



VISIBILIDADE DA PRODUÇÃO CIENTÍFICA DE GRUPOS DE PESQUISA:

um estudo de caso a partir do Grupo de Pesquisa de Estudos
Métricos em Informação da UNESP

Francielle Franco dos Santos

 <https://orcid.org/0000-0002-3686-595X>.

 franfranco.santos@gmail.com.

 Universidade Estadual Paulista (UNESP) |

 <https://ror.org/00987cb86> | Marília, Brasil.

Rafael Gutierrez Castanha

 <https://orcid.org/0000-0002-3117-1780>.

 r.castanha@gmail.com.

 UUniversidade de Marília (UNIMAR) |

 <https://ror.org/01w96sp43> | Marília, Brasil.

Eixo temático: Ciência, Tecnologia e Inovação

Modalidade: Resumo expandido

DOI: 10.22477/ix.ebbc.258

Resumo: Apresenta o processo de implantação de um modelo de estruturação, organização e visualização de dados de produção científica do grupo de pesquisa Estudos Métricos em Informação da UNESP, por meio do estudo de caso do *website* do grupo. Foram coletados os dados da produção científica dos 31 membros do grupo na plataforma Lattes, apresentando-os em formato de gráficos e tabelas interativas, e também gerando a rede de colaboração entre os membros do grupo e seus coautores. Conclui que o modelo auxilia na socialização do conhecimento científico gerado pelo grupo e serve como base de dados científicos do grupo.

Palavras-Chave: Grupos de pesquisa. Produção científica. Visibilidade. *Websites*.



1 INTRODUÇÃO

Os grupos de pesquisa, no contexto da ciência brasileira, destacam-se como centros para a geração de conhecimento, a promoção de colaboração entre seus membros e formação de recursos humanos. Desse modo, a consolidação desses grupos reflete tanto a capacidade investigativa das diferentes áreas do conhecimento que se inserem, quanto o reconhecimento institucional (Vanz; Caregnato, 2023). Assim, “grupos de pesquisa têm sido considerados espaços coletivos de aprendizado e pesquisa, liderados por agentes com papel fundamental na sua consolidação e permanência” (Bufrem *et al.* 2023, p. 212).

Dada a relevância conferida aos grupos de pesquisa, em virtude do papel crucial desempenhado na produção de pesquisa, compreende-se que a ausência de informação sobre o tema pode resultar em desafios e, conseqüentemente, pode comprometer a promoção e o financiamento da pesquisa. Portanto, a falta de dados acerca da produção em uma área pode acarretar prejuízos na formulação de programas e na avaliação de atividades relacionadas ao avanço científico e tecnológico no país (Moreira; Mueller; Vilan Filho, 2020).

Desse modo, avaliar e monitorar a produção científica de grupos de pesquisa pode favorecer a compreensão do desenvolvimento científico de áreas diversas em muitos aspectos como as principais fontes de publicação, internacionalização do grupo, desenvolvimento tecnológico e de inovação (produção de patentes e/ou programas de computador), principais temáticas de pesquisa, colaboração científica entre o membros do grupo ou entre membros externos, entre outros. Hilário e Freitas (2020) salientam que dados provenientes de grupos de pesquisa formalmente constituídos e registrados representam fontes valiosas para a caracterização de relações de colaboração científica, no entanto, é importante observar que a concretização de elementos analíticos derivados desses registros requer esforços para elaboração de pesquisas.

No âmbito dos Estudos Métricos da Informação (EMI), em busca realizada no diretório de grupos de pesquisa do CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico), foram encontrados 84 grupos distintos possivelmente relacionados aos EMI. Para a busca foram utilizadas duas estratégias, a primeira inserindo o termo *bibliometria cientometria informetria webmetria patentometria altmetria*, retornando 61 resultados e a segunda com o termo “estudos métricos”, resultando em 36 registros. Ao combinar ambas buscas, obteve-se 84 grupos distintos. Dentre estes, esta pesquisa dá foco ao grupo de pesquisa “Estudos Métricos em Informação” (GPEMI) vinculado ao Programa de Pós-Graduação em Ciência da Informação (PPGCI) da Universidade Estadual Paulista (UNESP), campus de Marília/SP liderado pelas pesquisadoras Ely Francina Tannuri de Oliveira e Maria Cláudia Cabrini Grácio e registrado no diretório do CNPq desde 2008.

Os resultados de pesquisa do grupo têm sido publicados em importantes periódicos e nos principais eventos científicos, nacionais e internacionais, da área da Ciência da Informação, especialmente da Bibliometria e da Cientometria. Somados a estes resultados destaca-se ainda a interação com pesquisadores internacionais que, por meio de suas reuniões científicas, vem permitindo o adensamento da rede de colaboração científica do GPEMI (Grácio *et al.*, 2020). Dessa maneira, o grupo, por meio de seus membros, ativos e egressos, apresenta produção científica relevante à área. Além disso, a visualização



desta produção, no que diz respeito às fontes de socialização do conhecimento, bem como a interação entre pesquisadores, poderá gerar insumos para compreender o real impacto e visibilidade do grupo perante a área.

Diante do exposto, esta pesquisa pretende responder ao seguinte problema de pesquisa: de que maneira a visualização de dados de produção científica do GPEMI pode auxiliar na socialização do conhecimento científico gerado pelo grupo de pesquisa? Tendo como objetivo, descrever a implementação de um modelo de estruturação, organização e visualização de dados de produção científica de grupos de pesquisa para possível ampliação da visibilidade de grupos de pesquisa. De forma específica, realizar um estudo de caso a partir do website que apresenta as produções científicas do GPEMI da UNESP.

2 METODOLOGIA

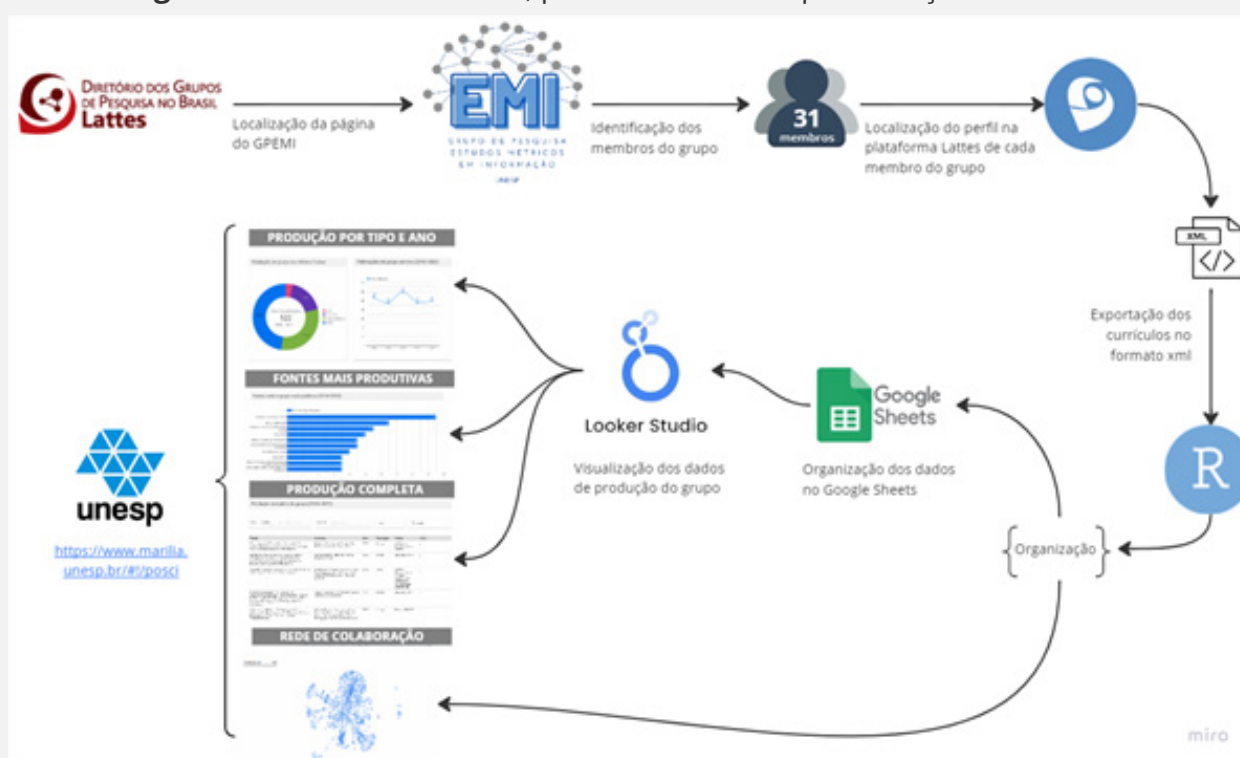
Para a proposição de um modelo de estruturação, organização e visualização de dados de produção científica, realizou-se um estudo de caso a partir do GPEMI, desenvolvido em quatro etapas presentes na Figura 1: i) coleta e extração de dados das produções científicas via plataforma de currículo Lattes; ii) organização de dados; iii) visualização de dados; iv) integração ao domínio *unesp.br*. O processo de coleta dos dados da produção científica do GPEMI foi iniciado a partir da identificação de todos os membros do grupo utilizando o sistema de busca do diretório de grupos de pesquisa do CNPq. Foram identificados 31 membros ativos, a partir da lista de membros localizou-se os 31 perfis na plataforma Lattes, por meio do campo de busca, para que os dados pudessem ser exportados no formato xml. Tanto a localização como a exportação foram realizadas manualmente, no mês de julho de 2023 e foram consideradas somente produções entre 2018 e julho de 2023.

A partir dos dados contidos nos arquivos xml foi implementado, em linguagem de programação R, um código capaz de extrair informações referentes às produções bibliográficas de cada autor, isto é, produções de artigos publicados em periódicos, trabalhos publicados em anais de evento, livros e capítulos de livro. Ainda, para cada tipologia documental, foram extraídos os seguintes metadados: título de cada produção, autores, ano de publicação, fonte de publicação e endereço eletrônico de cada publicação (ou DOI - *Digital Object Identifier*). O código é capaz de processar múltiplos arquivos xml de uma só vez e fornece, de maneira automática ao usuário, um arquivo em formato txt tabulado (organizado em colunas) com estas informações. Ademais, pelo fato de o código fornecer os dados de autoria, foi possível gerar, a rede de coautoria (em formato *html*) entre todos os autores identificados, ou seja, tanto os 31 membros dos grupos, quanto seus respectivos coautores, mapeando globalmente a relação entre os pesquisadores envolvidos em todas as publicações extraídas. É importante salientar que este código funciona para qualquer conjunto de currículos em formato xml, não sendo de uso exclusivo do GPEMI e executa, em um mesmo *script* a extração das produções e em seguida, a geração da rede de coautorias. Nesse sentido, o primeiro processo deve ser executado anteriormente ao segundo, visto que a geração da rede de coautorias irá processar o arquivo txt gerado no primeiro processo. O código está presente em Castanha e Franco (2024).

Após o processo de extração, organizou-se os dados de produção extraídos em uma planilha do

Google Sheets para conexão com o *Google Looker Studio*, ferramenta online e gratuita desenvolvida pelo *Google* que permite a criação de relatórios e painéis interativos personalizáveis. As informações são obtidas a partir de bancos de dados ou planilhas (*Google Looker Studio*, 2024). A partir da planilha do *Google Sheets*, gerou-se quatro visualizações desenvolvidos via *Google Looker Studio* incorporando-as na página do grupo, site do PPGCI Unesp/Marília, mais especificamente, dentro da página de produção: 1) Gráfico de setores contendo as frequências (em valores absolutos e relativos) das produções segundo tipologia documental: artigos publicados em periódicos, trabalhos publicados em anais de evento, livros e capítulos de livro; 2) Gráfico de linha contendo a evolução temporal da produção do grupo; 3) Gráfico de barras contendo as 10 fontes em que o grupo mais publicou; 4) tabela interativa contendo a produção completa do grupo. A Figura 1 sintetiza todo processo.

Figura 1 - Processo de coleta, processamento e apresentação dos dados.



Fonte: Elaboração Própria.

A incorporação se deu por meio de uma tag *iframe* inserida no código *html* da página. O site da UNESP é desenvolvido com base no *Foundation* que é uma estrutura de *front-end* responsiva que dispõe de componentes para desenvolvimento de *websites*. O *framework* oferece componentes para apresentação do conteúdo, contudo nenhum componente disponível atendeu a demanda e por isso o código foi inserido diretamente no *html* da página. Como última inserção tem-se a rede de coautorias, que foi gerada por meio código apresentado anteriormente e foi armazenado em servidor externo (fornecido pelo Rstudio) e incorporado ao *website* do grupo. Contudo, é importante mencionar que o código fornece a opção de salvar a rede em formato *html*. A utilização de servidor externo se deu pela inviabilidade de acesso ao servidor interno da UNESP, utilizou-se o recurso de inserção via *iframe* que possibilita a inclusão de um código *html* dentro de outro. Todos os dados mencionados anteriormente foram inseridos no *website* em agosto de 2023.

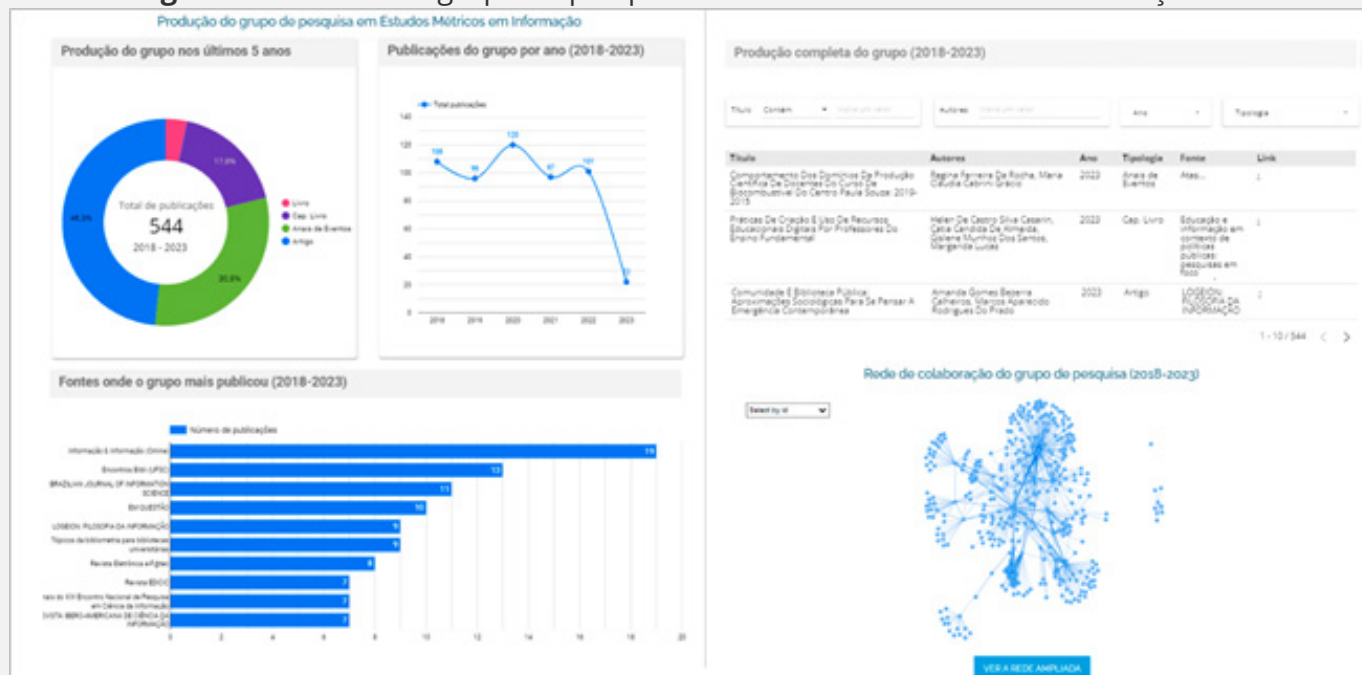
3 RESULTADOS

A primeira parte do website é constituída pelos gráficos de “Produção do grupo nos últimos 5 anos” e “Publicações do grupo por ano (2018-2023)”. No gráfico de setores, localizado na parte esquerda da página, é possível visualizar a quantidade de publicações segundo a tipologia documental (artigos publicados em periódicos, trabalhos publicados em anais de evento, livros e capítulos de livro) e no centro do gráfico tem-se o valor total de publicações. Destaca-se que todos os elementos são interativos, assim, ao sobrepor o mouse em uma das partes do gráfico é possível visualizar os valores absolutos e relativos, e ao clicar aplica-se um filtro que modifica todos os outros elementos da página de acordo com a seleção. A parte direita da mesma seção apresenta um gráfico de linha contendo a evolução temporal da produção. Assim, pode-se ter a visão de quais são os tipos mais recorrentes de produção desenvolvida pelo grupo ao longo dos anos e, conseqüentemente, em quais períodos o grupo esteve mais ativo em termos de produção.

Na rolagem seguinte é apresentado um gráfico de barras contendo as 10 fontes em que o grupo mais publicou. Esta visualização permite identificar quais são os periódicos e eventos onde o grupo teve presença consistente, e, por consequência, em quais espaços e/ou veículos de comunicação possui maior visibilidade. Quando o usuário clica em uma das barras aplica-se o filtro e todos os elementos são modificados, desse modo, navegando pelas outras seções, é possível visualizar quais os tipos documentais que mais se publicou na fonte selecionada (gráfico de setores), quais foram os anos em que o grupo esteve mais presente naquela fonte (gráfico de linha) e quais foram os trabalhos publicados na fonte (tabela interativa, localizada abaixo do gráfico de fontes mais produtivas).

Na tabela é possível visualizar a produção completa do grupo (no período de 2018-2023), por meio das colunas de título, autores, ano, tipologia, fonte e link. Acima da tabela existem opções de filtragem permitindo que o usuário interaja com o conteúdo realizando uma busca por títulos ou nome dos autores, além de poder filtrar a visualização por ano ou tipologia documental. Localizado entre os filtros e a tabela, existe um menu (denominado “Mais”) e representando pelo símbolo  onde é possível exportar os dados da tabela. Portanto, além de um recurso visual, esta solução pode ainda servir como uma base dados ao passo que permite a exportação de dados de produção atualizados de um grupo de pesquisa específico. Todos os elementos descritos podem ser observados na Figura 2.

Figura 2 - Website do grupo de pesquisa em Estudos Métricos em Informação/UNESP.



Fonte: Elaboração dos autores. Disponível em [EMI UNESP](https://emi.unesp.br).

Além dos recursos viabilizados pelo *Google Looker Studio*, destaca-se a rede de coautorias desenvolvida no Rstudio. A rede foi gerada pela segunda parte do código descrito anteriormente e conta as publicações relativas a artigos, trabalhos publicados em anais de evento, livros e capítulos de livro registrado nos 31 currículos Lattes evidenciando todas as parcerias entre os membros do grupo e seus respectivos coautores. Foram identificadas 544 produções: 263 artigos, 166 trabalhos publicados em anais de eventos, 97 capítulos de livros e 18 livros. Este montante foi produzido por 408 autores distintos.

4 CONSIDERAÇÕES FINAIS

Esta pesquisa apresentou o processo completo de implantação de um modelo de estruturação, organização e visualização de dados de produção científica do grupo de pesquisa “Estudos Métricos em Informação” (GPEMI) da UNESP, realizando um estudo de caso a partir do *website* que apresenta as produções científicas do grupo. Para isso, os dados dos 31 membros do grupo foram coletados e extraídos na plataforma de currículo Lattes para que pudessem ser organizados, visualizados e integrados ao domínio *unesp.br*.

Nesse sentido, ao responder o problema de pesquisa, o modelo proposto auxilia na socialização do conhecimento científico gerado pelo grupo de pesquisa na medida em que apresenta de maneira estruturada quais são os tipos mais recorrentes de produção desenvolvida pelo grupo ao longo dos anos; em quais períodos o grupo esteve mais ativo em termos de produção; quais as fontes onde o grupo tem maior presença e visibilidade; a produção e a rede de colaboração entre os membros do grupo de membros externos.

Além disso, o *website* do grupo pode ser considerado uma base de dados pois permite a exportação dos dados de produção atualizados de um grupo de pesquisa específico. As limitações para o de-



envolvimento do modelo foram identificadas nas etapas de coleta e processamento dos dados em função das atualizações dos currículos na Plataformas Lattes que, por vezes, são inconsistentes. Por fim, as perspectivas futuras incluem a atualização dos dados, bem como a inclusão de outros recursos como a visualização geográfica das produções do grupo e indicadores como a *score altmétrico* das produções.

REFERÊNCIAS

CASTANHA, Rafael Gutierrez; FRANCO; Francielle. **Código para extração de produções do currículo Lattes e gerador de rede de coautorias**. Zenodo [site], publicado em 7 fev. 2024. versão 1. DOI: <https://doi.org/10.5281/zenodo.10632648>. Disponível em: <https://zenodo.org/records/10632648>. Acesso em: 09 fev. 2024.

BUFREM, Leilah Santiago *et al.* Educação, pesquisa e produção científica: memória e trajetória de atuação de um grupo de pesquisa na Ciência da Informação no Brasil. Folha de Rosto: **Revista de Biblioteconomia e Ciência da Informação**, v. 9, n. 1, 2023, p. 211-229. DOI: <https://doi.org/10.56837/fr.2023.v9.n1.950>. Disponível em: <https://lume.ufrgs.br/bitstream/handle/10183/257638/001167106.pdf>. Acesso em: 09 fev. 2024.

GOOGLE LOOKER STUDIO. **Você está no Looker Studio** [site]. 2024. Disponível em: https://support.google.com/looker-studio/answer/6283323?hl=pt-BR&ref_topic=6267740. Acesso em: 04 fev. 2024.

GRÁCIO, Maria Cláudia Cabrini; OLIVEIRA, Ely Francina Tannuri de; MARTÍNEZ-ÁVILA, Daniel; ROSAS, Fábio Sampaio. Apresentação. In: GRÁCIO, Maria Cláudia Cabrini (Org.); OLIVEIRA, Ely Francina Tannuri de (Org.); MARTÍNEZ-ÁVILA, Daniel (Org.); ROSAS, Fábio Sampaio (Org.). **Tópicos da bibliometria para bibliotecas universitárias**. São Paulo: Oficina Universitária; Cultura Acadêmica, 2020. p 7-12. DOI: <https://doi.org/10.36311/2020.978-65-86546-91-0.p8-13>. Disponível em: https://ebooks.marilia.unesp.br/index.php/lab_editorial/catalog/book/187. Acesso em: 09 fev. 2024.

HILÁRIO, Carla Mara; FREITAS, Juliana Lazzarotto. Indicadores de colaboração científica: aspectos éticos, práticos e formas de mensuração. In: GRÁCIO, Maria Cláudia Cabrini (Org.); OLIVEIRA, Ely Francina Tannuri de (Org.); MARTÍNEZ-ÁVILA, Daniel (Org.); ROSAS, Fábio Sampaio (Org.). **Tópicos da bibliometria para bibliotecas universitárias**. São Paulo: Oficina Universitária; Cultura Acadêmica, 2020. p 71-92. DOI: <https://doi.org/10.36311/2020.978-65-86546-91-0.p8-13>. Disponível em: https://ebooks.marilia.unesp.br/index.php/lab_editorial/catalog/book/187. Acesso em: 9 fev. 2024.

MOREIRA, Jonathan Rosa; MUELLER, Suzana Pinheiro Machado; VILAN FILHO, Jayme Leiro. Produção científica dos membros dos grupos de pesquisa das áreas de informação no Brasil. **Informação & Informação**, v. 25, n. 1, 2020. DOI: <https://doi.org/10.5433/1981-8920.2020v25n1p1>. Disponível em: <https://ojs.uel.br/revistas/uel/index.php/informacao/article/view/35885>. Acesso em: 9 fev. 2024.

VANZ, Samile Andrea de Souza; CAREGNATO, Sônia Elisa. O grupo de pesquisa comunicação científica da universidade federal do rio grande do sul. Folha de Rosto: **Revista de Biblioteconomia e Ciência da Informação**, v. 9, n. 1, 2023. DOI: <https://doi.org/10.56837/fr.2023.v9.n1.950>. Disponível em: <https://periodicos.ufca.edu.br/ojs/index.php/folhaderosto/article/view/950>. Acesso em: 9 fev. 2024.