

problemas reais. A fim de medir a qualidade dos dados – sua “compreensibilidade” e “interessabilidade”, como cita Rezende [16] – aspectos subjetivos podem ser filtros para a (v) utilização do conhecimento, que é aplicação prática do que foi minerado.

MÉTODO PROPOSTO

As cinco etapas apresentadas por Rezende et al [17] serviram de base para o método proposto. O primeiro passo para a construção visual do método foi estabelecer sinais gráficos que representassem as cinco etapas. Utilizando formas gráficas que fazem menção ao processo do trabalho de extração de minerais – e de onde surge o termo “Mineração” – e, ainda, a dinâmica indicada por Rezende et al, os passos do método foram visualmente organizados (Figura 1).

Figura 1 – Processo de Mineração de Dados: definição de sinais gráficos com base na organização apresentada por Rezende et al.

Fonte: os autores.



As etapas do método foram, então, descritas de maneira aplicada. Em (i) conhecimento do domínio ficou estabelecido que o objetivo era selecionar as bases de dados que apresentavam um conjunto de informações relevantes, em termos de quantidade, frente ao interesse da pesquisa. As bases de dados podem variar de acordo com área de aplicação da pesquisa. No

caso de uma revisão sistemática que envolva a área da Saúde, por exemplo, os periódicos foco do estudo são diferentes daqueles em que a pesquisa objetiva áreas como Arte ou Marketing. As bases de dados a serem escolhidas serão aquelas que, portanto, contemplam os periódicos de interesse, ou seja, diferentes para cada caso.

Como (ii) pré-processamento, os termos descritores foram definidos, bem como selecionados os limitadores de pesquisa disponibilizados em cada base de dados. Os descritores ou palavras-chave devem ser estabelecidos visando todo o contexto da pesquisa e considerando a sua escrita em língua inglesa. O cuidado em observar sinônimos e diferentes formas de abordagem de um determinado assunto deve ser parte desta etapa. É também neste momento que o estudo das bases de dados deve ocorrer a fim de se selecionar – nas ferramentas de busca disponibilizadas – os limitadores, tais como áreas de interesse da pesquisa, seleção de periódicos específicos, textos completos e/ou resumos críticos, entre outros.

Para (iii) extração de padrões, propôs-se que operadores booleanos fossem utilizados em combinação aos termos descritores, seguindo a lógica de interesse da pesquisa. Os operadores booleanos são usados junto aos termos descritores para relacioná-los e, assim, estabelecer as condições que se deseja nos resultados da pesquisa. O operador “AND” faz o papel de intersecção dos termos, “OR” de soma e “NOT” de exclusão (Figura 2). Alguns sistemas de busca disponibilizam os operadores para seleção por parte do usuário, outros requerem que o usuário os escreva no próprio campo de busca [1].

Figura 2 – Operadores booleanos AND, OR e NOT.

Fonte: os autores.

