

GMP: Um Software para Cálculos de Indicadores de Qualidade e Quantidade de Produção Científica

Felipe de Melo Taveira

Graduando em Bacharelado em Ciência da Computação pela UNIFAL-MG - Universidade Federal de Alfenas. E-mail: felipe.taveira@bcc.unifal-mg.edu.br

Lucas Oliveira Nabeto

Graduando em Bacharelado em Ciência da Computação pela UNIFAL-MG - Universidade Federal de Alfenas. E-mail: lucas.nabeto@bcc.unifal-mg.edu.br

Eric Batista Ferreira

Doutor em Estatística e Experimentação Agropecuária. Professor do Departamento de Estatística da UNIFAL-MG - Universidade Federal de Alfenas. E-mail: eric.ferreira@unifal-mg.edu.br

Luiz Eduardo da Silva

Doutor em Ciências em Engenharia Elétrica. Mestre em Ciência da Computação. Professor do Departamento de Ciência da Computação da UNIFAL-MG - Universidade Federal de Alfenas. E-mail: luizedu@bcc.unifal-mg.edu.br

Flávio Bittencourt

Professor do Departamento de Estatística da UNIFAL-MG - Universidade Federal de Alfenas. E-mail: bflavioo@gmail.com

A escassez dos recursos de financiamento de pesquisa vem deixando o ambiente acadêmico cada vez mais concorrido. Editais de diversas naturezas classificam pesquisadores, individualmente, através de métricas indicadoras de sua produtividade, porém, faz-se necessário o gerenciamento dessas métricas para classificar um grupo de pesquisadores. Para tal fim, este artigo visa apresentar o software GMP, que utiliza o Google Acadêmico, uma forma gratuita e confiável de acessar informações dos trabalhos publicados por qualquer pesquisador, como fonte de dados para cálculos de métricas coletivas.

Palavras-chave: Google Acadêmico; Grupos de pesquisa; Métricas de produtividade; *Web crawler*.

GMP: A Software for Calculations of Quality and Quantity Indicators of Scientific Production

The scarcity of research funding has been making the academic environment increasingly competitive. There are edicts of various natures that individually rank researchers through some indicative metric of their productivity. However, it is necessary to manage such metrics to rank a group of researchers. To this end, this project aims to introduce the software GMP that uses Google Scholar, a free and reliable way to access the information of articles published by any researcher, if this allows, as data source for calculations of collective metrics.

Keywords: Google Scholar; Research groups; Productivity metrics; Web crawler.

1 Introdução

Com o passar do tempo, a quantidade de pesquisas científicas tem se tornado cada vez maior. Entretanto, nem sempre a quantidade de publicações de um pesquisador está diretamente relacionada com a qualidade das mesmas, pois *“a pressão para publicar deixa pesquisadores com tempo insuficiente para pensar sobre as questões importantes que exigem discussão aprofundada e reflexão”* (LOYOLA, 2013).

Visando classificar um pesquisador com base na qualidade de suas publicações, criaram-se algumas métricas baseadas na quantidade de vezes que cada publicação deste pesquisador foi citada por outros trabalhos científicos, pois elas adotam a ideia de que quanto melhor é a pesquisa que gerou a publicação, maior será o número de trabalhos que a mencionarão.

Os dados sobre as atividades científicas estão sendo cada vez mais utilizados para governar ciência e avaliações de pesquisa que foram projetados uma vez individualmente para seu contexto específico, agora são rotina e são baseados em métricas (HICKS et al, 2015).

Então, para que se pudesse ter um certo controle de qualidade sobre as publicações científicas dos pesquisadores, serviços como Google Acadêmico, *Research Gate*, Academia, dentre outros, foram criados. Nesses serviços, é possível que o pesquisador veja quantas vezes seu trabalho ou o trabalho de algum outro pesquisador foi citado por terceiros, mas eles não contam com o recurso de avaliar um determinado grupo de pesquisadores de acordo com as publicações de todos os seus integrantes.

O recurso em questão abriria um novo leque de oportunidades de uso para as informações de uma publicação. Algumas dessas oportunidades poderiam ser uma distribuição proporcional de bolsas de pesquisas estipulada através dos indicadores de qualidade utilizados, ou o ato de incentivar os pesquisadores a escreverem publicações de maior qualidade. *“Os cientistas nos países em desenvolvimento devem aumentar a qualidade de pesquisas publicando mais bons artigos, e não menos”* (LOYOLA, 2013).

Como forma de incentivo aos pesquisadores, permitir que saibam o indicador de qualidade de outros grupos de pesquisa pode resultar em uma competitividade sadia, onde estes fariam pesquisas cada vez melhores, focando mais na qualidade

do que na quantidade de suas publicações científicas. Tal competitividade já é inclusive uma estratégia muito utilizada em grandes empresas, por exemplo, e, segundo Degen e Mello (1989), pode ser considerada como um grande diferencial em diversos cenários. *"Competitividade é a base do sucesso ou fracasso de um negócio onde há livre concorrência. Aqueles com boa competitividade prosperam e se destacam"* (DEGEN; MELLO, 1989, p.106 - 107).

Dito isso, o objetivo deste trabalho é apresentar o software GMP, capaz de colher informações públicas encontradas na internet e permitir o agrupamento de pesquisadores de acordo com a necessidade da situação. Logo após, por meio de algumas métricas já conhecidas no meio acadêmico, é classificada a qualidade das pesquisas de cada um dos grupos selecionados de acordo com as informações obtidas anteriormente.

2 Metodologia

A linguagem de programação C# (C *Sharp*) reúne os melhores recursos de suas principais concorrentes (Java e C++), tornando-se uma linguagem poderosa e de fácil aprendizado.

O C# foi influenciado por inúmeras outras linguagens, dentre elas Delphi, C++ e Java. Com essa influência, o C# passou a ser uma linguagem poderosa, já que foram retirados os melhores recursos de cada uma destas linguagens, além de quebrar as limitações existentes. Por ser baseada em Delphi, C++ e Java, ela acaba sendo bastante atrativa para programadores iniciantes, pois estes têm uma base acadêmica fundamentada nestas linguagens (SANTOS, 2011).

Portanto, tendo o C# como a linguagem de programação a ser utilizada neste trabalho, são realizados procedimentos para buscar informações sobre os trabalhos científicos dos pesquisadores desejados. E, de posse de informações como título do trabalho, quantidade de vezes que o trabalho foi citado e ano de publicação do trabalho, são realizados cálculos de métricas que serão úteis para gerar uma classificação dos pesquisadores.

A técnica necessária para se obter as informações supracitadas é o *web crawler*, que pode ser definido por Olston e Narjork (2010) como um meio de se escolher porções específicas do conteúdo de uma página Web.

Web crawler é uma técnica para efetuar download a granel de páginas da web. O *web crawler* é usado para vários fins. Normalmente, eles são um dos principais componentes dos motores de busca na web, sistemas que reúnem uma coletânea de páginas da web, indexam-nas e permitem que os usuários emitam consultas [...] (OLSTON; NAJORK, 2010, p.176).

O *web crawler* é dividido em duas etapas: a de obtenção de todos os dados da página visitada e a de manipulação dos dados colhidos (SOARES, 2008).

No caso deste trabalho, as "páginas visitadas" são os perfis dos pesquisadores no Google Acadêmico, que, dentre todos os serviços mencionados na seção anterior, é uma forma gratuita e confiável de se obter os dados de cada uma das publicações científicas pertencentes aos pesquisadores que serão

avaliados, se este permitir. O Google Acadêmico pode ser definido como sendo “*um motor de busca especializado em informações científicas e na identificação das citações que os documentos acadêmicos recebem*” (TORRES-SALINAS; RUIZ-PÉREZ; DELGADO-LÓPEZ-CÓZAR, 2009).

A primeira etapa pode ser feita através da biblioteca de funções “*HttpClient*” do C#, a qual captura os *hyperlinks* (HTML) da página visitada, deixando-os disponíveis para que a segunda etapa extraia os dados que serão necessários futuramente.

Já na segunda etapa, entra em cena uma biblioteca de expansão de funções, que não é nativa do C#, mas é disponibilizada para ele por seus desenvolvedores, chamada “*Html Agility Pack (HAP)*”. Com o HAP, torna-se possível manipular, de forma prática e ágil, o HTML da página que foi capturada na primeira etapa, armazenando-se apenas os dados necessários para o cálculo das métricas de produtividade dos pesquisadores.

Todos os dados coletados através da técnica de *web crawler* são armazenados dentro de um banco de dados, que deve ser atualizado periodicamente conforme a vontade do responsável por administrar o sistema, e ficam a disposição para serem consultados.

Durante o processo de consulta, destaca-se o maior diferencial deste trabalho: o programa aqui proposto, chamado Gerenciador de Métricas de Produtividade (GMP).

3 Gerenciador de Métricas de Produtividade (GMP)

3.1 Definição e diferencial

O Gerenciador de Métricas de Produtividade (GMP) é um software desenvolvido para classificar e, então, graficamente comparar, com base em duas métricas distintas - índice h e fator de impacto - grupos de pesquisadores. Pode-se comparar instituições, unidades acadêmicas, subunidades acadêmicas (ou departamentos) e até mesmo pesquisadores específicos dessas entidades, sendo que é possível realizar agrupamentos diversos, como Instituição A e Instituição B versus Instituição C, por exemplo.

Enquanto que o índice h é definido por Barreto et al (2013) como uma métrica simples de ser estimada para caracterizar o impacto da produção científica de um pesquisador, o fator de impacto, segundo Silva (2014), “*busca medir a disseminação da produção científica através da observação da frequência das referências feitas*”.

Portanto, em suma, o GMP é capaz de utilizar as métricas definidas acima para calcular tanto indicadores individuais, tal como os serviços mencionados na primeira seção, quanto indicadores coletivos de qualidade e quantidade de produção científica, apresentando assim uma nova forma de se avaliar os pesquisadores.

3.2 Níveis de acesso

O Gerenciador de Métricas de Produtividade contém três níveis de acesso diferentes para seus usuários, cada um com uma variedade diferente de recursos a disposição. Os níveis são:

- **Regular:** É o nível básico do GMP, que apenas pode realizar consultas de indicadores individuais ou coletivos de qualidade e quantidade de produção científica de pesquisadores cadastrados no banco de dados do software.
- **Gerente:** Tem acesso aos recursos do nível regular e a novos recursos como adicionar novos pesquisadores e editar pesquisadores já existentes.
- **Desenvolvedor:** Tem acesso aos recursos dos demais níveis, além de também poder atualizar os dados de todas as publicações científicas contidas no banco de dados do GMP ou criar e editar instituições, unidades acadêmicas e subunidades acadêmicas (ou departamentos).

As permissões de desenvolvedor e gerente são checadas ao se realizar a autenticação no programa. Já o usuário regular não necessita de autenticação, pois a ele é dado apenas o direito de realizar consultas.

3.3 Registro do software e documentação

Com o objetivo de estar sempre em contínuo aprimoramento, o Gerenciador de Métricas de Produtividade está licenciado sob a licença *Creative Commons*, ou seja, o GMP é um software gratuito para ser utilizado e inclusive implementado, desde que não se remova os créditos de desenvolvimento dos desenvolvedores da versão original. Contudo, a licença escolhida para o software não permite que este seja modificado para usos comerciais, sendo que todas as novas versões criadas devem seguir a mesma licença da original, tal como indicado na Figura 1.



Figura 1 - Licença sob a qual o software se encontra

Fonte: *Creative Commons*

Portanto, por ser um software que autoriza que sejam feitas modificações em seu código de programação, ele conta com uma documentação detalhada explicando o raciocínio lógico utilizado em cada função disponível, facilitando o entendimento de possíveis programadores interessados em expandir suas funções atuais.

Além de ser bem documentado para novos desenvolvedores, o GMP também conta com guias de uso repletos de imagens e diretrizes para ajudar usuários “comuns” do software a utilizá-lo da melhor maneira possível.

São três guias, equivalentes aos níveis de acesso existentes no software: desenvolvedor, gerente e regular. Dessa forma, o usuário mantém-se informado apenas das funções que lhe foram dadas a permissão de utilizar.

3.4 Funcionalidades do software

Para permitir que seus usuários realizem consultas com base nos pesquisadores cadastrados no banco de dados do software, é necessário também fornecer meios de se preencher e estruturar tal banco. Então, além da função de

consultas, o GMP oferece funções gerenciais para interagir melhor com o banco de dados, portando-se quase como um sistema gerenciador de banco de dados padrão, porém simplificado para atender apenas as verdadeiras necessidades do software.

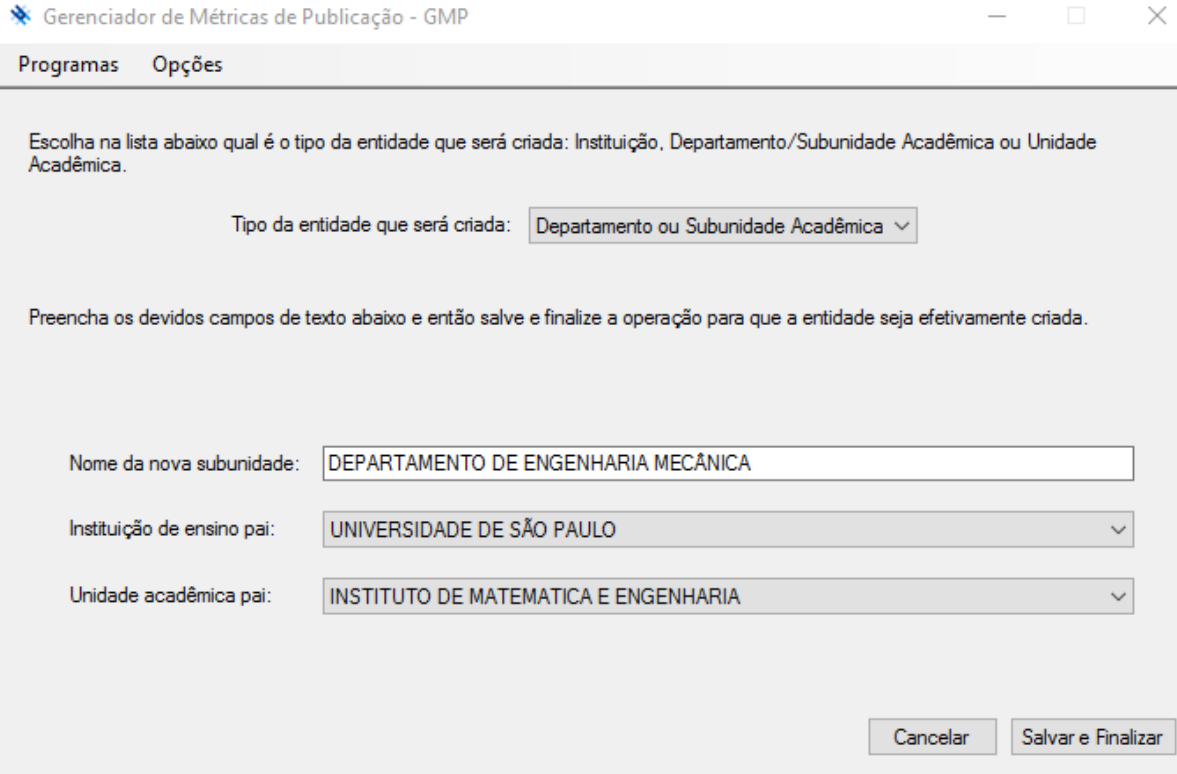
3.4.1 Atualizar dados do banco

Esta função conecta o GMP ao seu banco de dados, coleta os links do Google Acadêmico de cada pesquisador cadastrado no banco e realiza o processo de *web crawler* para atualizar todas as informações referentes aos trabalhos científicos desses pesquisadores.

Por ser uma ação desenvolvida para ser utilizada apenas de tempos em tempos, conforme for necessário, está disponível somente para usuários do nível de acesso “desenvolvedor”.

3.4.2 Criar entidades

Esta função, através da tela apresentada na Figura 2, que facilita o processo de interação com o banco de dados do software, permite que sejam criadas novas instituições, unidades acadêmicas e subunidades acadêmicas (ou departamentos).



The screenshot shows a web browser window titled "Gerenciador de Métricas de Publicação - GMP". The interface has a header with "Programas" and "Opções". Below the header, there is a text instruction: "Escolha na lista abaixo qual é o tipo da entidade que será criada: Instituição, Departamento/Subunidade Acadêmica ou Unidade Acadêmica." Below this is a dropdown menu labeled "Tipo da entidade que será criada:" with the selected option "Departamento ou Subunidade Acadêmica". A second instruction reads: "Preencha os devidos campos de texto abaixo e então salve e finalize a operação para que a entidade seja efetivamente criada." There are three input fields: "Nome da nova subunidade:" with the text "DEPARTAMENTO DE ENGENHARIA MECÂNICA"; "Instituição de ensino pai:" with a dropdown menu showing "UNIVERSIDADE DE SÃO PAULO"; and "Unidade acadêmica pai:" with a dropdown menu showing "INSTITUTO DE MATEMATICA E ENGENHARIA". At the bottom right, there are two buttons: "Cancelar" and "Salvar e Finalizar".

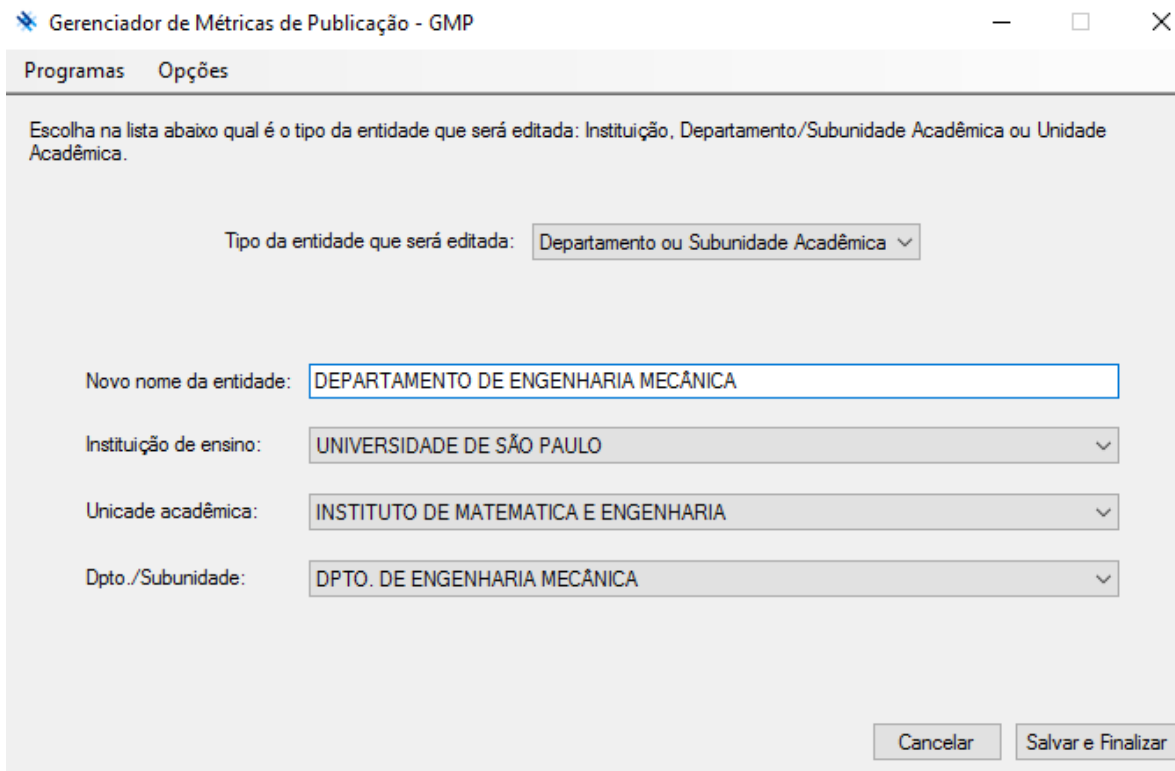
Figura 2 - Tela do GMP utilizada para criar novas entidades

Fonte: Imagem do software

Devido ao fato de trabalhar com estruturas importantes, esta função pode ser usada somente por usuários do nível de acesso “desenvolvedor”.

3.4.3 Editar entidades

Utilizando a tela apresentada na Figura 3, pode-se alterar o nome de entidades já cadastradas no banco de dados.



The screenshot shows a window titled "Gerenciador de Métricas de Publicação - GMP" with a standard Windows title bar (minimize, maximize, close). The window has a menu bar with "Programas" and "Opções". Below the menu bar, there is a text instruction: "Escolha na lista abaixo qual é o tipo da entidade que será editada: Instituição, Departamento/Subunidade Acadêmica ou Unidade Acadêmica." Below this instruction is a dropdown menu labeled "Tipo da entidade que será editada:" with the selected value "Departamento ou Subunidade Acadêmica". There are four input fields, each with a label and a value: "Novo nome da entidade:" with the value "DEPARTAMENTO DE ENGENHARIA MECÂNICA"; "Instituição de ensino:" with the value "UNIVERSIDADE DE SÃO PAULO"; "Unidade acadêmica:" with the value "INSTITUTO DE MATEMATICA E ENGENHARIA"; and "Dpto./Subunidade:" with the value "DPTO. DE ENGENHARIA MECÂNICA". At the bottom right of the window, there are two buttons: "Cancelar" and "Salvar e Finalizar".

Figura 3 - Tela do GMP utilizada para editar o nome de entidades já existentes
Fonte: Imagem do software

Assim como na função de criação de entidades, esta função pode ser usada somente por usuários do nível de acesso “desenvolvedor”, pois também trabalha com estruturas importantes.

3.4.4 Adicionar pesquisadores

Permite incluir novos pesquisadores no banco de dados do GMP, sendo que tal inclusão pode ser feita de duas formas: manual ou via arquivo de extensão “.csv”, através das telas apresentadas, respectivamente, nas Figuras 4 e 5. No caso da segunda opção, é necessário que sejam seguidas corretamente as instruções presentes nos guias para gerentes e desenvolvedores.

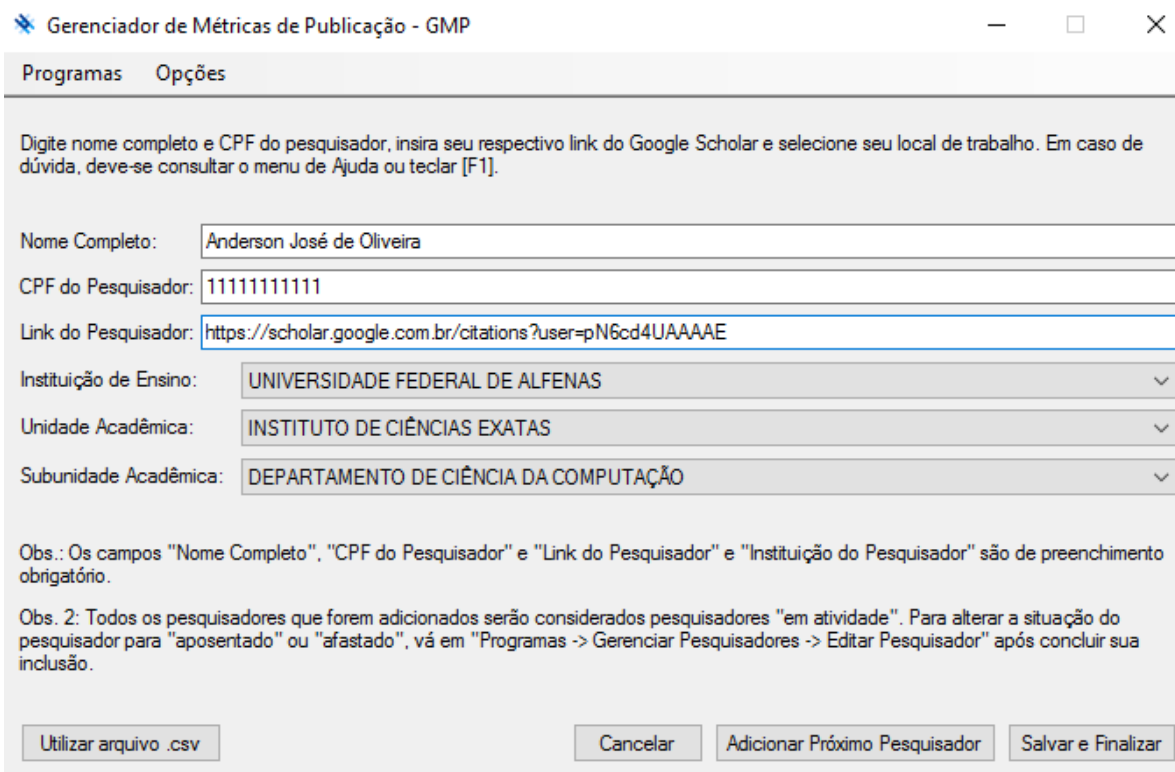


Figura 4 - Tela do GMP utilizada para adicionar manualmente novos pesquisadores no banco de dados

Fonte: Imagem do software

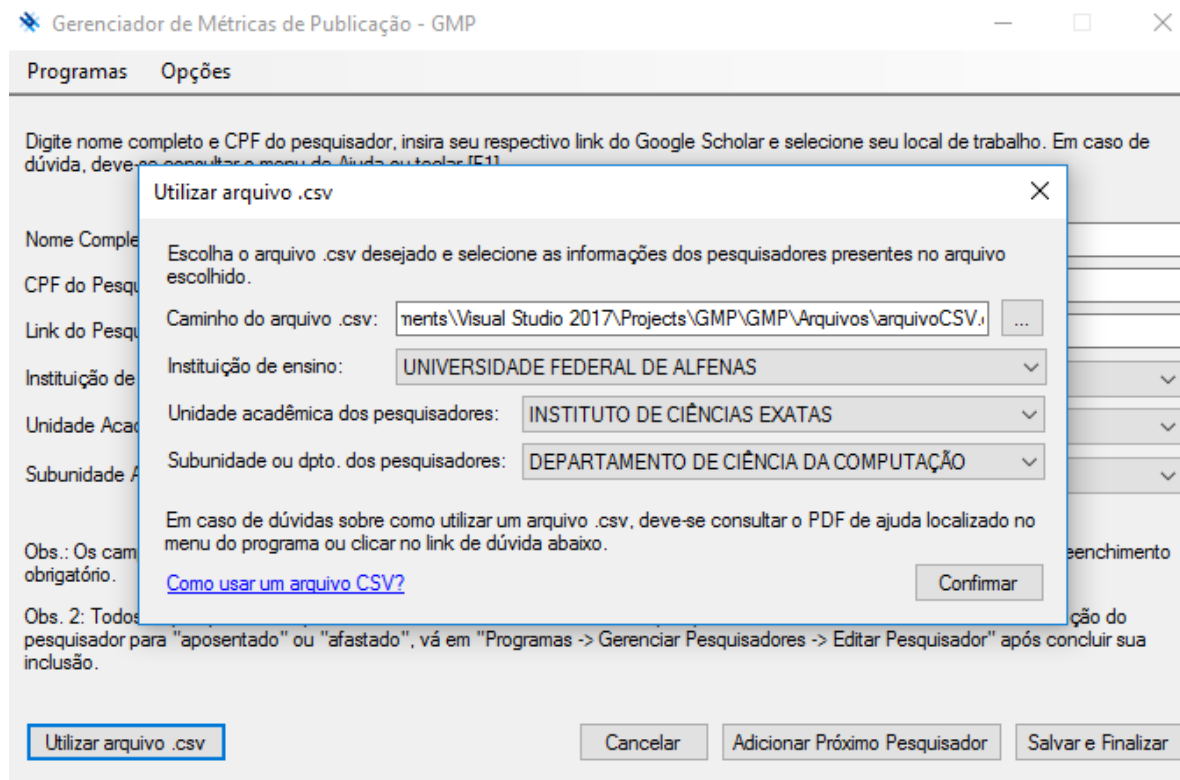


Figura 5 - Tela do GMP utilizada para adicionar, por meio de um arquivo de entrada de extensão ".csv", novos pesquisadores no banco de dados

Fonte: Imagem do software

Pode ser utilizada apenas por usuários pertencentes aos níveis de acesso “gerente” e “desenvolvedor”.

3.4.5 Editar pesquisadores

Esta função é capaz de editar diversos atributos dos pesquisadores cadastrados no banco de dados. Os atributos em questão, conforme pode ser visto na tela de edição apresentada na Figura 6, são: nome do pesquisador, link do perfil do pesquisador no Google Acadêmico, instituição, unidade acadêmica, subunidade acadêmica (ou departamento) e situação do pesquisador. O único atributo que não pode ser alterado após o cadastro é o CPF do pesquisador em questão.

CPF	Nome do Pesquisador	Instituição	Unidade Acadêmica
10987654321	MARIA BORGES COSTA TOKACELI	UNIVERSIDADE FEDERAL DE ALFENAS	INSTITUTO DE CIÊNCIAS EX

Nome do pesquisador: Seleção abaixo os atributos que serão editados e preencha os respectivos campos.

CPF do pesquisador que será modificado:

Nome completo:

Link:

Instituição:

Unidade:

Subunidade:

Situação:

Figura 6 - Tela do GMP utilizada para editar informações sobre os pesquisadores já cadastrados no banco de dados do software

Fonte: Imagem do software

Pode ser utilizada apenas por usuários pertencentes aos níveis de acesso “gerente” e “desenvolvedor”.

3.4.6 Realizar consulta

Na tela de realização de consultas, representada na Figura 7, o usuário pode agrupar instituições, unidades acadêmicas, subunidades acadêmicas ou pesquisadores em específico de inúmeras formas diferentes e, então, comparar seus respectivos indicadores de qualidade e quantidade de produção científica.

Como resultado da consulta, são gerados gráficos que facilitam visualmente a comparação entre os grupos definidos pelo usuário. Tais gráficos são gerados com base nas métricas de índice h e de fator de impacto, cabendo ao usuário escolher a que lhe for necessária.

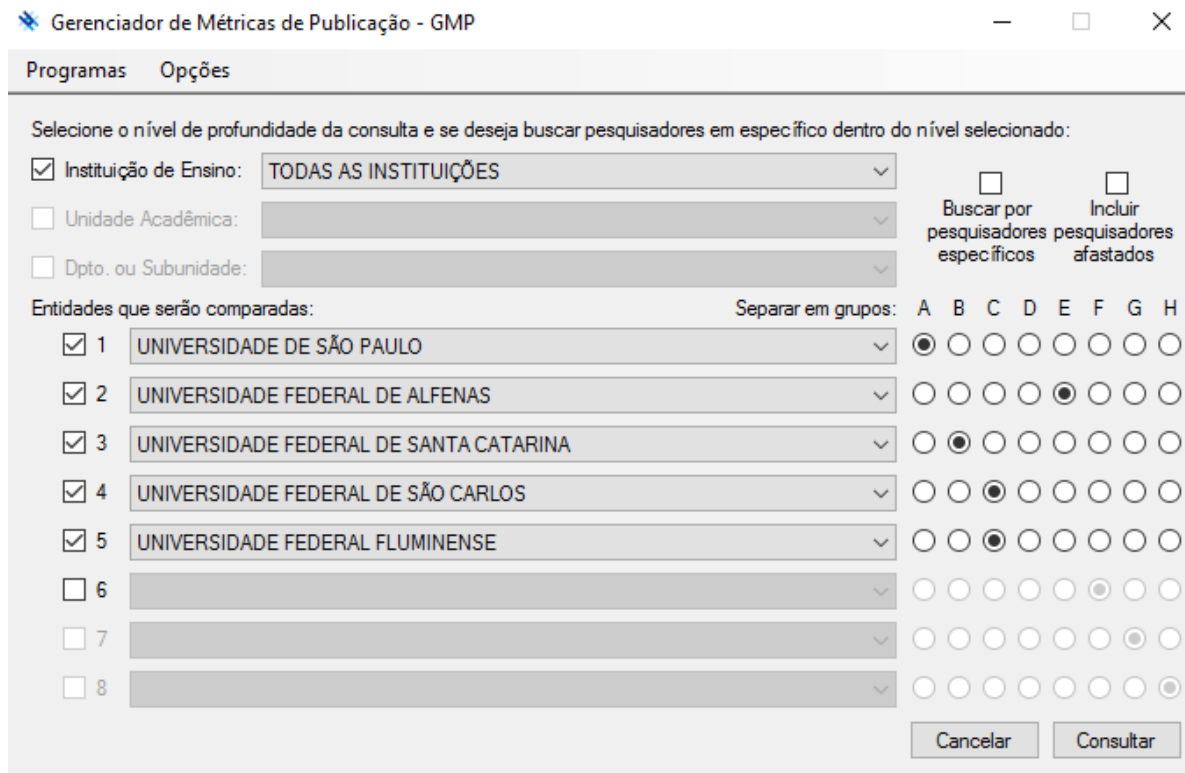


Figura 7 - Tela do GMP utilizada para realizar consultas individuais e agrupadas
Fonte: Imagem do software

O índice h é calculado por meio da seguinte lógica: tendo sido definidos os grupos a serem confrontados, ordenam-se, de forma decrescente, todos os trabalhos científicos publicados por seus integrantes de acordo com a quantidade de citações que cada um desses trabalhos recebeu, atribuindo-lhes, posteriormente, índices que começam a partir do 1, onde 1 é o índice do trabalho mais citado. Feito isso, o índice h é “definido como o número de artigos com número de citações $\geq h$ ” (HIRSCH, 2005).

A Figura 8, exibe o gráfico gerado por uma consulta que utiliza o índice h como indicador de produtividade de dois determinados grupos de pesquisa.

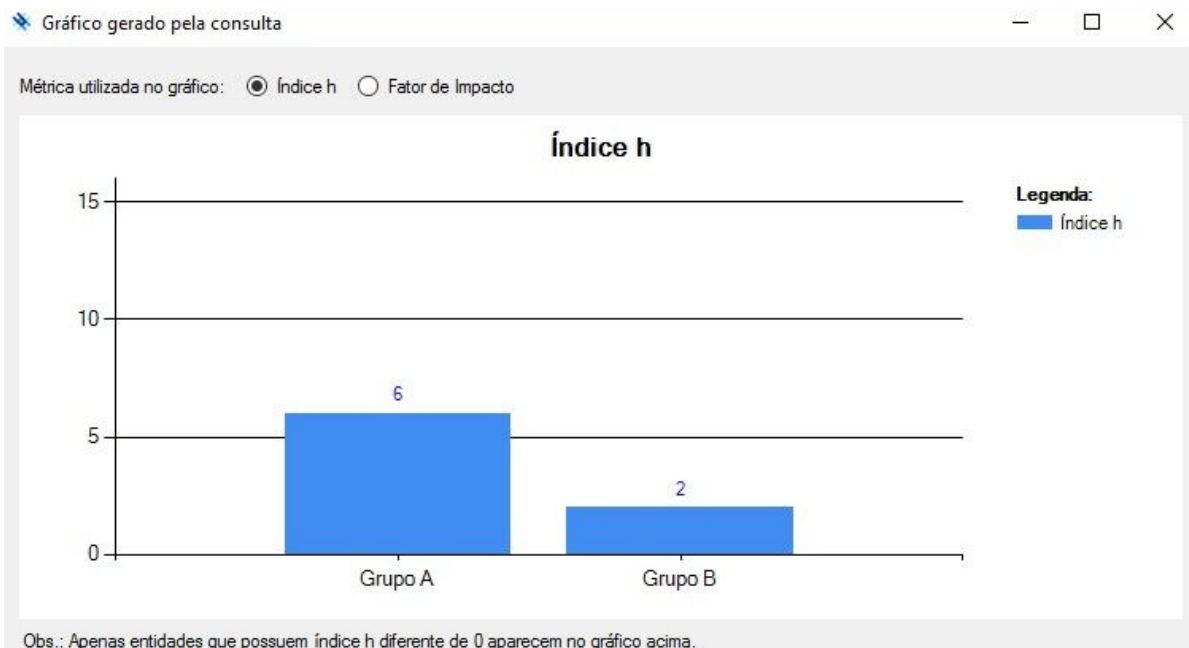


Figura 8 - Resultado de uma consulta realizada com o índice h como métrica para o cálculo dos indicadores de qualidade e quantidade de produção científica
 Fonte: Imagem do software

Diferentemente do índice h, o fator de impacto é uma métrica bem simples de ser utilizada: basta dividir o número total de citações recebidas nos trabalhos de cada grupo pelo número total de trabalhos publicados por cada um dos grupos em questão. Apenas os trabalhos publicados nos últimos X anos são levados em consideração, cabendo ao usuário do software definir o valor de X.

Entretanto, existem diversos sites e revistas que adotam um valor padrão para X, sendo 2 o valor mais conhecido, pois é o padrão do *Journal Citation Reports (JCR)*, a principal base de dados de citações de trabalhos científicos, quando este avalia o fator de impacto dos mesmos.

De acordo com Strehl (2005), o fator de impacto de um trabalho científico pode ser calculado como “a razão entre o número de citações feitas no corrente ano a itens publicados neste periódico nos últimos dois anos e o número de artigos publicados nos mesmos dois anos pelo mesmo periódico” (STREHL, 2005).

Na Figura 9, é exibido o gráfico resultante de uma consulta que utiliza o fator de impacto dos últimos 10 anos de dois grupos distintos.



Figura 9 - Resultado de uma consulta realizada com o fator de impacto dos últimos 10 anos como métrica para o cálculo dos indicadores de qualidade e quantidade de produção científica

Fonte: Imagem do software

3.5 Possíveis aplicações do software

O GMP pode ser usado para diversos fins, como, por exemplo, analisar o desempenho das produções científicas de um determinado grupo, administrar a distribuição de recursos financeiros dentro de uma instituição de ensino, incentivar uma competitividade sadia entre pesquisadores e assim por diante.

3.6 Melhorias futuras

Apesar de contar com uma boa gama de recursos, o GMP ainda tem muito potencial de desenvolvimento, podendo receber, em versões futuras, inúmeros novos recursos.

Dentre os possíveis recursos, podem ser citados o tratamento de autocitações para “penalizar” trabalhos científicos que referenciam outros trabalhos do mesmo pesquisador e a geração de relatórios, sejam eles impressos ou não, de operações ou consultas realizadas pelo usuário do software.

Outro aspecto que também seria interessante é a adição de novas fontes de dados para coleta de informações sobre os trabalhos científicos dos pesquisadores e a utilização de novas métricas, sejam elas existentes ou inéditas.

4 Considerações finais

Nesta primeira versão, o GMP já é capaz de cumprir os objetivos aqui propostos, isto é, torna possível a busca automática de informações dos trabalhos científicos dos pesquisadores cadastrados em seu banco de dados para que, então,

sejam feitas análises coletivas de qualidade e quantidade de produção científica com base nesses trabalhos.

Conseqüentemente, surge a oportunidade de se incentivar a realização de pesquisas e a escrita de produções científicas de qualidade, fazendo com que as unidades acadêmicas e suas subdivisões melhorem a qualidade de seus trabalhos e recebam maior visibilidade interna e externamente da Instituição a qual pertencem, ao mesmo tempo em que a própria Instituição se fortalece como um todo.

Referências

BARRETO, Maurício Lima et al. Diferenças entre as medidas do índice-h geradas em distintas fontes bibliográficas e engenho de busca. **Rev Saúde Pública**, São Paulo, v. 47, n. 2, p.231-238, abr. 2013. Disponível em: <<http://www.scielo.org/pdf/rsp/v47n2/0034-8910-rsp-47-02-0231.pdf>>. Acesso em: 03 ago. 2017.

DEGEN, P. J. ; MELLO, A. A. A. O empreendedor: fundamentos da iniciativa empresarial. São Paulo: McGraw-Hill, 1989

HICKS, Diana et al. El manifiesto de Leiden sobre indicadores de investigación. 2015. Disponível em: <http://www.ingenio.upv.es/sites/default/files/adjunto-pagina-basica/manifiesto_es.pdf>. Acesso em: 27 jul. 2017.

HIRSCH, J. E.. An index to quantify an individual's scientific research output. *Proceedings Of The National Academy Of Sciences*, [s.l.], v. 102, n. 46, p.16569-16572, 7 nov. 2005.

LOYOLA, Rafael. Países en desarrollo deberían evitar la 'ciencia lenta'. 2013. Disponível em: <<http://www.scidev.net/america-latina/i-d/opinion/pa-ses-en-desarrollo-deber-an-emitir-la-ciencia-lenta-.html>>. Acesso em: 28 jul. 2017.

OLSTON, Christopher; NAJORK, Marc. Web Crawling. **Foundations And Trends® In Information Retrieval**, [s.l.], v. 4, n. 3, p.175-246, 2010. Now Publishers. <http://dx.doi.org/10.1561/15000000017>.

SANTOS, Igor. **Aprendendo a programar em C# .Net: Por que escolher C# para desenvolvimento de uma aplicação?**. 2011. Disponível em: <<https://desenvolvedordotnet.wordpress.com/2011/09/13/por-que-escolher-c-para-deseenvolvimento-de-uma-aplicacao/>>. Acesso em: 20 jul. 2017.

SILVA, Bianca Rodrigues. **FATOR DE IMPACTO: ESTUDO DOS PERIÓDICOS CONTÁBEIS BRASILEIROS**. 2014. 61 f. TCC (Graduação) - Curso de Administração, Ciências Contábeis e Ciências Econômicas, Universidade Federal de Goiás, Goiânia, 2014. Disponível em: <<https://repositorio.bc.ufg.br/bitstream/ri/4176/6/TCCG - Ciências Contábeis - Bianca Rodrigues Silva.pdf>>. Acesso em: 04 ago. 2017.

SOARES, Fábio de Azevedo. **MINERAÇÃO DE TEXTOS NA COLETA INTELIGENTE DE DADOS NA WEB**. 2008. 120 f. Tese (Doutorado) - Curso de Engenharia Elétrica, Departamento de Engenharia Elétrica do Centro Técnico Científico da Puc-rio, Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2008. Disponível em: <<https://www.maxwell.vrac.puc-rio.br/eletricaonline/serieConsulta.php?strSecao=especifico&nrSeq=13212@1&ref=Teses%20e%20Disserta%C3%A7%C3%B5es>>. Acesso em: 02 ago. 2017.

STREHL, Letícia. O fator de impacto do ISI e a avaliação da produção científica: aspectos conceituais e metodológicos. **Ci. Inf.**, Brasília, v. 34, n. 1, p.19-27, abr. 2005. Disponível em: <<http://www.scielo.br/pdf/ci/v34n1/a03v34n1.pdf>>. Acesso em: 03 ago. 2017.

TORRES-SALINAS, Daniel; RUIZ-PÉREZ, Rafael; DELGADO-LÓPEZ-CÓZAR, Emilio. Google Scholar como herramienta para la evaluación científica. **El Profesional de La Información**, Barcelona, v. 18, n. 5, p.501-510, set. 2009. Disponível em: <<http://ec3.ugr.es/publicaciones/d700h04j123154rr.pdf>>. Acesso em: 02 jul. 2017.