

APLICAÇÃO DE SOFTWARE DE MINERAÇÃO DE TEXTO NA REPRESENTAÇÃO DA INFORMAÇÃO DE OBRAS ARTÍSTICO-PICTÓRICAS

Wagner Oliveira de Medeiros¹
UFPE.

Fabio Assis Pinho²
UFPE

Renato Fernandes Correa³
UFPE

Resumo

Esta pesquisa discute a aplicação da tecnologia da informação nos processos de representação da informação de obras artístico-pictóricas e teve por objetivo geral apresentar uma relação possível entre os softwares de mineração de textos e a sistematização de termos nas práticas de representação da informação a partir da documentação gerada pela análise de obras artístico-pictóricas pelo viés da Ciência da Informação. Para isso, a pesquisa consistiu de estudo de caso, utilizando o software online Sobek para processar uma amostra da documentação gerada pelas análises de obras artístico-pictóricas, a partir de três etapas: a) organização textual do conteúdo documental; b) inserção do conteúdo textual organizado no software Sobek e geração do grafo e; c) análise do grafo gerado pelo software Sobek. Os resultados indicaram que a utilização do Sobek demonstra indícios da viabilidade de uso da mineração de texto como auxílio aos processos de representação da informação de obras artístico-pictóricas, não obstante apontaram a necessidade de aprofundamento da questão para realização de futuras propostas metodológicas e de aperfeiçoamento do software a este fim documentário.

Palavras-chave: Obras artístico-pictóricas. Representação da Informação. Mineração de Textos. Mapas Conceituais. Mapas Mentais.

APPLICATION OF TEXT MINING SOFTWARE IN THE REPRESENTATION OF THE INFORMATION OF ARTISTIC-PICTURE WORKS

Abstract

This research discusses the application of information technology in the processes of information representation of artistic and pictorial works and aimed to present a possible relationship between text mining software and the systematization of terms in information representation practices from documentation generated by the analysis of artistic-pictorial works by Information Science view. For this, the research consisted of a case study, using the online software Sobek to process a sample of the documentation generated by the analysis of pictorial works, from three steps: a) textual organization of documentary content; b) insertion of the textual content organized in the Sobek software and generation of the graph and; c) analysis of the graph generated by the Sobek software. The results indicated that the use of Sobek shows evidence of the viability of using text mining as an aid to the processes of information representation of artistic and pictorial works, despite the need to deepen the issue for future methodological proposals and improvement of software to this documentary purpose.

Keywords: Artistic-pictorial works. Representation of Information. Text mining. Conceptual Maps. Mental maps.

¹ Doutorando em Ciência da Informação pelo PPGCI/UFPE.

² Docente do PPGCI/UFPE. Doutor em Ciência da Informação.

³ Docente do PPGCI/UFPE. Doutor em Ciência da Computação.



1 INTRODUÇÃO

As discussões sobre a representação da informação e do conhecimento de materiais informacionais não-convencionais abre, lentamente, um caminho de aperfeiçoamento das práticas de tratamento da informação na área de Organização da Informação e do Conhecimento. Este caminho aponta horizontes positivos ao avanço da área como um todo, no que diz respeito, principalmente, a sanar questões metodológicas que admitam as particularidades informativas de objetos complexos, como as obras de arte, tratadas em especificidade neste trabalho, através das obras artístico-pictóricas.

Vislumbrando este aperfeiçoamento, diversas pesquisas emergem na área de Ciência da Informação, ora apresentando e defendendo a importância destes tipos de suportes, ora propondo adaptações ou desenvolvendo metodologias para o tratamento informacional dos mesmos. Como exemplo deste primeiro caso, observamos os trabalhos que tratam a arte como fonte de informação, já comuns na área, reconhecendo as mais variadas formas de expressão artística como espaços de investigação para os estudos em Ciência da Informação. Como exemplo do segundo, podemos citar trabalhos como os de Agustín Lacruz (2006); Debastiani (2012; 2015); Maimone (2007; 2009); Maimone e Gracioso (2007); Maimone e Tálamo (2008); Oliveira (2014); Oliveira e Pinho (2014); Medeiros (2014; 2017) e Medeiros e Pinho (2017; 2018), que refletem e investigam aplicabilidades metodológicas para representação da informação e do conhecimento de obras artístico-pictóricas, como segue a linha deste trabalho.

Evidenciando e reconhecendo a complexidade das práticas de representação voltadas a estes suportes, que envolvem um alto grau de subjetividade, propomos a reflexão de alternativas possíveis a estes processos, no que tange a encontrar soluções práticas, com base em metodologias já existentes, que possam potencializar os resultados das atividades de representação voltadas a estes suportes.

Propõe-se considerar aqui, como universo específico de reflexão, a relação entre a mineração de textos, por meio do software Sobek⁴ e a representação da informação de obras artístico-pictóricas, constantes nas pesquisas de Medeiros (2017), e Medeiros e Pinho (2017). Estes autores revelam, nos estudos sobre as obras artístico-pictóricas, espaços para extração de assuntos que predizem uma investigação discursiva, do conteúdo e do contexto de produção, que ultrapassa os vieses analíticos já comuns à representação da informação e do conhecimento de suportes bibliográficos, tradicionalmente conhecidos na área de Ciência da Informação.

⁴ Disponível em: <http://sobek.ufrgs.br/>

Estes autores, em síntese, propõem o uso de mapas conceituais para representação da informação de obras artístico-pictóricas com representações de memória social. Contudo, seu trabalho não aponta caminhos práticos de uso dos conceitos e termos extraídos na construção dos mapas para, por exemplo, aplicação em sistemas de recuperação de informação, nem mesmo perspectivas que possibilitem uso de tecnologias como auxiliares no processo intelectual de representação.

Percebe-se que a documentação gerada, entretanto, pelos processos de representação, como, por exemplo, os dados textuais das análises das obras, textos de descrição e análise para delimitação de conceitos de Medeiros (2017), e textos interpretativos dos mapas, se apresentam como objeto importante que, além da própria delimitação conceitual relativa ao domínio da obra, podem ser utilizados como elementos de entrada para *softwares* de mineração de texto, como auxílio à indexação, através de processos automáticos, com usos possíveis que deem destinos ainda mais abrangentes a essas discussões metodológicas.

Para tanto, de modo a investigar o tratamento temático dessa documentação, considerando o viés metodológico tratado por Medeiros (2017), discutimos sobre as relações possíveis entre os *softwares* de mineração de texto, especificamente o Sobek, nos processos de representação, com base na documentação gerada pela análise de obras artístico-pictóricas, como recurso de sistematização de termos (representativos e/ou derivados dos conceitos) para produção de produtos documentários, tomando como base o estudo de caso envolvendo a documentação de três das obras artístico-pictóricas trabalhadas pelo autor.

Deste modo, entende-se que a mineração de texto possa auxiliar metodologicamente nos processos de indexação, possibilitando encontrar termos relevantes em documentos de texto, estabelecendo padrões e relacionamentos entre eles com base na frequência e temática dos termos encontrados (SERAPIÃO, 2010 apud PEZZINI, 2016, p. 59), onde se encontra a oportunidade na qual se embasa a presente reflexão sobre os *softwares* de mineração e a representação de obras artístico-pictóricas.

Submeter a documentação gerada na atividade de representação ao processamento automático do *software* de mineração perfaz-se, não como um desvio metodológico do que já se observa na pesquisa dos autores supracitados, mas como uma investigação inicial do uso de instrumentos tecnológicos na busca por efetividade e eficácia nas atividades de representação, na garantia de que o acesso à informação e a formação do conhecimento possam ser contínuos através das discussões e descobertas científicas da área de Ciência da Informação.

Deste modo, o objetivo geral deste trabalho foi apresentar uma relação possível entre os softwares de mineração de textos e a sistematização de termos nas práticas de representação da

informação a partir da documentação gerada pela análise de obras artístico-pictóricas tratadas por Medeiros (2017).

Uma vez que se faz ausente, ainda, tal relação de abordagem na Ciência da Informação, podem-se considerar propostas como estas oportunidades de se investigar novos caminhos para o uso de recursos tecnológicos na representação da informação de suportes não convencionais de informação, assumindo como justificativa de desenvolvimento da pesquisa, a necessidade de abrir espaços à discussão, investigando as possíveis relações através de aplicações práticas, admitindo, contudo, a constatação de possíveis hipóteses, às descobertas tanto positivas quanto negativas acerca do objeto estudado, desencadeando, deste modo, o surgimento de novos vieses de percepção tanto das atividades práticas (neste caso específico das práticas de representação da informação e do conhecimento), quanto do desenvolvimento e aprimoramento dos recursos tecnológicos diversos (neste caso, dos *softwares* de mineração de texto).

2 REPRESENTAÇÕES EM MAPAS CONCEITUAIS E MENTAIS DE OBRAS ARTÍSTICO-PICTÓRICAS

O entusiasmo de atender às necessidades de acesso à informação tem permitido, ao longo dos anos, e cada vez mais, que a Ciência da Informação e suas subáreas, se proponham a investigar e desenvolver instrumentos, modelos e tecnologias que permitam a representação e a recuperação da informação. Observamos, por meio desta constatação, a exemplo, os estudos sobre o uso de mapas conceituais como recursos e instrumentos na área de Organização e Representação da Informação e do Conhecimento - ORIC, como os trabalhos de Rodrigues e Cervantes (2013; 2014; 2016; 2017), Rodrigues, Cervantes e Lunardelli (2016), Cordovil e Francelin (2018).

Pode-se notar que na área de ORIC, parte dos trabalhos dedicados a compreender o uso dos mapas conceituais está relacionada à aplicação deste tipo de metodologia ou instrumento para aplicação em suportes informacionais de, conseqüentemente, áreas diversas. Destacamos, deste modo, os estudos sobre mapas conceituais para a representação de suportes não convencionais de informação, como as obras artístico-pictóricos, pelos trabalhos de Medeiros (2017) e Medeiros e Pinho (2017).

No caso desses últimos autores, em específico, dedicados ao suporte artístico de informação, ilustra-se quão complexa a representação da informação e do conhecimento pode se tornar quando ultrapassa as perspectivas tradicionais de indexação dos suportes bibliográficos. Ao discutirem vieses metodológicos possíveis para aplicação de mapas

conceituais a estes suportes, uma vez que é que intrínseca a estas pesquisas, os autores procuram, também, compreender as obras artístico-pictóricas sob os olhos da Ciência da Informação, respeitando e valorizando o seu domínio de origem.

Pesquisas como a acima citada nos dão indícios, na produção científica nacional, da busca por progresso em relação às metodologias, e a própria perspectiva de indexação de imagens. A complexidade que acompanha tal processo carece, ainda, de mais investigação e discussões, assim como mais perspectivas práticas que possam orientar os possíveis percursos para essa atividade voltada ao suporte artístico de informação.

Merece atenção, também, no âmbito da indexação, entender que, muito embora haja os processos de representação intelectuais e automáticos para a representação de conteúdo textual, quando o foco é a imagem de uma maneira geral, a automação na representação ainda é algo pouco discutido. Ainda que ambos os tipos de representação sejam processos distintos é, no fato de não serem excludentes, que estes processos de representação se fortalecem, aprimorando e facilitando as atividades dos profissionais de informação.

Observando para além da ação objetiva da indexação, as atividades de representação com processos intelectuais, exigem um maior contato e compreensão do profissional da informação com o domínio do conhecimento do qual pertence o objeto documental a ser representado. Os mapas conceituais, como trabalhados para a representação das obras artístico-pictóricas por Medeiros (2017) e Medeiros e Pinho (2017), quando utilizados como instrumentos de representação nos estudos de Ciência da Informação, mostram claramente esta necessidade. Além de serem, uma vez construídos, produtos da representação, os mapas conceituais são, em sua compreensão de processo, uma metodologia que possibilita a aproximação do profissional ao domínio, na formação desta compreensão.

Os mapas conceituais, criados pelo professor Joseph Novak, na década de 1970, são, segundo Tavares (2007), estruturas temáticas, que sistematizam conceitos pertencentes a uma mesma rede de proposições. Tais estruturas de sistematização oferecem a visualização de um domínio do conhecimento com base na organização lógico-cognitiva de quem a produziu, permitindo uma visualização que considera a profundidade e extensão da temática representada.

Moreira (2013), corroborando com Tavares (2007), aponta os mapas conceituais segundo a sua aplicação aos domínios do conhecimento diversos, apresentando-os como diagramas que estabelecem as relações de sentido entre conceitos, segundo uma ordenação hierárquica. Os mapas conceituais são, assim, conforme apresentam Rodrigues e Cervantes (2013, 2016), ferramentas fundamentadas na aprendizagem significativa, de David Ausubel, na busca por representações do conhecimento armazenado na estrutura cognitiva dos indivíduos.

Ressaltamos, também, os mapas mentais ou os *mind maps*, que não são diferentes dos mapas conceituais, foram desenvolvidos também na década de 1970, pelo inglês Tony Buzan, com o ideal de ser uma forma possível de se organizar o conhecimento através de uma representação livre e não necessariamente linear (CORDOVIL; FRANCELIN, 2018), sendo apresentados visualmente em formas de diagramas. Muito embora tenha propósitos similares aos dos mapas conceituais, os mapas mentais possuem uma forma mais livre de construção. No caso dos mapas mentais, as relações de associação entre as ideias não precisam, necessariamente, serem descritas com palavras de ligação, o que é mais comum nos mapas conceituais.

No caso da representação de obras artístico-pictóricas o uso de mapas conceituais como instrumentos de representação da informação e do conhecimento se dá, como explica e ilustra Medeiros (2017) e Medeiros e Pinho (2017), como mecanismo para análise, extração de conceitos e documentação do conteúdo narrativo das obras, criando novas formas de organização e apresentação desta narrativa para fins documentários diversos, desde o estudo e compreensão do domínio a usos diversos para recuperação e uso da informação.

Evidencia-se, com o uso de mapas conceituais para representação do conteúdo de obras artístico-pictórica, pela perspectiva metodológica dos autores acima, que esta atividade gera uma documentação que é desenvolvida no processo de transcrição do conteúdo visual da imagem para um formato textual nos processos de análise e interpretação. Neste caso, os autores registram, no momento de construção da sua análise de conteúdo e extração de conceitos, textos descritivos relativos a cada obra, que embasaram a construção de mapas conceituais para a representação e documentação de seus domínios.

Visualizamos nestes tipos de registros descritivos, relativos à representação das obras artístico-pictóricas de Medeiros (2017) e Medeiros e Pinho (2017), por mapas conceituais, uma perspectiva de evolução dos processos de análise e extração de conceitos. Dada à complexidade das atividades citadas, que demandam não apenas de tempo, mas, de grande esforço intelectual dos profissionais de informação, consideramos oportuna à discussão sobre as possibilidades de automação de algumas etapas destes processos, com o auxílio de ferramentas tecnológicas. Esta discussão pode, em um futuro não distante, evoluir ao desenvolvimento de sistemas mais complexos e completos para representação do conteúdo de imagens artísticas através de instrumentos para automação destas atividades.

É, a partir do exposto, que inserimos o uso de *softwares* de mineração de texto como uma possível porta de entrada para esta discussão sobre a possibilidade de inserção de ferramentas tecnológicas para auxílio na representação do conteúdo de obras artístico-

pictóricas. Este tipo de ferramenta – apresentada de forma mais detalhada na seção seguinte – utiliza documentação textual para sumarização e criação de índices de assuntos relevantes. Deste modo, nos é ponto de interseção com a perspectiva metodológica de Medeiros (2017) e Medeiros e Pinho (2017), a utilização de textos para extração de conceitos, o que nos permite a ideia de relacionar ambos os processos, de modo a vislumbrar uma possível interoperabilidade das práticas manuais e automáticas na representação da informação de suportes artístico-pictóricos de informação.

De modo a iniciar na literatura nacional tal discussão, este trabalho, de maneira prática, testa como um *software* de mineração de texto se comporta ao utilizar a documentação gerada pela análise do conteúdo de imagens artístico-pictóricas. Para este teste, foi utilizado um *software* de mineração de textos chamado Sobek, apresentado com detalhes no tópico seguinte. Este *software* gera, em seu processo de sumarização, apresentações gráficas em forma de mapas mentais, com relações estabelecidas pelo grau de relevância dos termos elencados por ele através de relações estatísticas de proximidade.

No tópico seguinte, apresentamos a mineração de textos, bem como os *softwares* de mineração de textos, especialmente o Sobek.

3 MINERAÇÃO DE TEXTO E *SOFTWARES* DE MINERAÇÃO DE TEXTO

A principal relação que propomos neste trabalho é a dos *softwares* de mineração de texto com as atividades de representação da informação e do conhecimento em obras artístico-pictóricas. No entanto cabe-nos apresentar, inicialmente a mineração de textos, com a ressalva de que, não é objetivo deste trabalho apresentar um panorama exaustivo com base na literatura de Ciência da Informação sobre os conceitos apresentados, mas, elucidar de forma breve estes conceitos para que seja compreensível a aplicação prática que se segue.

De antemão, conforme apresenta Kaur e Aggarwal (2013), a mineração de texto pode ser vista como uma técnica emergente da área de Mineração de Dados. Esta última, corresponde a uma área que lida com a análise de dados, para extração de padrões úteis e válidos, sendo, também, um processo que está no centro do KDD - Knowledge Discovery in Databases (KAUR; AGGARWAL, 2013). KDD pode ser entendido como um “conjunto de ações de exploração e análise de dados, envolvendo seleção, pré-processamento, limpeza, transformação [...] interpretação e avaliação”, incluindo a própria mineração de dados (STOROPOLI et al., 2016, p. 6).

A mineração de texto, segundo Woszezenki e Gonçalves (2013), se compreende como uma técnica utilizada para extração de informações importantes de textos. Pode ser entendida, também, como “o estudo e a prática de extrair padrões, regras e tendências a partir do texto completo [...] usando princípios da linguística computacional e métodos analíticos”. (SULLIVAN, 2001; FAIIAZEE et al., 2012 apud WOSZEZENKI; GONÇALVES, 2013).

Já Allahyari *et al.* (2017), traz uma definição de mineração que corrobora com a definição de Kaur e Aggarwal (2013), na qual a apresenta como uma tarefa que ganhou atenção nos últimos anos, visando extrair informações significativas de textos, fazendo uso de técnicas e algoritmos com intuito de descobrir padrões úteis.

Para Yoon, Phaal e Probert (2008), a mineração auxilia na identificação de assuntos relevantes, numa possível extração de conhecimentos, atentando-se à forma pela qual os resultados dessa extração podem ser apresentados, com o uso de palavras-chaves, com apresentações que resumem os conteúdos e estabelecem as relações entre os documentos.

A descoberta de padrões relevantes e úteis é, sem dúvidas, a principal característica da mineração de textos, e está presente em grande parte das definições que se encontra na literatura da área de Ciência da Informação. Do mesmo modo, Barrera (2014), faz ênfase a esta qualidade, corroborando com os demais autores e, especialmente com Yoon, Phaal e Probert (2008), no que se refere à extração de conhecimentos. Para Barrera (2014), a mineração de texto é um processo responsável pela descoberta de conhecimentos implícitos, emergentes do relacionamento de vários textos de uma coleção.

Em diálogo com representação da informação e do conhecimento em obras artístico-pictóricas desenvolvidas por Medeiros (2017) e Medeiros e Pinho (2017), observamos que a mineração de texto é uma alternativa que nos ajuda a pensar meios de manipulação de documentos textuais relativos ao processo intelectual de indexação para descoberta de assuntos relevantes. Importante, também, que, muito embora este trabalho se detenha a esta documentação, esta reflexão pode se ampliar às demais documentações relativas às imagens artísticas, inclusive, documentações textuais de outras fontes, que, combinadas com as da análise do profissional de informação podem, quem sabe, ampliar ou evidenciar novas relações importantes de assuntos, que mereçam ser indexadas.

A mineração de textos pode, assim como coloca Barrera (2014), contribuir não apenas na identificação de assuntos relevantes, mas ajudar a encontrar desvios e associações que passam despercebidas no manuseio de um documento ou de um conjunto deles. Assim como apresentam Gupta e Lehal (2009), estas descobertas possíveis através da mineração de textos

são realizadas por computadores, em um processo de extração automático de um ou diversos recursos escritos.

Embora os *softwares* de mineração de textos tenham muitas aplicabilidades, especialmente para quais fins forem utilizados, em relação ao uso da mineração de texto em unidades de informação, por exemplo, Barrera (2014), salienta a importância desta técnica para as bibliotecas digitais e seus usuários, situando o uso da mineração de textos, como recurso integrado aos sistemas de gerenciamento de coleções de texto. Existem muitos *softwares* de mineração de texto disponíveis e muitos deles livres. Dentre estes utilizados em unidades de informação temos, os citados por Barrera (2014), que são o Greenstone, que é utilizado para o gerenciamento de bibliotecas digitais, e libViewer que utiliza métodos de agrupamento, para organização de documentos em bibliotecas digitais e afins, com base em grupos temáticos.

De maneira geral, os *softwares* de mineração de texto estão, segundo Pezzini (2016), relacionados a um paradigma de programação que nasce da busca por entender as relações textuais, por vezes ambíguas, dos documentos virtuais, que até então não foram resolvidas por outros paradigmas de desenvolvimento de *software* estando, por sua vez, ligadas a diversas outras áreas, dentre elas à Recuperação da Informação.

Embora existam outros *softwares* de mineração de texto, recortamos para tratar neste trabalho o *software* Sobek, que é uma ferramenta de mineração de textos desenvolvida para criar representações em *grafos* dos termos mais relevantes extraídos de documentos textuais. O *software* foi desenvolvido pelo Programa de Pós-graduação em Informática na Educação, da Universidade Federal do Rio Grande do Sul - UFRGS, e se encontra disponível tanto a versão de uso online quanto a versão para download do *software*.

O Sobek se destaca entre os demais *softwares* de mineração de texto, dentro de nossa discussão, por já ter sido utilizado com textos em português em pesquisas anteriores e, especialmente, por gerar como resultado uma representação visual do texto, semelhante a um mapa mental. A forma de apresentação dos resultados do Sobek se assemelha, portanto, aos mapas conceituais desenvolvidos por Medeiros (2017) e Medeiros e Pinho (2017).

A mineração de texto desenvolvida pelo Sobek faz a tarefa de sumarização de texto, de forma não supervisionada, baseada em modelo de *grafo* e análise estatística do texto, sendo aplicado a textos individuais, visando gerar um sumário ou resumo gráfico do texto. O *software* utiliza o modelo de mineração textual conhecido como *n-simple*, que faz uso de um

[...] algoritmo que realiza uma análise estatística dos termos presentes no texto e os seleciona a partir do valor absoluto de sua ocorrência e [...] considera também as relações de proximidade entre os componentes do texto, ligando cada termo estatisticamente relevante a N subsequentes palavras também relevantes (SCHENKER, 2003 *apud* COSTA *et al* 2017, p. 94, grifo nosso).

Assim como os mapas conceituais e mentais, as apresentações em grafos geradas pelo Sobek têm sido utilizadas como ferramentas para compreensão de conteúdos diversos, na sua assimilação e estudo, também no âmbito educacional. De maneira muito simples, o software exige apenas a inserção do conteúdo escrito na sua caixa de texto e automaticamente, com base em cálculos estatísticos – que não perfaz como objetivo deste trabalho detalhar - extrai e relaciona os principais termos.

Autores como Klemann, Reategui e Lorenzatti (2009), definem os *grafos* gerados pelo Sobek como mapas conceituais. Consideramos que tal apresentação de *grafo* se assemelha mais às estruturas livres dos mapas mentais, o que não diminui a sua usabilidade, nem seu potencial. De certo modo, a sua apresentação como um mapa mental restringe, contudo, à características mais complexas de relacionamentos que são construídas nos mapas conceituais com uso de expressões e frases de ligação que direcionam a formação de sentido nas relações entre conceitos.

Assim, este trabalho ao iniciar a discussão sobre o uso da mineração de texto nas atividades de representação do conteúdo de obras artístico-pictóricas, vê em *softwares* como o Sobek, um forte candidato a aliado ao processo de manuseio da documentação gerada nas atividades intelectuais de representação, uma vez que, tendencialmente, este tipo de *software* já se propõem a apresentar seus resultados graficamente com diagramações que relacionam os assuntos extraídos em esquemas similares aos mapas mentais.

Deste modo, buscando visualizar o comportamento do Sobek ao ser utilizado com documentos oriundos das atividades de representação da informação e do conhecimento de obras artístico-pictóricas, propomos com um teste inicial, inserir uma amostra da documentação gerada por Medeiros (2017), no *software* e analisar os *grafos* gerados, fazendo uma breve comparação com os mapas conceituais resultantes das representações dos autores citados.

4 METODOLOGIA

Considerando a realidade das obras artístico-pictóricas tratadas por Medeiros (2017) e Medeiros e Pinho (2017), documentos visuais, para extração de assunto, se faz clara, para o processo de representação como um todo, a necessidade de um processamento manual de análise, para extração de assuntos relevantes, seguido da delimitação de conceitos e por fim, da criação de uma representação.

O processo de análise desenvolvido manualmente pelos autores gerou uma documentação importante, em formato textual, com registros descritivos e interpretativos dos conteúdos visuais das obras artístico-pictóricas das quais eles tratam. Dentre esta documentação, foram escolhidos do trabalho de Medeiros (2017) – visto que este é o trabalho mais abrangente, três (3) textos descritivos extraídos de um conjunto de 16, referentes à análise de obras artístico-pictóricas, de uma série intitulada Linguagens do Agreste, de 2014, do artista plástico pernambucano Joaz Silva. Estes textos descritivos formam as compilações das informações mais importantes pelas quais Medeiros (2017), pôde extrair as temáticas das obras, bem como delimitar os principais conceitos e suas relações para, a partir disso, confeccionar os mapas conceituais do seu trabalho.

O Sobek admite a inserção textual para extração de assuntos e delimitação de termos descritores. Neste caso, é possível verificar de que modo o *software* se comportaria e como seriam os *grafos* gerados por este, caso fossem inseridos os textos da documentação de Medeiros (2017). Assim, como metodologia, utilizamos um estudo de caso para verificar o comportamento do Sobek ao lidar diretamente com a documentação gerada pelas análises de Medeiros (2017).

O estudo de caso admite o estabelecimento de um contato inicial com a questão, tendo como propósito maior observar, com uma aplicação prática, o comportamento do objeto estudado para, em possíveis futuras pesquisas, tomar profundidade no desenvolvimento de aplicações mais complexas de manipulação da documentação, do *software* em questão ou mesmo, o desenvolvimento de novos *softwares* e ferramentas para mineração voltadas a tipos específicos de textos documentários.

Assim, tomou-se como corpus escolhido para realização do estudo de caso a documentação das 3 obras, propondo a sua inserção no *software* de modo a verificar também os *grafos* gerados. A documentação escolhida trata-se dos textos descritivos das análises das obras Aprendiz de Artesanato, Feira de Caruaru e Faroeste Caruaruense. Estes textos foram escolhidos sem um critério específico, uma vez que todos os textos da documentação de Medeiros (2017), seguem a mesma linha de construção, numa perspectiva descritiva que segue uma metodologia comum para todos. Da mesma forma, buscou-se fazer a aplicação com apenas 3 documentos textuais, visto que a ampliação deste número tornaria demasiado extensa a seção de análise deste trabalho, e possivelmente redundante, visto a similaridade estrutural e informativa dos documentos.

Os textos descritivos gerados pela análise de Medeiros (2017), reúnem 3 níveis de descrição de Panofsky (1991), para descrição de imagens artísticas, que foram registrados em

forma de textos corridos. Como cada uma das obras tratadas por Medeiros (2017), tratam de temáticas distintas, que foram analisadas pelo mesmo na sua individualidade, optou-se, também, por utilizar a documentação de 3 obras com referências visuais relativas a assuntos distintos, de modo a observar se a temática teria influência na extração de assuntos pelo Sobek.

Seguiram-se como etapas para realização do estudo de caso, para cada uma das documentações, as três seguintes: a) Organização textual do conteúdo documental; b) Inserção do conteúdo textual organizado no *software* Sobek e geração do *grafo* e; c) Análise do *grafo* gerado pelo *software* Sobek. A seguir seguem as descrições de cada um dos passos:

a) Organização textual do conteúdo documental – esta etapa consiste na verificação de linearidade do conteúdo textual da documentação a ser inserida no software, de modo que, caso necessário, se estabeleça uma organização com a remoção de espaços, *parágrafos*, imagens ou elementos diversos (não textuais) que possam vir a atrapalhar o processamento do software;

b) Inserção do conteúdo textual organizado no *software* Sobek e geração do *grafo* – uma vez realizada a etapa anterior, o conteúdo textual é inserido e processado pelo *software* de modo a gerar o *grafo* representativo do conteúdo, com termos descritores e suas relações;

c) Análise do *grafo* gerado pelo *software* Sobek – por fim, nesta terceira etapa, verificam-se os *grafos* gerados pelo *software* de modo a aferir constatações sobre o processo e sobre a apresentação visual do *grafo*, sobre as relações estabelecidas entre os termos, sua funcionalidade para este fim de representação e as possíveis relações com os mapas conceituais de Medeiros (2017).

As etapas foram aplicadas de forma individual para cada documentação, na ordem a, b e c, sendo que, uma vez que as três documentações⁵ utilizadas possuem a mesma forma de apresentação textual, a etapa a, ocorreu da mesma forma para todas, com a exclusão de recuos e do seccionamento em parágrafos, deixando o texto contínuo, de modo que não houve nenhum caso especial de mudança que exigisse especificação detalhada. A etapa b traz os *grafos* gerados para as três documentações, estando estes apresentados ao longo do texto à medida que se discorrem as considerações da etapa c, que está posta a seguir, considerando os três *grafos* simultaneamente, estabelecendo, por sua vez, considerações particulares para cada um dos *grafos* à medida que necessário.

⁵Disponível para consulta na dissertação de Medeiros (2017), nas páginas 94, para a documentação da obra “Aprendiz de artesanato”, 128, para a obra “Feira de Caruaru” e, 161, para a obra “Faroeste Caruaruense”.

Pôde-se observar que o processo de utilização do Sobek consiste em uma ação simples e muito prática, sem exigir muitas etapas. O *software* possui duas versões disponíveis na sua página, uma para download e uma para uso online, sendo esta segunda a versão utilizada. Em relação à etapa b, de inserção do conteúdo textual, a versão online não admite a inserção de arquivos diretamente, sendo necessário, assim, que o conteúdo seja copiado e colado para a área designada, o que dificulta por sua vez, que se trabalhe com conteúdos muito extensos.

5 ANÁLISE DOS RESULTADOS

Uma vez que a documentação que compôs o *corpus* deste trabalho foi inserida no *software* de mineração de textos Sobek, pontuamos neste tópico as principais observações que foram constatadas a partir dos *grafos* gerados, apresentados neste trabalho em apêndices.

De antemão, constatou-se que o *software* em questão possui um processamento rápido, gerando automaticamente os *grafos* a partir dos textos. Os *grafos*, por sua vez, se constroem de forma fluida, seguindo uma organização dinâmica de seus elementos, sejam estes os termos apresentados em balões. Do mesmo modo as ligações entre os termos acompanham a fluidez, de maneira em alguns pontos, as linhas se interceptam formando grandes emaranhados de ligações.

De uma forma positiva, observou-se como positiva a extração de termos, para uma possível avaliação de conteúdos/assuntos presentes no domínio representado. Sendo possível, deste modo, apontar termos a serem considerados ou não como descritores representativos de assuntos relevantes. Tais termos podem, de algum modo, não serem percebidos em um processo de indexação manual, o que já traz indícios de que a utilização do *software* de mineração de textos pode contribuir para o levantamento de conceitos dentro de grandes domínios. De outro modo, a visualização prévia do *grafo* pelo indexador poderia permitir maior agilidade na indexação dos documentos no momento da indexação das obras artísticas, fazendo com o que *software*, ainda que não automatizando por completo a atividade, oriente o processo de extração de conceitos.

O Sobek também apresenta os termos com tamanhos variados nos *grafos*. É positivo considerar que o destaque dos termos pode direcionar, também, o olhar do profissional indexador a termos importantes, como observando os três apresentados a seguir. Vê-se no *grafo* do Apêndice A, os termos obra, artesanato, produção e criança em destaque. Já no *grafo* do Apêndice B, têm-se os termos obra, feira e igreja, e, por fim no *grafo* do Apêndice C, os termos colagem, paisagem e cidade de Caruaru. Neste caso, o destaque dado a estes termos, pelo Sobek,

sinaliza que segundo os seus cálculos de relevância, estes termos possam expressar assuntos de maior importância em relação ao domínio do texto submetido.

As conexões entre os termos, por sua vez são naturalmente um ponto positivo apresentado nos *grafos* gerados pelo *software*, uma vez que explicita possíveis relações conceituais que possibilitam uma interpretação mais profunda do conteúdo na leitura do *grafo* e, na compreensão do domínio do conhecimento representado. Estas relações são importantes por enfatizarem sob quais esferas os temas se conectam dentro do domínio como um todo, dando-nos aporte a investigar tais conexões.

O tipo de relação entre termos, que o Sobek apresenta difere dos mapas conceituais de Medeiros (2017), por duas características específicas, que são a ausência de setas, que nos mapas conceituais direcionam o movimento que a leitura perfaz do mapa e a ordem da ligação entre conceitos e, a ausência de palavras ou frases para estabelecer quais as relações diretas entre os termos/conceitos são visíveis dentro do contexto do conteúdo do documento representado. Essa diferença se justifica pelos fins que ambas as representações possuem, e ainda assim, não diminui a complexidade da teia formada, pois pode, a qualquer momento, identificar relações que na construção manual dos mapas não esteja explícito.

159

O processo de organização automática dos termos deixa a desejar, por vezes ao permitir que os termos extraídos se apresentem de forma sobreposta, o que dificulta a visualização das linhas que apontam as relações, o que é possível observar nos *grafos* dos Apêndices A e C. Os *grafos* estão apresentados neste trabalho da forma que foram geradas.

Ainda que o Sobek gere os *grafos* automaticamente, este disponibiliza as funções de ajuste e edição dos termos, onde também se podem estabelecer novas conexões entre estes com possível exclusão e inserção deles, se preferível. Esta função de reorganização dos termos é importante, para, em casos de usos reais, possam ser solucionados os casos de arranjos onde os termos se sobrepõem, o que pode dificultar a leitura, como no *grafo* do Apêndice A, com os termos “criança” e “trabalho”.

Relacionando o Sobek com a extração e representação de documentações como as de Medeiros (2017), documentações estas que possuem um caráter muito particular, de serem descrições de conteúdos visuais, extraídos com base em análises e, neste sentido, observando uma utilização que busca atender a perspectivas da Ciência da Informação, é importante considerar algumas questões.

A primeira delas é que um teste, com uma aplicação como a deste trabalho, é ciente de que as funções do software, com base na sua proposta primária de desenvolvimento, não incluem, por exemplo, que este seja apto a se comportar positivamente ao caso pelo qual ele

está sendo testado. Sendo assim, o estudo de caso realizado busca observar no processo o comportamento do Sobek durante seu uso, para gerar possíveis discussões e, fomentar tal ideal de interoperabilidade para, em propostas futuras, serem testadas novas formas de aplicação e, desenvolvimento de outros softwares que desempenhem papeis similares.

A segunda questão é que, diferentemente dos mapas conceituais de Medeiros (2017), o Sobek gera esquemas que não admitem inferências intelectuais durante a construção primária do grafo, onde se estabelecem as relações. Uma vez que este processo admite um cálculo estatístico automático. Assim, não é possível, neste momento, a intervenção de quem manipula o processo, o que se exprime em arranjos menos densos em relação a abarcar as subjetividades dos textos e, com possíveis desvios temáticos, uma vez que, termos não necessariamente importantes da análise intelectual possam aparecer nos grafos, o que pode deixar a desejar, quando considerados como textos para representação os de Medeiros (2017), com intencionalidades definidas de apresentação de informações descritivas.

Outra questão evidente nos grafos gerados pelo Sobek é a aparição de termos isolados, ou seja, sem ligações estabelecidas diretamente a outros termos, o que pode permitir, dentro de um processo de compreensão do conteúdo representado, que estes termos, que nomeamos aqui como “flutuantes”, gerem possíveis confusões, ou seja, por vezes irrelevantes aos próprios grafos, o que se mostra visível nos três grafos gerados a partir das documentações deste trabalho, apresentados nos *grafos* dos Apêndices A, B e C, onde no *grafo* do Apêndice A, o termo flutuante “paisagem”, muito embora na documentação tenha um papel importante na análise da obra artístico-pictórica, sendo utilizada para designar um aspecto de localidade envolvida na narrativa visual, aparecendo sem relação nenhuma com os demais termos do grafo, perde, neste, seu sentido.

Já o *grafo* do Apêndice B, por sua vez, traz o mesmo caso com os termos “Senhora da Conceição” e “Comercialização”, que soltos também deixam de expressar sentidos qualitativos importantes para compreensão do grafo em sua leitura geral, o que também acontece com os termos “autor” e “obras” no *grafo* do Apêndice C. No entanto, de uma maneira geral, a aparição de termos isolados pode se dar pelo fato de que o software não encontrou coocorrência no texto deste termo com outros considerados importantes, o que dentro do processo estatístico o impediu de estabelecer conexões.

Um caso particular se apresenta no *grafo* do Apêndice B, que são os termos “Caruaru PE” e “Barracas”, que estão ligados entre si, porém desligados de todo o restante do grafo. Estes dois termos, dentro do contexto textual da documentação, têm ligação direta, pois, relacionam a cidade à prática de comércio representadas na descrição da obra, entretanto, a ausência de

ligação entre esses termos e o restante das ideias do grafo, consequentemente, reduz o potencial lógico da apresentação do gráfico como um todo, uma vez que as relações ausentes podem inibir constatações lógicas distintas.

Entre as questões destacadas, observa-se que o Sobek, ainda que de forma menos profunda que os mapas conceituais de Medeiros (2017), permitem uma visualização geral do conteúdo dos textos. Isto é um ponto positivo, principalmente, por este apresentar tal resultado mesmo com a especialidade do texto utilizado neste trabalho, que não é comumente utilizado nos estudos que envolvem o *software* em linhas gerais. Se faz, nesse sentido, admissível considerar estudos mais abrangentes sobre este tipo de aplicação.

Este estudo de caso nos aponta um potencial evidente de uso do Sobek como aliado às práticas de representação da informação e do conhecimento de obras artístico-pictóricas, demonstrado pela sua agilidade de produção de arranjos com termos, ainda que com as questões já citadas na sua apresentação visual. Esta agilidade pode, sem dúvidas, contribuir na visualização prévia das possíveis relações e dos possíveis caminhos de arranjo que o conteúdo documental das análises das obras pode ter para representações futuras, até mesmo para a construção de mapas conceituais mais densos.

A possibilidade de uso de uma ferramenta de mineração de texto como o Sobek, aliada aos processos intelectuais de representação, especialmente de representações que admitem níveis de subjetividades maiores, é um indício de novas perspectivas metodológicas que trazem praticidade aos profissionais da informação que lidam com suportes não convencionais, como as obras artístico-pictóricas, na utilização de ferramentas tecnológicas que auxiliam nos processos intelectuais demorados como o estabelecimento de relações entre termos e diagramação de mapas.

O estabelecimento de ligações conceituais, embora se perfaça como processo cognitivo inerente ao homem, pode ganhar auxílio à medida que as relações de relevância entre os termos possam ser visualizadas por esquemas automáticos como os grafos gerados pelo Sobek, numa ação precedente ao aprofundamento com outros tipos de representação.

Não se fazem justo estabelecer comparações equiparadas entre as infinitas possibilidades relacionais que os mapas conceituais e mentais permitem em processos intelectuais manuais, com esquemas gerados pelo Sobek, que tem limitações no estabelecimento de relações, presas às especificações do texto analisado, enquanto máquina, mas, não se pode negar que esse tipo de usabilidade é uma tendência aos processos de representação da informação e do conhecimento, como esse trabalho discute, especialmente na

busca a atender às particularidades de suportes como as obras artístico-pictóricas, no viés de tratamento pela Ciência da Informação, como o desenvolvido por Medeiros (2017).

Constata-se a viabilidade de aprofundamento na discussão de uso do Sobek nos processos de representação da informação e do conhecimento de obras artístico-pictóricas, como instrumento auxiliar no processo de delimitação de termos e estabelecimento de relações entre estes, admitindo, contudo, que esta discussão é, ainda, inicial, e exige maiores estudos para possíveis propostas metodológicas de uso e, por que não dizer, de possíveis adaptações do software ou desenvolvimento de novos, utilizando a lógica do Sobek, para fins de representações com base nas análises de obras artístico-pictóricas.

6 CONSIDERAÇÕES FINAIS

Consideramos, com a reflexão levantada neste trabalho, relevante compreender que há possíveis aplicações que softwares de mineração de textos e que estes podem, inclusive, contribuir nos processos de representação da informação de obras artístico-pictóricas. Abrindo, assim, espaço a novas investigações sobre alternativas de aprimoramento das práticas de tratamento de suportes não convencionais de informação.

Trabalhos como este buscam inovação às práticas de extração de assunto, delimitação de conceitos, sistematizações de domínios do conhecimento, com o diálogo entre técnicas já existentes na área de Ciência da Informação, como os processos de análise de conteúdo, comuns às atividades de representação da informação e do conhecimento e a mineração de textos, utilizada como técnica para sistematização de termos e sumarização de conteúdos relevantes de textos, sugerindo a inserção de alternativas tecnológicas como apoio ao tratamento temático dos registros de informação, como as obras artístico-pictóricas.

As particularidades do software Sobek somam muito às discussões de Medeiros (2017), por direcionarem também, além de um caminho possível de automação de parte dos processos de representação de obras artístico-pictóricas, novas possibilidades de, através da mineração, potencializar a delimitação de termos aplicáveis aos sistemas de busca, uma vez que esse tipo de técnica permite a criação de índices de assunto, por exemplo, que possam ser utilizados para ampliar os resultados na recuperação da informação desses suportes e, da documentação relacionada a estes, em sistemas de recuperação da informação das unidades de informação e memória, bem como de ambientes virtuais diversos, que agrupam e disseminam esses tipos de materiais, como galerias, museus, pinacotecas virtuais e repositórios e bibliotecas digitais.

Por fim, estimamos como passos seguintes, a realização de novos estudos, sobre perspectivas metodológicas de representação mista, que possam englobar a atividade analítica do profissional da informação e o uso de ferramentas tecnológicas como o Sobek ou outros instrumentos de mineração de textos, na representação de obras artístico-pictóricas, ou mesmo uma proposta de desenvolvimento de *software*, embasado nos princípios da mineração de texto voltado especificamente para representação de conteúdos documentais gerados do tratamento desse tipo de suporte artístico.

REFERENCIAS

AGUSTÍN LACRUZ, M. del C. **Análisis documental de contenido del retrato pictórico:** propuesta epistemológica y metodológica aplicada a la obra de Francisco e Goya. Cartagena: 3000 Informática, 2006. 271p. Disponível em: <http://www.brapci.inf.br/index.php/article/view/0000006258/65bc5fc4cfb65d3ca76c36aad3029772/>. Acesso em: 22 jul. 2018.

ALLAHYARI, M. *et al.* A brief survey of text mining: Classification, clustering and extraction techniques. **arXiv preprint**. 2017. (Submitted on 10 Jul 2017 (v1), last revised 28 Jul 2017 (this version, v. 2)). Disponível em: <https://arxiv.org/abs/1707.02919>. Acesso em: 18 jul. 2019.

BARRERA, M. C. Minería de texto: una visión actual. **Biblioteca Universitária**, México, v. 17, n. 2, p. 129-138, 2014. Disponível em: <http://biblio.unam.mx/rbu/index.php/rbu/article/view/72/68>. Acesso em: 18 jul. 2019.

CORDOVIL, V. R. S.; FRANCELIN, M. M. Organização e representações: uso de mapa mental e mapa conceitual. *In*: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 19., 2018, Londrina, PR. **Anais...** Londrina, PR: ANCIB, 2018.

COSTA, A. P. M. *et al.* Emprego de um software baseado em mineração de texto e apresentação gráfica multirrepresentacional como apoio à aprendizagem de conceitos científicos a partir de textos no ensino fundamental. **Ciência & Educação**, Bauru, v. 23, n. 1, p. 91-109, 2017.

DEBASTIANI, A. M. **Imagens:** polissemia versus indexação e recuperação da informação. 2015. Tese (Doutorado em Ciência da Informação) – Universidade de Brasília, Brasília, 2015. Disponível em: <http://repositorio.unb.br/handle/10482/19050>. Acesso em: 22 jul. 2018.

DEBASTIANI, A. M. **Obras de arte como fonte de informação:** uma análise da verossimilhança das informações contidas em obras pictóricas que representam o Estado do Rio Grande do Sul. 2012. 92f. Trabalho de Conclusão de Curso (Biblioteconomia) Universidade Federal do Rio Grande do Sul, Porto Alegre, 2012. Disponível em: <https://lume.ufrgs.br/handle/10183/54309>. Acesso em: 22 jul. 2018.

GUPTA, V.; LEHAL, G. S. A Survey of Text Mining Techniques and Applications. **Journal of Emerging Technologies in Web Intelligence**, Finland, v. 1, n. 1, 2009. Disponível em: <http://www.jetwi.us/uploadfile/2014/1230/20141230112729939.pdf>. Acesso em: 7 jul. 2019.

KAUR, R.; AGGARWAL, S. Techniques for mining text documents. **International Journal of Computer Applications**, New York, v. 66, n. 18, 2013. Disponível em: <https://www.ijcaonline.org/archives/volume66/number18/11184-6318>. Acesso em: 7 jul. 2019.

KLEMMANN, M.; REATEGUI, E.; LORENZATTI, A. O emprego da ferramenta de mineração de textos Sobek como apoio a produção textual. *In: SIMPÓSIO BRASILEIRO DE INFORMÁTICA NA EDUCAÇÃO*, 20., 2009, Florianópolis. **Anais...** Florianópolis: UFSC, 2009.

MAIMONE, G. D. **Estudo do tratamento informacional de obras artístico-pictóricas: cenário paulista: análises e propostas**. 2007. 140 f. Dissertação (Mestrado em Ciência da Informação), Pontifícia Universidade Católica de Campinas, 2007. Disponível em: <http://tede.bibliotecadigital.puc-campinas.edu.br:8080/jspui/handle/tede/808>. Acesso em: 22 jul. 2018.

MAIMONE, G. D. Imagem artística: fonte de informação ou contemplação. **Cultura em Recorte**: Revista Eletrônica de Museologia e Ação Cultural, Campinas, v. 1, n. 1, p.18-33, jan./jun. 2009. Disponível em: <http://www.culturaemrecorte.org/ojs-2.2/index.php/capa/article/view/34/37>. Acesso em: 26 jul. 2018.

MAIMONE, G. D.; GRACIOSO, L. S. Representação temática de imagens: perspectivas metodológicas. **Informação & Informação**, Londrina, v. 12, n. 1, jan./jun. 2007. Disponível em: <http://www.uel.br/revistas/uel/index.php/informacao/article/view/1760>. Acesso em: 22 jul. 2018.

MAIMONE, G. D.; TALAMO, M. F. M. Tratamento informacional de imagens artístico-pictóricas no contexto da Ciência da Informação. **Datagramazero**, Rio de Janeiro, v. 9, p. 2, 2008.

165

MEDEIROS, W. O. **A representação da informação em obras artístico-pictóricas como elemento de compreensão da memória**. 2017. Dissertação (Mestrado em Ciência da Informação) – Centro de Artes e Comunicação, Universidade Federal de Pernambuco, Recife, 2017. Disponível em: <https://repositorio.ufpe.br/handle/123456789/24626>. Acesso em: 22 jul. 2018.

MEDEIROS, W. O. **A obra pictórica como fonte de informação: a representação imagética das obras de Marcus Jussier**. 2014. 80p. Trabalho de conclusão de curso (Bacharel em Biblioteconomia) – UFCA, Juazeiro do Norte, 2014.

MEDEIROS, W. O.; PINHO, F. O uso de mapas conceituais na representação da informação memorialística de obras artístico-pictóricas. *In: PINHO, F. A.; GUIMARÃES, J. A. (org.). Memória, tecnologia e cultura na organização do conhecimento*. Recife: Ed. UFPE, 2017, p. 297-305. Disponível em: <http://isko-brasil.org.br/wp-content/uploads/2013/02/livro-ISKO-2017.pdf>. Acesso em: 22 jul. 2018.

MEDEIROS, W. O.; PINHO, F. Reflexões sobre a análise documental em séries artístico-pictóricas. *In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO*. 19., 2018, Londrina, PR. **Anais...** Londrina, PR: ANCIB, 2018.

MOREIRA, M. A. Aprendizagem significativa, organizadores prévios, mapas conceituais, diagramas e unidades de ensino potencialmente significativas. *In: ENCONTRO REGIONAL*

DE APRENDIZAGEM SIGNIFICATIVA (ERAS NORTE),1., 2013, Belém. **Anais...** Belém: UEPA, 2013.

OLIVEIRA, R. A. **Obras de arte e memória imagética**: uma análise dos métodos de representação. 2014. Dissertação (Mestrado em Ciência da Informação). Centro de Artes e Comunicação, Universidade Federal de Pernambuco, Recife, 2014.

OLIVEIRA, R. A.; PINHO, F. A. Análise dos métodos de representação em obras de arte. **DataGramZero**, Rio de Janeiro, v. 15, n. 4, out. 2014.

PANOFSKY, E. **Significado nas artes visuais**. São Paulo: Perspectiva, 1991.

PEZZINI, A. Mineração de textos: conceito, processos e aplicações. **REAVI**, Santa Catarina, v. 5, n. 8, dez. 2016. Disponível em: <http://www.revistas.udesc.br/index.php/reavi/article/viewFile/6750/6415>. Acesso em: 15 out. 2018.

RODRIGUES, M. R.; CERVANTES, B. M. N. Identificação de conceitos por meio de mapas conceituais no âmbito da organização e representação do conhecimento. *In*: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 18., 2017, Marília. **Anais...** Marília, SP: ANCIB, 2017.

RODRIGUES, M. R.; CERVANTES, B. M. N. Os mapas conceituais para a visualização de conceitos de áreas do conhecimento em unidades de informação. **Revista ACB: Biblioteconomia em Santa Catarina, Florianópolis**, v. 18, n. 1, p. 752-776, jan./jun., 2013.

RODRIGUES, M. R.; CERVANTES, B. M. N. Organização e representação do conhecimento por meio de mapas conceituais. **Ciência da Informação**, Brasília, DF, v. 41 n. 1, p.154-169, jan./abr., 2014.

RODRIGUES, M. R.; CERVANTES, B. M. N. Organização e representação do conhecimento por meio de mapas conceituais. **Ciência da Informação**, Brasília, DF, v. 41 n. 1, p.154-169, jan./abr., 2016.

RODRIGUES, M. R.; CERVANTES, B. M. N.; LUNARDELLI, R. S. A. Mapas conceituais para representação do conhecimento na perspectiva de Barité. *In*: COLÓQUIO EM ORGANIZAÇÃO, ACESSO E APROPRIAÇÃO DA INFORMAÇÃO E DO CONHECIMENTO, 2016, Londrina. **Anais...** Londrina: UEL, 2016.

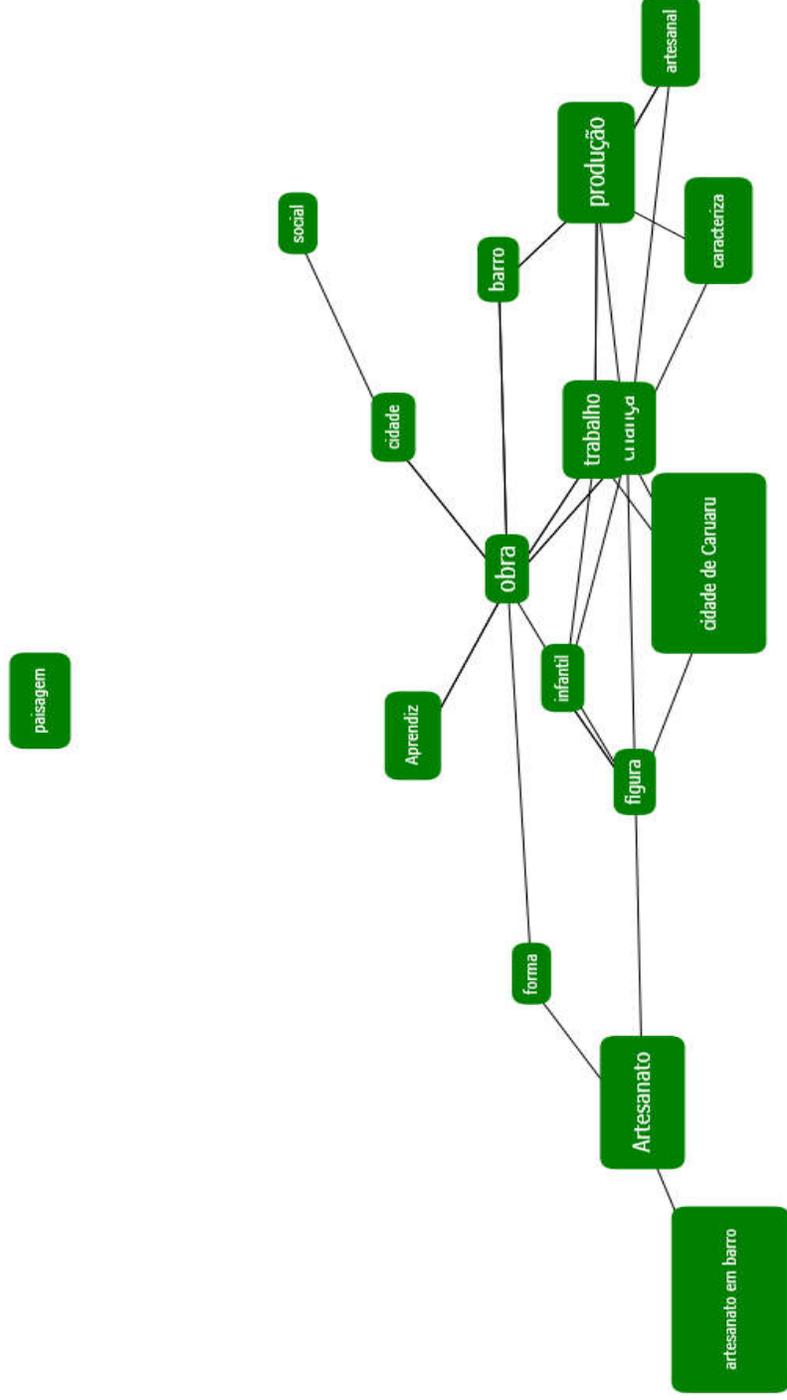
STOPOROLI, J. E. O uso do Knowledge Discovery in Database (KDD) de informações patentárias sobre o ensino a distância: contribuições para instituições de ensino superior. *In*: SIMPÓSIO INTERNACIONAL DE GESTÃO DE PROJETOS, INOVAÇÃO E SUSTENTABILIDADE, 5., 2016, São Paulo. **Anais...** São Paulo: FAPESP, 2016.

TAVARES, R. Construindo mapas conceituais. **Ciências & Cognição**, Rio de Janeiro, v. 12, p.72-85, 2007.

WOSZEZENKI, C. R.; GONÇALVES, A. L. Mineração de textos biomédicos: uma revisão bibliométrica. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 18, n. 3, p.2 4-44, jul./set. 2013.

YOON, B.; PHAAL, R.; PROBERT, D. Morphology analysis for technology roadmapping: application of text mining. **R&D Management**, v. 38, n. 1, p. 51-68, 2008.

APÊNDICE A – Grafo da documentação sobre a obra Aprendiz de Artesanato



APÊNDICE B – Grafo da documentação da obra Feira de Caruaru

