

O QUE A POLÍTICA CIENTÍFICA PODE APRENDER COM O DIÁRIO OFICIAL DA UNIÃO (DOU): UM OLHAR EXPLORATÓRIO¹

Paulo Eduardo Potyguara Coutinho Marques
pepcmarques@gmail.com

Maria Cristina Soares Guimarães
cristina.guimaraes@icict.fiocruz.br

Renato Rocha Souza
renato.souza@fgv.br

Janio Gustavo Barbosa
janioguga@gmail.com

Resumo

A Ciência da Política Científica emerge, recentemente, em resposta à preocupação crescente com a ausência de capacidade analítica das políticas que orientam o desenvolvimento da ciência. Uma das questões centrais é a tentativa de responder se os investimentos públicos no setor de pesquisa alcançaram os resultados esperados. A simplicidade dessa pergunta descortina um cenário complexo de interações, e a crescente produção de dados por parte dos financiadores abrem a oportunidade de análises de bases quantitativas para contribuir nesse esforço de entendimento. O texto lança um olhar exploratório e metodológico sobre o tema no contexto brasileiro, e toma o Diário Oficial da União (DOU) como fonte de dados potencial para análise do financiamento público de pesquisa em saúde no Brasil, especificamente em dengue. A metodologia desenvolvida, embora com limitações, desvela o potencial dessa fonte de dados para análises quantitativas sobre o esforço de pesquisa com financiamento público no Brasil.

Palavras-chave: Financiamento de pesquisa. Política de Ciência e Tecnologia. Diário Oficial da União (DOU). Dengue. Big Data. Análise Visual. Ciência da Política Científica.

WHAT SCIENTIFIC POLICY CAN LEARN FROM THE UNION'S OFFICIAL JOURNAL (DOU): AN EXPLORATORY LOOK

Abstract

Science of Science Policy has recently emerged in response to growing concern about the lack of analytical capacity of the policies that guide the development of science. One of the central questions is the attempt to answer whether public investments in the research sector have achieved the expected results. The simplicity of this question uncovers a complex scenario of interactions, and the growing production of data by funders opens the opportunity for quantitative analysis to contribute to this understanding effort. The text takes an exploratory and methodological look at the theme in the Brazilian context, and takes the Official Gazette (DOU) as a potential source of data for analysis of public funding for health research in Brazil, specifically in dengue. The methodology developed, albeit with limitations, reveals the potential of the DOU for quantitative analysis of the publicly funded research effort in Brazil.

Keywords: Funding for research. Science and Technologic Policy. Brazilian Official Gazette (DOU). Dengue. Big Data. Visual Analysis. Science of Scientific Policy.

¹ Essa pesquisa é resultante da tese de doutoramento de Marques (2017).



1 INTRODUÇÃO

A pesquisa no campo das políticas em ciência, tecnologia e inovação pode ser definida, de forma abrangente, como a aplicação das ciências sociais ao estudo das políticas para ciência, tecnologia e inovação. Não se configura, assim, como um campo disciplinar orientado por uma teoria ou um paradigma; antes, é orientado por problemas, e intrinsecamente interdisciplinar (MORLACCHI; MARTIN, 2009). O que nasceu há quase sete décadas atrás como Política Científica, ou, aquilo que Salomon (1973) circunscreveu como um conjunto de ações e decisões políticas em questões que afetam o empreendimento científico, e que trazia embutido uma substancial linearidade entre investimento em ciência e produção de bem-estar social, por meio da tecnologia e do processo de inovação industrial, chegou ao século XXI com novos matizes, maiores e complexos desafios. Ciência, tecnologia e inovação são agora como um todo interconectado, atravessado e moldado por diferentes políticas, e que desenham dinâmicas peculiares em tempos e geografias próprias.

Não sem surpresa, pesquisadores e *policy-makers* também reconhecem que tais políticas se encontram, na atualidade, em uma encruzilhada, traduzindo certo desencantamento com todo o esforço empreendido até aqui. O sonho de Vannevar Bush (1945) em *Science, The Endless Frontier*, que apostava que a ciência guiaria um futuro melhor para a humanidade, há muito se perdeu. Gibbons (2001) aponta para uma política de ciência e tecnologia (PCT) que está em estado de desânimo, de depressão, carecendo de ar fresco. Soete (2019) fala em crise da política, e sobre uma máquina quebrada; Metcalfe (2007) sugere que a mesma está focando no lugar errado, (AGHION et al., 2009) discutem sobre uma simplicidade dessas políticas, quase de tirar o ar, elegantemente (*breathtakingly elegant*).

Dentre todas as discussões que permeiam a literatura, há uma concordância na necessidade do aprendizado. É fundamental tomar cada “instrumento político” (*policy instruments*), por exemplo, o financiamento à pesquisa, como uma intervenção em um fluxo contínuo de eventos outros, que se co produzem ao longo do tempo. Todos os atores, especialmente os *policy-makers*, devem aprender com experiências passadas na tentativa de avaliar as relações de causa-efeito dos instrumentos políticos, em busca de elaborar e implementar políticas aprimoradas, ou minimamente mais reflexivas, em face das incertezas, interdependências e contingências que se colocam (FLANNAGAN; UYARRA, 2016). Iniciativas tanto de bases qualitativas como quantitativas são chamadas a contribuir nesse esforço.

Uma das iniciativas em curso é o programa *Science of Science and Innovation Policy (SciSIP)*, que nasceu da maior financiadora de pesquisa norte-americana, a *National Science Foundation (NSF)*, quando se questionou sobre a ausência de uma infraestrutura de dados que permitisse um olhar mais sistêmico, em perspectiva de país, sobre o alcance e as consequências de suas ações de fomento, procurando vinculá-los não só aos produtos (*outputs*) das pesquisas apoiadas, mas também às suas potenciais implicações para o processo de inovação e bem estar social.

A Ciência da Política Científica, segundo Valdez e Lane (2008), é um esforço que tem o objetivo de prover, a partir de pesquisa interdisciplinar, bases quantitativas de rigor científico para formuladores de políticas e pesquisadores com o intuito de analisar, principalmente, o financiamento da pesquisa, e avaliar seus resultados. Feller (2011, p. 131) cita o *Office of Management and Budget's* (<https://www.whitehouse.gov/omb/>) para situar um chamado para “(...) to develop a ‘science of science’ policy tools that can improve management of their research and development portfolios and better assess the impact of their science and technology investments”. Ou, a mesma base quantitativa de análise para a ciência que Derek de Solla Price iniciou com a cienciometria nos anos sessenta do século passado (SOLLA PRICE, 1963).

269

De fato, o Estado tem a preocupação de produzir e analisar dados de qualidade que proporcionem um olhar mais qualificado para o direcionamento dos recursos investidos na pesquisa. Isso se dá não só em um ambiente de recursos financeiros cada vez mais escassos para custear pesquisas cada vez mais grandiosas e dispendiosas, mas também com uma crescente preocupação social sobre os caminhos que a ciência vem tomando, especialmente nas respostas que ela não mais consegue apresentar. O que se sabe é que os dados estão disponíveis, mas dispersos, tanto por diferentes agências de fomento como por instituições de pesquisa. Um dos primeiros desafios que se impõe, portanto, diz respeito à coleta, reunião, organização e análise de dados.

Ainda nos Estados Unidos algumas iniciativas vêm sendo desenvolvidas, como o *Deep Insights Anytime, Anywhere (DIA2)* – <http://www.dia2.org/>), também por iniciativa da NSF. O projeto orientou a construção de uma ferramenta de visualização interativa que auxiliasse os tomadores de decisão, assim como os pesquisadores, na análise de dados sobre financiamento para pesquisa e sua relação com as publicações e resultados proveniente destes financiamentos.

A pesquisa aqui relatada partiu da indagação sobre o como uma estratégia dessa natureza poderia ser desenvolvida para o caso brasileiro. O financiamento público à pesquisa no Brasil se faz, quase majoritariamente, por meio de suas agências de fomento, notadamente o Conselho

Nacional de Desenvolvimento Científico e Tecnológico (CNPq), a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), e a Financiadora de Estudos e Projetos (FINEP). Estas agências disponibilizam os dados sobre o financiamento realizado isoladamente umas das outras, e em níveis de detalhamento diferentes, o que dificulta agregar e analisar os dados.

Entretanto, o país conta com uma fonte de dados onde são publicados todos os financiamentos para a pesquisa de nível federal, o Diário Oficial da União (DOU). O DOU, criado em 1808 teve como objetivo principal dar transparência às ações da família real. Esta premissa é mantida até hoje, ou seja, dar publicidade sobre todas as atividades do Estado, entre outros temas, sobre o financiamento das atividades de ciência e tecnologia, especificamente em sua Seção 3, onde são publicados as chamadas e os resultados de editais de financiamento, além de compras e editais de concursos. O DOU é disponibilizado na internet em formato PDF por páginas isoladas, podendo ser encontrado de forma impressa desde o ano de 1990. Assim, como todo o financiamento da esfera federal é ali publicado, este passa a ser uma fonte de dados valiosa e inequívoca para identificar o fomento à pesquisa científica no país.

Nesse sentido, a pesquisa aqui relatada foi exploratória, e teve tanto um caráter metodológico, ou seja, o desenvolvimento de uma metodologia de extração de dados sobre financiamento à pesquisa a partir do DOU, quanto uma vertente analítica, quando se buscou articular esse financiamento às instituições, pesquisadores e produtos da pesquisa, por meio da vinculação com dados do CV Lattes (<http://lattes.cnpq.br>) e o Google Acadêmico (<https://scholar.google.com.br/scholar>). O foco da análise é a pesquisa em dengue, cobrindo o período entre 2005 e 2014. A pesquisa se insere também tanto no campo do *Big Data em Saúde*, e da *Visual Analytics*, por meio do uso de soluções que auxiliem na visualização dos dados e de suas conexões. A expectativa é que se consiga, minimamente, delinear um conjunto de etapas que permitam a replicabilidade dessa estratégia para outras áreas de financiamento à pesquisa, contribuindo para o aprimoramento das políticas públicas de fomento à Ciência e Tecnologia.

No texto que se segue discute-se, sumariamente, o cenário de financiamento à pesquisa em saúde no Brasil, com foco nas agências de fomento. Na Seção 3 apresenta o DOU e descreve os passos da pesquisa. Os resultados são apresentados na quarta seção. O texto finaliza discutindo as limitações da pesquisa, e os caminhos que podem ser trilhados para o fortalecimento dessa estratégia.

2 BREVE CENÁRIO DO FINANCIAMENTO À PESQUISA EM SAÚDE NO BRASIL

O atual cenário do financiamento da pesquisa em saúde no Brasil começou a desenhar os seus contornos nos anos setenta do século passado após a criação das primeiras agências financiadoras – CNPq e FINEP. A criação do CNPq em 1951 surge como uma resposta a um modelo de inovação *science push*, ou seja, nasce da visão da ciência como o ponto de partida para o desenvolvimento de tecnologias e inovação voltadas ao bem-estar social. Este pensamento teve origem no período pós-segunda guerra quando se iniciou o entendimento de ciência como “motor do progresso” (RUIVO, 1994). De forma clara, o CNPq foi criado naquele modelo de linearidade anteriormente descrito, onde:

uma sequência de estágios, em que novos conhecimentos advindos da pesquisa científica levariam a processos de invenção que seriam seguidos por atividades de pesquisa aplicada e desenvolvimento tecnológico resultando, ao final da cadeia, em introdução de produtos e processos comercializáveis. (CONDE; ARAÚJO-JORGE, 2003, p.729).

Pesquisadores do campo da saúde registram, entretanto, que a criação do CNPq serviu como um estopim para o afastamento entre “núcleo hegemônico da pesquisa em saúde e as políticas de saúde, que se traduziu em um afastamento crescente entre a temática da pesquisa e as necessidades de saúde da população.” (GUIMARÃES, 2004, p.377). Este fato se deve ao fato do CNPq adotar o modelo da lógica de financiamento individual, para o pesquisador/cientista, o qual, supostamente, decide o que quer pesquisar, o que nem sempre pode estar alinhado com as demandas sociais.

No mesmo período, em 1965, é criado o Fundo de Financiamento de Estudos de Projetos e Programas (FINEP), quando a ciência passa a ser vista como “solucionadora de problemas” (RUIVO, 1994). Embora o modelo de inovação tecnológica continue sendo linear, aqui ele é regido pelo mercado (demanda). Em 1969 é criado o Fundo Nacional de Desenvolvimento de Científico e Tecnológico (FNDCT). O FNDCT que busca fortalecer o sistema nacional de Ciência, Tecnologia e Inovação (C, T&I):

[...] instituído pelo Decreto-Lei n 719, de 31 de julho de 1969, e restabelecido pela Lei n 8.172, de 18 de janeiro de 1991, é de natureza contábil e tem o objetivo de financiar a inovação e o desenvolvimento científico e tecnológico com vistas em promover o desenvolvimento econômico e social do País. (BRASIL, 2007).

Nas três décadas que se seguiram, ambos os modelos, orientados pela ciência (*science push*) e pelo mercado (*technology pull*) se seguiram acompanhados da criação das agências financiadoras de caráter local, ou estadual, as denominadas FAPs – Fundações de Amparo à Pesquisa. A partir dos anos noventa emerge um novo modelo de financiamento em C&T, com a criação dos Fundos Setoriais, vinculados ao FNDCT. O Fundo é agora composto por 16

Fundos Setoriais, sendo 14 referentes a setores específicos e dois transversais. De acordo com o sítio da FINEP (<http://www.finep.gov.br/>), “desde sua implementação nos anos recentes, os Fundos Setoriais têm se constituído no principal instrumento do Governo Federal para alavancar o sistema de C, T&I do País.” (FINEP, 2012).

O Fundo Setorial de Saúde (CT-SAÚDE) teve seus mecanismos de financiamento instituídos pela Lei nº 10.332, de 19 de dezembro de 2001, e regulamentados pelo Decreto nº 4.143 de 25 de fevereiro de 2002, e tem como foco: (i) a “capacitação tecnológica e inovação tecnológica nas áreas de interesse do Sistema Único de Saúde – SUS” e (ii) a “difusão e incorporação de novas tecnologias visando ampliação do acesso aos bens e serviços em saúde”. (BRASIL, 2002, p.3).

A Finep não disponibiliza diretamente os dados sobre seus investimentos em pesquisa. Esta agência de fomento redireciona o acesso ao Portal da Transparência do Governo Federal (<http://www.portaltransparencia.gov.br>). Não existe, portanto, um relatório consolidado com informações sobre o quanto ela distribui por ano, para que instituições ou para quais temáticas.

Além dos dois órgãos supracitados, CNPq e FINEP, o Brasil ainda conta no grupo das agências públicas de fomento em nível nacional, como a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), o Banco Nacional de Desenvolvimento Econômico e Social (BNDES), com as Fundações de Amparo à Pesquisa no âmbito estadual, que seguem a orientação do sistema nacional de ciência e tecnologia e visam estimular o desenvolvimento da ciência em caráter local e, no caso da saúde, o Departamento de Ciência e Tecnologia (DECIT) do Ministério da Saúde.

Estas agências apresentam, em seus *websites*, os dados sobre o financiamento realizado isoladamente umas das outras, e em níveis de detalhamento diferentes, o que dificulta a agregação e análise dos dados.

Todas estas diretrizes citadas acima compõem uma política de fomento à pesquisa que busca catalisar o Sistema Nacional de Inovação com interseção clara no Sistema Nacional de Saúde. Sendo assim, é apropriado dizer que esta política atende às necessidades de um Sistema Nacional de Inovação em Saúde (SNIS).

A despeito dessas políticas explícitas, muito pouco se sabe sobre a consequência direta dos investimentos em pesquisa no Brasil, de forma geral. Segundo dados provenientes do *Scimago Journal and Country Rank* (www.scimagojr.com) e do *Scimago Institution Rank* (www.scimagoir.com) a visibilidade da produção científica do Brasil cresceu mais de 30 vezes entre 1996 e 2018 na área da saúde, tendo sido capitaneado pelas Ciências Biológicas, com crescimento de cerca de 14 vezes, seguido da Medicina, com 7,6 vezes.

Reinaldo Felipe Nery Guimarães é, seguramente, o autor mais produtivo e combativo no campo do Planejamento, Gestão e Políticas de Ciência e Tecnologia e de Saúde no Brasil. O autor afirma que no Brasil, o financiamento em pesquisa é majoritariamente realizado pelo governo federal (GUIMARÃES, 2006) e, ainda que reconheça a “qualidade e a transparência nas ações de fomento, em particular as realizadas pelas agências do MCT², pela Capes, bem como pela maioria das agências estaduais” (GUIMARÃES, 2004, p.379), aponta que a transparência se dá de forma fragmentada, e não articulada. Nesta mesma linha, o autor cita que o modelo de fomento é descentralizado e plural; registrando que isso se traduz em uma das qualidades do modelo, uma vez que o fato de contar com diversas agências de fomento se torna “um instrumento de proteção dos executores de pesquisa contra eventuais obscurecimentos no que se refere à qualidade e à transparência nas ações de fomento.” (GUIMARÃES, 2004, p.380). Entretanto, esta característica é também um desafio pela “ausência de mecanismos de coordenação adequados entre as múltiplas instâncias de fomento, em especial entre os dois atores principais, o MCT e o MS³”.

As agências de fomento, tanto em nível federal quanto em nível estadual, tornam público o financiamento direcionado às pesquisas, particularmente por meio de suas páginas institucionais na Internet, mas não com o detalhamento necessário para que seja realizada uma unificação destes dados de forma a permitir uma descrição global sobre o financiamento à pesquisa no Brasil. (MARQUES, 2017).

A alternativa disponível para traçar um cenário do investimento público em pesquisa no Brasil é o G-Finder (<https://gfinder.policycuresresearch.org>), um levantamento realizado por uma organização internacional que tem por objetivo acompanhar o financiamento público, privado e filantrópico para pesquisa básica e desenvolvimento de produto para as prioridades globais da saúde. O foco principal do G-Finder é o financiamento para pesquisa e desenvolvimento em/para doenças negligenciadas.

Os dados compilados pelo G-Finder são auto declaratórios. Indo além, registra-se o fato de que 26,5% do montante apresentado pelo G-Finder é referente ao financiamento proveniente das Fundações de Amparo à Pesquisa (FAPs), não contemplando, portanto, o financiamento público de abrangência nacional.

² Ministério da Ciência e Tecnologia

³ Ministério da Saúde

3 O DOU COMO FONTE DE INFORMAÇÃO PARA C&T E O CAMINHO DA PESQUISA

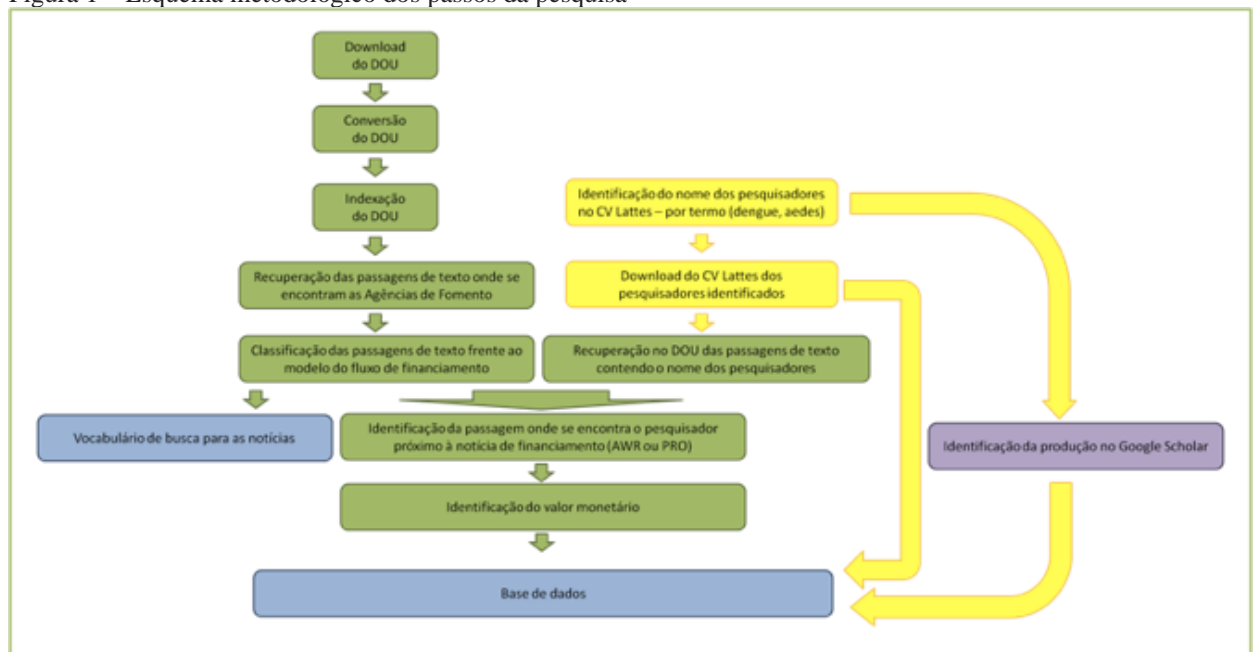
O Diário Oficial da União é organizado em três seções distintas, sendo que a Seção 1 apresenta as leis e decretos; a Seção 2 apresenta os atos da Administração Pública Federal (APF) e a Seção 3 publica as chamadas e os resultados de editais de financiamento, além de compras e editais de concursos. Análise prévia, realizada manualmente, indicou que as decisões relacionadas ao financiamento de pesquisa estavam registradas somente na Seção 3.

Os registros de financiamento contidas na Seção 3 possuem de um modo geral, os seguintes metadados: agência de financiamento, nome do pesquisador, e recurso liberado. Em alguns poucos registros é ainda possível encontrar o número do Edital lançado pela agência de financiamento, porém, na maioria das vezes, este número somente aparece quando da publicação do resultado do Edital, o que tornou impossível vincular os registros e criar os *links* do fluxo. Segue-se, a metodologia para extração dos mesmos.

Além do levantamento dos dados do DOU, lançou-se mão também do CV Lattes, de onde foram identificados os nomes dos pesquisadores doutores que possuíam em algum ponto de seu CV a palavra “dengue”, seguido de uma busca no Google Acadêmico para recuperação do quantitativo de artigos vinculados à temática. A Figura 1, a seguir, traz a sequência realizada para cada um destes processos, e como elas se interligam, tendo como desfecho a criação de uma base de dados.

274

Figura 1 – Esquema metodológico dos passos da pesquisa



Fonte: Marques (2017, p.65).

A primeira etapa realizada foi a coleta de dados efetuando o *download* do DOU, em formato PDF, para um servidor Linux capaz de armazenar os arquivos coletados. Tecnicamente, o DOU é disponibilizado na Internet por página, ou seja, é possível recuperar cada página do DOU independentemente, porém não é possível visualizar ou baixar todo um número do DOU em um arquivo único. A coleta e a tabulação dos dados foram realizadas automaticamente por *scripts*. Para fins desta pesquisa os *scripts* foram desenvolvidos em linguagem Python⁴.

O segundo passo foi a conversão de cada arquivo PDF em texto, de forma a promover uma etapa de busca e identificação dos metadados escolhidos. Para tanto, as páginas em texto foram indexadas. A indexação foi feita por meio do *Whoosh* (<https://pypi.org/project/Whoosh/>). A indexação de um texto permite uma rápida recuperação das páginas que contém o termo buscado. O *Whoosh* é totalmente integrado ao *Python*, necessitando apenas que o *script* que realiza a indexação acesse-o, e solicite a indexação.

Finalizada a indexação, com o objetivo de identificar e modelar o fluxo de informação sobre o financiamento público federal foi realizada uma busca englobando as três seções do DOU na base de dados tanto com os nomes das agências de fomento de nível federal quanto por suas siglas, a saber: Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Departamento de Ciência e Tecnologia (DECIT) e Financiadora de Estudos e Projetos (FINEP). Os resultados foram filtrados pelos anos de 2005 a 2014, inclusive.

Sendo assim, os trechos de texto que continham as agências de fomento foram salvos acrescentando a estes identificadores não sequenciais para cada agência de fomento, a saber: 1 (um) para CNPq, 2 (dois) para CAPES, 4 (quatro) para FINEP e 8 para DECIT. Esta decisão se deu pelo fato de onde aparecesse mais de uma agência de fomento, o identificador seria a soma dos identificadores individuais.

A partir das passagens de texto recuperadas acima, foram classificadas manualmente 1.200 passagens dentro das categorias: chamada para financiamento (CALL), publicação de resultado (RES), recebimento de financiamento (AWR), aditivo para o financiamento (PRO), prestação de contas (CON) e outro. Dentre o subgrupo “outro”, verificou-se a existência de passagens de texto que caracterizavam de forma bem clara “compras” e “concursos”. Essas passagens foram também salvas, como “Outro”, dada a possibilidade de, no futuro, iluminar

⁴ Python é uma linguagem orientada, ágil, possui um conjunto amplo de bibliotecas que a potencializam.

outras discussões. Esta classificação serviu de base para o treinamento do classificador automático de todas as outras passagens de texto.

Em paralelo ao processo de indexação, foi realizada uma busca na base no Currículo Lattes (<http://lattes.cnpq.br/>) com o objetivo de recuperar o nome dos pesquisadores que tivessem em seu currículo qualquer vinculação com a pesquisa em dengue. Para esta busca foram utilizados dois critérios: (i) busca exclusiva por doutores e, no campo “esta frase exata” foi digitada o termo “dengue”. A lista contendo 5.644 pesquisadores recuperados foi salva contendo o nome do pesquisador e o respectivo número identificador do CV Lattes, de 10 (dez) dígitos.

Neste ponto, buscou-se no DOU pelo nome de cada pesquisador existente na lista de pesquisadores recuperados do CV Lattes, e verificou-se a distância da página recuperada para a passagem de texto classificada como AWR ou PRO. Caso a distância fosse de até 3 (três) páginas, considerou-se que aquele pesquisador estaria potencialmente recebendo financiamento. Finalizou-se esta etapa buscando valores financeiros após o nome do pesquisador, dentro da mesma passagem de texto, e utilizou-se este valor como valor de financiamento.

Ainda nesta linha, buscando atividades científicas dos pesquisadores, foi utilizada a ferramenta *ScriptLattes* (<http://scriptlattes.sourceforge.net/>). Ao baixar os currículos dos pesquisadores, o *ScriptLattes* gera um arquivo contendo, de forma estruturada, em arquivo XML⁵, os dados existentes no currículo de cada pesquisador, além de dados processados pela própria ferramenta. Este arquivo foi acessado de modo a recuperar variáveis de análise para este estudo. Sendo assim, foram recuperados (i) a instituição à qual o pesquisador está vinculado; e (ii) dados geográficos (cidade, estado).

Embora o *ScriptLattes* também recupere a produção científica do pesquisador, optou-se por utilizar outra fonte de dados com vistas a explorar uma eventual relação entre financiamento e produção científica, o Google Acadêmico (<https://scholar.google.com.br>). Esta opção se deu pelo fato de que o CV Lattes somente possui o título do texto científico, o que não garantiria se a produção é realmente referente à dengue, ou outra temática. Sendo assim, foi realizado um levantamento, a partir do Google Acadêmico, da produção de cada um dos pesquisadores identificados no CV Lattes. Buscou-se, assim, a produção dos referidos pesquisadores utilizando as regras: (i) filtro de data: 2005 e 2015, inclusive; (ii) filtro “com a frase exata”

⁵ Do inglês *eXtensible Markup Language*, é uma linguagem de marcação para a criação de documentos com dados organizados hierarquicamente.

utilizando o termo “dengue”, podendo ser no título, resumo, ou texto completo; e (iii) não foram recuperados patentes ou citações. Este levantamento gerou dados como o nome do pesquisador, o ano e o quantitativo de produção científica para cada ano, totalizando 8120 artigos distribuídos ao longo dos dez anos.

4 RESULTADOS

Os 10 anos de análise do DOU contabilizaram um total de 967.736 páginas, somando as seções 1, 2 e 3 do mesmo. A Seção 3, onde são publicadas as notícias sobre financiamento para a pesquisa, somou um total de 483.825 páginas. A partir destas foram identificadas e classificadas um total de 101.030 notícias sobre financiamento de pesquisa. Para este período de dez anos foram identificados um total de cerca de R\$1.080.000.000,00, sendo que 58% (625 milhões de reais) destes distribuídos pelas cinco instituições que mais receberam recursos financeiros neste período.

Gráfico 1– Financiamento em nível federal para pesquisa em dengue

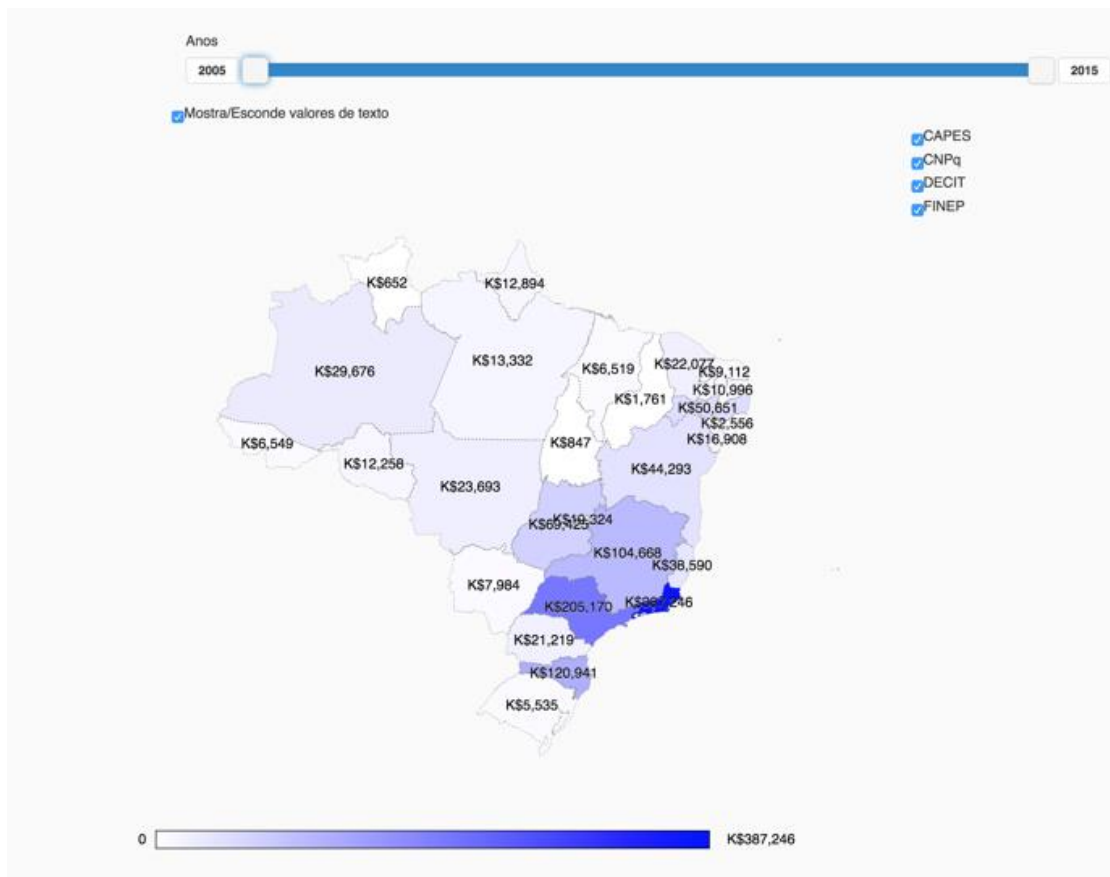


Fonte: Marques (2017, p. 112).

A partir da visualização dos resultados por meio de grafos e mapas, foi possível concluir que nos 10 anos pesquisados, a região Sudeste ocupou o primeiro lugar, sendo este fato previsível, pois a referida região possui cerca de 50% dos pesquisadores nesta temática. Já em relação às agências de fomento, o CNPq foi o órgão financiador que, sozinho, respondeu por 36% do financiamento para pesquisa em dengue longo dos dez anos, seguido pela Finep com 32%, e pela CAPES com 23%.

Seria oportuno investigar se os picos de financiamento, em anos específicos, estariam vinculados a episódios de recrudescimento da doença, em regiões específicas do país. E, claramente, isso poderia ser feito, também na perspectiva quantitativa, vinculando dados dos boletins epidemiológicos da Secretaria de Vigilância em Saúde⁶.

Figura 2 – Financiamento por região demográfica – 2005 a 2014



Fonte: Marques (2017, p.99).

Considerando que o financiamento para a pesquisa é majoritariamente público e que os pesquisadores estão, em sua grande maioria, vinculados à instituições públicas, não foi surpresa a concentração de financiamento em universidades e instituições de pesquisa públicas, exceção feita à Pontifícia Universidade Católica de Goiás, conforme apresentado abaixo:

⁶ <http://www.saude.gov.br/vigilancia-em-saude>

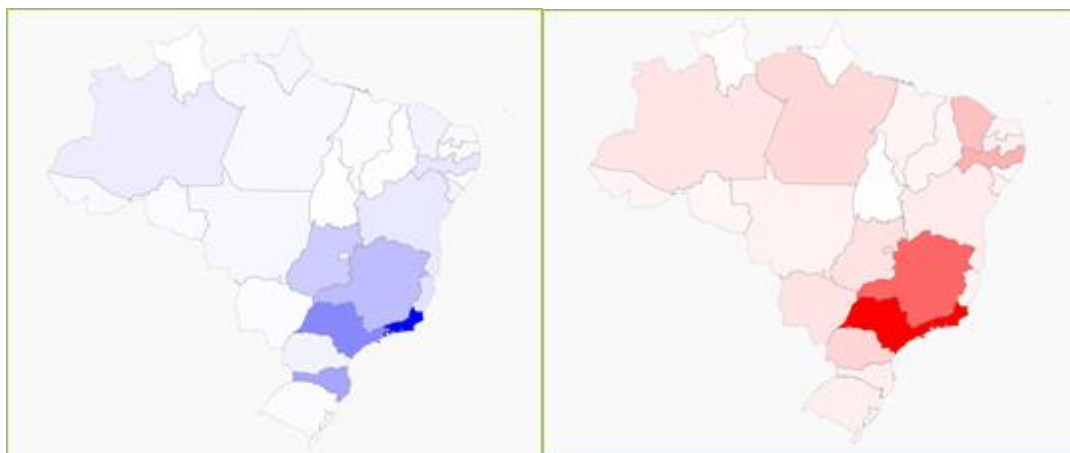
Quadro 1 – Financiamento recebido entre 2005 e 2014 pelas instituições (top 5)

Financiamento em reais (R\$)	(%)	Instituição
247.611.788,47	39,6	Fundação Oswaldo Cruz
174.285.880,33	27,9	Universidade Federal do Rio de Janeiro
80.626.510,15	12,9	Universidade Estadual Paulista Júlio de Mesquita Filho
66.329.039,28	10,6	Pontifícia Universidade Católica de Goiás
56.168.448,77	9,0	Universidade Federal de Minas Gerais

Fonte: Adaptado de Marques (2017, p. 113).

Quando se relaciona o financiamento à produção científica, o mesmo padrão de concentração se manteve, como apresentado na Figura 3, a seguir.

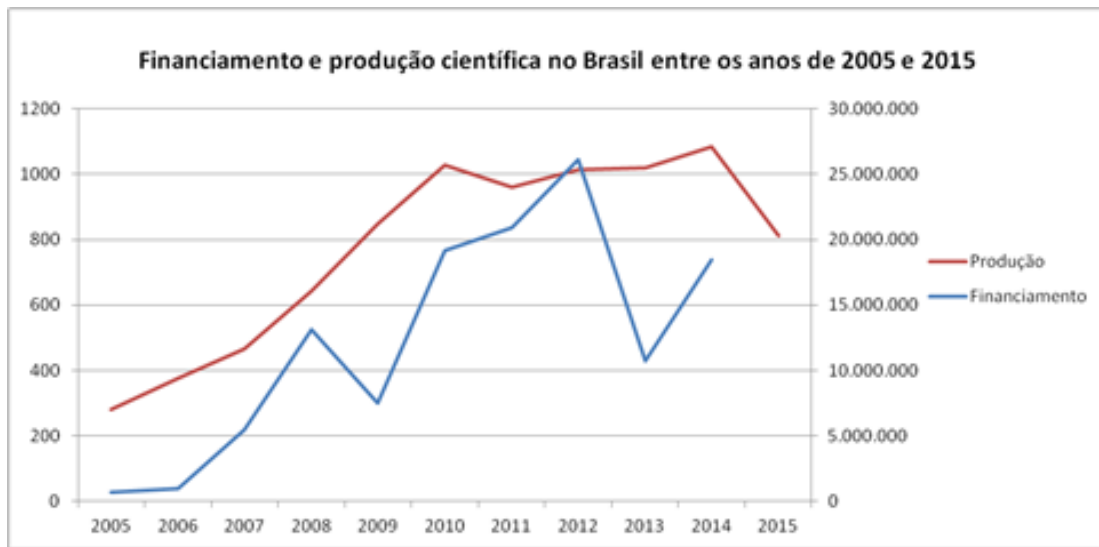
Figura 3 – Distribuição geográfica do financiamento (esquerda) e produção científica (direita) em dengue, no Brasil



Fonte: Marques (2017, p. 114).

De maneira geral, comparando as datas para financiamento à pesquisa (a partir do DOU) e a produção científica (por meio do Google Acadêmico) dos pesquisadores, há indícios que, em média, a produção científica se segue após 2 anos do financiamento. Em linhas gerais, quanto maior o volume de recursos para a pesquisa, maior o número de artigos publicados e referenciados no Google Acadêmico. Por exemplo, de 2008 para 2009, pode se observar uma queda no financiamento, o que se reflete em uma queda na produção científica no período entre 2010 para 2011. Esses são indícios que merecem ser melhor analisados.

Gráfico 2 – Financiamento público federal e produção acadêmica em dengue

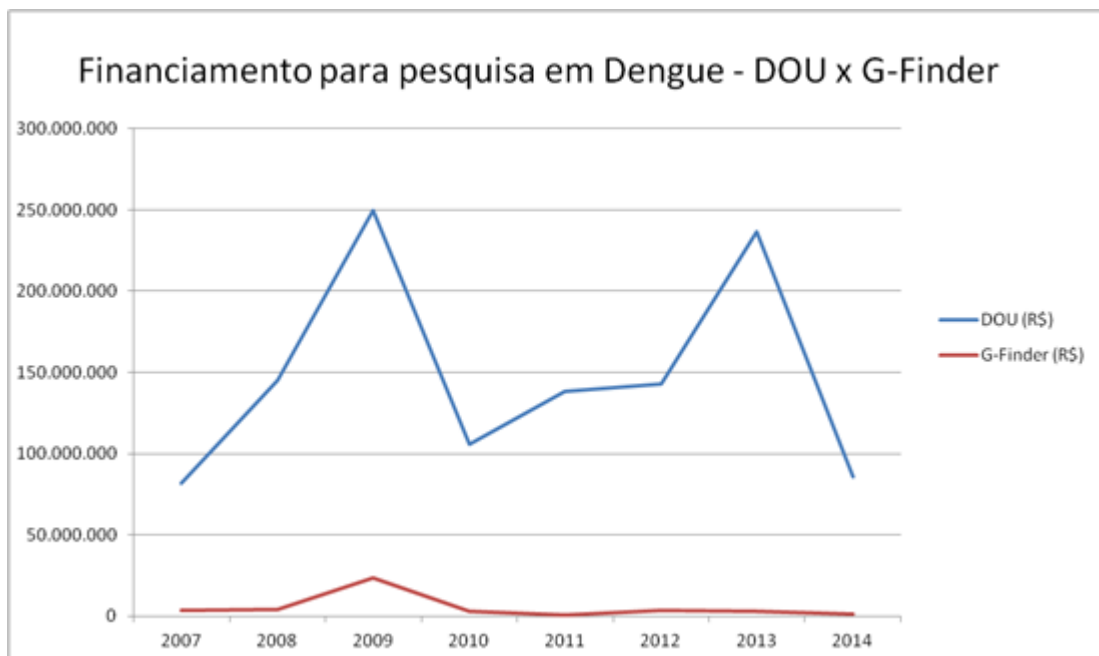


Fonte: Marques (2017, p. 112).

Tomando o G-Finder como padrão ouro, observa-se que os dados do DOU sobre financiamento destinado para pesquisadores em dengue no Brasil, entre os anos 2007 e 2014, é muito superior ao encontrado naquele agregador.

280

Gráfico 3 – Financiamento para pesquisa em dengue – DOU x G-Finder

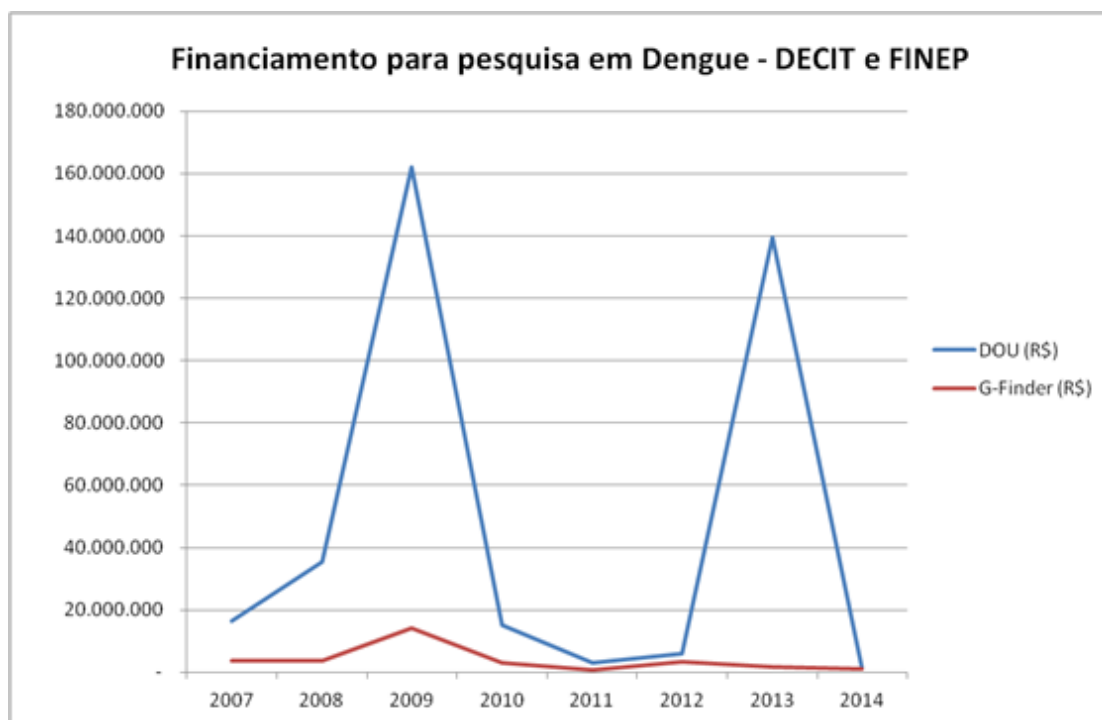


Fonte: Marques (2017, p. 120).

Ressalte-se que, das quatro agências de financiamento deste estudo – CNPq, CAPES, FINEP e DECIT – somente as duas últimas constam dos dados apresentados pelo G-Finder.

Neste caso, considerando apenas as duas agências, pode-se observar que, embora haja disparidade entre os valores, o padrão de variação ao longo dos anos se mantém semelhante. Verifica-se ainda que para os anos de 2011 e 2012 o gap entre as duas fontes é bem menor que para os demais anos.

Gráfico 4 – Financiamento para pesquisa em dengue – Decit e Finep



281

Fonte: Marques (2017, p.121).

O que a presente pesquisa aponta, claramente, é que o DOU parece ser fonte mais fidedigna para traçar os financiamentos públicos de pesquisa, em nível federal.

5 CONSIDERAÇÕES FINAIS

Embora este estudo tenha demonstrado uma metodologia válida para a recuperação do financiamento para a pesquisa em dengue, alguns desafios foram enfrentados.

Um deles foi de caráter técnico. A disponibilização do Diário Oficial da União (DOU) em arquivos no formato PDF é um limitador frente ao processo de recuperação de dados no mesmo. O processo de conversão descrito neste estudo pode servir de base para futuros estudos na conversão de textos do formato PDF para texto (TXT). No entanto, a diagramação do DOU torna o processo extremamente complexo, o que pode acarretar na perda de dados quando da tentativa de organizar o texto sequencialmente. Países como os Estados Unidos da América

disponibilizam sua publicação oficial, o *Federal Register*, nos formatos *Extensible Markup Language* (XML) e texto (TXT), o que possibilita a recuperação dos dados de forma mais fácil. O movimento técnico-político na direção de organização e disponibilização do DOU nestes formatos seria de imensa valia para os estudos das políticas públicas.

Outra questão que serviu como delimitador deste estudo, e que deve ser melhor aprofundada, é a identificação dos pesquisadores que estudam sobre uma determinada temática, no caso deste estudo, a dengue. Dos 5644 pesquisadores identificados no CV Lattes, 4247 não foram identificados como tendo recebido algum tipo de financiamento pelas agências de fomento de nível federal; e 3343 não publicaram nenhum artigo indexado no Google Scholar. Apenas 775 pesquisadores do grupo inicial identificado receberam financiamento e possuem publicação sobre o tema no referido buscador. Por se tratar de uma base declaratória que possui campos no formato aberto, o CV Lattes, embora se torne uma fonte de dados riquíssima em termos de volume de dados e de possibilidades de análise, também se torna de difícil cruzamento entre os inúmeros currículos. Mesmo assim, partir do CV Lattes para identificação de pesquisadores de uma determinada temática não deixa de ser uma boa opção. Todavia, filtrar pelo termo na própria busca avançada desse banco de currículos aumenta o número de pesquisadores recuperados, mas diminui a precisão, pois são recuperados todos os pesquisadores que possuem o termo – no caso deste estudo “dengue” – em seu CV, podendo este termo estar relacionado a uma palestra ou evento em que o pesquisador tenha participado.

Possíveis alternativas para aumentar a precisão na identificação de pesquisadores de uma determinada temática poderiam ser identificá-los: (i) por meio dos Grupos de Pesquisa, (ii) pela produção científica contida em Repositórios Institucionais ou em bases como o Scielo ou ainda (iii) pela mineração do CV dos pesquisadores nos campos referentes a projetos de pesquisa.

Por fim, ressalta-se que técnicas de processamento de linguagem natural, em conjunto com a análise visual dos dados, tornam possível a recuperação e análise de dados do Diário Oficial da União e possibilita um aprofundamento em estudos das políticas de ciência e tecnologia no Brasil. A metodologia aqui desenvolvida e empregada é de possível reprodução. E outros campos e temáticas de estudo podem ser beneficiados por ela.

REFERENCIAS

AGHION, P.; DAVID, P. A.; FORAY, D. Science, technology and innovation for economic growth: Linking policy research and practice in ‘STIG Systems’. **Research Policy**, [s.l.], v. 38, n. 4, p. 681-693, maio 2009. Disponível em:

<https://www.sciencedirect.com/science/article/abs/pii/S0048733309000213>. Acesso em: 08 set. 2019.

BRASIL. Decreto nº 4.143, de 25 de fevereiro de 2002. Regulamenta a Lei no 10.332, de 19 de dezembro de 2001, na parte que institui mecanismo de financiamento para o Programa de Fomento à Pesquisa em Saúde, e dá outras providências. Brasília, DF: Presidência da República, [2002]. Disponível em:

http://www.planalto.gov.br/ccivil_03/decreto/2002/D4143.htm. Acesso em: 09 set. 2019.

BRASIL. Lei nº 10.332, de 19 de dezembro de 2001. Institui mecanismo de financiamento para o Programa de Ciência e Tecnologia para o Agronegócio, para o Programa de Fomento à Pesquisa em Saúde, para o Programa Biotecnologia e Recursos Genéticos – Genoma, para o Programa de Ciência e Tecnologia para o Setor Aeronáutico e para o Programa de Inovação para Competitividade, e dá outras providências. Brasília, DF: Presidência da República, [2007]. Disponível em: http://www.planalto.gov.br/ccivil_03/LEIS/LEIS_2001/L10332.htm. Acesso em: 09 set. 2019.

BUSH, V. **Science, the Endless Frontier**: A Report to the President. Washington, DC: U.S. Government Printing Office, 1945.

283

CONDE, M. V. F.; ARAÚJO-JORGE, T. C. DE. Modelos e concepções de inovação: a transição de paradigmas, a reforma da C&T brasileira e as concepções de gestores de uma instituição pública de pesquisa em saúde. **Ciência & Saúde Coletiva**, v. 8, n. 3, p. 727–741, 2003.

CNPq. **Plataforma Lattes**. Disponível em: <http://lattes.cnpq.br/>. Acesso em: abr. 2016.

FELLER, I. Science of Science and Innovation Policy. In: FEALING, K. (ed.). et al. **The Science of Science Policy: A Handbook**. Stanford: Stanford Business Book, 2011. p. 131-155.

FINEP. **Finep - CT-Saúde, o que é?**. Disponível em:

http://www.finep.gov.br/pagina.asp?pag=fundos_ctorsaude. Acesso em: 11 jun. 2014.

FLANAGAN, K.; Uyarra, E. Four dangers in innovation policy studies – and how to avoid them. **Industry and Innovation**, v.23, n.2, p. 177-188, 2016.

GIBBONS, M. Governance and the new production of knowledge. In: DE LA MOTHE, J. **Science, technology and governance**. London: Continuum, 2001. p. 33-49.

GOOGLE. **Acadêmico**. Disponível em: <https://scholar.google.com.br>. Acesso em: abr. 2016.

GUIMARÃES, R. Bases para uma política nacional de ciência, tecnologia e inovação em saúde. **Ciência & Saúde Coletiva**, v. 9, n. 2, p. 375–387, jun. 2004.

GUIMARÃES, R. Pesquisa em saúde no Brasil: contexto e desafios. **Revista de Saúde Pública**, v. 40, n. spe, p. 3–10, ago. 2006.

MARQUES, Paulo Eduardo Potyguara Coutinho. **Padrão de financiamento à pesquisa em dengue a partir do Diário Oficial da União**. 2017. 142 f. Tese (Doutorado em Informação Científica e Tecnológica em Saúde) -Fundação Oswaldo Cruz, Instituto de Comunicação e Informação Científica e Tecnológica em Saúde, Rio de Janeiro, 2017.

MENA-CHALCO, J. P.; CESAR JUNIOR, R. M. ScriptLattes: an open-source knowledge extraction system from the Lattes platform. **Journal of the Brazilian Computer Society**, v. 15, n. 4, p. 31–39, dez. 2009.

METCALFE, J. S. Innovation systems, innovation policy and restless capitalism. In: MALERBA, F.; BRUSONI, S. (ed.). **Perspectives on Innovation**. Cambridge: Cambridge University Press, 2007. p. 411-454.

MORLACCHI, P.; MARTIN, B. R. Emerging challenges for science, technology and innovation policy research: A reflexive overview. **Research Policy**, v. 38, n. 4, p. 571–582, 2009. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0048733309000146>. Acesso em: 08 set. 2019.

RUIVO, B. “Phases” or “paradigms” of science policy? **Science and Public Policy**, v. 21, n. 3, p. 157–164, 1 jun. 1994.

SOETE, Luc. Science, technology and innovation studies at a crossroad: SPRU as case study. **Research Policy**, v. 48, n. 4, p. 849-857, 2019. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0048733318302695> Acesso em: 08 set. 2019.

SALOMON, J. J. **Science and Politics**. London: The Macmillan Press, 1973.

SCIMAGO Institution Rank. Brazil: Overall Rank. [2017]. Disponível em: <https://www.scimagoir.com/rankings.php?country=BRA&year=2011>. Acesso em: 08 set. 2019.

SCIMAGO. Brazil. [1996-2018]. Journal and Country Rank. Disponível em: <https://www.scimagojr.com/countryrank.php>. Acesso em: 08 set. 2019.

SOLLA PRICE, D. J de. **Little Science, Big Science**. New York: Columbia University Press, 1963.

VALDEZ, B.; LANE, J. **The science of science policy: a federal research roadmap**. Report to The Subcommittee on Social, Behavioral and Economics Sciences, Committee on Science, National Science and Technology Council. Washington, DC: Office of Science and Technology Policy, Executive Office of the President, 2008.