

PESQUISA PARA O DESENVOLVIMENTO DE MATERIAIS EM
DIFERENTES MÍDIAS PARA APOIO DIDÁTICO E DESENVOLVIMENTO DE
ATIVIDADES PARA A SNCT DE 2014

Autor: Caio Palla Marques. Docente da Universidade Federal de
Rondônia (UNIR), Departamento de Química. Campus José Ribeiro Filho - BR
364, Km 9,5. CEP 76801-059. Porto Velho/RO.

caio.palla@unir.br

Resultados do Projeto de Pesquisa - Processo 23118001303/2015-12

Resumo

Este texto apresenta as informações sobre o desenvolvimento, o funcionamento e os possíveis usos da plataforma de divulgação científica “Base Ciência”, que fornece uma base de dados para a busca seletiva por as notícias (artigos) sobre a produção científica e tecnológica das instituições públicas que publicam artigos de divulgação científica. A base de dados foi concluída e está disponível em www.baseciencia.com.br. O banco de dados do Base Ciência possui espaço disponível para registrar toda a produção de divulgação científica disponível on-line por todas as instituições de ensino e pesquisa nacional. No momento, o banco de dados do Base Ciência já tem o registro para a busca de artigos de divulgação científica de mais de 90 instituições públicas brasileiras, totalizando assim o registro até o momento, o registro de mais de 60 000 (sessenta mil) artigos de divulgação científica. Desta forma este projeto visa facilitar o acesso a materiais para apoio didático na forma de divulgação científica.

Palavras-Chave: Divulgação e Popularização da Ciência, Instituições Públicas. Base de dados.

Introdução, Problema e Justificativa

Há uma crescente necessidade por melhorias na popularização da ciência no sentido de intensificar o acesso, compreensão e acompanhamento da produção científica por toda sociedade.

Nos diversos países, incluindo os que compõem o Mercosul, já há uma vasta produção de textos e notícias de divulgação científica e tecnológica veiculados on-line em diversas páginas eletrônicas institucionais, como de: universidades, instituições de pesquisa, agências de fomento à pesquisa e outras entidades que públicas e privadas envolvidas com produção científica e inovação tecnologia. Essas notícias, ao contrário de um artigo científico, por serem menos densas e construídas numa linguagem menos técnica, tornam-se de mais fácil leitura àqueles que não especialistas da área, sem perder a credibilidade e qualidade. Por apresentarem tais características, as notícias,

portanto, cumprem o papel importante de divulgação e popularização do conhecimento científico.

A ampla e crescente quantidade de notícias disponíveis nessas diferentes páginas, além das diversas informações que estão dispostas em toda a web, culminou na necessidade de criação de uma plataforma, que hospede na sua base de dados sobre artigos de divulgação científica, que possuem as vantagens: i) serem mais acessíveis à leitura do que artigos de periódicos científicos. ii) possuírem credibilidade e respaldo institucional de publicação por instituições de ensino ou pesquisa ou fomento à pesquisa ou órgãos (ou entidades) governamentais, conforme TOLEDO (2014).

Uma nova base de dados desenvolvida e disponível on-line pode contribuir e facilitar a busca, o acesso e a leitura de textos de divulgação científica publicados on-line pelas instituições, portanto permite o acompanhamento e o acesso as informações sobre as pesquisas desenvolvidas no país à toda a sociedade.

Para a demonstração do funcionamento da base de dados desenvolvida, para este trabalho foram registrados mais de 60 mil notícias oriundas de mais de 30 instituições de pesquisa e inovação, que estão disponíveis para pesquisa.

A proposta do Base¹ Ciência está em consonância com o argumento de que os resultados da pesquisa devem ser públicos, portanto, o sistema de comunicação deve ser fundamental para aquela.

O desenvolvimento científico e tecnológico é acompanhado de uma crescente necessidade da comunicação entre pesquisadores e sociedade, de forma a se divulgar o conhecimento científico, objetivando, por exemplo, conforme apontado por

Entendo, como aponta Ildeu de Castro Moreira (2006), que é preciso uma formação permanente de toda a população, de modo a aumentar a qualificação geral científico-tecnológica para que se tenha maior participação destes e uma participação mais qualificada, uma vez que, segundo Othon Jambeiro citado em

¹ No início do projeto o nome seria ``Guia Ciência``, sendo modificado para ``Base Ciência``

TOLEDO(2014), os governos sofrem a pressão da população que causam impactos nas decisões que envolvem o direcionamento das políticas públicas de ciência.

Atento a essas observações, é possível identificar que o Brasil criou inúmeras ações, como atenta Ildeu de Castro Moreira (2006), para unir a população a produção científica no país, por meio de 1.um departamento para a popularização da C&T, no Ministério da Ciência e Tecnologia 2.centros e museus de ciência; 3.revistas e *websites*; 4. Livros, conferências populares e outros eventos por todo o país;5. Comitê Temático de Divulgação Científica e editais para divulgação científica e extensão universitária.

Apesar da importância crescente da ciência e tecnologia nos dias de hoje e dos crescentes esforços para a divulgação e popularização da ciência, em países como no Brasil, Carolina Rossetti de Toledo (2014) e também Ildeu de Castro Moreira (2006) ainda pontuam a(o) 1.falta de interesse da população pela Ciência e Tecnologia; 2.ampla desconhecimento da população sobre as instituições que produzem ciência no Brasil, entre pessoas de baixa a alta renda. (Ministério da Ciência e Tecnologia, 2010).

No entanto, Carolina Rossetti de Toledo (2014) aponta que conforme pesquisas de VOGT e col. (2013), e dados sobre percepção pública da ciência (Ministério da Ciência e Tecnologia, 2010), há importância da ciência ao reconhecer que “o conhecimento científico auxilia a capacidade das pessoas para decidir coisas importantes em suas vidas”, conforme citado por Carolina Rossetti de Toledo (2014), onde a autora cita pesquisas realizadas por meio de entrevista, tendo identificado que entre 36,7% dos entrevistados uma das principais razões para a falta de interesse pela Ciência e Tecnologia é o fato de não as compreenderem.

Nota-se, portanto, a necessidade de estratégias para o fortalecimento da divulgação e popularização da ciência no Brasil. “A visibilidade das atividades científicas está ligada à publicização do conhecimento produzido por elas. E há diversos meios de divulgação científica, bem como diferentes linguagens, levando-se em consideração o público que se quer atingir.” (SOYAMA, 2006).

Há, por outro lado, uma demanda por divulgação de informação mais qualificada, como ressalta a pesquisadora do Labjor Ana Paula Morales citada por Carolina Rossetti de Toledo (2014).

Sant'Ana (2008, p. 145) observa que a intensificação do uso de tecnologias de informação e comunicação está acompanhada do crescente aumento nos fluxos informacionais gerados pela evolução da internet, gerando um grande e quase ilimitado volume de informações armazenadas e disponíveis. Diante dessa vasta quantidade de dados, a autora adverte que é fundamental para o acesso da população, a melhoria dos mecanismos de buscas, que devem funcionar de modo cada vez mais eficientes.

Se saber procurar no acervo bibliotecas e livrarias era imprescindível para encontrar os textos impressos, no caso dos textos digitais, também é preciso saber buscar em ambientes virtuais, o que requer do leitor aprender modos de fazer boas buscas. (Moura e *col.*, 2012)

No ambiente digital, para MOURA e *col.*, um bom buscador deve possuir as seguintes características: i) avaliar a confiabilidade das fontes e selecionar informações das fontes escolhidas. ii) reconhecer o site como portador de informações confiáveis a respeito do assunto pesquisado. iii) ser crítico, sabendo distinguir o que vai ser relevante para as suas pesquisas. iv) escolher com cuidado e acuidade as palavras e expressões usadas nas buscas, para que elas sejam o mais precisas possível.

No contexto brasileiro, é notável a presença de fontes de divulgação científica impressa e on-line originárias tanto de instituições privadas quanto públicas. Com relação às fontes de divulgação de origem pública, aqui entende-se como revistas e portais vinculados a i) instituições públicas de ensino e pesquisa (universidades e institutos), ii) agências de fomento públicas de pesquisa ou ainda iii) órgãos/secretarias/entidades governamentais.

É relatado por SOYAMA (2006) que nas fontes de divulgação científica originárias de instituições públicas há uma maior aproximação entre divulgador e pesquisador, bem como nestas fontes pode haver outros aspectos que atendam às já citadas demandas populares por fontes “confiáveis” de informação científica na web (TOLEDO, 2014).

De fato, há uma grande variedade de periódicos de divulgação científica veiculados on-line por fontes públicas no Brasil e em outros países.

Disto, pode-se inferir que também pode haver uma reduzida leitura de artigos veiculados pelas já citadas fontes públicas de divulgação científica. Uma

das razões para isto pode ser a limitação dos motores de busca populares amplamente utilizados (conhecidos como “sites de busca”) em apontar páginas de artigos de divulgação científica vinculadas a universidades, institutos e agências de fomento à pesquisa ou órgãos governamentais.

Em buscadores populares, como o *Google*, por exemplo, utilizando como palavras-chave termos ligados a assuntos científicos de diferentes níveis de especificidade resultam em diversas ocorrências, não raro há poucas e nenhuma que vincule a uma fonte consolidada de divulgação científica no Brasil, como universidades, institutos e agências de fomento públicos, órgãos governamentais.

Apesar de haver na web uma grande variedade de bases de dados para a busca por artigos científicos de periódicos indexados, ao que consta após extensas pesquisas, não há em língua portuguesa uma base de dados para a busca específica por notícias de divulgação científica veiculadas por revistas, jornais e portais de instituições de ensino, pesquisa e inovação tecnológica públicas (ou privadas). Tampouco, é possível encontrar em países ibero-americanos uma base de dados tal qual a proposta pelo Base Ciência, que nada mais é do que um mecanismo para aproximar a sociedade de fontes de divulgação científica legítimas e de qualidade.

Objetivos

O objetivo principal deste trabalho é descrever a base de dados inovadora “Base Ciência” como uma fonte para apoio didático em diferentes contextos escolares, acadêmicos ou não acadêmicos. Esta base de dados deve hospedar dados (títulos/subtítulos de artigos, Fonte dos artigos, e Link para acesso à fonte do artigo) apenas de artigos de divulgação científica veiculados em publicações (revistas, jornais e portais), vinculados as instituições públicas brasileiras de ensino e/ou pesquisa, agências de fomento à pesquisa ou à órgãos governamentais, ou ainda de fontes privadas com notável consolidação na área de divulgação científica no Brasil.

A busca por palavra-chave deve resultar em uma exibição específica de títulos dos artigos, nome da fonte (Revista, Jornal, Portal ou Instituição) e links para acesso às páginas das fontes dos artigos de divulgação científica encontrados pela busca. A base de dados deve possuir acesso gratuito através do endereço www.baseciencia.com.br, de forma que o usuário digite no buscador a(s) palavra(s)-chave representativa do tema a ser pesquisado.

Com o desenvolvimento desta base de dados, é esperado facilitar a busca seletiva no nicho de fontes citadas por temas ligados a todas as áreas do conhecimento científica e tecnológico.

O objetivo do desenvolvimento desta base de dados é complementar as bases de dados científicas existentes, possibilitando uma opção de busca que exiba como resultados páginas que não seriam exibidas como primeiras ocorrências em motores de busca populares utilizando-se as mesmas palavras-chave.

Com isto, visa-se principalmente a intensificação da leitura de textos de divulgação científica veiculados nas fontes citadas, possibilitando um maior acesso e acompanhamento da sociedade diante das pesquisas realizadas ou financiadas por instituições públicas. Objetiva-se também tornar tais artigos de divulgação científica mais utilizados como materiais para apoio didático, conforme já relatada tal utilização na literatura (FATARELI e *col.*, 2015, GOMES e *col.*, 2016).

Apesar do presente trabalho ter como um dos objetivos registrar informações sobre instituições brasileiras, a base de dados desenvolvida deve possuir espaço disponível para que seja possível registrar as informações (títulos/subtítulos de artigos, Fonte dos artigos, e Link para acesso à fonte do artigo) de todos os artigos de divulgação científica publicados on-line por todas as instituições públicas de pesquisa, pertencentes a todos os países membros do Mercosul.

Neste trabalho, o objetivo é demonstrar o modo de funcionamento, funcionalidades e dados estatísticos para a base de dados desenvolvida (Ciência) com o registro de mais de 60 000 artigos de divulgação científica de 30

(noventa) instituições públicas brasileiras do nicho citado. A expansão do conteúdo (registros das informações de artigos) da base de dados pode tornar-se contínuo a medida que periodicamente são publicados mais artigos nas fontes citadas, ou ainda, como citado, podem ser incluídos registros de artigos de fontes de outros países do Mercosul na base de dados Guia Ciência.

Outro objetivo deste trabalho é realizar atividades de popularização da ciência na Semana Nacional de Ciência e Tecnologia.

Metodologia

A base de dados Guia Ciência constitui a estrutura lógico-matemática em que o sistema de busca foi desenvolvido e como foram registradas as informações, de que maneiras as informações sobre os artigos podem ser buscadas. Já o banco de dados que constitui os dados registrados dos artigos e que podem ser acessados através do buscado, ou seja, o banco de dados constitui o conteúdo que pode ser buscado (SILVA, 2010).

Desenvolvimento da Base de Dados:

O site www.baseciencia.com.br foi desenvolvido através de linguagem PHP (abreviação de Hypertext Preprocessor), utilizando Apache Tomcat (um software livre) e SSH (abreviação do programa e protocolo de rede Secure Shell). O site possui um banco de dados em planilha, onde são armazenados os dados de cada artigo registrado: Título (e subtítulo caso houver), Link (url), data (de publicação na fonte original) e nome da fonte (revista ou jornal ou portal). Assim, cada linha da planilha no banco de dados corresponde a um registro de um artigo. O banco de dados tem a capacidade para registro de 9 000 000 000 (nove bilhões) de linhas, podendo armazenar assim o registro dos dados (título, link, data e nome da fonte) de nove bilhões de artigos distintos.

O algoritmo do buscador foi criado a partir do início para uso especificamente no Guia Ciência. O algoritmo de busca possui dezenas de linhas em PHP com configuração Tomcat e nGinx. O algoritmo de busca do Guia Ciência é dividido e disponibilizado em três opções de busca para o usuário:

“Buscador palavra exata”, “Buscador por trecho de Palavra”, “Fonte” e “Data”. Há também disponível para o usuário o algoritmo de busca comercial GoogleWebSearch, outro algoritmo para busca no banco de dados do Guia Ciência.

Métodos e Materiais para a Aquisição dos Dados de cada artigo (título, link, data e nome da fonte)

Inicialmente foram determinadas as mais de 40 fontes de publicação de artigos de divulgação científica de onde seriam coletados os dados de artigos. O critério para a escolha das fontes (instituições, jornais, revistas, portais) de divulgação científica a constar no banco de dados do Base Ciência foi: fontes de divulgação científica vinculadas às principais instituições públicas de ensino e/ou pesquisa brasileiras e/ou agência de fomento à pesquisa brasileiras e/ou órgãos (ou entidades) governamentais brasileiras. Prioridade foi dispensada para a aquisição do maior número de registros de artigos ligados a fontes com maior destaque na publicação de artigos de divulgação científica no Brasil, com maior número de edições publicadas até 2015. Apesar de no momento o banco de dados do Base Ciência registrar artigos apenas de instituições públicas brasileiras, deve ser notado que artigos de instituições como a “Scientific American Brasil” e “Ciência Hoje” são interessantes para constar no banco de dados do Base Ciência devido à sua grande relevância, pioneirismo e notáveis contribuições à divulgação científica (FREIRE Jr., 2005).

Para a aquisição dos dados de cada artigo registrado (título, link, data e nome da fonte) no banco de dados, foi utilizado como material: Computador com processador Arquitetura X86 - Intel Core2Duo ou superior 2GB Ram, 1GB Livre de HD, Conexão à internet, Sistema Operacional, Ubuntu 14.04. Softwares: python 2.7 e bibliotecasos, sys, time, datetime, csv, re, click, pprint, requests, lxml, logging, pdftotext, LibreOfficeCalc.

Para a aquisição dos dados de cada artigo registrado (título, link, data e nome da fonte) no banco de dados, foram utilizados os seguintes procedimentos, que são representativos para todas as fontes pesquisadas:

Procedimento 1)

O algoritmo python baixa as páginas em formato XML (padrão RSS), iterando pelas páginas através de um valor na url, até que não haja mais notícias nessas páginas. Então transforma os dados baixados numa árvore XML e navega entre seus diversos nós, buscando as informações específicas de url, título/subtítulo, e data de publicação. Essas informações são salvas numa lista, e em seguida passam por um filtro para remoção de seções indesejadas ou incompletas, através de regras de expressão regular.

Finalizado esse filtro, a lista é salva em um arquivo CSV (commaseparatedvalues), e finalmente convertida manualmente em formato .xls (Planilha Microsoft Excel).

Procedimento 2)

Os algoritmos python, a partir de endereços iniciais fixos em seu código-fonte, baixa a página que contém a lista de edições dos sites, em formato html. Em seguida, transformam a página html em uma árvore navegável de html através de xpath (método de caminhos de navegação em árvore xml e html) e encontram os links para as edições. Assim, baixam então as páginas de cada edição e refazem a mesma tarefa de transformação do arquivo html em uma árvore e buscam blocos específicos de código, onde ficam as informações das notícias de cada edição. Encontrados esses blocos, os algoritmos fazem a extração dos nós pertinentes, retirando então o texto contido neles para o preenchimento de uma lista intermediária. Pronta a lista intermediária, esta tem seus registros filtrados através de expressões regulares, a fim de se retirar textos espúrios ou reportagens não necessários. Por fim, a lista é salva como um arquivo csv, que em seguida é conferido e transformado manualmente para .xls através do LibreOfficeCalc.

Procedimento 3)

Semelhante ao algoritmo da Unesp Ciência, porém com maiores nuances. Enquanto as tecnologias são as mesmas, são necessárias regras específicas para os diversos tipos de páginas que se apresentam no site do Jornal da Unicamp ao longo dos anos. Há pelo menos duas versões de páginas para se

obter as edições de cada ano. Obtidas as páginas com as listas de edições em html, estas são transformadas em árvores navegáveis através de xpath, para se encontrar os links que levam a cada edição. Encontrados esses links, ao fazer o download das páginas, é necessária uma conjunção de expressão regular e xpath para direcionar a página para a versão do algoritmo que atende àquela página (são três versões de extratores de links de notícias de dentro das páginas). Direcionadas as páginas aos seus respectivos extratores, estes transformam o arquivo obtido em uma árvore de html navegável e buscam por blocos html que contêm as informações de título e link. Obtidos todos os títulos e links numa lista, esta é iterada de modo que os algoritmos então possam baixar o arquivo html correspondente e buscar dentro desses arquivos, através de caminhos xpath e expressões regulares, os blocos de texto que correspondem ao subtítulo dessas notícias. Terminada essa etapa, a lista final de notícias passa por filtros de texto para remoção de caracteres espúrios, espaços e quebras de linha indesejados, e por fim salva num arquivo csv. Esse csv resultante então é conferido e transformado em um arquivo padrão Microsoft Excel (.xls) manualmente através do software LibreOfficeCalc.

Procedimento 4)

Semelhante ao Jornal da Unicamp em um dos algoritmos de extração (através de html), também faz uso de um algoritmo através de transformação de pdf em texto. É necessário nesse método novo que se baixe manualmente os pdfs da página de edições do site da Fapemig, assim como se criar uma lista com os links correspondentes a cada edição em pdf. Em seguida um algoritmo em python, fazendo uso do software pdftotext, transforma em texto o pdf das edições. Deve-se notar que a transformação em texto não é perfeita e algumas alterações manuais devem ser feitas nos arquivos em texto antes que sejam utilizados no algoritmo. Terminada a etapa de transformação e ajustes manuais, o algoritmo procura no texto, através de expressões regulares, as marcações de quebra de página e em seguida, qual a página de índice para que possa, a partir dessa, extrair as categorias e títulos das notícias da edição, também através de expressões regulares. A partir dos textos dos títulos brutos, o algoritmo filtra

algumas ocorrências encontradas, para que sejam removidos textos e números espúrios. Finalizada a extração, a lista gerada é enviada para um arquivo csv para que este em seguida, manualmente, seja transformado num arquivo xls.

Há de se notar que em todos os algoritmos, de todas as fontes, são necessárias várias iterações de análise dos resultados, visto que devido à natureza humana da geração das marcações html ou da diagramação dos textos em pdf, os arquivos não seguem exatamente os mesmos padrões, sendo necessárias então regras mais inclusivas e casos particulares para a obtenção de todos os textos de maneira correta.

Resultados e Discussão

Os métodos citados puderam ser utilizados de forma a se registrar informações sobre um total de mais de 60000 (sessenta mil) artigos distintos de divulgação científica sobre resultados de pesquisa científica. São quatro as informações registradas na base de dados sobre cada artigo são: 1) Título e Subtítulo, 2) Fonte (Nome da Revista, Jornal ou Seção do Portal em que os artigos são veiculados), 3) Link url e 4) data da publicação do artigo. Cada conjunto destas quatro informações citadas para cada artigo constitui o que é denominado neste trabalho de “registro”. Cada registro é, então, um conjunto de quatro informações de um artigo: Título/Subtítulo, Link url, Fonte e data.

A base de dados Base Ciência, desenvolvida no presente trabalho, foi programada de modo a funcionar através de um buscador onde o usuário realiza buscas ao digitar palavra(s)-chave. Após digitar a(s) palavra(s)-chave, a busca na base de dados é feita pressionando-se a tecla “enter” de um teclado computacional ou clicando-se no link à direita do buscador. No contexto de uma busca por palavra(s)-chave qualquer, cada “registro” é denominado de “ocorrência” quando aparece como resultado de uma determinada busca. Através da primeira opção de busca, automaticamente pré-selecionada ao se acessar o site, o algoritmo do buscador do Guia Ciência realiza as buscas verificando se a(s) palavra(s)-chave digitada no buscador também está presente em algum título ou subtítulo de algum artigo registrado na base de dados. Em

caso afirmativo, o registro do artigo é exibido como ocorrência através da exibição de seu título/subtítulo, fonte, data e também o link url, que está implicitamente inserido na mensagem “Clique para Ler”, que aparece após a data de cada ocorrência. O usuário, ao clicar em “Clique para Ler” será redirecionado, através da abertura de uma nova aba do navegador de internet, para a página original onde o artigo foi publicado. O mesmo acesso à página original é também conseguida através de um clique sobre o próprio título/subtítulo da ocorrência. Para uma pesquisa qualquer no buscador do Guia Ciência, as ocorrências foram programadas para serem listadas por prioridade de data, as mais recentes com prioridade, ou seja, aparecendo antes das menos recentes. Cada ocorrência recebe um número de 1 até o número total de ocorrências exibidas, que será o número total de ocorrências para uma determinada busca. Este número, chamado número da ocorrência, aparece seguido de um caracter “)”, antes (à esquerda) do título da ocorrência, e sua utilidade está em: i) fornecer um dado de organização para o usuário em suas pesquisas e na comunicação de suas pesquisa com outros; ii) fornecer de maneira prática o número total de ocorrências para uma determinada busca, que será o número da última ocorrência listada.

Conforme pode ser observado à direita do espaço onde se digita(m) a(s) palavra(s)-chave no buscador, há um selecionador de opções de busca, onde a busca pré-selecionada automaticamente é a busca “Palavra-Chave, Buscador 1”. Esta opção, como citado, busca a(s) palavra(s)-chave escrita no buscador nos títulos/subtítulos dos registros na base de dados. Porém, é possível buscar também por fonte, ou seja, nome do jornal ou revista ou portal em que artigos possam ser publicados. A utilidade da busca por fonte consiste, por exemplo, na possibilidade de qualquer usuário poder verificar o número total de artigos de uma determinada fonte que podem ser apontados pelo Guia Ciência. Assim, por exemplo, ao se digitar no buscador “Jornal da UNICAMP” (sem as aspas), selecionando a opção de busca “Fonte (Nome da Revista, Jornal ou Portal)”, qualquer usuário pode conferir o número total de registros de artigos do Jornal da UNICAMP que constam na base de dados Guia Ciência: basta observar o número da última ocorrência desta busca, conforme já explicado. O número total de registros de artigos por fonte está exibido na tabela a seguir, onde são

também listados outros dois dados para cada fonte cujas informações dos artigos estão registrados no Base Ciência: i) método de aquisição dos dados de cada registro para cada artigo, ou seja, método com que as informações Título/Subtítulo, Link url, Fonte, Data, foram extraídas dos artigos de cada fonte. ii) período em que os artigos de cada fonte foram publicadas na fonte, ou seja, data de publicação (na fonte) do artigo mais antigo e data do artigo mais recente registrado para cada fonte na base de dados. Note que os dados da segunda coluna podem ser conferidos por qualquer um que acessar www.baseciencia.com.br, bastando para tal: selecionar a busca por “fonte” no buscador do Base Ciência e digitar no buscador o nome da fonte que consta na primeira coluna da tabela. Ou seja, o número total de artigos que constam registrados na base de dados para uma fonte será igual ao número da última ocorrência listada ao se realizar a busca por nome da fonte (na opção “Fonte” do buscador), digitando-se o nome da fonte da primeira coluna.

Como citado, a utilidade do Base Ciência está, dentre outros objetivos, em auxiliar na busca específica por artigos de divulgação científica veiculados on-line por instituições públicas de ensino e/ou pesquisa, agências de fomento à pesquisa e órgãos governamentais. Assim, torna-se interessante uma análise do Base Ciência com relação ao número de ocorrências exibidas ao se buscar palavra(s)-chave representativas de temas científicos os mais variados possíveis. Conforme citado por TOLEDO (2014), estudos recentes apontam para a notável preferência de populações na pesquisa, em textos de divulgação científica, por temas ligados à área da saúde. Assim, a tabela em anexo expõe uma lista de palavra(s)-chave de temas ligados à área da saúde. Na segunda coluna da tabela, são exibidos os números totais de ocorrências listadas pelo Base Ciência ao se realizar uma busca por cada palavra-chave da primeira coluna (utilizando a opção do buscador “Palavra-chave, buscador 1”).

A partir da tabela anterior exposta, torna-se relevante analisar, para uma determinada palavra-chave, quais os títulos/subtítulos e fontes de cada um dos artigos apontados como ocorrências pelo Base Ciência. Após análises da duas tabela II em anexo, pôde-se concluir que para todas as palavras-chave listadas na tabela II, nenhum dos resultados de busca foi exibido por nenhum dos três grandes motores de busca popular como as primeiras 20 (vinte)

ocorrências ao se digitar as mesmas palavra(s)-chave em cada um destes 03 motores de busca. Assim, nota-se a especificidade do Base Ciência e exclusividade em encontrar (apontar) artigos de fontes de divulgação científica pertencentes aos padrões já citados. Contudo, dada a extensão do conhecimento científico e também dado o número total de registros de artigos no Base Ciência (~ sessenta mil), pode ser argumentado que a amostragem exibida nas tabelas X, Y, Z (ver também anexos) é insuficiente para conclusões sobre a exclusividade do Base Ciência, em relação a motores de busca populares, em encontrar artigos de divulgação científica de fontes pertencentes aos critérios já citados. Uma maneira de se verificar tal conclusão é o próprio leitor atribuir palavra(s)-chave representativas de tópicos (temas) de qualquer área do conhecimento científico, e realizar a busca comparativa no Base Ciência e em motores de busca populares, para cada uma das palavra(s)-chave atribuídas arbitrariamente pelo leitor. Para orientar o usuário da base de dados Base Ciência, torna-se relevante expor a lista de instruções e recomendações que consta na página inicial do Base Ciência, objetivando informar o usuário para o melhor uso possível da base de dados e maior eficácia nas pesquisas:

- Quando buscar por mais de uma palavra-chave, é recomendado buscar por cada uma das palavras-chave separadamente.
- Verifique o uso correto de acentos gráficos e realize a busca com os devidos acentos gráficos. Experimente também a busca da(s) mesma(s) palavra(s) sem acentos gráficos, pois de maneira menos frequente pode haver equívocos na escrita inclusive da artigo na fonte original.
- Ao buscar por uma determinada palavra-chave, sempre considere buscar separadamente por palavras-chave semelhantes, derivações da palavra-chave inicial ou sinônimos:
 - Exemplo: Ao buscar por Amazônia, analise as ocorrências encontradas e, após, busque separadamente também por derivações como “Amazônica”, “Amazônico”, “Amazônicas”, “Amazônicos”, dentre outras possíveis derivações, sinônimos e palavras semelhantes. Assim como, ao buscar, por exemplo, por: “boi”, analise as ocorrências encontradas nesta busca e, após, busque separadamente também por: “bois”, “bovino”, “bovina”, “gado”.

- Ao buscar por uma determinada palavra-chave, considere sempre também buscar adicionalmente, separadamente, por trechos da palavra-chave. Exemplo: ao buscar por “nanotecnologia”, analise os resultados e realize nova busca por “nano”; “nanotech”; “nanociência”.

- Em casos raros (menos que em 10 % das buscas testadas), o link da ocorrência pode não levar para a página correta de destino. Assim, basta que seja digitado o título da ocorrência no buscador de algum motor de busca popular (para encontrar a página correta logo nas primeiras ocorrências exibidas pelo motor de busca popular.

A base de dados Base Ciência foi desenvolvida de forma que as palavras-chave buscadas encontrarão resultados apenas se estiverem contidas, em todo ou em parte, nos títulos ou subtítulos dos artigos registrados na base de dados. Ou seja, o buscador do Base Ciência não realiza uma busca em cada palavra dos artigos registrados na base de dados. O buscador realiza busca no título e subtítulo dos artigos, e isto é condição satisfatória para grande utilidade devido aos artigos registrados no Base Ciência pertencerem, em sua totalidade, a um estilo literário de divulgação científica onde as palavras-chave do artigo constam no título e/ou subtítulo do artigo. Porém, ainda torna-se relevante para aperfeiçoamentos futuros da base de dados Base Ciência, que o buscador realize, além da busca pela palavra-chave no título/subtítulo do artigo, busca também em qualquer trecho dos artigos originais registrados na base de dados Base Ciência.

Outro resultado deste projeto foi a ministração do mini-curso “Noções de Espectroscopia e Percepção das Cores utilizando diferentes materiais para apoio didático”. Este mini-curso foi realizado durante a Semana Nacional de Ciência e Tecnologia conforme documentação em anexo, com a participação de discentes e técnicos da UNIR, no Campus José Ribeiro Filho, totalizando 08 horas de atividades teóricas e experimentais. As atividades teóricas ocorreram na forma 02 dias de palestra pelo coordenador do presente projeto. Já a atividade experimental ocorreu na Central Analítica do Departamento de Química da UNIR, utilizando um aparelho espectrofotômetro UV-Vis, tendo sido realizado um experimento demonstrativo de medida de coeficiente de

absortividade molar do permanganato de potássio, com projeção em Datashow ligado ao aparelho para exibição dos dados experimentais conforme os mesmos eram medidos. A documentação comprobatória esta em anexo.

Conclusões

Diante do exposto em todos os tópicos, algumas discussões conclusivas e perspectivas podem ser agrupadas a seguir:

- i) Até o momento, após 01 ano (um ano) de contínuas e extensas pesquisas semanais em páginas de artigos e notícias de dezenas de portais de universidades, portais de divulgação científica público e privados, portais de órgãos e institutos ligados à órgãos governamentais nas áreas de educação, ciência e tecnologia, agências de fomento à pesquisa, não foi encontrado nenhum relato de uma base de dados, em língua portuguesa ou espanhola, que tenha funcionalidade igual ou com similaridade à funcionalidade do Base Ciência: auxiliar na busca por artigos de divulgação científica veiculados em revistas, jornais e portais de universidades públicas e institutos públicos de pesquisa, agências de fomento (públicas) à pesquisa e órgãos governamentais (ou entidades ligadas à órgãos governamentais).

- ii) Após 08 meses com a base de dados funcionando nos endereços www.baseciencia.com.br ou em www.baseciencia.com.br/index.php, conclui-se que a base de dados e o site que a hospeda são estáveis, isto é, através de acessos semanais após 08 meses, em menos de 1 % dos acessos o site estava fora do ar ou precisou ser retirado do ar. Assim, o site www.baseciencia.com.br apresenta robustez e estabilidade desejável para um site que possa atender com eficácia e satisfação os mais exigentes públicos-alvo: desde cidadãos com interesse esporádico por ciência, até pesquisadores que fazem pesquisas diárias por notícias relacionadas a ciência e tecnologia.

- iii) A partir de todas as tabelas e dados expostos, a quantidade e diversidade dos artigos registrados permite buscas com resultados em todas as áreas do conhecimento, em tópicos com diferentes níveis de especificidade, como pode ser observado nas tabelas expostas neste trabalho. A base de dados é expansível, ou seja, podem ser adicionados até números na casa de bilhões de artigos. Assim, dentro da metodologia e procedimentos computacionais em que a base de dados foi criada, é possível que sejam registrados e disponíveis para a busca no Base Ciência a totalidade de artigos e/ou notícias de divulgação científica do nicho de instituições públicas de todo o bloco Mercosul ou do Brasil. Assim, pelos dados expostos, a utilização do Base Ciência pode contribuir para: i) Crescimento na procura, acesso e leitura de textos de divulgação científica veiculadas no nicho de instituições públicas citadas. ii) Crescimento no acesso de artigos de divulgação científica por cidadãos de nacionalidade distinta ao país sede da instituição que publica o artigo. Considerando o Mercosul, os países de língua espanhola podem realizar uma operação prévia de tradução espanhol-português com a palavra-chave a ser pesquisada. Para exemplificar, um cidadão de língua espanhola que objetivar pesquisas com a palavra-chave “naranja”, deve digitar esta palavra-chave em um dos eficazes tradutores on-line livremente disponíveis. O tradutor deverá retornar o termo “laranja”, que deverá ser então digitada no buscador do Base Ciência. É possível também que para trabalhos futuros sejam acoplados tradutores à base de dados BaseCiência. Porém, para o momento, a ágil operação prévia de tradução descrita e exemplificada deve ser eficaz na quase totalidade dos casos de pesquisa no Base Ciência.
- iv) O modo de programação como foi desenvolvida a base de dados Base Ciência também permite a emissão automática de relatórios que exibam as estatísticas completas de palavras-chave pesquisadas: que palavras-chave foram pesquisadas, quando foram pesquisadas, quantas vezes foram pesquisadas e em que região geográfica (cidade e país) foram pesquisadas. Estas informações podem ser coletadas

através de softwares e plataformas livremente disponíveis on-line. O modo de programação do Base Ciência permite também a emissão automáticas de relatórios estatísticos completos sobre onde, quando e quantas vezes foram acessados cada um dos links de todos os artigos registrados na base de dados. A coerência e autenticidade destes dados pode ser por completo verificadas através de comparação com as estatísticas de acesso nas fontes, ou seja, observando as estatísticas de acesso coletadas e atestadas pelos sites de cada uma das instituições (universidades, institutos, agências, órgãos e entidades governamentais) que hospedam o artigo completo original, que é acessado pelo usuário do Base Ciência pela respectiva url (link). Neste ponto, é interessante notar que o Guia Ciência não hospeda nenhum artigo completo, portanto o Base Ciência apenas facilita a procura e o acesso aos artigos cujos títulos/subtítulos constam na base dados, sempre acompanhados da devida citação da fonte institucional a que o artigo pertence. Deste modo, foram observadas e respeitadas todas as normas e leis de direito autoral, além de não haver nenhum anúncio de nenhuma instituição em nenhuma das páginas do Base Ciência. Deve ser citado também que não houve exploração comercial e nenhum fim lucrativo para o site do Base Ciência. O desenvolvimento do Base Ciência foi possível através do trabalho do presente autor com recursos aprovados em edital de fomento à pesquisa por instituição pública brasileira, descrito na seção “Agradecimentos”.

- v) A base de dados Base Ciência constitui um mecanismo de busca para a utilização de artigos de divulgação científica de qualidade como materiais para apoio didático.
- vi) As atividades da SNCT, objetivadas paralelamente neste projeto para a popularização da ciência, foram realizadas em 2015, de forma a se abordar tema relevante para a popularização da ciência.

Referências Bibliográficas

BARBOSA SILVA, E. R. O retrato das atividades no Brasil. *ComCiência*, n. 100, Campinas 2008.

FREIRE JR., O. Divulgação científica globalizada: potencialidades e riscos. *ScientiæStudia* (USP), São Paulo, v. 3, n. 4, p. 711-4, 2005.

FATARELI, E. F. ; MASSI, L. ; DE ABREU FERREIRA, L. N. ; QUEIROZ, S. L. Mapeamento de textos de divulgação científica para planejamento de debates no ensino de Química. *Química Nova na Escola*, v. 37, no 1, p. 11-18. 2015.

GERMANO, M. G. Uma nova ciência para um novo senso comum. Capítulo IV Popularização da ciência e tecnologia: limitações e possibilidades. Campina Grande: EDUEPB, 2011. 400 p.

GOMES, V. B. ; RIBEIRO DA SILVA, R. ; MACHADO, P. F. L. Elaboração de textos de divulgação científica e sua avaliação por alunos de Licenciatura em Química. *Química Nova na Escola*, v. 38, no 4, p. 387-403. 2016.

MOREIRA, I. C. A inclusão social e a popularização da ciência e tecnologia no Brasil. Ildeu de Castro Moreira. *Inclusão Social*, v. 1, n. 2, 2006.

MEZZACAPPA, M. Jornais latino-americanos não divulgam ciência dos países vizinhos. *ComCiência – Revista Eletrônica de Jornalismo Científico*. Notícias 07/02/2008. Disponível em <http://www.comciencia.br/comciencia/?section=3¬icia=393>. Acessado em 23/02/2015.

MINISTÉRIO DA CIÊNCIA E TECNOLOGIA. Percepção Pública da Ciência e Tecnologia no Brasil: Resultados da Enquete de 2010. Departamento de Popularização e Difusão da C&T/SECIS/MCT. Museu da Vida/COC/Fiocruz. 2011, disponível em: http://www.mct.gov.br/upd_blob/0214/214770.pdf. Acessado em 23 de fevereiro de 2015.

MOREIRA, I. C. A divulgação da ciência e da tecnologia no Brasil. *Revista da Universidade Federal de Minas Gerais*. Ano 7, n. 13, fevereiro/2008.

TOLEDO, C. R. Para melhor informar: Dois novos estudos, um da Unicamp e outro da Academia de Ciências da Bahia, ampliam a compreensão sobre a percepção pública da ciência no país. *Revista Pesquisa FAPESP*, ed. 217, março/2014.

MOURA, M. A. (org.) Educação científica e cidadania: abordagens teóricas e metodológicas para a formação de pesquisadores juvenis. Belo Horizonte: UFMG / PROEX, 2012. 280 p.: il. (Diálogos, 2).

SILVA, J. A. B. E. Agenda -Setting assente em bases de dados e algoritmos: bases conceituais e metodológicas para operacionalizar a relevância de temas, predicados e agendas entre usuários da web. Tese (Doutorado em Comunicação e Cultura Contemporâneas) - Faculdade de Comunicação. Universidade Federal da Bahia. Salvador, 2010.

SOYAMA, P. Diferentes Discursos encontram nicho na divulgação da ciência. *ComCiência – Revista Eletrônica de Jornalismo Científico*. Notícias 23/05/2006. Disponível em <http://www.comciencia.br/comciencia/?section=3¬icia=122&print=true>. Acessado em 23 de fevereiro de 2015.

VOGT, C.A. (Org.) ; DIAS, S. (Org.) ; FIGUEIREDO, S. P. (Org.) ; BARATA, G. F. (Org.) ; KANASHIRO, M. M. (Org.) . Comunicação, divulgação e percepção pública de ciência e tecnologia. 1. ed. Petrópolis/Brasília: De Petrus et Alii/Capes, 2013. v. 1. 180p.