

Paradigmas contemporâneos para difusão e consulta de artefatos da cultura visual

Contemporary paradigms for the dissemination and consultation of visual culture artifacts /
Paradigmas contemporâneos de difusión y consulta de artefactos de la cultura visual

Júlia Rabetti Giannella

Doutora em Design pela Universidade do Estado do Rio de Janeiro (UERJ). Professora substituta no curso de Comunicação Visual Design da Universidade Federal do Rio de Janeiro (UFRJ).

juliagiannella@gmail.com

RESUMO

Observa-se atualmente substancial debate sobre a atuação de metadados ricos e variados na estruturação, visualização e recuperação de coleções. Mediante a identificação de recentes pesquisas interdisciplinares envolvendo os campos das humanidades digitais, da ciência da informação, do patrimônio cultural e da visualização de dados, este artigo discorre sobre emergente tópico de reflexão e prática: a visualização de dados de coleções culturais.

Palavras-chave: humanidades digitais; patrimônio cultural; visualização de dados; cultura visual.

ABSTRACT

Currently, there is a substantial debate on how rich and varied metadata act on structuring, visualizing, and retrieving collections. By means of the identification of recent interdisciplinary research involving the fields of Digital Humanities, Information Science, Cultural Heritage, and Data Visualization, this article debates an emerging topic of reflection and practice: the visualization of cultural heritage collection data.

Keywords: digital humanities; cultural heritage; data visualization; visual culture.

RESUMEN

Actualmente, existe un debate sustancial sobre cómo los metadatos ricos y variados actúan sobre la estructuración, visualización y recuperación de colecciones. Por medio de la identificación de investigaciones interdisciplinarias recientes que involucran los campos de Humanidades Digitales, Ciencias de la Información, Patrimonio Cultural y Visualización de Datos en sus trabajos, este artículo debate un tema emergente de reflexión y práctica: la visualización de datos de colecciones culturales.

Palabras clave: humanidades digitales; patrimonio cultural; visualización de datos; cultura visual.

Coleções digitais como objeto e meio de análise

Enquanto o legado histórico-cultural é fruto da produção coletiva, sua preservação profissional na contemporaneidade é organizada, na maioria das vezes, por instituições como galerias, bibliotecas, arquivos e museus. Além da manutenção e preservação, essas instituições dedicam-se à digitalização, catalogação, disponibilização e difusão de coleções para apreciação geral e pesquisa.

Visões clássicas em campo historiográfico tendem a compreender clara distinção entre pesquisador e acervo: este forneceria a evidência “neutra” para que aquele o interpretasse, isto é, oferecesse sentido a esse material pelo rastreamento de sua narrativa histórica. Nessa linha de pensamento, arquivos são organizados por princípios básicos como *respects des fonds* (Duchemin, 1983),¹ o que garantiria a manutenção da ordem original dos artefatos e documentos de um arquivo e sugeriria que eles representam uma conservação neutra do passado.

Visões mais recentes, contudo, argumentam que “acervo” e “coleção” são termos repletos de ambiguidade, visto que são construções de indivíduos e instituições que os mantêm, organizam e difundem. A Norma Brasileira de Descrição Arquivística define coleção como um “conjunto de documentos com características comuns, reunidos intencionalmente” e acervo como a “totalidade de documentos de uma entidade custodiadora” (Conarq, 2006). O aspecto intencional, presente na definição de coleção, é o primeiro motivo pelo qual optamos por utilizar neste artigo “coleção” e não “acervo”. A aceitação de autoria e interpretação em agrupamentos de objetos documentais e/ou culturais traz subsídios à discussão a respeito da objetividade e da distância nas relações entre pesquisador e seus dados, instância importante no âmbito geral da historiografia. Ao destacar o tema da subjetividade das mediações, enfatizamos a presença de vieses curatoriais no ato de colecionar independentemente de haver ou não mediação tecnológica.

O segundo motivo determinante da opção por “coleção” é o uso restrito a países ibero-americanos do termo acervo. Segundo a *Enciclopédia Itaú Cultural*, “nas línguas inglesa e francesa, por exemplo, usa-se indiscriminadamente o termo *collection* para designar tanto o primeiro como o segundo sentido” (Acervo..., 2020).

A compreensão de que uma coleção é constituída mediante processos de seleção e agrupamento não é nova. Derrida (2001) já argumentava que o modo

¹ O princípio *respects des fonds* (respeito aos fundos) é um conceito elaborado pelo historiador francês Natalis de Wailly, em 1841, e amplamente difundido na arquivística.

como documentos são organizados interfere diretamente na leitura de quem os consulta. Segundo o autor, uma coleção não é uma seleção neutra e imparcial de objetos, e o ato de colecionar implica vieses curatoriais e subjetivos. Logo, o poder que indivíduos exercem sobre a definição dos critérios de preservação e difusão de uma coleção configura um aspecto importante a ser considerado.

A digitalização de coleções – e sua posterior disponibilização em ambientes digitais para acesso do público – traz outros ingredientes para essa discussão. A digitalização é “a conversão de um sinal ou código analógico em um sinal ou código digital” (Lee, 2000, p. 3). Atualmente considerada atividade habitual, se não obrigatória, na maioria das instituições dedicadas à cultura e à memória, a digitalização, por um lado, atende a requisitos de preservação em caso de perdas ou danos da coleção física original e, por outro, promove a difusão, em diferentes escalas, de coleções por meio de sua disponibilização em ambientes digitais e em rede.

Há defesas do uso preferencial de métodos “tradicionais” em vez dos digitais para investigação de coleções, o que inclui a examinação dos artefatos históricos e culturais com o mínimo de mediação tecnológica. Roussou (2006, p. 267) sustenta que “arqueólogos e estudiosos de áreas afins não reconhecem facilmente a validade de outras fontes para além das fontes primárias com as quais trabalham [...] e olham as inovações tecnológicas com desconfiança”. Se é admissível, no entanto, que evidências históricas contidas em coleções já estão sujeitas a influências individuais e institucionais, talvez seja necessário aceitar o fato de que nem mesmo o pesquisador que investigue uma coleção de artefatos físicos por meio de ferramentas analógicas esteja trabalhando com fontes primárias no sentido mais estrito.

Essa discussão torna-se mais acirrada com respeito a coleções digitais que, além da catalogação de dados, disponibilizam digitalizações de artefatos em alta resolução. Quando imagens em alta resolução permitem que pesquisadores examinem atributos e qualidade de pinceladas, ou a textura de um papel com mais acurácia do que permitiriam os originais físicos e “reais”, a implícita superioridade do artefato físico sobre o digitalizado para fins de examinação é colocada em xeque.

Grande parte dos debates seminais sobre o propósito, o mérito e o escopo de coleções digitais como objeto de análise encontra-se esgotada à medida que há considerável aceitação de instituições em oferecer recursos digitais para diferentes níveis de consulentes – do usuário geral ao pesquisador acadêmico (Deegan; Tanner, 2002). A atenção volta-se cada vez mais para os potenciais atributos que o meio digital oferece às coleções como meio de análise, isto é,

como ferramentas que podem ser empregadas para estender a compreensão sobre os artefatos históricos e culturais que as constituem.

Bibliotecas, arquivos e museus são instituições pioneiras na adoção de tecnologia digital para catalogar o conteúdo de suas coleções. Um catálogo funciona como um *index*: um modo eficiente de buscar um item específico a partir de informação conhecida e armazenada, por exemplo, origem, autoria, data, dimensão etc. Esses descritores são conhecidos como metadados, recorrentemente definidos como “dados sobre dados” (Pomerantz, 2015), ou seja, descrições sobre um dado original. Enquanto fichas catalográficas impressas só podem conter um conjunto limitado e estanque de metadados, suas instâncias digitais armazenam quantidades quase ilimitadas e podem ser atualizadas. Kallinikos, Aaltonem e Marton (2010) chamam essa qualidade do objeto digital de *editabilidade*.

Com o advento do banco de dados, foi possível expandir a criação e o armazenamento de metadados, que passam a descrever informações adicionais, por exemplo, material, estilo, assunto, movimento artístico, cor, ano de aquisição, doador, departamento, número da página, características de deterioração, histórico de utilização, localização física etc. Graças à consolidação de metadados ricos e variados, o catálogo digital torna-se mais que referência para um objeto, fazendo-se um recurso por si só: *um objeto e meio de análise*. O catálogo digital transforma-se em coleção digital, que não é uma reprodução da coleção analógica original. Coleções digitais, portanto, podem ser entendidas como versões digitais de catálogos de indexação e sistemas de busca e não como substitutos digitais de coleções físicas. Tal percepção, acredita-se, contribui para atenuar a confusão a respeito da terminologia em torno de coleções digitais e do processo de digitalização de uma coleção.

Somam-se a essa discussão as possibilidades que as diferentes soluções de interoperabilidade oferecem ao cruzamento de bases de dados. Sobre esse tema, Marcondes (2016) expõe como as tecnologias da web semântica e dos dados abertos interligados permitem a integração de coleções e destaca os novos tipos de relações significativas que podem ser estabelecidas entre objetos documentais e culturais mantidos por diferentes instituições.

Finalmente, a coleção digital não só permite consultas mais eficientes, mas possibilita maneiras complementares de explorar e interpretar artefatos históricos e culturais, promovendo descobertas. Neste artigo, sugerimos enquadrar coleções a partir da noção de objeto digital de Hui (2012), isto é, a partir da percepção simultânea das três fases que constituem um objeto digital: objetos (*objects*), dados (*data*) e redes (*networks*). O autor sugere que objetos podem ser formatados como

dados e que dados são unidades constitutivas das redes. Ele ilustra esses conceitos com o exemplo de uma visualização em estrutura de *rede* na qual o perfil de um usuário de rede social (*objeto*) é formatado na estrutura ontológica FOAF² (*dado*). Aproximando esse modelo conceitual do presente estudo, o objeto corresponde às reproduções digitais da cultura visual (fotografias, pinturas, capas de livro etc.), os dados equivalem à formatação dos objetos em esquemas, ou ontologias que podem ser generalizadas como metadados, e as redes representam as múltiplas configurações e arranjos que tornam as relações entre objetos e dados visíveis. Em última instância, a rede é a apreensão da macroestrutura da coleção. Uma interface gráfica que concretize esse modelo conceitual permite ao investigador transitar do geral (*rede*) ao específico (*objeto*) e vice-versa. A negociação entre visão micro e macroscópica é discutida a seguir.

Visualização de dados de coleções culturais

Nos últimos anos, técnicas e métodos digitais computacionais – mineração de dados textuais, processamento de imagens etc. – têm colaborado para gerar novas formas de se conceber, conduzir e comunicar pesquisas e projetos no campo das ciências humanas, sobretudo estudos cujo *corpus* seja grande. Nesse cenário, nota-se o crescente interesse no enquadramento conceitual de “coleções como dados” (Padilla, 2018), abordagem que consiste no emprego de métodos computacionais para acessar, processar, visualizar e reinterpretar coleções, sem necessariamente substituir métodos clássicos e qualitativos, como a leitura próxima. Leitura próxima ou leitura atenta (*close reading*) são expressões originadas na crítica literária para designar a interpretação cuidadosa de textos, priorizando aspectos particulares aos gerais. Embora remeta à linguagem escrita, a leitura próxima pode ser evidenciada na leitura de imagens. A análise iconográfica-iconológica proposta por Panofsky (1939), por exemplo, pode ser enquadrada como método de leitura próxima, posto que formaliza três níveis de interpretação de imagens dentro do prisma da história cultural: o pré-iconográfico, o iconográfico e o iconológico.

Franco Moretti (2000) introduziu a noção de leitura distante (*distant reading*) como prática de negociação entre métodos quantitativos e qualitativos para estudar literatura. Posteriormente, ela passou a ser utilizada no âmbito mais amplo da investigação em ciências humanas. Leitura distante é o uso de

² FOAF (acrônimo de *friend of a friend*) é uma ontologia, passível de ser lida por máquinas, que descreve pessoas, suas atividades e suas relações com outras pessoas e objetos.

métodos computacionais para analisar grandes coleções de objetos históricos e culturais a fim de obter apreensão micro e macroscópica do *corpus*. Moretti (2000, p. 57) destaca: “distância [...] é uma condição do conhecimento que lhe permite concentrar-se em unidades muito menores ou muito maiores do que o texto: dispositivos, temas, tropos – ou gêneros e sistemas”.

As tecnologias digitais possibilitam o processamento de conteúdos mais extensos pelo método de leitura distante. A leitura próxima exige envolvimento mais profundo e atento com o conteúdo, mas a seleção de uma pequena amostra de trabalho não é necessariamente uma escolha deliberada. O argumento de Moretti se baseia na constatação de que muitos projetos afirmam estudar a literatura em determinado nível de escala (global, nacional, regional etc.), quando, na verdade, lidam com pequena amostra da literatura existente devido a limitações das capacidades de processamento e análise humanas.³ Em contraste, a leitura distante permite ao pesquisador estudar um acervo em sua totalidade e a visualização de dados, compreendida como representações gráficas e abstratas (linhas do tempo, mapas, gráficos e diagramas), auxilia no processo de leitura distante.

O que denominamos *visualização de dados de coleções culturais* pode ser compreendido como uma emergente abordagem de pesquisa, interessada no desenvolvimento de ferramentas digitais baseadas em recursos de visualização de dados, para ampliar o acesso a coleções de valor histórico e/ou cultural e apoiar tanto sua análise acadêmica como sua apreciação casual. Os resultados alcançados com essa abordagem – os produtos e as reflexões que eles proporcionam – são relevantes para uma variedade de indivíduos: pesquisadores e profissionais na área de visualização; historiadores e estudiosos da cultura; acadêmicos na área de humanidades digitais; proprietários, curadores e guardiões de coleções; e apreciadores e curiosos da cultura. Em contraste às ferramentas convencionais para acessar coleções digitais (geralmente centradas na pesquisa por palavras-chave), essa abordagem alternativa é baseada na noção de *interfaces generosas* (Whitelaw, 2015), comentada a seguir.

A investigação na área de procura de informação se interessa pelo processo de busca do ponto de vista centrado no ser humano (Bates, 1989) e, embora relacionada, possui uma abordagem diferente dos estudos na área de recuperação de informação (Dörk, 2012, p. 8). Apesar do emprego corrente de expressões casuais como “surfando na internet”, a busca de informações é predominantemente

³ Moretti, contudo, não esteve isento de críticas. Historiadores de linhas de pensamento mais clássicas expressaram dúvidas sobre os novos métodos digitais, desafiando-os como reducionistas, frágeis e não pertinentes ao objetivo principal da narrativa histórica (Anderson, 2007).

tarefa assertiva, na qual precisão e eficiência são fatores cruciais. A pesquisa tradicional – inserção de palavras-chave pelo usuário e a consequente recuperação de resultados pelo sistema – constitui paradigma dominante de busca.

Essa evidência possui herança profunda na história da computação e, especialmente, nos campos da biblioteconomia e da ciência da informação. Como consequência, a pesquisa por palavras-chave está enraizada nos sistemas de consulta de coleções, mas não livre de posições contrárias. O trabalho fundamental de Taylor (1962) sobre as necessidades de informação argumenta que *query*⁴ é expressão comprometida e imperfeita de um sentimento que surge como “vaga insatisfação”.

Atualmente, há crescente interesse no projeto de sistemas para coleções cujas interfaces ofereçam uma experiência mais rica para pesquisa exploratória a partir de visualizações. Nessa linha de pensamento, as contribuições de Mitchell Whitelaw são representativas. Em 2015, o autor publicou um artigo defendendo uma abordagem “generosa” para interfaces destinadas a coleções no âmbito cultural. Com efeito, esse paradigma estende conceitos como *berrypicking*⁵ (Bates, 1989) e *flâneur de informação*⁶ (Dörk, Carpendale, Williamson, 2011) e representa modelos de interface nos quais a busca de informação não é apenas tarefa direcionada por objetivos específicos, mas, principalmente, atividade exploratória que resulta em descobertas afortunadas:

Neste artigo, argumento que a pesquisa textual, como estratégia dominante de navegação em coleções culturais, é inadequada. A pesquisa por palavra-chave é pouco generosa: exige consulta, desestimula a exploração e retém mais do que fornece [...]. Interfaces generosas fornecem representações ricas e navegáveis para grandes coleções digitais; proporcionam exploração e apoiam navegação casual, fazendo uso de visões gerais para estabelecer contexto e manter orientação, e visões detalhadas em várias escalas. (Whitelaw, 2015, p. 13)

O projeto de interfaces para visualização de dados de coleções culturais implica a articulação de dois princípios que guiam seu design e sua fruição: granularidade visual e múltiplos pontos de acesso.

4 Em computação, *query* é uma solicitação de informações feita a um banco de dados.

5 Analogia, em inglês, ao ato de colher mirtilos, um a um, na floresta. A metáfora da colheita representa busca por informação contínua, casual e com diversas paradas (Bates, 1989, p. 410).

6 Tradução de *information flâneur*. Dörk (2012, p. 208) explica que tomou o termo *flâneur* – errante, vadio – emprestado dos estudos culturais, de modo a enfatizar um “modelo positivo de busca de informações baseado em perspectivas centradas no ser humano sobre práticas de informação”.

Granularidade visual

Primeiramente, deve-se compreender de que maneira noções já discutidas de leitura distante e leitura próxima podem ser aplicadas ao processo de visualização. Bertin (2010) já mencionava a existência de diferentes níveis de leitura que uma informação gráfica pode oferecer: elementar, intermediário e geral. Foi Shneiderman (1996, p. 1), porém, que, no domínio das interfaces gráficas, consagrou a existência da visão geral, da visão detalhada e de outras tarefas na interação com visualizações ao propor o mantra “visão geral primeiro, zoom e filtro, depois detalhes sob demanda”. Greene et al. (2000) introduziram a distinção entre visualizações de tipo *preview* (substitutos visuais para objetos únicos) e de tipo *overview* (substitutos visuais para coleções inteiras) com o objetivo de conceituar o espaço informacional de interfaces de consulta para bibliotecas digitais. Essa distinção em dois níveis – objeto e coleção – equivale a uma passagem de granularidade no sentido estrito abordado na modelagem dimensional de Kimball e Ross (2013). Finalmente, Windhager et al. (2019) ampliaram essa distinção ao reconhecer níveis intermediários de leitura visual, que sugerem uma ordem gradativa de abstração, isto é, variações de um nível mais detalhado e próximo para um nível mais geral, sumarizado e distante.

Windhager et al. (2019) chamaram de *granularidade visual* esses diferentes níveis. Não há evidências de que os autores tenham se inspirado no conceito original de *granularidade* de Kimball e Ross (2013), mas a associação é inevitável. De toda forma, *granularidade* e *granularidade visual* não devem ser confundidas. Enquanto o termo *granularidade* utilizado isoladamente remete ao nível de agregação em nível conceitual (objeto e coleção), *granularidade visual* remete ao nível de agregação no domínio visivo. Windhager et al. (2019, p. 6) esclarecem: “sistemas de CH Infovis⁷ podem oferecer acesso a detalhes de artefatos individuais ou a visões gerais de coleções inteiras – ou a qualquer outro nível intermediário de agregação visual, ao qual chamamos granularidade visual”.

Os autores sugerem quatro níveis de granularidade visual. A exibição de um único objeto (*single-object preview*) corresponde a uma inspeção detalhada e individual de um registro, geralmente acompanhado de recursos de magnificação e ficha catalográfica. Em seguida, há a visão multiobjeto (*multi-object preview*), que representa uma seleção de objetos geralmente arranjados em forma de grade, na qual miniaturas permitem pré-visualizar a reprodução digital. O nível de visão geral da coleção (*collection overview*) utiliza diferentes estruturas

7 Abreviação usada pelos autores para designar *visualization of cultural heritage collection*.

de visualização de dados (linha do tempo, mapas, redes etc.) codificadas a partir de metadados disponíveis. Esse nível subdivide-se em dois, mas, neste artigo, serão tratados de modo unificado durante a análise das ferramentas.

Múltiplos pontos de acesso

Outro importante aspecto para discussão quando se fala em visualização de dados de coleções culturais é a possibilidade de conjugar diferentes pontos de acesso. Ao contrário de exposições físicas, nas quais uma obra ou documento é normalmente localizado em uma única parede ou *display*, as reproduções de coleções digitalizadas podem ser organizadas em diferentes configurações. Nesse sentido, autores sugerem que a complexidade de uma coleção não pode ser destilada em uma única visualização de dados, e a solução de múltiplas visualizações acessadas em diferentes telas permite perspectivas multifacetadas e complementares para exploração e análise de objetos culturais (Dörk, Pietsch, Credico, 2017).

A solução para tal desafio reside na multidimensionalidade subjacente aos objetos digitais, isto é, na disponibilidade de metadados ricos e variados por meio dos quais coleções podem ser organizadas e mapeadas em diferentes configurações visuais. Segundo Kimball e Ross (2013, p. 40), “as dimensões fornecem o contexto ‘quem, o que, onde, quando, por que e como’ em torno de um evento”. Com a escolha de cada dimensão, é possível atribuir valores discretos – na forma de listas controladas – ou valores contínuos – como números – capazes de descrever cada dimensão. No contexto de coleções culturais, as dimensões dependem dos metadados associados às reproduções digitais. Com efeito, podemos falar de dimensão temporal, espacial, temática, entre outras, que permitem apreender a organização e a variação da coleção segundo diferentes parâmetros de análise (data, local, cor, autoria, estilo etc.).

Ferramentas para visualização de dados de coleções culturais

Na perspectiva de visualização de dados de coleções culturais, surgem iniciativas teórico-práticas interessadas na concepção e no desenvolvimento de ferramentas digitais baseadas em recursos de visualização para melhorar o acesso, a exploração e a análise de coleções histórico-culturais. Nesta última seção, foram revisadas cinco ferramentas existentes e acessíveis via navegador da web com o objetivo de melhor ilustrar os princípios previamente definidos: granularidade visual e múltiplos pontos de acesso. A seleção faz parte de um *corpus*

mais amplo (Giannella, 2020) e justifica-se por representar uma variedade de manifestações da cultura visual (gravuras, fotografias, desenhos, mapas, pinturas e páginas ilustradas). Embora, no âmbito digital, objetos culturais tenham geralmente uma imagem a eles associada, é no domínio da cultura visual que a presença de reproduções digitais – em suas versões miniaturizadas e em alta resolução – se torna imprescindível para a exploração e a análise mais significativas da coleção.

Australian Prints and Printmaking⁸ é uma aplicação web lançada em 2014 por Mitchell Whitelaw e Ben Ennis Butler para uma coleção de cerca de quarenta mil impressões e gravuras da National Gallery of Australia. Originalmente desenvolvida para fornecer “três interfaces ‘generosas’ separadas, oferecendo múltiplas visualizações complementares” (Whitelaw, 2015), atualmente os consulentes podem explorar a coleção por cinco visualizações e cada uma apresenta diferentes granularidades visuais. Por exemplo, a visualização “resumo da década” (*decade summary*) organiza as obras por década (dimensão temporal) e tipos de impressão (dimensão categórica – entalhe, estêncil, relevo – relativa à especificidade do objeto) em um gráfico de barras que corresponde a uma granularidade de visão geral da coleção. O gráfico resultante apresenta o padrão cronológico da coleção e a distribuição relativa de diferentes tipos de impressão ao longo do tempo. Ele também funciona como um filtro de seleção de obras que são exibidas no painel à direita na forma de uma grade de miniaturas que corresponde à granularidade de visão multiobjeto. Finalmente, a ação de *mouse over* em cima de uma miniatura revela a imagem completa da obra (acompanhada de dados como título, autoria e ano) que corresponde à visão de um único objeto.

8 Disponível em: <http://www.printsandprintmaking.gov.au/explore/>. Acesso em: 20 out. 2019.



Figura 1 - Visualização de “resumo da década”, Australian Prints and Printmaking. Granularidades visuais: (a) visão geral da coleção; (b) visão multiobjeto; (c) visão de um único objeto. Fonte: criada pela autora, adaptado de Whitelaw (2015).

Photogrammar⁹ é uma aplicação web, lançada em 2014 por pesquisadores das universidades de Richmond, Stanford e Yale, para uma coleção de cerca 170 mil fotografias do Arquivo de Segurança e Informação de Guerra dos Estados Unidos (FSA-OWI),¹⁰ registradas entre 1935 e 1945, período que cobre a Grande Depressão e a Segunda Guerra Mundial (Arnold et al., 2020). O objetivo da FSA-OWI, ao decidir documentar a América nesse período, foi divulgar a administração de seus serviços governamentais. Fotógrafos foram então contratados e enviados para regiões norte-americanas. A coleção de fato cobre amplo espectro territorial dos Estados Unidos e constitui importante registro visual, tanto para acadêmicos como para o público em geral.

Na versão 3.0 da ferramenta, atualizada em 2020, os usuários acessam a coleção por meio de três visualizações de dados que permitem duas principais dimensões de análise: temática e espacial. A análise espacial da coleção fotográfica, por exemplo, é permitida pela visualização em estrutura de mapa coroplético, que exhibe a quantidade de fotografias registradas por condado (visão

⁹ Disponível em: <https://www.photogrammar.org/>. Acesso em: 20 out. 2019.

¹⁰ United States Farm Security Administration and Office of War Information (FSA-OWI) Archive.

geral da coleção). Uma segunda visualização de dados, no inferior da página, complementa a leitura distante da coleção, oferecendo outras dimensões – autorial e temporal – que podem ser relacionadas à análise espacial. As visualizações são interligadas entre si e permitem filtrar as obras exibidas no painel de miniaturas à esquerda (visão multiobjeto). Para acessar a fotografia em nível de detalhamento (visão de um único objeto), é necessário clicar no objeto que será carregado por cima das visualizações de dados.

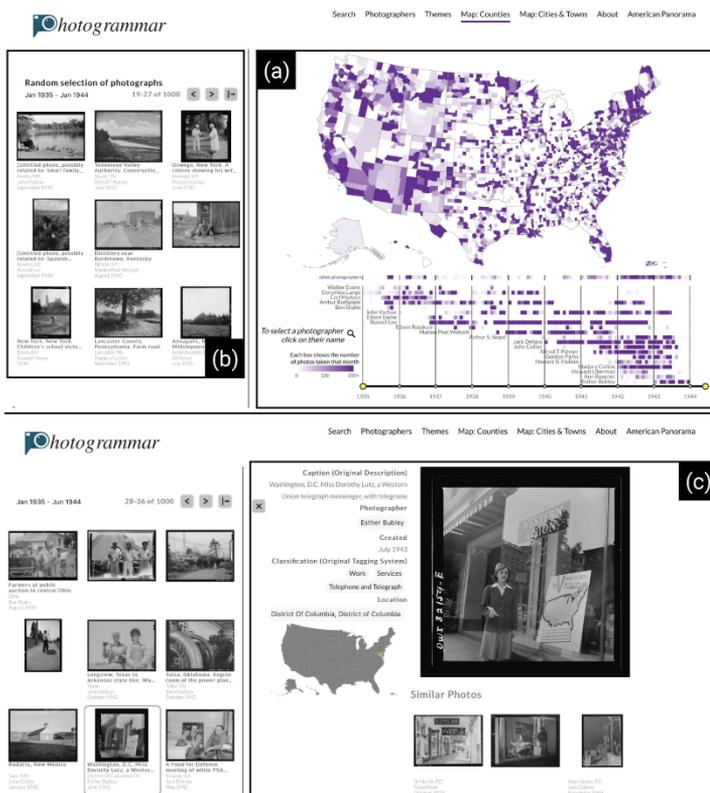


Figura 2 - Visualização mapa coroplético, Photogrammar. Granularidades visuais: (a) visão geral da coleção; (b) visão multiobjeto; (c) visão de um único objeto. Fonte: criada pela autora, adaptado de Arnold et al., 2020.

NYPL Public Domain Visualization¹¹ é uma aplicação web lançada em 2016 para uma coleção de cerca de 187 mil objetos culturais – cartões-postais,

11 Disponível em: <https://publicdomain.nypl.org/pd-visualization>. Acesso em: 20 out. 2019.

gravuras, fotografias, ilustrações etc. – mantidas pela Biblioteca Pública de Nova Iorque (NYPL) (Lafrance, 2016). A mesma técnica de visualização (grade) é utilizada para organizar fotografias de quatro maneiras distintas, que correspondem a diferentes dimensões categóricas de análise: século de criação, gênero, coleção e cor.¹² A visualização das miniaturas em grade permite uma visão multiobjeto da coleção, exibindo-as agrupadas por classes. Na dimensão cor, as classes são: vermelho, verde, azul, laranja, amarelo, marrom, cinza, preto e desconhecido. Um número absoluto ao lado do nome da classe quantifica os objetos, mas não é possível obter uma visão relativa da distribuição por classes. Como a visualização em grade é bastante vertical, causando a rolagem da página, há um componente auxiliar, localizado à direita da interface, que oferece uma noção do “comprimento” da coleção e funciona como uma espécie de mapa de navegação da visão multiobjeto. Por fim, é possível detalhar as fotografias dentro da visão multiobjeto acessando-as individualmente (visão de um único objeto). Ao passar o *mouse* sobre a miniatura, ela é magnificada e, ao clicá-la, o usuário é direcionado para nova página com mais detalhes. Essa ferramenta não oferece granularidade de visão geral da coleção.

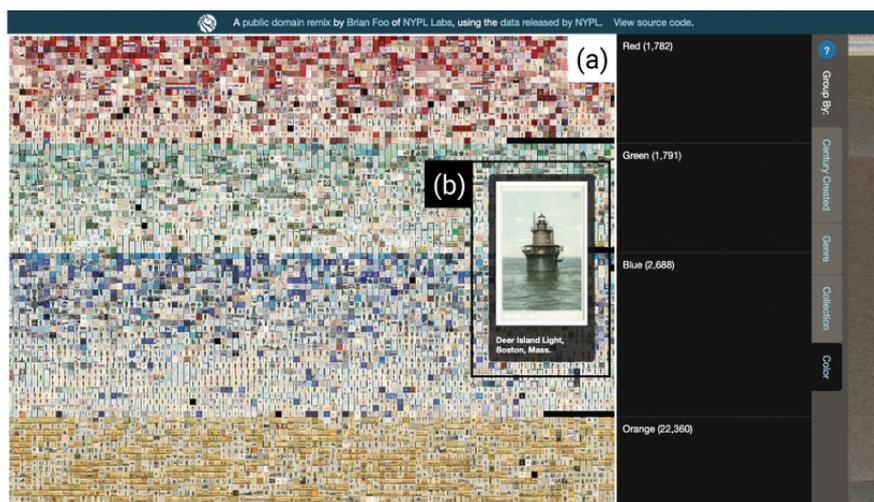


Figura 3 - Visualização em grade organizada por cor, NYPL Public Domain Visualization. Granularidades visuais: (a) visão multiobjeto; (b) visão de um único objeto. Fonte: criada pela autora, a partir de captura de tela.

¹² Cor dominante extraída por meio de um script para processamento de imagem.

Diferente das interfaces mostradas até aqui, Vikus Viewer¹³ é uma plataforma para criação de visualização de dados de coleções culturais. Desenvolvida por Christopher Pietsch entre 2017 e 2019, como parte de um projeto de pesquisa do Urban Complexity Lab da Universidade de Potsdam (Alemanha), organiza milhares de objetos culturais em uma tela e apoia a exploração de padrões temáticos e temporais de grandes coleções, além de fornecer acesso rápido a imagens em alta resolução. A plataforma foi instanciada com diferentes coleções, sendo a primeira a de 1.492 desenhos elaborados pelo rei Frederick William IV da Prússia (1795-1861)¹⁴ (Glinka; Pietsch; Dörk, 2017). A plataforma permite criar três visualizações independentes – linha do tempo, grade e t-SNE¹⁵ – com consistência. Na visualização de linha de tempo, a interface exibe uma configuração principal centralizada na tela, que se comporta ao mesmo tempo como visão geral da coleção (pois visualiza a distribuição dos objetos ao longo de datas de criação) e visão multiobjeto (pois os gráficos de barra são construídos a partir do empilhamento de miniaturas). Além disso, uma visualização auxiliar na estrutura de nuvem de palavras exibe palavras-chave associadas aos objetos de acordo com sua incidência. Assim como em Photogrammar, as visualizações são interligadas entre si e permitem filtrar as miniaturas exibidas na linha do tempo. A linha do tempo é implementada por meio de interface *zoomable*, que permite ao usuário se aproximar das miniaturas, transitando de uma visão geral e estrutural da coleção para uma visão próxima e detalhada da imagem (visão de um único objeto) em diferentes escalas.

13 Disponível em: <https://vikusviewer.fh-potsdam.de/>. Acesso em: 20 out. 2019.

14 Coleção mantida pela fundação alemã Stiftung Preußische Schlösser und Gärten Berlin-Brandenburg.

15 t-SNE (t-distributed stochastic neighbor embedding) é uma técnica de redução de dimensionalidade particularmente bem adaptada para visualizar bancos de dados com muitas dimensões. Basicamente, um t-SNE gera um arranjo para imagens a partir da identificação de similaridades entre elas. Imagens mais semelhantes entre si se aglomeram; imagens mais diferentes se distanciam.

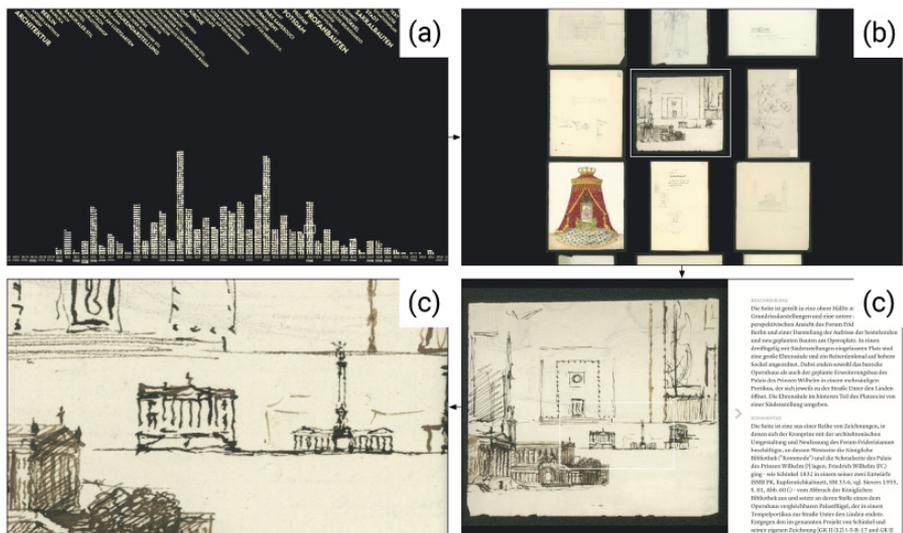


Figura 4 - Visualização de linha de tempo, Vikus Viewer. Granularidades visuais: (a) visão geral da coleção; (b) visão multiobjeto; (c) visão de um único objeto. Fonte: criado pela autora, adaptado de Glinka, Pietsch, Dörk (2017).

Codex Atlanticus¹⁶ é uma aplicação web lançada em 2019 para o Codice Atlantico, a coleção de documentos de Leonardo da Vinci, mantida pela biblioteca Ambrosiana, em Milão. Trata-se do conjunto de 1.119 páginas (frente e verso) que datam de 1478 a 1519 e abrangem grande variedade de assuntos, como anatomia, astronomia, botânica, química, geografia, matemática, mecânica etc. Algumas décadas após a morte de Leonardo, muitos de seus cadernos foram adquiridos pelo escultor Pompeo Leoni, que reorganizou e editou grande parte do material, recortando, colando desenhos e tentando classificar por tema, e posteriormente vendeu várias coleções do trabalho de da Vinci. O Codice Atlantico é a coleção mais relevante desse processo de edição conduzido por Leoni (Rendgen, 2019).

Comissionada pela biblioteca Ambrosiana e desenvolvida pelo estúdio italiano The Visual Agency, a aplicação possui visualização única em grade, que organiza as páginas por ordem de aparição no códex e ordem cronológica (data de criação). As miniaturas da grade não são reproduções das páginas, e sim retângulos cujas cores identificam os assuntos nelas contidos: “É possível ver facilmente alguns agrupamentos, como a intensa concentração de tópicos

16 Disponível em: <https://codex-atlanticus.it/>. Acesso em: 20 out. 2019.

relacionados a ferramentas e máquinas no início e novamente no final [...]. Esta é uma maneira realmente poderosa de fornecer uma visão geral de relance” (Rendgen, 2019). Assim, a visualização em grade oferece uma visão estrutural (visão geral da coleção) em termos temáticos (assuntos e tópicos) e temporal (1478 a 1518). É acompanhada de visualizações auxiliares, que permitem filtrar e relacionar essas dimensões de análise. Além disso, é possível transitar com facilidade entre leitura distante e leitura próxima: ao clicar na miniatura da página, o usuário é direcionado à nova tela para exame das páginas e obtenção de informações contextuais (visão de um único objeto). Codex Atlanticus não oferece visão de nível multiobjeto.

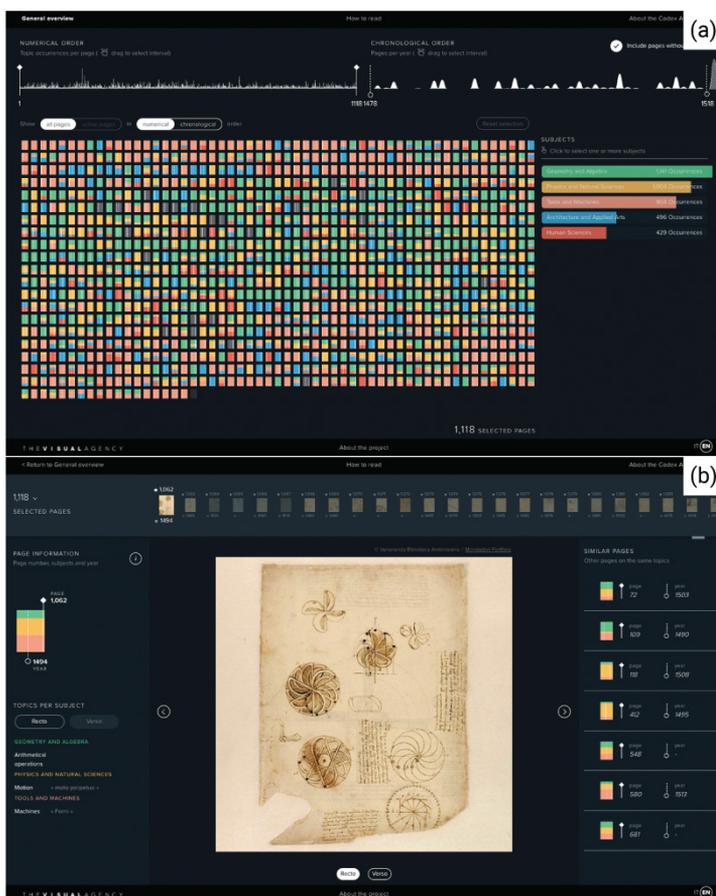


Figura 5 - Visualização em grade (a) e interface de inspeção detalhada (b), Codex Atlanticus. Granularidades visuais: (a) visão geral da coleção; (b) visão de um único objeto. Fonte: criado pela autora a partir de captura de tela.

Considerações finais

Pensar em coleções a partir do conceito de objeto digital (Hui, 2012) e do viés dos dados (Padilla, 2018) inaugura a oportunidade de aplicar métodos digitais a esses recursos, incorporando o processo de visualização de dados, que consiste no uso de representações visuais para apreensão da visão macroscópica de eventos e fenômenos, incluindo a evidenciação de padrões, tendências e anomalias.

Por meio da identificação de recentes pesquisas interdisciplinares envolvendo os campos das humanidades digitais, da ciência da informação e do patrimônio cultural, apresentamos e refletimos sobre a prática de visualização de dados de coleções culturais. O principal desafio que emergiu da revisão de literatura diz respeito à dificuldade de apreensão da complexidade e da riqueza de coleções, a partir de uma perspectiva humanística interessada tanto em generalizações como nas particularidades dos artefatos. As tradicionais ferramentas de consulta com base em pesquisa por palavra-chave mostram-se restritas e insuficientes. Como alternativa, destacamos a concepção de espaços informacionais guiados pelos princípios de granularidade visual e múltiplos pontos de acesso que, juntos, oferecem uma resposta a essas limitações.

O exame de ferramentas para visualização de dados de coleções culturais permitiu ilustrar esses princípios aplicados à consulta de artefatos histórico-culturais em cenários concretos de uso. A pesquisa também proporcionou observar diferentes modelos colaborativos na execução das cinco ferramentas analisadas: duas abordam empreendimentos investigativos conduzidos dentro da academia (Photogrammar e Vikus Viewer); uma origina-se de trabalho concebido e executado na instituição mantenedora (NYPL Public Domain Visualization); uma é fruto de realização comissionada pela instituição e executada por parceiro externo (Codex Atlanticus); e uma resulta de trabalho colaborativo entre a instituição e a academia (Australian Prints and Printmaking).

Referências

- ACERVO e Coleção. In: *Enciclopédia Itaú Cultural*. São Paulo: Itaú Cultural. Disponível em: <http://enciclopedia.itaucultural.org.br/termo14329/acervo-e-colecao>. Acesso em: 17 out. 2019.
- ANDERSON, M. Quantitative history. In: OUTHWAITE, W; TURNER, S. (ed.). *The sage handbook of social science methodology*. London: Sage, 2007.
- ARNOLD, T. et al. Visualizing a large spatiotemporal collection of historic photography with a generous interface. *arXiv*, 4 Sep. 2020.
- BATES, M. The design of browsing and berrypicking techniques for the online search interface. *Online Review*, Bradford, v. 13, n. 5, p. 407-424, 1989.
- BERTIN, J. *Semiology of graphics: diagrams, networks, maps*. California: Esri Press, 2010.
- CONSELHO NACIONAL DE ARQUIVOS (Conarq).

- Norma brasileira de descrição arquivística (Nobrade). Rio de Janeiro: Conarq, 2006.
- DEEGAN, M.; TANNER, S. The digital dark ages: digital preservation. *Library + Information Update*, London, v. 1, n. 2, p. 42-43, 2002.
- DERRIDA, J. *Mal de arquivo: uma impressão freudiana*. Rio de Janeiro: Relume Dumará, 2001.
- DÖRK, M.; PIETSCH, C.; CREDICO, G. One view is not enough: high-level visualizations of a large cultural collection. *Information Design Journal*, Amsterdam, v. 23, n. 1, p. 39-47, 2017.
- _____. *Visualization for search: exploring complex and dynamic information spaces*. Dissertation (Doctor of Philosophy) – Department of Computer Science, University of Calgary, Calgary, Canada, 2012.
- _____; CARPENDALE, S.; WILLIAMSON, C. The information flâneur: a fresh look at information seeking. In: SIGCHI CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS, 2011. *Proceedings...* 2011. p. 1.215-1.224.
- DUCHEIN, M. Theoretical principles and practical problems of respect des fonds in archival science. *Archivaria*, Ottawa, v. 16, 1983.
- GIANNELLA, J. R. *Design de interfaces para visualização, exploração e análise de coleções fotográficas*. 2020. Tese (Doutorado em Design) – Escola Superior de Desenho Industrial, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2020.
- GLINKA, K.; PIETSCH, C.; DÖRK, M. Past visions and reconciling views: visualizing time, texture and themes in cultural collections. *Digital Humanities Quarterly*, Boston, v. 11, n. 2, 2017.
- GREENE, S. et al. Previews and overviews in digital libraries: designing surrogates to support visual information seeking. *Journal of the American Society for Information Science*, Chapel Hill, v. 51, n. 4, p. 380-393, 2000.
- HUI, Y. What is a Digital Object? *Metaphilosophy*, v. 43, n. 4, p. 380-395, 2012.
- KALLINIKOS, J.; AALTONEN, A.; MARTON, A. A theory of digital objects. *First Monday*, v. 15, n. 6, 2010.
- KIMBALL, R.; ROSS, M. *The data warehouse toolkit: the definitive guide to dimensional modeling*. 3rd. Indianapolis: John Wiley & Sons, Inc., 2013.
- LAFRANCE, A. The New York Public Library's Remarkable New Digital Collection. *The Atlantic*, 2016. Disponível em: <https://www.theatlantic.com/technology/archive/2016/01/internet-rejoice/422883/>. Acesso em: 17 out. 2019.
- LEE, H.-L. What is a collection? *Journal of the American Society for Information Science*, Chapel Hill, v. 51, n. 12, p. 1.106-1.113, 2000.
- MARCONDES, C. H. Interoperabilidade entre acervos digitais de arquivos, bibliotecas e museus: potencialidades das tecnologias de dados abertos interligados. *Perspectivas em Ciência da Informação*, Belo Horizonte, v. 21, n. 2, p. 61-83, 2016.
- MORETTI, F. Conjectures on world literature. *New Left Review*, London, n. 1, p. 54-68, 2000.
- PADILLA, T. Collections as data: implications for enclosure. *College & Research Libraries News*, Chicago, v. 79, n. 6, 2018.
- PANOFSKY, E. *Studies in iconology*. New York: Oxford UP, 1939.
- POMERANTZ, J. *Metadata*. Cambridge, London: MIT Press, 2015.
- RENDGEN, S. *Interactive: Leonardo's Codex Atlanticus (The Visual Agency)*. 2019. Disponível em: <https://sandrendgen.wordpress.com/2019/05/15/interactive-leonardos-codex-atlanticus-the-visual-agency-2/>. Acesso em: 17 out. 2019.
- ROUSSOU, M. The components of engagement in virtual heritage environments. In: CONFERENCE ON CULTURAL HERITAGE AND NEW MEDIA, 2006. *Proceedings of new heritage: beyond verisimilitude*. 2006. p. 265-293.
- SHNEIDERMAN, B. The eyes have it: a task by data type taxonomy for information visualizations. In: IEEE SYMPOSIUM ON VISUAL LANGUAGES, 1996. *Proceedings...* 1996. p. 336-343.
- TAYLOR, R. The process of asking questions. *American Documentation*, Chapel Hill, v. 13, n. 4, p. 391-396, 1962.
- WHITELAW, M. Generous interfaces for digital cultural collections. *Digital Humanities Quarterly*, Boston, v. 9, n. 1, 2015.
- WINDHAGER, F. et al. Visualization of cultural heritage collection data: state of the art and future challenges. *IEEE Transactions on Visualization and Computer Graphics*, New York, v. 26, n. 6, p. 2.311-2.330, 2019.

Recebido em 15/1/2021

Aprovado em 9/9/2021