



ESTRATÉGIAS METODOLÓGICAS PARA A GERAÇÃO DE INDICADORES CIENTÍFICOS NA UNIVERSIDADE FEDERAL DO VALE DO SÃO FRANCISCO

Ana Paula Lopes da Silva

Mestre em Ciência da Informação pela Universidade Federal da Bahia, Brasil. Bibliotecária da Universidade Federal do Vale do São Francisco, Brasil.

E-mail: anapaula.llopes@hotmail.com

Francisco José Aragão Pedroza Cunha

Doutor em Difusão do Conhecimento pela Universidade Federal da Bahia, Brasil. Professor da Universidade Federal da Bahia, Brasil.

E-mail: franciscopetroza@ufba.br

Natanael Vitor Sobral

Doutorando em Ciência da Informação pela Universidade Federal da Bahia, Brasil. Professor da Universidade Federal da Bahia, Brasil.

E-mail: natanvsobral@gmail.com

Resumo

O objetivo deste trabalho é desenvolver estratégias metodológicas para a geração de indicadores científicos na Universidade Federal do Vale do São Francisco. As motivações para desenvolvimento do projeto vinculam-se ao desejo de economia de recursos na construção de indicadores, e também, fomentar os estudos cientométricos na instituição como forma de aperfeiçoar os instrumentos de avaliação das atividades científicas existentes na universidade. Partiu-se de um levantamento dos programas de pós-graduação *stricto sensu* da instituição, que juntamente com os docentes atuantes nestes programas, formaram o universo da pesquisa. Definiu-se como fonte inicial para construção da estratégia e congregação dos indicadores, a Plataforma Lattes do Conselho Nacional de Desenvolvimento Científico e Tecnológico, valendo-se do método de extração de dados da ferramenta ScriptLattes e dos meios de processamento e visualização pautados em tabulação estatística, análise de redes sociais e nuvem de tags. De modo geral, considerou-se a estratégia metodológica, construída a partir de ferramentas bibliométricas e cientométricas, viável de aplicação ao universo investigado, possibilitando estabelecer diversos indicadores científicos da pós-graduação da instituição selecionada. O objetivo de economia de recursos na estratégia foi um item prioritário, e para isto, foram priorizadas ferramentas gratuitas e de baixo custo. Enquanto limitação principal das estratégias formuladas, destaca-se a necessidade de realização de algumas atividades de natureza manual, o que torna os recursos metodológicos aqui empreendidos mais adequados a grupos menores. Aponta-se como diferencial das estratégias, a flexibilidade no tratamento dos dados, permitindo a personalização dos indicadores e a replicação dos modelos apresentados em instituições de natureza similar.

Palavras-chave: Indicadores de C&T. Indicadores bibliométricos. Universidade – Pós-graduação. Informação científica.

METHODOLOGICAL STRATEGIES FOR THE GENERATION OF SCIENTIFIC INDICATORS AT FEDERAL UNIVERSITY OF VALE DO SÃO FRANCISCO

Abstract

The objective of this paper is to develop methodological strategies for the generation of scientific indicators in Federal University of the São Francisco Valley. The motivations for development of the project are linked to the desire to save resources in the construction of indicators, as well as to promote the scientometrics studies in the institution as a way to improve the instruments for evaluating scientific activities in the university. It was based on a survey of the university's *stricto sensu* postgraduate programs, which together with the teachers working in these programs, formed the research universe. The Lattes Platform Database of the National Council for Scientific and Technological Development was defined as the initial source for the construction of strategy and congregation of indicators, using the method of data extraction of the tool ScriptLattes and the processing and visualization media ruled in statistical tabulation, social network analysis and tag clouds. In general, we considered the methodological strategy, built from bibliometric and scientometrics tools, feasible to apply to the universe investigated, making it possible to establish several scientific indicators of the postgraduate programs of the selected institution. The objective of saving resources in the strategy was a priority item, and for this, free and low-cost tools were prioritized. As a main limitation of the strategies formulated, it is necessary to carry out some activities of a manual nature, which makes the methodological resources undertaken here more suitable for smaller groups. It is pointed out as differential of the strategies, the flexibility in the treatment of the data, allowing the personalization of the indicators and the replication of the models presented in institutions of similar nature.

Keywords: Indicators of S & T. Bibliometric indicators. University - Postgraduate. Scientific information.

1 INTRODUÇÃO

Os indicadores científicos são instrumentos para a compreensão da dinâmica de produção de uma instituição, país ou área do conhecimento. Assim, entender como aplicá-los no contexto institucional é relevante para que adquiram utilidade, contribuindo com a definição de políticas, divulgação e gestão da produção institucional. Enquanto dispositivos que suportam a criação de indicadores, destacam-se as bases de dados acadêmicas e científicas, que armazenam registros da produção, indexando conteúdos segundo critérios específicos e típicos do universo acadêmico-científico.

Para as instituições que demandam o uso de indicadores, algumas dificuldades se impõem de forma concreta, dificultando o acesso a tais instrumentos de apoio à gestão. Dentre os principais obstáculos, destacam-se a ausência de recursos humanos qualificados para construir e analisar os indicadores, e também, o alto custo das informações que servem como insumo, exigindo das instituições gastos anuais para manutenção de dados acerca de sua produção. Logo, a recuperação de indicadores adequados à gestão acadêmica e científica, fica, por vezes, limitada às instituições que dispõem de altos recursos para investir na aquisição de *softwares* e bases de dados comerciais provedoras de dados científicos.

Em contrapartida, com a evolução dos estudos métricos e a importância atribuída à avaliação da produção científica, surgiram iniciativas de ferramentas cientométricas de baixo custo (*low cost*) que podem ser utilizadas por instituições que dispõem de pouco orçamento para a construção de indicadores. Neste sentido, é primordial trazer para o âmbito dessas instituições o entendimento sobre como se apropriar das ferramentas para o monitoramento de suas produções, gerando indicadores de atividade científica, possibilitando avaliações internas e externas. Para isto, é basilar dispor de um método de recuperação e sistematização desses indicadores de forma padronizada que possa ser executado de acordo com as

necessidades da instituição, bem como ser reproduzido por outras instituições com recursos financeiros semelhantes (SILVA, 2017).

Neste sentido, por motivo de vínculo institucional, definiu-se estabelecer um projeto com o intuito de proporcionar à instituição Universidade Federal do Vale do São Francisco (Univasf), economia quanto à aquisição de ferramentas de empresas proprietárias para extração de dados cientométricos que demandam custos com assinaturas mensais ou anuais. Além do custo, outro ponto desfavorável está na dependência institucional para a geração de dados, constituindo-se assim, um importante problema para a Universidade em questão.

Deste modo, enquanto problema de pesquisa pretende-se responder ao seguinte questionamento: quais são as alternativas de geração de indicadores científicos de baixo custo para a Univasf? Assim, o objetivo deste trabalho é desenvolver estratégias metodológicas para a geração de indicadores científicos na Univasf. As motivações para o desenvolvimento do projeto vinculam-se ao desejo de economia de recursos na construção de indicadores, e também, fomentar os estudos cientométricos na instituição como forma de aperfeiçoar os instrumentos de avaliação das atividades científicas desempenhadas por sua comunidade.

2 QUADRO TEÓRICO

Uma ferramenta bastante significativa e de ampla utilização na investigação cientométrica da produção nacional no Brasil, em relação aos pesquisadores e instituições, é a Plataforma Lattes do CNPq com seus currículos acadêmicos registrados em quantidade estimada em 3.520.867, segundo o Painel Lattes¹. A base de currículos Lattes funciona como uma biografia do pesquisador e mostra de forma cronológica toda a atuação, produção acadêmica e de pesquisa dos cientistas brasileiros, o que faz com que ela tenha grande potencial para a extração de dados científicos, inclusive, também, acadêmicos e tecnológicos. Um fato que contribui para tal aceção é que o preenchimento do currículo está atrelado ao financiamento, avaliação de pesquisas e desempenho dos pesquisadores e programas de pós-graduação (PERLIN *et al.*, 2017), o que implica na manutenção de dados atualizados acerca da produção científica no currículo, sendo uma das fontes mais abrangentes no que se refere ao registro das pesquisas nacionais e regionais, independente da área do conhecimento.

Entretanto, Silva (2007), aponta fatores que, de certa forma, fragilizam a utilização ampla dessa base. Entre eles estão: facilidade e flexibilidade do preenchimento das informações pelo pesquisador no currículo, tanto física quanto temática, e forma de representação e organização das informações na base. Embora apresente estes fatores, a Plataforma Lattes vem ganhando grande importância e credibilidade, sendo utilizada para a recuperação de indicadores em pesquisas nacionais. Perlin *et al.* (2017) reconhecem a necessidade de uma plataforma aberta e consistente para medir as atividades de produtividade acadêmica e a Plataforma Lattes é reconhecida internacionalmente como exemplo neste item, não existindo base semelhante que abranja currículos nacionais. A base registra produções e demais atividades dos pesquisadores que não se podem localizar em bases comerciais bibliográficas, todavia, encontram-se registrados em seus currículos.

“Realizar uma compilação ou sumarização de produções bibliográficas para um grupo de usuários cadastrados de médio ou grande porte (e.g. grupo de professores, departamento de pós-graduação) realmente requer um grande esforço manual suscetível a falhas” (MENA-CHALCO; CESAR JR., 2013, p. 3). Assim sendo, para facilitar a extração e interpretação dos dados contidos na Plataforma Lattes, automatizando o processo e com objetivo de explorar todo o manancial de informações que a base possui para a recuperação de indicadores,

¹ Disponível em: <http://estatico.cnpq.br/painelLattes/>. Dados extraídos em abril/2017.

existem algumas ferramentas auxiliares como a Plataforma Stela Experta² e ScriptLattes³, mais recorrentes na literatura.

A estrutura da Plataforma Lattes foi desenvolvida pelo Grupo Stela da Universidade Federal de Santa Catarina (UFSC). Essa mesma equipe criou uma ferramenta para extração e gerenciamento voltada para as instituições de ensino e pesquisa baseada nas informações presentes nos currículos acadêmicos da Plataforma Lattes, chamada Stela Experta. “A Plataforma Stela Experta foi desenvolvida a partir de uma demanda de mercado: trazer para dentro das IES as informações sobre os currículos Lattes de seus docentes e discentes para utilização nos processos de gestão, inicialmente da Pró-Reitoria de Pesquisa e Pós-Graduação” (STELA..., 2017).

A Stela Experta possibilita a recuperação de indicadores e relatórios sobre a instituição, docentes e discentes de forma ágil e facilitada, possibilitando combinação de dados. Permite a extração automática e periódica de dados dos grupos de pesquisa da instituição e dos currículos Lattes, disponibilizando por meio de uma página acesso às informações para análise e gestão da pesquisa, além de emitir relatórios da produção bibliográfica, técnica, artística, cultural, orientações, entre outras.

Embora baseada numa plataforma aberta, a Stela Experta não é um serviço gratuito, sendo comercializada às instituições que queiram manter e acessar seus dados acadêmicos, científicos e tecnológicos. Os custos variam conforme a quantidade de currículos a serem registrados para monitoramento. Para uma instituição que possua cerca de 500 docentes, por exemplo, os investimentos necessários para adoção da Plataforma, segundo orçamento fornecido pela empresa, seria de R\$ 19.500,00 (dezenove mil e quinhentos reais) para implantação e capacitação para uso do sistema, acrescido de um valor mensal de R\$ 2.800,00 (dois mil e oitocentos reais) referentes à manutenção dessas informações (orçamento obtido em 2016). A instituição deverá solicitar permissão ao CNPq para extrair seus grupos de pesquisa e os currículos Lattes de seus pesquisadores.

Outra ferramenta auxiliar para a extração de dados da Plataforma Lattes é o ScriptLattes, idealizado por Jesús P. Mena-Chalco e Roberto M. Cesar-Jr. Este *software* livre e gratuito, funciona no sistema operacional *Linux*, na perspectiva de não gerar custos. Passível de adaptações e redistribuição por ser um programa de código aberto, necessita de estrutura simples para sua execução e vem sendo utilizado por diversas instituições e departamentos de pesquisa no Brasil. O programa funciona de forma semiautomatizada, alimentado com dados prévios padronizados para a recuperação dos indicadores a serem investigados. É pertinente ressaltar que a qualidade dos indicadores gerados com o ScriptLattes dependerá do preenchimento adequado das informações pelos pesquisadores em seus currículos.

Ao longo do desenvolvimento do projeto que propiciou a produção deste artigo, identificaram-se diversos mecanismos para a produção, recuperação e elaboração de indicadores científicos concebidos como: fontes de registro/armazenamento de dados, ferramentas de indicadores e *softwares* de tratamento de dados, que estão sumariamente descritas abaixo para fins classificatórios.

² Disponível em: <http://site.stelaexperta.com.br/>

³ Disponível em: <http://scriptlattes.sourceforge.net/>

Quadro 1 – Mecanismos para recuperação de indicadores científicos

PLATAFORMAS DE REGISTRO/ARMAZENAMENTO DE DADOS	FERRAMENTAS DE INDICADORES	SOFTWARES DE TRATAMENTO DE DADOS
Web of Science	Stela Experta	The Vantage Point
Scopus	Script Lattes	UCINET
Google Scholar	Incites Journal Citation Reports - JCR	Netdraw
Scielo	SciVal	VOSviewer
Plataforma Lattes	SCImago	Pajek
Plataforma Sucupira		Bibexcel
Pascal		Microsoft Excel
Microsoft Academic		Publish or Perish
Repositórios Institucionais		CiteSpace
Brapci		Gephi
		Statistica

Fonte: Dados da pesquisa (2017)

3 PROCEDIMENTOS METODOLÓGICOS

A pesquisa, quanto aos seus objetivos, enquadra-se na categoria exploratória, busca explorar a questão dos indicadores no contexto da universidade como lócus de investigação. Também se caracteriza como descritiva ao ponto que descreve determinada população e busca registrar, analisar e correlacionar os dados obtidos acerca da mesma. No que se refere à sua proposta, o estudo é ainda uma pesquisa metodológica que resulta, não apenas na aplicação de um método a uma realidade concreta, mas uma construção, representação e avaliação de uma estratégia metodológica. Utiliza técnicas bibliométricas e cientométricas, com abordagem quantitativa e qualitativa de dados.

Os indicadores explorados nesta pesquisa têm como fonte principal a Plataforma Lattes, a partir da ferramenta ScriptLattes. Entende-se que a mencionada plataforma é o instrumento que melhor reflete a produção científica nacional, compreendendo, inclusive, as publicações não indexadas e realizadas em veículos de impacto circunscrito aos âmbitos local e regional. Sabendo que alguns docentes fazem parte de mais de um programa de pós-graduação, resolveu-se realizar uma coleta individualizada, considerando um programa por vez; e outra geral, com o nome de todos os docentes, porém, eliminando os repetidos. Tal procedimento visou não comprometer o resultado geral da coleta.

Sobre isto, ressalta-se que a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) permite a atuação como docente permanente em, no máximo 3 (três) programas de pós-graduação, conforme a Portaria nº 81, de 3 de junho de 2016 (COORDENAÇÃO DE APERFEIÇOAMENTO DE PESSOAL DE NÍVEL SUPERIOR, 2016). Desta feita, é comum na modelagem da tabela deste tipo de pesquisa, verificar a presença de multivalores nas relações “docente” versus “programa”.

Em relação à produção científica dos programas de pós-graduação, optou-se por elaborar inicialmente um quadro para retratar os números referentes aos quantitativos apresentados por cada programa em relação às tipologias de publicações. Tal indicador, apesar de sua simplicidade, abre a possibilidade de avaliação da evolução ou retrocesso das produções total e individual por tipologia. Dentre as possibilidades advindas desta pesquisa, destacam-se também a verificação de elementos como: qualis, fator de impacto, índice h, colaboração, entre outras análises.

O universo considerado compreende os programas de pós-graduação da Univasf, 13 no total, em nível de mestrado *stricto sensu* (Quadro 2). A instituição conta com 4 (quatro) programas de doutorado aprovados pela CAPES, mas que ainda não estão em funcionamento.

Quadro 2 – Programas de pós-graduação *stricto sensu* da Univasf (em funcionamento)

Categoria	Programa	Área de avaliação	Ano de implantação
Acadêmico	Pós-Graduação em Ciência dos Materiais/CPGCM	Materiais	2007
	Pós-Graduação em Ciência Animal/CPGCA	Zootecnia / Recursos Pesqueiros	2008
	Pós-Graduação em Engenharia Agrícola/CPGEA	Ciências Agrárias I	2011
	Pós-Graduação em Recursos Naturais do Semiárido/PGRNSA*	Farmácia	2011
	Pós-Graduação em Ciências da Saúde e Biológicas/CPGCSB	Interdisciplinar	2013
	Pós-Graduação em Ciências Veterinárias no Semiárido/CPGCVS	Medicina veterinária	2013
	Pós-Graduação em Agronomia - Produção Vegetal/PGAPV	Ciências Agrárias I	2014
	Pós-Graduação em Educação Física/PPGEF	Educação Física	2015
	Pós-Graduação em Psicologia/CPGPSI	Psicologia	2015
Profissional	Mestrado Profissional em Matemática em Rede Nacional/PROFMAT	Matemática / Probabilidade e Estatística	2011
	Mestrado Profissional em Ensino de Física/MNPEF	Astronomia / Física	2013
	Mestrado Profissional em Administração Pública em Rede Nacional/PROFIAP	Administração pública e de empresas, Ciências contábeis e Turismo	2016
	Mestrado Profissional em Extensão Rural/PPGExR	Interdisciplinar	2016

Fonte: Dados da pesquisa (2017)

* Teve sua denominação alterada para Programa de Pós-Graduação em Biociências (PPGB) e homologada na Plataforma Sucupira em 13 nov. 2018.

Destes programas, foram selecionados 158 docentes, que compõem o quadro permanente dos cursos de pós-graduação oferecidos na instituição. Foram observados os que desempenharam suas atividades como docentes permanentes nos programas entre os anos de 2007 a 2016. Esse período foi delimitado devido ao início das atividades da pós-graduação na universidade dar-se em 2007, se desenvolvendo progressivamente.

Nesse intervalo, foi considerado também o ano de implantação de cada programa de pós-graduação, tendo em vista que nem todos surgiram na mesma época. O levantamento foi realizado pela produção de cada docente atuante nos cursos *stricto sensu*, sendo especificado para cada pesquisador, a sua data de vinculação ao programa, considerando suas entradas e saídas no decorrer do período delimitado. Docentes permanentes vinculados aos programas no ano de 2017 não foram considerados para a coleta, tendo em vista o período delimitado.

A extração da produção alimentada nos currículos foi realizada em agosto/2017. Após isto, optou-se por apresentar a estratégia metodológica no formato de um fluxograma de processos, visando salientar as etapas desenvolvidas para fins de crítica, replicação e aprendizado institucional.

4 RESULTADOS E DISCUSSÃO

O objetivo das esquematizações que serão apresentadas não é de natureza comparativa, tendo em vista que os programas foram implantados em anos diferentes, pertencem a campos do conhecimento diversos e o quantitativo de docentes se altera em cada um deles, o que naturalmente faz com que os programas tenham indicadores de desempenho distintos. O foco da análise está centrado na perspectiva da obtenção dos dados, compreensão e percepção da sistematização dos indicadores.

Os dados aqui apresentados poderiam ser esquematizados de diversas formas, porém, escolheram-se os modelos abaixo descritos como forma de atender demandas diversas de produção de indicadores que surgem diariamente no contexto universitário, tais como, tabelas estatísticas, séries históricas, redes sociais, nuvens de *tags* organizadas por representatividade, gráficos de comparação e etc.

É possível perceber uma predominância em todos os programas da veiculação por meio de trabalhos em eventos, seja no formato completo ou resumo, e apresentações que somam 67,29% das tipologias identificadas (Tabela 1). Estes dados abarcam as áreas de Materiais, Zootecnia, Ciências Agrárias, Interdisciplinar, Educação Física, Medicina Veterinária, Administração Pública, Psicologia, Física, Matemática. Mesmo não sendo considerados para avaliações em algumas áreas do conhecimento, os trabalhos em eventos ainda são a forma mais difundida para compartilhar e expor as pesquisas desenvolvidas no âmbito dos programas de pós-graduação da Univasf.

Tabela 1 – Distribuição das publicações por veículo de publicação (2007-2016)

Veículo de publicação	Total dos programas	% por tipologia
Artigo	1691	28,36%
Livro	29	0,49%
Capítulo de livro	189	3,17%
Trabalho em evento	376	6,31%
Resumo em evento	2985	50,06%
Apresentação em evento	651	10,92%
Texto em jornais ou revistas	42	0,70%
TOTAL	5963	100,00%

Fonte: Dados da pesquisa (2017)

A Tabela 1 reflete dados das produções de todos os docentes atuantes nos programas de pós-graduação entre 2007 e 2016. É importante ressaltar que foram delimitadas no momento da recuperação as datas correspondentes ao intervalo de vinculação de cada docente em seu programa, recuperando dados apenas do período de sua real atuação. Por exemplo, docentes que atuaram entre 2007 e 2013 em determinado programa, tiveram apenas a produção desse período contabilizada.

No levantamento individual de cada programa foi gerada a tabela 2, que representa a produção bibliográfica distribuída ao longo dos anos. Considerando-se que apresentar detalhadamente os indicadores de cada programa deixaria o trabalho extenso e repetitivo, e o foco aqui é expressar os avanços metodológicos replicáveis obtidos, optou-se por tomar como exemplo a Pós-Graduação em Ciências dos Materiais (CPGCM), por ser a primeira implantada na Univasf, e que no ano de 2017 completou 10 anos de existência na instituição, com nota 4 na primeira avaliação quadrienal realizada pela CAPES.

Tabela 2– Produção científica dos Programas de Pós-graduação *Stricto Sensu* da Univasf (2007-2016) – Parte I

Categoria	Programa	Ano de implantação		Quant. docentes	Artigos publicados em periódico	Artigo aceito para publicação
Acadêmico	CPGCM	2007		13	170	2
	CPGCA	2008		14	428	2
	CPGEA	2011		15	166	1
	PGRNSA	2011		14	274	5
	CPGCSB	2013		19	157	0
	CPGCVS	2013		14	146	1
	PGAPV	2014		16	111	4
	PPGEF	2015		14	77	0
	CPGPSI	2015		17	62	2
Profissional	PROFMAT	2011		13	13	0
	MNPEF	2013		12	34	0
	PROFIAP	2016		10	36	4
	PPGExR	2016		23	17	1

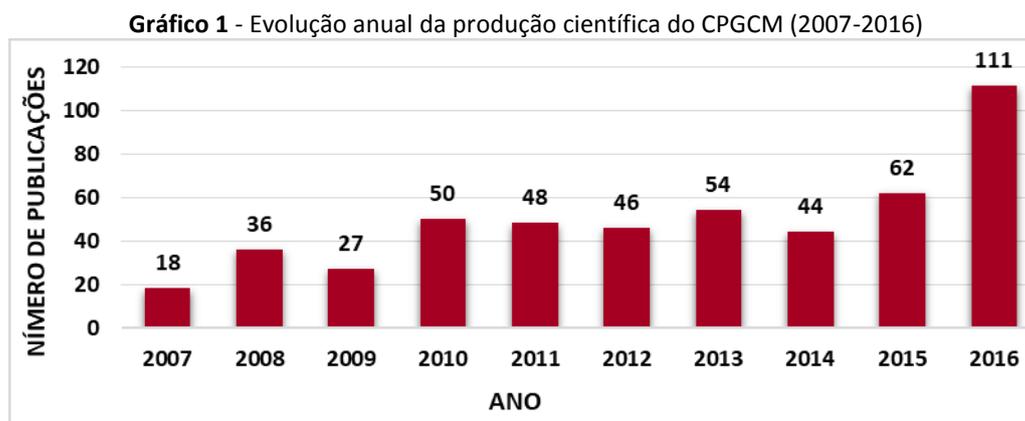
Fonte: Dados da pesquisa (2017)
*Soma de resumo e resumo expandido

Tabela 3– Produção científica dos Programas de Pós-graduação *Stricto Sensu* da Univasf (2007-2016) – Parte II (continuação)

Livro publicado /organizado	Capítulo de livro	Trabalho completo em eventos	Resumo em eventos*	Apresentação em evento	Texto em jornais ou revistas de notícias	Total
1	8	25	237	51	2	496
2	28	192	646	93	0	1391

3	39	55	530	60	0	854
3	33	16	450	39	1	821
6	27	15	256	143	4	608
1	3	16	313	66	1	547
2	10	17	281	11	2	438
1	3	8	95	59	0	243
5	24	8	86	51	18	256
1	1	1	17	28	0	61
0	3	7	61	25	0	130
1	6	13	1	7	13	81
3	4	3	12	18	1	59

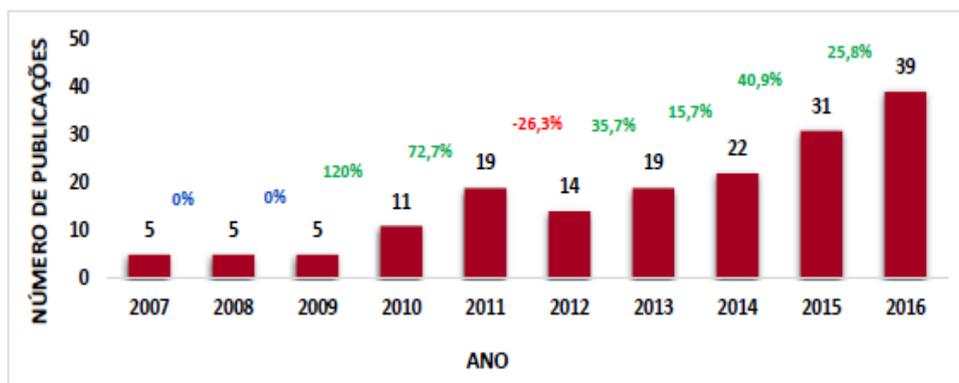
A partir desses dados pôde-se visualizar o desempenho anual da produção, concentração de publicações em determinados anos, oscilações e padrões de regularidade na produção. No caso do CPGCM, constata-se no Gráfico 1, avanços significativos na produção, que teve como grande pico o último ano investigado (2016).



Fonte: Adaptado do ScriptLattes (2017)

E esta mesma pode ser trabalhada individualmente por cada tipo de publicação, possibilitando acompanhar o desenvolvimento da produção nos diferentes veículos de divulgação da pesquisa. No Gráfico 2, constam os dados de desempenho anual apenas dos artigos de periódicos produzidos pelo CPGCM. O programa apresenta 170 artigos distribuídos em 10 anos, e a partir desses dados foi possível obter a taxa de crescimento e regressão das publicações ano a ano, complementada pela taxa de crescimento total do período de atuação do programa, considerando o valor do primeiro e último ano investigados.

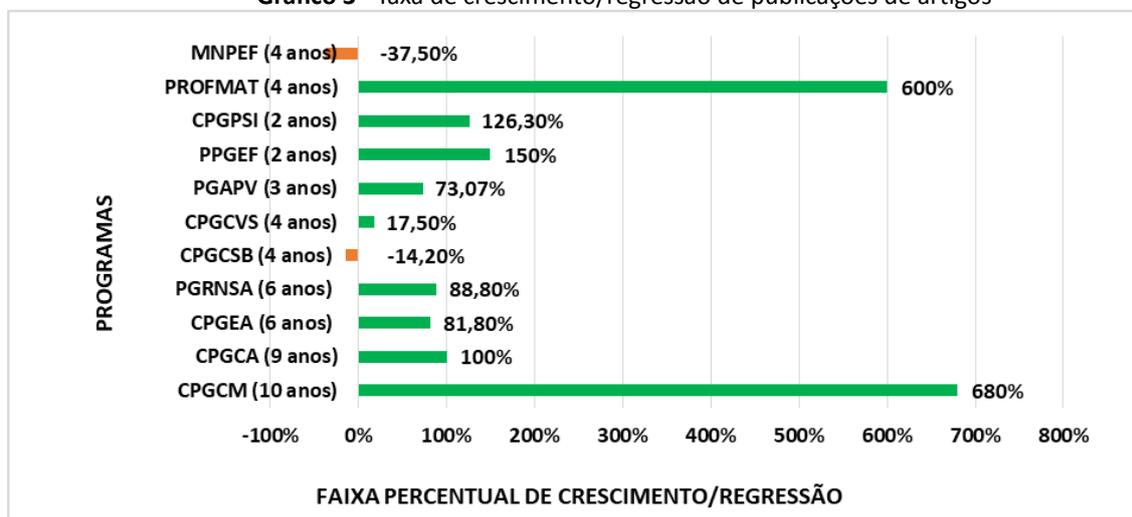
Gráfico 2–Produção anual de artigos publicados em periódicos CPGCM (2007-2016)



Fonte: Adaptado do ScriptLattes (2017)

Para o CPGCM, o programa avançou de 5 artigos publicados em 2007 para 39 em 2016, um aumento de 680% na produção total. Ao estender esta mesma análise individual de produção de artigos para os demais programas, obteve-se a taxa de crescimento ou regressão no total do período de anos investigados em cada programa, considerando seus respectivos períodos de implantação (Gráfico 3). Observa-se regressão na produção de artigos em dois programas, enquanto os demais tiveram aumento significativo desde sua implantação. Os programas PROFIAP e PPGExR, por terem sua implantação ainda em 2016, último ano de investigação da pesquisa, não dispõem de dados posteriores para comparação de crescimento ou regressão.

Gráfico 3 - Taxa de crescimento/regressão de publicações de artigos



Fonte: Dados da pesquisa (2017)

Por meio dos dados individuais dos programas, sistematizou-se a média por tipologia de publicações atingida pela pós-graduação e o desempenho dos programas em relação à mesma. Na Tabela 3, apresentam-se dados de publicação de artigos pelos programas no último ano investigado (2016), permitindo identificar que mais da metade dos programas de pós-graduação estão acima da média de publicações.

O PPGExR e PROFIAP, programas respectivamente abaixo e dentro da média, foram implantados em 2016, o que pode justificar uma menor produção por seu recente processo de estruturação. PROFMAT e MNPEF apresentam produções bem abaixo da média, ambos são mestrados profissionais que estão mais voltados à capacitação de profissionais do que à formação de pesquisadores, e cabe mencionar ainda, que em determinadas áreas o tempo para o desenvolvimento das pesquisas, e conseqüentemente, a frequência de publicações é um pouco mais demorada do que em outras.

Tabela 4 – Média de artigos publicados na pós-graduação (2016)

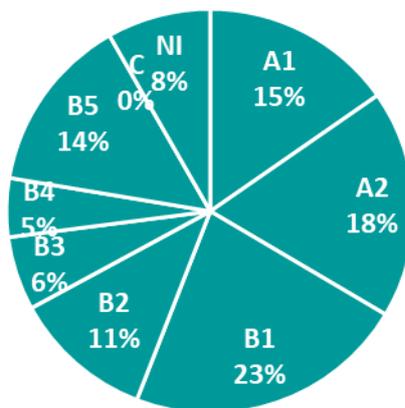
Média de artigos publicados (2016)	35,92308
CPGCM	↑ 39
CPGCA	↑ 40
CPGEA	↑ 40
PGRNSA	↑ 51
CPGCSB	↑ 42
CPGCVS	↑ 47
PGAPV	↑ 45
PPGEF	↑ 55
CPGPSI	↑ 43
PROFMAT	↓ 7
MNPEF	↓ 5
PROFIAP	→ 36
PPGExR	↓ 17

Fonte: Dados da pesquisa (2017)

Legenda: ↑ acima da média ↓ abaixo da média → dentro da média

Em relação ao estrato Qualis, dos 170 artigos retratados no Gráfico 2, publicados pelo CPGCM, tem-se a distribuição dos mesmos em cada estrato: A1 (26), A2 (31), B1 (38), B2 (19), B3 (10), B4 (8), B5 (24) e C (0), Qualis não identificado (NI) (14). Em uma análise percentual, a distribuição de qualificação da produção do programa pode ser observada no Gráfico 4.

Gráfico 4 – Distribuição do estrato Qualis para artigos publicados - CPGCM (2007-2016)

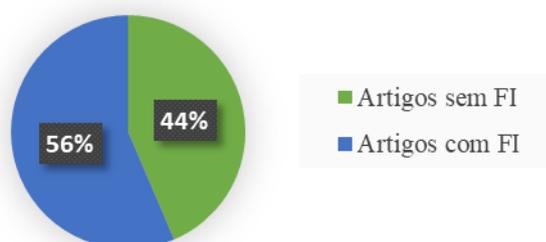


Fonte: Dados da pesquisa (2017)

Essa estratificação se propõe a aferir qualidade dos artigos produzidos pela qualificação atribuída aos periódicos em que estão publicados. Com isto, nota-se que mais da metade da produção do CPGCM está em periódicos com Qualis elevado (A1, A2 e B1), o que consequentemente atribui à pesquisa desenvolvida no programa um prestígio e reconhecimento na comunidade científica, bem como melhores pontuações nos processos de avaliação da pós-graduação.

Outra questão discutida nesta pesquisa é o Fator de Impacto. Por meio de consulta ao *Journal Citation Reports* (JCR), levantou-se, para exemplificação, o impacto dos periódicos no ano de 2016 do CPGCM (Gráfico 5).

Gráfico 5 – Percentual de artigos publicados com fator de impacto - CPGCM (2016)



Fonte: Dados da pesquisa (2017)

Legenda: FI = Fator de Impacto

Dos 39 artigos publicados pelo CPGCM no ano de 2016, 22 estão em periódicos com FI indexados pelo JCR, representando mais da metade (56%) da produção do programa neste ano. Estes artigos estão distribuídos nos seguintes periódicos e FI: *African Journal of Biotechnology*: 0.573; *Animal Reproduction Science*: 1.605; *Antimicrobial Agents and Chemotherapy*: 4.302; *Chemical Engineering Journal*: 6.216; *CrystEngComm*: 3.474; *Fibres & Textiles in Eastern Europe*: 0.626; *Fuel Cells*: 1.706; *Hyperfine Interactions*: 0.209; *Journal of Essential Oil Research*: 0.972; *Journal of Inclusion Phenomena and Macrocyclic Chemistry*: 1.095; *Journal of materials chemistry*: 5.256; *Journal of Medicinal Plants Research*: 0.879; *Journal of the Brazilian Chemical Society*: 1.198; *Latin American Journal of Pharmacy*: 0.298; *Materials Research Bulletin*: 2.446; *Materials Research*: 0.634; *Materials Science & Engineering C*: 4.164; *Radiation Measurements*: 1.442; *Recent Patents on Nanotechnology*: 0.912; *RSC Advances*: 3.108

Percebe-se a publicação do programa em revistas de grande destaque quanto ao fator de impacto como a *Chemical Engineering Journal*, com FI de 6.216 que ocupa 3ª posição no *ranking* específico da categoria Engineering, Environmental – Science que abrange um total de 49 periódicos; *Journal of materials chemistry*, FI de 5.256, 18ª posição na categoria Physics, Applied – Science com 147 periódicos no total; *Antimicrobial Agents and Chemotherapy*, FI de 4.302, 36ª posição na categoria Pharmacology & Pharmacy – Science com 256 revistas (dados de 2017).

O FI é calculado dividindo-se o número de citações recebidas pelos artigos publicados nos dois últimos anos da revista pelo número total de artigos publicados pela revista. Por isso mesmo, o FI vai se alterando anualmente para cada periódico, nesse levantamento considerou-se o FI atualizado do ano de 2016. O FI é uma métrica já utilizada em várias bases, porém, a mais reconhecida é a publicada pela Clarivate Analytics, editora da Web of Science. Neste caso, publicações que estejam em periódicos não indexados pelo JCR não têm a publicação sistematizada desta métrica, a exemplo de muitos periódicos nacionais, em língua

portuguesa, das áreas de ciências humanas e sociais que não possuem cobertura ampla das bases internacionais comerciais.

De mesmo modo que o FI, o índice *h* de autores, também se distingue em relação às bases onde é obtido, sendo calculado de acordo com a quantidade de trabalhos indexados e citações registradas em cada base. Este índice pode ser obtido, principalmente, por meio das bases *WoS*, *Scopus* e Google Acadêmico. No Quadro 3, expõe-se este índice para os pesquisadores do CPGCM recuperados nas três bases, a fim de expor esta distinção. O ScriptLattes já traz um *link* direto para citações no Google Acadêmico de cada pesquisador, onde é possível visualizar o índice obtido pelos docentes nesta base.

Quadro 3– Índice *h* dos pesquisadores do CPGCM

PPGCM	ÍNDICE H		
	Pesquisador	WoS	Scopus
P1	3	2	-
P2	1	1	2
P3	1	1	3
P4	5	6	-
P5	10	12	15
P6	24	26	-
P7	12	13	14
P8	7	7	9
P9	2	3	-
P10	19	20	-
P11	2	1	4
P12	2	3	12
P13	3	3	-

Fonte: Dados da pesquisa (2017)

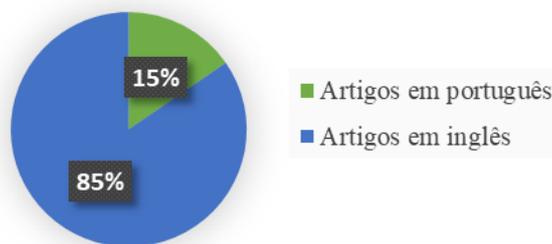
Legenda: P=Pesquisador

Os índices fornecidos pela *Scopus* e *WoS* apresentam maior proximidade pelas características semelhantes que estas bases possuem em relação à abrangência e indexação de publicações. O Google Acadêmico, por ser uma base mais ampla e englobar uma quantidade maior de trabalhos, em diversas áreas, traz números mais elevados. Para os pesquisadores que aparecem sem índice *h* na coluna referente ao Google Acadêmico no Quadro 3, não foi encontrado perfil na base.

A publicação em inglês é, muitas vezes, usada como estratégia para alavancar as publicações por meio de citações, ampliando as métricas de FI, índice-*h* e favorecendo a internacionalização da pesquisa, que são preocupações frequentes dos gestores de instituições acadêmicas. Nesta lógica, vê-se cada vez mais, periódicos nacionais editados em inglês e autores brasileiros publicando em inglês, mesmo em periódicos nacionais com o objetivo de dar visibilidade, internacionalizar suas pesquisas e obter reconhecimento na comunidade científica global.

No Gráfico 6, observa-se que 85% dos artigos produzidos em 2016 pelo CPGCM foram publicados em inglês e em periódicos internacionais. Identificou-se que nenhum dos artigos publicados em português, estão em periódicos indexados pelo JCR, e conseqüentemente, não possuem FI. Todos os periódicos com fator de impacto estão publicados em inglês.

Gráfico 6 – Percentual de artigos por idioma de publicação - CPGCM (2016)



Fonte: Dados da pesquisa (2017)

Outro fator que pode propiciar a promoção de pesquisas no meio científico é a colaboração, especialmente, por meio de coautoria. Todavia, admite-se que a colaboração em pesquisa pode se dar por outras formas, a exemplo do compartilhamento de equipamentos, laboratórios, correspondências, trocas informais de experiências, etc.

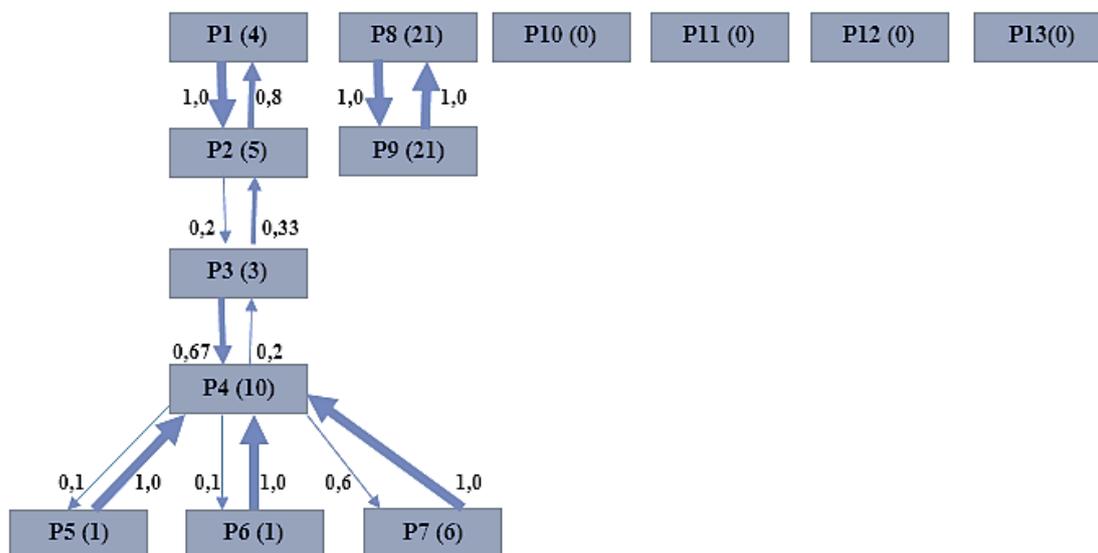
Inicialmente, por meio dos dados recuperados pela ferramenta ScriptLattes, foi possível identificar a colaboração em coautoria intraprogramas da Univasf, ou seja, aquela feita entre os docentes do mesmo programa. Isso foi possível por meio de levantamento individual dos dados de cada programa. Verificou-se também, a colaboração interprogramas, em que se pode visualizar a interação entre os pesquisadores dos diferentes programas da instituição, por meio de um levantamento geral com os dados de todos os docentes que fizeram parte do universo da pesquisa.

Como retrato da colaboração, expõem-se os dados do CPGCM no que diz respeito à colaboração intraprograma. Dos 13 pesquisadores permanentes atuantes no programa, 9 (nove) desenvolveram atividades de colaboração por meio de coautoria de trabalhos com outro membro do grupo. No Gráfico 7 está um esquema representativo desta colaboração adaptado do grafo gerado pelo ScriptLattes. Este grafo apresenta o *AuthorRank*, um valor numérico que mostra o grau de colaboração e impacto do pesquisador no grafo, a partir disso é possível visualizar a proporção da colaboração entre os pesquisadores e qual membro é mais influente na produção do grupo.

O número descrito entre parênteses ao lado de cada pesquisador representa o quantitativo de publicações desenvolvidas em colaboração com membros do programa, as setas direcionais representam o grau e intensidade de colaboração de um pesquisador para o outro de acordo com a quantidade de publicações desenvolvidas. A partir desses dados é possível conhecer o comportamento colaborativo dos docentes atuantes no programa.

Analisando-se, por exemplo, o pesquisador 4 que apresenta maior colaboração no grupo, vê-se que o mesmo desenvolveu pesquisas em coautoria com 4 membros (P3, P5, P6, P7) em diferentes graus de colaboração. O grau de colaboração de P4 para P3 é de 0,2, para o *AuthorRank* esse número, em percentual, corresponde a um peso de 20%, desta forma, das 10 publicações em colaboração de P4, 20%, ou seja, 2 (duas) foram com P3. A colaboração de P4 com P7 tem um grau de 0,6 (60%), assim, das 10 publicações em colaboração de P4 6 (seis) foram com P7. A colaboração de P4 com P5 e P6 possui o mesmo grau de 0,1(10%), o que significa que apenas 1 (um) trabalho de P4 foi escrito em coautoria com cada um destes membros. Inversamente, o grau de colaboração de P5, P6 e P7 para P4 é de 1,0, ou seja, 100% o que mostra que toda a produção em colaboração destes pesquisadores foi realizada com P4. Ainda é possível ver um laço profícuo entre P8 e P9 e os demais membros que não desenvolveram trabalhos em colaboração.

Gráfico 7 - Colaboração intraprograma - PPGCM (2007-2016)



Fonte: Adaptado do ScriptLattes (2017)
 Legenda: P = Pesquisador

De mesmo modo, o levantamento geral de todos os programas permitiu identificar a colaboração interprogramas pelas publicações em coautorias entre os docentes. Nesse levantamento, é possível visualizar um grafo de colaboração com as ligações entre os autores dos diferentes programas, em que cada programa é individualizado por meio de uma cor. O gráfico apresenta também quantitativo de publicações realizadas na colaboração entre os autores e o grau dessa relação. O grafo não foi aqui representado pela sua grande extensão.

Percorrendo o grafo fornecido pelo ScriptLattes, é possível identificar em qual programa atuam os atores mais e menos colaborativos, que programa se relaciona mais com outro, se há um isolamento de autores, entre outras possibilidades. Além do grafo, o sistema apresenta um *ranking* dos pesquisadores mais colaborativos de cada programa e ainda da pós-graduação no geral. Na colaboração interprograma, dos 158 docentes que compõem o quadro permanente dos cursos de pós-graduação no período investigado, 83 apresentaram colaboração com outro docente de algum programa da instituição. Por outro lado, para 75 docentes não foi identificada produção em coautoria no período analisado.

O ScriptLattes fornece um arquivo desta rede de colaboração interprogramas que pode ser trabalhado em *softwares* como o Gephi, possibilitando diversas análises e filtros dentro da rede, permitindo, por exemplo, visualizar a colaboração por determinado grau ou tipologia de publicação. Expõe-se no Gráfico 8 um exemplo de rede elaborada no Gephi com os autores que compõem o componente gigante da rede, ou seja, o grupo com mais conexões entre si. No centro da rede, destacaram-se os pesquisadores pertencentes a mais de um programa e que são responsáveis em grande parte por fazer a interligação e possibilitar a colaboração entre os diferentes programas da pós-graduação.

Por fim, exibe-se o registro do percurso metodológico em forma de processo, conforme o Gráfico 11, buscando possibilitar sua utilização pelos gestores em pesquisa da Univasf na obtenção de dados e indicadores da produção institucional, e ainda, por coordenadores de programas de pós-graduação e grupos de pesquisa ou pesquisadores que queiram realizar estudos cientométricos voltados à produção científica.

A estratégia metodológica explícita no Gráfico 11 foi elaborada em uma ferramenta gratuita para criar diagramas de processos, o HEFLO BPM⁴. Esta ferramenta possibilita a validação do processo registrado e gera automaticamente uma documentação com as etapas e descrições das mesmas.

5 CONSIDERAÇÕES FINAIS

A conjuntura encontrada na Univasf de incipiência no processo de obtenção de indicadores; indisponibilidade de orçamento para investimentos; pós-graduação em desenvolvimento configurou-se em universo propício à investigação, instigando e norteando a pesquisa e estratégia aqui apresentada.

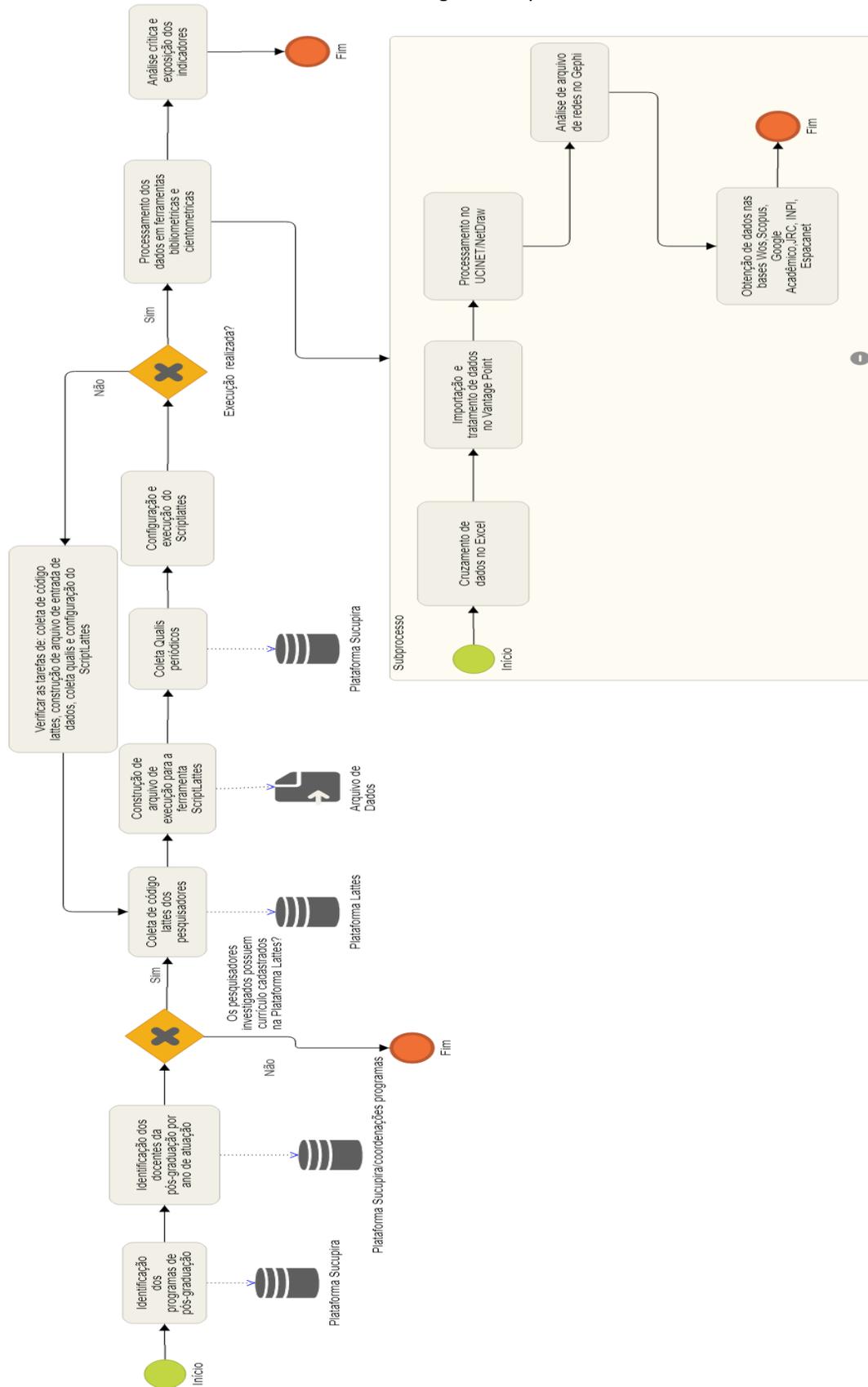
Quanto à estratégia metodológica utilizada, a quase totalidade dos indicadores científicos estabelecidos foi sistematizada a partir do cruzamento de dados fornecidos pela Plataforma Lattes. Aponta-se como um diferencial para a estratégia a flexibilidade de tratamento dos dados, possibilitando reelaborar e fazer ligações entre dados brutos para encontrar a informação desejada e interpretá-la. Pode o pesquisador ou gestor tratá-los de acordo com sua necessidade, permitindo a personalização dos indicadores, principalmente se o investigador domina *softwares* de tratamento e cruzamento de dados.

Por outro lado, etapas da estratégia, como a coleta de código Lattes dos pesquisadores e elaboração de arquivo com dados padronizados para processamento no ScriptLattes, base para todo o processo e a sistematização de alguns indicadores, a exemplo de índice h e FI, necessitaram de coleta e organização manual de dados, o que demanda tempo para o processo, não sendo possível fornecer dados para demandas imediatas. Considera-se esta como uma possível dificuldade para que a estratégia seja aplicada a uma ampla população, sendo mais adequada a estudos de grupos menores.

Conclui-se que as estratégias apresentadas são viáveis ao universo que se propôs investigar, pois possibilitou conhecer diversas vertentes da pós-graduação da Univasf, tendo como diferencial, o uso majoritário de ferramentas gratuitas, encaixando-se na realidade orçamentária da instituição. Com relação ao uso de ferramentas pagas, é possível conveniar-se às instituições que detêm as licenças, ou então, seguir fluxos alternativos amparados na ampla gama de *softwares* gratuitos disponíveis na comunidade científica bibliométrica e cientométrica, que tem crescido significativamente nos últimos anos.

⁴Disponível em: <https://www.heflo.com/pt-br/>

Gráfico 11 - Fluxograma do processo



Fonte: Dados da pesquisa (2017)

REFERÊNCIAS

COORDENAÇÃO DE APERFEIÇOAMENTO DE PESSOAL DE NÍVEL SUPERIOR. **Portaria nº 81, de 3 de junho de 2016**. Define as categorias de docentes que compõem os Programas de Pós-Graduação (PPG's) stricto sensu. 2016. Disponível em:

<http://capes.gov.br/images/stories/download/legislacao/06062016-PORTARIA-N-8-De-3-DE-JUNHO-DE-2016.pdf>. Acesso em: 11 maio 2018.

MENA-CHALCO, J. P.; CESAR-JR, R. M. Prospecção de dados acadêmicos de currículos Lattes através de scriptLattes. In: LETA, J.; HAYASHI, M. C. P. I. (Orgs.) **Bibliometria e Cientometria: reflexões teóricas e interfaces**. São Carlos: Pedro & João Editores, 2013, p. 109-128.

PERLIN, M. S. et al. The Brazilian scientific output published in journals: a study based on a large CV database. **Journal of Informetrics**, v. 11, n. 1, p. 18–31, 2007. Disponível em:

<http://www.sciencedirect.com/science/article/pii/S1751157716301559>. Acesso em: 15 fev. 2018.

SILVA, A. P. L. da. **Recuperação e sistematização de indicadores científicos, tecnológicos e acadêmicos**: uma proposta metodológica para a Univasf. 2017. 149 f. Dissertação (Mestrado em Ciência da Informação) - Universidade Federal da Bahia, Salvador, 2017. Disponível em: <https://repositorio.ufba.br/ri/bitstream/ri/25337/1/ANA%20PAULA%20LOPES%20DA%20SILVA.pdf>. Acesso em: 04 dez. 2018.

SILVA, F. M. E. **Organização da informação em sistemas eletrônicos abertos de Informação Científica & Tecnológica**: análise da Plataforma Lattes. 2007. 163 f. Tese (Doutorado em Ciência da Informação) - Universidade de São Paulo, São Paulo, 2007. Disponível em: <http://www.teses.usp.br/teses/disponiveis/27/27151/tde-17032008-095556/pt-br.php>. Acesso em: 07 jan. 2018.

STELA Experta. **A Stela Experta**. 2017. Disponível em: <http://site.stelaexperta.com.br/sobre/>. Acesso em: 07 fev. 2018.

Artigo recebido em 06/04/2018 e aceito para publicação em 31/01/2019
