

O Ciclo de Vida dos Dados aplicado às Imagens Digitais

Eveline Batista Rodrigues
Cristiane Lucy Rodolfo Bonfeti

Como citar: RODRIGUES, Eveline Batista; BONFETI, Cristiane Lucy Rodolfo. O Ciclo de Vida dos Dados aplicado às Imagens Digitais. *In*: MOREIRA, Fábio Mosso *et. al.* (org.). Transversalidade e verticalidade na Ciência da Informação. Marília: Oficina Universitária; São Paulo: Cultura Acadêmica, 2025. p.87-104. DOI: <https://doi.org/10.36311/2025.978-65-5954-613-8.p87-104>



All the contents of this work, except where otherwise noted, is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 (CC BY-NC-ND 4.0).

Todo o conteúdo deste trabalho, exceto quando houver ressalva, é publicado sob a licença Creative Commons Atribuição-NãoComercial-SemDerivações 4.0 (CC BY-NC-ND 4.0).

Todo el contenido de esta obra, excepto donde se indique lo contrario, está bajo licencia de la licencia Creative Commons Reconocimiento-No comercial-Sin derivados 4.0 (CC BY-NC-ND 4.0).

CAPÍTULO 5

O CICLO DE VIDA DOS DADOS APLICADO ÀS IMAGENS DIGITAIS

Eveline Batista Rodrigues¹ e Cristiane Lucy Rodolfo Bonfeti²

INTRODUÇÃO

As imagens acompanham o processo de socialização e evolução do homem desde a pré-história; delas decorrem a vida e organização social e colaboram com a relação entre os homens. Desde seu início, a imagem esteve relacionada à representação e à noção de imitação do real. Ela surge de uma troca simbólica e de uma imitação fabricada para enfrentar a eliminação/destruição provocada pela passagem do tempo (Santos; Madio, 2019).

A definição de imagem estabelecida pelo Tesouro Brasileiro de Ciência da Informação do Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict) é: um Tipo de Documento (categoria), icônico /iconográfico, com disponibilidade de acesso, por meios ou de forma física, usado para representações visuais, entre algumas espécies das imagens podem ser encontradas: diafilmes, fotografias, ilustrações, imagens digitalizadas e mapas (Pinheiro; Ferrez, 2014).

¹ Mestre em Engenharia Elétrica. Professora no Centro Paula Souza – Etec. E-mail: evelinerodrigues@yahoo.com. Lattes: <http://lattes.cnpq.br/4967341665335583>.

² Doutora em Ciência da Informação. Professora no Centro Paula Souza – Etec. E-mail: cristiane.bonfeti@etec.sp.gov.br. ORCID: <https://orcid.org/0000-0003-0612-607X>. Lattes: <http://lattes.cnpq.br/0859525247632138>.

Atualmente, em um mundo cada vez mais digital, estamos diariamente captando bilhões de imagens, como imagens de satélite, imagens médicas, imagens de paisagens e assim por diante. Também estamos inserindo essas imagens em sites sociais (*Facebook, Instagram, TikTok, WhatsApp*), aplicativos, sistemas web, etc., através dos mais diversos dispositivos eletrônicos e conectados em rede, gerando, dessa forma, um grande volume de imagens disponíveis nos diferentes ambientes informacionais digitais.

Neste cenário, para Simionato, Pinho Neto e Santos (2015) as imagens são recursos informacionais específicos com necessidades de representação e organização que devem ser tratadas apropriadamente, pois, há um grande volume de produção de imagens que faz com que seja necessário que a Ciência da Informação se ocupe do seu tratamento (desde o armazenamento até a disponibilização) com o objetivo de poder acessá-las e localizá-las.

Para além de ser um recurso informacional, a imagem pode ser um instrumento de conhecimento, conforme apontado por Joly (1996, p. 60): a função informativa (ou referencial), muitas vezes dominante na imagem, pode também ampliar-se em uma função epistêmica, proporcionando-lhe a dimensão de instrumento de conhecimento. Instrumento de conhecimento porque certamente fornece informações sobre os objetos, os lugares ou as pessoas, em formas visuais tão diversas quanto as ilustrações, as fotografias, os mapas ou ainda os painéis.

Desta forma, como a imagem é um recurso informacional e um instrumento de conhecimento, de acordo com o campo da Ciência da Informação (CI) que se caracteriza pela classificação e disseminação da informação, podemos compreender que, para que a informação gere conhecimento a um indivíduo, é necessário ser transmitida, e de preferência em um intervalo de tempo aceitável e em um formato compreensível (Dos Santos; Madio, 2018).

Para que as imagens digitais se tornem recursos disponíveis, recuperáveis e acessíveis é necessário desenvolvimento de processos para construção de representações. Um impulso na revolução das imagens digitais foi dado pela expansão da utilização dos computadores, onde ocorreu o

surgimento das técnicas para captura, armazenamento, processamento e transmissão de imagens (Santaella, 2015).

Por fim, o objetivo é examinar o Ciclo de Vida dos Dados (CVD) e adaptá-lo às características específicas dos dados de imagens. Esse enfoque busca proporcionar uma investigação das etapas envolvidas no acesso e uso intensivo dos dados imagéticos, que desempenham um papel central na sociedade contemporânea. Essa análise visa a oferecer uma estrutura que possa auxiliar tanto os usuários finais quanto às instituições na compreensão e gestão eficaz desses dados.

PROCEDIMENTOS METODOLÓGICOS

Os procedimentos metodológicos consistiram em pesquisa bibliográfica através da análise do Ciclo de Vida dos Dados aplicando-o aos dados imagéticos, proporcionando reflexões sobre a delimitação de fases envolvidas no acesso e uso desse recurso informacional.

IMAGEM DIGITAL COMO RECURSO INFORMACIONAL

Define-se recurso imagético digital, como “aquele que foi diretamente produzido por um processo digital, seja ele de qualquer forma de extensão armazenando informações visuais estáticas e que contenham propriedades e limites físicos e que possam ser organizadas” (Simionato, 2015, p.43). Ou seja, para que seja uma imagem digital, basta ela ter sido adquirida através de um processo digital como a digitalização em *scanner*, celulares, câmeras digitais ou até mesmo desenvolvidas através de softwares de criação e edição de imagens especializados.

De acordo com Gonzales (2010, p.21), a imagem é uma função bidimensional de intensidade da luz $f(x,y)$, onde x e y denotam as coordenadas espaciais e o valor de f em qualquer ponto (x,y) é proporcionado ao brilho, ou níveis de cinza, da imagem em um dado ponto. A imagem

digital é uma imagem $f(x,y)$ discretizada tanto em coordenadas espaciais quanto em brilho. Ela pode ser considerada como uma matriz de linhas e colunas cujos índices localizam unicamente um ponto e seu valor identifica o nível de cinza naquele ponto. Esses pontos únicos são denominados elementos da imagem, elementos da figura, “*pixels*” ou “*pels*”, sendo os dois últimos, abreviações do termo em inglês *picture elements* (elementos de figura) (Gonzales, 2010, p.21).

A imagem colorida é armazenada sob um modelo para o armazenamento de cores, cujo propósito é facilitar a especificação das cores em uma forma padrão e de aceite geral. Este modelo é uma especificação de um sistema de coordenadas tridimensionais e um subespaço dentro deste sistema onde cada cor é representada por um único ponto (Gonzales, 2010, p.22).

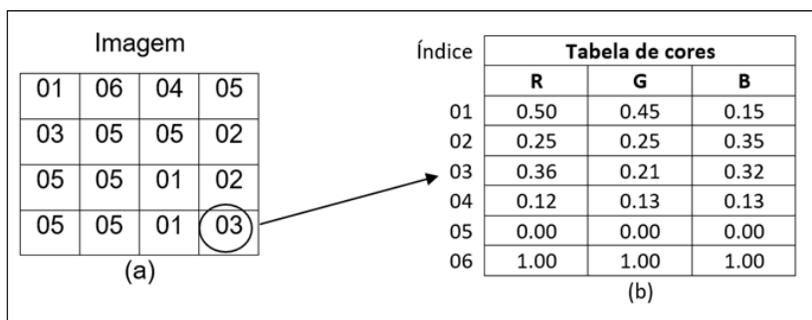
O modelo de cor mais frequentemente utilizado é o Modelo RGB, onde cada cor aparece nos seus componentes espectrais primários de vermelho (*red*), verde (*green*) e azul (*blue*). Neste modelo, para a cor desejada, é feita uma combinação das tonalidades ideais para cada uma das cores primárias.

De acordo com Brown (1995), o Modelo RGB pode ser implementado utilizando uma tabela de cores, onde cada pixel não conterà mais as informações da sua cor, e sim um índice associado a uma tabela de cores que determina as variações de cores possíveis na imagem. Este sistema possui a vantagem de economizar memória no armazenamento, mas, por outro lado, para que o sistema seja eficiente, cada pixel deverá ser representado por um único byte. Uma vez que o maior valor a ser representado por um byte é $2^8 = 256$, a tabela de cores ficará restrita a no máximo 256 cores diferentes.

A Figura 1 traz um exemplo da utilização de mapas de cores. Para um melhor entendimento a Figura 1 foi dividida em (a) e (b), onde (a) contém a representação *bitmap* de uma imagem colorida fictícia e (b) o mapa de cores da imagem representada por (a). Cada pixel de (a) não contém a informação de sua cor, e sim o índice de sua cor, o qual será encontrado na tabela de cores da imagem em (b).

As imagens digitais são obtidas de imagens contínuas através de dois processos de discretização, um para pontos espaciais e outro para cores (ou níveis de cinza), atribuídos a cada pixel. Esses dois processos são chamados de amostragem e quantização, respectivamente (Gonzales, 2010, p.25).

Figura 1 – Esquema para representar a utilização de uma tabela de cores



Fonte: Autoras.

A Ciência da Informação deve contribuir no tocante de que o acesso, uso e disseminação dos dados provenientes de imagens digitais se desenvolva com a experimentação e análise da vasta gama de informações que podem ser obtidas através das imagens digitais, buscando identificar e estruturar fatores e características que auxiliem nos processos de coleta, armazenamento, recuperação e descarte de dados imagéticos para obter a máxima otimização do uso desses dados.

CICLO DE VIDA DOS DADOS

O Ciclo de Vida dos Dados (CVD) concebido por Sant'ana (2016) propõe a utilização de uma delimitação de fases nos processos cíclicos dos dados, como o acesso, o uso e a manutenção dos dados, proporcionando uma estrutura que suporte os esforços, estudos e ações realizadas para obtenção, manutenção e uso de dados, tornando possível aproximar elementos semelhantes e distribuir teorias e metodologias em função de seu escopo.

A fase de coleta de dados envolve a obtenção dos dados necessários para atender a uma demanda específica de informações em um contexto particular. Isso requer o planejamento, a análise da viabilidade e a execução da coleta dos dados, com o objetivo de atender às necessidades identificadas. Nessa fase, surgem exigências por competências específicas relacionadas às necessidades informacionais que impulsionam o processo de coleta de dados (Sant'ana, 2016, p. 119).

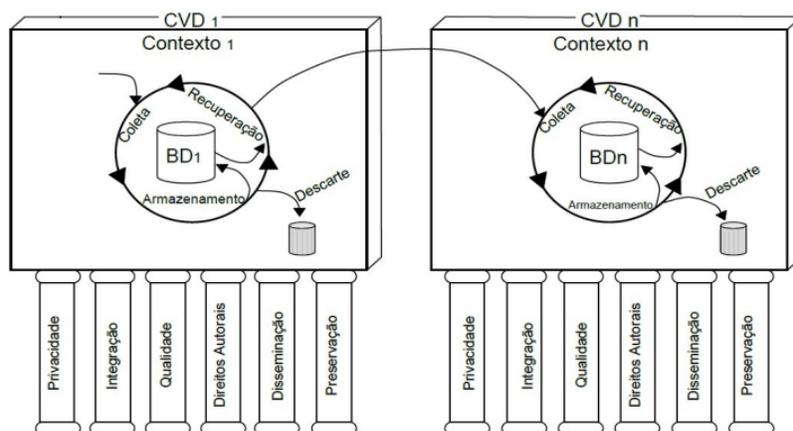
Após a coleta, os dados podem ser utilizados para um fim imediato e descartados ou se houver a necessidade, manter estes dados disponíveis para acesso futuro em um determinado suporte. Inicia-se o processo de armazenamento dos dados coletados, envolvendo planejamento e ações que demandam conhecimento avançado em Ciência da Computação, mas ainda oferecem oportunidades para a participação da Ciência da Informação. O papel do usuário fica mais distante, envolvendo-se principalmente na validação das estruturas de modelos definidas para os dados (Sant'Ana, 2016, p. 120).

Após a fase de armazenamento, pode surgir a necessidade de decidir pelo descarte dos dados ou por sua retenção. Geralmente, a busca é por alternativas que facilitem o acesso e a utilização contínua dos dados, levando a uma próxima fase a ser abordada. Começa, assim, uma nova fase – a recuperação – em que os esforços são voltados para que estes dados possam ser encontrados, acessados e interpretados (Sant'Ana, 2016, p. 121).

Também existe a possibilidade de se identificar que os dados já não são necessários ou que devem ser excluídos da base, o que leva a outra fase identificada como fase de descarte tem-se então a eliminação de parte dos dados (Sant'Ana, 2016, p. 122).

Desta forma, o CVD compreende quatro fases distintas e fatores comuns a todas elas. Esses fatores incluem questões de privacidade, integração, qualidade, direitos autorais, disseminação e preservação de dados. Cada fase é influenciada por esses elementos, que desempenham um papel fundamental ao longo do processo, conforme descrito na Figura 2 (Sant'Ana, 2013).

Figura 2 – Ciclo de Vida dos Dados para Ciência da Informação – (CVD-CI)



Fonte: Sant'Ana (2016, p.123).

Por fim, o CVD-CI será aplicado para as particularidades dos dados imagéticos, proporcionando um estudo sobre a delimitação de fases envolvidas no acesso, uso e manutenção dos dados imagéticos com o propósito de colaborar com as necessidades de representação e organização desse tipo de dados presentes na área da Ciência da Informação.

CICLO DE VIDA DOS DADOS APLICADO ÀS IMAGENS DIGITAIS

Nessa seção apresentamos a aplicação do Ciclo de Vida dos Dados (Sant'Ana, 2016) no contexto das Imagens Digitais, contemplando as mesmas fases, compostas pelos processos de coleta, armazenamento, recuperação e descarte de dados imagéticos evidenciando os diferentes momentos e fatores envolvidos nas particularidades e especificidades dos dados imagéticos.

COLETA DE DADOS IMAGÉTICOS

Em um primeiro momento, depara-se com a fase de obtenção da imagem digital. Esta fase se inicia com a aquisição da imagem digital e após a aquisição do objeto digital, a definição das necessidades informacionais que definirão quais dados imagéticos serão necessários e, portanto, deverão ser extraídos da imagem digital. Os processos ocorrem na ordem em que foram apresentados e constituem a fase de coleta de dados imagéticos.

A aquisição da imagem digital ocorre através do uso de equipamentos para obtenção de imagens digitais ou mecanismos de recuperação. No caso dos equipamentos para obtenção de imagens, podemos citar as câmeras digitais, *smartphones*, sensores, radares, satélites, digitalização 3D, câmeras de banco de dados, equipamentos médicos e odontológicos, etc. No caso dos mecanismos de recuperação, temos os motores de busca de imagens, banco de imagens, redes sociais, bibliotecas e arquivos digitais, software de design gráfico, aplicativos de edição de fotos, programas de modelagem e renderização 3d, *screenshots*, realidade virtual e aumentada, etc.

Além do equipamento ou mecanismo de aquisição da imagem digital, algumas premissas fundamentais devem ser levadas em consideração como a fonte de aquisição da imagem (objeto, pessoa ou lugar), como será a captura, o resultado esperado, se a imagem conseguiu captar os elementos ou dados necessários, o formato da imagem, os riscos de privacidade para indivíduos ou entidades presentes na imagem, permissão de coleta da imagem, procedência, acesso no futuro após armazenamento, constituem um elenco de questões que devem ser abarcadas segundo o CVD.

Na sequência à aquisição da imagem, é necessário definir as necessidades informacionais que nortearão as escolhas e definir quais dados imagéticos são necessários; estabelecer estratégias sobre como localizar, extrair, processar e avaliar estes dados; elaborar as metodologias e ferramentas necessárias para consecução destes dados.

A poli representação da imagem digital (Roa-Martinez, 2019, p. 202) conformada por diferentes tipos de conteúdo que representam a

imagem digital, sendo eles o sintático, o semântico e as informações não visuais e de contexto, os quais durante o pré-processamento dos recursos imagéticos são obtidos por diferentes técnicas ou métodos e armazenados. Essa polir representação pode ser utilizada em ambientes digitais com armazenamento em banco de dados e recuperada por mecanismos de busca com orientação no valor (sintaxe) dos dados, mais do que no significado (semântica) da informação proporcionando a disseminação desses dados.

Após a extração de atributos e características da imagem, é necessária a persistência desses dados em um banco de dados. Essa persistência é realizada através da associação da imagem a metadados, que são elementos descritivos ou atributos referenciais codificados que representam características próprias ou atribuídas às entidades sobre a imagem, como título, palavras-chave, autor, data de criação, entre outros (Alves, 2010, p. 47). Os registros de dados imagéticos são identificados inequivocamente através das chaves primárias ou candidatas, e seus correspondentes nas outras entidades através das chaves estrangeiras, definindo os relacionamentos necessários para que ela possa ser conectada a outras bases, garantindo dessa forma a integração entre diferentes entidades em um banco de dados de imagens (Sant'Ana, 2016, p. 125).

Muitos bancos de dados de imagens gerenciam os direitos autorais e as licenças das imagens. Isso permite que os usuários saibam como podem usar as imagens (por exemplo, para uso comercial, pessoal ou educacional) e quais restrições podem existir.

Um dos fatores a ser analisado na fase de coleta é a privacidade. É fundamental identificar em fontes usadas elementos que possam indicar violações de privacidade de indivíduos ou instituições ligadas aos dados imagéticos coletados. Tal cenário pode levar a responsabilidades futuras a partir da base de dados obtida, prejudicando as etapas subsequentes do ciclo de vida (Sant'Ana, 2016, p. 125).

A preservação dos dados coletados, inclusive os futuramente armazenados, pode requerer a inclusão de dados adicionais nos pré-requisitos de coleta. Isso possibilita uma identificação mais ampla dos dados e incorpora informações sobre as características dos dispositivos de origem. Essas mu-

danças podem incluir melhorias na precisão e granularidade dos dados ao longo do tempo.

ARMAZENAMENTO DE DADOS IMAGÉTICOS

Após a coleta de dados imagéticos, surge a perspectiva de seu uso futuro, incluindo novas análises diretas ou integração com outras bases de dados (Sant'Ana, 2016, p. 132). Esta fase é definida como fase de armazenamento de dados imagéticos.

Além dos dispositivos de armazenamento (disco rígido, *pen drive*, cartão de memória, SSDs, etc.), as premissas fundamentais que devem ser consideradas quando armazenamos uma imagem digital são: qual será a estrutura física e lógica utilizada para seu armazenamento, quais são os parâmetros de amostragem e quantização, quais metadados essas imagens possuem sobre definição e conteúdo, qual o tipo e grau de interferência nos dados causados por algoritmos de pré-processamento de imagens, como será realizado o armazenamento de representações e descrições da imagem, como persistir os dados complementares sobre a coleta para garantir o contexto de sua obtenção, se os dados obtidos através da imagem digital pode representar um risco a privacidade dos indivíduos ou instituições referenciados, se há o direito de armazenar estes dados imagéticos e se os fatores para sua utilização ao longo do tempo estão sendo mantidos.

As imagens digitais são compostas por *pixels*, que são os pontos individuais que formam a imagem. Cada pixel é composto por informações de cor e intensidade, que podem ser representadas usando diferentes modelos de cor, como RGB (vermelho, verde, azul) ou CMYK (ciano, magenta, amarelo, preto). As intensidades de cores são representadas por valores numéricos. Em resumo, imagens digitais são passíveis de serem armazenadas em banco de dados como uma matriz de *pixels* com informações de cor e intensidade.

Além da matriz de *pixels*, os metadados associados a esta imagem, como título, palavras-chave, origem, direitos de uso, informações sobre

a câmera, configurações de exposição, data de criação e autor da imagem também são armazenados.

Nesta fase tem-se um enfoque mais tecnológico e se definem aspectos que garantem a reutilização destes dados, por meio de especificações físicas e lógicas sobre como os dados serão registrados em um suporte. Algumas das definições necessárias nesta fase são:

- i. Definição do dicionário de dados e da estrutura lógica do banco de dados, ou seja, a representação abstrata e estruturada dos dados que serão armazenados em um sistema de gerenciamento de banco de dados (SGBD).
- ii. Definir a estrutura de como o conjunto de variáveis definida no item i precisa ser organizado em subconjuntos definidos de acordo com a semântica que os vincula a um elemento ou conceito do mundo real.
- iii. Definir autorizações de acesso levando em conta o tipo de acesso aos dados e sobre quem acessa esses dados.
- iv. Definir a forma de acesso aos dados.
- v. Definir o formato ou padrão de armazenamento e os metadados.
- vi. Definir o local de armazenamento, a arquitetura do banco de dados (se local, cliente-servidor ou distribuído).

Quanto à qualidade dos dados, as definições sobre seu armazenamento são fundamentais para garantir que estes dados mantenham sua integridade física e lógica.

Ao se armazenar dados, uma das preocupações deve ser com relação aos direitos autorais vinculados à fonte da qual os dados foram obtidos, buscando-se registrar, também, estas informações para que se mantenha a segurança institucional daqueles que respondem pelos dispositivos de armazenamento.

O armazenamento deve se preocupar com a disseminação dos dados, garantindo o acesso e a compreensão dos dados armazenados no fu-

turo. Dessa forma, devem ser incluídos elementos semânticos nos dados imagéticos para facilitar sua interpretação, preferencialmente por meio de processos automatizados. Ainda pensando no acesso futuro, devem ser elaboradas estratégias de execução de processos de atualização tecnológica e de verificação de integridade física e lógica, independentes dos dispositivos originais que o armazenaram, propiciando dessa forma a preservação dos dados imagéticos.

Destarte, ao permitir a disseminação dos dados imagéticos, não devemos esquecer que em meio ao maciço volume desses dados haverá dados sensíveis, e mais uma vez a privacidade deve estar planejada, com ações de proteção previstas no armazenamento.

RECUPERAÇÃO DE DADOS IMAGÉTICOS

Após a coleta e armazenamento dos dados o CVD inicia a fase de acesso e uso dos dados. Começa a fase em que os esforços estão voltados para que esses dados possam ser recuperados a partir da base de dados. Inicia-se a fase da recuperação de dados imagéticos.

Para a fase de recuperação deverá ser considerado o público-alvo, a necessidade informacional deste público, como será o acesso (se por base de dados ou software), os dados imagéticos a serem disponibilizados, quais metadados sobre a imagem serão disponibilizados, quais informações poderão ser obtidas através de algoritmos de processamento de imagens e algoritmos de visão computacional, se a recuperação de dados imagéticos será baseada em conteúdo (CBIR), por mineração de imagens (*image mining*) ou inteligência artificial, a frequência de atualização das imagens digitais para disponibilização, quem poderá acessar estes dados, os riscos à privacidade dos indivíduos ou entidades referenciados pelos conteúdos recuperados, como operacionalizar a integração entre os dados imagéticos e destes com outros conjuntos de dados, como garantir os elementos que sustentam a qualidade dos dados imagéticos que estão sendo disponibilizados, se há o direito de disponibilizar estes dados imagéticos, como viabi-

lizar que os dados imagéticos sejam recuperados (preferencialmente, e em muitos casos obrigatoriamente, por máquinas) e se os processos e procedimentos de recuperação estão estáveis o suficiente para que permaneçam polimorficamente utilizáveis ao longo do tempo.

Os dados coletados a partir de imagens podem ser organizados de várias maneiras, dependendo do contexto e das necessidades específicas, como descritos no Quadro 1.

Quadro 1 – Mecanismos de recuperação de dados imagéticos e sua descrição

| Mecanismos de recuperação | Descrição |
|-------------------------------|--|
| Metadados e Descritores | Cada imagem pode ser associada a metadados, como data, localização, autor e informações de câmera. Além disso, é possível extrair descritores visuais que representam características da imagem, como cor, textura e forma. Esses metadados e descritores auxiliam na categorização e na recuperação eficiente das imagens |
| Categorização e Classificação | As imagens podem ser organizadas em categorias ou classes com base em critérios predefinidos. Isso pode ser feito manualmente ou por meio de algoritmos de aprendizado de máquina. A categorização facilita a navegação e a busca em grandes conjuntos de imagens. |
| Hierarquias e Taxonomias | Se houver uma grande variedade de imagens, é possível criar hierarquias ou taxonomias para organizar as imagens em níveis de agrupamento. Por exemplo, em um banco de imagens de animais, você poderia ter categorias como “Mamíferos”, “Aves” e “Répteis”, e subcategorias mais específicas dentro de cada uma delas. |
| Indexação Visual | Utilizando técnicas de processamento de imagens, é possível indexar as imagens com base em conteúdo visual, como reconhecimento de objetos, faces, cenas etc. Isso permite a recuperação de imagens similares com base em características visuais. |
| Geolocalização | Se as imagens tiverem informações de localização incorporadas (por exemplo, coordenadas GPS), elas podem ser organizadas geograficamente em um mapa. |
| Data e Temporalidade | A organização por data ou cronologia é útil quando se lida com imagens ao longo do tempo, como em estudos de evolução de fenômenos naturais ou mudanças urbanas. |
| Tags e Palavras-Chave | Atribuir tags ou palavras-chave relevantes a cada imagem pode facilitar a busca e a recuperação. Essas tags podem ser adicionadas manualmente ou por meio de processos de identificação automática. |

| Mecanismos de recuperação | Descrição |
|----------------------------------|---|
| Coleções e Projetos | As imagens podem ser agrupadas em coleções ou projetos relacionados. Isso é particularmente útil quando se trabalha com diferentes conjuntos de imagens para diferentes propósitos. |
| Recursos Visuais | Criar miniaturas, pré-visualizações ou até mesmo resumos visuais das imagens pode ajudar na rápida identificação e seleção. |
| Bancos de Dados | Utilizar sistemas de gerenciamento de banco de dados para armazenar e organizar as informações associadas às imagens, permitindo consultas eficientes e estruturadas. |

Fonte: Autoras.

Geralmente, uma combinação de várias dessas abordagens é usada para criar um sistema de organização eficiente e acessível apresentando um grau de integração que propicie análises de entidades distintas.

Na fase de recuperação de dados imagéticos, devemos nos atentar à privacidade, considerando os envolvidos com os conteúdos recuperados, identificando estruturas e possíveis usuários. Também é importante indicar quem está autorizado a usar os dados e as formas permitidas de utilização, constituindo os direitos autorais.

São necessárias estratégias que permitam a localização dos dados imagéticos armazenados, não somente para acesso pelos próprios recursos de visualização de seus detentores, mas, também, por mecanismos automáticos que possam não só encontrá-los como ainda acessá-los em processos de coleta, garantindo dessa forma a disseminação.

A preservação na fase de recuperação garante que uma interpretação realizada em um determinado momento tenha a possibilidade de ser a mesma realizada em outro momento, desde que mantidos os critérios e objetivos originais.

A qualidade na recuperação de dados imagéticos está associada à interação do usuário, como os aspectos de arquitetura da informação na elaboração de recursos de recuperação e de elementos que ampliem a usabilidade e a acessibilidade dos recursos disponibilizados.

DESCARTE DE DADOS IMAGÉTICOS

No CVD também está descrita a fase de descarte de dados, que é a necessidade dos dados de serem excluídos da base de dados, quando estes já não são mais necessários, ou quando seu volume está acima da capacidade de tratá-los com eficiência, constituindo a fase do descarte de dados imagéticos.

O descarte de dados imagéticos deve ocorrer da sua polir representação (conteúdos sintáticos, semântico, informações não visuais e de contexto) e o descarte físico da imagem deve ocorrer em bloco, onde a exclusão será realizada pela exclusão total da imagem (exclusão de todas as linhas e colunas da imagem) pois a exclusão horizontal (de linhas da imagem) ou vertical verticalmente (de colunas da imagem), acarretaria alteração nos dados que compõem a imagem, gerando, conseqüentemente, uma nova imagem digital.

Desta forma, para a fase de descarte deverá ser considerado quais imagens já não são mais necessárias para que possam ser descartadas, se elas foram persistidas, em quais suportes, se estão replicadas em outras bases, como garantir que estes dados imagéticos foram realmente excluídos e não simplesmente ocultados, se a eliminação destes dados imagéticos não prejudicará a integridade ou interligação de outros dados imagéticos, se o descarte destes dados imagéticos não prejudicará a qualidade do conjunto de dados como um todo, se têm-se o direito de excluir este dado imagético, qual o impacto em sua encontrabilidade e acesso caso a imagem seja eliminada e se foi considerada a necessidade de preservação em seus diversos aspectos.

Nesta fase, a questão da privacidade se faz presente quando um indivíduo faz uso dos direitos autorais sobre uma imagem, exigindo que seja retirada de uma determinada base. Esse acesso sempre será mediado por um detentor e não se pode afirmar que a imagem será excluída ou se apenas será identificada como não acessível para visualização. Já com relação à integração dos dados imagéticos no momento do descarte, se uma imagem é excluída seus dados também o são, podendo causar a degeneração de re-

lacionamentos entre bases distintas levando a uma degradação do valor de uso da base como um todo, incidindo diretamente sobre a qualidade já que a possibilidade de relacionar e realizar análises comparativas entre conteúdos excluídos podem levar a conclusões equivocadas ou inexistentes. Por isso, processos de descarte devem manter informações registradas sobre os processos de eliminação.

Na fase de descarte são perdidas informações sobre autoria e seus registros, afetando questões relacionadas a direito autoral, assim, deve-se manter informações sobre dados que já foram disponibilizados e que podem ter sido utilizados por terceiros para não gerar insegurança legal.

Sobre a disseminação e sua relação com a fase de descarte, outra preocupação se refere ao risco de perder informações que, mesmo não estando relacionadas ao foco do conjunto de dados em questão, podem representar elementos-chave para encontrabilidade do conjunto.

A preservação deve ser buscada mesmo quando o dado não parece ser mais útil, já que sempre podem surgir novas necessidades, não previstas, que venham a requerer os dados eliminados.

Em função de custos de armazenamento cada vez menores, eliminar dados a favor da eficiência do sistema torna-se raro. Desta forma, manter uma cópia de dados imagéticos excluídos, muitas das vezes em formatos e estruturas diferentes das originais configura o início de um novo ciclo de vida dos dados para a imagem, já que será uma nova base de dados com suas características e objetivos específicos.

CONCLUSÃO

O contexto atual de disponibilidade de grandes volumes de dados imagéticos, ocasionados por crescentes usos dos mais diversos equipamentos de aquisição, de uma vasta gama de bancos de imagens, softwares de criação de imagens, compartilhamento em redes sociais, aplicativos e ultimamente com a crescente aplicação na Inteligência Artificial, assunto que

deve ser inclusive de interesse de pesquisas na área de CI, requer novos olhares para os processos de acesso e uso de dados imagéticos.

Transitamos por questões e reflexões presentes em cada uma das fases do Ciclo de Vida dos Dados, como a coleta, o armazenamento, a recuperação e o descarte, refletindo sobre os conceitos de privacidade, integração, qualidade, direitos autorais e preservação em cada uma das fases do CVD abarcando questões próprias aos dados imagéticos.

A temática de acesso a dados é ampla, e os assuntos não puderam ser esgotados, principalmente sobre os conceitos de privacidade, integração, qualidade, direitos autorais e preservação dentro de cada fase do CVD.

Objetivou-se criar uma estrutura que facilite as ações relacionadas à aquisição, manutenção e uso de dados sobre as imagens digitais. Essa estrutura busca agrupar elementos similares e disponibilizar teorias e métodos com base em seu âmbito, seja por fase do ciclo de vida dos dados ou de acordo com fatores identificados. Isso permite uma abordagem organizada para lidar com diferentes aspectos dos dados imagéticos em particular, ao longo de seu ciclo de vida. Pretendemos, assim, estreitar caminhos entre usuários e os dados imagéticos que necessitam.

REFERÊNCIAS

- ALVES, R. C. V. **Metadados como elementos do processo de catalogação**. Marília, SP: Universidade Estadual Paulista, 2010.
- BROWN, C.W.; SHEPHERD, B. J. **Graphics File Formats**. 2. ed. Greenwich, UK: Manning, 1995.
- GONZALEZ, R.C.; WOODS, R.C. **Processamento Digital de Imagens**. 3. ed. São Paulo: Pearson Prentice Hall, 2010.
- JOLY, M. **Introdução à análise da imagem**. São Paulo: Papirus, 1996.
- PINHEIRO, L. V. R.; FERREZ, H. D. **Tesouro Brasileiro de Ciência da Informação**. Rio de Janeiro; Brasília: Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), 2014.

ROA-MARTÍNEZ, S. M. **Da information findability à image findability:** aportes da polirrepresentação, recuperação e comportamento de busca. 235f. 2019. Tese (Doutorado em Ciência da Informação) — Universidade Estadual Paulista “Júlio de Mesquita Filho”, Faculdade de Filosofia e Ciências, Marília/SP, 2019. Disponível em: <http://hdl.handle.net/11449/182465>. Acesso em: 10 jul. 2023.

SANT'ANA, R. C. G. Ciclo de vida dos dados: uma perspectiva a partir da ciência da informação. **Informação & Informação**, Londrina, v. 21, n. 2, p. 116-142, 2016.

SANT'ANA, R. C. G.; ASSUMPCÃO, F. S.; SANTOS, P. L. V. A. C. Coleta de dados a partir de imagens: considerações sobre a privacidade dos usuários em redes sociais. **Em Questão**, Porto Alegre, v. 21, n. 2, p. 31-48, maio/ago. 2015. DOI: <http://dx.doi.org/10.19132/1808-5245212.31-48>.

SANTAELLA, L.; NOTH, W. **Imagem:** cognição, semiótica, mídia. São Paulo: Iluminuras, 2015.

SANT'ANA, R. C. G. Ciclo de vida dos dados: uma perspectiva a partir da ciência da informação. **Informação & Informação**, Londrina, v. 21, n. 2, p. 116-142, 2016. DOI: 10.5433/1981-8920.2016v21n2p116. Acesso em: 04 jul. 2023.

SANTOS, J. M. P.; MADIO, T. C. C. **Metadados:** a recuperação de imagens digitais baseada em conteúdo. *In:* ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB, 19., 2018. Disponível em: <https://brapci.inf.br/index.php/res/download/124657>. Acesso em: 09 jul. 2023.

SANTOS, J. M. P.; MADIO, T. C. C. Da câmera escura aos pixels: a importância do tratamento informacional imagético. **Acervo**, Rio de Janeiro, v. 32, n. 2, p. 101–116, 2019. Disponível em: <https://revista.an.gov.br/index.php/revistaacervo/article/view/1252>. Acesso em: 10 jul. 2023.

SIMIONATO, A. C. **Modelagem conceitual DILAM:** princípios descritivos de arquivos, bibliotecas e museus para o recurso imagético digital. 200 f. Tese (Doutorado em Ciência da Informação) — Universidade Estadual Paulista Júlio de Mesquita Filho, Faculdade de Filosofia e Ciências, Marília, 2015.

SIMIONATO, A. C.; PINHO NETO, J. A. S.; SANTOS, P. L. V. A. C. Ciência da informação, imagem e tecnologia. **Informação & Tecnologia**, João Pessoa, v. 2, n. 1, p. 53-65, 2015. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/41086>. Acesso em: 05 jul. 2023.