

MAPEAMENTO DE FONTE DE DADOS AGRONÔMICOS

RESUMO

1. Introdução

O setor agrícola brasileiro, um dos mais importantes para a economia nacional, demanda cada vez mais informações precisas e em tempo real para otimizar a produção, mitigar riscos e promover a sustentabilidade. A crescente digitalização do campo impulsiona a busca por dados agronômicos abertos e acessíveis via APIs, que permitem a integração automatizada em sistemas de inteligência. Este relatório responde a essa demanda, identificando e caracterizando as fontes de dados mais relevantes no cenário brasileiro.

A pesquisa detalhada visou identificar o máximo de fontes de dados possíveis, incluindo dados governamentais e de outras instituições e plataformas. Para cada fonte, foram levantadas informações essenciais como nome, tipo de dados oferecidos, método de coleta (API, web scraping, download direto) e links para acesso e documentação. A estrutura resultante visa facilitar a avaliação do volume, qualidade e acessibilidade de cada fonte, elementos fundamentais para o início da coleta de dados em projetos de inteligência agrônômica.

2. Instituto Brasileiro de Geografia e Estatística (IBGE)

O IBGE é uma das principais fontes de estatísticas e informações geocientíficas do Brasil, oferecendo uma gama de dados econômicos, sociais e territoriais.

2.1. Dados e Conteúdo

O IBGE disponibiliza informações econômico-financeiras, de produção, bens e serviços consumidos, e emprego, específicas das atividades da Agricultura, Pecuária, Produção Florestal, Pesca e Aquicultura.¹ Entre os principais produtos relacionados à agricultura, destacam-se:

- **Levantamento Sistemático da Produção Agrícola (LSPA):** Fornece previsões de safras agrícolas dos principais produtos, com revisões mensais.

- **Pesquisas de Estoques, Leite, Ovos de Galinha, Couro e Abate de Animais:** Detalham volumes e características da produção pecuária.
- **Produção da Extração Vegetal e da Silvicultura (PEVS):** Quantidade e valor da produção de florestas plantadas e nativas por município.
- **Pesquisa da Pecuária Municipal (PPM):** Quantidade e valor da produção de origem animal e efetivos das principais espécies por município.
- **Produção Agrícola Municipal (PAM):** Inclui quantidade, área colhida, rendimento médio e valor da produção agrícola por município.
- **Censo Agropecuário:** Abrange número e área de estabelecimentos, atividades agropecuárias e características do produtor.

Além disso, o IBGE oferece dados geográficos como divisões regionais, malhas territoriais e redes geodésicas.

2.2. Acesso e Documentação

- O IBGE oferece acesso programático aos seus dados através da [API SIDRA](#) (Sistema IBGE de Recuperação Automática). Esta API permite a recuperação de conjuntos de dados em formato JSON, ODS e XML. O IBGE também disponibiliza dados para consulta pública na área de Mapas do seu sítio.
- Bibliotecas como [sidrar \(para R\)](#) e [sidrapy \(para Python\)](#) foram desenvolvidas para facilitar a conexão e extração de dados da API SIDRA, oferecendo uma interface flexível para consultas a tabelas, variáveis, períodos e níveis geográficos.
- Para o uso da API, é possível que seja necessário um token de autenticação, seguindo o protocolo OAuth 2.0.
- A ampla gama de dados agrícolas e pecuários do IBGE, combinada com o acesso via API SIDRA, torna-o uma fonte fundamental para análises de tendências de produção, estimativas de safras e caracterização socioeconômica do setor. A granularidade municipal dos dados permite análises regionalizadas, enquanto a atualização mensal de previsões de safra contribui para a tomada de decisões em tempo hábil.

2.3. Precificação

Como o IBGE é uma instituição pública, os dados disponibilizados através do seu Portal de Dados Abertos são considerados públicos e de acesso livre e gratuito. O uso da API SIDRA para consulta e extração de dados não envolve

nenhuma taxa ou cobrança.

2.4. Avaliação de Volume e Qualidade

O IBGE disponibiliza uma extensa gama de dados agronômicos, incluindo:

- Levantamento Sistemático da Produção Agrícola (LSPA)
- Pesquisa da Pecuária Municipal (PPM)
- Produção da Extração Vegetal e da Silvicultura (PEVS)
- Produção Agrícola Municipal (PAM)
- Censo Agropecuário

Esses dados abrangem informações sobre produção, área colhida, rendimento médio, valor da produção, efetivos de animais, práticas agrícolas, entre outros aspectos.

2.4.1. Qualidade

O IBGE adota rigorosos padrões metodológicos e técnicos para garantir a confiabilidade de seus dados. A qualidade é assegurada por meio de:

- Metodologias padronizadas
- Processos de validação e verificação
- Adoção de normas internacionais

Além disso, o IBGE publica relatórios técnicos e notas metodológicas que detalham os processos de coleta e tratamento dos dados, permitindo uma avaliação crítica por parte dos usuários.

2.5. Considerações para Coleta de Dados e Inteligência

- **Metodologia Amostral:** O IBGE utiliza métodos amostrais e censitários, com amostras representativas por município, garantindo a precisão dos dados.
- **Periodicidade:** Pesquisas como o LSPA são realizadas mensalmente,

proporcionando atualizações frequentes.

- **Cobertura Geográfica:** Os dados são coletados em nível municipal, permitindo análises detalhadas por região.

2.6. Desafios e Oportunidades

2.6.1. Desafios:

- **Orçamentários:** Cortes no orçamento podem afetar a realização de grandes operações, como censos e pesquisas de larga escala.
- **Cobertura Completa:** Garantir a cobertura de todos os estabelecimentos agropecuários, especialmente em áreas remotas, pode ser desafiador.

2.6.2. Oportunidades:

- **Integração de Dados:** A colaboração com outras instituições e a integração de diferentes bases de dados podem enriquecer as análises e ampliar a compreensão do setor agropecuário.
- **Tecnologia e Inovação:** O uso de tecnologias emergentes, como inteligência artificial e aprendizado de máquina, pode aprimorar a análise e previsão de dados.
- **Abertura de Dados:** A contínua promoção da cultura de dados abertos facilita o acesso e o uso das informações pela sociedade e pesquisadores.

2.7. Recomendações para Integração

- **Utilização de APIs:** Integrar os dados do IBGE com outras fontes por meio de APIs facilita a atualização e a consistência das informações.
- **Padronização de Metadados:** Adotar padrões de metadados, como o Perfil de Metadados Geoespaciais do Brasil (Perfil MGB), assegura a interoperabilidade entre diferentes sistemas.
- **Capacitação:** Investir em treinamento e capacitação de profissionais para o uso eficiente das ferramentas e dados disponíveis.
- **Parcerias Institucionais:** Estabelecer parcerias com outras instituições

de pesquisa e governo para ampliar a cobertura e a qualidade dos dados.

3. Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA)

A EMBRAPA, como principal instituição de pesquisa agropecuária do Brasil, tem investido na disponibilização de suas vastas informações e modelos por meio de APIs.

A EMBRAPA consolidou suas APIs na [Plataforma AgroAPI](#), que visa o mercado de tecnologias de agricultura digital. Esta plataforma oferece informações e modelos gerados pela Embrapa e seus parceiros, abrangendo diversas áreas da agronomia. Entre as APIs disponíveis, destacam-se:

3.1. Acesso e Documentação

- Para começar a usar as APIs, [crie uma conta na plataforma AgroAPI](#). Assim, você pode escolher as APIs do seu interesse e criar aplicações para acessá-las.
- Depois de gerar os tokens de acesso para as aplicações, devem ser escolhidas as APIs que serão utilizadas. [Este tutorial irá auxiliá-lo na execução destes passos](#).

3.2. Precificação

O acesso às APIs freemium (Agritec, ClimAPI e SATVeg) da plataforma AgroAPI é gratuito por 1 mês para o máximo de 1.000 requisições por API. Após o consumo das 1.000 requisições ou o término do período de 1 mês (o que ocorrer primeiro), o acesso será interrompido e você só poderá continuar utilizando a API mediante assinatura de contrato para pagamento. Quer saber mais? Entre em contato conosco pelo e-mail agroapi@embrapa.br.

O acesso às APIs: Agrofit, Agrotermos, Bioinsumos, Blue Star Sting, Responde Agro, SmartSolos Expert e PlantAnnot é gratuito para um volume máximo de 100 mil requisições por mês, sem necessidade de assinatura de contrato.

3.3. AgriTec API

3.3.1 Dados e Conteúdo

A **API Agritec** fornece informações e modelos para gerenciamento da produção agrícola, especialmente para culturas como arroz, feijão, milho, soja e trigo.

Inclui dados como:

- Época ideal de plantio com base no Zoneamento Agrícola de Risco Climático (ZARC).
- Cultivares mais aptas para 12 culturas (arroz, algodão, amendoim, cevada, feijão, feijão caupi, girassol, mamona, milho, soja, sorgo e trigo).
- Recomendações de adubação e correção de solo conforme análise prévia, previsões de produtividade e condições climáticas.
- Estimativas de produtividade para arroz, feijão, milho, soja e trigo baseadas em modelos empíricos regionalizados e dados agro-meteorológicos municipais.
- Informações de balanço hídrico e condições climáticas antes e durante a safra.

3.3.2. Avaliação de Volume e Qualidade

3.3.2.1. Volume

A API oferece dados abrangentes para diversas culturas e sistemas de produção, com informações por decêndio para épocas de plantio. As informações de produtividade e balanço hídrico são vetores de dados diários.

3.3.2.2. Qualidade

Os dados são baseados em metodologias confiáveis, como o Zoneamento Agrícola de Risco Climático (Zarc) , o Registro Nacional de Cultivares (RNC) do Ministério da Agricultura, Pecuária e Abastecimento , e modelos empíricos ajustados para cada região do Brasil.

3.3.3. Considerações para Coleta de Dados e Inteligência

- As informações da API podem ser usadas em soluções que auxiliam a tomada de decisões no planejamento, monitoramento e gerenciamento da produção agrícola.

- A API beneficia agricultores, cooperativas, assistência técnica, bancos e seguradoras.
- Os dados de zoneamento de risco climático (Zarc) indicam a época ideal de plantio, com probabilidades de perda de produção de 20%, 30% e 40%.
- A funcionalidade de recomendação de adubação considera parâmetros de análise de solo e produtividade almejada para calcular a formulação e a quantidade de fertilizante.

3.3.4. Desafios e Oportunidades

3.3.4.1. Desafios

- A API está atualmente com status **DEPRECATED** desde agosto de 2024, isso significa que pode haver restrições de uso ou migração.
- Limitação de uso gratuito (ex. 1.000 requisições em 1 mês) exige planejamento para volume de consumo.
- Falta de detalhes abertos sobre frequência de atualização de dados, granularidade temporal ou espacial específica.

3.3.4.2. Oportunidades

- Dados amplamente validados e integráveis com sistemas de análise de risco, agroclima e planejamento de cultivo.
- Permite criação de soluções com impacto real, especialmente voltadas para crédito agrícola, seguro rural e monitoramento produtivo.
- Facilita desenvolvimento ágil de aplicações digitais (web ou mobile), ex: Zarc Plantio Certo

3.3.5. Recomendações para Integração

- **Verificar versão atual e roadmap:** como o status indica “deprecated” (parcialmente ou totalmente descontinuada desde 04/Ago/2024), é essencial confirmar com a Embrapa se há versões substitutivas ou migratórias, ou previsão de desativação definitiva.
- **Planejar volume de uso:** comece com conta de teste (freemium) e estimar volume de requisições conforme demanda. Para produção, considere assinatura via contrato, negociado com a Embrapa.

- **Modularizar integração:** separar chamadas por funcionalidade (ex: épocas de plantio, produtividade, adubação) para otimizar uso e facilitar manutenção.
 - **Criação de camada de caching:** para evitar chamadas repetidas, melhorar performance e reduzir custos e latências em produção.
 - **Monitorar qualidade e versionamento:** acompanhar documentos de release, changelogs ou comunicados da Embrapa sobre atualizações ou alterações na API.
 - **Fallback ou Redundância:** considerando o possível desligamento ou instabilidade da API, avalie integração com dados locais, outras APIs (como ClimAPI para clima ou SATVeg para vegetação), ou replicação offline conforme necessidade.
 - **Validação de dados e testes periódicos:** verificar consistência dos resultados, principalmente em regiões com menor cobertura amostral ou atualizações climáticas recentes.
-

3.4. Clim API

3.4.1 Dados e Conteúdo

A **ClimAPI** provê **dados agrometeorológicos de previsão do tempo** para todo o território nacional, sendo ideal para agricultura de precisão e suporte ao planejamento, decisão e monitoramento agrícola.

Contempla **17 variáveis climáticas**, entre elas:

- Temperatura do ponto de orvalho a 2m.
- Precipitação total.
- Velocidade e componentes do vento.
- Cobertura de nuvens (alta, média e baixa).
- Evaporação potencial.
- Umidade relativa e umidade volumétrica do solo (0-10cm, 10-40cm, 40cm-1m).
- Temperaturas máxima, mínima e da superfície.
- Duração da luz solar.

Os dados são processados com base no **Global Forecast System (GFS)** da **NOAA**, e disponibilizados com maior precisão espacial e temporal (latitude/longitude e áreas de ~25 km).

Atualizados a cada 6 horas, fornecendo **quatro atualizações diárias**, com previsões de **até 10 dias à frente** e histórico de **dados para os 10 dias anteriores**.

3.4.2. Avaliação de Volume e Qualidade

3.4.2.1. Volume

A API oferece dados atualizados a cada seis horas para um período de 10 dias para frente e 10 dias para trás. A abrangência é para todo o território nacional, com dados relativos a uma área de 25Km por latitude e longitude.

3.4.2.2. Qualidade

A maior precisão da ClimAPI está relacionada ao uso de dados gerados por latitude e longitude e à sua atualização frequente, a cada seis horas. A API utiliza dados do Global Forecast System (GFS) da NOAA.

3.4.3. Considerações para Coleta de Dados e Inteligência

- A ClimAPI pode fornecer dados **separados ou agregados**, facilmente integráveis em modelos ou sistemas por meio de equações, dashboards, apps ou websites.
- Permite apoio no **planejamento agrícola**, identificando **janelas ativas de drones, plantio, pulverização, colheita**, além de apoio logístico com base nas previsões climáticas.
- O acesso é via API RESTful com **token de acesso** obtido ao se cadastrar na **Plataforma AgroAPI**; há ambiente de sandbox e produção disponíveis.
- O formato padrão de resposta facilita a incorporação em soluções digitais, incluindo apps mobile.

3.4.4. Desafios e Oportunidades

3.4.4.1. Desafios

- O plano gratuito é limitado a **1.000 requisições durante o primeiro mês**, após o qual é necessário contratar um plano pago ao exceder o limite ou término do período.

- Não há esclarecimento público sobre detalhes como **latência**, **limites de taxa**, ou **nível de SLA**, exigindo testes em ambiente real.

3.4.4.2. Oportunidades

- Permite **rapidez no desenvolvimento** de soluções agrodigitais, diminuindo necessidade de processamento próprio de dados climáticos brutos.
- Ideal para casos de uso como **alertas agroclimáticos**, **manejo de irrigação**, **previsão de geadas**, **logística**, entre outros.
- Alta compatibilidade com apps de agricultura de precisão, startups, cooperativas e instituições de extensão rural.
- Uso estruturado e regionalizado (cobertura de todos os biomas e estados do Brasil) favorece soluções de escala nacional com customização local.

3.4.5. Recomendações para Integração

- **Planejar o volume de requisições**: iniciar com o período gratuito (1.000 requisições) e estimar demanda futura, optando por contrato pago conforme escalabilidade.
- **Implementar caching inteligente**: para evitar chamadas repetitivas, otimizar performance e evitar custos extras.
- **Modularizar uso da API**: separar chamadas conforme tipo de variável (ex: precipitação, temperatura, vento) e necessidade de período (histórico vs. previsão).
- **Monitorar e testar**: testar a qualidade dos dados, checar frequência de atualização, latência e consistência conforme localização e data.
- **Integrar com outras fontes** (como Agritec, SATVeg) para combinar dados climáticos com outros insumos agronômicos, enriquecendo Insights.
- **Criar fallback**: em caso de indisponibilidade ou limitações, considerar integração alternativa com fontes climáticas públicas ou sistemas locais.
- **Validar licenciamento**: seguir condições de uso, especialmente sobre direitos autorais da NOAA — não reivindicar propriedade ou

endosso, nem indicar como oficial — após manipulações dos dados.

- **Monitorar contratos e versões:** manter contato com a equipe da Embrapa, acompanhar mudanças na API ou políticas de acesso via e-mails de suporte técnico/Biz (como cnptia.climapi@embrapa.br)
-

3.5. SATVeg API

3.5.1 Dados e Conteúdo

A **SATVeg API** é derivada do Sistema de Análise Temporal da Vegetação (SATVeg), uma ferramenta web desenvolvida pela Embrapa Agricultura Digital. Ela oferece:

- **Curvas temporais dos índices vegetativos** NDVI (Normalized Difference Vegetation Index) e EVI (Enhanced Vegetation Index), com base em imagens do sensor MODIS (produtos MOD13Q1 e MYD13Q1) dos satélites Terra e Aqua.
- **Cobertura geográfica ampla**, abrangendo o Brasil e toda a América do Sul.
- Dados produzidos **desde o ano 2000 até a data mais recente disponível**, com resolução **temporal de 16 dias** e espacial de **250 metros**.

3.5.2. Avaliação de Volume e Qualidade

3.5.2.1. Volume

A API oferece uma série temporal completa de dados desde fevereiro de 2000 (para o satélite Terra) e julho de 2002 (para o satélite Aqua). Os dados são de alta resolução, com pixels de 250m x 250m, equivalentes a 6,25 hectares.

3.5.2.2. Qualidade

Os dados são baseados nas imagens multiespectrais dos satélites Terra e Aqua da NASA e no repositório oficial da LP DAAC (Land Processes Distributed Active Archive Center). A qualidade radiométrica do MODIS é de 12 bits, com 4.096 níveis de cinza.

3.5.3. Considerações para Coleta de Dados e Inteligência

- A API permite consultas automáticas ao banco de dados geográfico SATVeg, via **requisições espaciais** informando coordenadas geográficas (ponto ou polígono) de interesse.
- Inclui também funcionalidades de **pré-filtragem e suavização** das séries temporais (remoção de dados faltantes ou impactados por nuvens), utilizando métodos como FlatBottom, Wavelet–Coiflet4 e Savitzky–Golay.
- Usuários não precisam utilizar a interface gráfica da web; os dados são entregues via **API RESTful**, facilitando a integração automática em aplicações, dashboards e sistemas analíticos proprietários.

3.5.4. Desafios e Oportunidades

3.5.4.1. Desafios

- A API foi **deprecated** em 27 de setembro de 2024, conforme indicado na plataforma AgroAPI, o que pode exigir que os usuários considerem alternativas futuras ou migrações.
- Informações específicas como volume de dados acessíveis, limites de taxa (rate limits), SLA ou detalhes de licenciamento não estão amplamente divulgados publicamente.

3.5.4.2. Oportunidades

- **Monitoramento temporal da vegetação**, identificando ciclos de cultivo, vigor e biomassa.
- **Gestão territorial e ambiental**, incluindo análise de uso da terra, detecção de supressão vegetal e apoio a políticas públicas.
- **Setor financeiro e agrícola**, trazendo insights para seguro agrícola, crédito rural, análise de risco e avaliação histórica de produtividade.

3.5.5. Recomendações para Integração

- **Verificar a continuidade da API ou alternativas** caso o status deprecated se mantenha, contatar a equipe técnica da Embrapa para roadmap ou serviços substitutos.
 - **Planejar o uso sob a limitação freemium**, que, segundo a plataforma, oferece até 1.000 requisições no primeiro mês por API (igual ao Agritec e ClimAPI), após isso, é necessária assinatura.
 - **Implementar caching local** das séries temporais para reduzir chamadas repetidas, melhorar performance e controlar custos.
 - **Pré-processar dados via API** (suavização e limpeza) já que a API oferece filtros internos — isso simplifica a lógica no cliente final.
 - **Modularizar chamadas** por região e intervalo (ponto vs polígono), indexando resultados localmente para facilitar comparativos históricos ou análises espaciais.
 - **Combinar com outras APIs da AgroAPI** (como Agritec e ClimAPI) para enriquecer análises com clima, risco e vegetação — gerando insights mais completos.
 - **Monitorar contratos e versões** da API, mantendo contato via emails como agroapi@embrapa.br ou cnptia.satveg@embrapa.br para atualizações.
 - **Testar a qualidade dos dados output**, verificando alinhamento com dados de campo, imagens recentes e validação local, especialmente em regiões críticas.
-

3.6. AgroFit API

3.6.1 Dados e Conteúdo

A **Agrofit API** disponibiliza um **banco de dados completo** sobre produtos fitossanitários registrados no Brasil. Isso inclui defensivos agrícolas, agrotóxicos, produtos fitossanitários, bioinsumos e outros registros do Ministério da Agricultura e Pecuária (MAPA), com informações complementares da ANVISA (Ministério da Saúde) e IBAMA (Ministério do Meio Ambiente).

Por meio dela, é possível consultar:

- Produtos formulados e técnicos;
- Pragas (insetos e doenças) e plantas daninhas;
- Ingredientes ativos, modo de ação, formulação, e técnica de aplicação;
- Titular do registro, marca comercial;
- Classificações toxicológicas, ambientais, indicação para agricultura orgânica, inflamabilidade, corrosividade;
- Documentação associada (data de inclusão, tipo de documento e link).

3.6.2. Avaliação de Volume e Qualidade

3.6.2.1. Volume

A API oferece um banco de dados completo de todos os produtos registrados no Mapa. Inclui listagens e mecanismos de busca com filtros.

3.6.2.2. Qualidade

Os dados são obtidos diretamente do Ministério da Agricultura e Pecuária (Mapa) e são considerados "corretos e seguros". A API contribui para evitar o uso inadequado de agrotóxicos.

3.6.3. Considerações para Coleta de Dados e Inteligência

- A API é acessada via **RESTful API**, com autenticação por **token** gerado na plataforma AgroAPI, que oferece endpoints tanto para sandbox quanto produção.
- Consultas podem ser feitas utilizando filtros como praga, cultura, marca, titular, ingrediente ativo, classificação, entre outros, o que favorece flexibilidade e precisão nas integrações.
- Disponível em ambiente gratuito (até 100 mil requisições por mês), sem necessidade de contrato, ideal para prototipagem, POCs, apps e startups.

3.6.4. Desafios e Oportunidades

3.6.4.1. Desafios

- Embora tenha excelente cobertura de dados, a documentação técnica completa (como formatos de

resposta, exemplos de payload, limites de taxa, SLA) não está amplamente acessível, recomenda-se consultar diretamente na plataforma AgroAPI ou contatar suporte técnico.

- Possivelmente, estão sujeitos a atualizações regulatórias e operacionais que exigem acompanhamento ativo.

3.6.4.2. Oportunidades

- **Alta disponibilidade e volume gratuito** (100 mil requisições/mês) oferecem excelente custo-benefício para desenvolvimento de soluções digitais no agronegócio.
- Dados essenciais para **gestão fitossanitária**, escolha de defensivos seguros, conformidade com limites máximos de resíduos (LMR), prevenção de resistências e apoio a políticas agrícolas.

Relevante para aplicações como **assistentes de escolha de produtos, credenciamento técnico, inteligência agrícola ou aplicações de extensão rural**.

3.6.5. Recomendações para Integração

- **Registrar-se na plataforma AgroAPI**, gerar o **token de acesso** e testar no sandbox antes da produção.
- **Planejar volume de requisições**, começando com o patamar gratuito de 100 k/mês; se necessário, contatar agroapi@embrapa.br para ampliar o limite.
- **Implementar caching local** (por ex., armazenar consultas por produto ou classificação) para melhorar performance e evitar sobrecarga ou repetição de chamadas.
- **Modularizar chamadas** por filtro (por exemplo: por cultura, ingrediente ativo, classificação ambiental) para otimizar uso e clareza nas integrações.
- **Inspecionar os formatos de dados retornados** e validar em conformidade com requisitos de sistemas clientes.
- **Combinar com outras APIs da AgroAPI** (Agritec, ClimAPI, SATVeg) para enriquecer funcionalidades em plataformas de tomada de decisão agrícola.
- **Monitorar versões e atualizações** do Agrofit API, manter contato técnico por cnptia.agrofit@embrapa.br.

- **Garantir compliance com regulamentações**, especialmente na recomendação de uso; evitar sugestões que contrariem normas de saúde ou ambientais.
-

3.7. AgroTermos API

3.7.1 Dados e Conteúdo

- A **AgroTermos API** permite o acesso ao conteúdo completo do *AgroTermos*, que é uma plataforma de organização e qualificação terminológica e conceitual do conhecimento agropecuário, baseada em vocabulários controlados robustos como **Thesagro** e **Agrovoc**. Atualmente, reúne **cerca de 55 mil termos** — embora outra fonte oficial mencione “mais de 248 mil termos”.
- A API viabiliza **consultas automatizadas** ao banco terminológico, funcionando como uma ponte máquina-a-máquina, sem intervenção manual.
- São disponibilizados **seis serviços principais** que atendem diferentes formas de consulta terminológica e relacional:
 - Consulta Exata de Termos/Conceitos;
 - Consulta Parcial (por substring);
 - Consulta por ID;
 - Consulta de Termo com seus Relacionamentos;
 - Consulta de Todas as Relações;
 - Consulta de Relação Exata.

3.7.2. Avaliação de Volume e Qualidade

3.7.2.1. Volume

A plataforma Agrotermos é composta por cerca de 55 mil termos. A API oferece 6 serviços de consulta para acessar este vasto conteúdo

3.7.2.2. Qualidade

A plataforma utiliza Engenharia da Informação, PNL, Linguística de Corpus e modelagem semântica para organizar o conhecimento. Os termos são provenientes de tesouros agrícolas em português como o Thesagro e o Agrovoc.

3.7.3. Considerações para Coleta de Dados e Inteligência

- O acesso à API é realizado por meio da **plataforma AgroAPI**, com autenticação via token, disponível em ambientes **sandbox e de produção**.
- A API favorece **integração automática em aplicações digitais**, especialmente em sistemas que realizam **Processamento de Linguagem Natural (PLN)**, **mineração de texto**, **indexação semântica**, **aprendizado de máquina**, **web semântica e gestão de conhecimento** em contexto agropecuário.

3.7.4. Desafios e Oportunidades

3.7.4.1. Desafios

- Existem discrepâncias no número de termos reportados (55 mil vs. 248 mil), o que recomenda uma revisão direta na fonte ou na documentação técnica.
- A documentação pública carece de informações detalhadas sobre **exemplos de payload**, **formatos de resposta**, **limites de taxa (rate limits)** ou **SLA**. Esses detalhes precisam ser verificados diretamente via plataforma ou contato técnico.

3.7.4.2. Oportunidades

- A API abre caminho para enriquecer projetos com **vocabularização enriquecida**, favorecendo busca semântica, correção automática, recomendações inteligentes e interoperabilidade de dados agropecuários.
- É altamente adequada para sistemas de **IA aplicada ao agronegócio**, apoiando startups, instituições de pesquisa, sistemas de extensão ou repositórios de dados com metadados contextuais semânticos.

3.7.5. Recomendações para Integração

- **Registrar-se na plataforma AgroAPI** e gerar o **token de acesso**, realizando a subscrição no sandbox antes de migrar para produção.
- Identificar o volume real e a qualidade dos termos diretamente na API, verificando a discrepância entre os valores reportados.

- **Implementar cache local** para consultas recorrentes, especialmente naquelas com padrões de uso previsíveis (ex.: consultas por termo ou relacionamentos comuns).
 - **Modularizar chamadas** conforme necessidade: exata, parcial, por ID ou relacional, para otimizar performance e clareza no consumo.
 - **Validar o retorno dos serviços** e mapear formatos de resposta para modelagem em sistemas downstream (ex.: dashboards, modelos de PLN, bases de conhecimento).
 - **Combinar a AgroTermos com outras APIs da AgroAPI** (como Agritec, ClimAPI, Agrofit, SATVeg) para enriquecer dados com camadas semânticas e terminológicas.
 - Manter contato com o suporte técnico da Embrapa, especialmente com os proprietários **gtermos@embrapa.br** (negócios) e **leandro.oliveira@embrapa.br** (técnico), conforme listado na documentação.
 - Acompanhar atualizações da versão (atualizada em 11 de setembro de 2024) e sinalizações de novas funcionalidades ou documentação de uso.
-

3.8. Bioinsumos API

3.8.1 Dados e Conteúdo

- A API Bioinsumos reúne informações técnicas sobre bioinsumos registrados no Ministério da Agricultura e Pecuária (MAPA).
- As categorias de bioinsumos incluídas são "inoculante" e "produtos para controle de pragas".
- A API fornece uma listagem completa de produtos biológicos para controle de pragas, doenças, plantas daninhas e inoculantes.
- A listagem de produtos biológicos para controle de pragas inclui informações como número de registro, marca comercial, titular do registro, se é bioinsumo, classe agronômica, formulação, ingrediente ativo, modo de ação, técnica de aplicação, indicação de uso, se é aprovado para agricultura orgânica, classificação toxicológica e ambiental, se é inflamável e corrosivo, e links para documentos completos.

- A listagem de inoculantes apresenta dados como a identificação e data de registro do produto, razão social e UF do fornecedor, tipo, espécie, garantia, natureza física e culturas aplicadas ao produto.

3.8.2. Avaliação de Volume e Qualidade

3.8.2.1. Volume

A API oferece listagens completas de produtos biológicos para controle de pragas e inoculantes registrados no MAPA.

3.8.2.2. Qualidade

As informações são de origem oficial, sendo extraídas do Ministério da Agricultura e Pecuária (MAPA). A listagem de produtos biológicos para controle de pragas é proveniente da base de dados AGROFIT.

3.8.3. Considerações para Coleta de Dados e Inteligência

- A **autenticação** é baseada em **token**, obtido após cadastro e assinatura da API por meio da plataforma AgroAPI. Os endpoints estão disponíveis tanto em **sandbox** quanto em **produção**.
- A documentação pode incluir guias como “Primeiros passos” e instruções gerais, bem como possibilidade de geração de SDKs (como Java, Android) na versão v2.

A API se integra bem com sistemas que demandam acesso a dados fitossanitários, permitindo implementação em **aplicações móveis, dashboards, ferramentas de recomendação ou plataformas de conformidade**.

3.8.4. Desafios e Oportunidades

3.8.4.1. Desafios

- A versão **v1 está deprecada**, sugerindo que futuras integrações devem migrar para a versão **v2**, considerada estável e recomendada.
- A documentação pública ainda carece de exemplos detalhados de uso, **payloads, formatos de resposta, rate limits específicos** ou **SLA**, demandando acesso à plataforma ou contato direto para informações técnicas.

3.8.4.2. Oportunidades

O generoso limite de **100 mil requisições/mês gratuitas**, sem contrato, torna a API extremamente atraente para startups, POCs e iniciativas de pesquisa. Pode suportar importantes casos de uso como:

- **Recomendações de bioinsumos** adequados para culturas e pragas específicas.
- **Integração com sistemas de gestão agrícola**, permitindo seleção segura e baseada em dados regulamentares.
- **Análise comparativa e monitoramento de tendências** de uso de bioinsumos no Brasil.

3.8.5. Recomendações para Integração

- **Use a versão v2** da API, pois a v1 está depreciada; certifique-se de acessar pelos endpoints atualizados com autenticação via token (sandbox e produção).
 - **Se inscreva na plataforma AgroAPI**, obtenha o token e explore a documentação — especialmente os “Primeiros passos” e guias disponíveis.
 - **Implemente caching local** das consultas mais frequentes (como por ingrediente ativo ou tipo de produto) — isso otimiza o uso e reduz chamadas repetidas.
 - **Modularize chamadas** à API por tipo de filtro (ex: cultura, praga, tipo de bioinsumo) para clareza e eficiência.
 - **Teste os formatos de resposta** retornados pela API antes de integração completa — isso ajuda na modelagem em sistemas downstream.
 - **Combine com outras APIs da AgroAPI** (como Agritec, ClimAPI, SATVeg, Agrofit) para enriquecer sua solução com múltiplas camadas de dados.
 - **Monitore a migração entre versões** e mantenha contato com os responsáveis técnicos:
 - Dono técnico: cnptia.bioinsumos@embrapa.br
 - Plataforma AgroAPI: agroapi@embrapa.br
 - **Planeje o uso dentro do limite gratuito (100 k/mês)** e, se necessário, consulte possibilidades de expansão junto à Embrapa.
-

3.9. BlueStar Sting API

3.9.1 Dados e Conteúdo

A **BlueStar Sting API** serve como uma interface entre o banco de dados **STING_RDB** e aplicações externas, permitindo acesso global a dados sobre **descritores físico-químicos, geométricos, espaciais e estruturais de aminoácidos** em estruturas proteicas depositadas no **Protein Data Bank (PDB)**.

Oferece descritores detalhados como:

- **Air**: área de superfície acessível ao solvente por resíduo, usando sonda esférica de raio 1.4 Å.
- **Atom, Atom ASA**: identificação atômica, área acessível dos átomos.
- **Chain, Contact, Residue Contacts**: tipos de cadeia e contatos intra/intercadeias, incluindo energia e pares de átomos em contato.
- **Curvate e Curvate Atom**: curvatura dos resíduos em nível atômico, positiva, negativa ou zero conforme região côncava, convexa ou enterrada.
- **Density Sponge**: densidade local usando sondas esféricas de variados raios (3 a 7 Å).
- **Distances, DSSP, Stride, Structure e Rotamer**: distâncias estruturais, estrutura secundária, informações front-end, ângulos conformacionais, etc.
- Demais descritores incluem: **Electric Potential, Energy Density, FPocket / FPocket Atom, Graph Descriptor, HSSP, Hydrophobic Patch / Patch, Hydrophobicity, Ligand Pocket and Water Contact, Mouth / Mouth Atom, PROSITE, Unused Contacts, Weighted Contact Number, Space Clash** — cobrindo uma ampla gama de propriedades físicas, energéticas, conformacionais e interacionais em estruturas proteicas.

3.9.2. Avaliação de Volume e Qualidade

3.9.2.1. Volume

A API oferece acesso aos dados de descritores físico-químicos e estruturais de qualquer estrutura proteica depositada no Protein Data Bank (PDB). A plataforma Blue Star STING diferencia 14

tipos de contatos intra e intercadeias e 12 tipos de contatos entre cadeias proteicas e moléculas de DNA/RNA

3.9.2.2. Qualidade

Os dados são baseados em descritores calculados a partir de metodologias específicas, como o uso de sondas esféricas e algoritmos como DSSP e Stride. Os valores de choque estérico, por exemplo, são baseados em fatores de sobreposição com valores comuns em estruturas obtidas por difração de raios X.

3.9.3. Considerações para Coleta de Dados e Inteligência

- O acesso se dá via a **plataforma AgroAPI**, com autenticação mediante **token** e possibilidade de uso tanto em **sandbox** quanto em **produção**.
- A API permite consultas tanto simples quanto avançadas, possibilitando o uso de descritores selecionados conforme a necessidade, ideal para exportação de dados para sistemas analíticos, mineração de padrões e integração em pipelines de biologia computacional.
- O volume generoso de acesso gratuito, **até 100 mil requisições por mês**, viabiliza POCs, protótipos e pesquisas sem custo inicial.

3.9.4. Desafios e Oportunidades

3.9.4.1. Desafios

- A complexidade dos descritores pode demandar conhecimento técnico avançado para uso correto, especialmente no entendimento de parâmetros biológicos e estruturais.
- A documentação pública não detalha formatos de resposta, exemplos de payloads, limites específicos por endpoint (rate limits) ou SLAs; é recomendável consultar diretamente a plataforma para essas informações.

3.9.4.2. Oportunidades

- Potencial de aplicação em **biologia estrutural computacional**, **desenvolvimento de biofármacos**, **controle biotecnológico de pragas** e também na criação de **inseticidas de alto desempenho**.

- Permite pesquisas em áreas como **design de proteínas**, **estrutura-função**, **escore de interações moleculares**, com base em dados robustos e granulares.
- Excelente base para startups de **agrotecnologia**, **bioinformática** e instituições de pesquisa que buscam explorar dados estruturais aplicados à agricultura digital.

3.9.5. Recomendações para Integração

- **Registrar-se na plataforma AgroAPI**, obter o token de acesso e testar no ambiente sandbox antes de migrar para produção.
 - **Explorar descritores específicos** conforme requisitos do projeto, por exemplo, se precisar de propriedades geométricas ou eletrônicas, selecione os descritores correspondentes.
 - **Implementar caching local** de resultados recorrentes para reduzir latência e economia de chamadas.
 - **Pesquisar exemplos e guias técnicos** dentro da plataforma ou entrar em contato com o suporte técnico da Embrapa para obter FAQs, exemplos JSON e API specs.
 - **Modularizar chamadas** por tipo de consulta ou descritor para melhor organização e performance da aplicação.
 - **Integrar com outras APIs da AgroAPI** (como Agrofit, ClimAPI, etc.) se sua solução demandar combinação de fatores biológicos e agronômicos.
 - **Planejar conforme o limite gratuito (100k requisições/mês)** e, se necessário, discutir expansão via contato com agroapi@embrapa.br.
-

3.10. Responde Agro API

3.10.1 Dados e Conteúdo

- A **Responde Agro API** oferece um mecanismo de busca para consultar o conteúdo da série de publicações **Coleção 500 Perguntas 500 Respostas** da Embrapa. Ao fornecer um texto, a API retorna conjuntos ordenados de pares pergunta–resposta extraídos das obras indexadas. O conteúdo é entregue em formato **HTML**, pronto para exibição, incluindo imagens codificadas em **Base64**.

- A coleção abrange temas relacionados à agricultura e pecuária, com perguntas formuladas por produtores, cooperativas ou associações e respondidas por pesquisadores da Embrapa.

3.10.2. Avaliação de Volume e Qualidade

3.10.2.1. Volume

A API possibilita a busca em todos os livros indexados da Coleção "500 Perguntas 500 Respostas" ou em um livro específico. A coleção em si é vasta, cobrindo temas relacionados à agricultura e pecuária.

3.10.2.2. Qualidade

O conteúdo é de natureza técnico-científica e é gerado por pesquisadores da Embrapa. Os resultados são retornados em formato HTML, o que facilita a integração com as interfaces de usuário.

3.10.3. Considerações para Coleta de Dados e Inteligência

Acesso via **API RESTful** com autenticação por **token**, disponível nos ambientes **sandbox** e **produção** da plataforma AgroAPI. Os principais recursos incluem:

- Busca por uma pergunta específica dentro de um livro específico.
- Busca por texto em um livro ou em toda a coleção com controle de paginação (parâmetros `size`, `from`).
- Recurso de **autocompletar**, que sugere perguntas ou termos conforme o usuário digita, com base no índice.
- Consulta para listar todos os identificadores de livros indexados.

Essa inteligência permite criação de interfaces de busca intuitivas e responsivas, com sugestões automáticas enquanto o usuário digita a consulta.

3.10.4. Desafios e Oportunidades

3.10.4.1. Desafios

- A documentação pública não inclui exemplos completos de payloads, limites de taxa (`rate limits`), SLAs ou detalhes técnicos específicos, exigindo acesso direto via portal AgroAPI.

- A base restrita às obras da série 500 Perguntas 500 Respostas limita o escopo a esse conteúdo específico.

3.10.4.2. Oportunidades

- Permite incorporar um **mecanismo de busca semântico** em soluções digitais para técnicos agrícolas, produtores, agentes de extensão, cooperativas, APPs e chatbots.
- Favorece integração com plataformas de ATER (Assistência Técnica e Extensão Rural), como o Ater+ Digital, que já utiliza essa API.
- A funcionalidade de **autocomplete** enriquece UX e melhora a navegação por conteúdo técnico, potencialmente reduzindo barreiras de uso para usuários finais.

3.10.5. Recomendações para Integração

- **Registrar-se na plataforma AgroAPI** e gerar o token de acesso após assinar a API.
 - Testar os recursos disponíveis no sandbox, como busca por ID, texto geral, autocomplete e listagem de livros.
 - **Implementar cache local** para consultas frequentes, diminuindo latência e uso de requisições repetitivas.
 - **Modularizar chamadas** conforme tipo de uso — busca no livro específico, autocomplete, consulta por texto geral.
 - **Testar e validar** os formatos de resposta (HTML + Base64) para garantir compatibilidade com frontends (web, apps, chatbots etc.).
 - **Integrar com outras APIs da AgroAPI** (como ClimAPI, Agritec, Agrofit), criando soluções mais completas e contextualizadas para o agro.
 - Contatar os responsáveis técnicos para suporte ou esclarecimentos:
 - **Business Owner:** Embrapa Agricultura Digital (e-mail: agroapi@embrapa.br)
 - **Technical Owner:** Glauber Vaz (e-mail: glauber.vaz@embrapa.br)
-

3.11. Responde Agro API

3.11.1 Dados e Conteúdo

- **Responde Agro API** oferece um mecanismo de busca para consultar o conteúdo da série de publicações **Coleção 500 Perguntas 500 Respostas** da Embrapa. Ao fornecer um texto, a API retorna conjuntos ordenados de pares pergunta–resposta extraídos das obras indexadas. O conteúdo é entregue em formato **HTML**, pronto para exibição, incluindo imagens codificadas em **Base64**.
- A coleção abrange temas relacionados à agricultura e pecuária, com perguntas formuladas por produtores, cooperativas ou associações e respondidas por pesquisadores da Embrapa.

3.11.2. Avaliação de Volume e Qualidade

3.11.2.1. Volume

A API possibilita a busca em todos os livros indexados da Coleção "500 Perguntas 500 Respostas" ou em um livro específico. A coleção em si é vasta, cobrindo temas relacionados à agricultura e pecuária.

3.11.2.2. Qualidade

O conteúdo é de natureza técnico-científica e é gerado por pesquisadores da Embrapa. Os resultados são retornados em formato HTML, o que facilita a integração com as interfaces de usuário.

3.11.3. Considerações para Coleta de Dados e Inteligência

Acesso via **API RESTful** com autenticação por **token**, disponível nos ambientes **sandbox** e **produção** da plataforma AgroAPI.

Os principais recursos incluem:

- Busca por uma pergunta específica dentro de um livro específico.
- Busca por texto em um livro ou em toda a coleção com controle de paginação (parâmetros `size`, `from`).
- Recurso de **autocompletar**, que sugere perguntas ou termos conforme o usuário digita, com base no índice.
- Consulta para listar todos os identificadores de livros indexados.

Essa inteligência permite criação de interfaces de busca intuitivas e responsivas, com sugestões automáticas enquanto o usuário digita a consulta.

3.11.4. Desafios e Oportunidades

3.11.4.1. Desafios

- A documentação pública não inclui exemplos completos de payloads, limites de taxa (rate limits), SLAs ou detalhes técnicos específicos, exigindo acesso direto via portal AgroAPI.
- A base restrita às obras da série 500 Perguntas 500 Respostas limita o escopo a esse conteúdo específico.

3.11.4.2. Oportunidades

- Permite incorporar um **mecanismo de busca semântico** em soluções digitais para técnicos agrícolas, produtores, agentes de extensão, cooperativas, APPs e chatbots. Favorece integração com plataformas de ATER (Assistência Técnica e Extensão Rural), como o Ater+ Digital, que já utiliza essa API.
- A funcionalidade de **autocomplete** enriquece UX e melhora a navegação por conteúdo técnico, potencialmente reduzindo barreiras de uso para usuários finais.

3.11.5. Recomendações para Integração

- **Registrar-se na plataforma AgroAPI** e gerar o token de acesso após assinar a API.
- Testar os recursos disponíveis no sandbox, como busca por ID, texto geral, autocomplete e listagem de livros.
- **Implementar cache local** para consultas frequentes, diminuindo latência e uso de requisições repetitivas.
- **Modularizar chamadas** conforme tipo de uso — busca no livro específico, autocomplete, consulta por texto geral.
- **Testar e validar** os formatos de resposta (HTML + Base64) para garantir compatibilidade com frontends (web, apps, chatbots etc.).
- **Integrar com outras APIs da AgroAPI** (como ClimAPI, Agritec, Agrofit), criando soluções mais completas e contextualizadas para o agro.

Contatar os responsáveis técnicos para suporte ou esclarecimentos:

- **Business Owner:** Embrapa Agricultura Digital (e-mail: agroapi@embrapa.br)
 - **Technical Owner:** Glauber Vaz (e-mail: glauber.vaz@embrapa.br)
-

3.12. SmartSolos Expert API

3.12.1 Dados e Conteúdo

- A **SmartSolos Expert API** oferece infraestrutura para **classificação de perfis de solos no Brasil**, utilizando um sistema especialista baseado nas regras da **5ª edição do Sistema Brasileiro de Classificação de Solos (SiBCS)** — o sistema taxonômico oficial do país.
- A API permite tanto a **classificação automática** dos perfis quanto a **validação de classificações prévias**, comparando os resultados e fornecendo métricas de concordância.
- O sistema foi implementado em **SWI-Prolog** e se comunica via JSON, recebendo perfis com dados de horizontes (e.g., profundidades, textura, cor, pH, estrutura, nutrientes) e retornando os níveis de classificação: **ORDEM**, **SUBORDEM**, **GDE_GRUPO** e **SUBGRUPO**.

3.12.2. Avaliação de Volume e Qualidade

3.12.2.1. Volume

A API pode classificar múltiplos perfis de solo em uma única chamada. A entrada para cada perfil inclui uma lista de dados de horizontes, além de outros atributos do perfil.

3.12.2.2. Qualidade

O sistema especialista é baseado nas regras da 5ª edição do SiBCS, de 2018. A classificação é realizada com base em atributos detalhados de horizontes e perfis de solo. Os resultados podem ser diferentes para perfis classificados com versões anteriores do SiBCS. Se os dados de entrada não forem suficientes, o resultado pode ser incompleto ou incorreto.

3.12.3. Considerações para Coleta de Dados e Inteligência

A API está disponível através da **plataforma AgroAPI**, sendo necessário gerar um **token de acesso** após registro, com ambientes disponíveis tanto em **sandbox** quanto em **produção**.

O endpoint base é: <https://api.cnptia.embrapa.br/smartsolos/expert/v1>.

Utilizando JSON, a API suporta múltiplos perfis em uma única chamada, o retorno inclui os níveis de classificação mencionados e, em chamadas de validação, também compara com classificações prévias, gerando um resumo das divergências.

3.12.4. Desafios e Oportunidades

3.12.4.1. Desafios

- O poder do sistema depende da **completude e precisão dos dados de entrada**. Se parâmetros necessários estiverem ausentes ou incorretos, a classificação pode ser imprecisa ou incompleta [\[1\]](#).
- A classificação segue estritamente os critérios da **5ª edição do SiBCS (2018)**. Perfis classificados em versões anteriores podem gerar resultados discrepantes.
- A documentação pública e exemplos de uso são limitados, e para resolver inconsistências ou problemas, recomenda-se contato técnico: cnptia.smartsolosexpert@embrapa.br.

3.12.4.2. Oportunidades

- Facilita o **ensino de pedologia**, a **padronização de dados de solos** e a **curadoria de bases georreferenciadas**, especialmente dentro do Programa Nacional de Solos (Pronasolos).
- Ajuda a identificar e corrigir incoerências em bancos de dados históricos, promovendo decisões mais confiáveis e conformes com o SiBCS.
- A existência de um limite generoso e gratuito, até **100 mil requisições/mês**, sem necessidade de contrato, facilita a adoção em POCs, pesquisas e aplicações em larga escala.

3.12.5. Recomendações para Integração

- **Registrar-se na plataforma AgroAPI**, gerar o token de acesso e testar no ambiente sandbox antes de migrar para produção.

- **Mapear e enviar perfis bem estruturados:** incluir atributos obrigatórios como ID_PONTO, e dados de horizonte que cubram profundidades, cor, textura, nutrientes e outros, conforme exigido.
 - **Utilizar os dois recursos disponíveis:** classificação e validação, permitindo comparar resultados com classificações anteriores.
 - **Implementar caching** para perfis recorrentes, reduzindo chamadas repetidas e custos no limite gratuito.
 - **Analisar o summary das validações**, para identificar níveis de divergência e ajustar os dados de origem conforme necessário.
 - **Integrar com outras APIs da AgroAPI** (como Agritec, ClimAPI, SATVeg) para enriquecer soluções digitais com camadas climáticas, vegetacionais e fitossanitárias, além da pedológica.
- Acompanhar documentação, atualizações e roadmap técnico**, e manter contato com o responsável técnico, cnptia.smartsolosexpert@embrapa.br.
-

3.13. PlantAnnot API

3.13.1 Dados e Conteúdo

- A **API PlantAnnot** disponibiliza acesso a informações sobre **genes, transcritos e proteínas** de mais de **50 espécies de plantas**, mantidas no sistema PlantAnnot.
- Desenvolvida para identificar **proteínas de função desconhecida (PUFs)** que podem estar relacionadas com **mecanismos moleculares de estresses abióticos** em plantas.
- Possui diversas funcionalidades, incluindo recuperação de:
 - Localização cromossômica, referências externas (DBxRefs), sequências e artigos;
 - Anotações funcionais como similaridade (BLAST, Diamond), ontologia genética (Gene Ontology) e domínios proteicos (InterPro);
 - Expressão gênica em diferentes condições experimentais, incluindo co-expressão;
 - Grupos de ortólogos entre espécies.

3.13.2. Avaliação de Volume e Qualidade

3.13.2.1. Volume

A API oferece acesso a um banco de dados de mais de 50 espécies de plantas. As informações de expressão de genes estão disponíveis para organismos modelo como *Arabidopsis thaliana*, Soja, Arroz e Milho.

3.13.2.2. Qualidade

Os dados são utilizados em pesquisas do Laboratório Multiusuário de Bioinformática (LMB) e da UMIP GenClima da Embrapa. O sistema tem o potencial de gerar novas variedades de plantas geneticamente modificadas resistentes a estresses relacionados às mudanças climáticas.

3.13.3. Considerações para Coleta de Dados e Inteligência

- AAPI é acessada via **RESTful API** com autenticação por **token**, disponível nos ambientes de **sandbox** e **produção** [r](#).
- O endpoint base é <https://api.cnptia.embrapa.br/plantannot/v1>.
- A API oferece **SDKs** para **Java** e **Android**, além de documentação interativa (Swagger) para facilitar testes e integração.

3.13.4. Desafios e Oportunidades

3.13.4.1. Desafios

A documentação disponibilizada online é limitada; exemplos de payloads, formatos de resposta, limites de requisições (rate limits) e SLA não estão publicamente detalhados, recomenda-se consultar diretamente a documentação na plataforma ou buscar contato com os responsáveis técnicos.

3.13.4.2. Oportunidades

- Suporte avançado à bioinformática aplicada ao agronegócio, com destaque para identificação de PUFs ligados a estresses abióticos, essencial em projetos de melhoramento genético e resiliência climática.
- Disponível para prototipagem, pesquisa e desenvolvimento de soluções (como pipelines automatizados, web apps genômicos, ferramentas de anotação) com acesso gratuito generoso.

3.13.5. Recomendações para Integração

- **Cadastrar-se na plataforma AgroAPI** e gerar o **token de acesso**, testando inicialmente no ambiente sandbox.

- **Explorar a documentação via Swagger** e usar os SDKs disponíveis (Java e Android) para acelerar o desenvolvimento.
- **Testar os principais endpoints**, tais como buscas por gene ou proteína e recuperação de dados de anotação funcional, expressão e ortologia.
- **Implementar caching local** para informações estáticas ou consultadas com frequência (ex.: genes ou proteínas específicos).
- **Modularizar consultas** por tipo de dado (por exemplo, localização cromossômica, expressão, ortologia), otimizando requisições e lógica de integração.
- **Combinar com outras APIs da AgroAPI** (como AgroTermos, Agritec, ClimAPI) para integrar dados genômicos a classificações, clima e terminologia, construindo soluções interdisciplinares.
- **Contatar os responsáveis técnicos e de negócio em caso de dúvidas ou expansão do uso:**
 - Business Owner: mauricio.mudadu@embrapa.br
 - Technical Owner: adhemar.zerlotini@embrapa.br.
- **Planejar o uso dentro do limite gratuito (100 k requisições/mês)** e, se necessário, consultar possibilidades de expansão por e-mail via agroapi@embrapa.br.

4. Cadastro Ambiental Rural (CAR)

O Cadastro Ambiental Rural (CAR) é um registro público eletrônico obrigatório para todos os imóveis rurais no Brasil.

4.1. Dados e Conteúdo

O CAR integra informações ambientais de propriedades e posses rurais, incluindo Áreas de Preservação Permanente (APP), áreas de uso restrito, Reserva Legal, remanescentes de vegetação nativa e áreas consolidadas. Os dados abrangem informações georreferenciadas do perímetro do imóvel, áreas de interesse social e de utilidade pública, e localização de vegetação nativa. Embora o dataset oficial no Portal de Dados Abertos do Governo Federal seja classificado como "Não Aberto" e "Restrito - Informação Pessoal", ele fornece metadados e recursos em CSV sobre o número de imóveis por município/UF, sobreposição de registros e temas ambientais.

4.2. Acesso e Documentação

O acesso direto aos dados detalhados do CAR via API pública é restrito devido à natureza das informações pessoais envolvidas. No entanto, uma API de terceiros, a

[API CAR da Directd](#), oferece acesso a dados cadastrais e ambientais do imóvel, situação do cadastro e histórico de retificações. Esta API exige autenticação via TOKEN e pode ter custo adicional para geração de comprovantes em PDF.

A restrição de acesso direto aos dados brutos do CAR representa um desafio para a obtenção de informações detalhadas sobre propriedades individuais para fins de inteligência agrônômica. A necessidade de recorrer a APIs de terceiros ou a conjuntos de dados agregados limita a granularidade e a flexibilidade das análises.

4.3. Precificação

O acesso à API do CAR da Directd tem um modelo de precificação que varia de acordo com o volume de consultas. O preço é cobrado por consulta, e o valor unitário diminui à medida que o volume de uso mensal aumenta. Além disso, a plataforma exige um gasto mínimo mensal para uso dos serviços.

- **Custo por Consulta:** Varia de R\$ 0,20 a R\$ 0,05, dependendo do volume mensal de requisições.
- **Franquia Mínima Mensal:** A cobrança mínima por mês é de R\$ 100,00.
- **Serviços Adicionais:** A geração de um demonstrativo do CAR em PDF tem um custo extra por consulta, atualmente de R\$ 0,10.

A empresa também oferece um modelo pré-pago onde você pode comprar créditos, começando com recargas de R\$ 50,00.

4.4. Avaliação de Volume e Qualidade

4.4.1. Volume

O CAR é um registro obrigatório para todos os imóveis rurais no Brasil, com o objetivo de integrar informações ambientais para controle,

monitoramento e planejamento. Até abril de 2022, foram registrados aproximadamente 6.503.840 imóveis, totalizando cerca de 618 milhões de hectares cadastrados. Desses, 52% solicitaram adesão ao Programa de Regularização Ambiental (PRA), indicando passivos ambientais a serem recompostos, onde foi exibido no [Relatório Final de Exibição](#).

4.4.2. Qualidade

A qualidade dos dados do CAR é variável, com desafios significativos na análise e validação das informações. Até 2020, apenas 3,4% dos cadastros haviam sido analisados, devido à necessidade de verificação manual dos dados, que é demorada e custosa. Além disso, há inconsistências nos dados, como sobreposições de áreas e registros em zonas urbanas, que comprometem a precisão das informações, onde foi exibido no [Portal da Câmara dos Deputados](#).

4.5. Considerações para Coleta de Dados e Inteligência

4.5.1. Coleta de Dados

- **Georreferenciamento:** É essencial que os dados geoespaciais sejam precisos e atualizados, utilizando tecnologias como GPS de alta precisão e imagens de satélite de alta resolução. [CPI](#)
- **Fontes de Dados:** Utilizar fontes confiáveis, como o Sistema de Cadastro Ambiental Rural (SICAR) e o Cadastro Nacional de Imóveis Rurais (CNIR), para garantir a consistência e integridade dos dados. [Studies Publicações+2Repositório IFPE+2](#)

4.5.2. Inteligência

- **Análise Automatizada:** Implementar ferramentas de análise automatizada para acelerar o processo de validação dos dados, reduzindo custos e aumentando a eficiência. [Portal da Câmara dos Deputados](#)
- **Integração de Dados:** Integrar os dados do CAR com outras bases de dados ambientais e fundiárias para obter uma visão mais abrangente e precisa da situação ambiental dos imóveis rurais.

4.6. Desafios e Oportunidades

4.6.1 Desafios

- **Análise Manual:** A necessidade de verificação manual dos dados é um processo demorado e oneroso, limitando a capacidade de análise em larga escala.
- **Inconsistências nos Dados:** Há registros com sobreposições de áreas, dados em zonas urbanas e informações desatualizadas, comprometendo a qualidade dos dados.
- **Capacidade Técnica:** A falta de capacitação técnica em alguns estados e municípios dificulta a implementação eficaz do CAR.

4.6.2 Oportunidades

- **Automatização:** O desenvolvimento e implementação de ferramentas de análise automatizada podem melhorar significativamente a eficiência na validação dos dados.
- **Parcerias:** Estabelecer parcerias com instituições de pesquisa, universidades e organizações não governamentais pode fortalecer a implementação e monitoramento do CAR.
- **Capacitação:** Investir em programas de capacitação para técnicos e gestores ambientais pode melhorar a qualidade e eficiência na gestão do CAR.

4.7. Recomendações para Integração

- **Integração de Sistemas:** Integrar o CAR com outros sistemas de informação ambiental e fundiária, como o CNIR e o Sistema de Informação sobre a Biodiversidade Brasileira (SiBBr), para uma gestão mais eficiente e integrada.
- **Padronização de Dados:** Estabelecer padrões para a coleta, armazenamento e análise de dados, garantindo a consistência e interoperabilidade entre os sistemas.

Transparência e Acesso Público: Disponibilizar os dados do CAR de forma transparente e acessível ao público, promovendo a participação

social e o controle social na gestão ambiental.

- **Monitoramento Contínuo:** Implementar mecanismos de monitoramento contínuo para acompanhar a evolução dos cadastros e identificar áreas que necessitam de atenção especial.

5. Companhia Nacional de Abastecimento (CONAB)

A CONAB é responsável por monitorar e divulgar dados estratégicos sobre a produção agrícola e o abastecimento no país.

5.1. Dados e Conteúdo

A CONAB realiza o levantamento e avaliação das safras brasileiras de grãos, fibras, café e cana-de-açúcar. Os dados incluem análises fitotécnicas e econômicas, área plantada, produtividade, produção, monitoramento agrícola, prognóstico climático, oferta, demanda e análise de mercado. Os levantamentos de grãos e fibras são mensais, enquanto os de café e cana-de-açúcar são trimestrais.

Além dos dados de safras, a CONAB disponibiliza informações sobre:

- **Preços Agropecuários:** Preços de commodities agrícolas e cotações de mercado, incluindo séries históricas.
- **Custos de Produção:** Séries históricas e análises de rentabilidade para diversas culturas.
- **Estoques Públicos e Operações de Comercialização:** Dados sobre abastecimento.
- **Mapeamentos Agrícolas:** Distribuição geográfica de cultivos, auxiliando na estimativa de área e produtividade.

5.2. Acesso e Documentação

A CONAB disponibiliza seus dados e informações através do [Portal de Informações Agropecuárias](#). Embora o Portal da Transparência do Governo Federal liste uma "API de Dados" que inclui informações relacionadas à CONAB,

esta API é mais genérica para dados governamentais e requer cadastro para obter um token.

Para os dados específicos da CONAB, o acesso principal é via **download direto** de arquivos. Os mapeamentos agrícolas estão disponíveis para download no formato Shapefile. Séries históricas e outros dados de produção agrícola, custos e abastecimento podem ser baixados em formatos de arquivo não especificados, mas geralmente são planilhas.

Uma "API de Integração com dados e processos da CONAB - ADP-JAVA" foi desenvolvida, mas é descrita como um programa para tratamento e extração de dados de *data warehouses* internos da CONAB, com nível de maturidade tecnológica TRL 4, indicando que não é uma API pública para consumo externo.

A ausência de uma API pública robusta e dedicada para os dados agrônômicos da CONAB exige que os usuários dependam de downloads diretos e, possivelmente, de web scraping para obter informações atualizadas e em volume. Isso pode aumentar a complexidade da integração e a necessidade de processamento de dados para projetos de inteligência agrônômica.

5.3. Precificação

Como uma agência do governo, a CONAB disponibiliza suas informações para o público sem custo, com o objetivo de promover a transparência e a análise do mercado agropecuário.

5.4. Avaliação de Volume e Qualidade

5.4.1. Volume

A CONAB disponibiliza uma ampla gama de dados, incluindo informações sobre safras de grãos, fibras, café e cana-de-açúcar, preços agropecuários, custos de produção, estoques públicos e operações de comercialização. A base de dados Conab/Prohort, por exemplo, contempla informações de 117 frutas e 123 hortaliças, somando mais de mil produtos quando consideradas suas variedades.

5.4.2. Qualidade

Os dados são coletados por meio de pesquisas diretas com produtores, sensoriamento remoto, uso de pacotes tecnológicos, intenção de plantio, médias históricas, rede de informantes e acompanhamento agrometeorológico. Essas informações são processadas utilizando métodos estatísticos para garantir a confiabilidade e consistência das informações.

5.5. Considerações para Coleta de Dados e Inteligência

- **Fontes de Dados:** A CONAB utiliza diversas fontes para a coleta de dados, incluindo pesquisas diretas com produtores, sensoriamento remoto, uso de pacotes tecnológicos, intenção de plantio, médias históricas, rede de informantes e acompanhamento agrometeorológico.
- **Processamento e Análise:** Os dados coletados são processados utilizando métodos estatísticos para garantir a confiabilidade e consistência das informações. Além disso, a CONAB realiza um processo revisional para validar os dados antes de disponibilizá-los ao público.
- **Integração com Outras Fontes:** A CONAB também integra seus dados com outras fontes, como o Observatório da Agropecuária Brasileira, para fornecer uma visão mais abrangente do setor agropecuário.

5.6. Desafios e Oportunidades

5.6.1 Desafios

- **Integração de Dados:** A integração de dados de diferentes fontes pode ser desafiadora devido a diferenças nos formatos e padrões de dados. A falta de comunicação entre os diferentes sistemas e equipamentos de coleta de dados era um obstáculo importante no passado.
- **Recursos Limitados:** A gestão de recursos limitados, mudanças políticas e a necessidade de capacitação contínua impactam diretamente a atuação e eficácia da CONAB no setor agrícola.

5.6.2 Oportunidades

- **Tecnologia e Inovação:** A adoção de tecnologias avançadas para a integração e gerenciamento de dados pode melhorar a eficiência e a qualidade das informações.
- **Parcerias e Colaboração:** A colaboração com outras instituições e a participação em iniciativas de governo aberto podem ampliar o acesso e a utilização dos dados da CONAB.

5.7. Recomendações para Integração

- **Desenvolvimento de APIs Públicas:** A criação de APIs públicas robustas e dedicadas para os dados agronômicos da CONAB facilitaria a integração com outras plataformas e sistemas.
 - **Padronização de Dados:** A adoção de padrões de dados abertos e interoperáveis facilitaria a integração e o compartilhamento de informações.
 - **Capacitação e Treinamento:** Investir em capacitação contínua para os servidores da CONAB é essencial para lidar com as demandas do setor agrícola e melhorar a gestão de dados.
 - **Fortalecimento de Parcerias:** Estabelecer parcerias com outras instituições e participar de iniciativas de governo aberto pode ampliar o acesso e a utilização dos dados da CONAB.
-

6. Instituto Nacional de Meteorologia (INMET)

O INMET é o órgão responsável pela previsão do tempo e pela disponibilização de dados meteorológicos no Brasil.

6.1. Dados e Conteúdo

O INMET oferece dados meteorológicos diários históricos através do **BDMEP** (Banco de Dados Meteorológicos para Ensino e Pesquisa). O BDMEP contém milhões de informações de estações meteorológicas convencionais da rede INMET, referentes a medições diárias desde 1961. As variáveis incluem precipitação, radiação global, temperatura do ar (bulbo seco, máxima, mínima),

umidade relativa do ar (horária, máxima, mínima), velocidade e rajada máxima do vento, pressão atmosférica, e temperatura do ponto de orvalho. Os dados são atualizados a cada 90 dias.

6.2. Acesso e Documentação

O acesso aos dados do BDMEP é feito diretamente pelo [portal](#). Os dados podem ser baixados em formatos como CSV. Embora não haja uma API oficial do INMET para acesso programático em tempo real ou em massa, existe uma API não oficial no [Repositório do GitHub em Python](#) que permite extrair dados históricos de temperatura do BDMEP. Além de uma [Biblioteca em R](#) que permite o download de dados meteorológicos de estações do INMET em linguagem R.

A disponibilidade de dados históricos de longo prazo do INMET é crucial para modelos agrônômicos que dependem de séries temporais climáticas para análise de risco, zoneamento e previsão de produtividade. No entanto, a falta de uma API oficial para acesso direto e em tempo real dos dados brutos pode ser um limitador para aplicações que exigem alta frequência de atualização.

6.3. Precificação

O acesso aos dados do Instituto Nacional de Meteorologia (INMET) é gratuito.

O INMET é um órgão público e, como parte de sua missão, disponibiliza seus dados para a população sem nenhum custo. Isso inclui o acesso ao BDMEP (Banco de Dados Meteorológicos para Ensino e Pesquisa) e a todas as informações meteorológicas disponíveis em seu portal.

As APIs não oficiais que você encontrou no GitHub também são, por natureza, gratuitas, pois são projetos da comunidade que buscam facilitar o acesso aos dados públicos do INMET. O desenvolvedor dessas APIs não cobra pelo uso.

6.4. Avaliação de Volume e Qualidade

6.4.1. Volume

O BDMEP contém milhões de registros diários de estações meteorológicas convencionais desde 1961, abrangendo variáveis como

precipitação, temperatura, umidade relativa, pressão atmosférica, entre outras. A cobertura temporal vai até julho de 2025, com atualizações trimestrais.

6.4.2. Qualidade

Os dados são coletados conforme as normas da Organização Meteorológica Mundial (OMM), garantindo padronização internacional. Embora o BDMEP seja uma fonte confiável, a qualidade dos dados pode variar devido a fatores como manutenção das estações e condições climáticas adversas.

6.5. Considerações para Coleta de Dados e Inteligência

- O INMET opera uma rede de estações meteorológicas convencionais e automáticas, além de utilizar imagens de satélite para monitoramento.
- A coleta segue protocolos técnicos estabelecidos pela OMM, visando garantir a consistência e comparabilidade dos dados.
- O INMET possui sistemas de modelagem numérica para previsão do tempo e clima, como o Modelo Brasileiro de Alta Resolução (MBAR) e o COSMO, que utilizam os dados coletados para gerar previsões.

6.6. Desafios e Oportunidades

6.6.1 Desafios

- **Cobertura Espacial:** A rede de estações meteorológicas, embora abrangente, pode apresentar lacunas em áreas remotas ou de difícil acesso.
- **Manutenção das Estações:** A necessidade de manutenção constante das estações para garantir a precisão dos dados coletados.
- **Integração de Dados:** A integração de dados de diferentes fontes e formatos pode ser desafiadora, exigindo sistemas robustos de processamento e armazenamento.

6.6.2 Oportunidades

- **Expansão da Rede:** Ampliar a rede de estações meteorológicas para melhorar a cobertura espacial e a precisão das previsões.
- **Parcerias Tecnológicas:** Estabelecer parcerias com instituições de pesquisa e empresas de tecnologia para aprimorar os sistemas de coleta e processamento de dados.
- **Capacitação Profissional:** Investir em treinamento e capacitação dos profissionais envolvidos na coleta e análise de dados meteorológicos.

6.7. Recomendações para Integração

- **Desenvolvimento de APIs Oficiais:** Criar interfaces de programação de aplicativos (APIs) oficiais para facilitar o acesso e a integração dos dados em tempo real.
- **Padronização de Dados:** Adotar padrões internacionais para garantir a interoperabilidade dos dados com outras plataformas e sistemas.
- **Plataformas de Visualização:** Desenvolver plataformas interativas de visualização de dados para facilitar a interpretação e uso das informações por diversos públicos.
- **Colaboração Interinstitucional:** Fomentar a colaboração entre diferentes instituições governamentais, acadêmicas e privadas para promover a troca de dados e conhecimentos.

7. Agência Nacional de Águas e Saneamento Básico (ANA)

A ANA é a agência reguladora e gestora dos recursos hídricos no Brasil, fornecendo dados essenciais para o manejo da água na agricultura.

7.1. Dados e Conteúdo

A ANA disponibiliza dados e informações do **Sistema Nacional de Informações sobre Recursos Hídricos (SNIRH)** através de sua plataforma de Dados Abertos. O SNIRH é um instrumento de gestão que coleta, trata, armazena e recupera informações sobre o estado qualitativo e quantitativo da água no Brasil. A plataforma oferece 97 conjuntos de dados, fragmentados em dez grupos, incluindo:

- Divisão hidrográfica
- Quantidade e qualidade de água
- Usos da água
- Balanço hídrico
- Eventos hidrológicos críticos
- Planos de recursos hídricos
- Dados sobre áreas irrigadas por pivô central
- Índices de Segurança Hídrica (ISH)

7.2. Acesso e Documentação

A nova Plataforma de [Dados Abertos da ANA](#) oferece uma [Search API](#) que está em conformidade com a especificação OGC API - Records. Esta API permite consultar, filtrar e buscar o catálogo de dados programaticamente.

Os dados podem ser acessados em formatos como Feature Service (que implica compatibilidade com GeoJSON e Shapefile), e CSV. A API permite a obtenção de respostas em JSON.

A disponibilização de dados hídricos em formatos geoespaciais e via API é um diferencial significativo da ANA, permitindo a integração direta de informações sobre disponibilidade e uso da água em modelos de gestão agrícola e de risco hídrico. A conformidade com padrões OGC facilita a interoperabilidade com sistemas GIS.

7.3. Precificação

Não há **precificação** para o acesso aos dados da Agência Nacional de Águas e Saneamento Básico (ANA).

O acesso à plataforma de **Dados Abertos** e à **API** do Sistema Nacional de

Informações sobre Recursos Hídricos (SNIRH) é totalmente **gratuito**.

A cobrança que a ANA realiza se refere ao **uso da água** como recurso natural, o que é uma taxa paga por empresas e indústrias que utilizam a água em seus processos. Isso é uma política de gestão hídrica e não tem relação com a obtenção de dados por meio da API da agência.

7.4. Avaliação de Volume e Qualidade

7.4.1. Volume

A ANA possui uma plataforma com 97 conjuntos de dados, fragmentados em 10 grupos diferentes. A quantidade e a qualidade das informações disponíveis são bastante robustas e abrangem diversos aspectos relacionados à gestão de recursos hídricos no Brasil, como divisão hidrográfica, quantidade e qualidade da água, usos da água, entre outros.

7.4.2. Qualidade

A qualidade dos dados é alta, dado que a plataforma segue especificações da Open Geospatial Consortium (OGC), permitindo que os dados sejam bem estruturados e interoperáveis com sistemas GIS. Além disso, a plataforma oferece informações detalhadas e frequentemente atualizadas sobre os estados qualitativo e quantitativo da água.

7.5. Considerações para Coleta de Dados e Inteligência

- A ANA coleta dados de forma geoespacial e hidrológica, com foco na gestão de recursos hídricos e segurança hídrica. A coleta de dados envolve a utilização de estações de monitoramento espalhadas por diversas regiões do Brasil.
- A plataforma oferece recursos como APIs que permitem consultas e buscas de dados programaticamente. Isso facilita a integração e análise de dados para gestão agrícola, planejamento de uso de recursos hídricos e modelos preditivos relacionados ao risco hídrico.

7.6. Desafios e Oportunidades

7.6.1 Desafios

A integração de grandes volumes de dados de diferentes fontes e a necessidade de garantir a atualização constante desses dados são desafios consideráveis. Além disso, o fato de os dados serem geoespaciais pode exigir ferramentas e tecnologias específicas para análise e visualização.

7.6.2 Oportunidades

A disponibilização dos dados em formatos como GeoJSON e Shapefile permite uma vasta gama de análises para integração com sistemas GIS e modelos de previsão. A ANA tem uma excelente oportunidade de colaborar com outras entidades, como universidades e empresas, para avançar em soluções de gestão hídrica inteligente.

7.7. Recomendações para Integração

- **Integração com Sistemas de Gestão de Recursos Hídricos:** A utilização da plataforma de dados abertos da ANA pode ser integrada com sistemas de monitoramento em tempo real de consumo e qualidade de água. Utilizar os dados disponíveis em conjunto com sistemas de gestão agrícola pode melhorar a eficiência no uso de água em atividades agrícolas.
 - **Melhorias na Documentação e Acessibilidade:** Apesar de a plataforma oferecer uma documentação sólida, a ampliação das capacidades de personalização nas consultas via API poderia ajudar ainda mais na adaptação dos dados para diferentes contextos. Adicionar exemplos mais detalhados sobre o uso da API e oferecer suporte contínuo para desenvolvedores pode expandir a adoção da plataforma.
-

8. Ministério da Agricultura e Pecuária (MAPA)

O MAPA é o órgão central na gestão das políticas públicas e no fomento do agronegócio no Brasil.

8.1. Dados e Conteúdo

O Portal de Dados Abertos do MAPA (dados.agricultura.gov.br) é uma iniciativa que concentra diversas bases de dados relacionadas à agricultura e pecuária. Alguns dos conjuntos de dados e bases de conhecimento destacados incluem:

- **Sistema de Agrotóxicos Fitossanitários - Agrofit:** Acesso a informações sobre agrotóxicos e afins registrados federalmente para fins agrícolas.
- **Tábua de Risco - Zoneamento Agrícola de Risco Climático (ZARC):** Dados referentes aos períodos de plantio que constam nas Portarias de ZARC, indicando cultivares para plantio por safra, cultura, UF, grupo e região de adaptação.
- **Biblioteca Nacional de Agricultura (BINAGRI):** Informações da biblioteca nacional de agricultura.
- **Base de Conhecimento:** Informações e orientações sobre assuntos de interesse da sociedade relacionados à agropecuária.

8.2. Acesso e Documentação

O acesso ao Portal de Dados Abertos do MAPA é gratuito. O governo federal disponibiliza esses dados como parte da política de transparência e acesso à informação, e o uso da API CKAN não tem custo. Você pode usar os dados para pesquisa, desenvolvimento de aplicações e análises sem a necessidade de pagar taxas de acesso.

É importante não confundir o acesso à API com as taxas de serviços regulatórios que o MAPA cobra por outros procedimentos, como o registro de cultivares (que é o que foi encontrado em algumas buscas). O uso do portal de dados abertos para pesquisa e desenvolvimento não tem custo.

O [Portal de Dados Abertos do MAPA](#) é baseado no sistema de gerenciamento de dados abertos CKAN. Isso significa que ele oferece uma [API CKAN](#) para acesso programático e interação com os dados. A API CKAN é do tipo RESTful e permite diversas operações, como:

- Recuperação de listas de datasets, grupos e tags em formato JSON.
- Obtenção de representações JSON completas de datasets e recursos.
- Busca por pacotes ou recursos.³⁸
- Operações CRUD (Criar, Ler, Atualizar, Deletar) para datasets e recursos, embora as operações de criação, atualização e exclusão geralmente exigem autenticação via API token.

A documentação da API CKAN é abrangente e detalha como fazer requisições HTTP (POST/GET), a estrutura das respostas JSON, versões da API e métodos de autenticação. Os dados do ZARC, por exemplo, podem ser acessados via API e também estão disponíveis em formato CSV.

A utilização da plataforma CKAN pelo MAPA facilita a integração de seus dados em sistemas externos, dada a padronização da API. A disponibilidade de dados sobre agrotóxicos (Agrofit) e zoneamento de risco climático (ZARC) é de particular valor para aplicações de manejo de culturas e gestão de risco agrícola.

8.3. Precificação

O acesso aos dados do Ministério da Agricultura e Pecuária (MAPA) é gratuito.

O Portal de Dados Abertos do MAPA é uma iniciativa do governo federal, e sua API CKAN não tem custo. Você pode usar os dados para pesquisa, desenvolvimento de aplicações e análises sem a necessidade de pagar taxas.

O uso de um token de autenticação, que pode ser solicitado para algumas operações, serve para controle de uso, e não para cobrança. É uma prática comum para monitorar a frequência de acesso e garantir a estabilidade do serviço.

8.4. Avaliação de Volume e Qualidade

8.4.1. Volume

O Portal de Dados Abertos do MAPA apresenta **13 conjuntos de dados** disponíveis, conforme listado diretamente na página principal de datasets do portal.

8.4.2. Qualidade

Embora não haja métricas numéricas específicas publicadas no portal (como acurácia ou completude), alguns pontos reforçam a **qualidade da fonte**:

- **Licença clara e aberta:** todos os dados são disponibilizados sob **licença CC-BY**, o que assegura liberdade de uso com atribuição correta.
- **Diversidade de formatos estruturados:** os dados estão disponíveis em **CSV, PDF, XLSX** — formatos amplamente compatíveis com ingestão e análise automatizada.
- **Infraestrutura robusta (CKAN):** o portal utiliza API CKAN baseada em REST, renomada por sua padronização e confiabilidade, facilitando a exploração, paginação e acesso programático aos recursos.

8.5. Considerações para Coleta de Dados e Inteligência

- **Acesso programático (CKAN):** o portal é **CKAN-based** e expõe a API REST (/api/3/action/*) para **listar/buscar pacotes e recursos** (e CRUD com token quando aplicável). Use package_search com q, rows, start e filtros; atenção a paginação.
- **Formatação:** priorize **CSV/XLSX** para ingestão tabular; **PDF** costuma conter dicionários/explicações (úteis para schema mapping). **XML** (Thesagro) é ideal para taxonomias/glossários.
- **Atualidade operacional:** planeje rotinas distintas conforme o dataset: **semanal** (SIPEAGRO), **mensal** (SIGSIF) e **eventual** (ZARC conforme portarias). Trate “carimbos de atualização” do próprio CKAN como gatilho de coleta.
- **Contexto de negócio:** para analytics, combine:
 - **ZARC** (época/cultivar) + **INMET/CONAB (fora do MAPA)** para séries climáticas e safras.
 - **SIGSIF** para sanidade/abate e **SIPEAGRO** para insumos/registro.

8.6. Desafios e Oportunidades

8.6.1 Desafios

- **Heterogeneidade temporal:** variação na frequência de atualização e presença de conjuntos mais antigos (ex.: Base de Conhecimento, Thesagro) pode exigir curadoria e versionamento de schema.
- **Padronização semântica:** nomes de culturas/categorias podem divergir entre bases (ex.: ZARC vs. SIPEAGRO/SIGSIF); requer dicionários de equivalência.
- **Documentação dispersa:** parte das explicações está nos **PDFs de dicionário** ou em páginas institucionais (AGROFIT/ZARC), o que implica passos adicionais na automação.

8.6.2 Oportunidades

- **Dados licenciados em CC-BY:** favorece reuso comercial/analítico com atribuição.
- **Taxonomia agrícola (Thesagro):** padroniza termos para **data cataloging** e **NLP** (indexação/entidades).
- **Indicadores operacionais:** cruzar **ZARC** (janela de plantio) com **SIGSIF** (abates/doenças) e **SIPEAGRO** (insumos) para painéis de **risco, conformidade e produtividade**.

8.7. Recomendações para Integração

- **Descoberta & Catálogo:** Automatize `package_search` no CKAN para varredura diária/semanal por tags/organização “MAPA”, guardando **hash/last_modified** por recurso. (Doc CKAN API.). Use **Thesagro** para normalizar termos (culturas, pragas, insumos) e enriquecer metadados no seu data lake.
- **Ingestão & Qualidade:** Possuindo os Pipelines por domínio:
 - *Risco climático:* SISZARC/ZARC (CSV) com checagem de portaria e safra;
 - *Inspeção animal:* SIGSIF (CSV) + dicionários de dados;
 - *Registros/insumos:* SIPEAGRO (CSV) em lotes semanais;
 - *Agrotóxicos:* AGROFIT (estrutura de tabelas de produto/ingrediente/registro).

- Validações: schema tests (tipos, ranges), chaves compostas (ex.: UF+ano+espécie), e regras de *slowly changing dimensions* para cadastros (SIPEAGRO). (Baseie-se nos **dicionários SIGSIF** para constraints.)
 - **Modelagem Analítica:** Camada de conformidade com dimensões de *Cultura, UF/Município, Safra, Estabelecimento, Produto e Portaria ZARC*. Fatos: *Abates, Condenações, Registros, Janelas de Plantio*. (ZARC e SIGSIF são as âncoras de calendário e sanidade.)
 - **Orquestração & Monitoramento:** Agende **cron**/Airflow por frequência do dataset (**semanal** SIPEAGRO, **mensal** SIGSIF; **por evento** ZARC conforme novas portarias). Alerta quando last_updated mudar no CKAN.
 - **Governança & Licenciamento:** Publique o **atributo CC-BY** na sua documentação/linhas de rodapé de dashboard e retenha cópia dos **dicionários de dados** no repositório de metadados.
 - **APIs & Tokens:** Para leituras públicas, a API CKAN geralmente **não exige token**; mantenha opção de autenticação para operações administrativas (CRUD) ou portais com limitação. Siga a **documentação oficial** do CKAN e implemente **paginação**.
-

9. Ministério da Agricultura e Pecuária (MAPA)

O febr é uma iniciativa focada na padronização e disponibilização de dados de solo no Brasil.

9.1. Dados e Conteúdo

O [Repositório Brasileiro Livre para Dados Abertos do Solo \(febr\)](#), é um repositório centralizado para armazenar e fornecer dados abertos do solo em formato padronizado e harmonizado. Ele contém 14.477 observações de solo de 232 conjuntos de dados, sendo 42% do sul e sudeste do Brasil. A estrutura é orientada por conjunto de dados, com armazenamento em planilhas acessíveis via aplicação online.

Os dados do febr podem ser utilizados para derivar produtos como bancos de

dados especializados, funções de pedotransferência, guias de recomendação de fertilizantes, sistemas de classificação e mapas detalhados do solo.

9.2. Acesso e Documentação

O febr oferece múltiplas facilidades de download, incluindo formatos de arquivo padrão e conexões com R e QGIS através do pacote febr. Embora não haja uma API RESTful explícita, a conexão via pacotes de programação e a disponibilização de dados em planilhas e arquivos zip permitem o acesso programático. A documentação detalhada está disponível para guiar mantenedores e colaboradores. Os dados podem ser baixados diretamente do [Repositório GitHub](#).

A existência de um repositório dedicado a dados de solo, com foco em padronização e ferramentas de acesso via R e QGIS, é um recurso valioso para análises de aptidão agrícola, manejo de nutrientes e estudos de degradação do solo.

9.3. Precificação

O acesso ao febr é totalmente gratuito. Os dados são disponibilizados sob licenças abertas, permitindo uso, redistribuição e adaptação, desde que sejam citadas as fontes originais. Não há cobrança por download, utilização ou integração com ferramentas como R e QGIS.

9.4. Avaliação de Volume e Qualidade

9.4.1. Volume

O febr contém atualmente **14.477 observações de solo** provenientes de **232 conjuntos de dados**, cobrindo principalmente regiões do Sul e Sudeste do Brasil (42%), mas com representatividade também em outras regiões. Os dados estão organizados por conjunto e podem ser baixados integralmente ou filtrados.

9.4.2. Qualidade

Os dados passam por processos de **padronização e harmonização**, seguindo protocolos científicos. Há metadados detalhados para cada

conjunto, incluindo informações de localização, profundidade, atributos químicos e físicos. No entanto, a distribuição geográfica ainda é desigual, o que pode impactar estudos nacionais de alta resolução.

9.5. Considerações para Coleta de Dados e Inteligência

- A coleta é feita a partir de múltiplas fontes (instituições de pesquisa, universidades, levantamentos governamentais), exigindo processos de integração para uniformizar formatos e nomenclaturas.
- É importante considerar que a granularidade espacial e temporal varia entre conjuntos.
- Para inteligência de dados, recomenda-se uso de ferramentas como **pacote febr (R)**, **QGIS** ou scripts personalizados para cruzamento com variáveis ambientais, climáticas e de uso do solo.
- O alinhamento com padrões internacionais (ex. ISO, OGC) facilita interoperabilidade com outras bases globais de dados de solo.

9.6. Desafios e Oportunidades

9.6.1 Desafios

- Cobertura desigual do território brasileiro, com concentração no Sul e Sudeste.
- Ausência de API RESTful direta, o que limita integrações automatizadas em alguns cenários.
- Necessidade de constante atualização e inclusão de novos levantamentos.

9.6.2 Oportunidades

- Potencial de uso em modelagem preditiva de atributos de solo (Machine Learning).
- Integração com dados climáticos e agrícolas para zoneamento e manejo sustentável.
- Ampliação de parcerias para cobrir áreas pouco representadas, como Norte e Centro-Oeste.

9.7. Recomendações para Integração

- Utilizar o pacote **febr** no R para acesso direto e manipulação dos dados, permitindo automatização de rotinas.
 - Incorporar o dataset no fluxo de trabalho de SIG (como QGIS) para análise espacial e cruzamento com outras camadas.
 - Criar scripts que façam download periódico dos arquivos do GitHub, garantindo uso de versões mais atualizadas.
 - Estabelecer processos de normalização e validação para integração com bases internas ou outros repositórios externos.
 - Para aplicações de grande escala, considerar pré-processamento local para otimizar performance.
-

10. Ministério da Agricultura e Pecuária (MAPA)

A [Leaf Agriculture](https://withleaf.io) (withleaf.io) é uma infraestrutura de dados que oferece uma **API unificada** para conectar e padronizar informações de fazendas, operações de máquinas, limites de talhões, imagens (satélite/drone), clima, prescrições e ativos, a partir de múltiplos provedores.

10.1. Dados e Conteúdo

A Leaf é um *unified farm data API* que padroniza dados agrícolas de vários provedores (máquinas, clima, imagens, etc.) e devolve tudo em respostas **JSON** via REST. Os principais domínios de dados são:

- **Operações de máquinas (tillage, plantio, aplicação, colheita):** dados agregados e padronizados em JSON, com sumários e mapas renderizados.
- **Limites de talhões (Field Boundary Management):** sincroniza/gera/exporta limites entre marcas e mantém histórico de

geometrias (apenas 1 boundary ativo por campo; updates criam novas versões).

- **Arquivos de máquinas (upload manual & conversão):** upload de *zips* até **3 GB**; detecção automática do formato e conversão para um geojson padronizado; suporte a diversos formatos (John Deere 2630/4600+, CNHI, Trimble, AgLeader, Precision Planting, ISOXML, etc.).
- **Monitoramento de lavouras (imagens):**
 - **Sentinel-2** (público): 3–5 dias de revisita; gera 25 produtos por campo (bandas, NDVI/NDRE, RGB, GeoTIFF/PNG).
 - **PlanetScope** (comercial): integração a *assets* como *ortho_analytic_8b_sr*, *ortho_visual*, *UDM2*; *backfill/forward-fill* gerenciados pela Leaf. docs.withleaf.io
- **Clima (Weather):** previsões horárias/diárias e histórico recente (**–5 dias a +10 dias**) em formato unificado; modelos: **GFS, ICON, IFS, AROME/ARPEGE, MET Nordic** (seleção por campo).
- **Prescrições (write-back):** enviar/listar/baixar Rx em **John Deere, CNHI, Climate FieldView, Raven, Trimble, AgLeader** via endpoints beta.
- **Ativos (Assets):** endpoints para criar/obter/atualizar máquinas e associá-las a operações (filtro por fabricante, período, serial, etc.).
- **Irrigação:** dados “as-applied” agregados por dia, com geometrias multi-polígono por profundidade/volume e áreas irrigadas dentro do talhão; suporte a pivôs/sensores.
- **Validação de insumos (Input Validator):** *matching* automático de nomes de produtos (operações/arquivos) a bancos de insumos (ex.: CDMS, Agrian), com *status* e *score* de correspondência.
- **Alertas/Webhooks:** eventos como “nova imagem disponível”, “processamento concluído”, “credenciais expiradas” — com assinatura/segredo para verificação.

10.2. Acesso e Documentação

- **[Documentação principal](#) (REST/JSON):** guia de *quickstart*, amostras em cURL/Node/Python, e referência por serviço.
- **Autenticação ao API:** *Bearer token* (Admin Token) para seu **API Owner**; endpoints ficam sob <https://api.withleaf.io/services/...>
- **Conexão às contas/provedores dos produtores:**
 - **[Leaf Link](#)** (widget pronto) para **conectar provedores** (OAuth) e **fazer upload de arquivos** com pouquíssimo código.

- [Magic Link](#) para convidar usuários por e-mail/SMS e conduzir a autorização sem precisar embutir o widget no seu app.
- **Principais referências por domínio:**
 - [Limites de talhões](#) – Overview & Endpoints
 - [Operações de campo](#) – Endpoints
 - [Upload manual & conversão](#) – Overview & Endpoints
 - [Monitoramento de lavouras](#) – Endpoints / Sentinel / Planet
 - [Clima](#) – Overview & Endpoints
 - [Prescrições](#) – Endpoints (beta)
 - [Irrigação](#) – Overview & Endpoints
 - [Ativos](#) – Endpoints (beta)
 - [Alertas/Webhooks](#) – Overview & Lista de eventos
- **Status do sistema:** página pública de status/uptime e SLA oficial.
- **Para devs (centro de recursos):** *landing* de desenvolvedores e tutoriais passo-a-passo (ex.: John Deere/AgLeader/Trimble).
- **Formato & método de acesso:** **API REST** (JSON). Não há *web scraping*; quando arquivo, o upload é via **endpoints** ou **Leaf Link** (ZIPs de arquivos proprietários).

10.3. Precificação

- **Modelo:** [preço](#) por acre (volume-based), com descontos por compromisso de acres; *tiers* Core, Pro e Enterprise (add-ons como Planet Imagery e Leaf Workflows).
- **Contratação:** via [time comercial](#) ou AWS Marketplace (facilita *procurement*/faturamento em nuvem).
- **SLA:** disponibilidade mensal 99,95% conforme SLA.
- **Observação:** valores numéricos não são públicos no site; o contato comercial define faixas por uso/acre e *add-ons*.

10.4. Avaliação de Volume e Qualidade

10.4.1. Volume

- **Máquinas e operações:** suporta as principais [marcas](#) (John Deere, CNHI, Trimble, AgLeader, Raven, Climate FieldView, etc.), “*um só API*” para *multi-marcas*.
- **Imagens:** **Sentinel-2** (cobertura global, 3–5 dias de revisita) e [Planet](#) (alta resolução, assinaturas gerenciadas pela Leaf).
- **Clima:** múltiplos modelos globais/regionais (GFS, ICON, IFS, AROME/ARPEGE, MET Nordic) para qualquer coordenada/talhão.

- **Escalabilidade:** a Leaf reporta operar com ~99,5% de [uptime](#) e escala automática em picos (plantio/colheita).

10.4.2. Qualidade

- **Tradução e padronização:** converte formatos proprietários para **JSON padronizado**; **mescla** e **processa** arquivos em saídas consistentes.
- **Limites versionados** e histórico de geometrias, garantindo reprodutibilidade.
- **Imagens processadas** (NDVI/NDRE/RGB; GeoTIFF/PNG) e *cloud mask* integrada para cálculo de cobertura de nuvens.
- **Clima** com janela temporal clara (-5 a +10 dias) e variáveis horárias/diárias em formato único.
- **Input Validator** para normalizar nomes de produtos (insumos), com *score* e *status* (Predicted/Validated/etc.).

10.5. Considerações para Coleta de Dados e Inteligência

- **Modelo de dados centrado no “Leaf User”:** crie um **Leaf User** por cliente/agricultor para isolar permissões e dados.
- **Onboarding & consentimento:** use **Leaf Link/Magic Link** para autorização OAuth nas contas de provedores (reduz fricção e erros).
- **Qualidade de limites:** respeite a regra de **1 boundary ativo** por campo; novos *uploads* geram versões (ótimo para auditoria/ML).
- **Ingestão híbrida:** combine **integrações diretas** e **upload manual** (ZIP ≤ 3 GB) para cobrir frotas desconectadas.
- **Eventos & reatividade:** cadastre **webhooks** para eventos críticos (imagem nova, processamento concluído, credenciais expiradas) e evite *polling*.
- **Padronização semântica de insumos:** automatize *matching* com **Input Validator** para análises de uso de produtos, compliance e custos.
- **Clima/Imagens:** defina política de **cloud cover** e **modelo meteorológico preferido** por cultura/região (Leaf escolhe modelos por acurácia por campo, você pode combinar).

10.6. Desafios e Oportunidades

10.6.1 Desafios

- **Dependência de permissões dos provedores** (escopos OAuth, *orgId* no JD, *companyId* na CNHI, etc.) para “read” e “write-back” (prescrições).
- **Heterogeneidade de arquivos** (estruturas alteradas pelo usuário; nem sempre intactas), a Leaf tenta reparar, mas alguns *batches* podem falhar.
- **Custo de imagens comerciais (Planet)**, requer *add-on*.

10.6.2 Oportunidades

- **Tempo-de-mercado:** um único API para multi-marcas reduz meses de integração.
- **“Write-back” de Rx:** diferencial competitivo (mandar prescrições para vários OEMs).
- **Data-sharing entre empresas: Leaf Connect** para compartilhar dados com parceiros sem construir seu próprio API.
- **Compliance/transparência: Ag Data Transparent** (políticas claras de dados agrícolas).

10.7. Recomendações para Integração

- **Desenho inicial**
 - **Entidades:** *LeafUser* → *Farms* → *Fields* → *Boundaries* → *Operations/Files/Assets*. Use o ID do *Merged Field* como chave comum quando integrar múltiplos provedores. docs.withleaf.io
 - **Configurações por cliente:** utilize o **Configuration Service** para preferências (unidades, provedores ativos, etc.). withleaf.io
- **Onboarding & coleta**
 - **Leaf Link/Magic Link** para conectar **John Deere/CNHI/Trimble/AgLeader/Raven/Climate** (menos código e melhor UX). docs.withleaf.io+1
 - **Upload manual** para frotas sem telemetria; respeite o **limite de 3 GB** por ZIP e informe o **provider** quando souber (Leaf detecta se “Other”). docs.withleaf.io
- **Pipelines de dados**
 - **Webhooks (Alerts):** cadastre eventos para processar **assim que** imagens chegarem/arquivos concluírem/credenciais expirarem. Inclua verificação de assinatura. docs.withleaf.io

- **Operações & ativos:** cruze **Operations** com **Assets** para análises de eficiência de máquinas/operadores. withleaf.io/docs.withleaf.io
- **Imagens:** para **Sentinel**, armazene NDVI/NDRE (GeoTIFF) e PNGs “absolutos” para comparações temporais. Considere **Planet** para maior resolução. docs.withleaf.io+1
- **Clima:** escolha modelos padrão por região/cultura e habilite *fallback* entre modelos suportados. withleaf.io
- **Insumos:** aplique o **Input Validator** para limpar nomes de produtos de operações históricas e novas. withleaf.io
- **“Write-back” (Prescriptions)**
 - **John Deere:** exigir `organizationId`; **CNHI:** `companyId`; formatos aceitos (SHP/DBF/SHX em ZIP; variações por OEM). docs.withleaf.io
 - **Confiabilidade & governança**
 - **SLA e status:** monitore status.withleaf.io e ajuste *retries* conforme SLA 99,95%; use *backoff* nos *polls* e prefira webhooks. status.withleaf.io
 - **Transparência de dados:** consulte “Your data with Leaf” e o selo **Ag Data Transparent** nas políticas de dados. withleaf.io
 - **Procurement:** considere contratar via **AWS Marketplace** para simplificar billing.

11. OpenET

[OpenET](#) é uma plataforma sem fins lucrativos que disponibiliza estimativas de evapotranspiração (ET) baseadas em satélite em escala de talhão para apoiar manejo de irrigação e gestão hídrica; o consórcio é liderado por NASA, Desert Research Institute (DRI) e Environmental Defense Fund (EDF), com operação em Google Earth Engine.

11.1. Dados e Conteúdo

A plataforma fornece séries diárias e mensais de ET em 30 m (resolução Landsat), além do valor “ensemble” (média filtrada por MAD) de seis modelos (ALEXI/DisALEXI, eeMETRIC, geeSEBAL, PT-JPL, SIMS e SSEBop). Também

expõe variáveis de apoio como ETo (referência grama), ETr (referência alfafa), ETof/ETrf (fração de referência), NDVI, contagens de modelos e precipitação (PR), com unidades configuráveis (mm ou polegadas). Em Earth Engine, há coleções mensais públicas (licença CC-BY-4.0) cobrindo o CONUS com disponibilidade de 1999-10 a 2024-12. Na API, além de mensais, é possível consultar dados diários por geometria.

11.2. Acesso e Documentação

Para exploração manual, use o Data Explorer (visualização e download limitado).

Para uso programático, a [documentação oficial da API](#) traz “Quick Start”, endpoints Geodatabase (consultas pré-computadas) e Raster (consultas customizadas, inclusive GeoTIFF), parâmetros aceitos e exemplos. É necessário [criar conta](#) para obter a chave de API

Há integração opcional com Earth Engine e suporte a exportação para [Google Drive](#), além de escolher a fonte de referência ET (CIMIS na Califórnia ou gridMET) e usar a previsão de 7 dias de ETo (NOAA FRET) como referência. Para históricos mensais prontos, também é possível consumir diretamente as coleções públicas do OpenET no Google Earth Engine.

11.3. Precificação

A visualização no Data Explorer é **gratuita** e o download de volumes limitados também.

Para acesso em larga escala existe a API, com níveis de cota e possibilidade de níveis superiores ao vincular o Earth Engine. A página de FAQ deixa claro que o modelo busca manter o acesso público gratuito e que o uso em grande escala pela API é o caminho para operações pagas/sustentação do serviço.

Os detalhes comerciais além das cotas públicas devem ser tratados diretamente com a equipe.

11.4. Avaliação de Volume e Qualidade

11.4.1. Volume

O volume de dados na OpenET é considerável, abrangendo vastas áreas geográficas e um histórico de tempo significativo

11.4.2. Qualidade

A qualidade dos dados é considerada alta, pois a plataforma utiliza uma variedade de modelos validados por cientistas e especialistas, e os dados de satélite são de fontes confiáveis. A metodologia é revisada por pares e os dados são constantemente atualizados, garantindo precisão nas estimativas.

11.5. Considerações para Coleta de Dados e Inteligência

Para análises operacionais, privilegie o valor ensemble de ET e registre, junto, ETo/ETr e a fração (ETof/ETrf) para normalizar sazonalidade, além de precipitação (PR) da própria coleção. Isso facilita calcular ET líquida e orçar água com outros insumos locais (chuva efetiva, lençol raso, calendário fenológico).

Trate lacunas por nuvens conforme indicado em “Known Issues” e monitore a latência, mensais tendem a estabilizar até seis semanas após o mês de interesse. Padronize unidades (mm) na ingestão e use metadados da API (COUNT/MODEL_COUNT) para qualificar incerteza por período/local.

11.6. Desafios e Oportunidades

11.6.1 Desafios

O principal desafio para o seu caso no Brasil é a cobertura: OpenET está focado no oeste dos EUA, embora o consórcio pretenda expandir e alongar o arquivo histórico; diferenças entre modelos aumentam em ET muito baixa, corpos d'água e ambientes áridos, exigindo diagnóstico local e, às vezes, seleção de um subconjunto de modelos.

11.6.2 Oportunidades

há grande oportunidade de integrar ET de 30 m em plataformas de contas de água, trading, monitoramento de aquíferos e apoio à irrigação, com acesso gratuito para exploração e coleções públicas no Earth Engine para prototipagem reprodutível.

11.7. Recomendações para Integração

Para protótipos e estudos, consuma as coleções mensais do OpenET no Earth Engine (processamento e reprodutibilidade sem infraestrutura própria) e cruze com clima/solo/culturas locais.

Para operação, use a API Geodatabase quando precisar de agregados rápidos por talhão/ID e a API Raster quando precisar de diários/GeoTIFF ou geometrias customizadas, armazenando em seu data lake as séries de ET, ETo/ETr, ETof/ETrf e PR com carimbo de versão/modelo. Vincule sua conta do Earth Engine à OpenET para ampliar cotas, monitore limites mensais e trate os casos listados em “Known Issues” com regras de negócio (ex.: mascaramento de água, filtros por MODEL_COUNT).

Como a cobertura atual é dos EUA, se o objetivo for o agro brasileiro, use OpenET como referência metodológica e de benchmarking, e combine, no curto prazo, fontes regionais/alternativas para ET até que o escopo espacial da OpenET seja expandido oficialmente.

12. AgroMonitoring

O [AgroMonitoring](#) (Agro API) é um serviço da OpenWeather/Extreme Electronics Ltd. que entrega, via API, **imagens de satélite, índices de vegetação e dados meteorológicos/solo** para monitoramento agrícola em tempo quase real e histórico, operando por **polígonos** (talhões) definidos pelo usuário.

12.1. Dados e Conteúdo

A **API** cobre **satélites (Sentinel-2, Landsat-8/9, MODIS)** com geração de **NDVI, EVI, EVI2, NDWI, DSWI, NRI**, além de imagens **true/false color** e **estatísticas zonais** (mín., máx., mediana, desvio-padrão) por polígono.

As respostas incluem **PNG (imagens/tiles), GeoTIFF e JSON de metadados**. O **NDVI histórico** é atualizado **diariamente** combinando Landsat, Sentinel e MODIS, com **arquivo desde 2001** no plano corporativo. Também entrega **tempo atual, previsão de 5 dias/3 h, histórico, temperatura e umidade do solo** (superfície e 10 cm), **acumulados de precipitação e de temperatura (GDD)**, e **UV**. Em interface de dashboard, cada cena exibe **data, nebulosidade, resolução (m/px) e fonte**.

Na FAQ constam **resoluções mínimas de 10 m (Sentinel-2) e 30 m (Landsat)**.

12.2. Acesso e Documentação

O fluxo típico é **registrar-se e obter a [API](#) key, criar um polígono** (GeoJSON, ordem lon,lat) com a **Polygons API**, e então usar a **API de busca de imagens** (passo 1) seguida do **download de produtos/estatísticas** (passo 2).

Há guias “How to start” e documentação detalhada de **imagens, solos e tempo**; o retorno de busca já traz os **links “data” para GeoTIFF** de cada preset. A documentação oficial e o dashboard têm tutoriais passo a passo.

12.3. Precificação

Há camadas **Free (£0), Starter (£20/mês), Small Kit (£200/mês)** e **Corporate (sob consulta)**.

Os **planos** combinam **taxa fixa por área total de polígonos/mês** (ex.: 1 000 ha no Free; 4 000 ha no Starter; 20 000 ha no Small Kit) e **cobrança por excedente** (ex.: £0,02/ha no Starter.

£0,01/ha no Small Kit). Limites operacionais incluem **requisições por minuto a dados de satélite** (<60 Free; <600 Starter; <3 000 Small Kit) e **chamadas/dia a dados de clima** (<500 Free; <1 000 Starter; <10 000 Small Kit).

Recursos variam por plano: **NDVI histórico, solo histórico, acumulados e clima histórico** ficam disponíveis a partir dos pagos.

A profundidade padrão do **clima histórico** é **1 ano** (arquivo completo sob consulta/corporate). Licenças declaradas: **CC BY-SA 4.0** para mapas/APIs e **ODbL** para dados/bancos (há customização no corporativo).

12.4. Avaliação de Volume e Qualidade

12.4.1. Volume

Em **satélite**, o serviço reporta **dados quase em tempo real** e, no dashboard, disponibilidade típica de novas cenas a cada **2–4 dias** por polígono, dependendo da órbita/sensor e **cobertura de nuvens**; cada cena indica **fonte, resolução e nebulosidade**, permitindo filtrar qualidade antes do processamento.

12.4.2. Qualidade

A **qualidade espacial** é compatível com **10 m (Sentinel-2)** e **30 m (Landsat)**.

O **NDVI diário** combina múltiplas constelações, aumentando a cadência temporal. Em **clima/solo**, há **atualizações <1 h** nos planos pagos para clima e **solo 2×/dia** para temperatura/umidade, com **acumulados** e **UV**, limites diários de chamadas orientam o volume processável por dia.

12.5. Considerações para Coleta de Dados e Inteligência

Defina **polígonos** coerentes com a **resolução** ($\geq 10\text{--}30$ m) para evitar mistura de culturas/bordas.

Filtre por nebulosidade e cobertura antes de calcular indicadores; prefira **estatísticas zonais** da própria API para padronizar séries e reduzir ruído; **normalize** séries quando misturar sensores.

Trate unidades (ex.: **Kelvin** em solo/clima) e **timestamps UNIX/UTC**; gerencie **rate limits** e **quotas por área** com filas e **cache** de GeoTIFF/tiles. Para lacunas por nuvens, use **acumulados climáticos** e **modelos de interpolação** como features substitutas.

12.6. Desafios e Oportunidades

12.6.1 Desafios

Nuvens e sazonalidade orbital podem criar **buracos temporais** mesmo com fusão de sensores; no plano **Free**, **histórico/meteo** são limitados e podem restringir análises multi-safra.

12.6.2 Oportunidades

Por outro lado, a combinação de **NDVI diário**, **acumulados térmicos/chuva**, **umidade/temperatura do solo** e **estatísticas zonais** por talhão abre espaço para **alertas de estresse**, **estratégias de irrigação/fertilização**, **análise de vigor intra-talhão** e **modelagem de produtividade** com menor custo de aquisição de dados.

12.7. Recomendações para Integração

Modele o pipeline como “**polígono-centro**”: persistir **polygon_id**, agendar a **busca de cenas** e, quando houver janela sem nuvens, **baixar GeoTIFF** e salvar as **estatísticas zonais**.

Complemente dias sem imagem com **séries de clima/solo** e **acumulados**.

Padronize **CRS/escala** ao cruzar com camadas de gestão; monitore **consumo por área** e **limites de requisições**.

Registre **atribuição/licença** conforme **CC BY-SA/ODbL** e políticas do plano. Para começar rápido, siga o guia de “**How to start**”, **Polygons API (GeoJSON lon,lat)** e o **fluxo 2-passos** de imagens. Para análises avançadas, utilize os links “**data**” para **GeoTIFF** e processe no seu GIS/ETL.

13. FAOSTAT

[FAOSTAT](#) é o banco estatístico global da FAO que oferece acesso gratuito a séries históricas de agricultura e alimentos para mais de 245 países/territórios, cobrindo de 1961 até o ano mais recente disponível. É amplamente usado para produção, comércio, preços, uso da terra, emissões e indicadores climáticos.

13.1. Dados e Conteúdo

A base abrange dezenas de domínios temáticos (p.ex., **Crops and Livestock Products – QCL**, comércio, uso da terra, florestas, insumos, balança alimentar, segurança alimentar, preços, emissões e clima). Exemplos: o domínio **Temperature Change** publica tendências de anomalia de temperatura desde 1961 e foi atualizado até 2024.

Os domínios de **Preços** incluem séries **mensais e anuais** de preços ao produtor e índices; o domínio **Emissions Totals** traz estimativas anuais de CH₄, N₂O, CO₂ e equivalentes por país com notas metodológicas próprias. Em todos os domínios, metadados e notas metodológicas acompanham os dados.

13.2. Acesso e Documentação

Há duas formas principais de acesso:

- [Portal web \(exploração e download manual\)](#): Navegação por domínio com filtros e exportação; a página “Statistics” centraliza a porta de entrada do FAOSTAT.
- [API pública REST \(JSON\)](#): Baseada nos serviços “FENIX”, sem autenticação para consultas públicas, com endpoints por domínio e endpoints de **definitions** para descobrir códigos de áreas/itens/elementos. O repositório público descreve os serviços, e também no [diretório](#).

13.3. Precificação

O acesso é gratuito. As “FAO Statistical Database Terms of Use” indicam licença **CC BY 4.0**, exigindo **atribuição** adequada (ex.: “FAO. [ano]. FAOSTAT: [domínio]. Acessado em [data]”). Alguns conteúdos mais antigos podem ter políticas legadas (p.ex., orientações sobre licenças para bases anteriores a

30/11/2019); verifique sempre o rodapé/ícone de licença do domínio específico. Para usos comerciais fora do escopo, a FAO orienta solicitar autorização.

13.4. Avaliação de Volume e Qualidade

13.4.1. Volume

Cobertura **global (245+ países/territórios)** e séries **longas (1961–atual)**; a oferta inclui **dezenas de domínios** ($\approx 70+$) com [arquivos bulk](#) volumosos em CSV “normalized” (longo), geralmente acompanhados de colunas de **flags** e **notes**.

13.4.2. Qualidade

A produção segue o **FAO Statistical Quality Assurance Framework (SQAF)** (relevância, exatidão, tempestividade, comparabilidade, coerência, acessibilidade), com **notas metodológicas por domínio** (ex.: Emissions Totals) e calendários de revisão anual.

13.5. Considerações para Coleta de Dados e Inteligência

Para **ingestão em larga escala**, prefira **Bulk (CSV)** para *backfill* histórico e **API** para atualizações incrementais.

Use os **dicionários da API** para padronizar **códigos de área (M49/ISO3)**, **itens (FAO/CPC)** e **elementos**. Trabalhe no formato “normalized” (longo) e preserve **flags/metadados** (oficial, estimado, imputado) para análises de qualidade.

Documente as **unidades** por elemento (toneladas, hectares, índices, etc.) e trate **mudanças de classificação** ao longo do tempo (ex.: itens agregados).

13.6. Desafios e Oportunidades

13.6.1 Desafios

- **Defasagem temporal:** muitos domínios têm atualização **anual**.
- **Comparabilidade:** diferenças metodológicas nacionais e revisões históricas.
- **Classificações e códigos** evoluem (FAO/CPC), exigindo *crosswalks*.

- **Quebras de série** e agregações que podem levar a **dupla contagem** se o usuário não observar as notas (caso clássico com China em agregações históricas, tratado por utilitários em clientes).

13.6.2 Oportunidades

- **Ampla cobertura** para *benchmarks* e *nowcasting* híbrido (combinar FAOSTAT com satélite/clima/local).
- **Séries longas** para modelagem causal/estrutural.
- Integração com **SDG Data** e o novo **FAODATA Explorer (beta)** para exploração e *pipelines* de disseminação.

13.7. Recomendações para Integração

- **Arquitetura de ingestão em duas vias: Bulk** (job inicial) → armazene *staging* em formato longo com dicionários; depois **API** (jobs incrementais) por domínio/ano.
- **Dicionários e normalização**: mantenha tabelas mestras de **área (M49/ISO3)**, **itens (FAO/CPC)** e **elementos** vindas dos endpoints *definitions*; aplique *surrogate keys* estáveis e *versionamento* de classificações.
- **Metadados de qualidade**: propague **flags** e **notes** até as camadas analíticas. Construa *dashboards* de *data quality* alinhados ao **SQAF** (ex.: completude, revisões, pontualidade por domínio).
- **Governança de atualizações**: use campos de **release/year** expostos por clientes (ex.: `list_datasets_df()` do pacote Python) ou por catálogos/bulk para disparar *ingests* seletivos.
- **Camada semântica**: padronize **unidades** por elemento, crie *views* com medidas derivadas (produtividade, intensidades, balanços) e mantenha *crosswalks* para integrar FAOSTAT com outras fontes (INMET/ANA/CONAB/EO).

14. USDA Foreign Agricultural Service (FAS)

O [USDA Foreign Agricultural Service \(FAS\)](#) é a agência do governo dos EUA

dedicada à inteligência e ao comércio agrícola global; ele reúne bases como **PS&D** (produção, suprimentos e distribuição), **Export Sales** (vendas semanais dos EUA), **GATS** (comércio exterior dos EUA) e o **Crop Explorer** (condições agroclimáticas por satélite), formando um ecossistema de dados e análises para commodities e mercados. [FAS Apps+2FAS Apps+2ipad.fas.usda.gov](#)

14.1. Dados e Conteúdo

O **PS&D (Production, Supply & Distribution)** concentra as estimativas oficiais do USDA, produção, consumo, estoques, importações e exportações por commodity/país, revisadas **mensalmente** em alinhamento com o calendário do WASDE; os dados podem ser consultados online e baixados para planilhas. A cobertura histórica alcança **décadas** (para muitas commodities, desde 1960).

O **Export Sales Reporting** traz dados **semanais** de vendas líquidas e embarques por commodity/mercado (trigo, soja, milho, algodão, carnes, etc.), com relatórios e séries históricas em nível detalhado.

O **GATS (Global Agricultural Trade System)** disponibiliza estatísticas de **exportação e importação dos EUA** e integra fontes como **U.S. Census** e **UN Comtrade**; há consultas padrão/avançadas, relatórios e **API de dados**.

O **Crop Explorer (IPAD/FAS)** provê mapas e séries de **precipitação, temperatura, anomalias, número de dias de chuva/seca, umidade do solo (inclui produtos derivados como CPC e SMAP), e NDVI/MODIS**, oferecendo diagnóstico rápido de condições de safra no mundo. O **GAIN (Global Agricultural Information Network)** reúne inteligência de mercado produzida pelos adidos agrícolas do FAS (relatórios temáticos e por país). O pipeline publica milhares de relatórios desde a década de 1990 e serve como insumo qualitativo que complementa as séries do PS&D.

14.2. Acesso e Documentação

O ponto único é o [FAS Open Data Services \(Swagger\)](#), que documenta as APIs de **PS&D**, **Export Sales (ESR)** e **GATS**, além do [portal](#) da fas. Para uso programático é necessário **API key do api.data.gov**; a documentação indica como obter a chave e os principais endpoints (incluindo consulta de datas de

release). Para **PS&D**, além da API, existe o **PS&D Online** com consulta web e **download de arquivos** para planilhas/bancos.

Para **Export Sales**, a página oficial publica o relatório semanal, links para séries históricas, ferramentas de consulta e referência direta ao **ESR Data API**.

Para **GATS**, o portal oferece consultas e link para a **GATS Data API**.

Para **Crop Explorer**, o acesso é via **visualização web** de mapas e imagens com descrição das fontes/modelos; é ideal para análise contextual e monitoramento, não como feed numérico estruturado.

Como usar (API): com a chave do api.data.gov, as APIs aceitam autenticação por cabeçalho (ex.: *X-Api-Key*) ou parâmetro, conforme o manual do api.data.gov, e possuem limites de requisição. Consulte o Swagger do FAS para formatos/rotas específicas por produto.

14.3. Precificação

A **acesso é gratuito** (open data). O uso via api.data.gov está sujeito a **rate limits padrão** (por exemplo, 1.000 req/hora, podendo variar por serviço) e exige uma chave individual.

14.4. Avaliação de Volume e Qualidade

14.4.1. Volume

O PS&D cobre um amplo conjunto de commodities agrícolas e países, com séries históricas extensas (muitas desde 1960). O Export Sales fornece granularidade semanal por commodity/destino desde fins dos anos 1990; o GATS cobre todo o comércio agro dos EUA (exportações/importações). O GAIN adiciona ~2.000 relatórios/ano desde 1995, enriquecendo o contexto.

14.4.2. Qualidade

O PS&D é o **repositório oficial** do USDA, **revisado mensalmente** por comitê interagências (WAOB/FAS/ERS/FSA/AMS), com atualizações sincronizadas ao WASDE; GAIN traz análises de postos no exterior e **não é dado oficial**, embora seja insumo importante.

14.5. Considerações para Coleta de Dados e Inteligência

Trate com cuidado **anos-safra (Marketing Year)**, que variam por país/commodity e diferem de ano-calendário; as tabelas do PS&D equilibram fluxos nos respectivos anos locais e alguns relatórios trazem notas específicas de MY por produto/país. Registre o **Release Schedule** e as **datas de publicação** para reproducibilidade.

Para séries **semanais** (Export Sales), considere sazonalidades fortes e revisões; para **mensais** (PS&D), acompanhe cortes de produção e mudanças metodológicas divulgadas nos circulares/relatórios.

Ao usar **inteligência de GAIN**, lembre-se de que são análises de campo e podem divergir de números oficiais do PS&D; use-as para **contexto, risco regulatório, logística** e sinais de curto prazo.

14.6. Desafios e Oportunidades

14.6.1 Desafios

Esquemas diferentes entre APIs (**PS&D, ESR, GATS**), necessidade de chave e **rate limits**, mapeamento de **códigos de commodity** e harmonização de **MY vs. CY**. Além disso, o **Crop Explorer** é voltado a **imagens/mapas**, exigindo tratamento específico se você quiser derivar séries numéricas.

14.6.2 Oportunidades

Combinar **Export Sales (semanal)** com **PS&D (mensal)** para *nowcasting* de comércio/saldos; usar **NDVI/umidade do solo** do Crop Explorer para explicar revisões de produção; integrar o **GATS** para granularidade do comércio dos EUA por código HS/país, enriquecendo análises de competitividade e share.

14.7. Recomendações para Integração

Implemente um **modelo canônico** de dados por *commodity–país–ano* com chaves de mapeamento (códigos PS&D e HS), guardando **data/hora de release** de cada atualização para versionamento; o Swagger do FAS inclui endpoint para **datas de release**, útil para *backfill* e reprocessamento.

Programe uma **rotina semanal** (5ª feira, horário de Washington) para **Export Sales** e uma **rotina mensal** para **PS&D**, sincronizada às datas do WASDE; registre alterações e notas metodológicas dos circulares.

Para **acesso programático**, obtenha a chave no **api.data.gov** e autentique conforme o manual (ex.: cabeçalho **X-API-Key**); se atingir limites, solicite ajuste ao órgão responsável ou implemente **cache** e **fila** de chamadas.

Use o **GATS API** para validar séries de comércio dos EUA e o **GAIN** para sinais qualitativos (mudanças regulatórias, logística, clima local) que expliquem quebras de tendência. Para **Crop Explorer**, consuma **imagens e metadados** como camadas de contexto (não como feed tabular), integrando-as ao pipeline analítico/BI.

15. Copernicus Data Space Ecosystem (CDSE)

O [Copernicus Data Space Ecosystem](#) é o novo portal unificado da ESA/União Europeia para acesso aberto e instantâneo aos dados do programa Copernicus, oferecendo navegação, download e processamento em nuvem (APIs e ferramentas) e substituindo o antigo Copernicus Open Access Hub (SciHub).

15.1. Dados e Conteúdo

O CDSE distribui todo o acervo e as novas aquisições das missões Sentinel (S1, S2, S3 em terra/zonas costeiras e S5P), além de dados complementares como Copernicus DEM e mosaicos WorldCover; também agrega coleções federadas e de Missões Contribuintes, além de dados de Serviços Copernicus (ex.: Land, Atmosphere, Climate). Para agricultura, isso significa acesso a radar (S1) útil

para umidade de solo/estrutura, óptico multiespectral (S2, 10–60 m) para índices de vegetação (NDVI/EVI via processamento), temperatura de superfície/cores da terra (S3) e camadas temáticas de serviços Copernicus.

Os produtos Sentinel são gratuitos para uso científico e comercial. A publicação é contínua e muito rápida: os produtos são disponibilizados **até ~15 min** após o processamento pela ESA, e o ecossistema opera com armazenamento em escala de dezenas de petabytes.

15.2. Acesso e Documentação

Via navegador, consultar catálogos/baixar por [APIs](#) e processar sob demanda no próprio ambiente do CDSE:

- [Documentação Geral](#): visão de dados e APIs em um único lugar.
- [STAC API \(busca por área/tempo/propriedades\)](#): endpoint público do catálogo.
- [OData \(catálogo/metadata legado-compatível\)](#): REST para listar produtos e metadados; útil a quem migrou do SciHub.
- [S3 \(alto desempenho/parallelismo\)](#): acesso a objetos via protocolo S3 para ingestão/bulk em pipelines externos/HPC.
- [openEO \(processamento sob demanda com créditos\)](#): construa fluxos (visual ou via API) e rode no acervo completo.
- [Sentinel Hub APIs \(Process, Catalog, OGC WMS/WCS/WMTS\)](#): renderização, estatística e tiles prontos (ex.: NDVI) e integração direta em GIS.
- [QGIS/ArcGIS](#): plugin oficial do CDSE/Sentinel Hub para buscar e visualizar imagens direto no QGIS; ArcGIS Pro conecta via STAC/S3.

Observação de migração: o Copernicus Open Access Hub (SciHub) foi descontinuado em outubro/2023; todo acesso deve ocorrer pelo [CDSE](#).

15.3. Precificação

O [dado em si é livre e gratuito](#) (política open data Copernicus). Serviços do ecossistema oferecem **camadas gratuitas com cotas mensais** (requisições, Processing Units, produtos processados, **créditos openEO**) para uso regular.

Ao exceder certos limites, sua velocidade pode ser reduzida ou você precisará ampliar a cota conforme elegibilidade (ex.: usuários de Serviços Copernicus/projetos UE) ou via ofertas comerciais/Marketplace. *Exemplos oficiais:*

- Sentinel data “gratuito para todos os usuários” (científicos e comerciais).
- Quotas e limitações por conta (requisições/mês e/min, PUs/mês, produtos processados, créditos e volume em 30 dias).
- openEO: **10.000 créditos/mês** para “Copernicus General”; **20.000/mês** para perfis elegíveis.

15.4. Avaliação de Volume e Qualidade

15.4.1. Volume

infraestrutura em larga escala (ex.: ~120 PB de capacidade e ~84 PB de dados em 04/2024; “maior instalação Ceph do mundo” à época), com ingestão/publicação rápida (~15 min após processamento ESA).

15.4.2. Qualidade

produtos oficiais Sentinel (níveis L1/L2) com calibração/validação da ESA; S1 GRD também em **COG_SAFE** para acesso escalável; metadados STAC estão em evolução contínua (versão 1.0 implementada). Para S2, imagens em **JPEG2000** (SAFE) e, para S3, produtos baseados em **NetCDF** segundo especificações PDGS.

15.5. Considerações para Coleta de Dados e Inteligência

- **Escolha de missão/nível:** S1 (radar, todo tempo) para umidade/estrutura; S2 (óptico) para NDVI/EVI e mapeamentos; S3 para LST/OC/cores da terra; inclua DEM/WorldCover conforme necessário.
- **Latência/atualidade:** em geral muito rápida, mas acompanhe eventos (ex.: atrasos temporários S2 L2A noticiados e resolvidos).
- **Filtragem e recorte:** use STAC para recortes por **bbox/data/nuvem**; processe índices no openEO/Process API e exporte como GeoTIFF/COG quando possível.
- **Cotas e robustez:** monitore limites (requisições, PUs, créditos e tráfego 30 d) e implemente **retry/backoff** e cache.

- **Integração GIS:** para visualização/operacionalização, conecte **WMS/WCS/WMTS** ou o **plugin QGIS**; para pipelines de alto volume, prefira **S3**.
- **Rastreabilidade e compliance:** registre/cheque hashes e proveniência com o Traceability.

15.6. Desafios e Oportunidades

15.6.1 Desafios

Altíssimo volume (armazenamento/egresso), cotas compartilhadas entre usuários, **metadados STAC em evolução**, eventuais indisponibilidades pontuais; necessidade de migração de ferramentas antigas (ex.: *sentinel*sat do SciHub).

15.6.2 Oportunidades

Publicação quase em tempo real, **processamento sob demanda (openEO/SH Process API)**, formatos otimizados (**COG**), integração nativa com **QGIS/ArcGIS**, **Data Workspace/JupyterLab**, e **serviços comerciais** para escalar workloads quando necessário.

15.7. Recomendações para Integração

- Consulte **STAC** para listar itens/ativos e armazene apenas metadados/links; para lotes grandes, recupere via **S3**.
 - **Processamento:** rode NDVI/EVI/estatísticas usando **openEO** (créditos mensais) ou **Sentinel Hub Process API**, preferir exportar **COG** para análises repetíveis.
 - **Orquestração:** trate **cotas** (limites por minuto/mês e por volume 30d) com fila/retry e logging; configure alertas para **service news**.
 - **Visualização e entrega:** publique camadas **WMS/WCS/WMTS** para equipes e conecte o **plugin QGIS**; para auditoria, gere **traces** dos produtos distribuídos.
 - **Migração:** se usava SciHub/sentinelsat, troque para **STAC + pystac-client** e/ou **sentinelhub-py** conforme a necessidade de processamento e OGC.
-

16. FAOLEX (Organização das Nações Unidas para Alimentação e Agricultura – FAO)

O [FAOLEX](#) é o banco de dados jurídico-político da FAO que reúne leis, regulamentos, políticas e acordos internacionais sobre agricultura, alimentação e recursos naturais do mundo todo, com resumos, indexação temática e, quando disponível, o texto completo em PDF/TXT. É atualizado de forma contínua e cobre mais de 200 países/territórios.

16.1. Dados e Conteúdo

O acervo abrange legislação e políticas sobre agricultura e desenvolvimento rural, plantas cultivadas, pesca e aquicultura, pecuária, floresta, solos e água, alimentação e nutrição, meio ambiente, biodiversidade/ecossistemas, resíduos/perigosos, clima e energia, além de constituições e acordos internacionais.

Cada registro traz metadados padronizados (ID LEX-FAOC..., URLs do registro e do documento, título original e traduzido, datas, status de revogação, tipo de texto, domínios temáticos, palavras-chave e **abstract**), e links para o PDF e para uma versão TXT útil para NLP.

Em 13 mai 2025, o “Complete Collection” somava **213.579** registros (CSV ~318 MB), com subconjuntos temáticos (por exemplo, “Agriculture”, “Food and Nutrition”, “Land and Water”, “Environment and Ecosystems”, “Policies” e “Constitutions”).

16.2. Acesso e Documentação

O acesso pode ser feito pela interface web (busca simples/avançada) e por [downloads em CSV](#) na página de Open Data. A FAO fornece um **Guia de Busca** com operadores de campo e uma página de **Metadados** descrevendo cada coluna do CSV.

Não há API pública documentada específica do FAOLEX, para ingestão programática, a via oficial é o download dos pacotes CSV (atente às datas de atualização na página).

Documentação e termos relevantes:

- Página principal/Busca avançada (navegação e filtros), FAOLEX.
- **Open Data** (links de download, tamanhos, última atualização, citação recomendada/licença).
- **Metadados FAOLEX** (dicionário de dados/colunas).
- [Guia de Busca](#) (sintaxe/operadores).
- **Termos e Condições** e **Statistical Database Terms of Use** (licenças/reuso).

16.3. Precificação

O acesso é **gratuito**. Os datasets corporativos da FAO são disponibilizados sem custo; no caso do FAOLEX Open Data, a página de download apresenta citação recomendada com **licença CC BY-NC-SA 3.0 IGO** (restrição de uso comercial), enquanto os termos gerais dos bancos estatísticos FAO indicam **CC BY 4.0** “salvo especificado de outra forma” — portanto, para FAOLEX prevalece o aviso da própria página de Open Data.

16.4. Avaliação de Volume e Qualidade

16.4.1. Volume

Coleção grande e crescente: **>213 mil** registros no dump de 13 mai 2025; média de **~8.000** novas entradas/ano; cobertura de **200+** países/territórios e documentos em **40+** línguas.

16.4.2. Qualidade

Curadoria do **FAO Legal Office**, com base em gazetas oficiais enviadas pelos Estados-Membros (art. XI da Constituição da FAO), resumos e indexação trilingue (EN/FR/ES), e links ao texto oficial. Há variação natural de qualidade/legibilidade entre países e épocas; a presença de TXT facilita processamento automatizado.

16.5. Considerações para Coleta de Dados e Inteligência

Para pipelines de IA/BI, trate o FAOLEX como **dados normativos** (não numéricos): modelem o esquema com a chave `record_id`, normalizem

jurisdição (ISO-3166/ONU), tipo de texto (constituição/política/lei/regulamento/acordo), domínios e palavras-chave.

Priorize a ingestão via CSV do Open Data e o versionamento por “última atualização” do arquivo, fazendo **ingestão incremental** por comparação de IDs.

Use os **abstracts** e o **TXT** para NLP (classificação temática, NER de setores/commodities/agências, extração de obrigações e prazos).

Relacione documentos a eventos regulatórios e a dados operacionais (ex.: licenças, pesticidas, pesca, florestas) para análises de risco/conformidade setorial.

16.6. Desafios e Oportunidades

16.6.1 Desafios

- Licenciamento do FAOLEX Open Data indica **NC** (não-comercial), o que limita produtos comerciais diretos sem autorização.
- Heterogeneidade linguística e de formatos originais.
- Possíveis PDFs digitalizados com OCR imperfeito.
- Diferenças de estrutura legal entre países que dificultam comparabilidade.
- Atualização contínua exige monitoramento e reprocessamento.

16.6.2 Oportunidades

- Base global para **monitoramento regulatório** agroambiental.
- Insumo para **mapas de risco** (ex.: agrotóxicos, pesca IUU, desmatamento, segurança de alimentos).
- Enriquecimento de bases agronômicas com **contexto jurídico**.
- Uso do TXT para **LLMs/IE** (resumos executivos, extração de requisitos, timelines de alterações).

16.7. Recomendações para Integração

- **Ingestão:** baixe o “Complete Collection” e/ou subconjuntos temáticos do Open Data; armazene brutos (data lake) e normalize para um modelo

relacional/documental com índices por `record_id`, país, tipo de texto, domínios e datas (original/amenda).

- **Atualizações:** automatize checagem periódica da página de Open Data (comparando “Last updated” e contagem de registros) e aplique *upserts* por `record_id`.
 - **NLP/Busca:** gere embeddings do **abstract** e, quando possível, do **TXT** do documento (campo Text URL) para busca semântica e classificação; guarde o link ao **PDF oficial** para auditoria.
 - **Governança & Licenças:** registre a origem e a licença no nível do dataset (FAOLEX = **CC BY-NC-SA 3.0 IGO** na página de Open Data) e avalie necessidades de autorização para usos comerciais; quando o caso for estatístico FAO geral, verifique se aplica **CC BY 4.0**.
 - **Enriquecimento:** cruze com listas nacionais (ex.: ministérios/agências), códigos ISO, temas SDG e *taxonomies* internas para relatórios comparáveis por país/tema; explore as **Thematic/Associated Databases** (p.ex., **Family Farming Lex**) quando seu caso exigir recortes específicos.
-

17. OpenWeatherMap

[OpenWeatherMap](#) é uma plataforma privada de dados meteorológicos que fornece condições atuais, previsões, alertas e históricos globais via APIs REST, com produtos especializados (incluindo Agro API) e pacotes de histórico em alta resolução temporal.

17.1. Dados e Conteúdo

A base “Weather API” cobre tempo atual, previsões minuto-a-minuto (1 h), horárias (até 48 h), diárias (até 8 dias), alertas governamentais e histórico; pela One Call API 3.0 há acesso consolidado a atual, previsões, histórico de 46+ anos e previsões agregadas de longo prazo (até ~1,5 ano). Respostas em JSON (e, em alguns endpoints, XML/HTML).

Para análises em mapas, há **Weather Maps 2.0** (camadas de temperatura, precipitação, vento, pressão etc.) com dados atuais, previsão e histórico, disponibilizados como tiles para Leaflet/OpenLayers/Google Maps.

Para uso agrônomo, a **Agro API** permite cadastrar polígonos de talhões e obter imagens de satélite (verdadeira/falsa cor) e índices de vegetação **NDVI** e **EVI**, além de séries históricas desses índices para cada polígono.

Para backfill e modelagem, o **History API / History Bulk** disponibiliza observações horárias de ~15 variáveis desde 1º jan 1979, entregues em JSON/CSV (download em lote via Marketplace).

Há também **Geocoding API** (direto e reverso) para resolver nomes/CEPs em coordenadas, útil para padronizar entradas de locais.

17.2. Acesso e Documentação

A integração é via [APIs REST](#) com **API key** (APPID). O fluxo típico: criar conta, obter a chave e chamar endpoints HTTPS com lat/lon (ou usar o Geocoding API para converter nomes em coordenadas).

17.3. Precificação

A **One Call API 3.0** opera em modelo “[pay as you call](#)”: inclui **1.000 chamadas/dia grátis**; excedentes são cobrados (a OpenWeather divulga a tarifa como **£0,0012 por chamada**).

Planos mensais maiores (Developer/Professional/Expert/Enterprise) definem limites de **chamadas por minuto/mês** e acesso a coleções/produtos adicionais (p.ex., mapas avançados).

Produtos históricos **em lote** (History Bulk) e mapas avançados podem ter **precificação específica** via Marketplace/“Get a quote”.

17.4. Avaliação de Volume e Qualidade

17.4.1. Volume

Cobertura global com amplo portfólio (tempo atual, previsões diversas, mapas, histórico desde 1979 com passo horário), além de camadas especializadas (p.ex., precipitação histórica em passo de 10 minutos em mapas de alta resolução).

17.4.2. Qualidade

Os dados são **fundidos** de múltiplas fontes: modelos globais/locais, radares, satélites e uma rede extensa de estações, combinados por modelos proprietários/ML para gerar previsões e nowcasting. Isso assegura atualização frequente e boa robustez, embora a acurácia em microclimas possa variar.

17.5. Considerações para Coleta de Dados e Inteligência

Para aplicações agronômicas, use **One Call 3.0** para unificar atual+previsão+histórico por coordenada e **Agro API** para análises por talhão (polígonos e NDVI/EVI).

Faça **backfill** com **History Bulk** para treinar modelos de rendimento/risco. Estruture **cache** e **limitação de taxa** (rate-limit) conforme o plano, padronize **unidades** (metric/imperial) e **fuso/UTC** nos timestamps, e resolva localidades via **Geocoding API**.

Para inteligência, cruze precipitação, temperatura e índices de vegetação por janela fenológica e aplique features agregadas (acumulados/degree-days).

17.6. Desafios e Oportunidades

17.6.1 Desafios

Custo variável em alto volume (pay-as-you-go), necessidade de orquestrar múltiplos produtos (One Call, Agro, History), heterogeneidade espacial/temporal vs. microclima local, e gestão de tiles/mapas se você embutir camadas cartográficas.

17.6.2 Oportunidades

Cobertura histórica longa para modelagem, NDVI/EVI por polígono para monitorar vigor e estresse, precipitação/vento/temperatura com alta cadência para alertas operacionais, e fácil integração com stacks de mapas (Leaflet/OpenLayers/Google Maps).

17.7. Recomendações para Integração

Priorize **One Call 3.0** como fonte “única” por coordenada e **habilite cache** (e.g., TTL de 10–30 min para atual/nowcast e 1–3 h para previsões horárias) para otimizar custos.

Use **Geocoding API** para entrada do usuário; armazene identificadores de talhões e suas **geometrias** e integre **Agro API** para NDVI/EVI e imagens; faça **backfill** inicial com **History Bulk** e atualize diariamente.

Normalize unidades/UTC; monitore consumo e **quotas** no painel da OpenWeather; e, para visualização, sobreponha **Weather Maps 2.0** como tiles (com chave) no seu WebGIS.

18. BuiltWith Datasets

[BuiltWith Datasets](#) é um repositório estruturado que disponibiliza, via download, metadados sobre tecnologias presentes em domínios web, cobrindo diversos setores como agricultura digital para análises estratégicas.

18.1. Dados e Conteúdo

- Oferece informações sobre quais tecnologias estão implementadas em websites, categorizadas por tipo (como IoT, rastreamento por satélite, CMS, APIs, etc).
- Contém dados de domínio, tecnologias, categorias, contagem de registros e histórico de uso.
- Exemplos incluem datasets como “Category Datasets” (agrupados por categoria tecnológica) com milhões de registros, e “Entire Internet Datasets” com cobertura de todos os sites ativos .

18.2. Acesso e Documentação

- Os datasets são acessados por **download**, geralmente em **ZIP** contendo arquivos JSON.

- É possível baixar um [dataset de exemplo](#) (ZIP) que oferece visão prática da estrutura dos dados.
- A documentação sobre os campos presentes nos arquivos JSON está disponível no site da BuiltWith, explicando cada propriedade e sua finalidade.
- Embora este pacote específico seja por download, a BuiltWith também oferece acesso [via API](#) para consultas dinâmicas e detalhadas, conforme os recursos disponíveis na seção “API Access”.

18.3. Precificação

- A BuiltWith não publica preços fixos para os datasets em seu site; é necessário **solicitar orçamento personalizado** para acesso completo.
- Planos de API são estruturados com base em volume de consultas e tipos de dados, com valores escaláveis conforme necessidade.

18.4. Avaliação de Volume e Qualidade

18.4.1. Volume

O volume de dados da BuiltWith é massivo. A empresa afirma rastrear mais de 673 milhões de websites ativos, o que representa uma cobertura praticamente completa da internet. Para o setor agrícola, o desafio não é a falta de dados, mas sim a capacidade de filtrar e identificar os domínios relevantes.

18.4.2. Qualidade

A qualidade dos dados de identificação de tecnologia é considerada alta e é a principal fortaleza da plataforma. A precisão na detecção de scripts, widgets e frameworks é elevada. No entanto, a qualidade de metadados como informações de contato (e-mail, telefone) e localização pode variar, uma vez que depende da disponibilidade pública dessas informações nos próprios websites. A atualização mensal garante que os dados sejam relativamente recentes.

18.5. Considerações para Coleta de Dados e Inteligência

É uma fonte valiosa para análises setoriais em agricultura digital: permite filtrar domínios relevantes e entender quais tecnologias estão sendo aplicadas.

Pode ser combinada com outras bases (como meteorologia, sensores de campo, mercado) para gerar inteligência competitiva e identificar tendências tecnológicas no agronegócio.

Como os dados são amplos, criar filtros por categoria ou domínio relacionados à agricultura será fundamental para extração de insights relevantes.

18.6. Desafios e Oportunidades

18.6.1 Desafios

- Volume massivo exige infraestrutura robusta para armazenamento e processamento.
- Filtragem manual ou automatizada necessária para isolar domínios realmente relacionados ao agronegócio.
- Custo pode ser elevado, dependendo do volume e profundidade dos dados necessários (via API ou downloads completos).

18.6.2 Oportunidades

- Mapeamento do ecossistema tecnológico agrícola: identificação de fornecedores, plataformas IoT, rastreamento satelital etc.
- Insights para prospecção de clientes, parceiros ou benchmarking de soluções tecnológicas.
- Potencial para análise de adoção tecnológica global ou regional com atualização periódica (mensal ou conforme frequência dos datasets).

18.7. Recomendações para Integração

- **Obtenha e estude o dataset de exemplo** (via ZIP/JSON) e a documentação de campos, para entender a estrutura e potenciais filtros.
- **Solicite orçamento personalizado** para os datasets específicos de interesse (categoria ou “Entire Internet”) e avalie planos API se desejar acesso dinâmico.

- **Desenvolva pipelines de ingestão de dados**, com filtragem por tecnologia relevante ao agronegócio e armazenamento em data warehouse ou sistema de análise.
- **Enriqueça os dados**, cruzando com outras fontes (mercado, satélite, clima, locais) para gerar dashboards e relatórios estratégicos.
- **Estabeleça um processo de atualização periódico**, alinhando com a frequência dos dados (geralmente mensal) para manter insights atualizados.