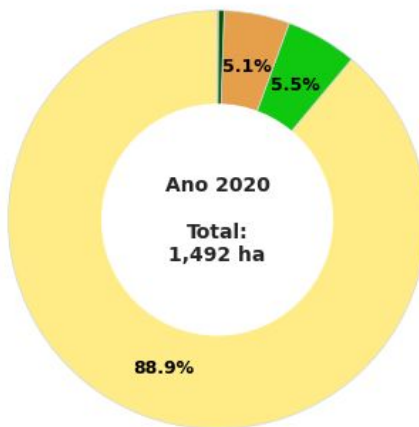
Resultados: Trajetória de Uso e Ocupação da Terra (Amazônia)

Trajetória da Cobertura da Terra - Polígono PRODES: amazonia_5

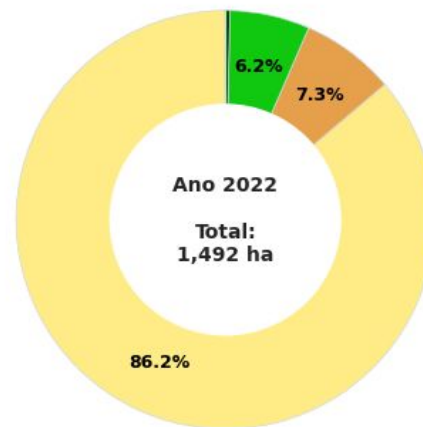
Cenário 2018



Cenário 2020



Cenário 2022



DESFLORESTAMENTO NO ANO
PASTAGEM ARBUSTIVA/ARBOREA

PASTAGEM HERBACEA
VEGETACAO NATURAL FLORESTAL PRIMARIA

VEGETACAO NATURAL FLORESTAL SECUNDARIA

Discussão: Questões Norteadoras do Projeto

Q1

ANÁLISE ARQUITETURAL

Quais os pontos fortes e fracos de uma arquitetura Lakehouse local (single-node) ao processar e unificar dados geoespaciais heterogêneos comparada aos paradigmas tradicionais?

Q2

DINÂMICA DO FENÔMENO

Como os polígonos de desmatamento alteraram seu padrão de uso e cobertura da terra ao longo do tempo (2018, 2020, 2022) nos biomas Amazônia e Cerrado?

Q3

PRÁTICA E COMPARAÇÃO

De que maneira a adoção do ecossistema Lakehouse unificado facilitou ou impôs barreiras para a extração de trajetórias espaço-temporais complexas quando comparada aos fluxos tradicionais de SIG?

Discussão: Perspectivas e Métodos de Avaliação

Validação da Arquitetura

Mapeamento do comportamento do protótipo frente aos 8 requisitos técnicos independentes de fornecedor de Schneider et al. (2024), avaliando desde o armazenamento unificado (R1) até o suporte a processamento híbrido (R8).

Análise Quanti-Qualitativa do Uso do Solo

Cruzamento espacial e temporal de dados do PRODES/DETER com rasters do TerraClass para identificar tendências, sazonalidades e trajetórias de transição ecológica na Amazônia e no Cerrado.

Relato de Experiência Prática

Sistematização dos aprendizados locais, documentando gargalos de desempenho de hardware (nó único) e o nível de maturidade atual das ferramentas de Big Data para o ecossistema geográfico.

Discussão: O Lakehouse Geoespacial

Q1

Quais os pontos fortes e fracos de uma arquitetura Lakehouse local (single-node) ao processar e unificar dados geoespaciais heterogêneos comparada aos paradigmas tradicionais?

- **Armazenamento e Formato Unificados e CRUD (R1, R2):**
 - Unificação realizada no nível de *Storage* (mesmo ambiente físico local acessado pelo Spark);
 - Operações CRUD e processamento incremental;
 - Flexibilização do *Data Format* único (**Delta para vetores** vs. **COG para rasters**) devido à maturidade das ferramentas geoespaciais. Metadados do raster podem ser indexados em tabelas Delta.
- **Representação Relacional, Suporte SQL e Acesso (R3, R4, R7):**
 - Sedona e Spark acessam os arquivos diretamente no disco sem camadas intermediárias como GeoServer (R7);
 - Os dados são expostos em formato relacional (R3) e processados via consultas SQL com extensões espaciais nativas (R4);
- **Consistência, Atomicidade e Isolamento (R5, R6):**
 - O log de transações do Delta garantiu consistência (R5), atomicidade/isolamento (R6) a nível de arquivo no ambiente controlado de nó único (*single-node*).
 - O formato de tabela sozinho não escala concorrência multiusuário. Para cenários reais, a arquitetura **exige um Catálogo de Metadados** (*Unity Catalog, Polaris ou Hive Metastore*) como árbitro central (*Lock Manager*) e governança.
- **Processamento Híbrido Batch/Streaming (R8):**
 - Requisito **não avaliado** no protótipo. O projeto operou estritamente em modo *Batch* devido à natureza histórica e estática dos dados de desmatamento e uso do solo (PRODES/TerraClass).

Discussão: Trajetórias Amazônia vs. Cerrado

Q2

Como os polígonos de desmatamento alteraram seu padrão de uso e cobertura da terra ao longo do tempo (2018, 2020 e 2022) nos biomas Amazônia e Cerrado.

- **Cerrado: A Fronteira Agrícola Acelerada**
 - **Padrão Observado:** Transição rápida e direta
 - Desmatamento contínuo ao longo do ano → Pastagem Temporária → Cultura Agrícola de 1 Ciclo (Soja/Milho).
 - **Contexto na Literatura:** O relevo plano e a facilidade de mecanização do Cerrado atraem o agronegócio extensivo, encurtando o tempo de transição da terra. Destaque histórico para a fronteira agrícola na região do MATOBIBA, mas principalmente para Goiás, Mato Grosso, Mato Grosso do Sul e Minas Gerais.
- **Amazônia: Degradação Progressiva**
 - **Padrão Observado:** Transição progressiva. Sazonalidade no desmatamento (alto em período de seca). Presença marcante de Vegetação Secundária nos primeiros dois anos pós-desmatamento, seguida pela dominância de Pastagem Arbustiva/Arbórea e Herbácea.
 - **Savanização da Amazônia:** A floresta tenta se regenerar (origina a secundária), mas o ciclo recorrente de fogo e o pastoreio extensivo degradam a capacidade de resiliência do bioma, empurrando a área para uma fisionomia mais aberta e savanizada.

Discussão: Os Impactos da Tecnologia (Positivos)

Q3

De que maneira a adoção do ecossistema Lakehouse unificado facilitou ou impôs barreiras para a extração de trajetórias espaço-temporais complexas quando comparada aos fluxos tradicionais de SIG?

- **Vantagens Nativas do Delta Lake sobre Outros Formatos:**
 - **Suporte Geográfico:** Integração nativa com tipos *geometry/geography*, superando concorrentes como o Apache Iceberg (v2), que ainda exigem conversões custosas para WKB/WKT.
 - **Enriquecimento de Metadados:** Uso de *Table Properties* para embutir linhagem e metadados de negócio diretamente no cabeçalho das tabelas.
- **Time Travel e Auditoria Temporal (Change Data Feed):**
 - Capacidade nativa de reconstruir o passado com precisão matemática. Permite responder de forma declarativa a perguntas complexas como: *"Qual era o status de homologação de uma Terra Indígena específica em março de 2025?"*.
- **Arquitetura Cloud-Native Pragmática (R7):**
 - A ingestão direta de rasters em formato **COG** na camada Bronze eliminou a necessidade de servidores de mapas legados (como GeoServer), permitindo que o motor de processamento consumisse o dado bruto direto do armazenamento.

Discussão: Os Impactos da Tecnologia (Negativos)

Q3

De que maneira a adoção do ecossistema Lakehouse unificado facilitou ou impôs barreiras para a extração de trajetórias espaço-temporais complexas quando comparada aos fluxos tradicionais de SIG?

- **O Gargalo do Processamento Vetor-Raster:**
 - O Apache Sedona particionou os vetores com eficiência, mas o Spark tentou tratar a matriz contínua do raster (TerraClass) como linhas tabulares, gerando extrema lentidão e falha de paralelização local.
- **Imaturidade do Ecossistema de Big Data Geográfico (Apache Sedona):**
 - A função `ST_Transform` não suporta projeções cônicas como a **EPSG:5880** - Policônica do Brasil.
 - Incapacidade de Reprojeção de Raster: O Sedona ainda não possui funções maduras para reprojeter rasters.
 - A conversão dos vetores para Albers (EPSG:10857) gerou ruídos matemáticos de ponto flutuante (64 bits), exigindo correções agressivas via `ST_MakeValid` e `ST_Buffer(0)`.
- **Impacto Metodológico no Cálculo de Áreas:**
 - Para contornar a limitação do raster, os Vetores do PRODES foram reprojutados para coordenadas geográficas (EPSG:4674), estimando a área pela contagem de pixels multiplicada pela resolução aproximada (10 m ~ 0.00009°).
 - O cálculo em graus decimais gera distorção por **convergência dos meridianos**, subestimando as áreas à medida que os polígonos se afastam da linha do Equador.

Considerações finais: O Balanço do Protótipo

Por um lado:

- O formato lakehouse ainda é projetado para vetores. A manipulação de rasters no Lakehouse exige contornos cartográficos que ferem o rigor metodológico e aumentam superfície de erros.
- O ambiente *single-node* serve para validação de conceitos, mas encontra limites em álgebra de mapas complexa. O formato Delta, isolado e sem catálogo, não sustenta produção multiusuário concorrente.

Pelo outro:

- A capacidade de unificar dados heterogêneos (PRODES, DETER, TerraClass) em uma única linguagem declarativa (SQL Espacial) com auditoria temporal (*Time Travel*) pode superar os fluxos fragmentados do passado.

Próximos Passos e Futuros Aprimoramentos:

3. Migrar para ambiente multi-nó (*cluster*) na nuvem para testar o real poder de paralelização de rasters.
4. Integrar um catálogo aberto (*Apache Polaris* ou *Unity Catalog*) para validar o controle de acesso e concorrência ACID.
5. Testar a indexação espacial de rasters (padrão STAC) dentro de tabelas Delta para evitar cálculos de área em graus decimais.

Referências

ARMBRUST, M., Ghodsi, A., Xin, R., & Zaharia, M. Lakehouse: a new generation of open platforms that unify data warehousing and advanced analytics. 2021. In Proceedings of CIDR (Vol. 8, No. 1, p. 28).

ERRAMI, A., Hajji H, Ait El Kadi K, Badir H. Spatial big data architecture: from data warehouses and data lakes to the lake-house. J Parall Distrib Comput. 2023;176:70–9. <https://doi.org/10.1016/j.jpdc.2023.02.007>.

JANSSEN, N., Ilayperuma, T., Jayasinghe, J. et al. The evolution of data storage architectures: examining the secure value of the Data Lakehouse. J. of Data, Inf. and Manag. 6, 309–334 (2024). <https://doi.org/10.1007/s42488-024-00132-1>

SCHNEIDER, J., Gröger, C., Lutsch, A. et al. The Lakehouse: State of the Art on Concepts and Technologies. SN COMPUT. SCI. 5, 449 (2024). <https://doi.org/10.1007/s42979-024-02737-0>.

ZIOTI, F.; Ferreira, K. R.; Queiroz, G. R.; Neves, A. K.; Carlos, F. M.; Souza, F. C.; Santos, L. A.; Simoes, R. E. O. A platform for land use and land cover data integration and trajectory analysis. International Journal of Applied Earth Observation and Geoinformation. V 106, P 102655, Feb 2022.