



MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E INOVAÇÃO
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

sid.inpe.br/mtc-m21d/2023/04.26.13.37-TDI

PROPOSTA DE INTEGRAÇÃO DO SISTEMA DE MONITORAMENTO E MAPEAMENTO DE CULTIVO SEN2-AGRI E A PLATAFORMA BRAZIL DATA CUBE

Yuri Domaradzki Moreira Nunes

Dissertação de Mestrado do Curso de Pós-Graduação em Computação Aplicada, orientada pela Dra. Lúbia Vinhas, aprovada em 30 de março de 2023.

URL do documento original:

<<http://urlib.net/8JMKD3MGP3W34T/48UURQH>>

INPE
São José dos Campos
2023

PUBLICADO POR:

Instituto Nacional de Pesquisas Espaciais - INPE
Coordenação de Ensino, Pesquisa e Extensão (COEPE)
Divisão de Biblioteca (DIBIB)
CEP 12.227-010
São José dos Campos - SP - Brasil
Tel.:(012) 3208-6923/7348
E-mail: pubtc@inpe.br

CONSELHO DE EDITORAÇÃO E PRESERVAÇÃO DA PRODUÇÃO INTELLECTUAL DO INPE - CEPPII (PORTARIA Nº 176/2018/SEI-INPE):

Presidente:

Dra. Marley Cavalcante de Lima Moscati - Coordenação-Geral de Ciências da Terra (CGCT)

Membros:

Dra. Ieda Del Arco Sanches - Conselho de Pós-Graduação (CPG)
Dr. Evandro Marconi Rocco - Coordenação-Geral de Engenharia, Tecnologia e Ciência Espaciais (CGCE)
Dr. Rafael Duarte Coelho dos Santos - Coordenação-Geral de Infraestrutura e Pesquisas Aplicadas (CGIP)
Simone Angélica Del Ducca Barbedo - Divisão de Biblioteca (DIBIB)

BIBLIOTECA DIGITAL:

Dr. Gerald Jean Francis Banon
Clayton Martins Pereira - Divisão de Biblioteca (DIBIB)

REVISÃO E NORMALIZAÇÃO DOCUMENTÁRIA:

Simone Angélica Del Ducca Barbedo - Divisão de Biblioteca (DIBIB)
André Luis Dias Fernandes - Divisão de Biblioteca (DIBIB)

EDITORAÇÃO ELETRÔNICA:

Ivone Martins - Divisão de Biblioteca (DIBIB)
André Luis Dias Fernandes - Divisão de Biblioteca (DIBIB)



MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E INOVAÇÃO
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

sid.inpe.br/mtc-m21d/2023/04.26.13.37-TDI

PROPOSTA DE INTEGRAÇÃO DO SISTEMA DE MONITORAMENTO E MAPEAMENTO DE CULTIVO SEN2-AGRI E A PLATAFORMA BRAZIL DATA CUBE

Yuri Domaradzki Moreira Nunes

Dissertação de Mestrado do Curso de Pós-Graduação em Computação Aplicada, orientada pela Dra. Lúbia Vinhas, aprovada em 30 de março de 2023.

URL do documento original:

<<http://urlib.net/8JMKD3MGP3W34T/48UURQH>>

INPE
São José dos Campos
2023

Dados Internacionais de Catalogação na Publicação (CIP)

Nunes, Yuri Domaradzki Moreira.

N922p Proposta de integração do sistema de monitoramento e mapeamento de cultivo Sen2-Agri e a plataforma Brazil data cube / Yuri Domaradzki Moreira Nunes. – São José dos Campos : INPE, 2023.

xxii + 74 p. ; (sid.inpe.br/mtc-m21d/2023/04.26.13.37-TDI)

Dissertação (Mestrado em Computação Aplicada) – Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2023.

Orientadora : Dra. Lúbia Vinhas.

1. Sensoriamento remoto. 2. Dados de observação da terra. 3. Brazil Data Cube. 4. Sen2-Agri. 5. Mapas para aplicações agrícolas. I.Título.

CDU 528.8:631



Esta obra foi licenciada sob uma Licença [Creative Commons Atribuição-NãoComercial 3.0 Não Adaptada](https://creativecommons.org/licenses/by-nc/3.0/).

This work is licensed under a [Creative Commons Attribution-NonCommercial 3.0 Unported License](https://creativecommons.org/licenses/by-nc/3.0/).



MINISTÉRIO DA
CIÊNCIA, TECNOLOGIA
E INOVAÇÃO



INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

DEFESA FINAL DE DISSERTAÇÃO YURI DOMARADZKI MOREIRA NUNES BANCA Nº 057/2023, REG. 531802/2020

No dia 30 de março de 2023, as 10h, no auditório da OBT, o(a) aluno(a) mencionado(a) acima defendeu seu trabalho final (apresentação oral seguida de arguição) perante uma Banca Examinadora, cujos membros estão listados abaixo. O(A) aluno(a) foi APROVADO(A) pela Banca Examinadora, por unanimidade, em cumprimento ao requisito exigido para obtenção do Título de Mestre em Computação Aplicada, com a exigência de que o trabalho final a ser publicado deverá incorporar as correções sugeridas pela Banca Examinadora, com revisão pelo(s) orientador(es).

Novo título: "Proposta de Integração do Sistema de Monitoramento e Mapeamento de Cultivo Sen2-Agri e a Plataforma Brazil Data Cube"

Membros da Banca:

Dr. Gilberto Ribeiro de Queiroz – Presidente – INPE
Dra. Lúbia Vinhas – Orientadora – INPE
Dra. Karla Donato Fook – Membro Externo – INPE



Documento assinado eletronicamente por **Lúbia Vinhas, Chefe da Divisão de Projeto Estratégico 2**, em 31/03/2023, às 09:19 (horário oficial de Brasília), com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Gilberto Ribeiro de Queiroz, Tecnologista**, em 31/03/2023, às 11:34 (horário oficial de Brasília), com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Karla donato fook (E), Usuário Externo**, em 03/04/2023, às 10:54 (horário oficial de Brasília), com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site <https://sei.mcti.gov.br/verifica.html>, informando o código verificador **10955322** e o código CRC **BD3761F3**.

“Aprender é a única coisa de que a mente nunca se cansa, nunca tem medo e nunca se arrepende”.

LEONARDO DA VINCI

*A minha mãe **Monica**, a meu irmão **Arthur**, ao meu
tio **Paulo**, a minha madrinha **Santa**, a minha avó
Sônia e a meu padrinho **Herbert***

AGRADECIMENTOS

Agradeço a Deus por estar sempre presente em minha vida.

Gostaria de agradecer à minha família por todo apoio nessa caminhada. Em especial, a minha mãe Mônica, meu irmão Arthur e meu tio Paulo por sempre estarem presentes me apoiando e encorajando a estudar. Eu sou muito grato a eles por sempre estarem ao meu lado acreditando em meu potencial.

Gostaria de agradecer à minha madrinha Santa, que sempre esteve presente durante esses anos me apoiando a seguir estudando e sempre acreditando em meu potencial. E também a sua família por todo suporte que me deram.

Agradeço à minha orientadora Dra. Lúbia Vinhas pela competência, dedicação, suporte, paciência e tempo que teve para me auxiliar no desenvolvimento deste trabalho. Muito obrigado por todos os conselhos que serviram para o meu desenvolvimento. Sou muito grato por todas as oportunidades e disponibilidade.

Gostaria de agradecer aos meus amigos, em especial ao Paulo, Hellen e Maria, por todo suporte, incentivo, momentos de lazer e conversas. Além deles, gostaria de agradecer também a Juliana Balera, que me incentivou a ingressar no programa de Pós-Graduação em Computação Aplicada e sempre esteve presente para me auxiliar durante esse anos.

Agradeço à toda equipe do *Brazil Data Cube* por contribuir direta ou indiretamente no desenvolvimento do trabalho, estando sempre disponível para esclarecer dúvidas que surgiam.

Agradeço aos meus colegas Marcos, Adeline e Leonardo por todas as conversas e ajuda na reta final do presente trabalho.

Agradeço à instituição Instituto Nacional de Pesquisas Espaciais (INPE), que disponibilizou suporte de infraestrutura durante o desenvolvimento desta dissertação.

Agradeço à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo apoio financeiro para o desenvolvimento do presente trabalho.

Agradeço a todos que contribuíram de alguma forma para o desenvolvimento deste trabalho.

RESUMO

Os avanços nas tecnologias de sensoriamento remoto possibilitaram a aquisição de grandes volumes de dados de Observação da Terra (EO) com melhor resolução espacial e taxa de revisita, aumentando o nível de detalhamento da superfície terrestre. Além disso, a adoção de políticas abertas para o compartilhamento desses dados impulsionaram o desenvolvimento de pacotes de *software*, serviços *web* e plataformas de processamento para extrair informações desses grandes volumes de dados. Com isso, aplicações são desenvolvidas com diversos temas, como, por exemplo, para desenvolver sistemas de monitoramento de atividades agrícolas. As imagens obtidas pela missão *Sentinel-2* e o pacote de software Sen2-Agri permitem a geração de mapas sobre diferentes aspectos das culturas e áreas cultivadas, que podem ser usados para a implementação de sistemas de monitoramento agrícola. Por outro lado, a plataforma *Brazil Data Cube* (BDC) fornece um ambiente de acesso, visualização e processamento de cubos de dados de imagens de vários sensores para toda a extensão do território brasileiro. Este trabalho propõe o desenvolvimento de uma arquitetura de integração do Sen2-Agri com a plataforma BDC a fim de aproveitar este grande repositório de imagens sobre o Brasil para a temática de monitoramento de safras agrícolas. Um estudo de caso foi conduzido para a região de Luís Eduardo Magalhães para validar a arquitetura de integração proposta. Os resultados mostram a viabilidade e as vantagens desta integração, através de um serviço *web* que fornece uma interface amigável para acessar as melhores recursos de ambos os sistemas.

Palavras-chave: Sensoriamento Remoto. Dados de Observação da Terra. Brazil Data Cube. Sen2-Agri. Mapas para Aplicações Agrícolas.

PROPOSAL FOR THE INTEGRATION OF THE SEN2-AGRI CROP MONITORING AND MAPPING SYSTEM TO THE BRAZIL DATA CUBE PLATFORM

ABSTRACT

Advances in remote sensing technologies have enabled the acquisition of large volumes of Earth Observation (EO) data with better spatial resolution and revisit rate, increasing the level of detail on the Earth's surface. In addition, the adoption of open policies for sharing this data has driven the development of *software* packages, *web* services and processing platforms for extracting information from these large volumes of data. With this, applications are developed in various themes, such as, for example, to develop monitoring systems for agricultural activities. The images obtained by the Sentinel-2 mission and the Sen2-Agri *software* package allow the generation of maps on different aspects of crops and cultivated areas, which can be used for the implementation of agricultural monitoring systems. On the other hand, the Brazil Data Cube (BDC) platform provides an environment for accessing, visualizing, and processing image data cubes from various sensors for the entire extension of the Brazilian territory. This work proposes the development of an integration architecture of Sen2-Agri with the BDC platform in order to take advantage of this large repository of images about Brazil in the theme of monitoring agricultural crops. A case study for the region of Luís Eduardo Magalhães was conducted to validate the proposed integration architecture. The results show the feasibility and advantages of this integration, through a *web* service that exposes a friendly interface to access the best features of both systems.

Keywords: Remote Sensing. Earth Observation Data. Brazil Data Cube. Sen2-Agri. Maps for Agricultural Applications.

LISTA DE FIGURAS

	<u>Pág.</u>
2.1 Sensoriamento Remoto Orbital.	5
2.2 Composição colorida da região de Luís Eduardo Magalhães, BA. Satélite Amazonia-1, Sensor WFI, obtida em 30/06/2020.	6
2.3 Ciclo de vida da soja.	8
2.4 Sequência de imagens Landsat de uma área agrícola.	9
2.5 Exemplo de mapa de uso e cobertura da Terra produzido através da análise de imagens:mapa de uso e cobertura do solo no Bioma Cerrado, em 2020, produzido pelo projeto TerraClass.	10
2.6 Duas imagens obtidas por 2 sensores diferentes: Amazonia-1 WFI e CBERS-04A MUX.	12
2.7 Estrutura do Sen2-Agri.	16
2.8 Fluxo do Processador L3A.	17
2.9 Parâmetros utilizados pelo NDVI Metrics.	19
2.10 Fluxo do Processador L3B.	20
2.11 Fluxo do Processador L3E.	21
2.12 Fluxo do Processador L4A.	22
2.13 Captura de tela da página inicial do sistema Sen2-Agri.	25
2.14 Fluxo de funcionamento do BDC STAC.	27
2.15 Exemplo de um arquivo JSON com fontes de dados acessíveis pelo WLTS.	28
2.16 Fluxo de funcionamento do WLTS.	29
2.17 Modelo de banco de dados do WLCSS.	30
2.18 Esquema do banco de dados do BDC.	31
3.1 Orquestrador do Sen2-Agri com indicações dos módulos disponíveis na plataforma <i>Brazil Data Cube</i>	34
3.2 Arquitetura da integração entre a plataforma BDC e o Sen2-Agri.	35
3.3 Processo de obtenção dos dados in situ utilizando o WLTS.	38
3.4 Definição de retorno das operações de visualização.	48
4.1 Fluxograma de definição da área de monitoramento.	50
4.2 Fluxograma de geração de dados de amostra.	50
4.3 Fluxograma de geração dos produtos Sen2-Agri.	51
4.4 Fluxograma de consulta dos produtos.	52
4.5 Área de monitoramento do estudo de caso.	53
4.6 Exemplo de como importar a biblioteca cliente através do <i>Jupyter Notebook</i>	54
4.7 Utilização da operação <code>get_wlts_collections</code> para obter as coleções WLTS.	56

4.8	Verificando se a coleção mapbiomas-v6 intercepta as extensões do sítio.	56
4.9	Obtendo as classes da coleção mapbiomas-v6.	57
4.10	Criação do sítio LEM_SUMMER.	60
4.11	Tabela com o total de <i>pixels</i> rotulados para cada uma das classes LULC oriundo dos dados de amostra do sítio.	61
4.12	Mapas do Estado vegetativo com os índices LAI e NDVI obtidos para o sítio LEM_SUMMER.	62
4.13	Mapas do Estado vegetativo com índice fenológico obtidos para o sítio LEM_SUMMER.	63
4.14	Máscaras Dinâmicas de Cultivo obtidas para o sítio LEM_SUMMER.	64
4.15	Mapas do Tipo de Cultivo obtidas para o sítio LEM_SUMMER.	64
4.16	Mapa de Cultivo obtido na metade da estação de cultivo para o sítio “LEM_SUMMER”.	65
4.17	Contagem das classes rotuladas pelo classificador para o mapa da primeira metade da estação do sítio “LEM_SUMMER”.	65
4.18	Mapa de Cultivo obtido ao final da estação de cultivo para o sítio “LEM_SUMMER”.	66
4.19	Contagem das classes rotuladas pelo classificador para o mapa da segunda metade da estação do sítio “LEM_SUMMER”.	66

LISTA DE TABELAS

	<u>Pág.</u>
3.1 Parâmetros da operação <i>list_collections</i>	38
3.2 Parâmetros da operação <i>describe_collection</i>	39
3.3 Parâmetros da operação <i>verify_collection_extent</i>	41
3.4 Parâmetros da operação <i>get_collection_classes</i>	42
3.5 Parâmetros da operação <i>create_site</i>	44
3.6 Parâmetros das operações de visualização.	45

LISTA DE ABREVIATURAS E SIGLAS

API	–	<i>Application Programming Interface</i>
BDC	–	<i>Brazil Data Cube</i>
CBERS	–	<i>China-Brazil Earth Resources Satellite</i>
EMBRAPA	–	Empresa Brasileira de Agropecuária
ESA	–	European Space Agency
HTTP	–	<i>Hypertext Transfer Protocol</i>
IBGE	–	Instituto Brasileiro de Geografia e Estatística
INPE	–	Instituto Nacional de Pesquisas Espaciais
JSON	–	<i>JavaScript Object Notation</i>
LAI	–	<i>Leaf Area Index</i>
MUX	–	<i>Multispectral Camera</i>
NDVI	–	<i>Normalized Difference Vegetation Index</i>
NDVI Metrics	–	<i>Normalized Difference Vegetation Index Metrics</i>
Sen2-Agri	–	<i>Sentinel-2 for Agriculture</i>
STAC	–	<i>SpatioTemporal Asset Catalogs</i>
WFI	–	<i>Wide Field Imager</i>

SUMÁRIO

	<u>Pág.</u>
1 INTRODUÇÃO	1
1.1 Motivação	3
1.2 Objetivo	3
1.3 Organização	3
2 REFERENCIAL TEÓRICO	5
2.1 Sensoriamento Remoto para mapeamento de uso e cobertura da Terra . .	5
2.2 Big Data e imagens de sensoriamento remoto	9
2.3 <i>Food Security TEP</i>	14
2.4 Sen2-Agri	15
2.4.1 Reflectância de superfície - L2A	15
2.4.2 Compostos temporais - L3A	16
2.4.3 Indicadores de estado da vegetação - L3B	18
2.4.4 Máscara dinâmica de cultivo - L4A	20
2.4.5 Mapa do tipo de cultivo - L4B	23
2.4.6 Arquitetura do sistema	23
2.4.7 Modos de uso	24
2.5 Brazil Data Cube	25
2.5.1 <i>Brazil Data Cube SpatioTemporal Asset Catalog</i>	26
2.5.2 <i>Web Land Trajectory Service</i>	27
2.5.3 <i>Web Land Classification System Service</i>	29
2.5.4 <i>Brazil Data Cube Image Metadata Catalog</i>	31
3 INTEGRAÇÃO DO SEN2-AGRI AO BRAZIL DATA CUBE .	33
3.1 Dados de amostra	35
3.2 list_collections	37
3.3 describe_collection	38
3.4 get_wlts_collections	40
3.5 verify_collection_extent	41
3.6 get_collection_classes	42
3.7 create_site	43
3.8 Operações de visualização	44
3.9 Implementação	47

4	RESULTADOS E ESTUDO DE CASO	49
4.1	Fluxograma de uso do pacote	49
4.2	Estudo de caso	52
5	CONCLUSÕES	69
	REFERÊNCIAS BIBLIOGRÁFICAS	71

1 INTRODUÇÃO

A Organização para a Alimentação e Agricultura das Nações Unidas (FAO) estimou que a população mundial em 2050 irá alcançar 9,1 bilhões de pessoas. Nesse cenário, para garantir a segurança alimentar da população será necessário um aumento de 70%, em relação a 2017, na produção de alimentos. Ao mesmo tempo, as mudanças ambientais globais ameaçam e pressionam a sustentabilidade dos sistemas agrícolas. O aumento na produção de alimentos deve ser acompanhado por um manejo sustentável de terras agrícolas para diminuir ao máximo os impactos negativos sobre a qualidade e quantidade dos recursos hídricos e do solo, a degradação da terra, as emissões de gases de efeito estufa e o impacto sobre a biodiversidade. Esse desafio está expresso na Agenda 2030 do Desenvolvimento Sustentável das Nações Unidas, no seu objetivo 2: “*Acabar com a fome, alcançar a segurança alimentar e melhoria da nutrição e promover a agricultura sustentável*” (NAÇÕES UNIDAS BRASIL, 2021).

O Brasil está entre os líderes mundiais na produção de soja, milho e suco de laranja, sendo o terceiro maior exportador de produtos agroalimentares, depois da União Europeia e dos Estados Unidos. Dois terços do valor total da produção agrícola são produtos agrícolas e um terço produtos de origem animal. O principal produto das exportações brasileiras é a soja (grão, farelo e óleo), que representa quase 50% da exportação de produtos agroalimentares. Segundo os dados da Produção Agrícola Municipal 2019 (PAM), fornecido pelo Instituto Brasileiro de Geografia e Estatística (IBGE) (INSITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE), 2021), a produção agrícola nacional de cereais, leguminosas e oleaginosas somou, em 2019, o volume de 243,3 milhões de toneladas. Considerando somente a soja, estima-se uma área plantada de quase 36 milhões de hectares.

As tecnologias digitais emergentes, como a da Internet das Coisas (IOT) e as tecnologias relacionadas à *Big Data* têm impactado a sociedade em diversas temáticas de atuação. Por exemplo, no caso da agricultura de precisão, elas podem contribuir para tornar as fazendas mais inteligentes através do monitoramento e análise dos aspectos físicos das plantações. A partir disso, as fazendas inteligentes fornecem aos agricultores dados que auxiliam na melhoria do gerenciamento de toda a estrutura, permitindo, por exemplo, tomar decisões em tempo real sobre a entrega da taxa apropriada de insumos (água, fertilizante, pesticida) (DELGADO et al., 2019). Assim, os fazendeiros podem utilizar as informações obtidas através do processamento dos dados para combater os desafios da produção agrícola em termos de produtivi-

dade e impacto ambiental com sustentabilidade para garantir a segurança alimentar da população. Segundo (KAMILARIS et al., 2017), “*Big Data é menos uma questão de volume de dados do que a capacidade de pesquisar, agregar, visualizar, e referenciar grandes conjuntos de dados em um tempo razoável. É sobre a capacidade de extrair informações e ideias onde previamente era econômica ou tecnicamente inviável*”.

Outra fonte de dados extremamente volumosa e que se encaixa dentro do conceito de *Big Data* são os sistemas de sensoriamento remoto orbital da superfície da Terra. O uso de dados de sensoriamento remoto em aplicações relacionadas à agricultura têm crescido muito graças ao progresso nas tecnologias de obtenção e tratamento desses dados. Isso inclui desde o desenvolvimento de sensores com múltiplas combinações de resoluções temporais, espectrais e espaciais, a bordo de satélites de observação da Terra ou de pequenas plataformas que carregam esses sensores, como os nano-satélites ou drones.

O processamento de grandes de dados de sensoriamento remoto requer uma infraestrutura computacional de alto desempenho em termos de armazenamento e de processamento de dados, a fim de que se possa obter dados mais precisos e em tempo hábil. Nessa dimensão, pode ser notado o aumento das capacidades de computação em nuvem e das técnicas de aprendizado de máquina para extração de informações (WEISS et al., 2020).

Essas tecnologias permitem desenvolver aplicações para monitorar culturas, gerando mapas de áreas de cultivo e sobre essas mostrar se estão saudáveis e produtivas e até mesmo estimar a produtividade de certas culturas. Esses dados formam a base para a construção de sistemas de monitoramento contínuo de áreas cultivadas, por exemplo. Algoritmos diversos são implementados e disponibilizados em pacotes e sistemas de *software* para essa área de aplicação específica, que neste trabalho é chamada de aplicações agrícolas.

Essa dissertação visa a contribuição no tema do uso de imagens de sensoriamento remoto em aplicações agrícolas, com foco nas técnicas de computação que permitem o processamento de dados junto ao seu armazenamento, através da proposição e prototipação de uma arquitetura de integração de sistemas.

1.1 Motivação

A pesquisa sobre a utilização de imagens de sensoriamento remoto para mapeamento de tipo de culturas e de práticas agrícolas é extensa e tende a crescer em função do aumento e variedade dos sensores produzindo imagens em diferentes resoluções espaciais e temporais. Isto motiva a integração de pacotes específicos para geração de produtos para aplicações agrícolas baseados em imagens de satélites às plataformas de organização e processamento de grandes volumes de imagens de sensoriamento remoto.

1.2 Objetivo

Este trabalho tem o objetivo de desenvolver uma arquitetura de integração entre a plataforma do projeto *Brazil Data Cube* (BDC) com o pacote de *software* Sen2-Agri com o intuito de gerar e visualizar mapas para aplicações agrícolas. Além disso, podem ser destacados os seguintes objetivos específicos:

- Propor e desenvolver uma arquitetura de integração do Sen2-Agri com a plataforma BDC através do uso de serviços web;
- Definir as interfaces mínimas do serviço web de integração de sistemas;
- Exercitar o uso do pacote de integração em ambientes de computação interativa;
- Utilizar tecnologias do projeto Brazil Data Cube para padronizar o acesso e armazenamento dos produtos em uma base de dados;
- Implementar um protótipo e um estudo de caso.

1.3 Organização

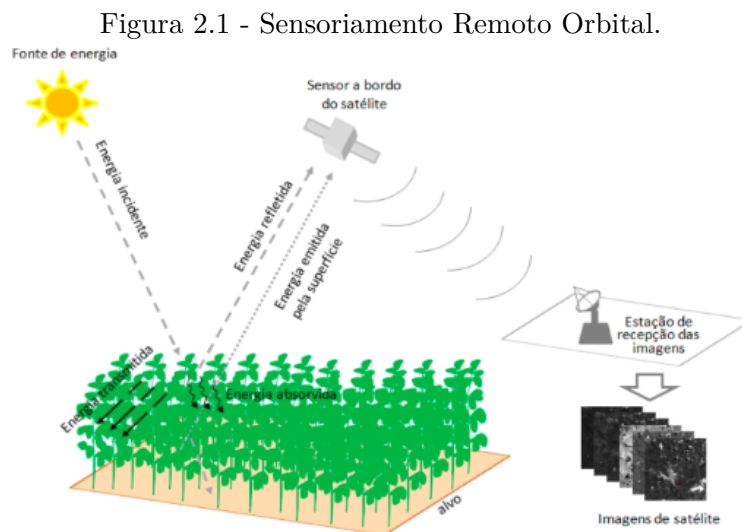
No Capítulo 2 são apresentados os conceitos e definições utilizados como base para o desenvolvimento do trabalho. No Capítulo 3, é apresentada a metodologia de desenvolvimento do trabalho e da arquitetura de integração proposta. Neste capítulo são abordadas as tecnologias utilizadas para o desenvolvimento da integração proposta e como o serviço foi construído. No Capítulo 4 são evidenciados os resultados obtidos por meio de estudo de caso. Por fim, no Capítulo 5 são apresentadas as conclusões do trabalho.

2 REFERENCIAL TEÓRICO

Este capítulo aborda um resumo de discussões realizadas por autores em relação ao tema do trabalho, desempenhado com o intuito de adquirir embasamento e conhecimento acerca destes assuntos.

2.1 Sensoriamento Remoto para mapeamento de uso e cobertura da Terra

Sensoriamento Remoto se refere à disciplina e ao processo de detecção e monitoramento das características físicas de uma área, medindo a radiação refletida e emitida à distância, normalmente por sensores embarcados em satélites ou aeronaves. Neste trabalho trata somente de imagens da superfície, obtidas por sensoriamento remoto orbital, ou seja, a bordo de satélites. E também somente daqueles com sensores óticos, ou seja, que dependem da radiação solar incidente refletida pelos alvos na superfície da Terra (CHUVIECO, 2016). A Figura 2.1 representa o processo do sensoriamento remoto orbital.



Fonte: Formaggio e Sanches (2017).

As imagens de sensoriamento remoto são caracterizadas por quatro parâmetros chamados de resolução, que dependem das características dos sensores. A resolução espacial indica o tamanho da menor área do terreno que gera uma medida do sensor, e reflete o grau de detalhamento das feições na imagem (quanto maior a resolução menor, maior é o detalhamento); a resolução temporal indica o tempo decorrido

entre a obtenção de duas imagens do mesmo ponto; a resolução espectral se refere a capacidade do sensor de distinguir diferentes intervalos do espectro eletromagnético, as bandas; finalmente, a resolução radiométrica, se refere ao número de *bits* usados para codificar os valores na imagem. A primeira análise feita sobre as imagens é a visual, ou seja, através da associação de bandas aos canais de cores de um dispositivo de apresentação, podem ser identificadas as feições do terreno. Como exemplo, pode-se observar a Figura 2.2, que mostra uma composição colorida das bandas do sensor de Campo de Imageamento Largo (WFI - do inglês *Wild Field Imager*), a bordo do satélite Amazônia-1, desenvolvido e operado pelo Instituto Nacional de Pesquisas Espaciais (INPE). A imagem é da região do município de Luís Eduardo Magalhães, na Bahia, uma área muito usada para atividades de agricultura. Na imagem podem ser visualmente identificados talhões e pivôs de irrigação de culturas.

Figura 2.2 - Composição colorida da região de Luís Eduardo Magalhães, BA. Satélite Amazônia-1, Sensor WFI, obtida em 30/06/2020.



Fonte: Instituto Nacional de Pesquisas Espaciais (2023b).

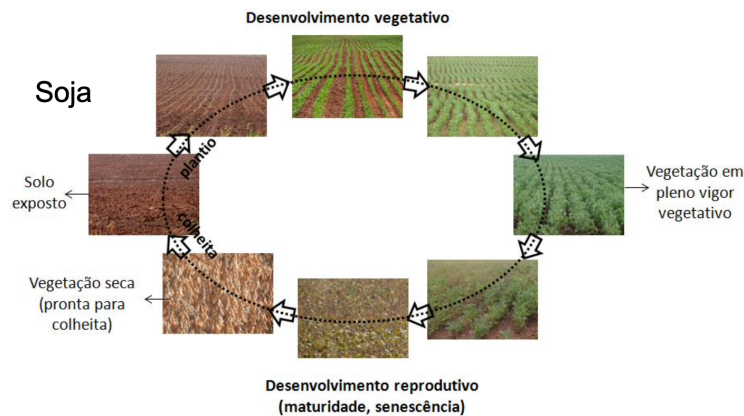
As imagens são insumos para aplicações na área ambiental, como a quantificação de parâmetros biofísicos e a execução de mapeamentos do uso e cobertura da Terra em diferentes escalas. A cobertura da Terra se refere à identificação do tipo de cobertura, natural ou artificial, em uma unidade geográfica, é o que se pode observar nas imagens de sensoriamento remoto. O uso da Terra, refere-se aos arranjos, atividades e na forma como a área é ocupada. Os dois termos são relacionados e usados concomitantemente na definição de temas, ou classes, atribuídos a segmentos homogêneos do terreno, que resultam em mapas de uso e cobertura. Em (CHAVES et al., 2020) é possível encontrar uma revisão sistemática da literatura recente sobre o uso de imagens de média resolução espacial, tipicamente entre 10 e 30 metros para o mapeamento de uso e cobertura da Terra.

Dentre as possíveis classes de uso e cobertura destacam-se aquelas relacionadas às práticas agrícolas. As plantas, assim como os outros alvos na superfície da Terra, interagem com a luz do sol, ou seja, com todo o espectro da radiação eletromagnética emitida pelo sol, de maneira diferente, dependendo do comprimento de onda observado. A radiação eletromagnética que é refletida pelas plantas contém informações sobre sua composição biofísica e seu estado fisiológico. Com o aumento do número e da variedade desses sensores, as propriedades espectrais e temporais da superfície da Terra, resultantes de práticas agrícolas, podem ser capturadas e monitoradas em diferentes escalas espaciais e temporais. Analisando imagens obtidas em datas distintas, busca-se identificar os diferentes estágios do ciclo de vida da cultura plantada em uma determinada área. A Figura 2.3 ilustra esse ciclo para uma cultura de soja, enquanto a Figura 2.4 mostra um conjunto de imagens do mesmo sensor em diferentes datas de um talhão de cultura de diferentes plantas (FORMAGGIO; SANCHES, 2017).

Um exemplo de dado de uso e cobertura da Terra pode ser visto na Figura 2.5. Esse projeto desenvolvido em parceria pelo INPE e a Empresa Brasileira de Agropecuária (EMBRAPA) tem por objetivo gerar mapas de uso e cobertura da Terra para os biomas Amazônia e Cerrado, a partir da análise de imagens de satélites. Nesse projeto, para o Cerrado, são mapeadas, entre outras classes, as áreas de cultura do tipo perene, semi perene, temporária de um ciclo e de mais de um ciclo, confirmando a importância dessa categoria de classes nos dados de uso e cobertura da Terra.

As imagens de satélites, com a resolução apropriada, podem ser usadas para identificar o sistema de cultivo usado, que refere ao tipo de colheita, sequência e arranjo do plantio, e às técnicas de manejo usadas em uma determinada área de cultura ao

Figura 2.3 - Ciclo de vida da soja.



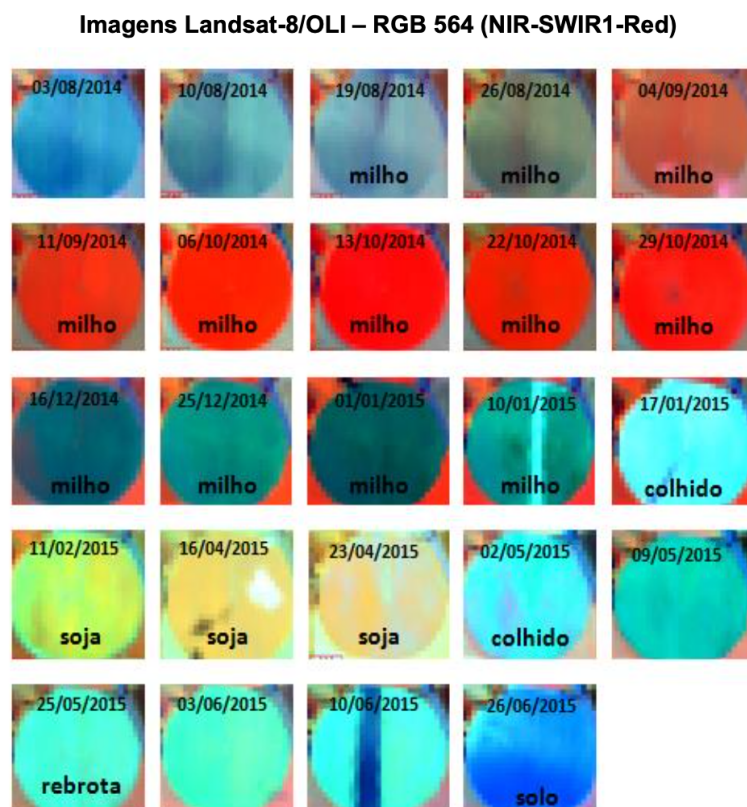
Fonte: Formaggio e Sanches (2017).

longo de um período de anos, como no caso das classes do TerraClass não indicam qual é a cultura específica mas sim as áreas de cultivo e os tipos de manejo. Uma revisão de trabalhos que usam sensoriamento remoto para identificar o sistema de cultivo pode ser encontrada em (BÉGUÉ et al., 2018).

Para se acompanhar atividades relacionadas à prática da agricultura não é suficiente a análise de uma única imagem, é de fundamental importância capturar a dinâmica do processo através de sequências de imagens no tempo (Figura 2.4), ou séries temporais de imagens. Gómez e colaboradores (GÓMEZ et al., 2016) apresentaram uma revisão de trabalhos que utilizam séries temporais de imagens óticas para gerar mapas anuais de agricultura.

Para transformar as imagens de sensoriamento remoto na produção de dados em escala regional e nacional, em modo operacional, ainda existem desenvolvimentos a serem feitos, principalmente no Brasil, com sua grande extensão territorial, diversidades regionais em aspectos físicos (clima, solo, relevo, cobertura vegetal e disponibilidade hídrica) e econômicos (fontes de financiamento e nível de investimento em insumos e tecnologia). Para solucionar esses desafios, o uso de plataformas de computação apropriadas para gerenciar grandes volumes de dados e facilitar o desenvolvimento de aplicações torna-se imprescindível. A grande disponibilidade de imagens de sensoriamento remoto e a melhora nas capacidade de computação e armazenamento de dados, têm permitido organizar grandes coleções de imagens, que formam composições densas no tempo e no espaço, com pouca ou nenhuma lacuna ou falta. Esses conjuntos de dados permitem que grandes extensões da superfície da Terra possam ser analisadas, para o mapeamento de temático do uso e da cobertura

Figura 2.4 - Sequência de imagens Landsat de uma área agrícola.



Fonte: Formaggio e Sanches (2017).

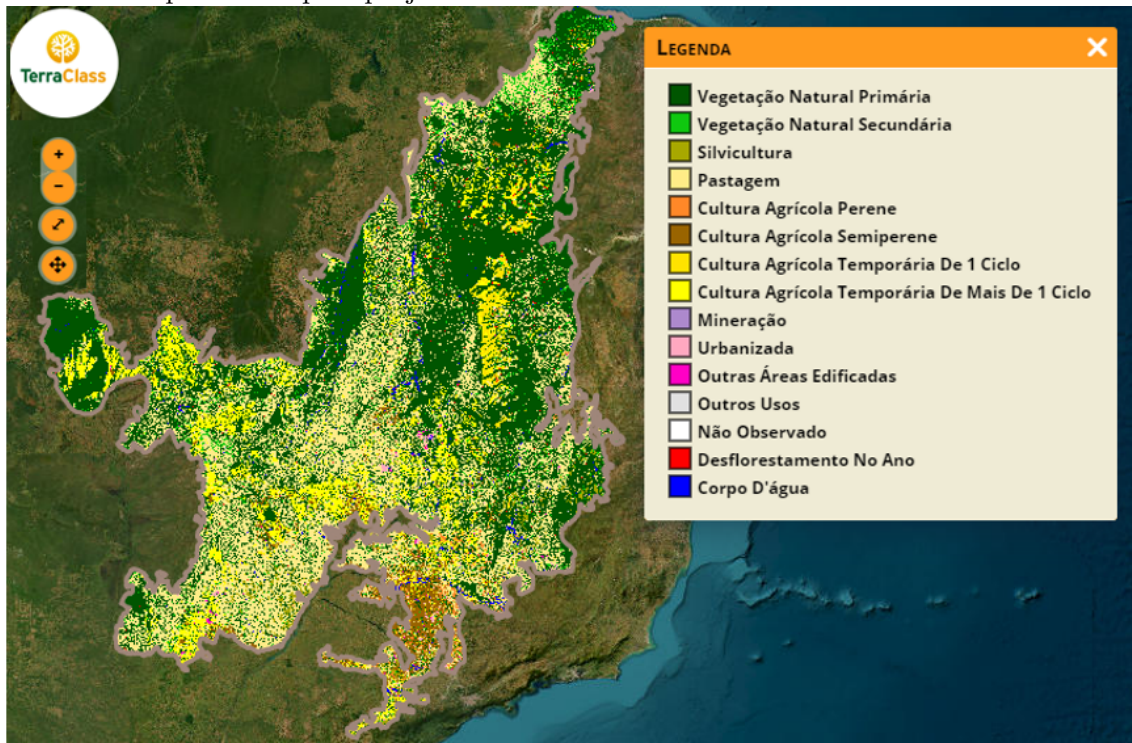
da Terra em escalas nacionais ou até mesmo globais. Nesse contexto, introduz-se o termo “*big data*”, que tem sido usado para tratar dados com características de grande variedade, volume, velocidade de aquisição e valor, no contexto particular de grandes volumes de imagens de sensoriamento remoto.

2.2 Big Data e imagens de sensoriamento remoto

Segundo (GEOSPATIAL..., 2016), “*Big Data ainda não é um termo claramente definido, tendo conceitos diferentes estabelecidos pelas perspectivas tecnológicas, industriais, de pesquisa ou acadêmicas. Em geral, é considerada como conjuntos de dados estruturados ou não que comportam grandes volumes de dados que não podem ser facilmente capturados, armazenados, manipulados, analisados, gerenciados e apresentados pelas tecnologias tradicionais de software, hardware e banco de dados*”. Uma fonte de Big Data são os dados geoespaciais.

Os dados geoespaciais possuem informações qualitativas e/ou quantitativas junto com posicionamento explícito, detalhes de localização e atributos adicionais que

Figura 2.5 - Exemplo de mapa de uso e cobertura da Terra produzido através da análise de imagens: mapa de uso e cobertura do solo no Bioma Cerrado, em 2020, produzido pelo projeto TerraClass.



Fonte: TerraClass (2023).

descrevem recursos específicos encontrados em um conjunto de dados. Estes dados são digitalmente representados em estruturas vetoriais ou matriciais. Os dados vetoriais são representados a partir de primitivas geométricas simples, pontos, linhas, polígonos e suas combinações. Dados com representação matricial são caracterizados como uma matriz de células de tamanho igual organizadas em linhas e colunas. A cada célula está associado um valor que pode ser multidimensional (WORBOYS; DUCKHAM, 2004).

O sensoriamento remoto orbital produz, possivelmente, o maior volume de dados geoespaciais com representação matricial. A geração atual de sensores espaciais está gerando fluxos quase contínuos de produção de dados de Observação da Terra (EO), fornecendo imagens com diferentes resoluções espaciais, espectrais e temporais. Esses parâmetros, que caracterizam esse tipo de dado, são:

- resolução espacial: representa a menor área no terreno correspondente a um valor da matriz. É tipicamente dada em metros;

- resolução temporal: representa a taxa de revisita a um mesmo ponto no terreno (ex. 16 dias, 20 dias);
- resolução espectral: está associada as faixas do espectro eletromagnético que são medidas pelo sensor. São caracterizadas por um intervalo dado de valores em μm (micrómetro, onde $1\mu m = 10^{-6}m$). Cada faixa é também chamada de *banda* e podem ser nomeadas característicos do intervalo ou seu uso (ex. a banda 05 do sensor MUX a bordo do CBERS-04 também é chamada de azul, por corresponder ao intervalo do espectro eletromagnético correspondente a percepção dessa cor).

É importante notar que existe um compromisso entre a resolução espacial, a largura da faixa imageada e a resolução temporal ou taxa de revisita do sensor ao mesmo ponto. A Figura 2.6 exemplifica duas imagens com diferentes características de resolução espacial e largura de faixa imageada. A câmera MUX, do CBERS 4A, possui uma largura de faixa de 95km, resolução espacial de 16m, taxa de revista de 5 dias e 4 bandas espectrais na faixa do visível: azul, verde, vermelho e infravermelho próximo (INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS, 2023d). A câmera WFI, do Amazonia-1, possui uma largura de faixa de 850km, resolução espacial de 64m e também três bandas do visível (azul, verde e vermelho) e uma banda no infravermelho próximo (INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS, 2023c). Nota-se que, quão maior for a largura da banda, menor é a resolução espacial e vice-versa. Isto impacta diretamente na taxa de revisita do satélite. Caso a resolução seja baixa e a largura de banda maior, a taxa de revisita será menor, uma vez que a câmera cobre grandes áreas da superfície terrestre com um nível menor de detalhamento. Agora, se a resolução for maior e a largura menor, a taxa de revisita será maior.

Dentre os produtores de imagens de sensoriamento remoto que disponibilizam seus dados de maneira aberta, destacam-se o INPE, o Serviço Geológico dos Estados Unidos (USGS) e a Agência Espacial Européia (ESA). O INPE foi responsável pela construção, lançamento e operação dos satélites China-Brazil Earth Resource Satellite (CBERS-1, CBERS-2, CBERS-3, CBERS-4, e CBERS-4A) feitos em cooperação com a China, e também do satélite Amazonia-1, construído somente pelo Brasil. O CBERS-4, o CBERS-4A e o Amazonia-1 estão operacionais e carregam sensores ópticos. O INPE destaca-se por ter sido a primeira agência produtora de imagens de satélites a adotar uma política de acesso livre.

O USGS é responsável pelo programa *Landsat*, que produz imagens óticas de média

Figura 2.6 - Duas imagens obtidas por 2 sensores diferentes: Amazonia-1 WFI e CBERS-04A MUX.



Fonte: Instituto Nacional de Pesquisas Espaciais (2023a).

resolução espacial a 50 anos. O acesso livre ao acervo *Landsat* foi implementado em 2008, gerando um aumento substancial no uso desses dados e no desenvolvimento de novas aplicações. Mais recentemente, a USGS começou a processar seu acervo pretérito, gerando novas coleções que permitem a implementação de análises baseadas em séries temporais de imagens ao invés de imagens únicas (WULDER et al., 2022). O volume estimado de dados Landsat publicados diariamente é estimado em 1,5 TB.

Como neste trabalho é de especial interesse as imagens *Sentinel-2*, a seguir são descritas com mais detalhes as características deste sensor. Estima-se que a Agência Espacial Européia (ESA, do inglês *European Space Agency*) publica diariamente 4TB de dados Sentinel-2.

A missão *Sentinel-2* conta com dois satélites chamados (A e B) gera imagens utilizando um sensor óptico, multiespectral, com foco na coleta de dados sobre o ambiente terrestre, incluindo vegetação, solos e umidade, rios e áreas costeiras. Ao todo, o sensor *Multispectral Imager* (MSI) a bordo do Sentinel-2 gera 13 bandas espectrais, sendo quatro na faixa do visível e infravermelho; seis pertencem ao “*reg edge*” e infravermelho de ondas curtas; e três são para correções atmosféricas. As resoluções espaciais das imagens podem variar de 10, 20 e 60 metros a depender da banda e a

capacidade de revisita do sensor é de aproximadamente 5 dias.

As bandas de resolução espacial de 10m são as bandas da faixa do visível e do infravermelho, numeradas como B02, B03, B04 e B08, nomeadas, respectivamente, como Azul, Verde, Vermelho e Infravermelho próximo. As bandas de resolução espacial de 20m são as bandas no espectro do “*red edge*” e infravermelho próximo, numeradas como B05, B06, B07, B08A, B11 e B12, nomeadas, respectivamente, como *Red Edge 1*, *Red Edge 2*, *Red Edge 3*, *Red Edge 4*, SWIR 1 e SWIR 2. As bandas de resolução espacial de 60m são as utilizadas para correções atmosféricas, numeradas como B01, B09 e B10, nomeadas, respectivamente, como *Aerossol*, *Water Vapor* e *Cirrus*.

Além das bandas originais, combinação entre bandas podem ser usadas para realçar feições e características extraídas das imagens, como características geológicas, relacionadas a práticas agrícolas ou a vegetação (GIS GEOGRAPHY, 2022). A seguir são apresentados algumas combinações de bandas típicas das imagens *Sentinel-2*:

- a combinação B04/B03/B02 no canal RGB gera uma imagem em cor verdadeira (natural), ou seja, representa uma imagem similar a percebida pelo olhar humano;
- B08/B04/B03 no canal RGB gera uma imagem composição de falsa cor infravermelha, comumente utilizada para distinguir a vegetação sadia, uma vez que a banda 8 é muito refletida pela clorofila presente nas plantas, de maneira que as áreas de muita vegetação verde são destacadas em vermelho;
- B11/B8/B2 no canal RGB é uma combinação típica para aplicações agrícolas, pois pode ser usada para monitorar a saúde das culturas;
- Já o índice dado pela razão $(B8-B4)/(B8+B4)$ produz indicador de quantidade de vegetação verde com valores altos sugerindo áreas de dossel denso e valores baixos ou negativos indicam feições urbanas e aquáticas;

A partir da disponibilização destes grandes volumes de dados de observação da terra, plataformas e pacotes de *software* estão sendo disponibilizados para analisar e visualizar informações relevantes sobre estes dados. Gomes e colaboradores (GOMES et al., 2020) apresentam uma visão geral de sete plataformas para gerenciamento e análise de grandes dados de observação e uma comparação dessas plataformas de acordo com critérios que representam capacidades de interesse da comunidade.

As plataformas analisadas podem ser consideradas genéricas, uma vez que podem atender ao desenvolvimento de aplicações em diversos domínios. Nesse trabalho será de interesse a plataforma Brazil Data Cube, descrita mais adiante.

2.3 *Food Security TEP*

Em um contexto mais específico, o *Food Security TEP* é uma das sete plataformas de exploração de dados de observação da Terra suportadas pelo programa Copernicus. Uma plataforma de exploração é pensada como um espaço de trabalho virtual, fornecendo a uma comunidade de usuários acesso a (i) grande volume de dados, (ii) desenvolvimento de algoritmos e ambiente de integração, (iii) processamento de software e serviços (por exemplo, bibliotecas de algoritmos de processamento, rotinas de visualização), (iv) recursos de computação, (v) ferramentas de colaboração (por exemplo, fóruns, wiki, etc.). Quando uma plataforma é dedicada a um Tema específico, torna-se uma Plataforma de Exploração Temática (TEP) (ROMEO et al., 2017).

O Food Security TEP é uma plataforma para a extração de informações de dados de Observação da Terra para produzir dados e serviços no setor da segurança alimentar principalmente da Europa e África. Desse modo, visa promover aplicações agrícolas e aquícolas inteligentes e com uso intensivo de dados no domínio científico, privado e público (MUERTH et al., 2020).

Esta plataforma foi estruturada como um serviço *web* que faz a utilização da computação em nuvem para processar os dados de sensoriamento da Terra acessando diretamente o repositório de dados do projeto *Copernicus*, da Agência Espacial Europeia, e a outras fontes de dados fornecidos por setores públicos e privados. Assim, a plataforma é alimentada por um grande conjunto de dados associado aos temas de agricultura e aquíicultura. A plataforma permite ao usuário encontrar e manipular dados oriundos de várias fontes distintas sobre a área a ser monitorada. Além disso, o *Food Security TEP* permite adicionar novos algoritmos à sua plataforma base com o intuito de adequar toda a etapa de processamento dos dados às características particulares da região de estudo que o usuário configurou.

Isso possibilitou o desenvolvimento de aplicativos configurados para as necessidades específicas de algumas localizações da África, processo este realizado em conjunto com os membros do *Food Security TEP* e usuários africanos (que possuem maior domínio sobre a área a ser monitorada). Assim, foram desenvolvidas aplicações para mapear derrubadas de manguezais e produção de algas na costa marinha em Zan-

zibar; melhorar a segurança alimentar para os agricultores no Quênia; e melhorar o acesso a créditos para pequenos agricultores no Quênia através do monitoramento constante dos campos agrícolas.

Portanto, o desenvolvimento do *Food Security TEP* facilitou o acesso e o compartilhamento de grandes volumes de dados de sensoriamento da Terra para monitoramento agrícola e aquícola, já que oferece uma plataforma em nuvem que faz toda a manipulação dos dados e fornece ao usuário a informação desejada, dispensando que o utilizador do sistema tenha que ter uma super infraestrutura local para processar os dados e obter as informações necessárias para o monitoramento da área desejada.

2.4 Sen2-Agri

Em um contexto ainda mais específico, a ESA apoia o projeto *Sentinel-2 for Agriculture* (Sen2-Agri), que é um sistema de código-livre baseado em métodos genéricos de análises de séries temporais para monitoramento e mapeamento de cultivo. As metodologias e processos implementados no Sen2-Agri foram pensados para levar em conta a disponibilidade global de séries temporais de imagens de observação da Terra obtidas pelos satélites *Sentinel-2* e *Landsat-8*. Ou seja, podem ser aplicadas em qualquer localização do planeta (DEFOURNY et al., 2019).

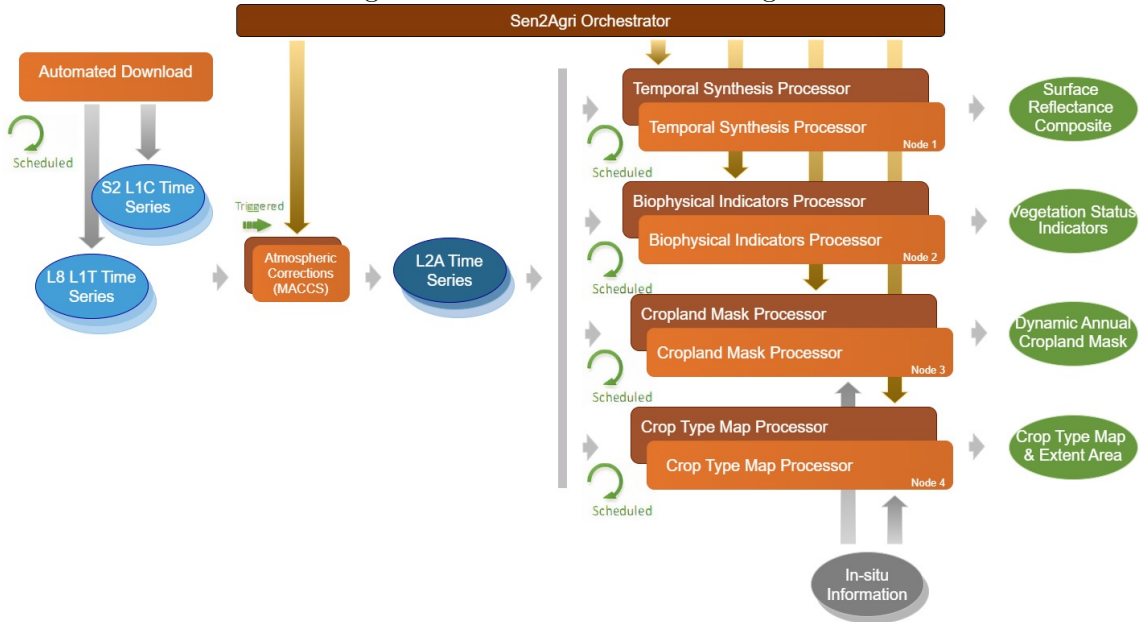
Segundo (BONTEMPS et al., 2015), “o objetivo do projeto Sen2-Agri é fornecer à comunidade internacional algoritmos validados, códigos-fonte abertos e melhores práticas para processar dados Sentinel-2 de forma operacional em produtos agrícolas relevantes para os principais sistemas agrícolas representativos mundiais”.

O Sen2-Agri é estruturado em cinco módulos, ou processadores, para gerar diferentes produtos. Esses módulos podem ser chamados por um orquestrador, visando implementar um sistema de monitoramento contínuo (Figura 2.7). Os módulos do Sen2-Agri são o L2A, que é responsável por gerar produtos de correção atmosférica; o L3A, que obtém os produtos de composição mensal livre de nuvens; o L3B, que gera os mapas de estado vegetativo com os índices NDVI, LAI e NDVI Metrics; o L4A, que gera a máscara dinâmica de cultivo; e o L4B, que obtém os mapas de tipos de cultivos. Esses módulos estão descritos com mais detalhes abaixo.

2.4.1 Reflectância de superfície - L2A

O processador L2A é responsável por executar correções atmosféricas para transformar o valor das imagens em reflectância de superfície. Esses processos utilizam parâmetros do satélite, bem como as condições atmosféricas no momento da aquisi-

Figura 2.7 - Estrutura do Sen2-Agri.



Fonte: Defourny et al. (2019).

ção visando ter uma medida de reflectância mais precisa da superfície. Além disso, produz máscaras de pixels com nuvens, sombras de nuvem ou gelo.

2.4.2 Compostos temporais - L3A

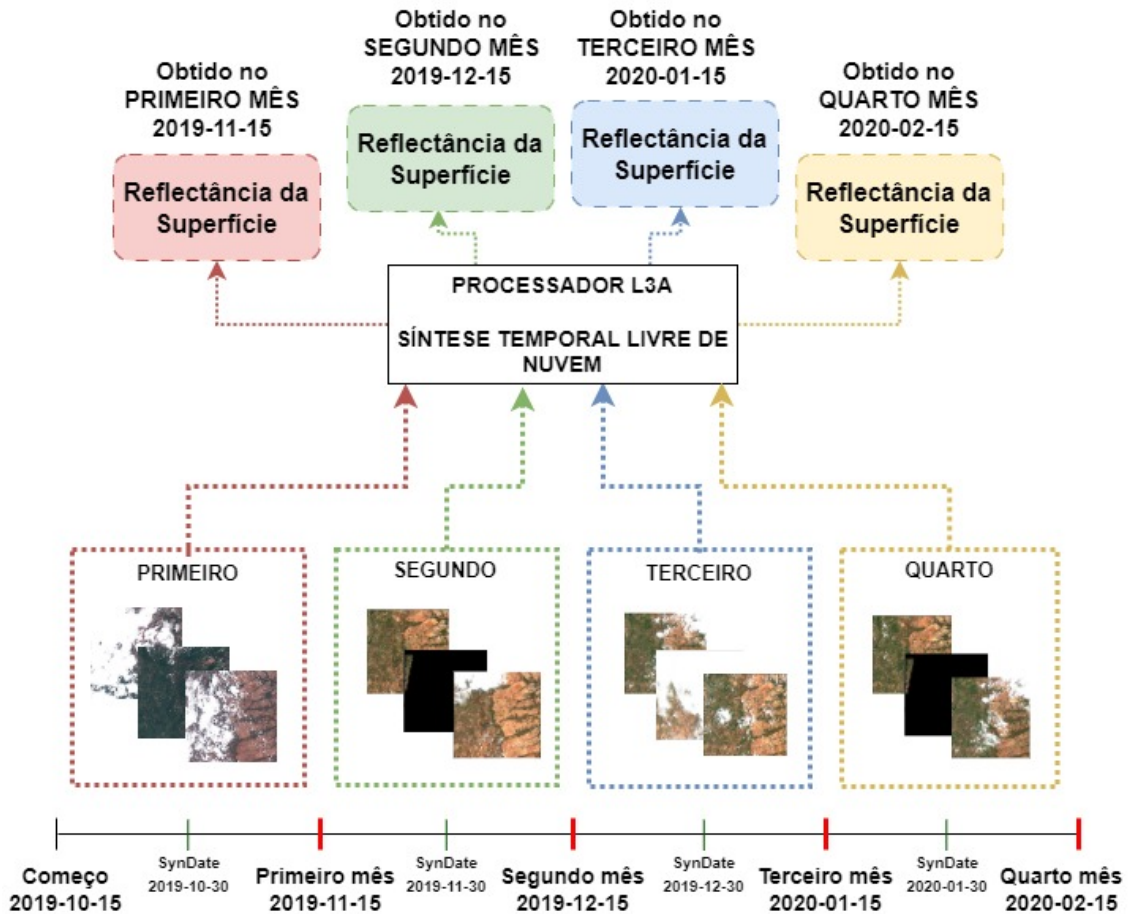
O processador L3A fornece sínteses temporais mensais de dados livres de nuvens, a partir das imagens de reflectância de superfície, para todas as 10 bandas *Sentinel-2*. Elas são geradas usando um algoritmo de média ponderada, que consiste em calcular a média dos valores de reflectância da superfície livre de nuvens em um determinado período. O método baseia-se na repetitividade das observações para reduzir estatisticamente os erros que podem ocorrer devido a nuvens ou sombras de nuvens não detectadas ou erros de correção atmosférica (DEFOURNY et al., 2019).

As composições mensais podem ser geradas para períodos de 30 a 50 dias, o que permite uma maior flexibilidade em relação à cobertura de nuvens do local. Assim, um conjunto de amostras é adquirido, a partir das imagens deste intervalo temporal, para analisar e gerar produtos com números menores de *pixels* de nuvens. Após definir a quantidade de dias do mês, o sistema calcula e configura a data que marca a metade do intervalo temporal definido para o sítio, denominada de *synthesis date*, e separa as imagens para processamento, obtendo as composições livre de nuvens para cada mês.

Os produtos podem ser gerados com resoluções espaciais de 10 e 20m. Para a resolução espacial de 10 metros, as bandas *Blue* (B2), *green* (B3), *red* (B4) e *near infrared* (B8) são utilizadas, e, para os produtos com resolução de 20 metros, são utilizadas as bandas *Vegetation red-edge* (B5, B6, B7, B8a) e *shortwave infrared* (B11, B12).

Para exemplificar o funcionamento do processador L3A, foi definida uma estação de cultivo com quatro meses de duração, com início em 15/10/2019 e término em 15/02/2020, e que possui uma janela temporal mensal de 30 dias. Assim, o *synDate* é definido com a data 15 dias após o começo do mês. Cada mês possui seu *synDate*. Por fim, o processador separa as imagens de cada mês e obtém os compostos. A figura 2.8 ilustra esse exemplo.

Figura 2.8 - Fluxo do Processador L3A.



Fonte: Produção do autor.

2.4.3 Indicadores de estado da vegetação - L3B

Este produto consiste em um conjunto de mapas de indicadores descrevendo a evolução da vegetação verde. Três tipos de indicadores são computados: o conhecido *Normalized Difference Vegetation Index* (NDVI), o *Leaf Area Index* (LAI) e as métricas de fenologia derivadas dos perfis temporais NDVI (BONTEMPS et al., 2015).

O NDVI, o indicador mais popular operacionalmente usado para monitoramento de vegetação, é fornecido para garantir continuidade com séries temporais de longo prazo e, assim, permitir a detecção de anomalias. O LAI é uma variável primária intrínseca ao dossel que não deve depender das condições de observação. Esta variável determina o tamanho da interface foliar para troca de energia e massa entre o dossel e a atmosfera. Índices de fenologia, também conhecidos como métricas de NDVI, informam sobre parâmetros-chave específicos da estação de crescimento (data de início, duração da estação e data da taxa máxima de crescimento) (UDROIU et al., 2020).

Os mapas LAI, gerados pelo processador L3B, são gerados para toda a região de interesse com uma resolução espacial de 10 metros. O sistema oferece 3 formas para obtê-los:

- Uma estimativa mono data para cada nova aquisição;
- Uma estimativa de reproprocessamento multidade em tempo real usando “n” aquisições anteriores;
- Uma estimativa gerada ao final da estação de cultivo ajustando um modelo fenológico, que deve permitir obter uma série temporal de maior qualidade;

Como padrão, o Sen2-Agri utiliza o processamento mono data, que pode variar entre 7 e 10 dias para obter o mapa para cada aquisição. Caso a região de monitoramento tenha cobertura dos satélites *Sentinel-2* A e B, a resolução temporal é de aproximadamente 7 dias. Porém, se a região possui cobertura de apenas um dos satélites, essa resolução aumenta para aproximadamente 10 dias.

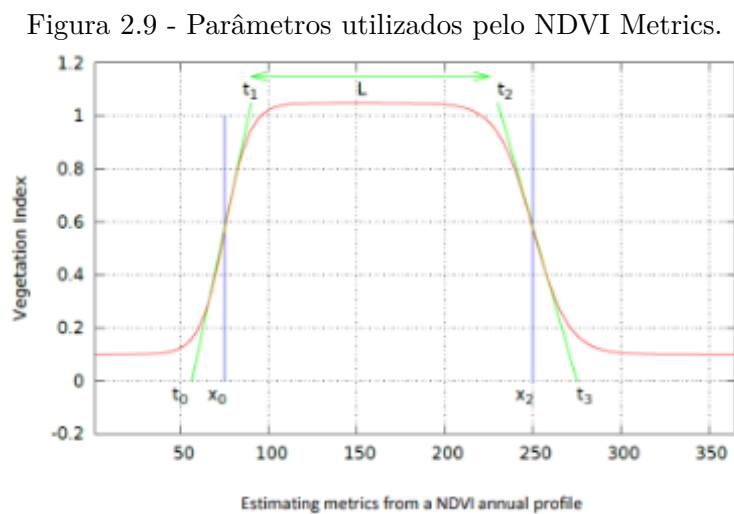
Cada um dos produtos LAI são obtidos através de uma aprendizagem de máquina que é treinado utilizando as bandas 3, 4, 5, 6, 7, 8, 9, 12, 13 para construir um modelo de regressão não-linear estimado utilizando os valores de reflectância simulados usando o modelo *ProSail* (UDROIU et al., 2020).

Os mapas NDVI, obtidos pelo processador L3B, são gerados para toda a região de interesse com uma resolução espacial de 10 metros. Seu processamento é feito através de uma estimativa mono data para cada nova aquisição das séries temporais, e, assim como o LAI, pode variar entre 7 e 10 dias dependendo da cobertura de satélite para a região. Os mapas NDVI são obtidos através da fórmula padrão aplicada sobre as bandas red (B4) e vegetation red-edge (B8a).

Os mapas utilizando os índices fenológicos, gerados pelo processador L3E, são entregues para toda a região de interesse com uma resolução espacial de 10 metros. Seu processamento é feito através da extração de perfis temporais de NDVI por meio de uma função logística dupla ajustada a cada *pixel* (UDROIU et al., 2020). Esta função ajustada é utilizada para estimar os seguintes dados do cultivo:

- A data inicial para a estação de cultivo (t_0);
- A data da taxa máxima de crescimento (x_0);
- O comprimento do platô de maturidade (L);
- A data de término da estação de crescimento correspondente à senescência ou colheita (x_3);

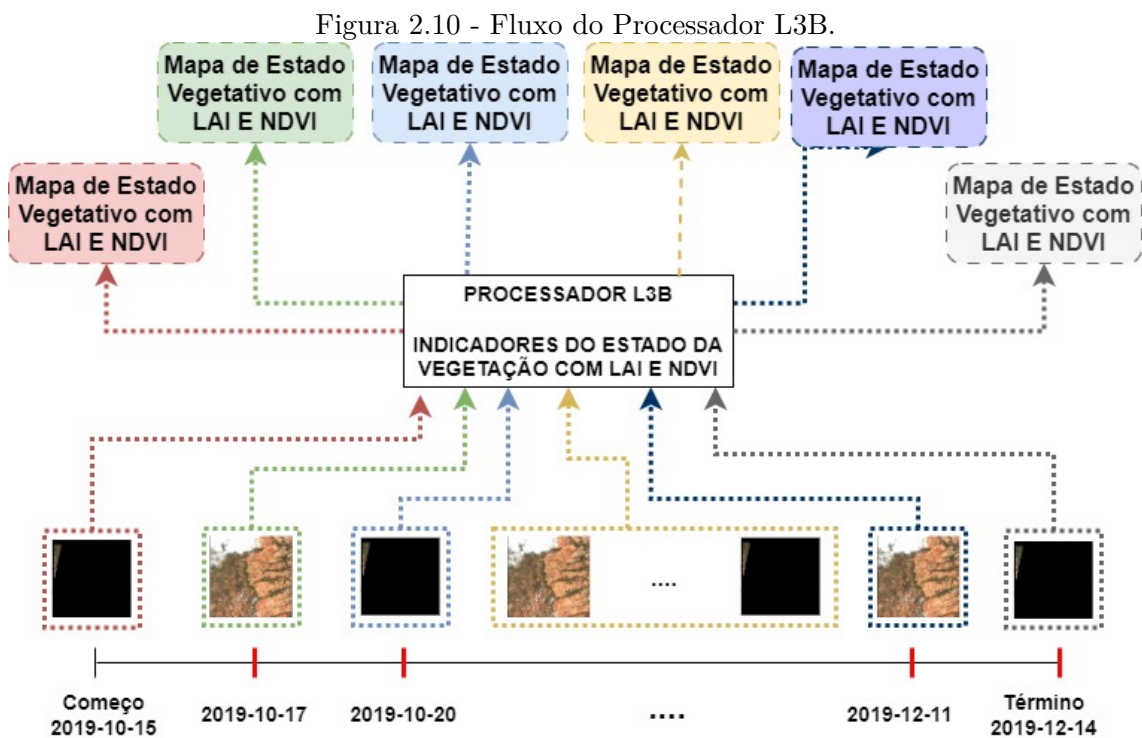
A Figura 2.9 mostra uma função ajustada e a estimativa dos dados do cultivo utilizando o *NDVI metrics*.



Fonte: ESA (2016c).

Diferente dos mapas LAI e NDVI, o índice de fenologia gera dois mapas, um na metade e outro no final da estação de cultivo.

Para exemplificar o funcionamento dos processadores L3B e L3E, foi definida uma estação de cultivo com quatro meses de duração, com início em 15/10/2019 e término em 15/02/2020. A Figura 2.10 representa o processamento de cada uma das imagens disponíveis para o intervalo temporal e, nos resultados, são esquematizados os mapas NDVI e LAI de cada uma das imagens obtidas. Na Figura 2.11, pode-se visualizar a esquematização de dois mapas de estado vegetativo obtido com as métricas fenológicas, um na metade da estação, em 15/12/2019, e o outro ao término, em 15/02/2020.

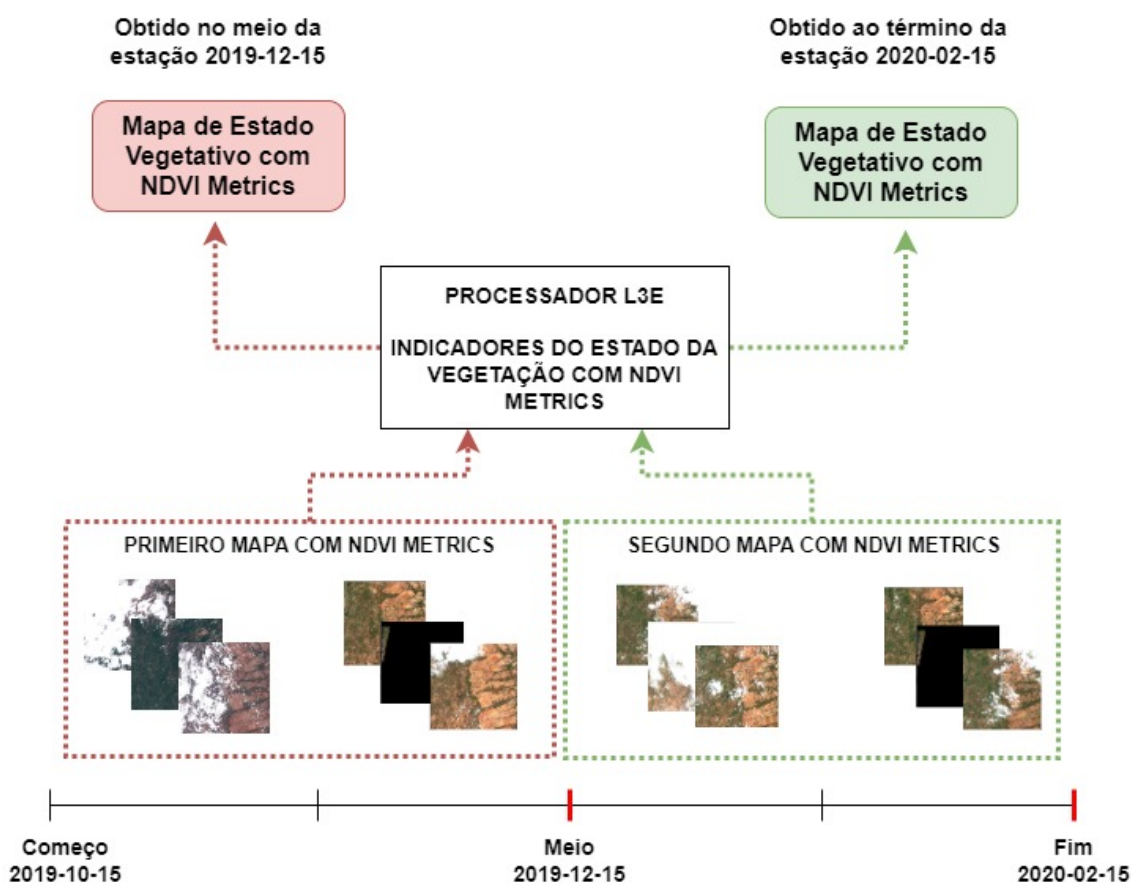


Fonte: Produção do autor.

2.4.4 Máscara dinâmica de cultivo - L4A

A máscara dinâmica de cultivo consiste em um mapa binário com resolução espacial de 10 metros que separa áreas agrícolas anuais de outras áreas (BONTEMPS et al., 2015). Segundo (ESA, 2016b), “uma área anual é definida como um pedaço de terra com no mínimo 0,25 ha de área, atualmente semeada/cultivada e colhível pelo menos uma vez dentro do ano seguinte à data de semeadura”.

Figura 2.11 - Fluxo do Processador L3E.

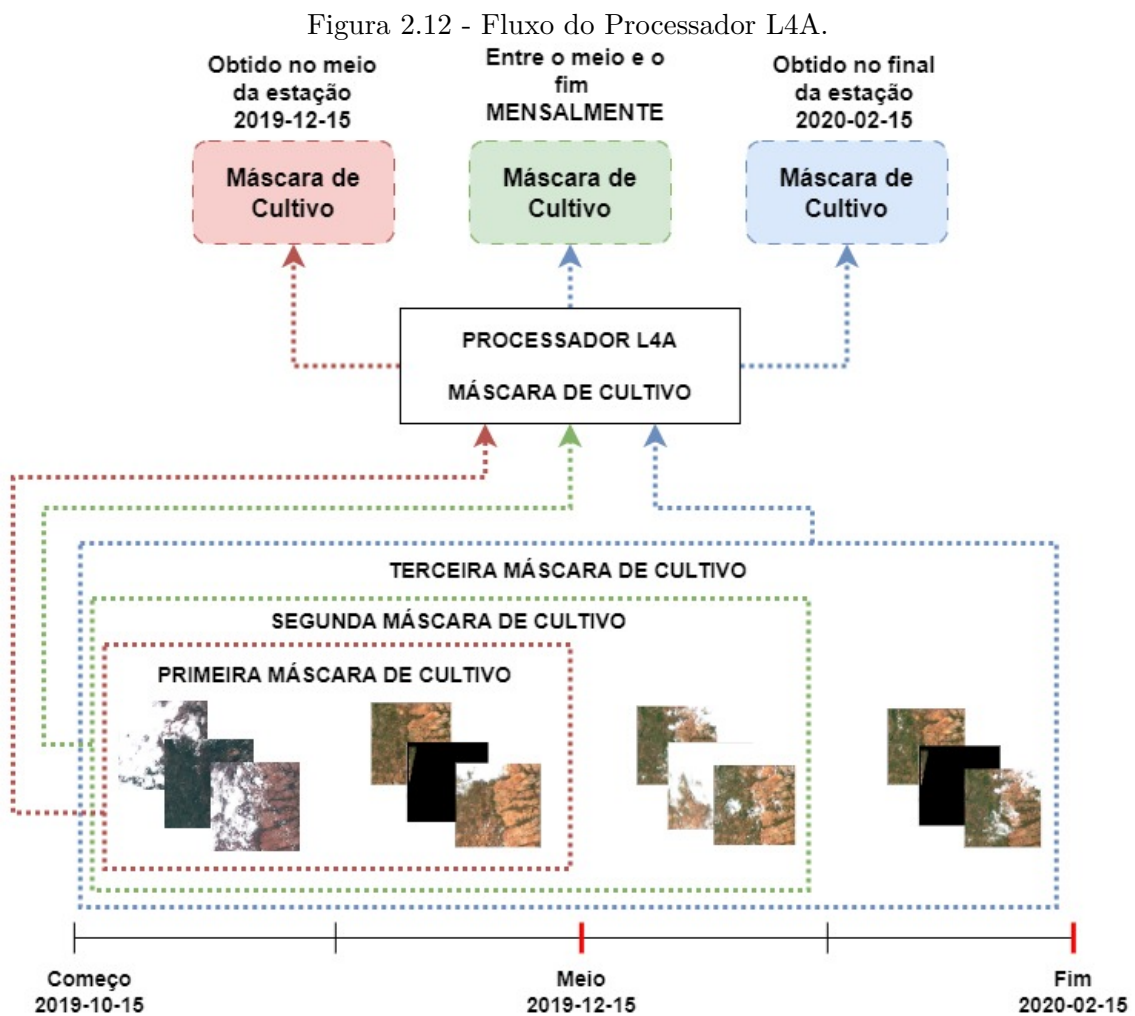


Fonte: Produção do autor.

As máscaras de cultivo podem ser geradas a partir de amostras de campo (*in situ*) representativas de áreas de culturas e de não-culturas. Alternativamente, o sistema pode extrair amostras a partir um “mapa de cultura de referência”. Um mapa de cultura de referência padrão existe no sistema, mas o usuário pode fornecer seu próprio mapa. Na primeira abordagem, os dados *in situ* são obrigatórios para permitir a execução do processador. Em ambos os casos, o produto é gerado usando um classificador *Random Forest*, complementado sob demanda por uma filtragem a posteriori para suavizar a máscara resultante (UDROIU et al., 2020).

Cada máscara obtida possui resolução espacial de 10 metros e são geradas mensalmente após alcançar a metade da estação de cultivo. Por fim, ao decorrer do tempo, o aprendizado de máquina tem acesso a uma quantidade maior de imagens, permitindo que a taxa de acurácia aumente e o produto fique com maior qualidade e assertividade.

Para exemplificar o funcionamento do processador L4A, foi definida uma estação de cultivo com quatro meses de duração, com início em 15/10/2019 e término em 15/02/2020. A partir da metade da estação, mensalmente, é feito o processamento das imagens e são geradas as máscaras. A primeira máscara gerada utiliza as imagens do início até a metade da estação. Após isso, é feita a utilização das mesmas imagens utilizadas na primeira máscara, adicionando as imagens captadas durante o mês que decorreu. Este processo é feito sucessivamente. Assim, quanto mais tempo passar, mais imagens serão utilizadas pelo processador para gerar as máscaras. A figura 2.12 ilustra este exemplo dado acima.



Fonte: Produção do autor.

2.4.5 Mapa do tipo de cultivo - L4B

O mapa do tipo de cultivo, gerado pelo processador L4B, é um mapa dos principais tipos de culturas (ou grupos de culturas) para uma determinada região com resolução espacial de 10 metros (BONTEMPS et al., 2015).

Os tipos de cultura são classificados apenas para a área cultivada identificada na máscara de cultivo. Os principais tipos de cultivo são definidos como (i) aqueles que cobrem uma área mínima de 5% da área cultivada anual e (ii) cuja área acumulada representa mais de 75% da área cultivada anualmente. A justificativa para esses limites de 5% e 75% é evitar tipos de cultura que são fracamente representados e, portanto, muito caros para campanha de campo. Um máximo de 5 tipos de culturas são consideradas por local (BONTEMPS et al., 2015).

Os mapas de cultivo são processados seguindo uma abordagem de classificação *Random Forest* supervisionada, baseada em dados *in situ* a serem coletados durante a estação de crescimento em questão. As informações *in situ* são divididas entre amostras de treinamento usadas para alimentar o processo de classificação e amostras de validação usadas para avaliar a precisão do produto por meio do cálculo de vários critérios (Precisão geral, *F-Score*) (ESA, 2016a).

Os mapas de cultivos são entregues duas vezes ao longo da estação - na metade e ao final. Este produto tem um aumento na taxa de acurácia ao decorrer da estação devido a uma maior quantidade de imagens disponíveis para análise. Para gerar os mapas, este processador requer como entrada as imagens de observação da terra, um dado de amostra e uma máscara de cultivo da região. Este último serve para auxiliar o classificador a identificar as áreas cultivadas das não-cultivadas, permitindo a classificação dos tipos de cultura para as áreas destacadas como cultivadas.

2.4.6 Arquitetura do sistema

Após a descrição sobre os produtos fornecidos pelo Sen2-Agri, o próximo passo é entender qual é o fluxo de funcionamento deste pacote. O início deste fluxo ocorre a partir da criação de uma região de monitoramento (*site*), que é definida através dos seguintes parâmetros:

- **Área:** define qual a extensão espacial que delimita a região de monitoramento. Pode ser informada através de um arquivo *shapefile*;
- **Estação de cultivo (*Season*):** é o intervalo temporal no qual será feito

o monitoramento da região;

- **Dado *in Situ*:** são arquivos que contém polígonos espaciais indicando a presença e o tipo de cultivo obtido através de observações do campo durante a principal estação de cultivo (DEFOURNY et al., 2019).

A Figura 2.7 ilustra o fluxo de funcionamento do Sen2-Agri. Na primeira etapa, após a definição do site, o sistema automaticamente cria uma agenda para baixar as imagens de Observação da Terra, provenientes dos satélites *Sentinel-2* e *Landsat-8*, que atendam as extensões espacial e temporal definidas pelo usuário. Caso a data final da estação de cultivo seja uma data futura, o sistema utiliza essa agenda para, ao passar do tempo, adquirir novas e gerar produtos em tempo quase real.

Após a aquisição das imagens, cada uma delas passa por um processamento para converter as imagens que representam a refletância do topo da atmosfera em imagens com correção atmosférica (L2A), ou seja que representam a reflectância na superfície. Estes produtos serão utilizados como dado de entrada para os demais processadores. Isso permite que o orquestrador do Sen2-Agri possa agendar e iniciar o processamento das imagens para gerar os produtos agrícolas desejados (L3A, L3B, L4A ou L4B). Caso o usuário deseje gerar os produtos de máscara ou mapa de cultivo, no processo de definição do sítio ele tem que alimentar o sistema com o dado *in situ* para a região de monitoramento.

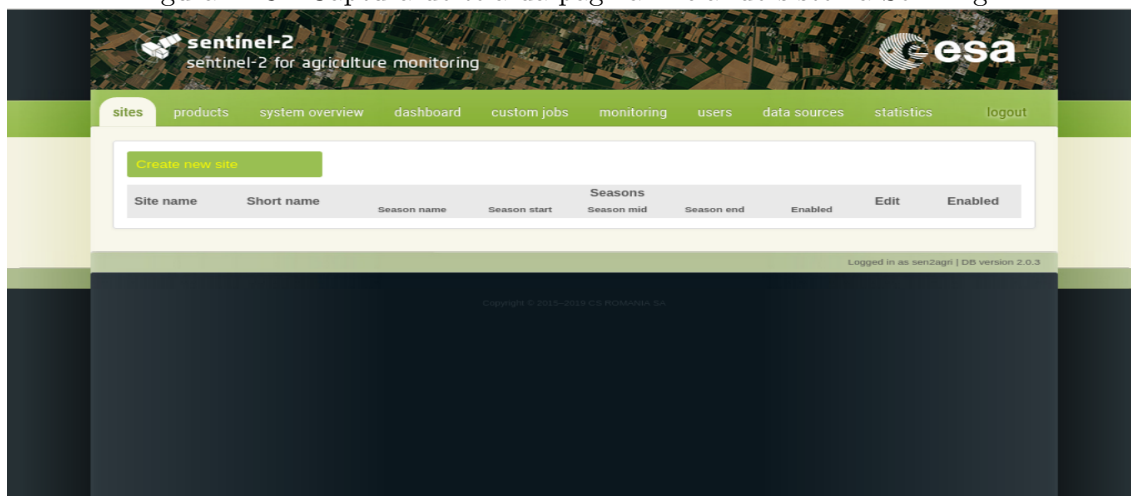
2.4.7 Modos de uso

Para criar uma rotina de monitoramento para uma sítio, o Sen2-Agri dispõe de duas alternativas: a manual e a automática.

O modo automático disponibiliza uma interface *web*, representada na Figura 2.13, que fornece ao usuário um conjunto de funcionalidades para ele definir uma região de interesse, informar dados de amostra, acompanhar a aquisição e processamento das imagens L1C e visualizar os produtos obtidos para o sítio definido.

A utilização do modo manual do sistema permite ao usuário realizar um processamento manual independente para cada um dos processadores, sem necessitar da instalação de todo o sistema. Para isso, são disponibilizadas duas opções: via terminal ou *Sentinel-2 Toolbox / Sentinel Application Platform* (SNAP). Ambas as opções podem ser executadas utilizando scripts escritos na linguagem de programação *Python*.

Figura 2.13 - Captura de tela da página inicial do sistema Sen2-Agri.



Fonte: Produção do autor.

O presente trabalho realizou a instalação e configuração do sistema de maneira que permita a sua utilização tanto no modo manual quanto no automático, que é o fluxo de configuração padrão para o Sen2-Agri e que permite utilizar o melhor dos dois modos.

2.5 Brazil Data Cube

O *Brazil Data Cube* (BDC) é um projeto do Instituto Nacional de Pesquisas Espaciais (INPE). Esse projeto está produzindo dados a partir de grandes volumes de imagens de sensoriamento remoto de média resolução para todo o território nacional e desenvolvendo uma plataforma computacional para processar e analisar esses dados usando inteligência artificial, aprendizado de máquina e análise de séries temporais de imagens. (BRAZIL DATA CUBE, 2023a).

Segundo (FERREIRA et al., 2020), “o *Brazil Data Cube* surgiu através de projetos do INPE que tinham por objetivo explorar as imagens inéditas de sensoriamento remoto disponíveis gratuitamente hoje em dia utilizando tecnologias de big data, ambientes de computação em nuvem e métodos de aprendizagem de máquina. Um destes projetos é o *E-Sensing*, que investigou novos métodos e técnicas para melhorar substancialmente a extração de informações de mudança de uso e cobertura da terra de grandes conjuntos de dados de Observação da Terra de maneira aberta e reproduzível. Baseado nas informações adquiridas no projeto *E-Sensing* e nas demandas nacionais de monitoramento de uso e cobertura da terra, foram definidos os requisitos para construir os cubos de dados do Brasil”.

Este projeto tem quatro objetivos principais: criar coleções de imagens que requerem esforço mínimo para serem analisadas com imagens de média resolução para todo o território brasileiro; modelar essas coleções em cubos de dados com três ou mais dimensões para permitir a análise de séries temporais; utilizar, propor e desenvolver tecnologias de *big data*, computação em nuvem e ambientes de processamento distribuídos para criar, processar e armazenar estes cubos; e, por último, criar informações de uso e cobertura da terra dos cubos de dados de todo o território nacional.

Para consumir a infraestrutura do projeto, foram criados *softwares* gratuitos e livres, que se encaixam em duas categorias, serviços e aplicações. Todos os conjuntos de dados e metadados são acessados e processados através de serviços *web*. A categoria de aplicações se refere a produtos de *software* para o usuário final incluindo sistemas com interface gráfica e Interfaces de Programação de Aplicações (API, do inglês Application Programming Interface) para as linguagens de programação R e *Python*. A categoria de serviços *web* promove a interface entre os conjuntos de dados do projeto e as aplicações (FERREIRA et al., 2020).

Dentre os serviços e ferramentas disponibilizados pelo *Brazil Data Cube*, o presente trabalho utilizou: o *Brazil Data Cube SpatioTemporal Asset Catalog* (BDC STAC), o *Web Land Trajectory Service* (WLTS), o *Web Land Classification System Service* (WLCSS) e o *Brazil Data Cube Image Metadata Catalog* (BDC CATALOG).

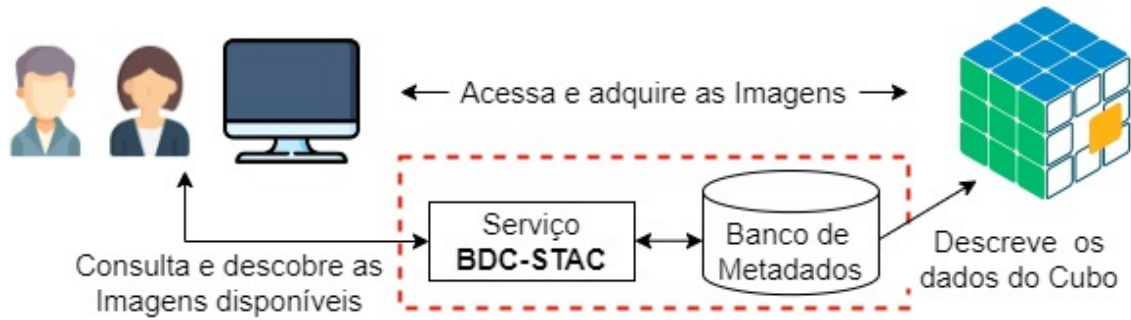
2.5.1 *Brazil Data Cube SpatioTemporal Asset Catalog*

Segundo (ZAGLIA et al., 2019), “O STAC é um padrão que especifica como metadados de recursos geoespaciais [...] são organizados, consultados e disponibilizados na *web*. Seu principal foco é a catalogação de metadados de imagens provenientes de observações da Terra por sensores orbitais”.

O usuário pode acessar os recursos disponibilizados pelo STAC através de duas maneiras: acesso estático e acesso dinâmico. A forma estática ocorre através da navegação por meio de *links* (manipulação de arquivos JSON). O acesso dinâmico é feito através de requisições para serviços *web* que atendam as especificações do STAC API.

Um protótipo de um serviço *web* foi desenvolvido pelo BDC para que o protocolo STAC pudesse ser adaptado as necessidades do projeto. Assim, foi gerado o *Brazil Data Cube SpatioTemporal Asset Catalog* (BDC-STAC). Sua implementação utiliza o microframework *Flask*, do *Python*, para tratamento das requisições e respostas

Figura 2.14 - Fluxo de funcionamento do BDC STAC.



Fonte: Adaptado de Zaglia et al. (2019).

HTTP. A partir das requisições, são realizadas consultas em um banco de dados MySQL onde estão contidos os metadados que descrevem o BDC, ilustrado na Figura 2.14.

Para atender as necessidades do BDC, foi criada uma extensão que define propriedades específicas. Estas propriedades são utilizadas no STAC *Collections* e STAC *Items* e podem ser identificadas através do prefixo “**bdc:**”.

Com o desenvolvimento do serviço BDC-STAC, usuários e aplicações conseguem pesquisar e descobrir todas as coleções de imagens e cubos de dados armazenados nos bancos de dados do projeto a partir de uma interface padrão fornecida pelo serviço STAC (FERREIRA et al., 2020).

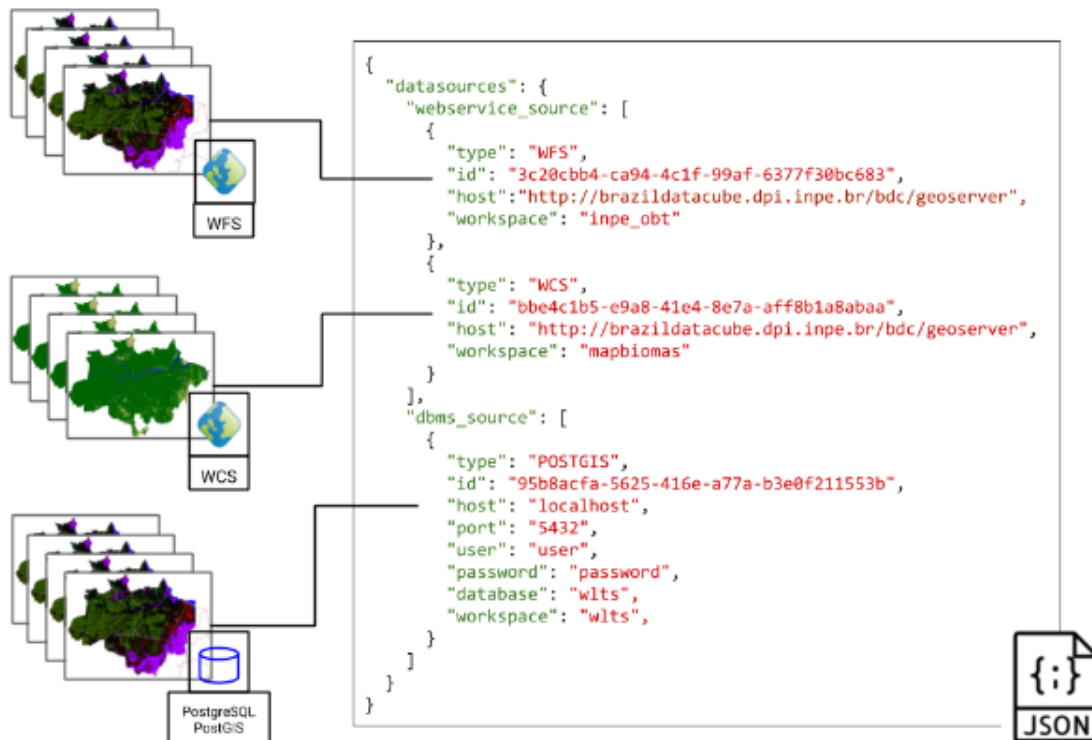
2.5.2 *Web Land Trajectory Service*

A trajetória de uso e cobertura da terra é o conceito que se refere a identificar qual é o uso da terra numa determinada localização num intervalo de tempo. A partir deste conceito, (ZIOTI et al., 2022) desenvolveu um serviço, denominado *Web Land Trajectory Service* (WLTS), que permite acessar, integrar e retornar trajetórias de uso e cobertura da Terra de diferentes fontes de dados.

Para consultar fontes de dados distintas, o WLTS acessa arquivos de metadados, em formato JSON, que descrevem como as fontes são fornecidas e armazenadas. Este padrão de acesso permite a adição de novas fontes de dados. A Figura 2.15 ilustra um exemplo de fontes de dados preenchidas.

Através de uma API simples, o usuário pode indicar uma localização espacial e o WLTS é capaz de acessar diferentes fontes de dados para aquela localização e retornar sua classificação, ao longo do tempo, rotulada de acordo com cada uma das

Figura 2.15 - Exemplo de um arquivo JSON com fontes de dados acessíveis pelo WLTS.



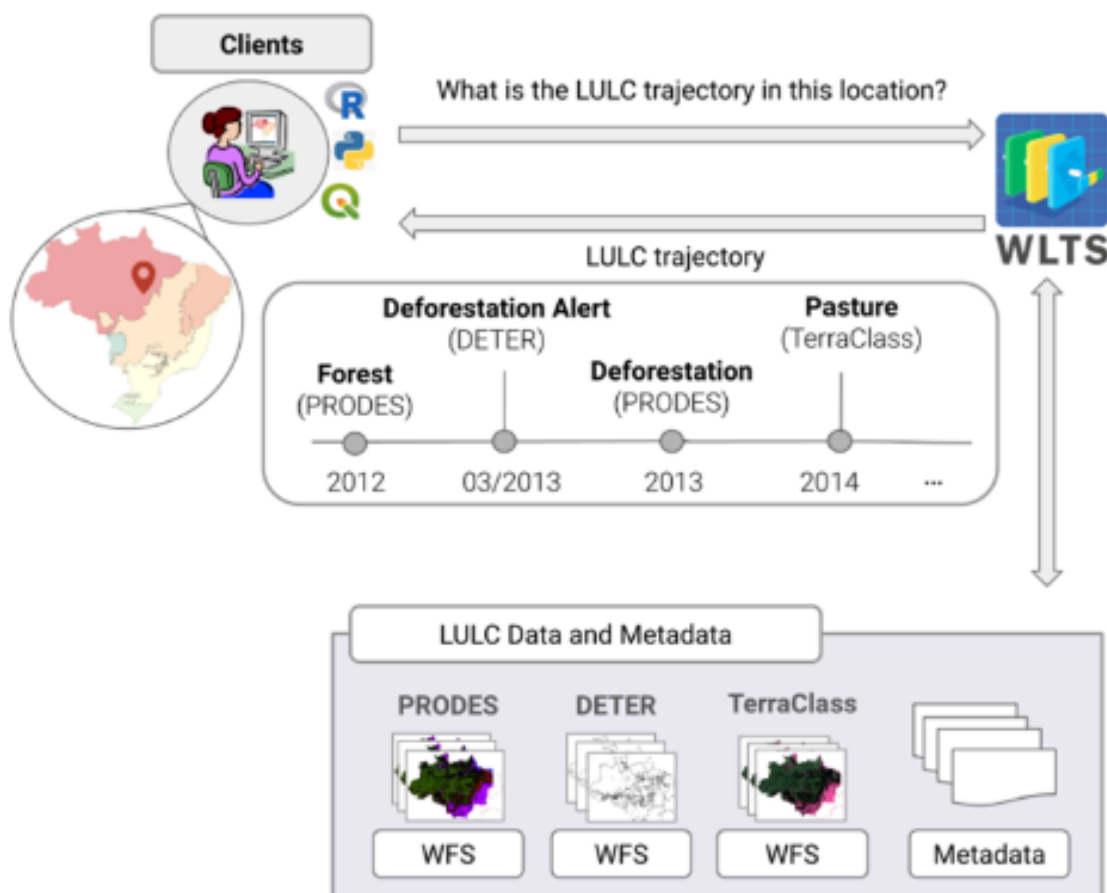
Fonte: Zioti et al. (2022).

fontes. Esta operação descrita anteriormente é denominada de **trajectory**. Além dela, outras duas operações são disponibilizadas pela API do *Web Land Trajectory Service*:

- **list_collections**: retorna uma lista com os nomes das coleções disponíveis no repositório acessado pelo serviço. Estes nomes são utilizados como identificadores das coleções nas operações *trajectory* e *describe_collection*, detalhada a seguir;
- **describe_collection**: retorna o metadado de uma determinada coleção selecionada pelo usuário;

A Figura 2.16 ilustra o funcionamento da operação *trajectory*. O usuário fornece uma localização espacial ao serviço. Esta localização serve de entrada para que o WLTS consulte as fontes de dados disponíveis no projeto BDC e obtenha a trajetória de uso e cobertura da terra para aquela determinada localização.

Figura 2.16 - Fluxo de funcionamento do WLTS.



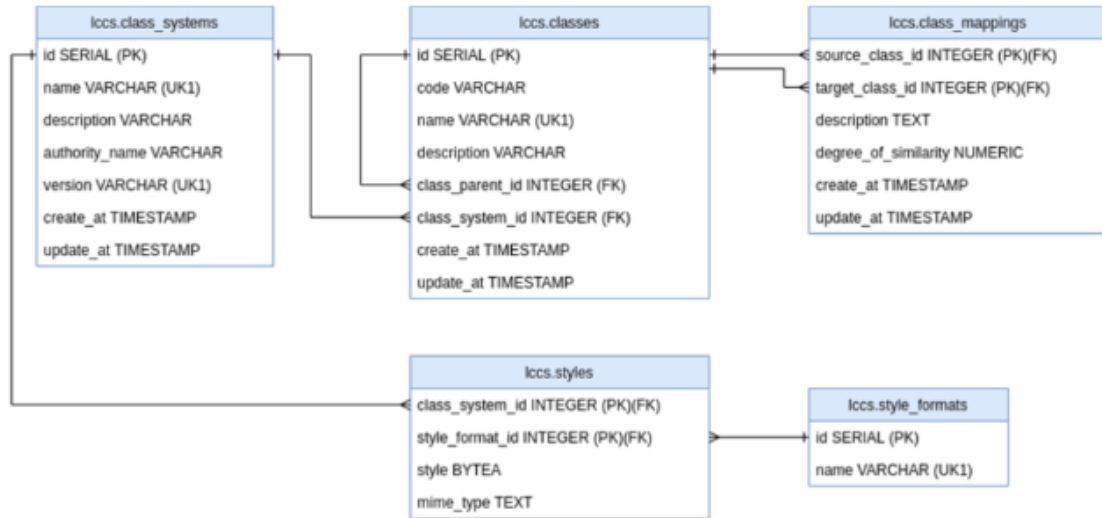
Fonte: Zioti et al. (2022).

2.5.3 Web Land Classification System Service

Segundo (ZIOTI et al., 2022), “o Web Land Classification System Service (WLCCS) é designado para representar e mapear sistemas de classificação para coleções de dados de uso e cobertura da Terra (LULC, do inglês Land User Land Cover). Ele fornece uma interface simples para usar e consultar o banco de dados de metadados, apresentado na Figura 2.17, que representa diferentes sistemas de classificação, hierárquicos ou não, e suas respectivas classes”.

O WLCCS foi projetado considerando as questões de reutilização de dados. Ele fornece um ambiente que permite os pesquisadores e especialistas descreverem sistemas de classificação associados com diferentes coleções de dados de uso e cobertura da Terra e como eles podem se relacionar, ou seja, este serviço ajuda os usuários a integrar, harmonizar e comparar distintas coleções de dados LULC (ZIOTI et al.,

Figura 2.17 - Modelo de banco de dados do WLCSS.



Fonte: Zioti et al. (2022).

2022).

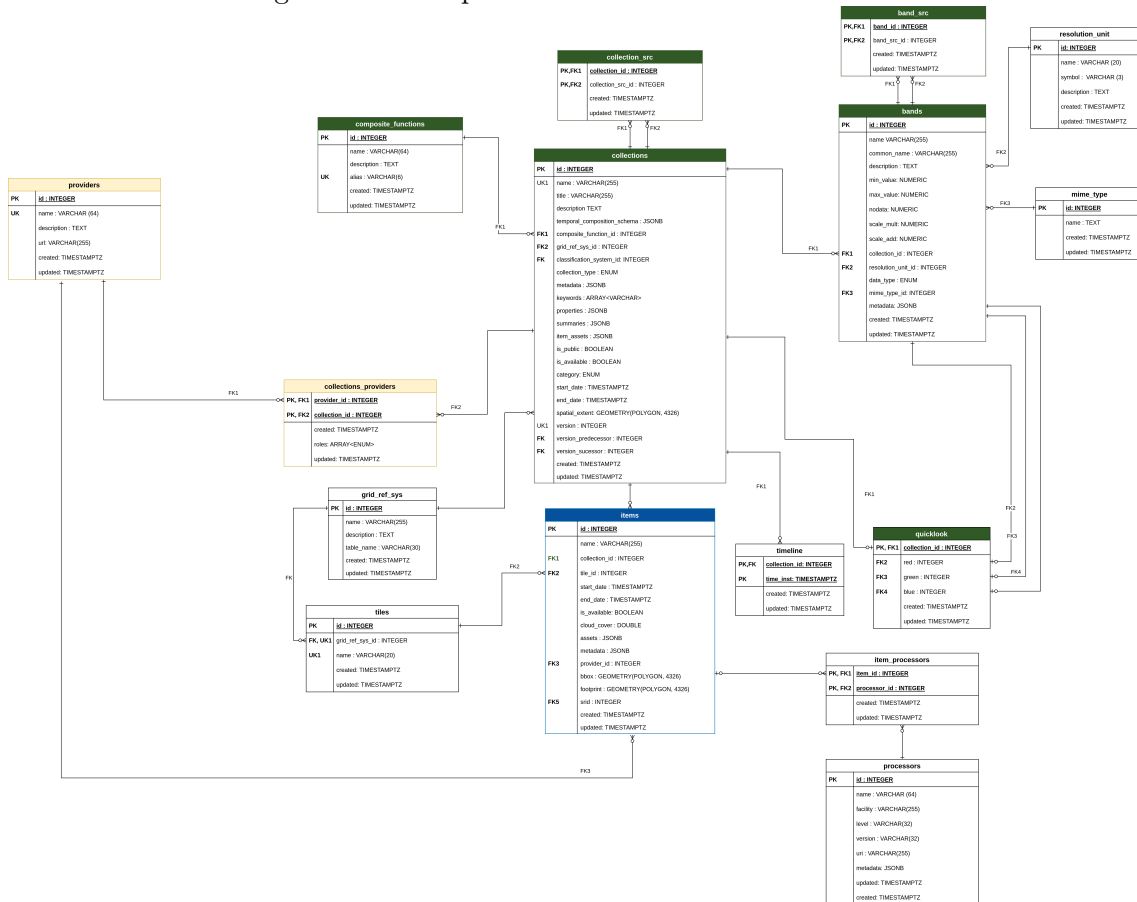
O WLCSS disponibiliza uma API, escrito nas linguagens R e *Python*, com um conjunto de operações para consultar e mapear as coleções de dados LULC. Abaixo serão apontadas algumas das operações fornecidas pelo serviço:

- ***classification_systems***: retorna uma lista com os sistemas de classificação de uso de cobertura da Terra disponíveis no repositório do BDC. Os valores retornados por essa operação podem ser utilizados nas demais funções disponibilizadas pela API;
- ***classification_systems/system_id***: retorna os metadados de um determinado sistema de classificação;
- ***classification_systems/system_id/classes***: retorna a lista de classes pertencentes a um determinado sistema de classificação;
- ***classification_systems/system_id/styles***: retorna uma lista com os formatos de estilos disponíveis para a classificação;
- ***classification_systems/system_id_source/system_id_target***: retorna o mapeamento entre dois sistemas de classificação;

2.5.4 Brazil Data Cube Image Metadata Catalog

O *Brazil Data Cube Image Metadata Catalog* (BDC Catalog) é um pacote base de outros *softwares* desenvolvidos pelo *Brazil Data Cube* que serve para criar uma instância local com o esquema do banco de dados utilizado pelo projeto, representado na Figura 2.18. A criação da base de dados localmente permite o armazenamento de metadados das imagens de Observação da Terra (BRAZIL DATA CUBE, 2023b).

Figura 2.18 - Esquema do banco de dados do BDC.



Fonte: Brazil Data Cube (2023b).

3 INTEGRAÇÃO DO SEN2-AGRI AO *BRAZIL DATA CUBE*

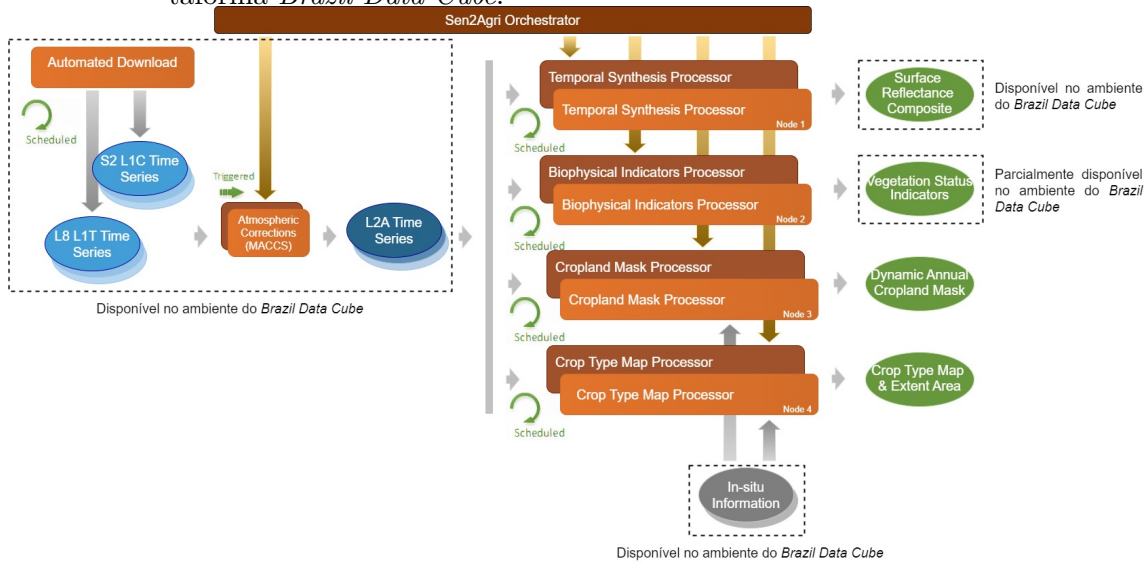
A integração de sistemas de *software* diferentes é em geral buscada para atingir vantagens como aperfeiçoamento de processos, ganho de desempenho e redução de retrabalho. As dificuldades enfrentadas para realizar essa junção são: os diferentes tipos de sistemas, a falta de um formato único dos dados, a interconexão com outros sistemas, questões de gestão, etc. (MILEV, 2019).

Este processo de integração de sistemas pode ser feito através da troca de dados, ou seja, um agente intermediário captar informações de diferentes sistemas e exportar para um sistema único, através do compartilhamento de um banco de dados único entre diferentes sistemas, ou ainda via Interfaces de Programação de Aplicações (API) que fornecem um protocolo conhecido para troca de dados e operações entre sistemas.

O presente trabalho tem por objetivo a integração entre o pacote de *software* Sen2-Agri e a plataforma *Brazil Data Cube*. O Sen2-Agri possui um orquestrador para gerenciar o fluxo do *software*, que tem como primeira tarefa baixar as imagens de Observação da Terra (EO) nos repositórios de dados disponíveis (por padrão, *United States Geological Survey* - USGS - e *Sentinels Scientific Data Hub* - SciHub -) e processá-las utilizando o algoritmo MAJA. Nesta etapa são realizadas correções atmosféricas (produto L2A) que produzem valores de reflexão do fundo da atmosfera com neve, água, máscaras de nuvens e sombra de nuvens presentes nas imagens, com o intuito de melhorar a qualidade delas, uma vez que este produto serve de entrada para os demais processadores do pacote. Após isso, o orquestrador agenda a produção dos produtos agrícolas respeitando as características temporais de cada um deles. Outra tarefa do orquestrador é fornecer ao classificador *Random Forest* as amostras, disponibilizadas pelo usuário, para treinar o algoritmo de aprendizado de máquina, processar as imagens L2A e obter as máscaras e mapas de cultivo. Este processo está representado na Figura 2.7 da subseção Sen2-Agri.

Ao estudar o processo que o Sen2-Agri realiza para obter os seus produtos, percebe-se que a infraestrutura do BDC fornece opções para otimizar o fluxo de funcionamento do sistema, que são elas: imagens e séries temporais L2A já processadas e prontas para uso, compostos mensais da reflectância da superfície livre de nuvens, produtos de estado vegetativo obtidos com o NDVI e os dados *in situ* utilizados como amostras para os classificadores *Random Forest*. Este levantamento está representado na Figura 3.1.

Figura 3.1 - Orquestrador do Sen2-Agri com indicações dos módulos disponíveis na plataforma *Brazil Data Cube*.



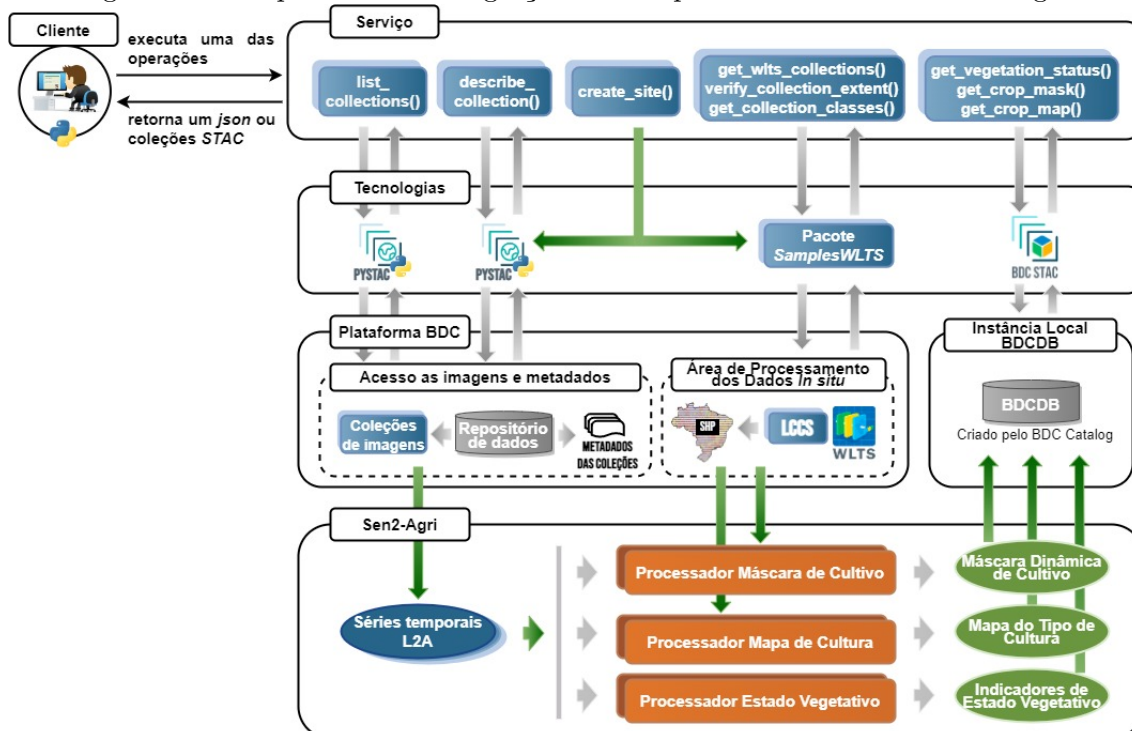
Fonte: Produção do autor.

A partir deste estudo, outra questão observada é que a partir da utilização do Sen2-Agri, é possível obter uma coleção de mapas para acompanhamento dos cultivos bem completa, com métodos publicados na literatura e potencialmente simples para se implantar sistemas de monitoramento contínuo de culturas. No entanto, essa implantação exige uma infraestrutura computacional robusta, incluindo a obtenção e o armazenamento das imagens de entrada. Além disso, a implementação corrente apresenta pouca flexibilidade quanto ao sistema operacional onde roda.

Considerando, então, que alguns componentes da arquitetura do Sen2-Agri já estão disponíveis na plataforma BDC e que o ambiente do projeto fornece uma infraestrutura para otimizar o fluxo do pacote, foi desenvolvida uma arquitetura para realizar a integração, da plataforma BDC com o pacote de *software* Sen2-Agri. Essa arquitetura, mostrada na Figura 3.2, é baseada em APIs e fornece para o usuário um cliente em *Python* com nove operações: *create_site()*, *list_collections()*, *describe_collection()*, *get_wlts_collections()*, *verify_collection_extent()*, *get_collection_classes()*, *get_vegetation_status()*, *get_cropland_mask()* e *get_cropland_type()*. Essas operações serão descritas com mais detalhes a partir da subseção 3.2.

Além das operações, é importante destacar as áreas de acesso aos dados proposta na arquitetura, rotuladas como “Plataforma BDC” e “Instância Local BDCDB”. A área de “Plataforma BDC” é responsável por: obter imagens de observação da

Figura 3.2 - Arquitetura da integração entre a plataforma BDC e o Sen2-Agri.



Fonte: Produção do autor.

Terra e seus metadados através da consulta ao repositório do BDC via `PySTAC`¹; e obter as trajetórias de LULC para gerar os dados de amostra através de um pacote chamado `SamplesWLTS`. O processo de obtenção dos dados de amostra está descrito com mais detalhes a seguir, na subseção 3.1. A área de “Instância Local BDCDB” dispõe de um banco de dados local estruturado de acordo com o esquema do BDC através do BDC Catalog. Este banco armazena os produtos “Sen2-Agri” e os dados de amostra obtidos. O acesso a esta área é feito através do serviço BDC STAC.

3.1 Dados de amostra

Os dados de amostra (*in situ*) são conjuntos de dados que identificam como a terra tem sido utilizada em uma determinada região num intervalo de tempo. Para utilizar esses conjuntos de amostras, os processadores Sen2-Agri exigem que alguns atributos específicos estejam presentes no dado, que são eles:

- **ID:** é um valor inteiro utilizado como indicador único para cada *pixel*;

¹`PySTAC` é uma biblioteca escrita em *Python* utilizada para trabalhar com a especificação STAC.

- **LC**: é uma cadeia de caracteres que representa o nome da classe de uso e cobertura da Terra atrelada a cada um dos *pixels*;
- **CODE**: é um valor inteiro utilizado para enumerar as classes de uso e cobertura da Terra. Portanto, cada uma das classes recebe um “*code*” único que serve para identificá-la;
- **CROP**: é um atributo binário que identifica se a classe associada ao *pixel* é uma classe de cultivo ou não (valor 1 para classes de cultivo e 0 para as de não cultivo).

Estes conjuntos de dados são utilizados para treinar e validar o classificador *Random Forest* utilizado pelos processadores de máscara e mapa de cultivo. A adição dos atributos acima mencionados serve para ajustar o algoritmo de aprendizado de máquina do classificador empregado, disponibilizando informações sobre classes previamente identificadas na região e quais delas estão atreladas ao conceito de classe de cultivo.

Tendo em vista que esses dados de amostra identificam a trajetória de uso e cobertura da Terra para o sítio definido, a arquitetura desenvolvida no presente trabalho elaborou um módulo que contém um conjunto de funções, denominado *SamplesWLTS*, que utiliza: o *Web Land Trajectory Service* para obter essas as trajetórias de uso e cobertura da Terra e gerar as amostras de acordo com as especificações necessárias; e o *Land Cover Classification System* (LCCS) para mapear as classes de uso e cobertura da Terra para uma determinada coleção WLTS.

Para gerar as amostras, o *SamplesWLTS* pega o par de coordenadas mínima e máxima do *bouding-box* da extensão espacial do sítio ((**xmin**, **ymin**), (**xmax**, **ymax**)) e gera uma grade com resolução espacial de 80x80. Para cada um dos “*tiles*” dessa grade é extraído o centróide e obtido uma lista com todas as localizações espaciais da região. Em sequência, é definida qual coleção WLTS será utilizada. Para isso, o sistema tem dois caminhos a seguir: consultar as coleções e definir uma delas ou utilizar uma coleção escolhida pelo usuário. Partindo da alternativa no qual o usuário não informa uma coleção, o sistema analisa quais são as coleções que interceptam a extensão espacial e temporal do sítio e escolhe uma delas aleatoriamente.

Escolhida a coleção WLTS a ser utilizada, o sistema ajusta o período de consulta para o ano anterior a data de início da janela temporal do sítio definida pelo usuário. Por exemplo, caso seja definida a extensão espacial com início em Junho de 2019,

o sistema estabelece o período de busca no WLTS para 2018. Esta abordagem visa gerar dados de amostras o mais atual possível, respeitando o intervalo de tempo definido pelo usuário para monitorar a região.

Com os parâmetros da operação *trajectory* configurados, o *SamplesWLTS* utiliza o WLTS para consultar as trajetórias LULC para cada uma das localizações da área de monitoramento e armazena as respostas em um objeto *dataframe* (que é uma tabela com os dados separados por colunas).

O dataframe gerado já possui dois dos quatro atributos requisitados pelas especificações do Sen2-Agri (o **id** e o **lc**). Para rotular cada um dos registros do *dataframe* para o atributo *code*, o script pega todas as classes obtidas ao consultar as trajetórias e atribui um valor inteiro, partindo do 1, para cada uma delas. Em relação a propriedade *crop*, ela atribui o valor 0 ou 1 para cada um dos registros de acordo com a classe que ele representa. Esta verificação é feita através de uma lista de classes que representam cultivo. Caso a classe do registro esteja presente na lista de classes de cultivo, é configurado o valor 1. Se não, é associado o valor 0.

Essa lista que contém as classes que representam o cultivo pode ser obtida através de duas maneiras: informada pelo usuário ou obtida pelo sistema. A primeira alternativa ocorre quando o usuário informa uma coleção WLTS de interesse e uma lista com classes que ele quer que sejam relacionadas a cultivo. A outra maneira ocorre quando o usuário não informa uma coleção WLTS e nem a lista de classes. Quando isso ocorre, o sistema utiliza uma lista pré-definida com classes relacionadas a cultivo.

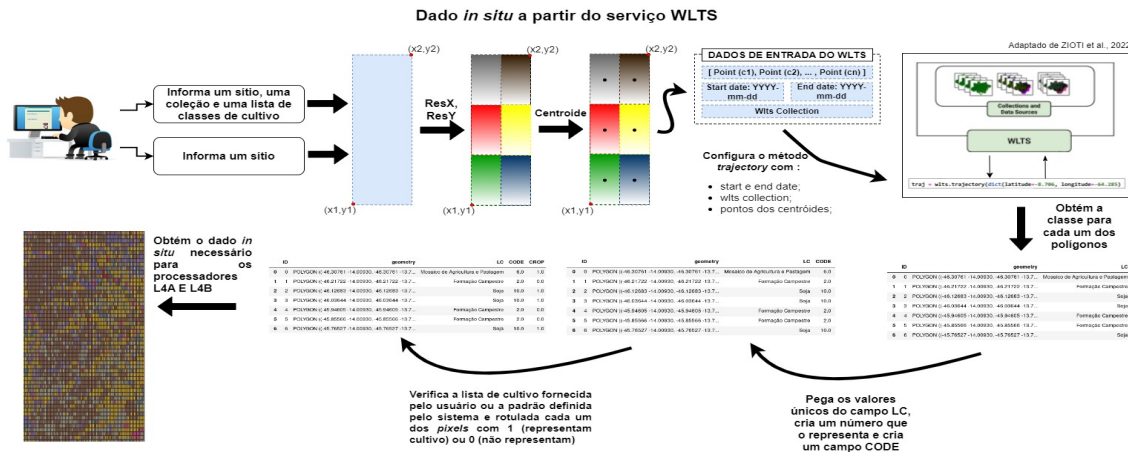
Por fim, o último passo do fluxo módulo *SamplesWLTS* é exportar esse dataframe em um dado vetorial (arquivo *shapefile*), obtendo, assim, o dado de amostra personalizado para o sítio. O procedimento citado nesta subseção está representado na Figura 3.3.

Esta seção abordou o módulo *SamplesWLTS* projetado para ajustar e obter os dados de amostra requisitados pelos processadores de máscara e mapa de cultivo. A partir da próxima seção, serão abordadas as operações idealizadas na arquitetura de integração proposta.

3.2 list_collections

Esta operação recupera uma lista das coleções disponíveis para consulta. A tabela 3.1 apresenta os parâmetros suportados por essa operação. Quando executada, ela utiliza o cliente *PySTAC* para se comunicar com o ambiente do BDC e consultar

Figura 3.3 - Processo de obtenção dos dados in situ utilizando o WLTS.



Fonte: Produção do autor.

todas as coleções disponíveis.

Tabela 3.1 - Parâmetros da operação `list_collections`.

Parâmetro	Descrição	Requerido
<code>result_cols</code>	Booleano que modifica o comportamento da operação. Caso seu valor seja verdadeiro, o método recupera uma lista com todas as coleções “Sen2-Agri” disponíveis. Se não, é recuperada uma lista com as coleções e séries temporais disponíveis no repositório de dados do BDC. Por padrão, seu valor é configurado como falso.	Não

Como resposta, a operação retorna um *array* de string. Este *array* é composto pelos nomes das coleções disponíveis. A Listagem 1 apresenta o retorno desta operação.

3.3 describe_collection

Esta operação recupera uma lista das coleções disponíveis para consulta. A tabela 3.2 apresenta os parâmetros suportados por essa operação. Quando executada, essa operação utiliza o cliente *PySTAC* para se comunicar com o ambiente do BDC e obter os metadados para a coleção escolhida pelo usuário.

```
1 {  
2   "collections": [ "string" ]  
3 }
```

Listing 1 - Definição de retorno da operação *list_collections*.

Tabela 3.2 - Parâmetros da operação *describe_collection*.

Parâmetro	Descrição	Requerido
name	Nome da coleção	Sim

Como resposta esta operação retorna um arquivo JSON com as informações de uma coleção, informada pelo parâmetro “name”. A listagem 2 apresenta a definição de retorno deste método.

```

1 { "Collection <name>": {
2     "type": "string",
3     "id": "string",
4     "description": "string",
5     "properties": {
6         "eo:bands": "list",
7         "outros properties associados..."
8     } ,
9     "bdc:tiles": "list",
10    "extent": {
11        "spatial": {
12            "bbox": [[
13                xmin: 0,
14                ymin: 0,
15                xmax: 0,
16                ymax: 0
17            ]]
18        },
19        "temporal": {
20            "interval": [[
21                start_date: "string",
22                end_date: "string"
23            ]]
24        },
25    },
26    "outros metadados associados..."
27 }
28 }

```

Listing 2 - Definição de retorno da operação *describe_collection*.

3.4 get_wlts_collections

Recupera uma lista com as coleções disponíveis no serviço WLTS. Quando o usuário executa essa operação, ela utiliza o conjunto de funções *SamplesWLTS* para acessar uma instância do *Web Land Trajectory Service* e consultar quais as coleções LULC disponíveis.

Como resposta, ela retorna um *array* de *string*. Este *array* é composto pelos nomes das coleções disponíveis. A Listagem 3 apresenta o retorno desta operação.

```

1 {
2   "WLTS collections": [ "string" ]
3 }

```

Listing 3 - Definição de retorno da operação *get_wlts_collections*.

3.5 verify_collection_extent

Recupera uma mensagem informando se as extensões espacial e temporal da coleção interceptam ou não as extensões do sítio. A tabela 3.3 apresenta os parâmetros suportados por essa operação.

Tabela 3.3 - Parâmetros da operação *verify_collection_extent*.

Parâmetro	Descrição	Requerido
collection_name	Nome da coleção	Sim
period	Extensão espacial do sítio	Sim
site_extent	Extensão temporal do sítio	Sim

Quando executada, essa operação utiliza o *SamplesWLTS* para acessar uma instância do *Web Land Trajectory Service* e buscar os metadados da coleção escolhida para verificar se as extensões espacial e temporal da coleção interceptam as do sítio.

Como resposta, ela retorna uma *string* que notifica se a coleção contém ou não dados para o sítio, permitindo, assim, que o usuário consiga ter informações para definir qual coleção ele irá utilizar para gerar dados de amostra personalizados. A Listagem 4 apresenta o retorno desta operação.

```

1 {
2     "Mensagem": "string"
3 }

```

Listing 4 - Definição de retorno da operação *verify_collection_extent*

3.6 get_collection_classes

Recupera uma lista com as classes de uso e cobertura da Terra disponíveis para a coleção escolhida dentre as disponibilizadas pelo WLTS para consulta. A tabela 3.4 apresenta os parâmetros suportados por essa operação.

Tabela 3.4 - Parâmetros da operação *get_collection_classes*.

Parâmetro	Descrição	Requerido
name	Nome da coleção escolhida	Sim

Quando executada, essa operação utiliza o módulo *SamplesWLTS* para acessar uma instância do *Web Land Trajectory Service*, associa uma coleção a essa instância e utiliza a operação *classification_system* para obter o nome do sistema de classificação associado a coleção informada. Em seguida, um objeto do serviço LCCS é instanciado para obter as classes do sistema de classificação desejado.

Como resposta, ela retorna um *array* de *string*. Este *array* é composto pelos nomes das classes utilizadas pelo sistema de classificação escolhido. A Listagem 5 apresenta o retorno desta operação.

```
1 {
2     "Classes da coleção <name>": ["string"]
3 }
```

Listing 5 - Definição de retorno da operação *get_collection_classes*.

3.7 create_site

Define um sítio e gera os produtos “Sen2-Agri” para ele. A tabela 3.5 apresenta os parâmetros suportados por essa operação.

Em relação aos parâmetros da operação *create_site*, na tabela 3.5 é especificado que o *collection_name* e o classes não são requeridos. Isso quer dizer que a função funciona normalmente caso o usuário não informe-os. Eles servem como modificadores do comportamento da operação. Caso o usuário informe-os, a função irá obter dados de amostras personalizados utilizando a coleção WLTS escolhida e rotular as classes informadas através do parâmetro *classes* como classes que representam cultivo.

Quando executada, essa operação utiliza o cliente PySTAC para se comunicar com o ambiente BDC e obter as imagens de Observação da Terra para o sítio delimitado. Em seguida, o processamento dos produtos escolhidos pelo usuário são agendados. Caso a máscara ou mapa de cultivo tenham sido selecionados, é feita uma verificação no banco de dados local, utilizando o BDC STAC, para consultar se há dados de amostra disponíveis que atendam as necessidades da operação. Caso exista, é apontado o diretório de armazenamento deste dado. Se não, é agendada a geração do dado in situ personalizado de acordo com as propriedades do sítio. Por fim, são gerados os produtos Sen2-Agri e armazenados na instância local para consultas posteriores.

Como resposta, ela retorna um conjunto de informações relacionadas ao sítio definido pelo usuário, que são: o nome; a coleção de imagens *Sentinel* escolhida; as extensões espacial e temporal e uma resposta se a coleção possui imagens para a região na janela de tempo definida; os *tiles* disponíveis para a área selecionada; a quantidade de imagens encontradas; se a coleção possui produtos de reflectância da superfície e NDVI disponíveis para a região; e uma lista com os rótulos de cada uma das imagens a serem baixadas para o sítio. A Listagem 6 apresenta o retorno desta operação.

Tabela 3.5 - Parâmetros da operação *create_site*.

Parâmetro	Descrição	Requerido
site_info	<p>Dicionário, informado pelo usuário, que contém as propriedades do sítio definido:</p> <ul style="list-style-type: none"> • name: nome do sítio; • collection: nome da coleção de imagens <i>Sentinel</i> desejada; • spatialExt: uma <i>string</i> que contém a extensão espacial da região; • temporalExt: uma <i>string</i> que contém a extensão temporal da região; • tiles: uma lista de <i>string</i> que permite ao usuário gerar os produtos apenas para os <i>tiles</i> escolhidos. Este parâmetro é opcional; • products: uma lista de <i>string</i> que permite ao usuário indicar quais produtos ele deseja obter para este sítio. Este parâmetro é opcional 	Sim
collection_name	Nome da coleção WLTS utilizada para gerar dados de amostra para o sítio.	Não
classes	Lista de <i>string</i> que contém as classes escolhidas pelo usuário para serem associadas às classes de cultivo.	Não

3.8 Operações de visualização

As operações *get_vegetation_status*, *get_crop_mask* e *get_crop_map* recuperam, respectivamente, uma coleção estruturada de acordo com o protocolo STAC para

```

1 {
2     "Site": "string",
3     "Coleção": "string",
4     "Intervalo temporal": {
5         "Janela Temporal fornecida": "string",
6         "A coleção estende ao longo da janela de tempo do site?": "string"
7     },
8     "Extensão espacial": {
9         "Área fornecida": "string",
10        "A coleção intercepta a extensão espacial do site?": "string"
11    } ,
12    "Tiles": "list",
13    "Quantidade de imagens disponíveis": "string",
14    "A coleção tem produto de SR disponível?": "string",
15    "A coleção tem produto NDVI disponível?": "string",
16    "Arquivos a serem baixados": ["string"]
17 }

```

Listing 6 - Definição de retorno da operação *create_site*.

os produtos de mapa de estado de vegetação utilizando os índices LAI, NDVI e Fenológico; de máscara dinâmica de cultivo; e de mapa de tipo de cultura. A tabela 3.6 apresenta os parâmetros suportados por essas operações.

Tabela 3.6 - Parâmetros das operações de visualização.

Parâmetro	Descrição	Requerido
time_interval	Intervalo temporal para consultar os produtos desejados.	Sim
indice	Lista de quais índices serão pesquisados. Pode receber combinações de valores entre LAI, NDVI e NDVI Metrics.	Não
site_extent	Extensão espacial para delimitar a busca dos produtos.	Sim

Em relação aos parâmetros das operações de visualização, na tabela 3.6 é especificado que o *indice* não é requerido. Isso quer dizer que este parâmetro só precisa ser fornecido pelo usuário caso ele utilize a operação *get_vegetation_status*. Para as demais operações de visualização este parâmetro não é indicado.

Quando executada, essa operação utiliza o BDC STAC para acessar o banco de dados local e consultar os produtos disponíveis para a busca definida pelo usuário.

Como resposta, ela retorna um objeto da classe *ItemCollection*. Essa classe é disponibilizada pelo PySTAC. Este objeto pode conter uma lista de produtos de mapa de estado da vegetação, máscara dinâmica de cultivo ou mapa de cultura. Este retorno varia de acordo com a operação de visualização utilizada. A Figura 3.4 apresenta o retorno desta operação.

3.9 Implementação

O desenvolvimento da arquitetura de integração considerou um conjunto de tecnologias dentro de um cenário tecnológico de serviços *Web*. A linguagem *Python* foi adotada como linguagem de programação primária de codificação. Considerando o cenário *Web*, foram utilizados três frameworks: *Flask*, *Flask SQLAlchemy* e *Flask RESTful*.

A instalação do Sen2-Agri é restrita apenas ao Sistema Operacional (SO) *CentOS* 7. Devido a isso, foi preparado um ambiente em uma máquina virtual para instalar este pacote de *software*.

O *GitHub* foi utilizado para disponibilizar os pacotes e códigos desenvolvidos neste trabalho. Ele é uma das principais ferramentas para trabalhar com versionamento e repositório de códigos, além de permitir a integração de diversas tecnologias.

O *Celery* e o *RabbitMQ* foram utilizados para agendar e gerenciar as tarefas demandadas pela aplicação. A instalação do *RabbitMQ* foi realizada utilizando uma imagem *Docker*, disponibilizando um container para executar o aplicativo de mensageria. A utilização destas ferramentas permite que o fluxo de funcionamento da aplicação não trave esperando algum processamento ou algo do tipo. Assim, quando o usuário realizar uma requisição que demande tempo para ser executada, o *Celery* coloca-a na fila de tarefas e o *RabbitMQ* vai gerenciando o processamento dessas tarefas.

Para integrar a plataforma BDC com o pacote de software Sen2-Agri, as aplicações BDC Catalog, WLTS e BDC STAC foram utilizadas para, respectivamente, criar uma instância local do banco de dados esquematizado para as necessidades do *Brazil Data Cube*, gerar dados de amostra personalizados para o sítio e consultar os itens obtidos nos processamentos.

O *Jupyter Notebook* foi adotado como ferramenta para apresentação de um modelo de documentação onde são apresentados exemplos de códigos das ferramentas desenvolvidas.

Figura 3.4 - Definição de retorno das operações de visualização.



Fonte: Produção do autor.

4 RESULTADOS E ESTUDO DE CASO

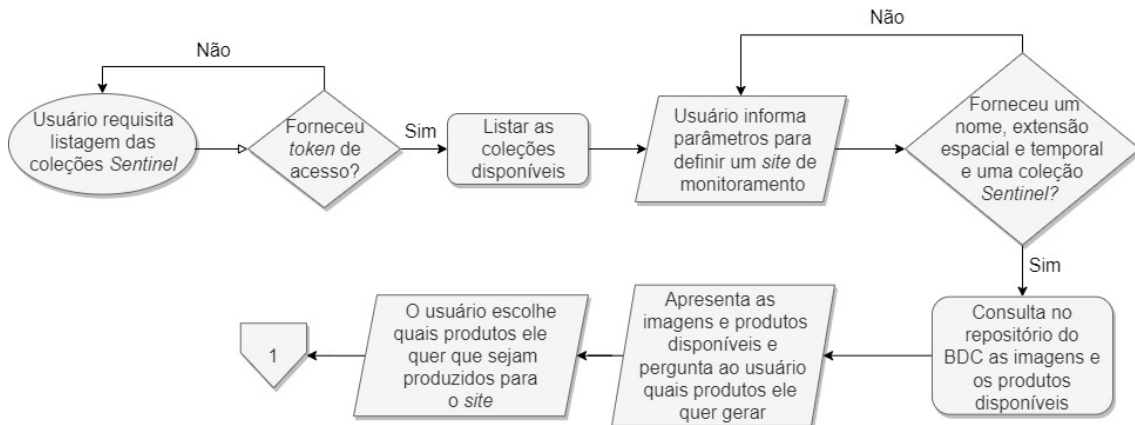
A integração do pacote de *software* Sen2-Agri com a plataforma *Brazil Data Cube* visa aproveitar o melhor de ambos. Para isso, foi implementado um conjunto de métodos que suportam um serviço de extração de produtos do tipo Sen-2Agri de maneira integrada à plataforma BDC. Esses métodos foram implementados em um pacote de funções escritas na linguagem *Python*. O pacote permite ao usuário definir uma região de monitoramento, obter produtos de reflectância da superfície, estado vegetativo das culturas representados pelos índices LAI, NDVI e *NDVI Metrics*, máscaras dinâmicas de cultivo e mapas de tipo de cultura para essa região. As seções a seguir apresentam fluxogramas que expressam o fluxo de funcionamento do pacote desenvolvido e um estudo de caso realizado para testar a implementação como um todo.

4.1 Fluxograma de uso do pacote

O pacote, escrito em Python, foi desenvolvido para implementar a arquitetura proposta, fornecendo, assim, um serviço web que disponibiliza um conjunto de operações que permite a geração dos produtos de monitoramento agrícola. Seu fluxo de funcionamento tem início com uma consulta feita pelo usuário para listar quais são as coleções *Sentinel* disponíveis no repositório de dados do BDC. A partir desta informação, o usuário escolhe uma das coleções para prover os dados de entrada para os processadores Sen2-Agri e verificar se ela disponibiliza produtos NDVI e de reflectância da superfície. Em seguida, o usuário informa os parâmetros de definição da área de monitoramento, o sistema consulta o repositório para obter as imagens e produtos disponíveis para a região definida e apresenta-os ao usuário junto com um conjunto de informações que servem de suporte para que ele escolha quais são os produtos que devem ser processados. Este fluxo é representado na Figura 4.1.

A partir da escolha dos produtos desejados, o sistema verifica se o produto de máscara dinâmica de cultivo ou mapa de tipo de cultura foi selecionado. Caso tenha sido, o sistema questiona ao usuário se ele deseja personalizar os dados de amostra ou não. Essa etapa tem duas respostas possíveis: sim ou não. Se o retorno for afirmativo, o usuário consulta quais os sistemas de classificação disponíveis no repositório que o WLTS acessa, retorna uma lista com os identificadores dessas coleções e pede para o usuário escolher uma. Após o usuário selecionar e informar uma delas, o sistema verifica se as extensões espacial e temporal da coleção cobrem a área de monitoramento. Se cobrirem, ele apresenta uma listagem com as classes disponíveis no sistema de classificação e pede para o usuário criar uma lista com as classes que

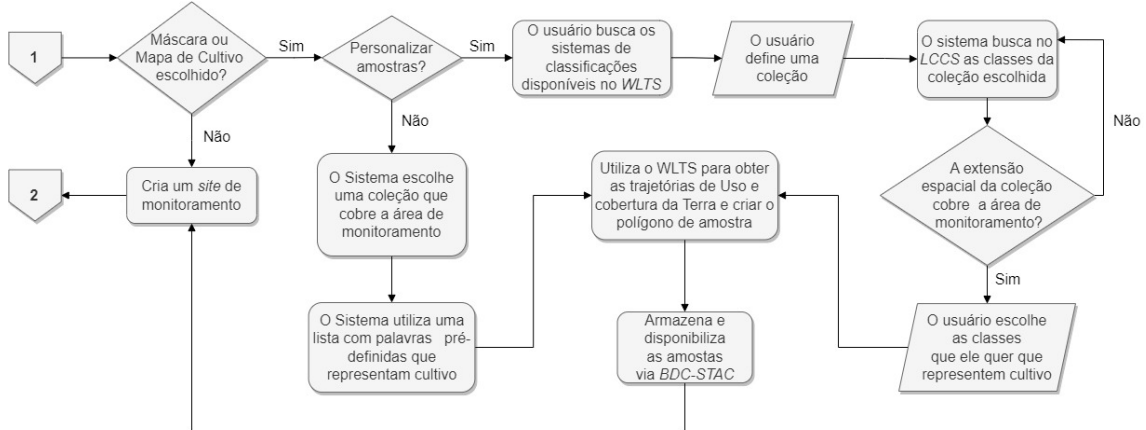
Figura 4.1 - Fluxograma de definição da área de monitoramento.



Fonte: Produção do autor.

ele quer que representem cultivo. Porém, caso não cubram, ele pede para o usuário escolher outra coleção e repete o procedimento de verificação de cobertura da coleção para a área de monitoramento. Agora, se o retorno do usuário for negativo para a personalização dos dados de amostra, o sistema realiza a escolha aleatoriamente de um sistema de classificação que cubra o sítio e utiliza uma lista pré-definida com classes que representam cultivo. Em sequência, o sistema consulta as trajetórias de uso e cobertura da Terra para as coordenadas da área monitorada, filtrando a busca para o ano anterior ao definido na data de início da janela temporal. Essa consulta é feita através do WLTS. Por fim, a partir das trajetórias, é criado o polígono de amostra, que é armazenado em um banco de dados estruturado utilizando o BDC Catalog e disponibilizado via BDC STAC. Este fluxo é representado na Figura 4.2.

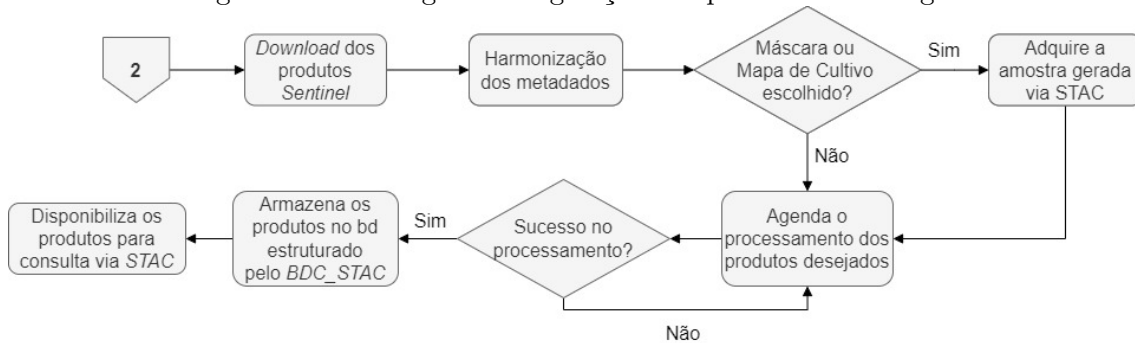
Figura 4.2 - Fluxograma de geração de dados de amostra.



Fonte: Produção do autor.

A criação da área de monitoramento é a próxima etapa do fluxo de funcionamento do pacote implementado. Ela ocorre após a obtenção dos dados de amostra caso o produto de máscara dinâmica de cultivo ou mapas de cultura tenha sido selecionado, ou após a definição do sítio, caso nenhum dos produtos comentados anteriormente tenham sido selecionados. Nesta fase, as imagens e produtos *Sentinel* disponíveis no repositório do *Brazil Data Cube* são armazenados localmente durante o processamento e é feita uma harmonização dos metadados das imagens a fim de alinhar as informações presentes no arquivo com a estrutura do diretório de imagens. Em seguida, o sistema faz uma consulta para verificar se o usuário escolheu gerar ou não produtos de máscara de cultivo ou mapa de cultura. Caso afirmativo, é adquirida a amostra gerada via STAC e é agendado o processamento dos produtos desejados (esta etapa é executada automaticamente caso o usuário não escolha gerar produtos de máscara ou mapa de cultivo). Por fim, o sistema verifica se o processamento foi bem sucedido. Caso tenha sido, os produtos gerados são armazenados no banco de dados e são disponibilizados para a consulta do usuário. Se não, é repetido o agendamento dos processamentos até que tenha êxito na tarefa, seguindo, assim, o fluxo afirmativo dessa condição. Este fluxo é representado na Figura 4.3.

Figura 4.3 - Fluxograma de geração dos produtos Sen2-Agri.

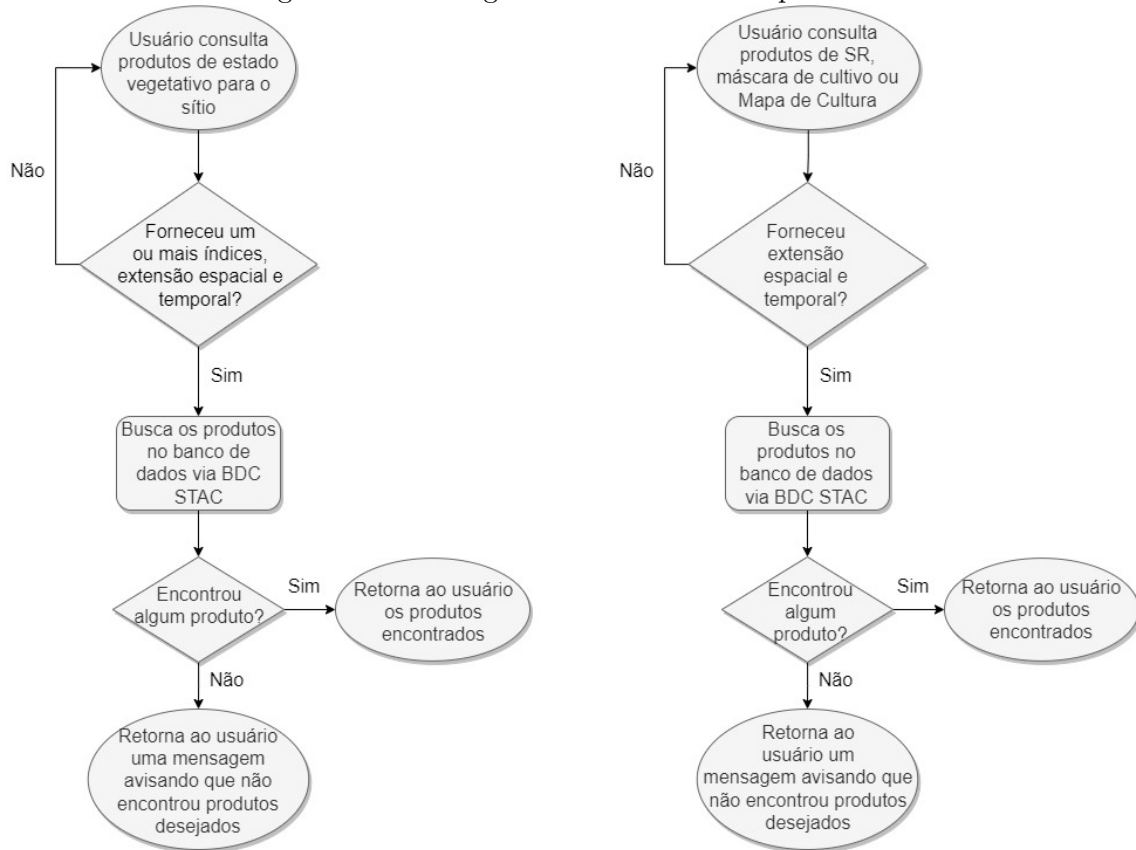


Fonte: Produção do Autor.

A etapa final do fluxo de processamento é a visualização dos produtos obtidos. Nela, o usuário utiliza um dos métodos de consulta e informa uma extensão espacial e temporal para filtrar e obter os produtos que atendam a essas condições. Caso o usuário queira buscar produtos de estado vegetativo das culturas, ele precisa informar também um ou mais índices dentre os disponíveis para consulta. A partir disso, o sistema realiza uma consulta, via STAC, nos metadados dos produtos armazenados no banco de dados e verifica se há produtos disponíveis para a consulta feita. Se afirmativo, retorna os produtos escolhidos para o usuário visualizar. Caso contrário, o usuário

é informado de que não há nenhum produto disponível para a consulta feita. Este fluxo é representado na Figura 4.4.

Figura 4.4 - Fluxograma de consulta dos produtos.



Fonte: Produção do autor.

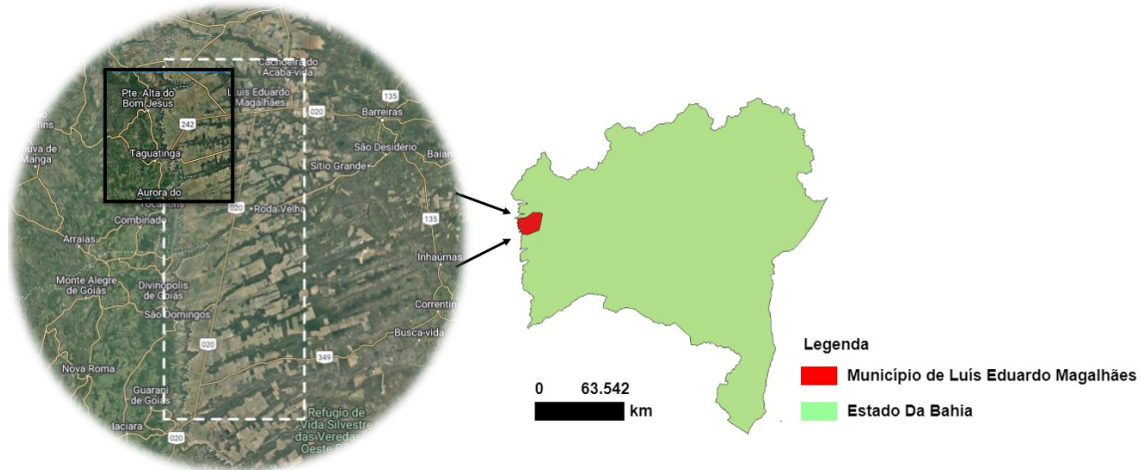
A partir do fluxo de funcionamento adotado ao pacote, descrito na presente seção, um estudo de caso foi realizado para testar a sequência lógica de funcionamento da integração proposta.

4.2 Estudo de caso

Como prova de conceito, delimitamos uma área de monitoramento, ilustrada na Figura 4.5, que é situada no município de Luís Eduardo Magalhães, localizado no estado da Bahia, região nordeste do Brasil. A escolha desta área se deu por ser um polo de agronegócio de seu estado, permitindo, assim, testar os produtos Sen2-Agri e verificar a capacidade do pacote de software em identificar, acompanhar e mapear os cultivos. A janela temporal escolhida possui quatro meses de duração, tendo início em 15 de outubro de 2019 e término em 15 de fevereiro de 2020. Além disso, foi

atrelado a este sítio um identificador único (nome) denominado “LEM-SUMMER”.

Figura 4.5 - Área de monitoramento do estudo de caso.



Fonte: Produção do autor.

Além de definir o nome e as extensões espacial e temporal do sítio, foi definido também a coleção, os produtos desejados e os tiles que serão processados. Para isso, foram utilizadas as operações *list_collections()* e *describe_collection()*, fornecidas pelo cliente desenvolvido no presente trabalho.

Para utilizar as operações citadas acima, o usuário precisa primeiro instanciar um objeto do cliente. Para isso, ele indica um *token* de acesso, que é único para cada usuário e que valida/permite o acesso a plataforma do *Brazil Data Cube*. Este *token* pode ser gerado ou visualizado na “área do usuário” da plataforma *web BDC Explorer*, hospedado no link <<https://brazildatacube.dpi.inpe.br/portal/explore>>. A Figura 4.6 apresenta este procedimento descrito previamente.

A escolha da coleção pode ser definida ao utilizar a operação *list_collections()*. A Listagem 7 apresenta um exemplo de como utilizar esta operação e o retorno que ela tem ao ser executada. Para o presente estudo, a coleção escolhida foi a *Sentinel-2 - MSI - Level-2A (S2_L2A-1)*, que fornece produtos “*monotemporais de refletância da superfície (Bottom-of-Atmosphere) e mapas de Espessura Óptica do Aerossol (AOT), de Vapor d’ Água (WV) e de Classificação de Cena (SCL) através do processamento dos produtos de dados Sentinel-2 Level-1C utilizando o Sen2Cor*” (LOUIS et al., 2019).

Figura 4.6 - Exemplo de como importar a biblioteca cliente através do *Jupyter Notebook*.



Fonte: Produção do autor.

```
1 import cliente
2 service = cliente.Cliente(token='access_token')
3 service.list_collections()
4 {
5     "Collections": [
6         "S2_L2A-1"
7     ]
8 }
```

Listing 7 - Listar as coleções usando o cliente desenvolvido no presente trabalho.

Definida a coleção de imagens de entrada, o passo seguinte foi criar uma lista com os produtos “Sen2-Agri” escolhidos para o sítio. Para isso, utilizou-se a função `list_collections()` com o parâmetro `result_cols` configurado com o valor `True`. Isso permite que apenas os nomes das coleções referentes aos produtos “Sen2-Agri” sejam exibidas. Assim, para o sítio “LEM-SUMMER”, foram escolhidos os produtos `S2AGRI_SLAIMONO`, `S2AGRI_SPHENO`, `S2AGRI_CM` e `S2AGRI_CT`. A listagem 8 demonstra a utilização da operação `list_collections()` para obter a lista de produtos desejados para o sítio.

Dentre os produtos escolhidos para o sítio “LEM-SUMMER” estão a máscara e o mapa de cultivo. Estes produtos necessitam dos dados de amostras para serem processados. Para gerar esses dados de amostras, a primeira etapa foi consultar quais são as coleções disponibilizadas pelo WLTS, através da operação `get_wlts_collections()` (ilustrado na Figura 4.7). Como resultado, uma lista com estas coleções foi obtida.

```

1 import cliente
2 service = cliente.Cliente(token='access_token')
3 service.list_collections(result_cols=True)
4 {
5     "Result Collections": [
6         "S2AGRI_CM-1",
7         "S2AGRI_CT-1",
8         "S2AGRI_SLAIMONO-1",
9         "S2AGRI_SPHENO-1"
10    ]
11 }

```

Listing 8 - Listar as coleções dos produtos disponíveis usando o cliente desenvolvido no presente trabalho.

A partir desta lista resultante, a coleção “mapbiomas-v6” foi escolhida para verificar se a sua extensão espacial e temporal cobre a extensão do sítio. Este processo é feito através da operação `verify_collection_extent()` (representado na Figura 4.8). A resposta resultante da execução desta operação foi positiva, permitindo que a coleção fosse utilizada para gerar os dados de amostra para o sítio. Por fim, o último passo é definir quais as classes LULC representam cultivo. A operação `get_wlts_collection_classes()` foi utilizada para obter a classificação da coleção “mapbiomas-v6” (este processo está representado na Figura 4.9). Com isso, uma lista com todas as classes foi retornada, permitindo a criação de uma lista com as classes que o usuário deseja que representem cultivo. A Listagem 9 apresenta as classes de cultivo configuradas para o presente estudo de caso.

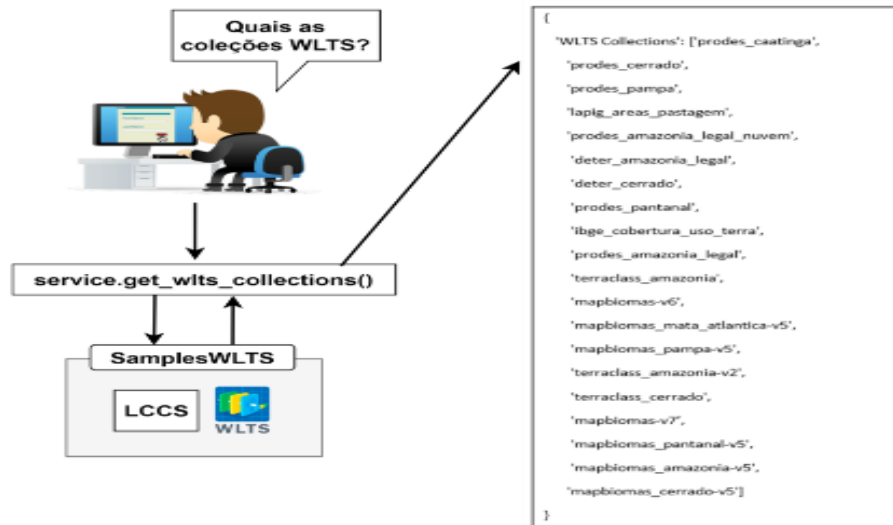
```

1 crop_list = ['Arroz (beta)', 'Outras Lavouras Temporárias', 'Café (beta)',
2             'Citrus (beta)', 'Outras Lavouras Perenes', 'Agropecuária',
3             'Agricultura', 'Pastagem', 'Lavoura Temporária', 'Soja', 'Cana',
4             'Mosaico de Agricultura e Pastagem', 'Lavoura Perene']

```

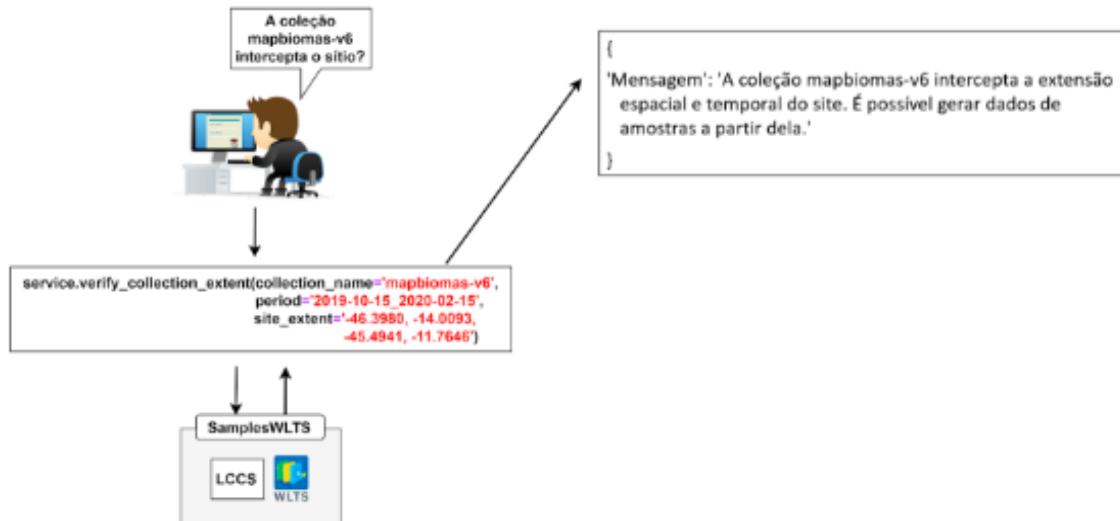
Listing 9 - Lista de classes que representam cultivo.

Figura 4.7 - Utilização da operação `get_wlts_collections` para obter as coleções WLTS.



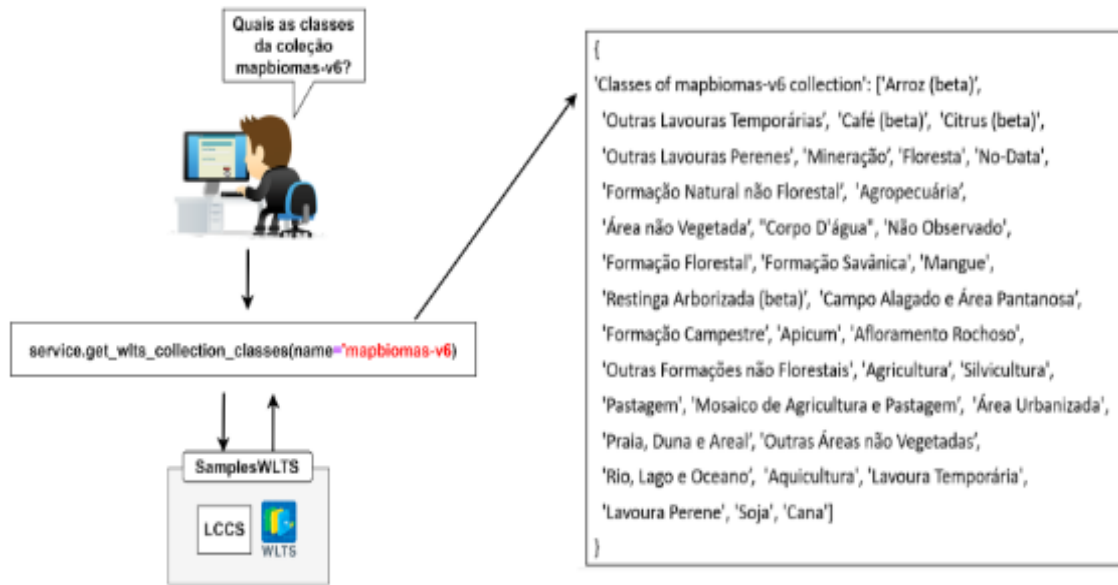
Fonte: Produção do autor.

Figura 4.8 - Verificando se a coleção `mapbiomas-v6` intercepta as extensões do sítio.



Fonte: Produção do autor.

Figura 4.9 - Obtendo as classes da coleção mapbiomas-v6.



Fonte: Produção do autor.

Após a escolha da classe e da lista de coleções para gerar os dados de amostra, o parâmetro *tiles* foi configurado com uma lista que indica o *tile* 23LLG para o sítio do presente estudo de caso. Para verificar os *tiles* que a coleção de entrada disponibiliza, é utilizada a operação `describe_collection()`. Neste caso, o parâmetro `name` desta operação foi configurado como “S2_L2A-1”. A partir disso, obteve-se os metadados da coleção e uma das propriedades informadas era uma lista com os *tiles* disponíveis. A listagem 10 demonstra a utilização da operação `describe_collection()` para obter os metadados e escolher o *tile*.

```

1 import cliente
2 service = cliente.Cliente(token='access_token')
3 service.describe_collection(name= "S2_L2A-1")
4 {
5     "Collection S2_L2A-1": {
6         "collection properties...",
7         "bdc:tiles": [
8             "Demais tiles...",
9             "23LLG",
10        ]
11    }

```

Listing 10 - Consultar os metadados de uma coleção usando o cliente desenvolvido no presente trabalho.

Portanto, para o presente estudo de caso, os parâmetros definidos para a operação `create_site` foram:

- **site_info:** um dicionário que contém o nome, as extensões espacial e temporal, a coleção de imagens de entrada e uma lista dos produtos e dos *tiles* desejados;
- **collection_name:** uma string que contém o nome da coleção WLTS escolhida para gerar os dados de amostra;
- **classes:** um *array* de *string* que contém as classes LULC escolhidas para representarem cultivo. Este *array* está representado na Listagem 10;

A partir destes parâmetros, o sítio “LEM_SUMMER” foi delimitado e criado. Este processo de criação é distribuído em tarefas, gerenciado pelo *Celery* e *RabbitMQ*. Estas tarefas são criadas para obter as imagens de entrada, os dados de amostra e os produtos “Sen2-Agri”. Cada uma delas é armazenada em uma fila para ser executada posteriormente. Isto permite que o usuário possa utilizar normalmente outras operações sem ter que esperar o processamento finalizar. Com as tarefas enfileiradas, o sistema retorna ao usuário informações sobre o sítio definido. A Listagem 11 apresenta esse conjunto de informações retornado.

A execução das tarefas é realizada em cadeia pelo sistema. Primeiro todas as imagens de entrada encontradas para o sítio são baixadas. Em seguida, é gerado o dado

```

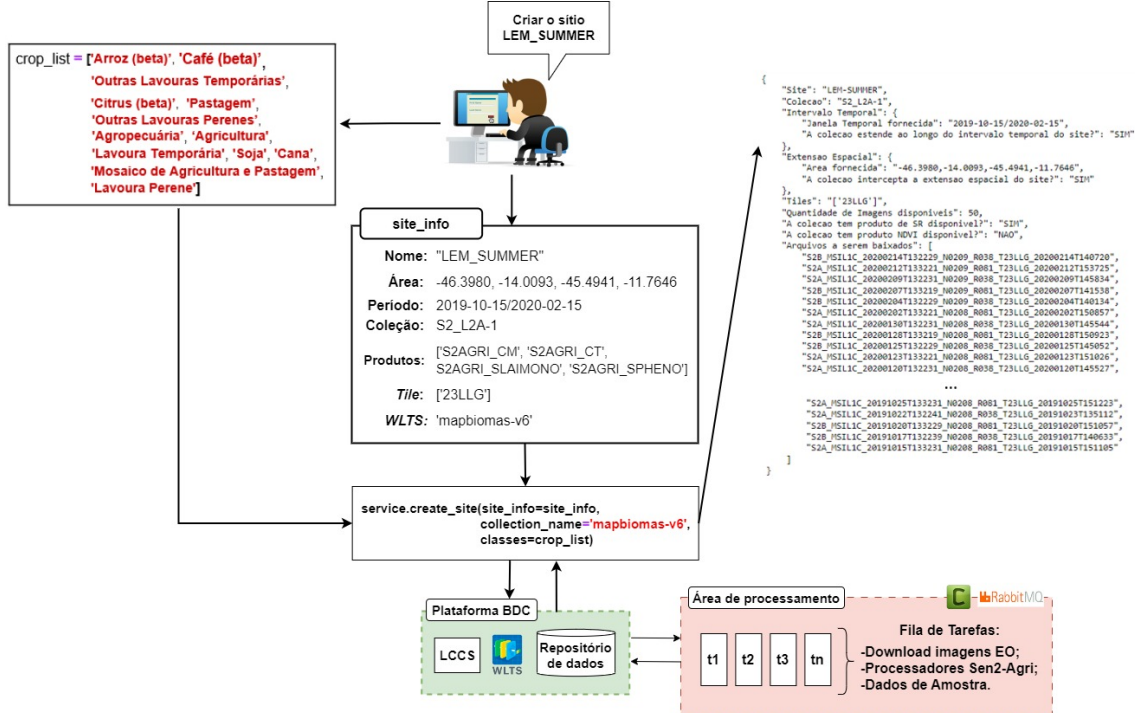
1 {
2   "Site": "LEM-SUMMER",
3   "Colecao": "S2_L2A-1",
4   "Intervalo Temporal": {
5     "Janela Temporal fornecida": "2019-10-15/2020-02-15",
6     "A colecao estende ao longo do intervalo temporal do site?": "SIM"
7   },
8   "Extensao Espacial": {
9     "Area fornecida": "-46.3980,-14.0093,-45.4941,-11.7646",
10    "A colecao intercepta a extensao espacial do site?": "SIM"
11  },
12  "Tiles": "['23LLG']",
13  "Quantidade de Imagens disponiveis": 50,
14  "A colecao tem produto de SR disponivel?": "SIM",
15  "A colecao tem produto NDVI disponivel?": "NAO",
16  "Arquivos a serem baixados": [
17    "S2B_MSIL1C_20200214T132229_N0209_R038_T23LLG_20200214T140720",
18    "S2A_MSIL1C_20200212T133221_N0209_R081_T23LLG_20200212T153725",
19    "S2A_MSIL1C_20200209T132231_N0209_R038_T23LLG_20200209T145834",
20    "Demais imagens...",
21    "S2B_MSIL1C_20191020T133229_N0208_R081_T23LLG_20191020T151057",
22    "S2B_MSIL1C_20191017T132239_N0208_R038_T23LLG_20191017T140633",
23    "S2A_MSIL1C_20191015T133231_N0208_R081_T23LLG_20191015T151105"
24  ]

```

Listing 11 - Retorno da operação create_site().

de amostra personalizado para a região. Caso os produtos que necessitam destes dados não sejam escolhidos, esta etapa não é executada. Por fim, são executados os processadores Sen2-Agri na seguinte ordem (caso todos os produtos tenham sido escolhidos): inicia com a obtenção dos mapas de cultivo utilizando os índices LAI, NDVI e fenológico, em seguida é gerado as máscaras de cultivo, e, por fim, os mapas de cultura. A partir disso, nota-se que há algumas dependências neste fluxo de processamento, que são elas: os processadores são executados após o sistema baixar as imagens localmente; o processador de mapa de cultura executa após a obtenção das máscara de cultivo; e, ambos os processadores são executados após a criação de um dado de amostra para o sítio. O usuário fica livre para escolher os produtos que ele deseja obter para o sítio. Para o presente estudo de caso, a execução das tarefas ocorreu da maneira descrita acima, selecionando todos os produtos para geração. Este processo está representado na Figura 4.10.

Figura 4.10 - Criação do sítio LEM_SUMMER.



Fonte: Produção do autor.

Visto que os produtos de máscara e mapa de cultivo foram selecionados para obtenção, o dado de amostra foi criado para o sítio. Este dado foi obtido utilizando a classificação da coleção mapbiomas-v6 e tendo como referência o ano de 2018. Uma análise foi feita a partir da amostra para levantar qual a quantidade de *pixels* rotulados para cada uma das classes LULC presentes. A Figura 4.11 apresenta uma tabela desse levantamento. Ao analisar o resultado, nota-se que a classe predominante é a Soja, com 1985 pixels, seguida por Formação Savânica, Outras Lavouras Temporárias, Formação Campestre e Mosaico de Agricultura e Pastagem, rotulados, respectivamente, com 1657, 1004, 860 e 395 *pixels*. Isto evidencia a escolha da área para realizar o estudo de caso, uma vez que das cinco classes mais predominantes na amostra, três estão relacionadas a classes de cultivo.

Figura 4.11 - Tabela com o total de *pixels* rotulados para cada uma das classes LULC oriundo dos dados de amostra do sítio.

```
gpd.GeoDataFrame(samples.groupby(by=['LC']).count().sort_values(by=['ID'], ascending=False)['ID'])
```

LC	ID
Soja	1985
Formação Savânica	1657
Outras Lavouras Temporárias	1004
Formação Campestre	860
Mosaico de Agricultura e Pastagem	395
Formação Florestal	164
Pastagem	157
Campo Alagado e Área Pantanosa	133
Café (beta)	17
Outras Áreas não Vegetadas	12
Área Urbanizada	11
Rio, Lago e Oceano	3
Outras Lavouras Perenes	1
Silvicultura	1

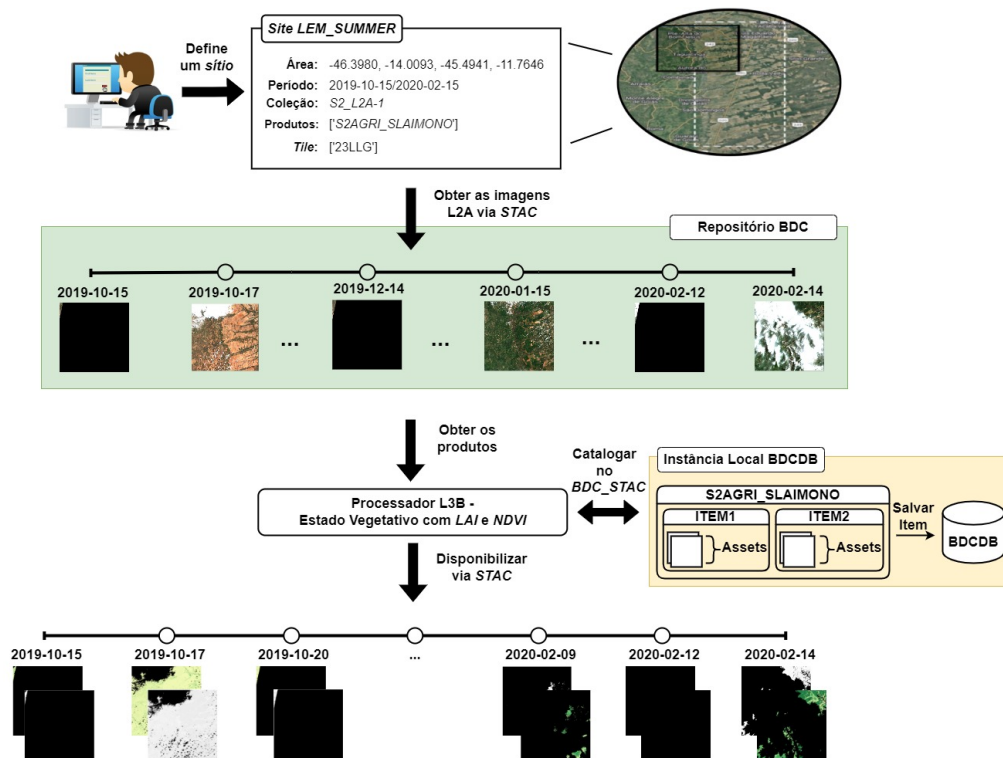
Fonte: Produção do autor.

Ao final do processamento das tarefas geradas pela operação `create_site()`, os produtos “Sen2-Agri” foram obtidos.

A Figura 4.12 apresenta o mecanismo de obtenção dos mapas de estado vegetativo utilizando os índices LAI e NDVI para o sítio. Para cada uma das imagens de entrada foi feito o processamento utilizando estes índices. Como resultado, obteve-se mapas de estado vegetativo monodatas. Os metadados destes produtos foram armazenados no banco de dados estruturado pelo BDC-Catalog e podem ser acessados ao utilizar a operação `get_vegetation_status()` com o parâmetro *indice* configurado como ['LAI', 'NDVI'].

A Figura 4.13 apresenta o mecanismo de obtenção dos mapas de estado vegetativo utilizando os índices fenológicos para o sítio. Este produto foi obtido duas vezes na estação: uma na metade (14/12/2019) e outra ao final (14/02/2020). Como resultado, obteve-se dois mapas de estado vegetativo gerado a partir do índice fenológico. Os metadados destes produtos foram armazenados no banco de dados estruturado pelo BDC-Catalog e podem ser acessados ao utilizar a operação `get_vegetation_status()` com o parâmetro *indice* configurado como ['NDVI Metrics'].

Figura 4.12 - Mapas do Estado vegetativo com os índices LAI e NDVI obtidos para o sítio LEM_SUMMER.

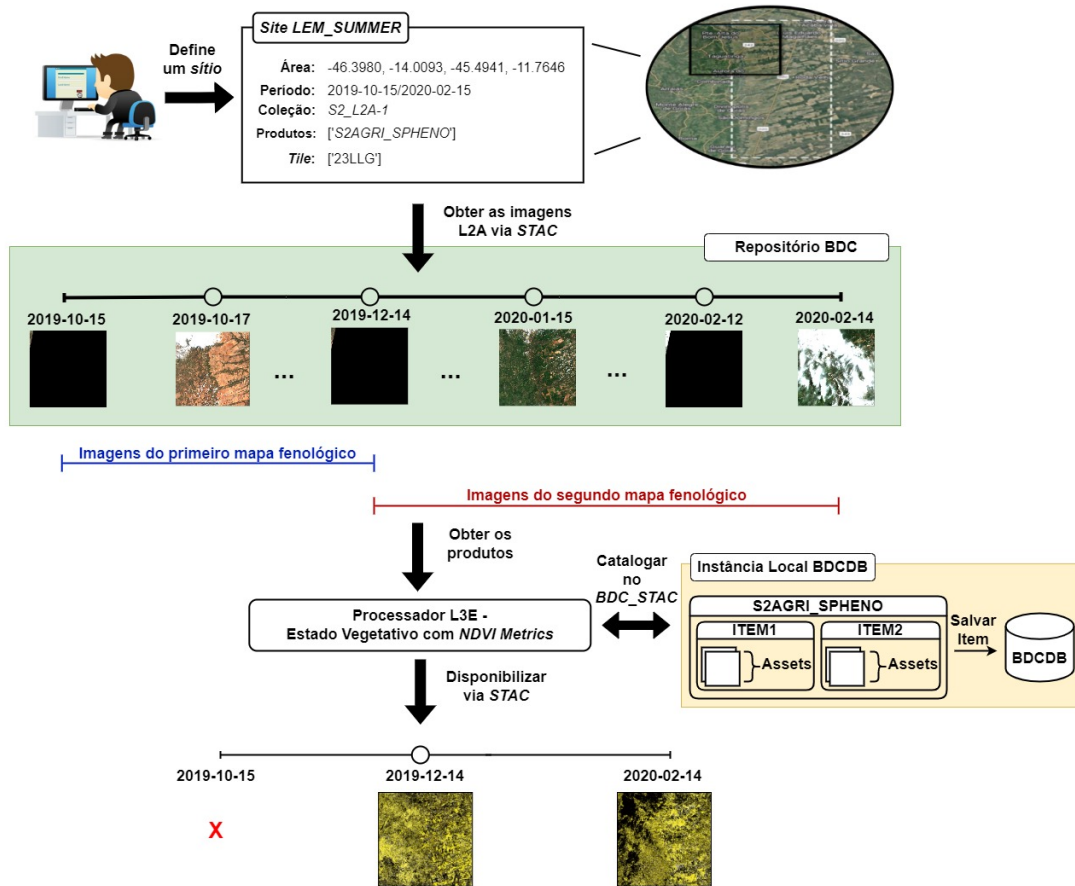


Fonte: Produção do autor.

A Figura 4.14 apresenta o mecanismo de obtenção das máscaras dinâmicas de cultivo para o sítio. Este produto foi obtido três vezes na estação: uma na metade (14/12/2019), outra um mês após a metade (14/01/2020) e a última ao final (14/02/2020). Como resultado, obteve-se três máscaras de cultivo. Os metadados destes produtos foram armazenados no banco de dados estruturado pelo BDC-Catalog e podem ser acessados ao utilizar a operação `get_crop_mask()`.

A Figura 4.15 apresenta o mecanismo de obtenção dos mapas de tipo de cultura para o sítio. Este produto foi obtido duas vezes na estação: uma na metade (14/12/2019) e outra ao final (14/02/2020). Como resultado, obteve-se dois mapas de cultura. Os metadados destes produtos foram armazenados no banco de dados estruturado pelo BDC-Catalog e podem ser acessados ao utilizar a operação `get_crop_map()`.

Figura 4.13 - Mapas do Estado vegetativo com índice fenológico obtidos para o sítio LEM_SUMMER.

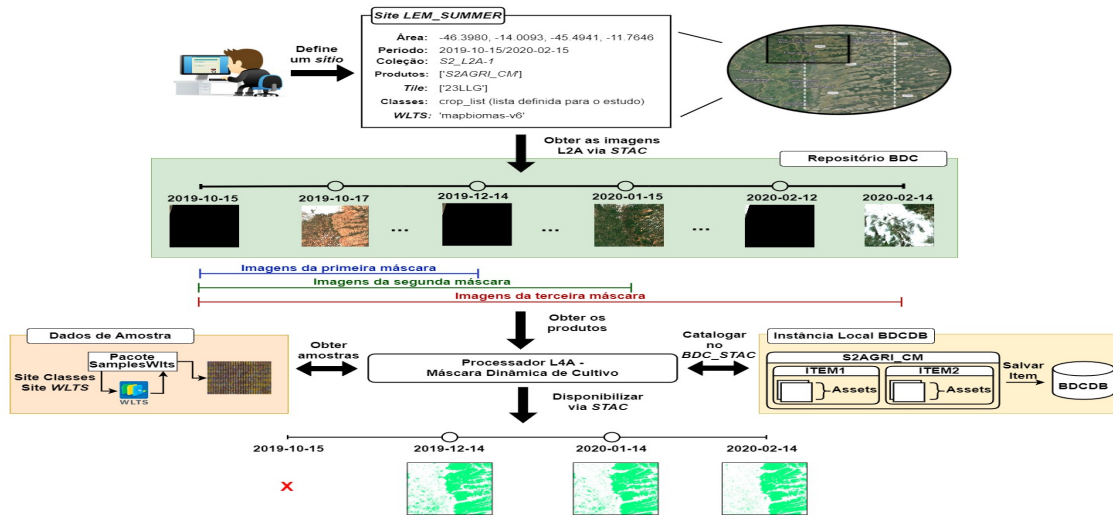


Fonte: Produção do autor.

Obtidos os mapas de cultura para o sítio “LEM_SUMMER”, algumas manipulações foram feitas sobre os produtos para analisar o resultado obtido. Esta análise utilizou três *assets* dos dois *Items* disponíveis: o **CT_CT**, que é um dado raster em formato “.tif” que contém o mapa de cultivo; o **MTD**, que é um arquivo “.xml” que contém os metadados do produto gerado; e o **LUT**, que é um arquivo “.txt” que armazena os valores e cores de representação de cada uma das classes de uso e cobertura da terra presente no mapa.

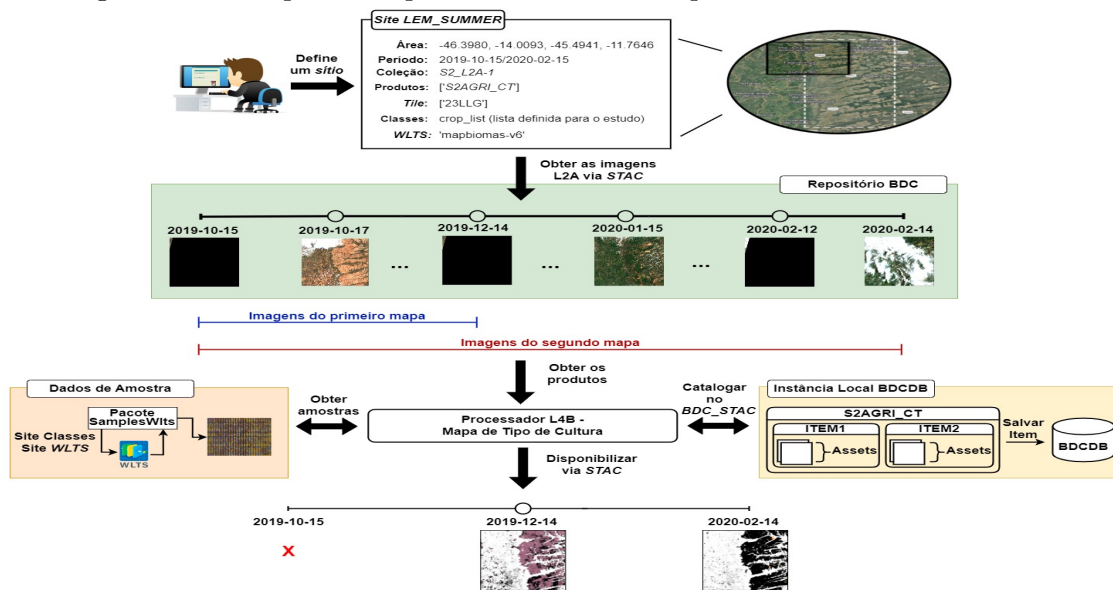
A Figura 4.16 apresenta o mapa de cultura da primeira metade da estação de cultivo. Nela é possível identificar, visualmente, a classificação de uso e cobertura da terra desempenhada pelo classificador “Sen2-Agri”. A taxa de acurácia obtida pelo classificador para este mapa foi de aproximadamente 67%. Além do mapa, uma agre-

Figura 4.14 - Máscaras Dinâmicas de Cultivo obtidas para o sítio LEM_SUMMER.



Fonte: Produção do autor.

Figura 4.15 - Mapas do Tipo de Cultivo obtidas para o sítio LEM_SUMMER.



Fonte: Produção do autor.

gação por classes foi feita para verificar a quantidade de *pixels* em que cada uma delas estava associada. Este processo é ilustrado na Figura 4.17. Ao analisar o mapa e a tabela de agregação, nota-se que a classe de cultivo de predominância foi a Soja, seguida por Outras Lavouras Temporárias e Pastagem com 28.841.204, 10.206.977 e 5.431.460 *pixels* rotulados, respectivamente.

Figura 4.16 - Mapa de Cultivo obtido na metade da estação de cultivo para o sítio “LEM_-SUMMER”.



Fonte: Produção do autor.

A Figura 4.18 apresenta o mapa de cultura da segunda metade da estação de cultivo. Nela é possível identificar, visualmente, a classificação de uso e cobertura da terra desempenhada pelo classificador “Sen2-Agri”. A taxa de acurácia obtida pelo classificador para este mapa foi de aproximadamente 72%. Além do mapa, uma agregação por classes foi feita para verificar a quantidade de *pixels* em que cada uma delas estava associada. Este processo é ilustrado na Figura 4.19. Ao analisar o mapa e a tabela de agregação, nota-se que a classe de cultivo de predominância

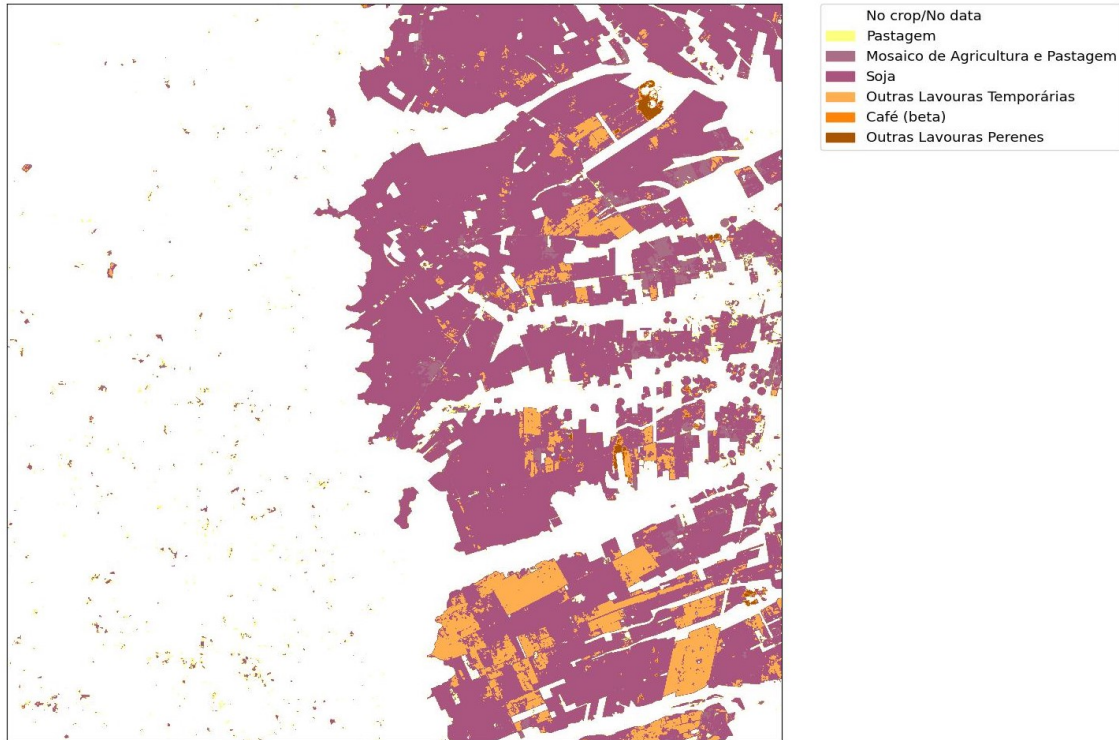
Figura 4.17 - Contagem das classes rotuladas pelo classificador para o mapa da primeira metade da estação do sítio “LEM_SUMMER”.

CODE	COUNT	LC
0	71363496	No crop/No data
5	28841204	Soja
3	10206977	Outras Lavouras Temporárias
4	5431460	Pastagem
2	4093547	Mosaico de Agricultura e Pastagem
1	623716	Café (beta)

Fonte: Produção do autor.

foi a Soja, seguida por Outras Lavouras Temporárias e Mosaico de Agricultura e Pastagem com 30.233.461, 8.053.181 e 3.391.619 *pixels* rotulados, respectivamente.

Figura 4.18 - Mapa de Cultivo obtido ao final da estação de cultivo para o sítio “LEM_SUMMER”.



Fonte: Produção do autor.

Figura 4.19 - Contagem das classes rotuladas pelo classificador para o mapa da segunda metade da estação do sítio “LEM_SUMMER”.

CODE	COUNT	LC
0	76577520	No crop/No data
3	30233461	Soja
4	8053181	Outras Lavouras Temporárias
2	3391619	Mosaico de Agricultura e Pastagem
1	1980883	Pastagem
5	323736	Café (beta)

Fonte: Produção do autor.

Ao analisar a taxa de acurácia, nota-se que houve um aumento de 5% entre o primeiro e o segundo mapa. Isso acontece devido a uma maior quantidade de dados de entrada para treinar e alimentar o classificador, uma vez que o primeiro mapa utiliza imagens referentes aos dois primeiros meses da estação de cultivo definida para o sítio e o segundo utiliza as imagens para todo o período. Portanto, quão maior for o volume dos dados de entrada para o classificador, maior será a taxa de acurácia, já que o algoritmo dispõe de mais imagens para aprender os padrões e gerar máscaras e mapas de cultivos mais precisos.

5 CONCLUSÕES

Neste trabalho foi desenvolvido um serviço que integra a plataforma *Brazil Data Cube* ao pacote de *software* Sen2-Agri para obter produtos de auxílio ao monitoramento agrícola. Esta integração permite otimizar o fluxo de funcionamento do Sen2-Agri, uma vez que dados e produtos necessários na etapa de processamento encontram-se disponíveis e prontos para uso no BDC. O serviço fornece uma semântica de alto nível, mais próximo do domínio de aplicação, e se mostra apropriado para a obtenção de produtos que auxiliam no monitoramento de culturas.

Este trabalho apresentou os desafios para manipular grandes volumes de dados, vantagens e desvantagens da utilização do Sen2-Agri e uma proposta de arquitetura para otimizá-lo. Nesta arquitetura, os serviços, bibliotecas, dados e a infraestrutura do *Brazil Data Cube* são utilizados para suprir as necessidades (dados de entrada e amostra) que os processadores Sen2-Agri possuem para obter e disponibilizar os produtos. Assim, essa proposta de arquitetura faz uso do melhor dos dois ambientes, uma vez que o *Brazil Data Cube* já cuida do *download* das imagens e de alguns produtos, porém, é um ambiente genérico em termos das aplicações sobre esses dados; e o pacote Sen2-Agri já tem uma maturidade tecnológica e de uso específico na aplicação de desenvolvimento de monitoramento de cultivos. Além disso, um serviço foi implementado seguindo a arquitetura proposta para fornecer operações aos usuários, permitindo, assim, obter e visualizar os produtos Sen2-Agri desejados.

O trabalho demonstrou a utilização da integração através de um estudo de caso realizado sobre a região de Luís Eduardo Magalhães utilizando a coleção de imagens S2_L2A-1. A arquitetura proposta mostrou-se capaz de gerar e disponibilizar os produtos de monitoramento agrícola para um sítio definido pelo usuário, otimizando os meios de acesso às imagens e aos dados de amostra feitos comumente pelo Sen2-Agri.

Por fim, em relação aos trabalhos futuros, pode-se levantar algumas melhorias e possibilidades de outros desenvolvimentos. A coleção de imagens utilizada como dado de entrada, na versão atual, é restrita à Sentinel S2_L2A-1. Este quesito pode ser tratado futuramente, adotando um protocolo padrão de organização dos diretórios das imagens de coleções distintas e obter arquivos de metadados que apontem essa estrutura adotada para que os processadores Sen2-Agri consigam identificar a localização dos arquivos necessários para o processamento.

REFERÊNCIAS BIBLIOGRÁFICAS

BONTEMPS, S.; ARIAS, M.; CARA, C.; DEDIEU, G.; GUZZONATO, E.; HAGOLLE, O.; INGLADA, J.; MATTON, N.; MORIN, D.; POPESCU, R.; RABAUTE, T.; SAVINAUD, M.; SEPULCRE, G.; VALERO, S.; AHMAD, I.; BÉGUÉ, A.; WU, B.; ABELLEYRA, D. D.; DIARRA, A.; DUPUY, S.; FRENCH, A.; AKHTAR, I. U. H.; KUSSUL, N.; LEBOURGEOIS, V.; PAGE, M. L.; NEWBY, T.; SAVIN, I.; VERÓN, S. R.; KOETZ, B.; DEFOURNY, P. Building a data set over 12 globally distributed sites to support the development of agriculture monitoring applications with sentinel-2. **Remote Sensing**, v. 7, n. 12, p. 16062–16090, 2015. ISSN 2072-4292. Disponível em: <<<https://www.mdpi.com/2072-4292/7/12/15815>>>. 15, 18, 20, 23

BRAZIL DATA CUBE. **Brazil Data Cube**. 2023. Disponível em: <<<http://brazildatacube.org/sobre-o-brazil-data-cube/>>>. Acesso em: 10 fev. 2023. 25

_____. **Brazil Data Cube Image Metadata catalog: about**. 2023. Disponível em: <<<https://bdc-catalog.readthedocs.io/en/latest/>>>. Acesso em: 27 jan. 2023. 31

BÉGUÉ, A.; ARVOR, D.; BELLON, B.; BETBEDER, J.; ABELLEYRA, D. D.; FERRAZ, R. P. D.; LEBOURGEOIS, V.; LELONG, C.; SIMÕES, M.; VERÓN, S. R. Remote sensing and cropping practices: a review. **Remote Sensing**, v. 10, n. 1, 2018. ISSN 2072-4292. Disponível em: <<<https://www.mdpi.com/2072-4292/10/1/99>>>. 8

CHAVES, M. E. D.; PICOLI, M. C. A.; SANCHES, I. D. Recent applications of landsat 8/oli and sentinel-2/msi for land use and land cover mapping: a systematic review. **Remote Sensing**, v. 12, n. 18, 2020. ISSN 2072-4292. Disponível em: <<<https://www.mdpi.com/2072-4292/12/18/3062>>>. 7

CHUVIECO, E. **Fundamentals of satellite remote sensing: an environmental approach**. [S.l.]: CRC Press, 2016. 5

DEFOURNY, P.; BONTEMPS, S.; BELLEMANS, N.; CARA, C.; DEDIEU, G.; GUZZONATO, E.; HAGOLLE, O.; INGLADA, J.; NICOLA, L.; RABAUTE, T.; SAVINAUD, M.; UDROIU, C.; VALERO, S.; BÉGUÉ, A.; DEJOUX, J.-F.; El Harti, A.; EZZAHAR, J.; KUSSUL, N.; LABBASSI, K.; LEBOURGEOIS, V.; MIAO, Z.; NEWBY, T.; NYAMUGAMA, A.; SALH, N.; SHELESTOV, A.; SIMONNEAUX, V.; TRAORE, P. S.; TRAORE, S. S.; KOETZ, B. Near real-time agriculture monitoring at national scale at parcel resolution: performance assessment of the sen2-agri automated system in various cropping systems around the world. **Remote Sensing of Environment**, v. 221, p. 551–568, 2019. ISSN 0034-4257. Disponível em: <<<https://www.sciencedirect.com/science/article/pii/S0034425718305145>>>. 15, 16, 24

DELGADO, J. A.; SHORT, N. M.; ROBERTS, D. P.; VANDENBERG, B. Big data analysis for sustainable agriculture on a geospatial cloud framework. **Frontiers in Sustainable Food Systems**, v. 3, p. 54, 2019. ISSN 2571-581X. Disponível em: <<<https://www.frontiersin.org/article/10.3389/fsufs.2019.00054>>>. 1

EUROPEAN SPACE AGENCY (ESA). **Crop type map: mapping crop types using an automated approach**. 2016. Disponível em: <<<http://www.esa-sen2agri.org/wp-content/uploads/resources/technical-documents/Sen2-Agri-Product-Crop-Type-Map.pdf>>>. Acesso em: 08 jun. 2022. 23

_____. **Dynamic cropland mask: delineating crop areas in the most efficient way**. 2016. Disponível em: <<<http://www.esa-sen2agri.org/wp-content/uploads/resources/technical-documents/Sen2-Agri-Product-Cropland-Mask.pdf>>>. Acesso em: 08 jun. 2022. 20

_____. **Vegetation status indicators: achieving multi-temporal monitoring of crop status and development**. 2016. Disponível em: <<<http://www.esa-sen2agri.org/wp-content/uploads/resources/technical-documents/Sen2-Agri-Product-Vegetation-Status.pdf>>>. Acesso em: 08 jun. 2022. 19

FERREIRA, K. R.; QUEIROZ, G. R.; VINHAS, L.; MARUJO, R. F.; SIMOES, R. E.; PICOLI, M. C.; CAMARA, G.; CARTAXO, R.; GOMES, V. C.; SANTOS, L. A. Earth observation data cubes for Brazil: requirements, methodology and products. **Remote Sensing**, v. 12, n. 24, p. 4033, 2020. 25, 26, 27

FORMAGGIO, A. R.; SANCHES, I. D. **Sensoriamento remoto em agricultura**. [S.l.]: Oficina de Textos, 2017. 5, 7, 8, 9

GEOSPATIAL big data handling theory and methods: a review and research challenges. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 115, p. 119–133, 2016. 9

GIS GEOGRAPHY. **Sentinel 2 bands and combinations**. 2022. Disponível em: <<[GOMES, V. C.; QUEIROZ, G. R.; FERREIRA, K. R. An overview of platforms for big earth observation data management and analysis. **Remote Sensing**, v. 12, n. 8, p. 1253, 2020. 13](https://gisgeography.com/sentinel-2-bands-combinations/#:~:text=Sentinel-2%20carries%20the%20Multispectral,have%20a%2010-meter%20resolution.>>. Acesso em: 6 mar. 2023. 13</p></div><div data-bbox=)

GÓMEZ, C.; WHITE, J. C.; WULDER, M. A. Optical remotely sensed time series data for land cover classification: a review. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 116, p. 55–72, 2016. ISSN 0924-2716. Disponível em: <<<https://www.sciencedirect.com/science/article/pii/S0924271616000769>>>. 8

- INSITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). **Produção Agrícola Municipal - PAM**. 2021. Disponível em: <<<https://www.ibge.gov.br/estatisticas/economicas/agricultura-e-pecuaria/9117-producao-agricola-municipal-culturas-temporarias-e-permanentes.html>>>. Acesso em: 5 maio 2021. 1
- INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS. **Catálogo de imagens**. 2023. Disponível em: <<<http://www2.dgi.inpe.br/catalogo/explore>>>. Acesso em: 2 fev. 2023. 12
- _____. **Galeria de imagens**. 2023. Disponível em: <<<http://www.dpi.inpe.br/galeria/>>>. Acesso em: 2 fev. 2023. 6
- _____. **Missão Amazonia-1**. 2023. Disponível em: <<http://www.inpe.br/amazonia1/sobre_satelite/>>. Acesso em: 2 fev. 2023. 11
- _____. **Missão CBERS**. 2023. Disponível em: <<<http://www.cbbers.inpe.br/>>>. Acesso em: 2 fev. 2023. 11
- KAMILARIS, A.; KARTAKOULLIS, A.; PRENAFETA-BOLDÚ, F. X. A review on the practice of big data analysis in agriculture. **Computers and Electronics in Agriculture**, v. 143, p. 23–37, 2017. ISSN 0168-1699. Disponível em: <<<https://www.sciencedirect.com/science/article/pii/S0168169917301230>>>. 2
- LOUIS, J.; PFLUG, B.; MAIN-KNORN, M.; DEBAECKER, V.; MUELLER-WILM, U.; IANNONE, R. Q.; CADAU, E. G.; BOCCIA, V.; GASCON, F. **Sentinel-2 global surface reflectance leve 2a product generated with Sen2Cor**. [S.l.: s.n.], 2019. 8522-8525 p. ISBN 9781538691540. 53
- MILEV, P. Integration of software solutions via an intermediary web service. **Trakia Journal of Sciences**, v. 17, p. 181–185, 2019. Disponível em: <<<http://tru.uni-sz.bg/tsj/Volume%2017,%202019,%20Supplement%201,%20Series%20Social%20Sciences/2/za%20pe4at/30.pdf>>>. 33
- MUERTH, M.; MIGDALL, S.; HODRIUS, M.; NIGGEMANN, F.; HOLZAPFEL, M.; BACH, H.; GILLIAMS, S.; ROEY, T. van; CUOMO, A.; HARWOOD, P. Food security tep - supporting sustainable intensification of food production from space. **IOP Conference Series: Earth and Environmental Science**, v. 509, p. 3, 2020. ISSN 012038. Disponível em: <<<https://iopscience.iop.org/article/10.1088/1755-1315/509/1/012038>>>. 14
- NAÇÕES UNIDAS BRASIL. **Objetivo de desenvolvimento sustentável 2 fome zero e agricultura sustentável**. 2021. Disponível em: <<<https://brasil.un.org/pt-br/sdgs/2>>>. Acesso em: 5 maio 2021. 1
- ROMEO, A.; PINTO, S.; LOEKKEN, S.; MARIN, A. Cloud based earth observation data exploitation platforms. In: AGU FALL MEETING ABSTRACTS, 2017. **Proceedings...** [S.l.], 2017. p. IN21F–03. 14

- TERRACLASS. **Portal WebGis - Bioma Cerrado**. 2023. Disponível em: <<<https://www.terraclass.gov.br>>>. Acesso em: 2 fev. 2023. 10
- UDROIU, C.; NICOLA, L.; CARA, C.; SAVINAUD, M.; BELLEMANS, N.; MALCORPS, P.; BONTEMPS, S.; DEFOURNY, P. **Sentinel-2 agriculture software user manual**. 2020. 146 p. Disponível em: <<<http://www.esa-sen2agri.org/wp-content/uploads/resources/technical-documents/Sen2-Agri-Software-User-Manual-3.2.pdf>>>. Acesso em: 07 jun. 2022. 18, 19, 21
- WEISS, M.; JACOB, F.; DUVEILLER, G. Remote sensing for agricultural applications: A meta-review. **Remote Sensing of Environment**, v. 236, p. 111402, 2020. ISSN 0034-4257. Disponível em: <<<https://www.sciencedirect.com/science/article/pii/S0034425719304213>>>. 2
- WORBOYS, M. F.; DUCKHAM, M. **GIS: a computing perspective**. [S.l.]: CRC Press, 2004. 10
- WULDER, M. A.; ROY, D. P.; RADELOFF, V. C.; LOVELAND, T. R.; ANDERSON, M. C.; JOHNSON, D. M.; HEALEY, S.; ZHU, Z.; SCAMBOS, T. A.; PAHLEVAN, N.; HANSEN, M.; GORELICK, N.; CRAWFORD, C. J.; MASEK, J. G.; HERMOSILLA, T.; WHITE, J. C.; BELWARD, A. S.; SCHAAF, C.; WOODCOCK, C. E.; HUNTINGTON, J. L.; LYMBURNER, L.; HOSTERT, P.; GAO, F.; LYAPUSTIN, A.; PEKEL, J.-F.; STROBL, P.; COOK, B. D. Fifty years of Landsat science and impacts. **Remote Sensing of Environment**, v. 280, p. 113195, 2022. ISSN 0034-4257. Disponível em: <<<https://www.sciencedirect.com/science/article/pii/S0034425722003054>>>. 12
- ZAGLIA, M. C.; VINHAS, L.; QUEIROZ, G. R. de; SIMOES, R. Catalogação de metadados do cubo de dados do brasil com o spatiotemporal asset catalog. **International Journal of Applied Earth Observation and Geoinformation**, p. 280–285, 2019. Disponível em: <<<http://mtc-m16d.sid.inpe.br/rep/8JMKD3MGPDW34R/3UFEFD8>>>. 26, 27
- ZIOTI, F.; FERREIRA, K. R.; QUEIROZ, G. R.; NEVES, A. K.; CARLOS, F. M.; SOUZA, F. C.; SANTOS, L. A.; SIMOES, R. E. A platform for land use and land cover data integration and trajectory analysis. **International Journal of Applied Earth Observation and Geoinformation**, v. 106, p. 102655, 2022. ISSN 1569-8432. Disponível em: <<<https://www.sciencedirect.com/science/article/pii/S0303243421003627>>>. 27, 28, 29, 30