



Metodologia para a construção de portfólios tecnológicos agrícolas a partir de publicações técnico-científicas

Maria Fernanda Moura¹, Celina Maki Takemura², Ivan Lucas Cabral Silva³, Luisa Miyashiro Tápias³, Carolina Tavares de Oliveira³, Luis Henrique Bassoi⁴, Stanley Robson de Medeiros Oliveira¹

¹ *Pesquisador(a) da Embrapa Informática Agropecuária, Campinas, São Paulo, Brasil, {maria-fernanda.moura, stanley.oliveira}@embrapa.br*

² *Pesquisadora da Embrapa Monitoramento por Satélite, Campinas, São Paulo, Brasil, celina.takemura@embrapa.br*

³ *Estagiário da Embrapa Informática Agropecuária, Campinas, São Paulo, Brasil, mtapias.luisa@hotmail.com, caroli.agro@gmail.com*

⁴ *Pesquisador da Embrapa Instrumentação, São Carlos, São Paulo, Brasil, luis.bassoi@embrapa.br*

RESUMO

Neste trabalho é apresentada a proposta uma metodologia semi-automática para construir portfólios tecnológicos, a partir de um grande número de publicações técnico-científicas. Para delimitar o conjunto de dados foi utilizado um sistema de busca confiável sobre uma base validada de publicações, selecionadas do Sistema Aberto e Integrado de Informação em Agricultura (SABIIA), que cobre todo o domínio de interesse. Para a construção do portfólio foram adotados: alguns recursos linguísticos - como dicionários construídos pelos especialistas no domínio; técnicas de extração de informações nos textos por similaridade; técnicas de análise de estatística descritiva (percentis, gráficos); e, técnicas de reconhecimento de padrões associativos (análise de regras de associação). O experimento conduzido mostrou a aplicabilidade da metodologia, tendo resultado em um portfólio de tecnologias do uso da água na agricultura.

PALAVRAS-CHAVE: mineração de textos, extração de informação em textos, dicionários, regras de associação, cobertura geoespacial de textos, desambiguação de topônimos.

ABSTRACT

In this paper, a semi-automatic methodology to build technological portfolios is proposed from on a large number of technical-scientific publications. To delimit the data set, a reliable search system was used on a validated database of publications, selected from the Open and Integrated Information System in Agriculture (SABIIA), which covers the whole domain of interest. For the construction of the portfolio some linguistic resources were adopted, to wit: dictionaries built by specialists in the field; techniques for extracting information in texts by similarity; techniques of descriptive statistics analysis (percentiles, graphs); and, associative pattern recognition techniques (association rule analysis). The experiment conducted showed the applicability of the methodology, resulting in a portfolio of technologies associated with water use in agriculture.

KEYWORDS: text mining, information extraction, dictionaries, association rules, geospatial coverage of texts, toponym disambiguation

INTRODUÇÃO

Neste trabalho, entende-se portfólios tecnológicos como planilhas com a relação de tecnologias, localidades de aplicação, tipo de solo, culturas e outros atributos de interesse para caracterizar o uso da tecnologia. Esses portfólios, após identificados, devem ser analisados por especialistas nas tecnologias, a fim de verificar a importância de cada relação encontrada. Um exemplo é o portfólio construído com o uso da metodologia aqui proposta, no qual o interesse é listar as tecnologias agrícolas, utilizadas no Brasil, e relativas ao uso da água na agricultura, com as regiões brasileiras e culturas agrícolas (OLIVEIRA et al., 2016). Após a identificação desse portfólio e análise das tecnologias de uso mais significativo (neste caso, as que mais aparecem nas publicações técnico-científicas de agricultura no Brasil), os especialistas de domínio verificaram quais dessas tecnologias podem ser soluções para problemas decorrentes da avaliação e da adaptação de tecnologias para os biomas brasileiros, frente às mudanças climáticas. Neste caso, a construção do portfólio é o passo inicial e fundamental para as demais análises.

Considerando-se que a construção de um portfólio é a base de dados para outras análises, é interessante que ele seja construído a partir de um grande volume de referências bibliográficas, com ampla cobertura espacial, que tenham sido validadas e componham uma fonte confiável de informações. Dado que o domínio do trabalho é agrícola, optou-se por buscar informações via o Sistema Aberto e Integrado de Informação em Agricultura (SABIIA) (VACARI et al., 2011). A vantagem do SABIIA é reunir dados de vários provedores

(repositórios institucionais e temáticos, e periódicos científicos) previamente coletados e selecionados de acordo com as áreas de interesse da Empresa Brasileira de Pesquisa Agropecuária (Embrapa); o que nos oferece uma variada fonte de publicadores e de material altamente confiável. Essa variedade resulta em grandes conjuntos de resposta de busca, o que inviabiliza a análise humana desses resultados, em um curto intervalo de tempo.

Assim, neste trabalho, apresenta-se uma metodologia semi-automatizada com base em processos de mineração de dados textuais. Semi-automatizada porque há a dependência humana para: a) a definição das expressões de busca; b) construção de um vocabulário e sua sinonímia, caso não exista um disponível na literatura; e, claro, c) a avaliação e seleção dos resultados – tarefa de pós-processamento. A automatização do processo se dá na construção das planilhas a partir dos textos, extração/identificação de geolocalidades e extração de relações, por meio de regras de associação. As regras de associação buscam encontrar o relacionamento entre itens de dados que ocorram com uma certa frequência, ou seja, identificar padrões em dados históricos (AGRAWAL et al., 1994). Por exemplo, uma regra encontrada para o portfólio de tecnologias agrícolas ligadas ao uso da água foi “se (região=nordeste) e (cultura=uva) então (tecnologia=irrigação por gotejamento)”. Os resultados obtidos com as regras de associação resumem a compreensão das principais relações entre os atributos (que no experimento são tecnologias, culturas, solos e regiões) e essas são as relações que interessam para indicar a versão final do portfólio. Deve-se observar que, o contexto semântico das palavras e suas sinonímias não são considerados na extração das relações, o que pode resultar em falsas relações associativas, mesmo que elas apresentem uma alta frequência.

Nos próximos itens, detalha-se a metodologia proposta, discutem-se os principais resultados obtidos com o experimento das tecnologias agrícolas relacionadas ao uso da água na agricultura brasileira e apresentam-se as conclusões.

MATERIAL E MÉTODOS

A Figura 1 ilustra as etapas de construção do portfólio, cujos detalhes são apresentados nas próximas subseções.

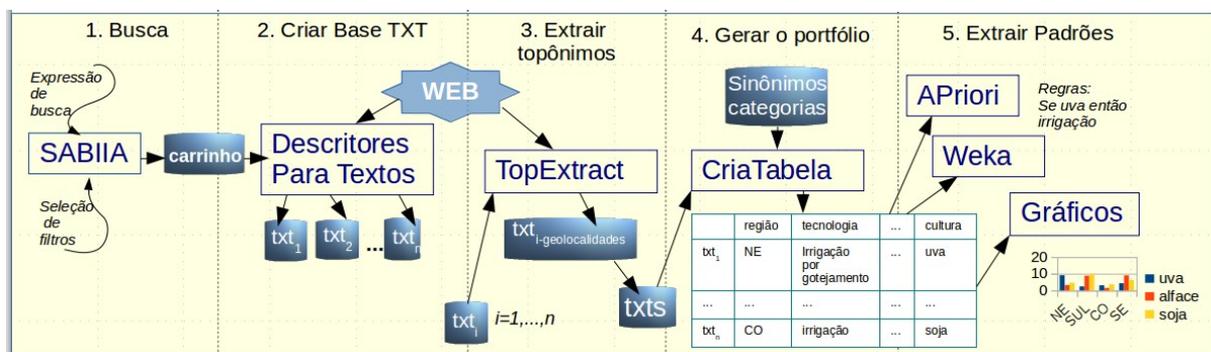
Busca

A busca é realizada pelo SABIIA, devido à variedade de publicações de acesso livre, validação dos domínios (publicadores) pelas bibliotecárias da Embrapa e, conseqüentemente, confiança nos dados.

O primeiro passo nesta etapa é construir as expressões de busca. Essa tarefa é realizada por especialistas do domínio das tecnologias junto aos mineradores de textos. Dessas expressões depende o resultado da análise, dado que elas delimitam o conjunto de dados sobre o qual os resultados serão obtidos. Logo, pode ser que se volte a esta etapa, se nas próximas etapas os dados não fornecerem uma cobertura de documentos aceitável.

Os filtros, “seleção de filtros” na Figura 1, são recursos que o SABIIA oferece, tais como filtrar os resultados de busca de acordo com: provedores de dados, autores, palavras-chave presentes nos resultados, tipo de publicação (relatórios, artigos, etc), ano de publicação e idioma. Esses filtros ajudam a delimitar os resultados de uma busca.

Figura 1: Metodologia semi-automática para a construção de portfólios tecnológicos



Criar base TXT

Quando o minerador e o especialista estiverem satisfeitos (este é um julgamento subjetivo) com o resultado da busca, eles devem salvar os documentos correspondentes no “carrinho do SABIIA”. O carrinho é um arquivo RTF, que pode ser convertido para txt plano, e que carrega todos os metadados de cada publicação recuperada. Nesta etapa, o carrinho é convertido para n arquivos textos.

A ferramenta de software¹ que faz essa conversão, permite que, se for interesse da análise de dados, sejam incluídos os textos completos na base TXT, delimitados por descritores XML e se, e somente se, os textos completos forem de acesso livre e, estiverem facilmente disponíveis nas urls indicadas em seus metadados.

Extrair Topônimos

Normalmente, é de interesse nesse tipo de análise, que se saiba os locais onde essas tecnologias são utilizadas. Para isso tem sido usada a ferramenta TopExtract², que utiliza uma biblioteca Python poliglota - que identifica entidades nomeadas como topônimos nos textos.

¹ Nomeada decritores2txt.py, em python, que pode ser requisitada aos autores.

² A ferramenta TopExtract também é desenvolvida em Python e pode ser requisitada aos autores.

Após identificar essas entidades, a ferramenta executa um processo de desambiguação, que leva em conta todo o conjunto de topônimos encontrados em cada texto. Por exemplo, se o Rio São Francisco aparece no texto, é necessário saber qual o polígono envolvente do trecho presente no texto (TAKEMURA et al., 2016), que pode ser considerado como o mais próximo a um outro topônimo também presente no texto, por exemplo, Petrolina. A partir do polígono encontrado, delimita-se a que polígono das regiões brasileiras ele pertence (Norte, Sul, etc).

Gerar o Portfólio

Essa etapa envolve: i) extração de termos dos textos, por similaridade e sinonímia; ii) criação das transações presentes nos textos; e iii) eliminação de dados redundantes ou que não contribuem para os resultados.

Na i) extração de termos dos textos, primeiro resolve-se a questão de quais são os termos de interesse e suas sinonímias. Normalmente essa tarefa é manual, a menos que já exista a relação de termos e seus sinônimos para o domínio em questão. Após essa identificação escolhe-se um grau de similaridade que se ache adequado à aplicação, por exemplo, 75% de coincidência - utilizando-se “edit distance” para calcular a similaridade. Por exemplo, ao comparar palavras como “alho” e “milho”, é conveniente que se utilize um grau de similaridade acima de 80%. Também é aconselhável que se construam dicionários de categorias; por exemplo, para a categoria “controle fitossanitário” pode-se ter um conjunto de termos tais como “{pulverização, pesticida, agrotóxico}” e outros de interesse. As categorias também são colocadas no portfólio e podem ser objeto de análise.

Para a ii) criação das transações geram-se combinações dos atributos. Por exemplo, se um documento apresenta para o atributo “cultura” os valores “alface” e “uva” e, para o atributo “tecnologia” os termos “pulverização” e “irrigação”, criam-se as combinações: (“alface”, “pulverização”); (“alface”, “irrigação”); (“uva”, “pulverização”); e, (“uva”, “irrigação”). Cada combinação dá origem a uma transação, ou seja, uma linha do portfólio.

Para a etapa de iii) eliminação de dados redundantes ou desnecessários, considera-se, primeiramente, colunas (atributos) que estejam todos vazios. A eliminação de linhas, isto é, de documentos, segue a ideia de que o número de combinações para cada documento não pode ser muito discrepante entre os demais documentos. Para isso, após observações empíricas de alguns casos, decidiu-se eliminar os documentos de acordo com os percentis das frequências de combinações que eles geram. Um número razoável a ser eliminado, de acordo com os experimentos realizados, é do percentil 75% ou superior - porém, manteve-se a escolha do percentil como parâmetro do processo dado que essa escolha depende de cada resultado de

busca. Uma outra forma de eliminar dados é considerar apenas as combinações que obrigatoriamente contenham uma instância de um determinado atributo, por exemplo, “tecnologia” não pode estar vazia na combinação. Esses filtros podem ser aplicados de forma independente e ao longo das análises, sem que seja necessário gerar um novo portfólio.

Extrair Padrões

Nesta etapa analisa-se o portfólio obtido, gerando-se regras de associação entre os seus atributos e gráficos descritivos da distribuição de seus atributos, de modo a visualizar um resumo dos dados. As regras de associação ajudam a identificar quais são as relações estatisticamente mais significantes entre regiões, tecnologias e culturas, e, também as relações mais raras, que, em muitos casos, podem ser as mais importantes. De acordo com a análise dessas relações, mantém-se ou não linhas no portfólio encontrado. Essas decisões são dos especialistas nas tecnologias junto aos mineradores de dados.

A opção pelo uso de regras de associação é a sua facilidade de interpretação, e a existência de algoritmos bastante eficientes frente a um grande conjunto de dados. As regras de associação podem ser representadas da forma $X \rightarrow Y$, em que X representa o antecedente e Y o conseqüente, tal que $X \cap Y = \emptyset$. Para cada regra estão associadas duas medidas tradicionais: confiança (Conf) e suporte (Sup). Sup representa o número de tuplas que contêm X e Y , ao passo que Conf constitui a razão entre o número de tuplas que contêm X e Y sobre o número de tuplas que contêm X e: $\text{Sup} = P(X \cup Y)$; $\text{Conf} = P(X/Y)$. Do ponto de vista conceitual, o suporte representa a significância estatística dos itens (termos) nas tuplas, ao passo que a confiança determina a força da regra; uma confiança alta implica em regras mais comuns (fortes) e uma confiança baixa em regras mais raras.

Para a geração das regras de associação foi utilizado o algoritmo Apriori (LIU et al., 1998), disponível no software Weka (HALL et al., 2009). A versão permite ao usuário definir o conseqüente das regras que serão geradas.

RESULTADOS E DISCUSSÃO

O experimento aqui mostrado é o resultado da construção de um portfólio de tecnologias agrícolas relacionadas ao uso da água na agricultura. A fonte de dados são os documentos técnico-científicos disponibilizados no SABIIA. As expressões de busca foram inicialmente propostas por especialistas do domínio (empregados da Embrapa) e, ao longo das análises, foram sendo aperfeiçoadas e evoluídas (OLIVEIRA et al., 2016). Ainda, os especialistas trabalharam na construção de dicionários de sinônimos para as tecnologias e para as culturas,

a partir dos termos dos thesaurus agrícolas Thesagro³ e do Agrovoc⁴ combinados aos termos presentes nas expressões de busca e retornados junto aos filtros de assunto pelo SABIIA. Por exemplo, para cultura “MANDIOCA” o dicionário construído apresenta as opções “mandioca, cassava, manioc”, e, para tecnologia “IRRIGACAO POR GOTEJAMENTO” o dicionário construído apresenta “irrigacao por gotejamento, irrigação por gotejamento, drip irrigation”, ou seja, inclusive as variações produzidas pela retirada de caracteres latinos (ç, ...).

Após a realização de algumas análises e observação da distribuição de frequências das culturas e tecnologias, os especialistas identificaram categorias de culturas e tecnologias presentes nos resultados de busca. Um exemplo é a classe “Base alimentar e commodities” que engloba as culturas “arroz, feijao, milho, soja, cana-de-acucar, limao, laranja, algodao, trigo, cafe, sorgo, mandioca, chuchu, pepino, batata” e, um outro, a classe de tecnologia “Melhoramento genetico” que engloba “cultivar, metodos, propagacao vegetativa, hibridacao, variacao genetica, variabilidade genetica, micropropagacao, germoplasma, melhoramento genetico “.

A base de dados final, após várias consultas e reformulações e filtros de consultas, foi formada por 2.185 documentos recuperados do SABIIA. Para esses documentos foram geradas 126.565 combinações (linhas do portfólio consideradas como transações). Como o objetivo inicial do portfólio era identificar tecnologias que envolvam o uso da água na agricultura utilizadas nas regiões do Brasil, foram eliminadas as linhas do portfólio que não apresentavam informação sobre o atributo tecnologia, resultando em 99.668 transações. Em seguida, foi observado que algumas linhas do portfólio eram constituídas de várias combinações de tecnologias com culturas, geradas de uma mesma publicação (mesmo documento txt). Para lidar com esse problema, foi aplicado um filtro que eliminou as linhas do portfólio cujas frequências foram superiores ao percentil 95%. Com isso, o portfólio foi reduzido a 84.204 transações. Por último, foram eliminadas do portfólio as linhas que não apresentaram informações sobre regiões do Brasil, pois tanto tecnologia como região eram objeto de estudo desse trabalho, resultando em 2.185 combinações (linhas).

Deve-se salientar que os procedimentos adotados acima foram realizados depois de várias tentativas empíricas de se obter um recorte razoável dos dados. Esse recorte razoável

³ Disponível em: http://snida.agricultura.gov.br:81/binagri/html/Cen_Thes1.html. Consultado em 23 de maio de 2017.

⁴ Disponível em: <http://aims.fao.org/vest-registry/vocabularies/agrovoc-multilingual-agricultural-thesaurus>. Consultado em 23 de maio de 2017.

deveria ser capaz de apresentar regras interessantes e também regras raras, envolvendo tecnologias, culturas e regiões. Essas decisões, embora baseadas em medidas, são subjetivas.

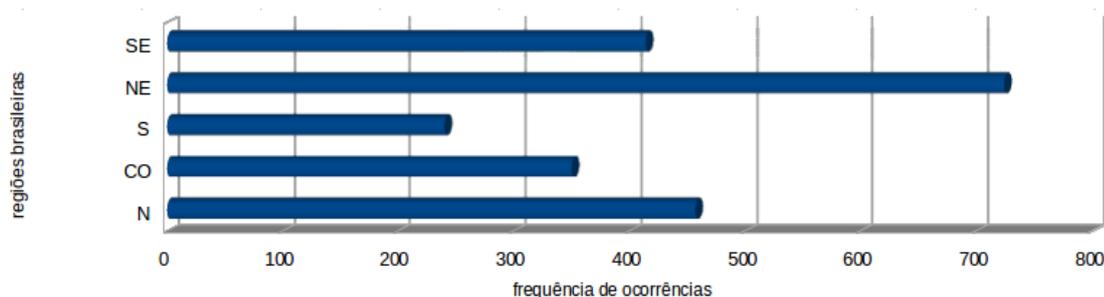
Dessa forma, obteve-se um portfólio, para o qual algumas linhas são ilustradas na Figura 2. No portfólio guardou-se a informação do arquivo (documento) que gerou cada transação, pode-se observar que o A_54904.201406150000.txt gerou duas transações, pois contém a tecnologia - “plântio direto” relacionada às classes (“solo” e “sistemas de producao”), ambas atreladas à cultura da soja.

Figura 2: Visão parcial do portfólio de tecnologias.

Arquivo	tecnologia	tecnologiaClasse	cultura	culturaClasse	solo	regiao	uf
A_13595.2015061500000.txt	irrigacao	'engenharia agricola'	uva	frutiferas		NE	BA
A_43443.2013061500000.txt	'manejo fitotecnico'	'controle fitossanitario'	alho	olericulturas		NE	BA
A_54904.2014061500000.txt	'plântio direto'	solo	soja	'base alimentar e commodities'		NE	BA
A_54904.2014061500000.txt	'plântio direto'	'sistemas de producao'	soja	'base alimentar e commodities'		NE	BA
A_20638.1982061500000.txt	germoplasma	'melhoramento genetico'	arroz	'base alimentar e commodities'		SE	MG
A_20638.1982061500000.txt	'variabilidade genetica'	'melhoramento genetico'	couve	olericulturas		SE	MG
A_20638.1982061500000.txt	'variabilidade genetica'	'melhoramento genetico'	arroz	'base alimentar e commodities'		SE	MG
A_61753.2015061500000.txt	manejo	solo	mandioca	'base alimentar e commodities'		CO	GO
A_44806.2000061500000.txt	'manejo do solo'		alho	olericulturas	latossolo	CO	GO
A_51410.2000061500000.txt	manejo	solo	feijao	'base alimentar e commodities'		CO	GO
A_55263.1984061500000.txt	fertilizante	solo	seringueira	extrativas		CO	GO
A_55263.1984061500000.txt	adubacao	solo	seringueira	extrativas		CO	GO
A_53347.2004061500000.txt	'plântio direto'	solo				S	RS
A_53347.2004061500000.txt	'plântio direto'	'sistemas de producao'				S	RS
A_33803.2014061500000.txt	manejo	solo			argissolo	S	RS
A_55657.2011061500000.txt	fertilizante	solo	alho	olericulturas	chernossolo	S	RS
A_23013.1998061500000.txt	irrigacao	'engenharia agricola'	pimenta	olericulturas		N	RR
A_23013.1998061500000.txt	adubacao	solo	pimentao	olericulturas		N	RR
A_25681.2010061500000.txt	'melhoramento genetico'	'melhoramento genetico'	alho	olericulturas		N	RR
A_25681.2010061500000.txt	'melhoramento genetico'	'melhoramento genetico'	arroz	'base alimentar e commodities'		N	RR

Para a geração das regras, foram conduzidos vários experimentos em que a tecnologia foi fixada como o consequente da regra (Y) e os demais atributos foram combinados no antecedente da regra (X). Desta forma, as regras identificaram qual o grau de associação entre tecnologias, tipos de solo culturas e UF em cada região do Brasil. No entanto, devido ao grande desbalanço entre os dados de cada região, como mostrado no gráfico da Figura 3, as 2.185 linhas do portfólio (instâncias) foram segmentadas por região, e cada região analisada individualmente. Dessa forma, a segmentação permitiu que regras de associação das regiões com menor frequência (Sul e Centro-Oeste) fossem identificadas no processo de mineração. As regras geradas apresentaram baixos valores de suporte e altos valores de confiança.

Figura 3: Dispersão das regiões brasileiras nos dados obtidos



Após a segmentação, para a região Norte foram encontradas 32 regras, considerando o suporte de 2% e a confiança de 90%, sendo parte delas redundantes ou não apresentava novidade para o especialista do domínio. Um dos exemplos de uma regra relevante foi (*Se tecnologiaClasse = planta & culturaClasse = base alimentar e commodities & solo = latossolo ==> tecnologia = fertilizante*). Outro exemplo foi (*Se tecnologiaClasse = sistemas de producao & culturaClasse = base alimentar e commodities ==> tecnologia = plantio direto*).

Para dados disponíveis das demais regiões do Brasil foram geradas regras similares. Por exemplo, para a região Nordeste, com suporte de 2% e confiança de 90%, foram geradas 40 regras. Exemplos dessas regras incluem (*Se tecnologiaClasse = engenharia agricola & cultura = uva & culturaClasse = frutiferas & regioao = NE ==> tecnologia = irrigacao*) e (*Se cultura = maca & culturaClasse = frutiferas & regioao = NE & uf = BA ==> tecnologia = manejo dos recursos hidricos*).

Para a região Centro-Oeste foram geradas 18 regras, considerando suporte igual a 3% e confiança igual a 90%. Exemplos dessas regras incluem: (*Se cultura = trigo & culturaClasse = base alimentar e commodities & uf = DF ==> tecnologia = fertilizante*) e (*Se tecnologiaClasse = solo & culturaClasse = olericulturas & solo = latossolo ==> tecnologia = plantio direto*).

Por fim, para a região Sul foram geradas 36 regras de associação para o suporte igual a 2% e confiança igual a 90%, com destaque para as regras (*Se cultura = trigo & culturaClasse = base alimentar e commodities & uf = SC ==> tecnologia = adubacao*) e (*Se culturaClasse = base alimentar e commodities & solo = gleissolo ==> tecnologia = fertilizante*).

CONCLUSÕES

Este trabalho apresentou uma metodologia semi-automatizada para a construção de portfólios tecnológicos a partir de um grande volume de publicações técnico-científicas com o uso de ferramentas de mineração de textos. A metodologia é baseada em: i) busca e delimitação do domínio de resultados, o que depende de um trabalho minucioso com auxílio dos especialistas de domínio; ii) criação automática de base de dados; iii) criação do portfólio, com decisões de filtros a serem aplicados dependentes dos mineradores de dados; e, iv) análise dos dados a partir de regras de associações automaticamente geradas. O processo é retroalimentável, pois observações em uma etapa podem sugerir aos analistas o retorno a uma ou mais etapas.

Deve-se observar que, nesta metodologia não foram considerados os contextos semânticos nos quais os termos dos dicionários aparecem. Os termos são considerados apenas

por presença, frequência e combinações destas no mesmo texto. Isto permite construir uma representação *bag of words* das transações, que pode resultar em regras não verdadeiras, que devem ser avaliadas pelos especialistas do domínio. Sendo que, o experimento ilustrativo mostra que a metodologia serviu ao seu propósito, tendo colaborado na construção de um portfólio de tecnologias relacionadas ao uso da água na agricultura.

Adicionalmente, a proposta da metodologia permite que ela seja adaptada a vários outros problemas semelhantes, desde que se trabalhe os recursos linguísticos (dicionários de sinônimas e categorias) dependentes do domínio e do conhecimento dos especialistas.

REFERÊNCIAS

- AGRAWAL, R.; SRIKANT, R. Fast algorithms for mining association rules in large databases. In: INTERNATIONAL CONFERENCE ON VERY LARGE DATA BASES, 20., 1994, Santiago. Proceedings... San Francisco: Morgan Kaufmann, 1994. p. 478-499.
- HALL, M. A.; FRANK, E.; HOLMES, G.; PFAHRINGER, B.; REUTEMANN, P.; WITTEN, I. H. **The WEKA Data Mining Software: An Update**. SIGKDD Explorations, v. 11, n. 1. 2009.
- LIU, B., HSU, W., MA, Y. Integrating Classification and Association Rule Mining. In: Fourth International Conference on Knowledge Discovery and Data Mining, 80-86, 1998.
- OLIVEIRA, S. R. de M.; MOURA, M. F.; TÁPIAS, L. M.; OLIVEIRA, C. T. de; BASSOI, L. H. Técnicas de mineração de textos para auxiliar a construção de portfólio de tecnologias de irrigação a partir de publicações científicas. Campinas: Embrapa informática Agropecuária, 2016. 96 p. il. (Embrapa informática Agropecuária. Documentos, 142).
- TAKEMURA, C. M.; SILVA, G. B. S. da; OLIVEIRA, S. R. de M.; MOURA, M. F. Desambiguação de topônimos usando dicionários geográficos. In: SEMINÁRIO DA REDE AGROHIDRO, 4., 2016, Brasília, DF. Água e agricultura: incertezas e desafios para a sustentabilidade frente às mudanças do clima e do uso da terra: anais. Planaltina, DF: Embrapa Cerrados, 2016. p. 67-74.
- VACARI, I.; VISOLI, M. C.; GONZALES, L. E. Acesso aberto a informação científica agropecuária na internet: caso do sistema aberto e integrado de informação em agricultura (Sabiia). In: CONGRESSO BRASILEIRO DE AGROINFORMÁTICA, 8., 2011, Bento Gonçalves. Anais... Florianópolis: UFSC; Pelotas: UFPel, 2011. Não paginado.