

# Fusão de dados para análise de imagens registradas por satélites: proposta de modelo de metadados

## Isaque Katahira

Doutorando em Ciência da Informação pela Universidade Estadual Paulista Júlio de Mesquita Filho (Unesp) - Palmital, SP - Brasil. Mestre em Bioinformática pela Universidade Tecnológica Federal do Paraná (UTFPR) - Brasil. Professor da Faculdade de Tecnologia (FATEC) - Pompeia, SP - Brasil.

<http://lattes.cnpq.br/9340096523226667>

<https://orcid.org/0000-0001-5800-9890>

E-mail: [isaque.katahira@fatec.sp.gov.br](mailto:isaque.katahira@fatec.sp.gov.br)

## Danilo Camargo Dias

Mestrando em Ciência da Informação pela Universidade Estadual Paulista Júlio de Mesquita Filho (Unesp) - Marília, SP - Brasil. Especialização em Formação Pedagógica (Licenciatura) pela Centro Estadual de Educação Tecnológica Paula Souza (CEETEPS) - Brasil. Especialização em Engenharia de Sistemas pela Escola Superior Aberta do Brasil (ESAB) - Brasil. Professor da Escola Técnica Estadual (ETEC) - Itapeva, SP - Brasil.

<http://lattes.cnpq.br/4387186879667385>

<https://orcid.org/0000-0002-9838-8861>

E-mail: [danilo.dias@etec.sp.gov.br](mailto:danilo.dias@etec.sp.gov.br)

## Danilo Dolci

Mestrando em Ciência da Informação pela Universidade Estadual Paulista Júlio de Mesquita Filho (Unesp) - Marília, SP - Brasil. Especialização em Pós-graduação em Redes de Computadores pela Faculdade Estácio de Sá de Ourinhos (FAESO) - Brasil. Especialização em Ciências da Computação pela Universidade Estadual de Londrina (UEL) - PR - Brasil. Professor e coordenador do Curso de Análise e Desenvolvimento de Sistemas da Faculdade de Tecnologia de Garça (FATEC) - Garça, SP - Brasil.

<http://lattes.cnpq.br/1419975657681241>

<https://orcid.org/0000-0003-0415-6240>

E-mail: [danilo.dolci@fatec.sp.gov.br](mailto:danilo.dolci@fatec.sp.gov.br)

## Mariângela Spotti Lopes Fujita

Livre-docência pela Universidade Estadual Paulista Júlio de Mesquita Filho (UNESP) - Brasil. Pós-Doutorado pela Universidad de Murcia (UM) - Espanha. Doutora em Ciências da Comunicação pela Universidade de São Paulo (USP) - São Paulo, SP - Brasil. Membro permanente do Programa de Pós-Graduação em Ciência da Informação da Universidade Estadual Paulista Júlio de Mesquita Filho (Unesp) - Marília, SP - Brasil.

<http://lattes.cnpq.br/6530346906709462>

<http://orcid.org/0000-0002-8239-7114>

E-mail: [fujita@marilia.unesp.br](mailto:fujita@marilia.unesp.br)

## Leonardo Castro Botega

Pós-Doutorado pelo Instituto de Ciências Matemáticas e de Computação (ICMC-USP) - Brasil. Doutor em Ciência da Computação pela Universidade Federal de São Carlos (UFSCAR) - SP - Brasil. Professor da Universidade Estadual Paulista Júlio de Mesquita Filho (Unesp), Brasil.

<http://lattes.cnpq.br/6027755717265622>

<https://orcid.org/0000-0003-1495-5935>

E-mail: [leonardo.botega@unesp.br](mailto:leonardo.botega@unesp.br)

## Isidoro Gil Leiva

Doutor em Técnicas y Métodos Actuales en Información pela Universidad de Murcia (UM) - Espanha. Professor da Universidad de Murcia (UM) - Campus de Espinardo - Murcia - Espanha

<http://lattes.cnpq.br/4334024093986852>

<https://orcid.org/0000-0002-7175-3099>

E-mail: [isgil@um.es](mailto:isgil@um.es)

Submetido em: 01/12/2020. Aprovado em: 09/03/2021. Publicado em: 28/07/2021.

## RESUMO

Diversas tecnologias surgiram nas últimas décadas, que propiciaram um crescimento considerável no volume e na complexidade dos repositórios digitais de imagem. Diante desse desenvolvimento, também se faz necessária a criação de novas estratégias de tratamento da informação, sobretudo aquelas ligadas à organização e recuperação. Historicamente, o uso de metadados é uma técnica utilizada para identificação de arquivos de imagem. Nesse sentido, a fusão de dados pode ser utilizada como um método para conectar parâmetros em conjuntos de dados heterogêneos para indexação em repositórios de imagem com vistas a estabelecer conexões. Sendo assim, o objetivo deste trabalho é demonstrar as potencialidades da fusão de dados para a organização de imagens registradas por diferentes satélites, especificamente, desenvolver um modelo de metadados para a representação da informação na fusão de dados por sistemas de IA. Esta pesquisa é uma revisão bibliográfica exploratória, utilizada para demonstrar as potencialidades da fusão de dados a partir de imagens registradas por diferentes satélites, oriundas de dois sistemas de monitoramento da NASA - *National Snow & Ice Data Center* e o *GES DISC Data Archive* -, com posterior desenvolvimento de proposta de modelo de metadados para auxiliar a fusão de dados através da observação de parâmetros comuns e da exploração da estrutura dos metadados. Ao concluir este trabalho, pôde-se comprovar que a fusão de dados de imagem se destaca como uma estratégia eficiente de representação e recuperação da informação e colabora com a ampliação da compreensão das alterações climáticas no planeta.

**Palavras-chave:** Repositórios digitais de imagem. Metadados de imagens. Ciência de dados.

## ***Fusion of data for analysis of images recorded by satellites: proposal of metadata model***

### **ABSTRACT**

*Several technologies have emerged in recent decades that have provided considerable growth in the volume and complexity of digital image repositories. In view of this development, it is also necessary to create new information treatment strategies, especially those related to organization and recovery. Historically, the use of metadata is a technique used to identify image files. In this sense, data fusion can be used as a method to connect parameters in heterogeneous data sets for indexing in image repositories, in order to establish connections. In this sense, the objective of this work is to demonstrate the potential of data fusion for the organization of images recorded by different satellites, specifically, to develop a metadata model for the representation of information in data fusion by AI systems. This research is an exploratory bibliographic review used to demonstrate the potential of data fusion from images recorded by different satellites, coming from two NASA monitoring systems: National Snow & Ice Data Center and the GES DISC Data Archive, with further development proposal of a metadata model to assist data fusion, by observing common parameters and exploring the metadata structure. At the end of this work, it was possible to prove that the fusion of image data stands out as an efficient strategy for the representation and recovery of information and collaborates to expand the understanding of climate change on the planet.*

**Keywords:** Digital image repository. Image metadata. Data science.

## **Fusão de dados para el análisis de imágenes grabadas por satélites: propuesta de modelo de metadatos**

### **Resumen**

*Han surgido varias tecnologías en las últimas décadas que han proporcionado un crecimiento considerable en el volumen y la complejidad de los repositorios de imágenes digitales. En vista de este desarrollo, también es necesario crear nuevas estrategias de tratamiento de la información, especialmente las relacionadas con la organización y la recuperación. Históricamente, el uso de metadatos es una técnica utilizada para identificar archivos de imagen. En este sentido, la fusión de datos puede usarse como un método para conectar parámetros en conjuntos de datos heterogéneos para indexar en repositorios de imágenes, a fin de establecer conexiones. Mientras tanto, el objetivo de este trabajo es demostrar el potencial de fusión de datos para la organización de imágenes grabadas por diferentes satélites, específicamente, para desarrollar un modelo de metadatos para la representación de información en fusión de datos por sistemas de IA. Esta investigación utiliza la revisión bibliográfica exploratoria para demostrar el potencial de fusión de datos de imágenes grabadas por diferentes satélites, provenientes de dos sistemas de monitoreo de la NASA: el Centro Nacional de Datos de Nieve y Hielo y el Archivo de Datos GES DISC, con un mayor desarrollo propuesta de un modelo de metadatos para ayudar a la fusión de datos, observando parámetros comunes y explorando la estructura de metadatos. Al final de este trabajo, fue posible demostrar que la fusión de datos de imágenes se destaca como una estrategia eficiente para la representación y recuperación de información y colabora para expandir la comprensión del cambio climático en el planeta.*

**Palabras clave:** Repositorios de imágenes digitales. Metadatos de imagen. Ciencia de datos.

### **INTRODUÇÃO**

Com o crescimento dos repositórios digitais de imagens, as técnicas de fusão de dados têm conquistado uma importância cada vez maior. A intensidade e a heterogeneidade com que arquivos, especialmente os de imagens, são produzidos, traz novos e grandes desafios, pois à medida que os dados gerados se multiplicam, há também a necessidade crescente de criar estratégias para sua organização e recuperação. Uma forma de aprimorar a descrição e a identificação de documentos de imagens é a utilização de metadados.

O indexador deve considerar tanto a rapidez quanto a precisão proporcionada pelos metadados catalogados. No entanto, observando diferentes fontes, percebe-se que nem sempre os metadados são convergentes, fato que pode trazer dúvidas aos usuários ou explicitar lacunas de observação. Nesse sentido, a fusão de dados configura-se como uma estratégia importante para extração de parâmetros mais representativos (HALL; JORDAN, 2010; BOTEGA, 2016).

Para Juan e Tauler (2019), a fusão de dados implica com frequência na concatenação de conjuntos de dados que apresentam uma enorme diversidade em termos de informação, tamanho e comportamento. As informações conectadas refletem a variação atribuída por componentes, eventos ou fontes, agregando dados representativos que podem ser analisados simultaneamente.

É notório que as últimas décadas presenciaram o surgimento de novas aplicações para a tecnologia da informação, sobretudo no que tange à organização de dados dos repositórios de imagens digitais. Somado a isso, o compartilhamento de documentos aumentou, principalmente, quando se trata de fotografias. Contudo, segundo Oliveira e Vital (2015), não há, na literatura científica da área da Ciência da Informação, um acordo sobre como as imagens devem ser tratadas, pois diversos fatores implicam nessa análise, incluindo o tipo de imagem (pinturas, fotografias etc.) e o suporte no qual ela é divulgada.

Nesse contexto, a hipótese que orientará a investigação realizada é que a execução do registro de um fenômeno a partir de diversos critérios colabora para a sua compreensão mais global. Assim, a fusão de dados se mostraria relevante à obtenção de um resultado superior de recuperação, na medida em que viabilizaria a comparação, a extração e a correlação de dados sob diferentes perspectivas.

A fim de confirmar a validade da hipótese formulada, utiliza-se a Ciência de Dados para reunir e comparar informações sobre eventos ambientais, bem como identificar os registros mais significativos para sua recuperação. A proposição desta pesquisa é evidenciar as potencialidades da fusão de dados para a organização de imagens registradas por diferentes satélites e, assim, desenvolver um modelo de metadados para a representação da informação na fusão de dados por sistemas de inteligência artificial (IA). As imagens selecionadas são resultado do monitoramento do dióxido de carbono livre na troposfera, da temperatura dos oceanos e da temperatura atmosférica.

A relevância da proposta se confirma pela inegável utilidade das imagens de satélite para monitoramento de grandes áreas, permitindo a compreensão e a avaliação de alterações ambientais (tanto positivas quanto negativas), ao correlacionar, por exemplo, os índices de emissão de CO<sub>2</sub> (dióxido de carbono) à temperatura média dos oceanos e à temperatura atmosférica. Ademais, a atualidade temática confirma a validade da pesquisa, já que iniciativas de aprimorar o uso da ciência em benefício do meio ambiente se mostram urgentes.

Nesse ínterim, a fusão de dados revela-se fundamental ao potencializar as informações obtidas pela ciência, permitindo avaliações capazes de orientar políticas baseadas em evidências, a fim de contribuir para a proposição de ações efetivas que auxiliem na manutenção da vida humana (SRIVASTAVA *et al.*, 2019; VENKATESH *et al.*, 2019).

Para alcançar os objetivos pretendidos, inicialmente, realiza-se revisão bibliográfica exploratória sobre padrões de metadados e fusão de dados de imagens. Na sequência, consulta-se dois sistemas de monitoramento da NASA, o *National Snow & Ice Data Center* e o *GES DISC Data Archive*, especificamente com relação aos metadados catalogados sobre as imagens de interesse, para, finalmente, propor a fusão dos dados encontrados com vistas a melhor compreender os efeitos dos altos índices de emissão de CO<sub>2</sub> para a qualidade de vida no planeta.

## FUNDAMENTAÇÃO TEÓRICA

*A priori*, a revisão bibliográfica exploratória subsidiou a pesquisa por permitir o desenvolvimento de hipóteses, o levantamento da fortuna crítica disponível e a compreensão organizada dos conceitos (MARCONI; LAKATOS, 2017). De modo preliminar, para destacar a importância da fusão de dados como estratégia que congrega informações relevantes, é oportuno mencionar a criação de metadados.

Segundo Alves (2012), os metadados são dados que possibilitam a recuperação, descrição, avaliação, interpretação e manipulação de documentos eletrônicos. Nesse sentido, os metadados são considerados essenciais em diversas fases do ciclo informacional. Para Simionato (2012), não há como recuperar ou preservar um recurso informacional sem que exista um registro com base em metadados, pois são eles que representam documentalmente os arquivos selecionados.

No âmbito das produções de imagens, o registro de metadados para identificação e recuperação de informações é ainda mais desafiador, tendo em vista a necessidade de registrar, por meio de palavras, dados obtidos por intermédio da observação da linguagem não verbal (YAMANE; CASTRO, 2018).

Há diferentes formas de representação ou reprodução de uma imagem: pinturas, ilustrações, animações, ícones, fotografias entre outras.

Sobre a fotografia, pode-se defini-la como um gênero textual que apresenta certa complexidade para o processo de indexação, uma vez que é o resultado da captação e fixação de uma determinada ação, em determinado momento e lugar, por determinado agente. Diante dessa captação, muitíssimo peculiar, o indexador encontra a difícil tarefa de registrar verbalmente o que fora registrado com todas as nuances da linguagem não verbal. Diante desse desafio, fica evidente a interferência da subjetividade, pois, se a captação foi feita em determinado tempo-espço, também o contexto da leitura feita pelo indexador é relevante e pode interferir nos termos atribuídos durante a indexação.

Dessa forma, o uso de metadados e descritores em imagens fotográficas mostra-se fundamental, visto que existem informações significativas nesse tipo de conteúdo que podem ser perdidas, caso os processos de catalogação e indexação não sejam realizados conforme os critérios validados.

A catalogação e a indexação são mecanismos que condensam a informação, a fim de sua representação, para a sua recuperação. A partir disso, a catalogação e a indexação de fotografia em ambientes digitais se fazem importantes, não podendo ser elaboradas de qualquer maneira (FELIPE; FELIPE, 2016, p. 5).

Ainda de acordo com Felipe e Felipe (2016), a propósito da indexação de imagens, é necessário manter em perspectiva que esse tipo de registro é carregado de subjetividade e, por isso, sua descrição por meio de metadados é essencial à recuperação de seu conteúdo.

Na esfera científica, as fotografias obtidas a partir do sensoriamento remoto se materializam como valiosas fontes de informação para pesquisas relacionadas ao clima do planeta (ZANOTTA; FERREIRA; ZORTEA, 2019). Recortando intervalos espaço-temporais, as fotografias permitem tanto o acompanhamento sincrônico quanto diacrônico das questões ambientais, fato que permite um olhar contextualizado sobre as principais mudanças naturais e os fenômenos provocados pela interação humana.

Segundo Zanotta, Ferreira e Zorteia (2019), a crescente facilidade para se conseguir imagens da superfície terrestre em alta resolução, graças ao uso de satélites, possibilitou uma aproximação inovadora entre a tecnologia e a sociedade. As imagens fixadas trazem um imenso conteúdo informacional que, coletado e processado continuamente em diversas frentes de pesquisa científica, pode contribuir de maneira decisiva para a compreensão da vida no planeta.

Chino, Romani e Traina (2010) afirmam que, como os dados são gerados por diversos tipos de satélites, com imagens registradas em resolução e periodicidades diferentes, é de suma importância investir na integração das informações obtidas. Para que essa integração se concretize exponencialmente, a fusão de dados – conceituada como um conjunto de processos contínuos de refinamento e avaliações que lidam com múltiplos dados e informações, de fontes únicas ou diversas – é uma estratégia vital (HALL; JORDAN, 2010).

Segundo Botega (2016), a fusão de dados ou informações corresponde à rotina de transformação de dados ou informações para produzir estimativas e predições de estados de entidades, visando a maximizar o valor da informação e estimular a consciência situacional de analistas sobre um ambiente de interesse.

Sobre a questão, a título de exemplo, o trabalho de Srivastava *et al.* (2019) busca detectar áreas com alta probabilidade de incêndio florestal em uma região da Índia, com base na diferença da *Normalized Burn Ratio* (NBR), entre as condições pré e pós-incêndio. O estudo compara *datasets* gerados pelo satélite LANDSAT TM 5<sup>1</sup> por dois modelos de captura, o *Geographical Information Systems* (GIS)<sup>2</sup> e o *Earth Observation* (EO)<sup>3</sup>. O resultado é uma avaliação multicritério, incorporando atributos como fontes antropogênicas e naturais, de modo a fundir os modelos e realizar a predição de zonas com alta probabilidade de incêndio.

<sup>1</sup> Disponível em: <http://landsat.gsfc.nasa.gov/>. Acesso em: 17 jun. 2020.

<sup>2</sup> Disponível em: <https://grass.osgeo.org/>. Acesso em: 17 jun. 2020.

<sup>3</sup> Disponível em: <https://earthobservatory.nasa.gov/>. Acesso em: 17 jul. 2020.



A partir do exposto, são de suma relevância estudos que abordem a fusão de dados em relação a todas as áreas do conhecimento científico, sobretudo no tratamento de imagens fotográficas. A coleta e a organização de informações oriundas de diversas fontes configuram-se como estratégia eficiente em busca da qualidade e da assertividade na tomada de decisões.

## METODOLOGIA

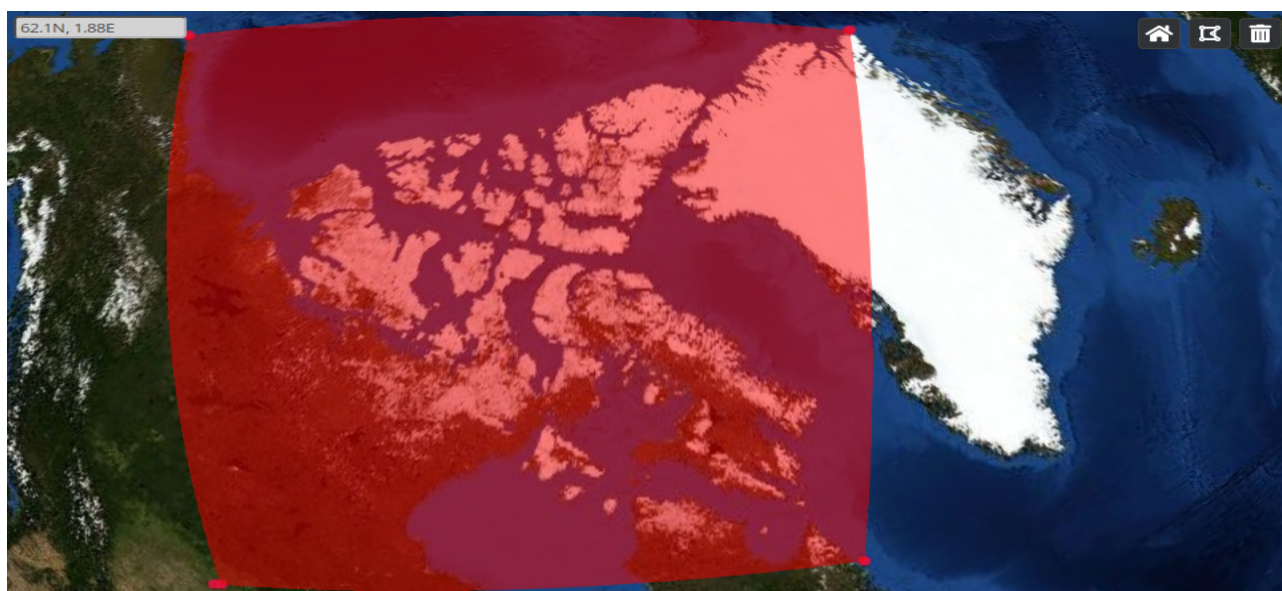
Com o intuito de confirmar a fusão de dados como ferramenta capaz de reunir, selecionar e evidenciar dados relevantes, é realizada a consulta e seleção de imagens e metadados de dois sistemas de monitoramento da *National Aeronautics and Space Administration* (NASA)<sup>4</sup>, o *National Snow & Ice Data Center* (NSIDC)<sup>5</sup> e o *GES DISC Data Archive*<sup>6</sup>, especificamente com relação à temática de interesse.

A partir dos registros selecionados, evidencia-se as potencialidades da fusão dos dados encontrados, a fim de melhor compreender como os altos índices de emissão de CO<sub>2</sub> reverberam a qualidade de vida na Terra.

As ferramentas *National Snow & Ice Data Center* e *GES DISC Data Archive* foram acessadas pela área de dados abertos da NASA para extrair metadados de imagens obtidas por satélites que monitoram o dióxido de carbono livre na troposfera, a temperatura dos oceanos e a temperatura atmosférica.

O NSIDC é um centro de pesquisas da NASA que gerencia e distribui dados científicos sobre a criosfera, isto é, regiões da superfície terrestre cobertas permanentemente por gelo e neve. Aos usuários cadastrados, o sistema permite a solicitação de informações de uma área que deve ser delimitada pelo requerente. Diversas informações são geradas e compactadas num arquivo que fica disponível na área do usuário, conforme figura 1.

Figura 1 – A área destacada em vermelho é a localização da qual foram extraídos os metadados



Fonte: National Snow & Ice Data Center, 2019.

<sup>4</sup> Disponível em: <https://www.nasa.gov/>. Acesso em: 17 jul. 2020.

<sup>5</sup> Disponível em: <https://nsidc.org/>. Acesso em: 17 jul. 2020.

<sup>6</sup> Disponível em: <https://disc.gsfc.nasa.gov/>. Acesso em: 17 jul. 2020.

Ao se consultar uma imagem fotográfica disponibilizada pelo NSIDC em seu ambiente digital, pode-se ver que são utilizados metadados para as descrições. Os campos que compõem a catalogação são: identificador do *datacenter*, data de registro, data da última alteração; nome do arquivo compartilhado; tamanho (MB); data e hora da captura; espaço de tempo inicial da captura; espaço de tempo final da captura; coordenadas da área observada (área retangular - norte, sul, leste e oeste); nome da plataforma; instrumento de sensoriamento utilizado; nome da campanha da NASA e identificador da aeronave utilizada.

A partir da figura 1, é possível obter treze registros de metadados, explicitados no quadro 1.

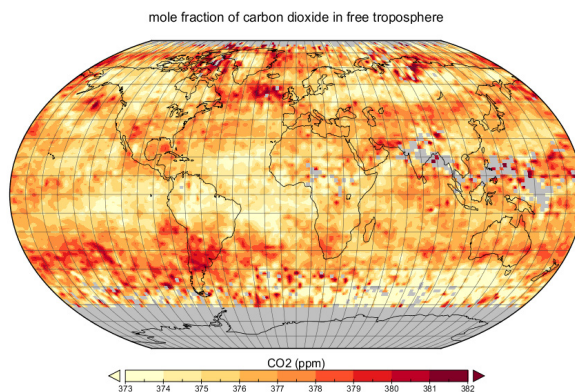
Quadro 1 – Derretimento das calotas polares

National Snow & Ice Data Center	
Identificador do datacenter	NSIDC
Data de registro	2017-08-07T15:50:19Z
Data da última alteração	2017-08-10T15:23:31Z
Nome do arquivo compartilhado	ILATM1B_20180419_181100.ATM6AT6.h5
Tamanho (MB)	5.62564
Data e hora da captura	2017-12-18 03:18:19.000
Espaço de tempo inicial da captura	2017-04-19T18:11:00Z
Espaço de tempo final da captura	2017-04-20T08:25:00Z
Coordenadas da área observada (Área retangular – Norte, Sul, Leste e Oeste)	76.54604 76.53789 76.53673 76.54488
Nome da plataforma	P-3B
Instrumento de sensoriamento utilizado	ATM
Nome da campanha da NASA	2017_GR_NASA
Identificador da aeronave utilizada	N426NA

Fonte: Elaborado pelo autor com base nos dados obtidos na ferramenta *National Snow & Ice Data Center*, 2019.

O *GES DISC Data Archive* é o sistema que fornece dados, informações e serviços de ciências da Terra para pesquisadores, cientistas de dados, usuários de aplicativos e estudantes, incluindo: composição atmosférica, ciclos de água e energia, e variabilidade climática. É possível observar que também são arquivados conjuntos de dados aplicáveis ao Ciclo de Carbono e ao Ecossistema. Assim, o *GES DISC* disponibiliza diversos *datasets* sobre a situação climática do planeta. O sistema gera e renderiza imagens e metadados sobre regiões em diversos padrões concernentes à emissão de CO<sub>2</sub>, à temperatura dos oceanos e à temperatura atmosférica. Quanto à emissão de CO<sub>2</sub>, selecionou-se a seguinte imagem:

Figura 2 – Dióxido de carbono livre na troposfera



Fonte: [https://docsserver.gesdisc.eosdis.nasa.gov/public/project/Images/AIRS3C28\\_005.png](https://docsserver.gesdisc.eosdis.nasa.gov/public/project/Images/AIRS3C28_005.png), 2019.

Sobre a figura 2, os treze metadados relacionados na catalogação são: data de registro; data da última alteração; nome da coleção a que pertence; identificador da coleção; tamanho (MB); período (diurno / noturno); data e hora da captura; espaço de tempo inicial da captura; espaço de tempo final da captura; coordenadas da área observada (área retangular - norte, sul, leste e oeste); parâmetros medidos; parâmetro de qualidade automatizada e URL de acesso online ao recurso. Nesta pesquisa, optou-se pelo padrão Nativo, que permitiu a identificação dos metadados, conforme quadro 2.

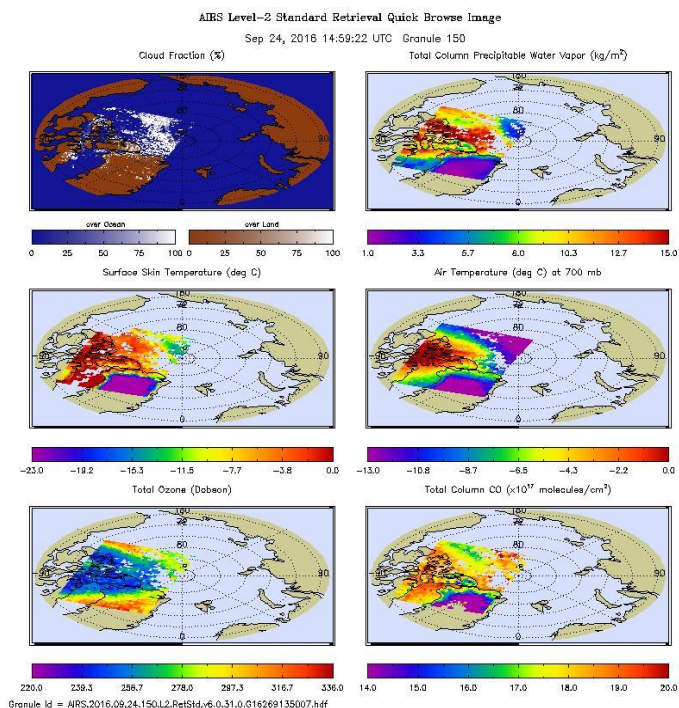
No que tange à temperatura dos oceanos, selecionou-se, a título de exemplo, a figura 3, com dezessete metadados descritos.

Quadro 2 – Emissão de CO<sub>2</sub>

GES DISC Data Archive	
Data de registro	2017-08-07T15:50:19Z
Data da última alteração	2017-08-09T15:50:19Z
Nome da coleção a que pertence	AIRS3C28
Identificador da coleção	5
Tamanho (MB)	0.503100395202637
Período (Diurno/Noturno)	Day
Data e hora da captura	2017-03-09T22:51:31.000Z
Espaço de tempo inicial da captura	2017-03-08T00:00:00.000000Z
Espaço de tempo final da captura	2017-03-06T23:59:59.999999Z
Coordenadas da área observada (Área retangular - Norte, Sul, Leste e Oeste)	-180.0 90.0 180.0 -90.0
Parâmetros medidos	Carbon Dioxide Column Density
Parâmetro de qualidade automatizada	Passed
URL de acesso online ao recurso	<a href="http://acdisc.gesdisc.eosdis.nasa.gov/data//Aqua_AIRS_Level3/AIRS3C28.005/2017/AIRS.2017.02.14.L3.CO2Std_IR008.v5.9.14.0.IRonly.X17068145128.hdf">http://acdisc.gesdisc.eosdis.nasa.gov/data//Aqua_AIRS_Level3/AIRS3C28.005/2017/AIRS.2017.02.14.L3.CO2Std_IR008.v5.9.14.0.IRonly.X17068145128.hdf</a>

Fonte: Elaborado pelo autor com base nos dados obtidos na ferramenta GES DISC Data Archive, 2019.

Figura 3 – Temperatura dos oceanos



Fonte: [https://airsl2.gesdisc.nasa.gov/data/Aqua\\_AIRS\\_Level2/AIRX2RET.006/2016/268/AIRS.2016.09.24.150.L2.RetStd.v6.0.31.0.G16269135007.hdf.jpg](https://airsl2.gesdisc.nasa.gov/data/Aqua_AIRS_Level2/AIRX2RET.006/2016/268/AIRS.2016.09.24.150.L2.RetStd.v6.0.31.0.G16269135007.hdf.jpg), 2019.



Sobre a figura 3, os metadados relacionados na catalogação são: data de registro; data da última alteração; nome da coleção a que pertence; indicador da coleção; tamanho (MB); tipo de camada atmosférica observada; período (diurno/noturno); data e hora da captura; espaço de tempo inicial da captura; espaço de tempo final da captura; domínio espacial vertical (máximo e mínimo);

coordenadas da área observada (área retangular - norte, sul, leste e oeste); URL de acesso online ao recurso, descrição; tipo de dado (aberto ou não); *mime type* e navegação em imagens associadas (URL, tamanho e descrição), como registrado no quadro 3.

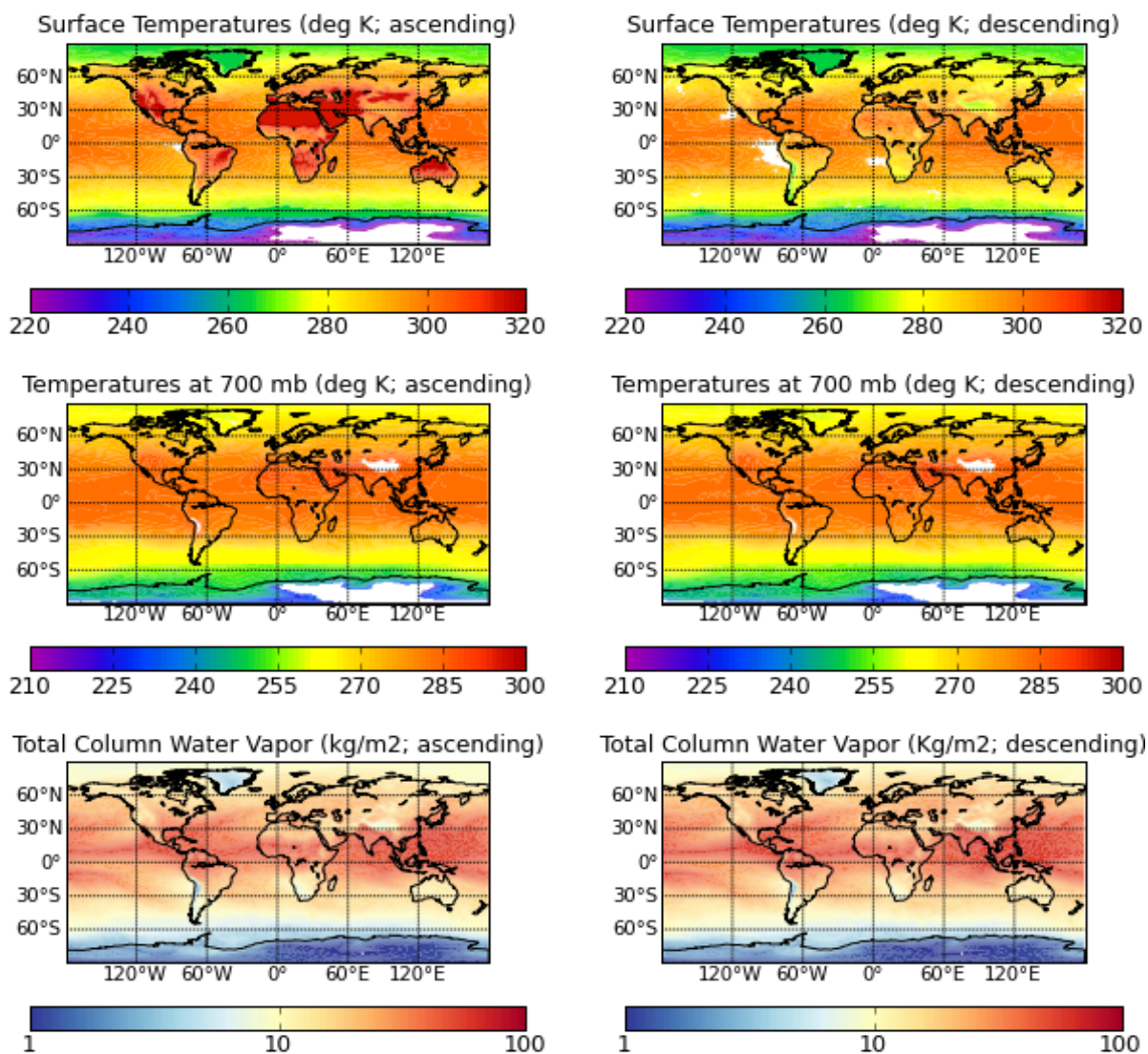
A última imagem selecionada para apresentação de metadados representativos faz referência à temperatura atmosférica.

Quadro 3 – Temperatura dos oceanos

GES DISC Data Archive	
Data de registro	2017-08-07T15:50:19Z
Data da última alteração	2017-08-08T15:54:19Z
Nome da coleção a que pertence	AIRX2RET
Identificador da coleção	6
Tamanho (MB)	3.583529472351074
Tipo de camada atmosférica observada	Atmosphere Profile
Período (Diurno/Noturno)	BOTH
Data e hora da captura	2017-04-25T17:50:08.000Z
Espaço de tempo inicial da captura	2017-04-24T14:59:22.000000Z
Espaço de tempo final da captura	2017-04-25T14:59:22.000000Z
Domínio espacial vertical (máximo e mínimo)	0.1mb, or appr. 50km
Coordenadas da área observada (Área retangular – Norte, Sul, Leste e Oeste)	160.906860351562 90.0 -22.1724796295166 68.8115463256836
URL de acesso online ao recurso	<a href="https://airs12.gesdisc.eosdis.nasa.gov/data/Aqua_AIRS_Level2/AIRX2RET.006/2016/268/AIRS.2016.09.24.150.L2.RetStd.v6.0.31.0.G16269135007.hdf">https://airs12.gesdisc.eosdis.nasa.gov/data/Aqua_AIRS_Level2/AIRX2RET.006/2016/268/AIRS.2016.09.24.150.L2.RetStd.v6.0.31.0.G16269135007.hdf</a>
Descrição	The OPENDAP location for the granule
Tipo de dado (Aberto ou não)	GET DATA : OPENDAP DATA
Mime type	application/x-hdf
Navegação em imagens associadas (URL, tamanho e descrição)	<a href="https://airs12.gesdisc.eosdis.nasa.gov/data/Aqua_AIRS_Level2/AIRX2RET.006/2016/268/AIRS.2016.09.24.150.L2.RetStd.v6.0.31.0.G16269135007.hdf.jpg">https://airs12.gesdisc.eosdis.nasa.gov/data/Aqua_AIRS_Level2/AIRX2RET.006/2016/268/AIRS.2016.09.24.150.L2.RetStd.v6.0.31.0.G16269135007.hdf.jpg</a> 170797 Browse image for AIRS.2016.09.24.150.L2.RetStd.v6.0.31.0.G16269135007.hdf

Fonte: Elaborado pelo autor com base nos dados obtidos na ferramenta GES DISC Data Archive, 2019.

Figura 4 – Temperatura atmosférica



Local Granule Id = AIRS.2013.08.01.L3.RetStd031.v6.0.9.0.G13247185301.hdf

Fonte: [https://docserver.gesdisc.eosdis.nasa.gov/public/project/Imag es/AIRH3SPM\\_006.png](https://docserver.gesdisc.eosdis.nasa.gov/public/project/Imag es/AIRH3SPM_006.png), 2019.

Para descrição da figura 4, são fixados quinze metadados: data de registro; data da última alteração; nome da coleção a que pertence; identificador da coleção; tamanho (MB); período (diurno / noturno); data e hora da captura; espaço de tempo inicial da captura; espaço de tempo final da captura; domínio espacial vertical (máximo e mínimo); coordenadas da área observada (área retangular - norte, sul, leste e oeste); URL de acesso online ao recurso; descrição; tipo de dado (aberto ou não) e *mime type*, como registrado no quadro 4:

Quadro 4 – Temperatura atmosférica

GES DISC Data Archive	
Data de registro	2017-08-07T15:50:19Z
Data da última alteração	2017-08-09T15:55:32Z
Nome da coleção a que pertence	AIRIBQAP_NRT
Identificador da coleção	5
Tamanho (MB)	2.480497360229492
Período (Diurno/Noturno)	DAY
Data e hora da captura	2017-05-24T13:26:06.000Z
Espaço de tempo inicial da captura	2017-05-23T12:29:21.000000Z
Espaço de tempo final da captura	2017-05-24T10:35:20.999999Z
Domínio espacial vertical (máximo e mínimo)	TOA SFC
Coordenadas da área observada (Área retangular – Norte, Sul, Leste e Oeste)	-98.5035018920898 90.0 17.6566963195801 63.0190010070801
URL de acesso online ao recurso	<a href="https://discnrt1.gesdisc.eosdis.nasa.gov/data//Aqua_AIRS_NRT/AIRIBQAP_NRT.005/2019/236/AIRS.2019.08.24.125.L1B.AIRS_QaSub.v5.0.23.0.R19236092606.hdf">https://discnrt1.gesdisc.eosdis.nasa.gov/data//Aqua_AIRS_NRT/AIRIBQAP_NRT.005/2019/236/AIRS.2019.08.24.125.L1B.AIRS_QaSub.v5.0.23.0.R19236092606.hdf</a>
Descrição	The OPENDAP location for the granule
Tipo de dado (aberto ou não)	GET DATA: OPENDAP DATA
Mime type	application/x-hdf

Fonte: Elaborado pelo autor com base nos dados obtidos na ferramenta GES DISC Data Archive, 2019.

Assim, tendo destacado as imagens selecionadas para a amostra e exposto os respectivos metadados, apresenta-se a análise dos resultados e um modelo para a fusão de dados.

## ANÁLISE E DISCUSSÃO DOS RESULTADOS

Diante da seleção das imagens e da enumeração-descritiva dos respectivos metadados apresentados nos quadros 1 a 4, é possível constatar que a complementaridade das informações catalogadas permite, ao pesquisador, maior abrangência de análise do fenômeno de interesse. Sobre o derretimento das calotas polares, a identificação de diversas variáveis, possivelmente associadas, tais como, a concentração de CO<sub>2</sub> na troposfera, a temperatura dos oceanos e a temperatura atmosférica, permite um olhar mais sistêmico sobre o problema.

Após a seleção e identificação das imagens da amostra, a organização da fusão de dados foi realizada em três etapas: i) verificação dos metadados que uma base e outra utilizam, ou seja, exploração da estrutura dos metadados, da quantidade, da descrição, do tipo entre outros; ii) verificação da seleção de termos comuns – quantidade e qualidade –, a fim de explicitar os parâmetros de fusão de dados selecionados; e iii) seleção das imagens após a realização da fusão de dados, com posterior tabulação para análise comparada.

A propósito da última etapa, cumpre destacar que a comparação se deu por meio da análise dos dados obtidos separadamente em relação aos dados reunidos.

Pela seleção e descrição das imagens, por intermédio dos metadados identificados, é possível gerar uma visão global da qualidade de vida na Terra. O quadro 5 exhibe uma proposta de fusão de dados. Para tanto, foi selecionada uma área retangular do norte do globo até o extremo sul, de modo a fixar esse eixo.

Concernente à fusão de dados, destacando os parâmetros comuns e a verificação das estratégias que uma e outra utilizam de metadados, ou seja, exploração da estrutura dos metadados, da quantidade, da descrição, do tipo entre outros, os resultados, na primeira etapa, são apresentados conforme disposto no quadro 5.

Com base no quadro 5, é possível inserir os dados em um determinado algoritmo de IA (OSÓRIO; BITTENCOURT, 2000) que faça a leitura e a interpretação dos dados, em um determinado período, e posteriormente, estabeleça a correlação existente entre eles. Para tanto, uma técnica potencial é a de redes neurais, que são amplamente utilizadas em softwares de análise de imagens (GONÇALVES *et al.*, 2008; MOREIRA *et al.*, 2002). Destaca-se o Modelo de *Hopfield* de redes neurais artificiais (RNA) para analisar as imagens obtidas por satélites, haja vista apresentar os conceitos de redes com realimentação e comportamento dinâmico (HAYKIN, 2007).

Outra proposta é a utilização do algoritmo de MAXVER, ou Máxima Verossimilhança, que, pela classificação supervisionada, “[...] considera a ponderação das distâncias entre as médias dos níveis digitais das classes e o pixel, utilizando parâmetros estatísticos, isto é, considerando a distribuição de probabilidade normal para cada classe [...]” (RIBEIRO; CENTENO, 2001, p. 1342).

Nesse sentido, o algoritmo especialista de IA analisa a poluição do ar e a temperatura dos oceanos, estabelecendo, dessa forma, uma correlação com a elevação da temperatura global. Tal procedimento acrescenta novas informações aos metadados das imagens e viabiliza uma série de análises automatizadas da situação do planeta embasando-se na fusão dos dados (que estão separados e armazenados nos metadados) das imagens.

Quadro 5 – Modelo para visão global da qualidade do planeta Terra a partir da fusão de parâmetros

National Snow & Ice Data Center e GES DISC Data Archive
Data de registro
Data da última alteração
Nome da coleção a que pertence
Identificador da coleção
Tamanho (MB)
Período (Diurno/Noturno)
Data e hora da captura
Espaço de tempo inicial da captura
Espaço de tempo final da captura
Coordenadas da área observada (Área retangular – Extremo Norte Global Fixo, Extremo Sul Global Fixo, Leste Variável e Oeste Variável)
Quantidade de CO2 registrado
Camada atmosférica onde foi registrada a medição da temperatura dos oceanos
Registro da temperatura média atmosférica continental
Dimensão das calotas polares registradas na faixa capturada
National Snow & Ice Data Center e GES DISC Data Archive
Correlação entre as médias das temperaturas aferidas
Correlação da quantidade de CO2 e a temperatura média atmosférica
Correlação da dimensão das calotas polares e a temperatura média atmosférica
URL de acesso aos recursos imagéticos

Fonte: Elaborado pelo autor com base nos dados obtidos nas ferramentas *National Snow & Ice Data Center e a GES DISC Data Archive*, 2019.



## CONSIDERAÇÕES FINAIS

Em meio à heterogeneidade, ao volume e à complexidade de dados disponibilizados diariamente, as estratégias para seleção de informações têm ganhado destaque. No cenário da produção de imagens, essa realidade não é diferente, visto que o volume de dados proporcionados pela comunidade em geral, embora impulse avanços em todas as áreas do saber, também torna necessário o desenvolvimento de estratégias para sua seleção e utilização eficiente.

Diante das imagens e dos metadados apresentados com foco no monitoramento de grandes áreas, foi possível ampliar a compreensão e a avaliação de alterações climáticas (tanto positivas quanto negativas), ao correlacionar, por exemplo, os índices de emissão de CO<sub>2</sub> à temperatura dos oceanos e à temperatura atmosférica.

Considerando as potencialidades dos novos recursos disponibilizados pelas ferramentas computacionais, foi possível comprovar que o registro de um fenômeno feito com base em diversos critérios colabora com a sua compreensão mais global, como proposto pela utilização da fusão de dados. A propósito das limitações da pesquisa com imagens, tem-se as restrições técnicas dos sensores, as dificuldades de descrição verbal e possíveis ruídos na catalogação.

Assim, a fusão de dados da área Espacial nas bases de dados do repositório da NASA mostrou-se relevante para a obtenção de um resultado superior de recuperação, pois evidenciou a complementaridade dos dados a partir das correlações estabelecidas. Logo, foi possível destacar os registros mais significativos para sua recuperação.

Por fim, convém sublinhar que a fusão de dados de imagens e os novos métodos de seleção emergem no cenário atual como estratégias eficientes de representação e recuperação da informação. Para trabalhos futuros, o método poderia ser testado em um *corpus* maior de imagens, bem como em outras áreas do conhecimento, de modo a ampliar a validação do comparativo, incluindo análise por meio de RNAs.

## REFERÊNCIAS

- ALVES, C. D. Metadados para a Recuperação de Imagens na WEB: utilizando o software ADOBE BRIDGE. *PontodeAcesso*, Salvador, v. 6, n. 1, p. 32-48, 2012.
- DOI: <http://dx.doi.org/10.9771/1981-6766rpa.v6i1.5131>. Disponível em: <https://periodicos.ufba.br/index.php/revistaici/article/view/5131>. Acesso em: mar. 2021.
- BOTEGA, L.C. *Modelo de fusão dirigido por humanos e ciente de qualidade de informação*. Orientadora: Regina Borges de Araújo. 2016. Tese (Doutorado em Ciência da Computação) – Programa de Pós-Graduação em Ciência da Computação, Universidade Federal de São Carlos, São Paulo, 2016. Disponível em: <https://aberto.univem.edu.br/handle/11077/1483>. Acesso em: 17 jul. 2020.
- CHINO, D. Y. T.; ROMANI, L. A. S.; TRAINA, A. J. M. Construindo séries temporais de imagens de satélite para sumarização de dados climáticos e monitoramento de safras agrícolas. *Revista Eletrônica de Iniciação Científica*, [s.l.], v. 10, n. 3, p. 1-20, 2010.
- FELIPE, C. B. M.; FELIPE, A. A. C. *Análise do uso de metadados no auxílio à recuperação da informação em ambientes digitais*. [s.l.], 2016. Disponível em: [http://www.liber.ufpe.br/home/wp-content/uploads/2016/09/04-Analise-no-uso-de-metadados\\_Felipe.pdf](http://www.liber.ufpe.br/home/wp-content/uploads/2016/09/04-Analise-no-uso-de-metadados_Felipe.pdf). Acesso em: 2 mar. 2020.
- GONÇALVES, M. L. *et al.* Classificação não-supervisionada de imagens de sensores remotos utilizando redes neurais auto-organizáveis e métodos de agrupamentos hierárquicos. *Revista brasileira de cartografia*, [s.l.], v. 60, n. 1, abr. 2008. Disponível em: <http://www.seer.ufu.br/index.php/revistabrasileiracartografia/article/view/44880>. Acesso em: mar. 2021.
- HALL, D.; JORDAN, J. *Human-centered information fusion*. [S.l.]: Artech House, 2010.
- HAYKIN, S. *Redes neurais: princípios e prática*. 2. ed. Porto Alegre: Bookman, 2007.
- JUAN, A.; TAULER, R. Data Fusion by Multivariate Curve Resolution. In: COCCHI, M. (ed.). *Data Handling in Science and Technology*. Netherlands: Elsevier, 2019. v. 31, cap. 8, p. 205-233.
- MARCONI, M. A.; LAKATOS, E. M. *Metodologia do trabalho científico*. 8. ed. São Paulo: Atlas, 2017.
- MOREIRA, F. C. *et al.* *Reconhecimento e classificação de padrões de imagens de núcleos de linfócitos do sangue periférico humano com a utilização de redes neurais artificiais*. 2002. Dissertação (mestrado) - Programa de Pós-Graduação em Ciência da Computação, Universidade Federal de Santa Catarina, Florianópolis, 2002. Disponível em: <https://repositorio.ufsc.br/handle/123456789/82305>. Acesso em: 17 jul. 2020.
- NORMALIZED BURN RATIO. *NBR*. USGS, [2020]. Disponível em: <https://www.usgs.gov/land-resources/nli/landsat/landsat-normalized-burn-ratio>. Acesso em: 20 abr. 2020.

OLIVEIRA, R. A.; VITAL, L. P. Análise e indexação de imagens na rede Flickr. *Em Questão*, Porto Alegre, v. 21, n. 2, p. 7-30, maio/ago. 2015. DOI: <http://dx.doi.org/10.19132/1808-5245212.7-30>. Disponível em: <https://seer.ufrgs.br/EmQuestao/article/view/50968/33977>. Acesso em: 1 dez. 2019.

OSÓRIO, F. S.; BITTENCOURT, J. R. Sistemas Inteligentes baseados em redes neurais artificiais aplicados ao processamento de imagens. In: WORKSHOP DE INTELIGÊNCIA ARTIFICIAL, 1., 2000, Rio Grande do Sul. *Anais [...]*. Rio Grande do Sul: UNISC, 2000.

RIBEIRO, S. R. A.; CENTENO, J. S. Classificação do uso do solo utilizando redes neurais e o algoritmo MAXVER. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 10., 2001, Foz do Iguaçu. *Anais [...]*. Foz do Iguaçu: INPE, 2001. p. 1341-1348. Disponível em: <http://mar.tecnico.unicamp.br/col/dpi.inpe.br/lise/2001/09.20.17.56/doc/1341.1348.074.pdf>. Acesso em: 17 abr. 2020.

SIMIONATO, A. C. *Representação, acesso, uso e reuso da imagem digital*. 142 f. 2012. Dissertação (Mestrado em Ciência da Informação) – Programa de Pós-Graduação em Ciência da Informação da Faculdade de Filosofia e Ciências, Universidade Estadual Paulista Júlio de Mesquita Filho, Marília, São Paulo, 2012. Disponível em: [https://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/Simionato%20A.C.\\_mestrado\\_C.I.\\_2012.pdf](https://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/Simionato%20A.C._mestrado_C.I._2012.pdf). Acesso em: mar. 2021.

SRIVASTAVA, P. K. *et al.* Deriving forest fire probability maps from the fusion of visible/infrared satellite data and geospatial data mining. *Modeling Earth Systems and Environment*, [s.l.], v. 5, n. 2, p. 627-643, 2019. Disponível em: <https://link.springer.com/article/10.1007/s40808-018-0555-5>. Acesso em: mar. 2021.

VENKATESH, V. *et al.* Precision centric framework for activity recognition using Dempster Shaffer theory and information fusion algorithm in smart environment. *Journal of Intelligent & Fuzzy Systems*, [s.l.], v. 36, n. 3, p. 2117-2124, 2019. DOI: 10.3233/JIFS-169923.

YAMANE, G. A. C.; CASTRO, F. F. O estudo e a identificação dos padrões de metadados para a representação e a recuperação da imagem digital na perspectiva da web. *Em Questão*, Porto Alegre, v. 24, n. 1, p. 145-173, 2018. DOI: <https://doi.org/10.19132/1808-5245241.145-173>. Disponível em: <https://seer.ufrgs.br/index.php/EmQuestao/article/view/71475>. Acesso em: mar. 2021.

ZANOTTA, D. C.; FERREIRA, M. P.; ZORTEA, M. *Processamento de imagens de satélite*. São Paulo: Oficina de Textos, 2019.