



**UNIVERSIDADE FEDERAL DO OESTE DO PARÁ**  
**PRÓ-REITORIA DE PESQUISA, PÓS-GRADUAÇÃO E INOVAÇÃO TECNOLÓGICA**  
**MESTRADO PROFISSIONAL EM PROPRIEDADE INTELECTUAL E**  
**TRANSFERÊNCIA DE TECNOLOGIA PARA INOVAÇÃO**

**FABIANO PAULO NHOATTO**

**SCIENSH: SISTEMA DE APOIO À GESTÃO DA PESQUISA E INOVAÇÃO BASEADO**  
**EM CIENTOMETRIA**

**SANTARÉM - PA**

**2022**

**FABIANO PAULO NHOATTO**

**SCIENSH: SISTEMA DE APOIO À GESTÃO DA PESQUISA E INOVAÇÃO BASEADO  
EM CIENTOMETRIA**

Trabalho apresentado ao curso de Mestrado Profissional em Propriedade Intelectual e Transferência de Tecnologia para Inovação na Universidade Federal do Oeste do Pará como pré-requisito para obtenção do Título de Mestre.

Orientador(a): Prof<sup>o</sup>. Dr. Fábio Manoel França Lobato

Coorientador(a): Prof<sup>o</sup>. Dr. Éfren Lopes de Souza

**SANTARÉM - PA**

**2022**

**Dados Internacionais de Catalogação-na-Publicação (CIP)**  
**Sistema Integrado de Bibliotecas – SIBI/UFOPA**

- 
- N576s    Nhoatto, Fabiano Paulo  
        Sciensh: Sistema de Apoio à Gestão da Pesquisa e Inovação baseado em cientometria./ Fabiano Paulo Nhoatto. – Santarém, 2022.  
        65 p.: il.  
        Inclui bibliografias.
- Orientador: Fábio Manoel França Lobato.  
        Coorientador: Éfren Lopes de Souza.  
        Dissertação (Mestrado) – Universidade Federal do Oeste do Pará, Instituto de Engenharia e Geociências, Programa de Mestrado Profissional em Rede Nacional em Propriedade Intelectual - PROFNIT.
1. Cientometria. 2. Gestão Pública. 3. Transparência. I. Lobato, Fábio Manoel França, *orient.* II. Souza, Éfren Lopes de, *coorient.* III. Título.

CDD: 23 ed. 378.005

---

Bibliotecária - Documentalista: Renata Ferreira – CRB/2 1440

---

Fabiano Paulo Nhoatto

## **SCIENSH: Sistema de apoio à gestão da pesquisa e inovação baseado em cientometria**

Trabalho apresentado ao curso de Mestrado Profissional em Propriedade Intelectual e Transferência de Tecnologia para Inovação na Universidade Federal do Oeste do Pará como pré-requisito para obtenção do Título de Mestre.

Trabalho apresentado. Santarém - PA, 03 de março de 2022:



---

**Prof<sup>o</sup>. Dr. Fábio Manoel França Lobato**  
(Orientador - Universidade Federal do Oeste do Pará)

---

**Prof<sup>o</sup>. Dr. Éfren Lopes de Souza**  
(Coorientador - Universidade Federal do Oeste do Pará)

---

**Prof. Dr. Antônio do Socorro Ferreira Pinheiro**  
(Membro do PROFNIT/ Universidade Federal do Oeste do Pará)



---

**Dr. Carlos Patrick Alves da Silva**  
(Analista de TI no Tribunal de Contas do Estado do Pará e consultor independente em análise de Dados)

Santarém - PA

2022

## **AGRADECIMENTOS**

Os agradecimentos principais são direcionados à UFOPA por permitir minha qualificação no papel de servidor da instituição.

A PROPPIT pelo apoio na implementação e execução do projeto. Especialmente ao Prof. Dr. Bruno Batista e a Prof<sup>a</sup>. Dr<sup>a</sup>. Lenise Vargas Flores.

Ao CTIC pelo apoio e fornecimento da infra-estrutura para a implantação do projeto.

Ao meu coorientador Prof. Dr. Éfren Lopes pela dedicação na orientação e condução.

Ao Adrielson Ferreira nas colaborações com os estudos sobre prospecção e na coautoria de um artigo científico, produto deste trabalho.

Aos meus amigos Adisson Marinho e Samuel Souza pela ajuda em questões técnicas em que cada um possui domínio.

E ao Prof. Dr. Fábio Lobato pela dedicação tanto na orientação como no apoio de toda a implementação do projeto.

Este trabalho foi parcialmente financiado pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) DT-308334/2020 e 154551/2021-0; pela Fundação Amazônia de Amparo a Estudos e Pesquisas (FAPESPA) -PRONEM-FAPESPA/CNPq nº 045/2021.

## RESUMO

Considerando que as Universidades têm como base os pilares de ensino, pesquisa e extensão, faz-se imperativo o uso de métodos para avaliação de sua produção técnico-científica. Por meio da cientometria, mensura-se a contribuição técnico-científica em determinadas áreas, identificando potencialidades e oportunidades institucionais. Com o estudo de caso a Universidade Federal do Oeste do Pará (UFOPA), observou-se os seguintes problemas na gestão de suas pesquisas: dificuldade de coleta e organização dos dados de pesquisa da instituição, integração com sistemas institucionais, publicidades e transparência das produções técnico-científicas. Tais fatos impactam na melhor gestão da pesquisa em relação à alocação de recursos, por exemplo. Visando resolver os problemas supraditos, o presente trabalho propõe o desenvolvimento de um artefato de *software* para gestão e publicização de dados referentes à produção técnico-científica dos pesquisadores vinculados à UFOPA. O artefato possibilita uma maior transparência, inserção social da universidade e capacidade de tomada de decisão orientada a dados, fornecendo auxílio à gestão da pesquisa e inovação e avaliação de políticas tecnocientíficas. O presente trabalho foi conduzido de acordo com o método *Design Science Research*, incluindo o desenvolvimento de módulos para coleta e análise automatizada de dados da Plataforma Lattes, bem como de visualização de Indicadores Cientométricos de produção tecnocientífica. Espera-se que o produto possibilite uma maior eficiência na gestão de recursos públicos, fomentando parcerias inter e intra institucional e aumento do impacto das pesquisas. O sistema é escalável, podendo ser adotado por outras Instituições de Ciência e Tecnologia.

**Palavras-chaves:** Decisão orientada a dados. Cientometria. Gestão pública. Transparência. Visualização da Informação.

## ABSTRACT

Considering that Universities are based on teaching, research, and extension pillars, it is imperative to use methods to assess their technical-scientific production. Through scientometrics, the technical-scientific contribution in certain areas is measured, identifying institutional potentials and opportunities. In the Federal University of Western Pará (UFOPA) case study, the following problems were observed in the management of their research: difficulty in collecting and organizing the institution's research data, integration with institutional systems, advertising, and transparency of productions technical-scientific. Such facts impact research management concerning resource allocation, for example. Aiming to solve the above issues, this work proposes the development of a *software* artifact for the management and publication of data related to the technical-scientific production of researchers linked to UFOPA. The artifact enables greater transparency, social inclusion of the university, and data-oriented decision-making capacity, assisting in the research and innovation management and evaluating technoscientific policies. The present work was carried out according to the Design Science Research method, including the development of modules for automated data collection and analysis from the Lattes Platform and visualization of Scientometric Indicators of technoscientific production. It is expected that the product will enable greater efficiency in public resources management, fostering inter and intra-institutional partnerships and increasing the research impact. The system is scalable and can be adopted by other Science and Technology Institutions.

**Keywords:** Data-oriented decision. Scientometrics. Public management. Transparency. Information visualization.

## LISTA DE ILUSTRAÇÕES

Figura 1 – Etapas do DSR. . . . .	23
Figura 2 – Histograma da última atualização dos currículos Lattes dos pesquisadores da UFOPA. . . . .	26
Figura 3 – Fluxo do <i>software</i> a ser desenvolvido. . . . .	29
Figura 4 – Infográfico dos meios de divulgação e tipos de afirmações de outras IFES. . .	32
Figura 5 – Infográfico dos tipos de informações divulgados pelas IFES. . . . .	33
Figura 6 – Infográfico do edital 03/2020 de pesquisa. . . . .	38
Figura 7 – Exemplo da resultado detalhado de uma pontuação. . . . .	40
Figura 8 – <i>Dashboard</i> com perfil das produções da instituição. . . . .	41
Figura 9 – Filtros e exibições descritivas. . . . .	41
Figura 10 – Perfil do pesquisador. . . . .	42
Figura 11 – Página pública da instituição. . . . .	43



## LISTA DE TABELAS

Tabela 1 – Levantamento sobre <i>softwares</i> de cientometria baseados na plataforma Lattes	20
Tabela 2 – Quantitativo pós-fusão de dados . . . . .	36

## LISTA DE ABREVIATURAS E SIGLAS

API	<i>Application Programming Interfaces</i>
ASCII	<i>American Standard Code for Information Interchange</i>
CNPQ	Conselho Nacional de Desenvolvimento Científico e Tecnológico
DPE	Diretoria de Pesquisa
DSR	<i>Design Science Research</i>
HTML	<i>HyperText Markup Language</i>
IFES	Instituição Federal de Ensino Superior
INPI	Instituto Nacional da Propriedade Industrial
ISI	<i>International Scientific Indexing</i>
JSON	<i>JavaScript Object Notation</i>
NoSQL	<i>Not only SQL</i>
PDI	Plano de Desenvolvimento Institucional
PET	Programa de Ensino Tutorial
PL	Plataforma Lattes
PROPPIT	Pró-Reitoria de Pesquisa, Pós-graduação e Inovação Tecnológica
REUNI	Programa de Apoio a Planos de Reestruturação e Expansão das Universidades Federais
SIGAA	Sistema Integrado de Gestão de Atividades Acadêmicas
UFOPA	Universidade Federal do Oeste do Pará
UFRA	Universidade Federal Rural da Amazônia
UFPA	Universidade Federal do Pará
UFRN	Universidade Federal do Rio Grande do Norte
XML	<i>eXtensible Markup Language</i>
WOS	<i>Web of Science</i>

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO . . . . .</b>	<b>15</b>
<b>1.1</b>	<b>Contextualização . . . . .</b>	<b>17</b>
<b>1.2</b>	<b>Justificativa . . . . .</b>	<b>19</b>
<b>1.3</b>	<b>Objetivos . . . . .</b>	<b>20</b>
<b>1.4</b>	<b>Impactos Esperados . . . . .</b>	<b>21</b>
<b>1.5</b>	<b>Produtos e Artefatos Desenvolvidos . . . . .</b>	<b>21</b>
<b>1.6</b>	<b>Organização do documento . . . . .</b>	<b>22</b>
<b>2</b>	<b>MATERIAIS E MÉTODOS . . . . .</b>	<b>23</b>
<b>2.1</b>	<b>Método de Pesquisa . . . . .</b>	<b>23</b>
<b>2.2</b>	<b>Materiais . . . . .</b>	<b>24</b>
2.2.1	Plataforma Lattes . . . . .	24
2.2.2	Google Scholar . . . . .	26
2.2.3	Sistemas Institucionais . . . . .	27
<b>2.3</b>	<b>Tecnologias . . . . .</b>	<b>28</b>
<b>3</b>	<b>RESULTADOS . . . . .</b>	<b>29</b>
<b>3.1</b>	<b>Artefatos Construídos . . . . .</b>	<b>29</b>
3.1.1	Desafio 1: Edital PIBIC . . . . .	30
3.1.2	Desafio 2: Visualização dos dados . . . . .	30
<b>3.2</b>	<b>Objetivos das soluções . . . . .</b>	<b>31</b>
3.2.1	Desafio 1: Edital PIBIC . . . . .	31
3.2.2	Desafio 2: Visualização dos dados . . . . .	32
<b>3.3</b>	<b>Projeto e Desenvolvimento dos artefatos . . . . .</b>	<b>34</b>
<b>3.4</b>	<b>Demonstração . . . . .</b>	<b>38</b>
3.4.1	Desafio 1: Edital PIBIC . . . . .	38
3.4.2	Desafio 2: Visualização dos dados . . . . .	40
<b>3.5</b>	<b>Avaliação . . . . .</b>	<b>43</b>
<b>3.6</b>	<b>Comunicação . . . . .</b>	<b>44</b>
<b>4</b>	<b>CONCLUSÕES . . . . .</b>	<b>46</b>
	<b>REFERÊNCIAS . . . . .</b>	<b>48</b>

<b>APÊNDICES</b>		<b>51</b>
<b>APÊNDICE A</b>	<b>ARTIGO PUBLICADO NA REVISTA <i>E-TECH: TECNOLOGIAS PARA COMPETITIVIDADE INDUSTRIAL</i> EM 2021 . . . . .</b>	<b>52</b>
<b>APÊNDICE B</b>	<b>FORMULÁRIO DE AVALIAÇÃO . . . . .</b>	<b>66</b>
<b>ANEXOS</b>		<b>69</b>
<b>ANEXO A</b>	<b>EDITAL PESQUISA 2020 . . . . .</b>	<b>70</b>
<b>ANEXO B</b>	<b>QUESTIONÁRIO DE SATISFAÇÃO . . . . .</b>	<b>72</b>
<b>ANEXO C</b>	<b>REGISTRO DE SOFTWARE . . . . .</b>	<b>75</b>

## 1 INTRODUÇÃO

Considerando que as Universidades têm como base os pilares ensino, pesquisa e extensão e que os professores devem desenvolver atividades nestes três eixos, faz-se imperativo o uso de métodos para avaliação de sua produção técnico-científica (HICKS *et al.*, 2015; CARPENTER; CONE; SARLI, 2014). Neste ramo de análise e mensuração da ciência é possível encontrar: a bibliometria, a cientometria, a informetria e a webmetria (MINGERS; LEYDESDORFF, 2015).

Dentre a bibliometria e a cientometria pode-se afirmar que a primeira é focada em análises voltadas à gestão de bibliotecas quantificando os livros, literatura, documentos e revistas científicas. No entanto, a cientometria aplica análises à produção científica, principalmente artigos, a circulação e o consumo da produção científica, visando a dinâmica da ciência: utilidade científica e impacto (PARRA; COUTINHO; PESSANO, 2019). Segundo Hayashi (2013), a cientometria se restringe à mensuração da ciência por meio de produções estritamente científicas.

O advento da Internet e dos bancos de dados eletrônicos tem sugerido a variação do termo bibliometria para os termos: webmetria e informetria, conforme sugerido por Schneider, Larsen e Ingwersen (2009). A informetria considera a aplicação dos princípios bibliométricos a contextos e bases não-acadêmicos, já a webmetria tem sido voltada a bases acadêmicas como o sistema Google Scholar (COSTA, 2010).

Adentrando especificamente no âmbito da pesquisa, a cientometria é o campo de estudo que emprega análise quantitativa para se mensurar o impacto e a produtividade de pesquisas científicas (SILVA; BIANCHI, 2001). Os dados estatísticos elaborados por meio dos estudos bibliométricos mensuram a contribuição do conhecimento científico derivado das publicações em determinadas áreas, representando tendências de uma determinada pesquisa e identificação de temas para novas investigações (SU; LEE, 2010; ZENG *et al.*, 2017).

A cientometria nasce, portanto, como a conhecemos hoje, a partir dos anos 60 do século 20, do cruzamento da documentação científica, da sociologia da ciência e da história social da ciência, com a finalidade de investigar a atividade científica como fenômeno humano, social e mediante parâmetros e indicadores baseados em modelos matemáticos (PARRA; COUTINHO; PESSANO, 2019). Conforme descrito por Price (1963), devido ao crescimento exponencial das produções científicas foi necessário utilizar estatística e métricas matemáticas para analisar e acompanhar essas produções.

Segundo Parra, Coutinho e Pessano (2019) é a partir desses indicadores que se guia o processo de investigação e mensuração da ciência. Eles são úteis no planejamento e avaliação de políticas públicas voltadas para pesquisas (SANTOS; KOBASHI, 2009). A aplicação de indicadores é uma necessidade inequívoca na otimização de recursos, sempre limitados, que se destinam ao fomento da ciência e tecnologia (SILVA; BIANCHI, 2001). Conforme relata Lopes

(2012), a busca pela excelência em áreas de pesquisa e a competição pelos recursos financeiros de agências de fomento são fatores que impulsionam a realização destes estudos e a adoção de métricas cientométricas. Com eles é possível ainda identificar relações de colaboração, de forma a vincular dois temas normalmente abordados separadamente: análise de redes sociais e cientometria (DIGIAMPIETRI *et al.*, 2012).

Existem inúmeras formas de utilização das técnicas cientométricas, e segundo Spinak (1998), a importância pode ser vista analisando a seguinte lista de possíveis aplicações:

1. Identificar tendências e crescimento do conhecimento em todas as disciplinas;
2. Estimar a cobertura de revistas secundárias;
3. Identificar usuários de diferentes disciplinas;
4. Identificar autores e tendências em diferentes disciplinas;
5. Mensurar a utilidade de disseminação seletiva de informações;
6. Prever publicação tendências;
7. Identificar revistas do núcleo em cada disciplina;
8. Formular políticas ajustadas às aquisições de orçamento;
9. Adaptar as políticas públicas.

Além dessas possibilidades informadas, análises cientométricas podem permitir identificação de: (i) autores mais relevantes na área; (ii) autores com mais publicações em determinada área; (iii) maturidade de um certo tema ou área de um autor, grupo ou instituição; (iv) tendências de áreas de pesquisa; (v) produtividade de autores; (vi) conexões de autores e redes de colaboração; (viii) autores importantes para as redes de pesquisa, ou seja, autores que conectam e produzem pesquisas com diversos grupos; (ix) artigos relevantes para iniciar novas pesquisas; (x) árvores genealógicas de autores por orientações ou citações.

Esses indicadores podem ajudar a mensurar a contribuição dada por cada país ao acervo mundial de pesquisa, além de localizar geograficamente as pesquisas por autor, assunto e acompanhar o crescimento em cada área da ciência – exemplos desses indicadores incluem: a quantidade de trabalhos, quantidade de coautorias, quantidade de publicações por autor, por instituição ou país, quantidade de patentes, quantidade de artigos publicados, fator de impacto (ZENG *et al.*, 2017).

De acordo com Dias, Moreira e Dias (2018), editais para financiamento de projetos de pesquisa por agências de amparo à produção de pesquisas no Brasil consideram dados dos currículos cadastrados na plataforma Lattes como uma das formas de avaliação. A Plataforma

Lattes (PL) possui uma vasta fonte de informação para a criação e análise de redes sociais de pesquisadores, tornando-a uma fonte extremamente rica para análise da produção científica brasileira (BALANCIERI *et al.*, 2005). Tais conceitos são pertinentes para o entendimento do contexto do desenvolvimento do trabalho, o qual é apresentado a seguir.

## 1.1 Contextualização

Este trabalho surge como demanda espontânea da Universidade Federal do Oeste do Pará (UFOPA), que apresentou suas dificuldades na gestão da pesquisa e inovação. Assim, todo o estudo é baseado no contexto da instituição, sendo essa aplicada como estudo de caso deste trabalho.

A Universidade Federal do Oeste do Pará (UFOPA) é a primeira Instituição Federal de Ensino Superior (IFES) localizada no interior da Amazônia brasileira, no oeste do Pará. Criada com uma abordagem regional e multi-campi, com sede na cidade de Santarém e com campi localizados nos municípios de Alenquer, Itaituba, Juruti, Monte Alegre, Óbidos e Oriximiná. Foi criada pela Lei nº 12.085, de 5 de novembro de 2009, por desmembramento e integração dos campi da Universidade Federal do Pará (UFPA) e da Universidade Federal Rural da Amazônia (UFRA), envolvida no Programa de Apoio a Planos de Reestruturação e Expansão das Universidades Federais (REUNI).

Segundo o Plano de Desenvolvimento Institucional (PDI) 2019 a 2023 da UFOPA<sup>1</sup>, sua estrutura é composta por 7 institutos e 9 campi, espalhados por 7 municípios da região. Desenvolve um processo de formação de 7.909 alunos ativos, destes 6.924 são da graduação e 985 da pós-graduação, excluindo alunos trancados. Seu quadro funcional é constituído por 1.095 servidores efetivos, sendo 485 docentes e 607 técnicos administrativos em educação, distribuídos nas atividades-meio e atividades-fim. Oferece para a sociedade formação em 42 cursos de graduação<sup>2</sup> e 16 programas de pós graduação<sup>3</sup>. Envolve 49 linhas de pesquisas e 22 áreas cadastradas na plataforma Sucupira<sup>4</sup>. Dentre elas: biotecnologia, biodiversidade, matemática aplicada e propriedade intelectual e transferência de tecnologia para inovação. E seus pesquisadores coordenam 88 grupos de pesquisa<sup>5</sup>.

A gestão da pesquisa na instituição é gerenciada pela PROPPIT, mais especificamente pela Diretoria de Pesquisa (DPE). Esta unidade gerencia a pesquisa por meio de editais de

<sup>1</sup> Disponível em: <http://www.ufopa.edu.br/media/file/site/proplan/documentos/2020/766eald5a36f6bde3acdc4b46199218e.pdf>. Acesso em 29 de julho de 2021.

<sup>2</sup> Disponível em: <http://www.ufopa.edu.br/ufopa/ensino/cursos-de-graduacao-2/>. Acesso em 29 de julho de 2021.

<sup>3</sup> Disponível em: <http://www.ufopa.edu.br/proppit/programas/>. Acesso em 29 de julho de 2021.

<sup>4</sup> Disponível em: <https://sucupira.capes.gov.br/sucupira/public/consultas/coleta/linhaPesquisa/listaLinhaPesquisa.xhtml>. Acesso em 29 de julho de 2021.

<sup>5</sup> Disponível em: <http://www.ufopa.edu.br/proppit/diretoria-de-pesquisa/>. Acesso em 29 de julho de 2021.

fomento e políticas públicas. Para isso utiliza programas como Stella Experta<sup>6</sup> para analisar as produções da instituição e melhorar sua estratégia visando otimizar recursos e melhorar resultados em prol da produção de conhecimento.

A complexidade de realização dessas análises baseia-se na multitude de produções técnico-científicas e docentes envolvidos, especificamente no caso da UFOPA que conta com 485 docentes ativos, mais de 70 mil produções e média 4.769 produções anuais nos últimos dez anos. É patente da necessidade de filtros temporais no cálculo de índices de produtividade dos docentes para distribuição de recursos de editais de fomento, além de outros filtros como área ou seleção de docentes pertencentes a um grupo de pesquisa *etc.* Tal funcionalidade é difícil de ser articulada com o uso de planilhas eletrônicas, uma vez que os dados ficam dispersos (*e.g.*: artigos publicados por pesquisadores da instituição em uma planilha e membros de grupo de pesquisa em outra).

Neste contexto, o presente estudo visa automatizar o processo de coleta e análise da produção técnico-científica dos docentes através de uma ferramenta computacional denominada: Sciensh. Para isso, tal ferramenta combina análise cientométrica e análise de redes sociais, tendo como base os dados obtidos a partir da Plataforma Lattes. Além disso, o Sciensh funciona em sincronia com os sistemas institucionais – Sistema Integrado de Gestão de Atividades Acadêmicas (SIGAA) e Intellectus –, reaproveitando os investimentos já feitos e melhorando a integração institucional.

Alguns desafios operacionais e de gestão foram elencados por servidores atuantes na DPE a fim de que o sistema proposto pudesse auxiliá-los, dentre os quais, destacam-se:

- **Desafio 1:** em busca de agilidade e viabilidade por meio da automatização de processos, o primeiro desafio apresentado pela PROPPIT foi a mensuração de índices de produtividade personalizáveis dos docentes para distribuição de recursos de pesquisa por meio de editais. Sua justificativa baseia-se na grande quantidade de produções e pesquisadores que participam dos processos de seleção. Assim, almeja-se maior eficácia no processo de distribuição e alocação de recursos;
- **Desafio 2:** a PROPPIT evidenciou que para gerir a pesquisa na Instituição é de fundamental importância ter conhecimento sobre a produtividade, áreas de pesquisa, tópicos, tendências, produções recentes, pesquisadores e produções relevantes. Assim, a demanda por uma forma de visualização, publicidade e conhecimento destas informações foi tratada por meio de uma ferramenta computacional que unifica e automatiza a atualização dessas informações.

---

<sup>6</sup> Disponível em: <http://site.stelaexperta.com.br>. Acesso em 29 de julho de 2021.



## 1.2 Justificativa

Considerando a necessidade de avaliação da produção técnico-científica de pesquisadores com intuito de fomentar a pesquisa na instituição, juntamente com a inviabilidade de análise manual dado o volume de informações e a multitude de atividades desenvolvidas como: artigos científicos publicados em periódicos, livros, capítulos de livros, projetos de pesquisa, patentes, *softwares*, entre outras. Requer-se a utilização de ferramentas computacionais para auxiliar a coleta de dados da Plataforma Lattes, e prover indicadores quantitativos e visualização da produtividade e desempenho de instituições e pesquisadores. Destaca-se ainda que a análise manual é uma tarefa demorada e altamente suscetível a erros (ALVES; YANASSE; SOMA, 2012; CESSO; FERRAZ, 2017).

O uso dessas ferramentas proporciona vários benefícios para os usuários, que vão desde a automatização, praticidade e maior velocidade na execução de tarefas até a redução de custos (PRESSMAN; MAXIM, 2016; SILVEIRA; NUNES, 2013). Nesse cenário, encontra-se uma gama de *softwares* disponíveis, a exemplo do ScriptLattes<sup>7</sup>, Stela Experta<sup>8</sup> e Intellectus<sup>9</sup>. O próprio CNPq disponibiliza às Instituições a ferramenta Lattes Extrator<sup>10</sup> que permite acesso aos dados da plataforma Lattes para que implementem ferramentas de análises de suas produções.

No entanto, algumas ferramentas não podem ser acessadas ou distribuídas livremente, outras carecem de alguma funcionalidade, como a geração de redes de colaboração, perfil do pesquisador ou indicadores específicos de uma instituição (MEGLIORANSI, 2019). Na Tabela 1 consta uma comparação de ferramentas disponíveis em correlação aos requisitos levantados como solução para o problema relatado.

Como em um cenário de contingenciamento de verbas torna inviável a aquisição de um novo sistema, surge uma demanda para o desenvolvimento de uma ferramenta unificada que automatize tarefas na instituição e facilite a visualização e o acompanhamento das produções técnico-científicas, a fim de manter a continuidade dos serviços. Este *software* tem como requisitos suprir a necessidade de indicadores específicos da UFOPA, integração com os sistemas institucionais, modelos de pontuação, publicização dos dados de pesquisa, além de dados pontuais para gestão da pesquisa na instituição, possibilitando análises de produção, redes de coautoria, perfis e tendências das pesquisas.

Este projeto encontra-se engajado com o PROFNIT na área de Gestão da inovação e Políticas Públicas, almejando maior agilidade e conexão de processos e dados institucionais e divulgação pública de produções técnico-científicas. Enquadrando-se em ações de Políticas Públicas em âmbito global por meio do objetivo 16.6 da Organizações das Nações Unidas - ONU

<sup>7</sup> Disponível em: <http://scriptlattes.sourceforge.net>. Acesso em 06 de julho de 2021.

<sup>8</sup> Disponível em: <http://site.stelaexperta.com.br>. Acesso em 06 de julho de 2021.

<sup>9</sup> Disponível em: <https://intellectus.ufrn.br>. Acesso em 06 de julho de 2021.

<sup>10</sup> Disponível em: <http://memoria.cnpq.br/web/portal-lattes/extracoes-de-dados>. Acesso em 06 de julho de 2021.

Tabela 1 – Levantamento sobre *softwares* de cientometria baseados na plataforma Lattes

Funcionalidades	SL	LL	LT	eL	SS	SE
Produções Bibliográficas por grupo	+	+	+	+	+	+
Indicadores do membro do grupo	—	—	+	—	+	+
Propriedade intelectual	+	+	+	—	+	+
Prêmios científicos/Artísticos	+	+	+	—	—	+
Gráficos quantitativos de produção	+	+	—	+	+	+
Grafos de Coautoria	+	+	+	+	—	—
Gráficos interativos	—	—	—	—	+	+
Livre Acesso	+	+	+	+	+	—

SL - ScriptLattes; LL - LucyLattes; LT - Lattices; eL - e-Lattes; SS - Sistema Somos; SE - Stela Experta.

Fonte: Elaboração própria.

que “Propõe o desenvolvimento de instituições eficazes, responsáveis e transparentes em todos os níveis”. Em âmbito nacional encontra-se enquadrado no Plano Plurianual 2020-2023<sup>11</sup> por meio do artigo 3º, inciso I e II, em que são definidas como diretrizes: o aprimoramento da governança, da modernização do Estado e da gestão pública federal, com eficiência administrativa, transparência da ação estatal, digitalização de serviços governamentais e promoção da produtividade da estrutura administrativa do Estado; e a busca contínua pelo aprimoramento da qualidade do gasto público, por meio da adoção de indicadores e metas que possibilitem a mensuração da eficácia das políticas públicas. Além de fatores relativos a inovação no setor público, em que há melhoria no atendimento sem ocorre aumento expressivo dos recursos públicos despendidos pelo poder público. (FRANÇA, 2017)

### 1.3 Objetivos

O objetivo geral deste trabalho é otimizar a gestão da pesquisa e inovação na UFOPA com a utilização de uma ferramenta computacional que colete, processe e disponibilize informações sobre as produções tecnocientíficas dos pesquisadores da instituição e permita a tomada decisões orientada a dados. Os objetivos específicos são:

1. Identificar os perfis e tendências de pesquisa da Instituição a partir de dados extraídos da plataforma Lattes, utilizando algoritmos de análise de mídias sociais e de cientometria;
2. Automatizar a obtenção dos dados de produção dos pesquisadores e suas relações de cooperação;

<sup>11</sup> Disponível em: <https://www.gov.br/economia/pt-br/assuntos/planejamento-e-orcamento/plano-plurianual-ppa/arquivos/Lein13.971de27dedezembrode2019.pdf>. Acesso em 29 de julho de 2021.

3. Prover informações e auxiliar no planejamento e na avaliação de políticas voltadas para pesquisa e inovação;
4. Fortalecer as conexões pesquisador-pesquisador e pesquisador-sociedade através da divulgação das atividades de pesquisa para a comunidade;
5. Difundir as produções técnico-científicas da Instituição, integrando em uma ferramenta *web* os artefatos desenvolvidos neste trabalho.

## 1.4 Impactos Esperados

Este trabalho visa primeiramente gerar impactos locais, no ambiente da UFOPA onde esta sendo desenvolvido e implantado o *software*. Nesse âmbito espera-se possibilitar uma melhor gestão da pesquisa da instituição, ao consolidar e difundir dados das produções de forma fácil e centralizada, possibilitando que gestores avaliem e tomem decisões e que pesquisadores e a sociedade de modo geral possam interagir e obter o conhecimento produzido pela Instituição.

Em consonância aos objetivos da ONU e ao Plano Plurianual já citados na Seção 1.2, juntamente com o artigo 70 da constituição federal<sup>12</sup>, almeja-se gerar inovação no serviço público, proporcionando economia e celeridade. Visa-se também fornecer um processo eficaz e transparente, além de indicadores que possibilitem a mensuração da eficácia de políticas públicas nas instituições de pesquisa, por meio da automatização do processo de coleta, quantificação, estratificação e divulgação das produções de cada instituição.

A possibilidade de melhorar a gestão é evidente após a identificação dos pontos fortes de pesquisa da instituição, permitindo maior controle para a gestão de recursos a grupos de pesquisa e editais de fomento à pesquisa e a busca de novos financiadores e parceiros.

Por fim, em consequência, em caso de sucesso deste artefato, pretende-se prover escalabilidade de seus benefícios objetivando alcançar outras instituições de Ciência, Tecnologia e Inovação (ICT's), conforme a lei de Inovação (10.973/2004), que venham a ter interesse no sistema a ser construído.

## 1.5 Produtos e Artefatos Desenvolvidos

Durante o desenvolvimento deste estudo foram produzidos artefatos como soluções, estudos e meios para sua conclusão. Esses são descritos abaixo:

1. SCIENSH: sistema de apoio à gestão da pesquisa e inovação baseado em cientometria;
2. Dipson: uma biblioteca escrita em python para buscas profundas de forma recursiva em estrutura Json;

---

<sup>12</sup> Constituição Federal. Disponível em: [http://www.planalto.gov.br/ccivil\\_03/constituicao/constituicao.htm](http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm). Acesso em 29 de julho de 2021.

3. Desenvolvimento de um artigo sobre a prospecção da utilização e divulgação de dados centimétricos pelas IFES brasileiras e as ferramentas disponíveis;<sup>13</sup>
4. Desenvolvimento desta dissertação como consolidação dos estudos e formalização deste Mestrado para a obtenção do Título;
5. Disponibilização de uma base de dados centimétricos das produções da UFOPA;
6. Disponibilização de manuais, desenvolvidos durante a disciplina de Oficina Profissional, para orientação de preenchimento das plataformas de pesquisas utilizadas nas coletas de dados e utilização do sistema desenvolvido;

## 1.6 Organização do documento

No Capítulo1 foi apresentado um contexto do projeto, os desafios de pesquisa, objetivos e impactos esperados. Os Materiais e Métodos, com ênfase no método de pesquisa e desenvolvimento, fontes de dados e tecnologias, foram apresentados no Capítulo 2. Os resultados foram apresentados e discutidos no Capítulo 3, tendo sido organizado de acordo com os desafios de pesquisa. Por fim, as conclusões são apresentadas no Capítulo 4, bem como as dificuldades encontradas no decorrer do projeto, sugestões de trabalhos futuros e ameaças a validade da pesquisa.

---

<sup>13</sup> Prospecção de ferramentas para gestão de pesquisa e desenvolvimento em instituições de ciência e tecnologia. Disponível em: <http://etech.sc.senai.br/edicao01/article/view/1173/76>

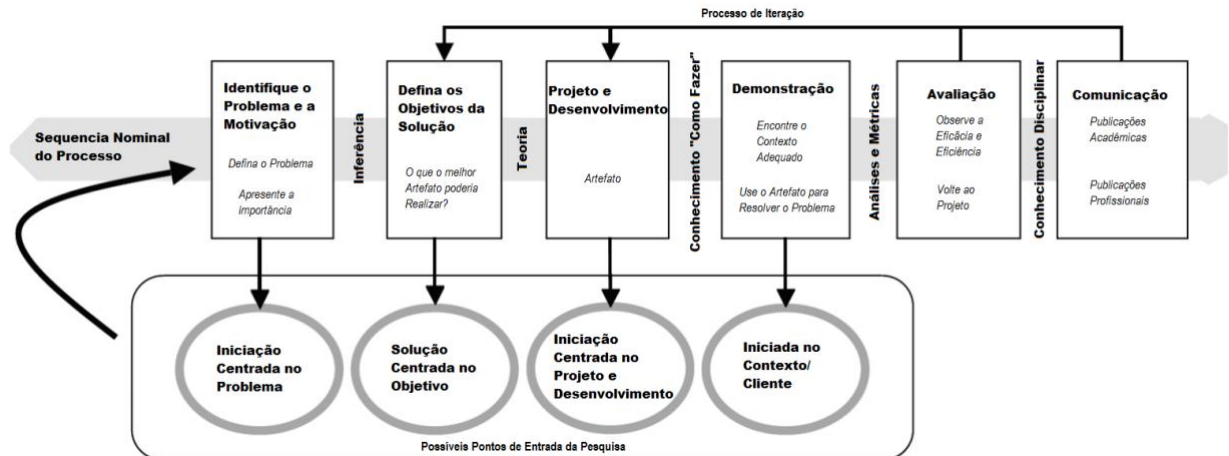
## 2 MATERIAIS E MÉTODOS

Neste capítulo descreve-se a forma como o artefato foi desenvolvido, justificando a escolha do método de pesquisa e desenvolvimento utilizado, descrevendo-o detalhadamente e os passos executados juntamente com a expectativa de resultados de cada passo. Também descreve-se o método de coleta, escolha e justificativa das fontes de dados.

### 2.1 Método de Pesquisa

Este trabalho possui caráter exploratório e descritivo e trata-se de uma pesquisa aplicada. Seu objetivo está menos voltado para o desenvolvimento de teorias e mais voltado para a aplicação imediata numa realidade circunstancial (GIL, 2008). Foi utilizado o método *Design Science Research* (DSR), que consiste na construção de conhecimento científico voltado à inovação por meio de artefatos, com a intenção de resolver problemas reais, com contribuição científica de caráter prescritivo (DRESCH; LACERDA; JÚNIOR, 2015). As fases deste método serão baseadas na definição utilizada por (PEFFERS *et al.*, 2007), tal como disposto na Figura 1.

Figura 1 – Etapas do DSR.



Fonte: Adaptado de Peffers *et al.* (2007).

Conforme visto na Figura 1, o DSR é um processo iterativo composto por seis etapas, a saber: (i) Identificação do problema e motivação de sua solução; (ii) Definição dos objetivos da solução; (iii) Projeto e desenvolvimento do artefato; (iv) Demonstração do artefato quanto a resolução de uma ou mais instâncias dos problemas levantados; (v) Avaliação da eficiência da solução na resolução dos problemas; e, por fim, (vi) Comunicação dos resultados técnico-científicos alcançados. Importante notar que este é um processo iterativo e que permite início em uma das quatro primeiras fases.

Para o presente estudo, o início foi centrado no problema que foi apresentado pela Diretoria de Pesquisa. A fase inicial consiste no entendimento de negócio e visa a identificação e a motivação do problema, juntamente com setor responsável pela pesquisa da UFOPA. Considerando que eram vários desafios apresentados, selecionaram-se os mais relevantes, tal como destacado na Seção 1.1. A partir do estudo da relevância e viabilidade, dois foram selecionados e passou-se para a definição dos objetivos, já apresentados na Seção 1.3. Importante notar que a definição dos objetivos foi realizada conjuntamente com uma revisão literária a fim de investigar técnicas e procedimentos para geração de indicadores de produção.

Na fase de projeto e desenvolvimento, definiu-se que o artefato a ser desenvolvido será a implementação de uma ferramenta computacional para coletar e permitir analisar os dados das produções e geração dos indicadores.

Após a implementação do artefato, ocorreu a demonstração do artefato como solução para as classes de problemas levantadas. Nesta etapa ocorreu a implantação do sistema para testes e correção de eventuais falhas na implementação. A avaliação da solução foi executada por meio de análises qualitativas e quantitativas através do desenvolvimento e aplicação de questionários de satisfação aos *stakeholders* envolvidos. Em conclusão, a fase final consiste na comunicação contendo desenvolvimento da dissertação com a apresentação da solução desenvolvida e a análise dos dados gerados pela ferramenta para identificar os perfis e tendências de pesquisa, além da divulgação nos veículos de comunicação institucionais para a ciência da comunidade. Esta etapa também foi realizada por meio da publicação do artigo em periódico e no registro de programa de computador junto ao Instituto Nacional de Propriedade Industrial (INPI).

## 2.2 Materiais

Nesta Seção são apresentados os materiais utilizados para o desenvolvimento do artefato, mais especificamente, as fontes de dados e tecnologias utilizados. A etapa de coleta de dados permite a seleção das informações relevantes para a formação do banco de fonte de dados necessária para o processamento das análises e visualizações propostas neste estudo.

### 2.2.1 Plataforma Lattes

A Plataforma Lattes foi disponibilizada pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) com a finalidade de unificar o cadastro nacional de pesquisadores. Em 2012, a Plataforma Lattes contava com mais de dois milhões de currículos cadastrados (ALVES; YANASSE; SOMA, 2012). É a plataforma com maior e mais relevante repositório curricular de pesquisadores acadêmicos da Ciência Brasileira (LANE, 2010).

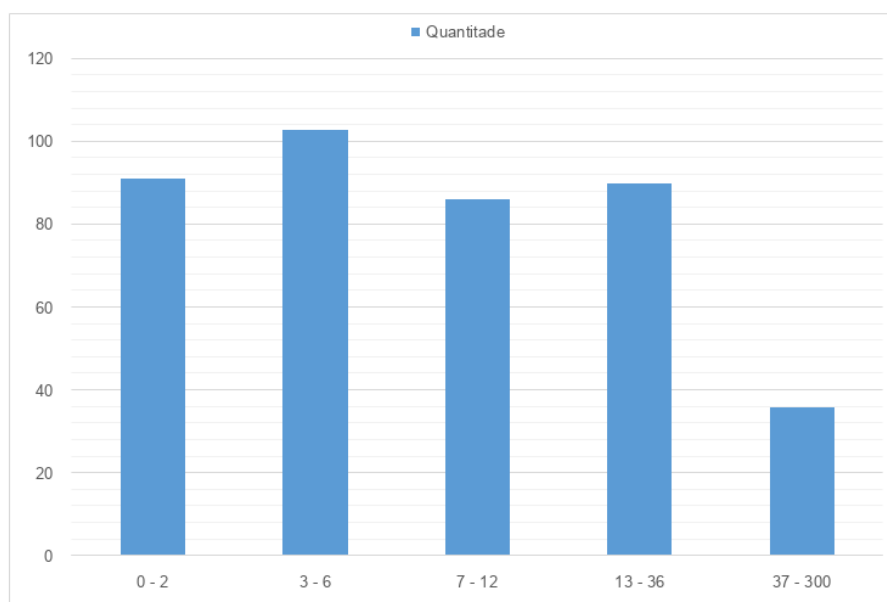
Plataformas com as mesmas características não são encontradas em outras partes do mundo, o que se tem disponível são bases de dados referenciais, que permitem extração de fontes de informação secundárias. O Brasil é o único país que vincula obrigatoriamente as informações

constantes em uma plataforma curricular com as solicitações de fomento aos órgãos competentes. Para tal, além da concessão de bolsas e auxílios à pesquisa, há também a avaliação da qualidade dos cursos de Pós-Graduação Stricto Sensu. Embora a plataforma seja usada por outros países, como Portugal e países da América latina, ainda não se tem informações sobre sua utilização e obrigatoriedade para decisões de fomento (MACCARI; KNISS; FERRAZ, 2019).

A avaliação dos currículos cadastrados na Plataforma Lattes permite a obtenção de uma série de informações que não estão disponíveis nas bases de caráter referencial, como, por exemplo, artigos que já estão aceitos para publicação, projetos de pesquisa submetidos, orientações de mestrado e doutorado, supervisões de pós-doutorado em andamento, publicações de matérias em revistas e jornais de alcance popular, técnicas, produtos ou processos tecnológicos desenvolvidos, dentre outras modalidades de produção, bem como para avaliar as redes de relação existentes entre os pesquisadores (BALANCIERI *et al.*, 2005) e redes de colaboração em produção científica (MARTINS *et al.*, 2014). A coleta de dados em larga escala da plataforma Lattes foi restringida pelo CNPq, permitindo somente esta funcionalidade a instituições e pesquisadores que solicitem autorização especial.

Devido ser uma plataforma de inserção obrigatória para pesquisadores que necessitem participar de editais de fomento à pesquisa, se torna uma rica fonte de informações unificadas que viabiliza o desenvolvimento deste projeto, permitindo análises de dados e automatização de processos, uma vez que a instituição conquistou permissão do CNPq para coletar dados da Plataforma Lattes por meio de seu *web service* (LEITE; MUGNAINI; LETA, 2011; MUGNAINI; LEITE; LETA, 2011). Após efetuar a coleta de dados, foi produzido um histograma, conforme disposto na Figura 2, baseado na última data de atualização dos currículos dos docentes da Instituição.

Figura 2 – Histograma da última atualização dos currículos Lattes dos pesquisadores da UFOPA.



Fonte: Elaborado pelo autor.

Conforme visto na Figura 2, observa-se que a grande maioria dos docentes atualizou seus currículos em menos de 12 meses. Apesar de sua completude, alguns dados não estão disponíveis no Lattes, como o número de citações, por este motivo, buscou-se no *Google Scholar* uma alternativa, tal como descrito a seguir.

### 2.2.2 Google Scholar

Com o surgimento de novos índices de citação quebra-se o que parecia monopólio absoluto, desde a década de 1960, do *Institute of Scientific Information* (ISI), atualmente Thomson/ISI (CAREGNATO, 2011). Da mesma forma, houve um crescimento da aplicação dos métodos e técnicas bibliométricas no estudo da ciência para analisar seu impacto e produtividade por meio de índices de citações. Até recentemente, para calcular o impacto das citações era necessário utilizar a base ISI, comercializada como *Web of Science* (WOS). Porém, com a crescente disponibilidade de publicações científicas em formato eletrônico disponível na *web* novas plataformas começaram a ser disponibilizadas. A WOS é ainda a fonte mais utilizada e conhecida para a análise de citações, embora receba também muitas críticas quanto ao reduzido número de títulos indexados (CAREGNATO, 2011).

O Google Scholar, diferentemente da base de dados multidisciplinares WOS, é uma ferramenta gratuita, que permite localizar trabalhos acadêmicos de vários tipos (por exemplo, artigos de congressos, teses e dissertações, além de artigos de periódicos de acesso aberto ou pagos), em múltiplas línguas (inclusive português), disponibilizadas em repositórios na *web* ou sites acadêmicos, além de determinar a frequência com que foram citados em outras publicações acadêmicas (CAREGNATO, 2011).



Essa plataforma surgiu na *web* no final de 2004 com o propósito de ser uma ferramenta de busca de informações de caráter acadêmico. No entanto, por incorporar dados extraídos das referências bibliográficas dos documentos, passou a trazer informações relativas ao impacto dos trabalhos medido pelo número de citações recebidas (CAREGNATO, 2011). Assim, tem possibilidade de ser utilizado como referência principal em índices de citações para estudos bibliométricos (NORUZI, 2005), em mensuração do impacto das pesquisas (VAUGHAN; SHAW, 2008) e periódicos de acesso aberto (MUKHERJEE, 2009).

Seu processo de coleta de dados é feito por meio de um *software* que rastreia a *web* e reconhece automaticamente os campos que compõem os documentos científicos e suas referências. Com isso, surge uma ferramenta sem precedentes provendo acesso gratuito a dados bibliográficos e a milhões de artigos na íntegra (JACSÓ, 2010).

Frente ao exposto, optou-se por utilizar dados provenientes do *Google Scholar* para obter índices de impacto e citações das pesquisas dos docentes da instituição. Permitindo assim um gerenciamento e visualização de forma rápida e unificada das produções e pesquisadores mais relevantes na instituição. Para identificar os pesquisadores tanto no Lattes quanto no Google Scholar, utilizam-se os dados armazenados nos sistemas institucionais. Esta integração é abordada a seguir.

### 2.2.3 Sistemas Institucionais

Um dos diferenciais deste trabalho é a obtenção de uma solução que consiga integrar dados institucionais, objetivando automatizar a sincronização de dados e fazer jus aos recursos já investidos. Assim, alguns dados que não se pode obter pelas fontes anteriormente citadas devem ser obtidas por meio dos sistemas institucionais. A UFOPA utiliza um sistema acadêmico bem difundido pelas IFES brasileiras, no caso o sistema SIGAA da instituição UFRN. Segundo site da UFRN<sup>1</sup>, 73 instituições utilizam seus sistemas.

A base de dados do SIGAA fornece dados sobre os docentes, servidores técnicos e discentes da instituição, suas unidades de lotação, inícios e termos de vínculo, facilitando assim a alimentação de maneira automática da ferramenta proposta. O SIGAA permite, por meio do módulo pesquisa, o cadastro e gerenciamento de editais de pesquisa. Assim, a coleta de dados buscará os inscritos nestes editais para produzir resultados de pontuações.

É por meio do Sistema Intellectus<sup>2</sup>, sistema da UFRN, que se obtém os dados dos currículos Lattes importados dos servidores do CNPq. Este sistema conecta à base de dados do SIGAA e busca os docentes da instituição para poder obter seus currículos.

<sup>1</sup> Disponível em: <https://portalcooperacao.info.ufrn.br/pagina.php?a=parceiros>. Acesso em 29 de julho de 2021.

<sup>2</sup> Disponível em: <https://sistemas.ufrn.br/>. Acesso em 29 de julho de 2021.

## 2.3 Tecnologias

Primeiramente a escolha da linguagem de programação para o desenvolvimento do projeto foi Python<sup>3</sup>, baseou-se na escolha de outros projetos disponíveis, como o *ScriptLattes*, visando facilitar a integração entre as ferramentas e aprendizado (MENA-CHALCO; JÚNIOR, 2013). Além de ser uma linguagem bastante utilizada no meio científico, com muitas bibliotecas especializadas em análises de dados e aprendizado de máquina, o que expande o horizonte para trabalhos futuros. Uma dessas bibliotecas de análises de dados utilizada neste projeto para o processamento dos dados foi a biblioteca Pandas<sup>4</sup>.

O artefato faz comunicação com o sistema gerenciador de banco de dados Postgres<sup>5</sup> para sincronização com dados institucionais do Sistemas Sigs. Mas para armazenamento de dados de sua base própria foi escolhido o SGBD MongoDB<sup>6</sup>, pois apresenta vantagens de escalabilidade, manipulação da quantidade massiva dos dados, flexibilidade e facilidade para consultas (SOUZA; OLIVEIRA, 2019).

Para a entrega dos dados por meio de páginas *web*, o *framework* escolhido foi o FastAPI<sup>7</sup>. Essa escolha se definiu devido tecnologias atualizadas deste *framework*, oferecendo ferramentas que agilizam o desenvolvimento e melhoram a manutenção e desempenho, como documentação automática através do Swagger<sup>8</sup>, funções assíncronas, tipagem de dados com retornos amigáveis para validações.

Consoante a tecnologia padrão do FastAPI, MongoDB e os formatos de dados do Intellectus, utilizou-se o formato JSON para armazenar e distribuir a base de dados. A tecnologia *JavaScript Object Notation* (JSON) é uma sintaxe, baseada em Javascript, para serialização de objetos, matrizes, números, *strings*, booleanos, e *NULL*<sup>9</sup>. Para facilitar a disponibilização da base desenvolvida pelo artefato, foi implementado através do FastAPI, *web services* com rotas para acesso a projetos, artigos, pesquisadores entre outros por meio da especificação OpenAPI<sup>10</sup>. Juntamente com o Swagger permitem fácil visualização e testes para conexão do artefato com outros sistemas.

<sup>3</sup> Disponível em: <https://www.python.org/>. Acesso em 29 de julho de 2021.

<sup>4</sup> Disponível em: <https://pandas.pydata.org/>. Acesso em 29 de julho de 2021.

<sup>5</sup> Disponível em: <https://www.postgresql.org/>. Acesso em 29 de julho de 2021.

<sup>6</sup> Disponível em: <https://docs.mongodb.com/manual/>. Acesso em 29 de julho de 2021.

<sup>7</sup> Disponível em: <https://fastapi.tiangolo.com/>. Acesso em 29 de julho de 2021.

<sup>8</sup> Disponível em: <https://swagger.io/>. Acesso em 29 de julho de 2021.

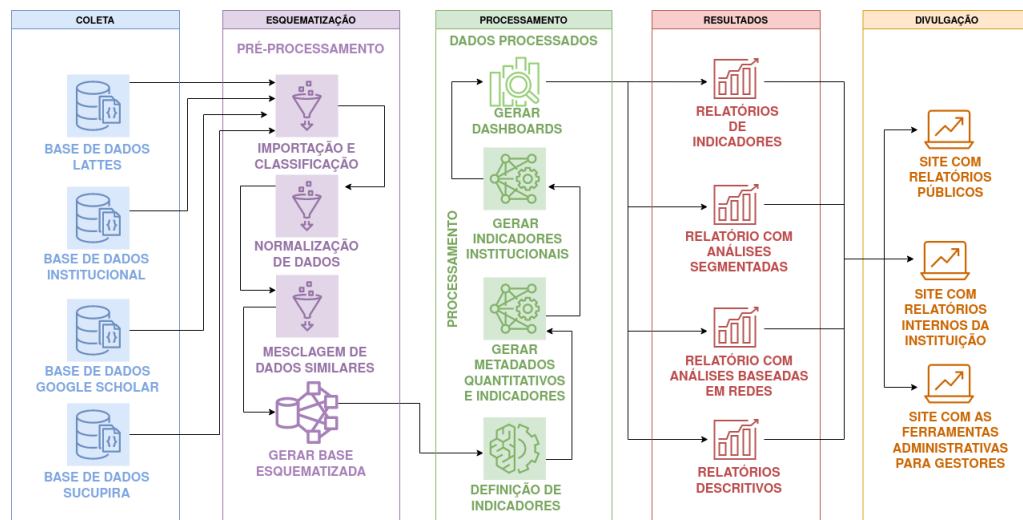
<sup>9</sup> Disponível em: [https://developer.mozilla.org/pt-BR/docs/Web/JavaScript/Reference/Global\\_Objects/JSON](https://developer.mozilla.org/pt-BR/docs/Web/JavaScript/Reference/Global_Objects/JSON). Acesso em 29 de julho de 2021.

<sup>10</sup> Disponível em: <https://github.com/OAI/OpenAPI-Specification>. Acesso em 29 de julho de 2021.

### 3 RESULTADOS

A partir dos materiais e métodos apresentados, consoante aos objetivos definidos, alguns artefatos foram construídos. Neste capítulo são apresentados os resultados obtidos, chegando até a avaliação da solução. Tais artefatos culminaram no amadurecimento da estrutura do sistema, o qual tem seu fluxo apresentado na Figura 3.

Figura 3 – Fluxo do *software* a ser desenvolvido.



Fonte: Elaborado pelo autor.

Neste fluxo podemos observar a coleta de dados obtida de diferente base de dados, como plataforma lattes, sistemas institucionais, *Google Scholar* e plataforma Sucupira. Após a coleta são efetuados processos de classificação e normalização dos dados para obter uma fusão das informações para criação de uma nova base esquematizada. Os artefatos se baseiam nesta última etapa para processar os dados quantitativos e descritivos sobre as produções. Assim é possível identificar e quantificar indicadores, dados quantitativos e descritivos sobre pesquisadores e instituições, para então gerar os *dashboards* para expor e gerir esses dados, permitindo gerar relatórios. A etapa de divulgação foca na publicidade das informações por meio de páginas na internet e intranet.

#### 3.1 Artefatos Construídos

Esta Seção consiste na apresentação dos artefatos e demonstração dos problemas solucionados, em que se descreve a origem da demanda, o levantamento e a identificação das classes de problema.

### 3.1.1 Desafio 1: Edital PIBIC

Este artefato surge com uma demanda espontânea da instituição em levantar dados de forma ágil e eficiente sobre as produções desenvolvidas por seus pesquisadores. O problema inicial foi a dificuldade em se quantificar um índice de produtividade para distribuição de bolsas de pesquisa por meio de editais de fomento interno. Tal qual outras IFES, para distribuir recursos de pesquisas, a universidade estabeleceu um índice de produtividade dos pesquisadores, em que tipos específicos de produção têm pontuações específicas e ainda, classes de produtos têm limitação de pontuação, conforme Anexo A.

As categorias de produção incluem: (i) artigos completos publicados segmentados por periódicos com maior Qualis; (ii) livros e capítulos de livros produzidos ou organizados segmentados por nível de edição internacional ou nacional; (iii) trabalhos publicados em eventos segmentados por evento internacional ou nacional; (iv) trabalhos técnicos segmentados por tipo: registro de patente, registro de *software*, processos e design de produto; (v) orientações de graduação e *Stricto Sensu* segmentadas por nível de ensino; e (vi) outras produções como por exemplo prêmios e coordenação de projetos de pesquisa.

Em editais anteriores, este cômputo era realizado de forma manual, em que os docentes se inscreviam nos editais de pesquisa por meio do SIGAA e enviavam planilhas informando a pontuação individual, calculada por eles mesmos, contendo a quantificação segmentada de sua produção. Após isso, três servidores da Diretoria de Pesquisa, em um prazo de 2 semanas, conferiam as pontuações solicitadas item por item, buscando na plataforma Lattes de cada pesquisador e cruzando os dados com a plataforma Sucupira, onde se encontra o extrato Qualis de cada periódico.

Este processo gerava gasto de recursos para quantificar e ranquear os pesquisadores, tanto dos docentes que precisavam produzir uma planilha preliminar quanto dos técnicos administrativos que precisavam validá-las, um processo moroso e propenso a erros.

Conforme mencionado no Capítulo 2, a UFOPA tem disponível o sistema *Intellectus* que permite criar modelos de pontuação com base nos dados da plataforma Lattes, porém estes modelos não permitem a implantação de algumas regras específicas da instituição, como por exemplo a limitação de pontos por seção. Dessa forma, identificou-se a necessidade de automatizar este processo, modularizando o método de quantificação de produções, de uma forma genérica, capaz de calcular com base em atributos de cada produção e permitir controles de limites de pontos e unidades por Seção e tipo de produção. A posposta de solução para este problema será apresentada na Seção 3.2.1.

### 3.1.2 Desafio 2: Visualização dos dados

Após o sucesso na automatização do desafio mencionado na Subseção anterior, os gestores do setor apresentaram uma segunda demanda, que diz respeito à dificuldade de mapeamento

das produções feitas pelos pesquisadores da universidade. Esse mapeamento visa quantificar e identificar as áreas de pesquisa das produções da universidade, descobrindo tendências, tópicos pesquisados, produções relevantes, pesquisadores relevantes, conexões entre pesquisadores e projetos desenvolvidos, por exemplo. Além de publicizar esses dados para a instituição e para a sociedade.

O desafio nesta demanda dá-se devido à multitude de produções e ao elevado número de pesquisadores da instituição. O que acarreta em muito trabalho para unir e manter os dados sobre o tema atualizados.

Essas informações são necessárias para: (i) avaliar a produção técnico-científica; (ii) a produtividade de pesquisadores, grupos de pesquisa e institutos; (iii) avaliar e gerenciar políticas de fomento de pesquisa, permitindo definir novas estratégias baseando-se nos resultados visualizados. Além de que conhecer as características de pesquisa da instituição possibilita maior conexão intra e extra instituição, permitindo mais conexões entre os pesquisadores, entidades de financiamento e sociedade.

Importante destacar que a demanda por transparência de dados e publicidade das produções técnico-científicas não é restrita à UFOPA, uma vez que a gestão da pesquisa e a transparência destes dados é uma necessidade de qualquer instituição de pesquisa. Portanto uma solução para este problema se torna escalável e viável para muitas instituições de ciência e tecnologia do país.

## 3.2 Objetivos das soluções

Nesta Seção serão apresentadas as propostas de solução para os problemas identificados na Seção anterior, descrevendo os requisitos necessários para a efetividade no processo e como serão implementados.

### 3.2.1 Desafio 1: Edital PIBIC

Dado o primeiro desafio discutido no Capítulo 1, buscou-se uma forma de tornar eficiente e viável o processo de coleta, quantificação de índices personalizáveis e ranqueamento automático dos docentes baseados em suas produções.

A proposta consistiu na implementação de um *software* que se integrasse com os editais cadastrados no SIGAA e buscasse os docentes participantes quantificando suas produções conforme um modelo de pontuação previamente cadastrado no próprio *software*. Este modelo precisou ser genérico, permitindo buscas recursivas em itens chave-valor do currículo Lattes e definição de condições para encontrar itens e contabilizá-los. Além de permitir limites de pontuação e quantidade de itens por tipo e grupos.

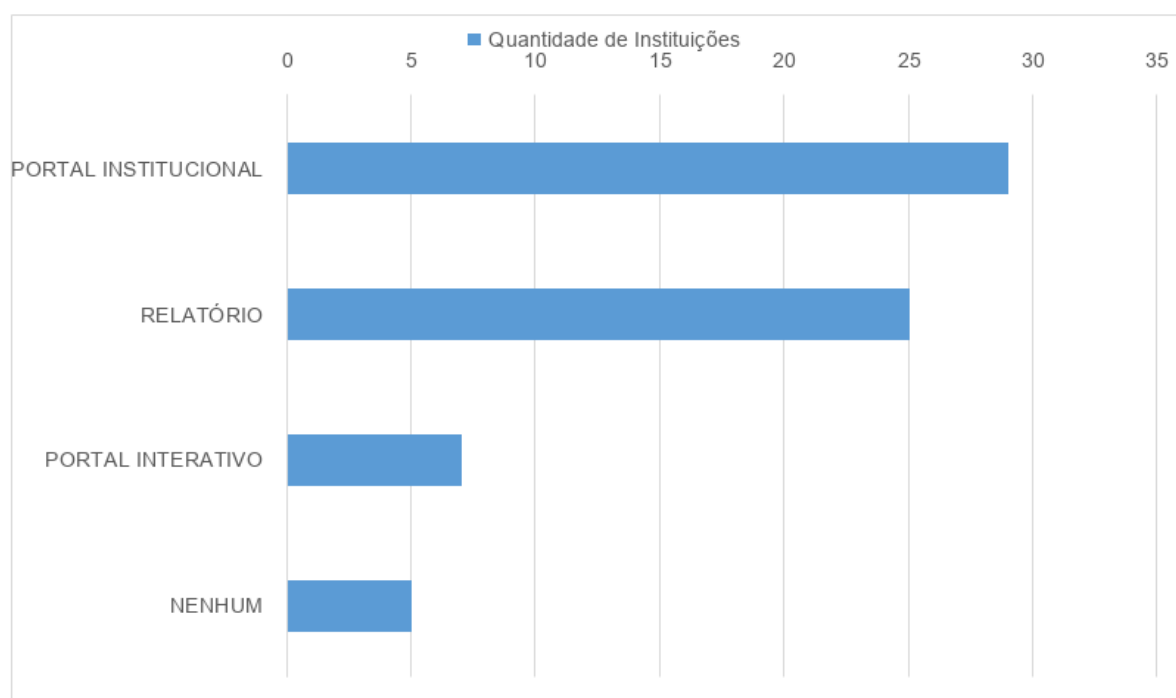
Identificou-se como necessidade a disponibilização de resultados por página *web*, que ordene os concorrentes do edital por pontuação e forneça a pontuação detalhada, identificando

os itens contabilizados – Permitindo uma maior transparência dos resultados dos editais. Essa solução visa melhorar o gerenciamento de índices personalizados com base na produção dos pesquisadores e reduzir o tempo para apresentar os resultados dos editais. Na Seção 3.4.1 é abordada em detalhes este desafio, mostrando os itens avaliados nos editais de pesquisa da instituição e os valores para cada tipo de produção.

### 3.2.2 Desafio 2: Visualização dos dados

Em relação ao segundo desafio apresentado no Capítulo 1, iniciou-se um estudo sobre como as universidades federais do Brasil divulgam dados sobre suas pesquisas. Assim, foi efetuado um levantamento dos meios de divulgação e os tipos de informações divulgadas por outras universidades, conforme Apêndice A . Essa pesquisa baseou-se em buscas nos sites institucionais e pró-reitorias de pesquisa das universidades e pesquisas no Google, pelas palavras-chaves: (i) anuário, (ii) indicadores, (iii) indicadores de pesquisa e (iv) indicadores de produção. Por meio da análise dos resultados das buscas, observou-se que a maioria das instituições divulgam seus dados de pesquisa por meio de relatórios e portais, tal como disposto na Figura 4.

Figura 4 – Infográfico dos meios de divulgação e tipos de afirmações de outras IFES.



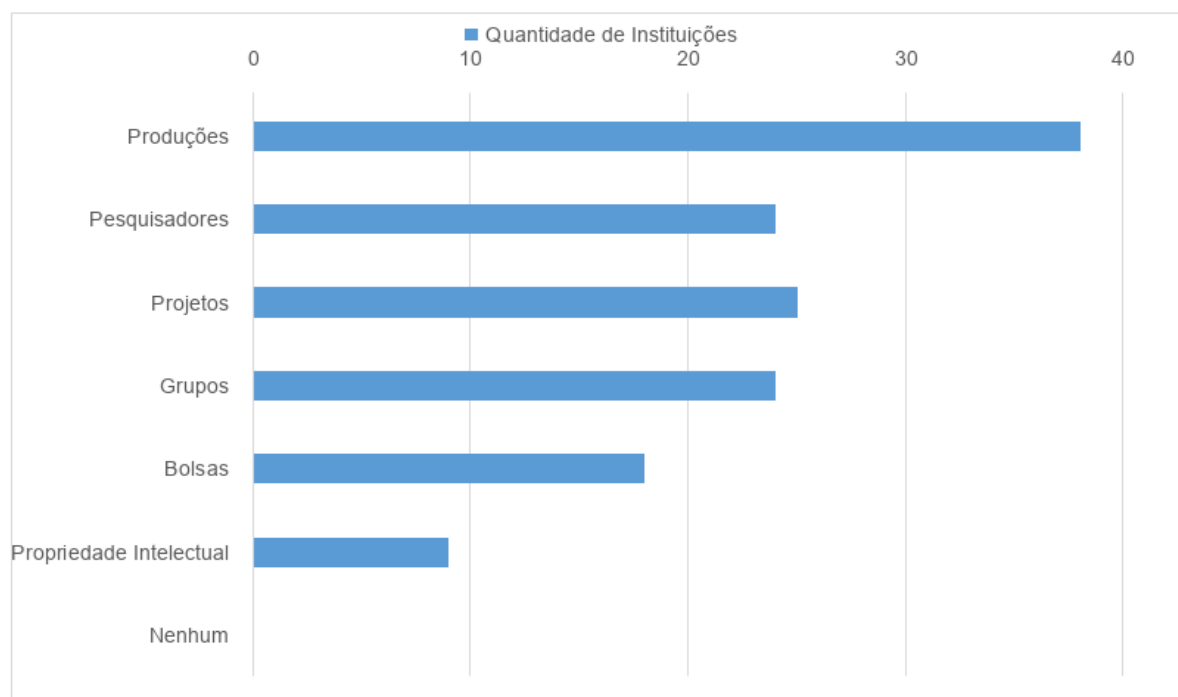
Fonte: Elaborado pelo autor

Também foi levantado um estudo sobre os tipos de informações que essas IFES divulgam, para identificar os itens publicizados e que demonstram importância, conforme Figura 5.

Por meio da análise crítica dos resultados dispostos na Figura 4, percebe-se que os meios de divulgação empregados dificultam a visualização por serem estáticos. Consequentemente, tal

formato não permite filtros de informações e atualização dos dados para maior transparência e interatividade. Algumas poucas universidades utilizam portais interativos que solucionam, parcialmente, estas últimas questões apresentadas.

Figura 5 – Infográfico dos tipos de informações divulgados pelas IFES.



Fonte: Elaborado pelo autor

A partir desta constatação, executou-se uma pesquisa por soluções disponíveis que abordam as necessidades identificadas. Esse resultado pode ser observado na Tabela 1, apresentada no Capítulo 1, em que constatou-se uma carência quanto a ferramentas de código-livre ou gratuitas para obtenção e análise de dados advindos da Plataforma Lattes. Além disso, as ferramentas proprietárias analisadas não contemplavam algumas produções importantes, como prêmios, títulos e grafos de coautorias.

Assim, se propôs o desenvolvimento de uma ferramenta interativa, denominada de Sciensh, capaz de auxiliar na avaliação da pesquisa institucional, abalizando o processo de tomada de decisão, com acesso simples e flexível à produção técnico-científica institucional. Baseada nos dados da plataforma Lattes, *Google Scholar*, plataforma Sucupira e sistemas institucionais utiliza algoritmos de fusão de dados para unificar essas bases e exibir os dados por meio de *dashboards* objetivos e atualizados. Além de permitir compartilhar essa nova base para outras aplicações institucionais.

### 3.3 Projeto e Desenvolvimento dos artefatos

Esta Seção busca descrever o processo de desenvolvimento dos artefatos de *software*. Define-se nesta etapa as decisões para tornar viável a entrega de um protótipo do *software* como solução proposta. Após a definição das tecnologias utilizadas, o passo seguinte foi o entendimento do negócio e das dificuldades já existentes. Para começar o desenvolvimento foi necessário fazer um mapeamento dos itens do currículo Lattes. A primeira coisa a ser feita foi a importação dos currículos no formato JSON e armazená-los no banco MongoDB. Observou-se uma grande dificuldade para o mapeamento e manutenção dos caminhos específicos de cada item na estrutura de arquivo. Então foi desenvolvida um algoritmo com buscas recursivas em um JSON, que agilizou e viabilizou o desenvolvimento da obtenção dos itens nos currículos.

Assim foi possível cadastrar vários modelos de pontuação, de forma genérica, permitindo buscas de itens por chave-valor e comparando-os com as condições especificadas no modelo, para então contabilizá-los. Os dados do modelo são cadastrados no banco e uma lógica de negócio implementa a validação do modelo com o currículo.

Para o Desafio 1, bastou analisar os currículos individualmente. Porém, para o segundo problema uma grande dificuldade foi encontrada, a duplicidade de dados. O primeiro problema percebido para obter uma base confiável de dados foi a duplicidade das produções, pois a base Lattes é uma base de currículos, ou seja, os registros são focados na individualidade de cada autor. Assim, registros de coautorias dentro da instituição são coletados uma vez para cada autor, gerando redundância. Então, desenvolveu-se um algoritmo para o tratamento de redundâncias que permite a detecção de produções acadêmicas iguais ou similares, conforme Algoritmo 1. Por uma questão de performance decidiu-se comparar apenas os primeiros trinta caracteres para depois comparar o texto completo.

Outro problema encontrado foi em relação a não completude de informações nos currículos, uma vez que nem todos os campos do Lattes são obrigatórios e não se tem um identificador obrigatório para cada item, o que dificulta a fusão dos dados entre currículos. Essas informações registradas não mantêm um padrão e, em muitos casos, estão incompletas ou em campos trocados. Por exemplo registros com nomes de autores abreviados, ausência do identificador, títulos com atributos de formatação e ausência de acentos. Assim, a obtenção de uma base unificada representa um problema não-trivial para uma melhor análise das produções.

Nesse contexto, optou-se pela comparação dos nomes das produções, porém devido a inconsistências, como erros de digitação ou falta de padronização na escrita de cada currículo, não é eficaz comparar textos diretamente. Para isso, a solução adotada foi a busca de proximidade de cadeias de caracteres. Para a definição de similaridade entre registros foi escolhido o algoritmo de distância de *Levenshtein*. Esse método compara duas cadeias de caracteres e retorna um valor correspondente à quantidade de alterações para um texto se transformar em outro, como inserções, eliminações ou substituições de caracteres (MENA-CHALCO; JÚNIOR, 2013).



---

**Algoritmo 1: ALGORITMO UTILIZADO PARA IDENTIFICAR PRODUÇÕES SIMILARES**


---

**Entrada:** itemBuscado, listaDeProducoes, atributo**Saída:** itemsSimilares

```

1 início
2   itemsSimilares = []
3   para cada indice, itemComparado em listaDeProducoes faça
4     se itemComparado['nomeAutorAscii'] == itemBuscado['nomeAutorAscii']
5       então
6         continuar
7     fim
8     se itemComparado não contém atributo então
9       continuar
10    fim
11    se ratio(itemComparado[atributo][:30],itemBuscado[atributo][:30]) > 0.8
12      então
13        se ratio(itemComparado[atributo],itemBuscado[atributo]) > 0.90 então
14          adiciona itemComparado a itemsSimilares
15        fim
16    fim
17 retorna itemsSimilares
18 fim

```

---

Assim, dois itens são considerados iguais se possuírem um coeficiente de similaridade superior a 90%, diferentemente dos 80% utilizados por Mena-Chalco e Júnior (2013). Esta decisão foi baseada em uma avaliação qualitativa com servidores da DPE. A Tabela 2 foi construída com base nos currículos coletados, e possibilita observar a quantidade de registros encontrados por tipo de produção e a quantidade obtida após fusão de dados, juntamente com o percentual de itens similares obtido. Por fim, este procedimento executa comparações dois a dois entre todas as produções de um conjunto de dados segmentado por tipo de produção.

Para melhorar os resultados foram aplicadas técnicas para normalizar textos antes de compará-los. Os procedimentos adotados para o pré-processamento foram: (i) conversão para caixa baixa; (ii) remoção de espaços em branco extras; (iii) remoção de *tags* Linguagem de Marcação de Hipertexto, do inglês, *HyperText Markup Language* (HTML); (iv) remoção de elementos de formatação; e (v) remoção de caracteres não padronizados de acordo com o *American Standard Code for Information Interchange* (ASCII).

Um dos benefícios do uso desta estratégia é a geração de uma base de dados mais completa após fusão de registros de produções similares (vinda de currículos diferentes), pois as informações de registros similares são mescladas, assim informações existentes em um registro complementam as que faltavam em outros. Para casos em que ocorreram divergências de informações foi implementada e aplicada uma função, conforme Algoritmo 2 para identificar

Tabela 2 – Quantitativo pós-fusão de dados

Tipo de produção	Registros	Únicos	Similares (%)
Artigos Completos	5.223	4.554	12,80
Trabalhos em eventos	13.225	12.214	7,64
Livros	455	399	12,30
Capítulos de livro	1.464	1.337	8,67
Patentes	29	25	13,79
Softwares	86	84	2,32
Projetos	4.467	3.770	15,60
Projetos de Pesquisa	3.005	2.608	13,21
Projetos de Extensão	1.061	820	22,71
Projetos de Desenvolvimento	93	73	21,50
Projetos de Ensino	62	57	8,06

Fonte: Elaborado pelo autor.

o autor principal ou responsável pela produção, e então utiliza-se o Algoritmo 3 para mesclar os dados dos registros dando prioridade para o registro do autor principal, caso seja docente da instituição, caso contrário utiliza-se como referência o primeiro registro de um autor da instituição.

---

**Algoritmo 2:** ALGORITMO UTILIZADO PARA IDENTIFICAR RESPONSÁVEL PELA PRODUÇÃO.

---

```

Entrada: dados
Saída: Autor principal da produção
1  início
2      autorPrincipal = vazio
3      para cada dado em dados faça
4          se dado contém 'participantes' então
5              para cada participante em dado faça
6                  se participante contém 'ordemdeautoria' então
7                      se participante['ordemdeautoria'] == 1 então
8                          autorPrincipal = participante
9                          break
10                     fim
11                 fim
12             senão
13                 autorPrincipal = participante
14                 break
15             fim
16         fim
17     fim
18 fim
19 retorna autorPrincipal
20 fim

```

---

---

**Algoritmo 3: ALGORITMO UTILIZADO PARA FUSÃO DE DADOS DE REGISTROS SIMILARES.**


---

**Entrada:** producoesSimilares, curriculos  
**Saída:** dados de produção mesclado

```

1 início
2   se existem 'participantes' em producoesSimilares então
3     registroMesclado = vazio responsavel = obterAutorPrincipal(dados)
4     indiceRegistroResponsavel = vazio se responsavel não é vazio então
5       identificar o indice do responsavel na lista producoesSimilares se
6         indiceRegistroResponsavel não é vazio então
7           registroMesclado = producoesSimilares[indiceRegistroResponsavel]
8           remove registro de producoesSimilares
9         fim
10      fim
11     para cada dado em dados faça
12       mesclar objetos dado e registroMesclado
13     fim
14   retorna registroMesclado
15 fim

```

---

A construção dos *dashboards* foi baseada em boas práticas para mostrar conhecimento por meio de dados, fornecendo visualizações eficientes para a tomada de decisões, conforme Knaflitz (2019) orienta. Algumas destas orientações foram aplicadas para entregar conhecimento por meio de infográficos simples e objetivos, evitando problemas comuns que dificultam a transmissão da mensagem. Por exemplo, evitar gráficos em formato de pizza quando existem muitos dados ou pouca diferença entre os valores, evitar gráficos 3D caso a terceira dimensão não apresente uma nova informação e eliminar elementos apenas visuais que não colaboram com a passagem da mensagem.

Foram adicionadas também funções de exportações de dados para cada gráfico, objetivando manter a integridade das informações, assim é possível verificar os dados reais sem distorções visuais que alguns gráficos pode gerar. E para melhorar a visualização dos dados os *dashboards* também foram implementados para visualização em dispositivos móveis.

Outro ponto, relevante foi a fusão de dados junto à base *Google Scholar*. Para isso utilizou-se a biblioteca *scholarly*<sup>1</sup>, escrita em Python, para coleta de dados. Para efetuar a fusão dos dados, uma vez que o identificador *Google* do pesquisador não é encontrado no currículo Lattes, utilizou-se o nome completo do pesquisador e nome em citações, juntamente com uma validação do domínio de e-mail igual a “ufopa.edu.br”.

Por fim, foi criada uma arquitetura baseada em OpenAPI para que a base compilada

---

<sup>1</sup> Disponível em: <https://pypi.org/project/scholarly/>. Acesso em 29 de julho de 2021.

pudesse ser acessada como serviço *web*, garantindo transparência e flexibilidade de dados já processados pelo sistema. Além disso, é possível estender novas exibições por meio de *iframe*<sup>2</sup>, uma funcionalidade adicionada que permite que *dashboards* e relatórios possam ser gerados externamente, como por exemplo em Microsoft PowerBI, Apache Superset e Metabase, e disponibilizadas por páginas internas no sistema centralizando os dados em um único portal.

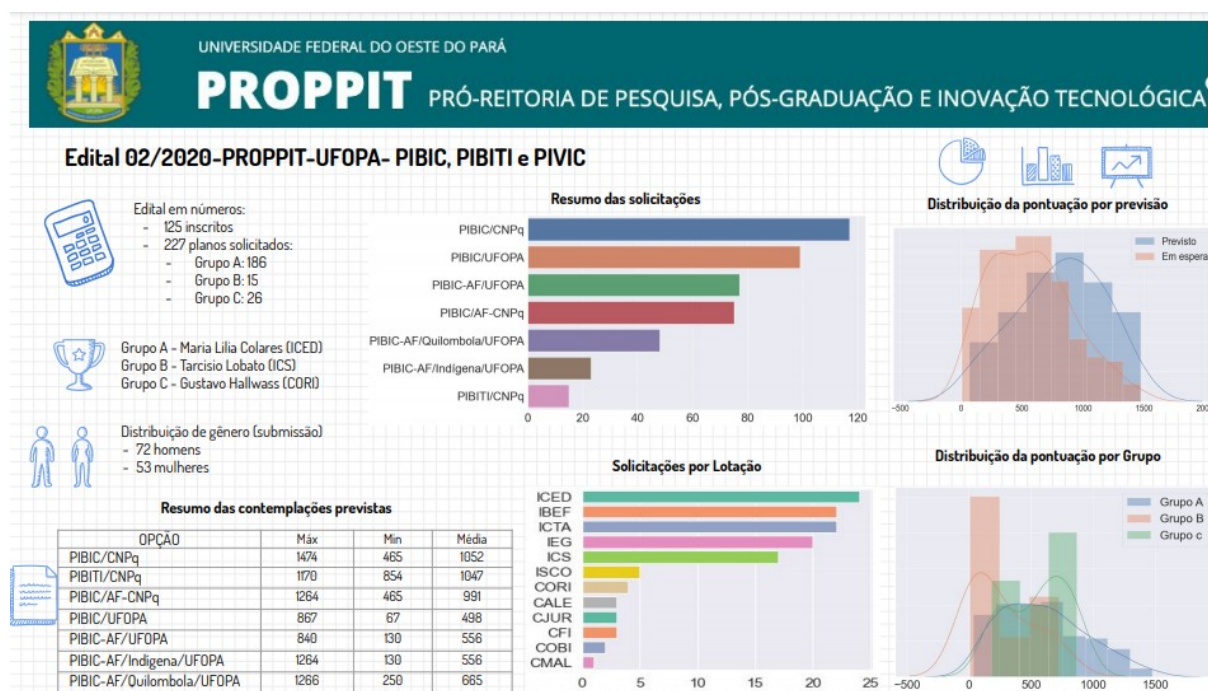
### 3.4 Demonstração

Nesta Seção os artefatos construídos são apresentados, demonstrando a solução dos desafios apresentados e o cumprimento dos objetivos traçados.

#### 3.4.1 Desafio 1: Edital PIBIC

A proposta de solução para o Desafio 1, descrita na Seção 3.2.1, foi apresentada para a DPE e aplicada no edital de pesquisa 03 de 2020 da UFOPA. Neste Processo 125 docentes estavam concorrendo às cotas. Como se pode observar na Figura 6, que mostra os pesquisadores mais bem pontuados, a distribuição das bolsas por setores, e distribuição da pontuação por grupo dos pesquisadores.

Figura 6 – Infográfico do edital 03/2020 de pesquisa.



Fonte: Disponibilizado pela Diretoria de Pesquisa.

<sup>2</sup> Um elemento HTML que incorpora outra página HTML na página atual. fonte: <https://developer.mozilla.org/pt-BR/docs/Web/HTML/Element/iframe>

Ressalta-se a automatização do processo seletivo, que desonera o setor, provê maior agilidade e confiabilidade ao pleito.

Nesse processo observamos, conforme disposto na Tabela ??, que houve redução na quantidade de servidores envolvidos no processo, possibilitando alocar estes recursos para outras atividades – isto sem citar os 125 docentes que tiveram que construir as tabelas de pontuação individual. Além de reduzir drasticamente o tempo de processamento das pontuações e divulgação do resultado do edital.

O processo manual contava com a participação de 22 servidores<sup>3</sup>, destes 2 servidores da DPE atuando de forma dedicada e mais 20 docentes participantes da comissão de pesquisa, para validar as planilhas de pontuação enviadas pelos docentes. Este processo era concluído em um prazo de 10 dias. Com a automatização do processo, apenas 1 servidor é necessário para gerenciar o sistema e o tempo de processamento foi reduzido para apenas 12 minutos com a quantidade de candidatos do edital específico. A substituição do sistema contratado pelo Sciensh, resultaria numa economia de recursos financeiros de significativa se considerar o valor aproximado desde o início da contratação do sistema Stela Experta.

Para confirmar a validade do artefato, as pontuações computadas foram confrontadas com planilhas produzidas manualmente pelos servidores. O edital foi aplicado e concluído, foi aberto período de recurso sobre as pontuações. Houve três recursos, sendo apenas um detectado como uma falha no sistema – referente à pontuação de orientação de Programa de Ensino Tutorial (PET). Como não há este campo explícito no Lattes, sugeriu-se que em próximo edital este item fosse modificado para “*Outras orientações*”. Outros recursos eram decorrentes de falha de preenchimento do Lattes, como por exemplo, o docente solicitou que o livro fosse computado como internacional, mas o local de publicação estava preenchido como Brasil; ou ainda, o periódico considerado como A2, mas o ISSN cadastrado estava relacionado com o Estrato B1. A Figura 7 apresenta um exemplo da visualização da pontuação e discriminação por item.

Conforme mostrado na Figura 7, para melhorar a validação dos processos seletivos dos editais de fomento à pesquisa, foi criada uma página para visualização detalhada das pontuações, discriminando os valores pontuados por grupos e os elencando os itens encontrados e utilizados no computador. Assim, permite que docentes validem suas pontuações a fim de corrigir erros no processo.

Com o auxílio do artefato foi possível simular novos modelos de pontuação, considerando a agilidade no processo de cálculo das pontuações. Permitindo avaliar os resultados para otimizar modelos que mais se adéquem aos perfis de pesquisa da instituição – este item foi particularmente importante para a construção da nova política de pesquisa da instituição.

<sup>3</sup> Destes 2 servidores dedicados e 20 docentes participantes da comissão de pesquisa.

Figura 7 – Exemplo da resultado detalhado de uma pontuação.

PONTUAÇÃO DA PRODUÇÃO CURRICULAR INDIVIDUAL DOCENTE - 2015 A 2020				
<div> <div> <div>SCIENSH</div> <div>Pontuações</div> <div>Produções</div> <div>Projetos</div> <div>Pesquisadores</div> </div> <div>Ajuda Admin</div> </div>				
<div> <div>nome: MARIA LILIA IMBIRIBA SOUSA COLARES</div> <div>identificador: 9671465461954562</div> <div>pontuação: 1474</div> </div> <div> <div>edital: 2020.1</div> <div>ano: 2020</div> <div>Categoria: Docente</div> </div>				
Produção	Valor Item	Máx. Pontos	Qtd.	Pontuação
<input checked="" type="checkbox"/> Artigos Publicados Em Periódicos De Acordo Com O Maior Qualis Da Revista		500	25	500
<input checked="" type="checkbox"/> Artigo Completo Em Periódico A1	50	0	1	50
<input checked="" type="checkbox"/> Artigo Completo Em Periódico A2	45	0	8	360
1 Políticas Educacionais Para A Formação Docente Na Educação Básica 151171949 2016				
2 Educação Infantil: O Papel Do Gestor Escolar Para A Promoção Da Educação Integral Da Criança 21782229 2016				
3 Desafios Da Gestão Da Escola De Tempo Integral 21782229 2017				
4 A Escola Transformando Vidas De Mulheres Negras, Ribeirinhas, Na Região Fronteiriça Brasil-bolivia Em Me... 21776210 2019				
5 Política Educacional No Âmbito Municipal: Experiência No Interior Da Amazônia 23581425 2019				
6 Reforma Do Ensino Médio: Desafios E Possibilidades Da Educação Integral 21782229 2019				
7 Sistemas Municipais De Educação: Autonomia E Gestão Democrática Como Premissas 19836597 2018				
8 Editorial 22379460 2017				
<input checked="" type="checkbox"/> Artigo Completo Em Periódico B1	35	0	13	455
<input checked="" type="checkbox"/> Artigo Completo Em Periódico B2	25	0	1	25
<input checked="" type="checkbox"/> Artigo Completo Em Periódico B3	20	0	1	20
<input checked="" type="checkbox"/> Artigo Completo Em Periódico B4	15	0	1	15
<input checked="" type="checkbox"/> Artigo Completo Em Periódico B5	10	0	0	0
<input checked="" type="checkbox"/> Artigo Completo Em Periódico C Ou Sem Extrato	5	0	0	0
<input checked="" type="checkbox"/> Livros E Capítulos De Livros		200	28	200
<input checked="" type="checkbox"/> Autoria De Livro Especializado (Edição Internacional)	60	0	0	0
<input checked="" type="checkbox"/> Autoria De Livro Especializado (Edição Nacional/focal)	50	0	1	50
<input checked="" type="checkbox"/> Autoria De Capítulo De Livro Especializado (Edição Internacional)	40	0	0	0
<input checked="" type="checkbox"/> Autoria De Capítulo De Livro Especializado (Edição Nacional/focal)	30	0	16	480
<input checked="" type="checkbox"/> Organização De Livro Especializado (Edição Internacional)	20	0	0	0
<input checked="" type="checkbox"/> Organização De Livro Especializado (Edição Nacional/focal)	10	0	11	110
<input checked="" type="checkbox"/> Trabalhos Em Eventos*		300	33	174
<input checked="" type="checkbox"/> Trabalho Completo Em Evento Internacional	10	0	3	30
<input checked="" type="checkbox"/> Trabalho Completo Em Evento Nacional	7	0	6	42
<input checked="" type="checkbox"/> Trabalho Completo Em Evento Regional/focal	5	0	8	40
<input checked="" type="checkbox"/> Resumo Ou Resumo Expandido Em Evento Internacional	7	0	3	21
<input checked="" type="checkbox"/> Resumo Ou Resumo Expandido Em Evento Nacional	5	0	1	5

version 0.0.1

Fonte: Elaborado pelo autor.

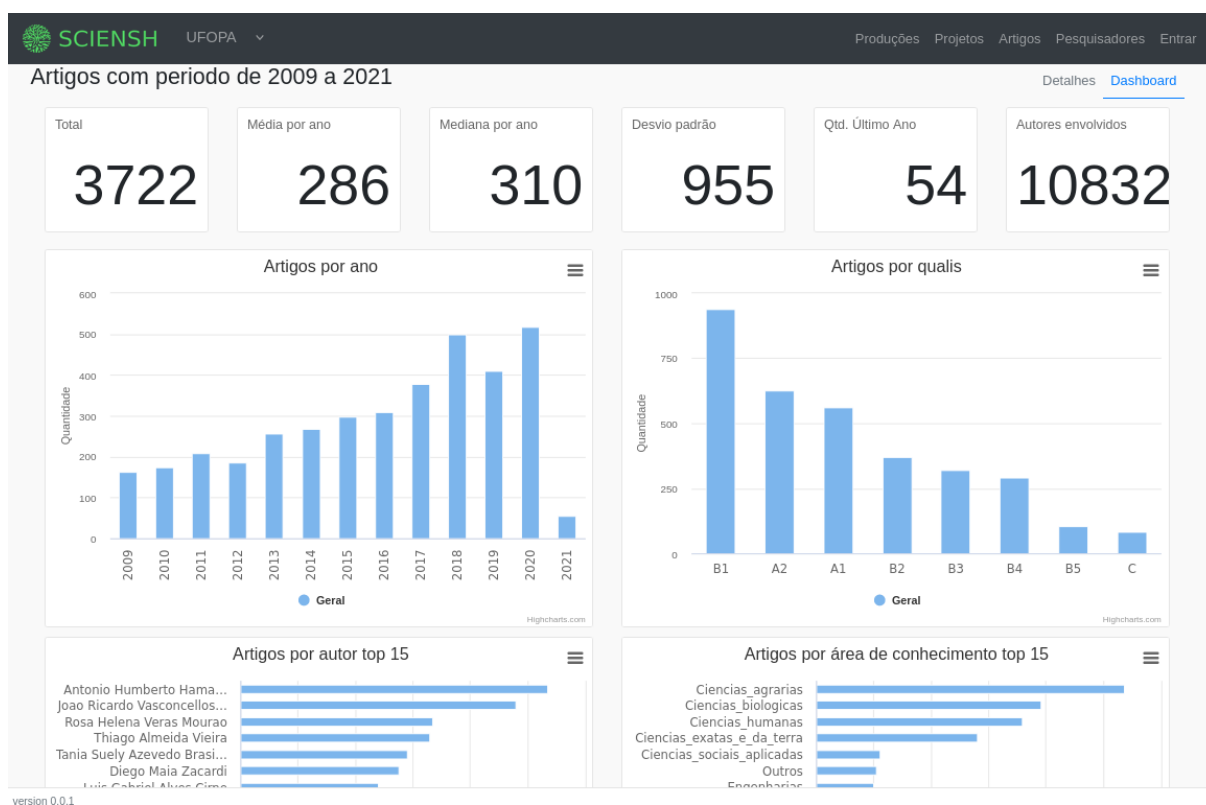
### 3.4.2 Desafio 2: Visualização dos dados

Nesta etapa, pôde-se demonstrar a validade do artefato. Uma versão do protótipo está disponibilizada para utilização da PROPPIT. Nesta versão, o processo de coleta de dados está executando de forma automática e disponibilizado por páginas públicas e internas, permitindo a visualização das produções por meio de *dashboards*, conforme Figura 8.

A Figura 8 apresenta a contabilização de produções por ano e por Qualis, identificando os autores mais ativos e as áreas de conhecimento Capes mais pesquisadas. É possível identificar os institutos que mais geram produções, palavras-chave com base nos títulos e descrições das produções. Além de fornecer estatísticas como média, desvio padrão, mediana, total e quantidade no último ano por tipo de produção.

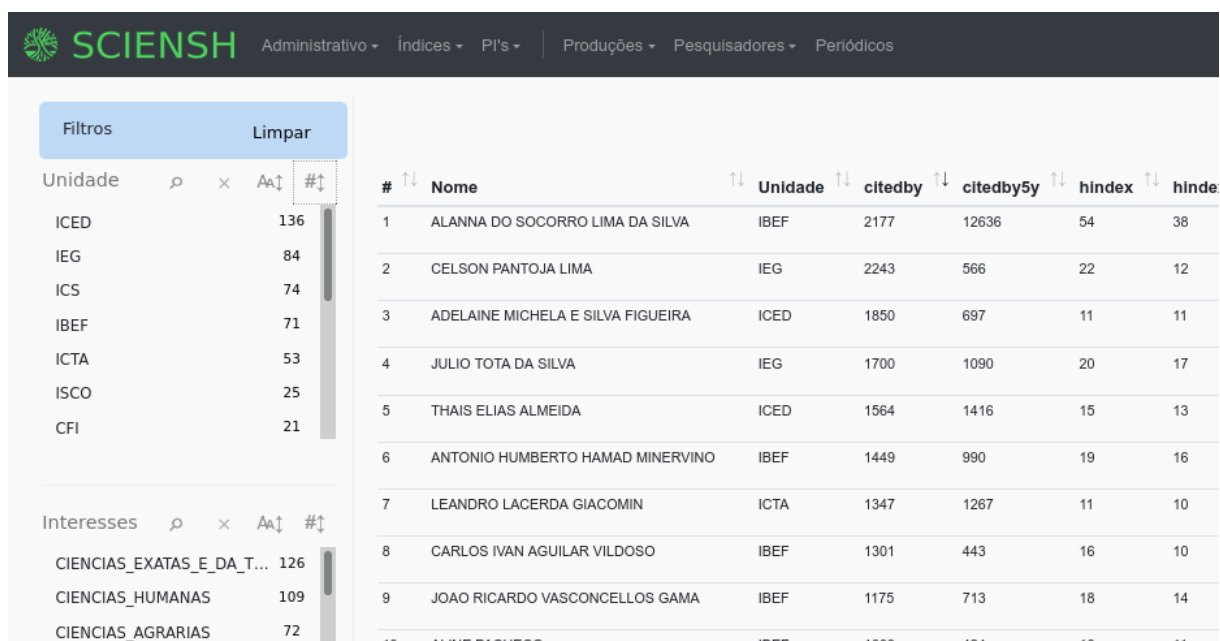
Além de *dashboards* quantitativos, é possível obter o conhecimento detalhando das produções através de páginas descritivas, permitindo vários filtros, ordenamentos e funções de exportação destes dados, conforme a Figura 9.

Figura 8 – Dashboard com perfil das produções da instituição.



Fonte: Elaborado pelo autor.

Figura 9 – Filtros e exibições descritivas.

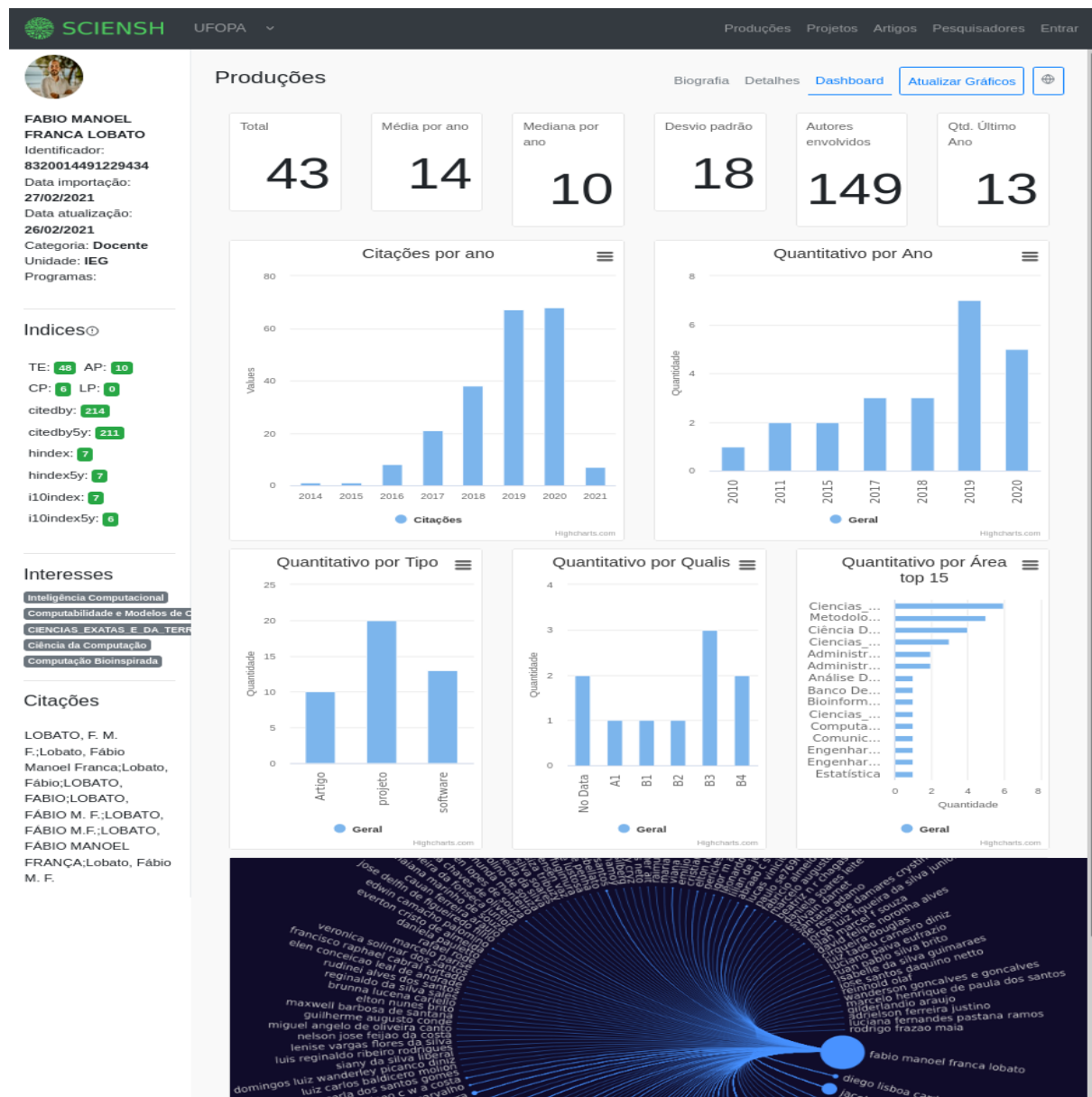


Fonte: Elaborado pelo autor.

Para facilitar a procura de pesquisadores, foi criado um *dashboard* com o perfil individual.

Nele é possível identificar pesquisadores por área de conhecimento, áreas de interesses, palavras-chave de suas produções, ordenamento por índices de produção, seus coautores, estatísticas de produtividades, entre outros.

Figura 10 – Perfil do pesquisador.



Fonte: Elaborado pelo autor.

Aborda-se também a quantificação e estratificação de vários tipos de produção, artigos, trabalhos em eventos, prêmios, orientações, registros de patentes, registros de *softwares* e projetos. Essas quantificações foram segmentadas por pesquisador gerando índices individuais e facilitando ordenações. Por exemplo, quantitativo de artigos completos, de livros, capítulos, trabalhos em eventos, índice-H e citações, por exemplo.

Necessidades de transparência foram solucionadas por meio de páginas públicas, como a



página da instituição, mostrada na Figura 11. Este requisito permite ver o perfil da instituição por meio das palavras-chave, seus pesquisadores, seus projetos, artigos, pesquisadores mais citados, maiores financiadores, últimos projetos, entre outros. Visando ampliar alcance e visibilidade das produções da instituição, são mostrados *links* de cada artigo completo para buscas dos arquivos no *Google Scholar* e *links* diretos dos perfis dos pesquisadores na Plataforma Lattes. Para avaliar a produtividade da instituição foi desenvolvido um gráfico segmentado por ano e tipo de produção, capaz de mostrar o quanto foi produzido de cada tipo de produção ao longo do tempo.

Com a utilização da plataforma Lattes como base de dados para mensurar a produtividade dos pesquisadores da instituição permitiu-se descentralizar o trabalho e garantir a validade das informações, uma vez que os próprios pesquisadores são os responsáveis pelo cadastro. Isso permitiu fornecer anuário e relatórios dinâmicos de forma automatizada.

Figura 11 – Página pública da instituição.



Fonte: Elaborado pelo autor.

### 3.5 Avaliação

A etapa de avaliação consiste na resposta dos *stakeholders* sobre as soluções propostas e demonstradas, com o intuito de determinar se o artefato supera os desafios apresentados. Adicionalmente, identificam-se novos requisitos para melhorias dos artefatos construídos. Para tal, foi necessário estabelecer uma forma de avaliação e coleta de informações para, se possível, adaptações da solução oferecida. Assim, a condução desta etapa se executou em uma abordagem quantitativa-qualitativa.

Do ponto de vista qualitativo, conduziu-se entrevistas não-estruturadas com os gestores no momento da demonstração do artefato. Esta etapa auxiliou na eliciação de novos requisitos.

Convém destacar que o edital PIBIC foi avaliado internamente pela diretoria de pesquisa por meio de pesquisa de satisfação com os docentes participantes. Como tal item não foi desenvolvido no escopo do presente estudo, os resultados constam como Anexo B.

Durante a apresentação da segunda etapa do sistema, sobre o desafio 2, que consiste na necessidade de visualizar e publicizar as produções da instituição. Foram coletados os *feedbacks* dos *stakeholders* as soluções do sistema foram aprovadas e alguns novos requisitos foram solicitados para tornar a solução mais completa.

Dentre os requisitos solicitados estão: acesso diferenciado para perfis de pesquisador, gestor, administrador; permissão para simulação da própria pontuação conforme modelos de pontuações utilizados na instituição; *dashboards* de livros; *dashboards* de prêmios; *dashboards* de artigos publicados em conferencias; padronização da nomenclatura das áreas de conhecimento; vitrine virtual de produtos contendo os interesses negociais da instituição e os serviços que a instituição oferece, tais como transferência de tecnologia, patentes e softwares; mecanismo de troca de mensagens para captar interessados em financiar, desenvolver ou apoiar projetos com a instituição.

Após a etapa de desenvolvimento dos novos requisitos identificados foi aplicado um questionário de satisfação na escala *Likert*, que consiste numa escala para mensuração da concordância com as afirmações propostas (FILARDI; TRAINA, 2008). O formulário aplicado foi desenvolvido com 10 afirmações e utilizando a escala *Likert* de cinco pontos invertida, conforme análises de (DALMORO; VIEIRA, 2013), pois, em média, apresenta maior precisão, facilidade e agilidade em ser respondida. Este formulário está disposto no Apêndice B. O resultado de satisfação dos *stakeholders* em relação as premissas constantes no questionário foi positivo, concordando plenamente e parcialmente com os requisitos levantados nos estudos.

### 3.6 Comunicação

A etapa de comunicação consiste na apresentação e divulgação dos dados coletados e produzidos durante esta pesquisa. E se desenvolveu na publicação de um artigo de prospecção sobre as ferramentas de análises de produtividade técnica-científica e as informações e meios que as instituições de nível superior divulgam. Esse artigo encontra-se publicado no periódico “*E-Tech: Tecnologias para Competitividade Industrial*” e está disponível no Apêndice A.

Outra produção aderente ao PROFNIT é a construção de um relatório técnico, contendo a pesquisa de qualidade e infográficos produzidos durante o edital de pesquisa 2020, em que o artefato produzido foi aplicado. Este infográfico foi publicado pela página institucional da diretoria de pesquisa<sup>4</sup>. Destaca-se ainda o registro dos artefatos (*softwares*) construídos junto ao Instituto Nacional de Propriedade Industrial. O certificado de registro de software consta no Anexo C.

<sup>4</sup> Disponível em: <http://www.ufopa.edu.br/media/file/site/proppit/documentos/2020/\5b4ac5bdbed47c76d67b207d0223125e.pdf>. Acesso em 29 de julho de 2021.

Adicionalmente, cita-se a dissertação, descrevendo os estudos desenvolvidos nesta temática e a disponibilização do artefato para a instituição publicizar os dados coletados e produzidos, buscando maior divulgação dentro da universidade e maior transparência para a sociedade. Além do desenvolvimento de materiais de apoio, durante a disciplina de Oficina Profissional, que auxiliavam na utilização da ferramenta e preenchimento de informações importantes em plataformas de pesquisa.

## 4 CONCLUSÕES

Nesta dissertação o processo de projeto, construção e avaliação do *SCIENSH: Sistema de apoio à gestão da pesquisa e inovação baseado em cientometria* foi apresentado. O artefato foi desenvolvido sob a ótica do método DSR e teve como produtos um artigo publicado no periódico *E-Tech: Tecnologias para competitividade industrial*; Além do próprio sistema, que resultou no Registro de Programa de Computador junto ao INPI, caracterizando produções científica e técnica, respectivamente. Alguns subprodutos também valem ser mencionados, como o vídeo de apresentação disponibilizado aos gestores a fim de auxiliá-los no uso inicial e na exploração da ferramenta; a biblioteca para buscas recursivas em arquivos JSON, formato que permite explorar profundamente os currículos cadastrados na plataforma Lattes.

O Sciensh usa diversas fontes de dados, tais como a plataforma Lattes, o *Google Scholar*, o SIGAA e o Sucupira, permitindo assim, a fusão de dados e maior completude de dados cientométricos a serem disponibilizados aos gestores institucionais. As principais funcionalidades do Sciensh incluem a geração de *dashboards* e indicadores institucionais, disponibilização de metadados e também a customização de indicadores a fim de se adequarem à realidade institucional. Além disso, utilizam-se técnicas de análises baseadas em redes e em mineração de texto para prover informações com maior acuidade e de fácil entendimento, possibilitando a tomada de decisões baseadas em dados. Destaca-se ainda que o Sciensh provê três níveis de acesso, um portal público, outro para membros internos da instituição e um específico para os gestores.

Dos impactos esperados do projeto, vale mencionar a melhora na eficiência da gestão por meio do processo de mapeamento das áreas e tendências de pesquisa, integração com dados institucionais, acompanhamento da produtividade e possibilidade de avaliação e direcionamento das pesquisas e políticas públicas da instituição. Com o término desse projeto, será criada uma base de dados única das produções acadêmicas da instituição que visa divulgar o atual estado da produção acadêmica para auxiliar em tomadas de decisões.

Como próximos passos deste estudo pretende-se continuar com o aperfeiçoamento do protótipo, junto com o setor responsável pela gestão da pesquisa na instituição e com outras instituições parceiras, durante a utilização do artefato na instituição, aperfeiçoando funcionalidades como redes de coautorias, análises de tópicos e grupos de pesquisa. Vislumbra-se também a disponibilização de artefatos como a biblioteca para buscas recursivas e material de apoio. Por fim, estuda-se a possibilidade de integração com outras fontes de dados como o *Research Gate* e a implementação de um módulo para comparação interinstitucional.

Vale ressaltar ainda que o enfoque do projeto até o momento consistiu no desenvolvimento do artefato e para as próximas etapas almeja-se a construção de um modelo de negócios

que possibilite a transferência de tecnologia para outras instituições e monetização do produto desenvolvido.

Algumas dificuldades foram enfrentadas no decorrer do projeto. O entendimento do domínio de aplicação tão amplo que é a cientometria, aliada a multiplicidade e dinamicidade das fontes de dados, foi o quesito que mais consumiu o tempo deste projeto. A pandemia da Covid-19 também impactou no agendamento de reuniões com os *stakeholders* e no próprio processo de implementação e de implantação do sistema - considerando o sigilo dos dados (e.g.: uso do CPF como chave de busca na PL, por exemplo), o desenvolvimento ocorreu majoritariamente na própria UFOPA, ou seja, sempre que estes dados precisaram ser utilizados, sendo assim, o processo foi prejudicado durante as medidas de restrição de acesso.

Algumas ameaças à validade do estudo também precisam ser mencionadas. Por exemplo, embora o Sciensh automatize e forneça relatórios com fusão de dados das produções, os resultados referem-se aos dados cadastrados na Plataforma Lattes, sendo recomendada uma verificação cuidadosa nos relatórios automaticamente gerados para que possíveis erros de preenchimento ou ausências de dados sejam corrigidos. Ressalta-se, que os dados quantitativos produzidos por ferramentas similares a este protótipo, precisam ser analisado com cuidado, levando em consideração contextos sociais, econômicos e regionais para não haver dissociação com a realidade e por fim contribuir de forma produtiva com a gestão do conhecimento.

## REFERÊNCIAS

ALVES, Alexandre Donizeti; YANASSE, Horacio Hideki; SOMA, Nei Yoshihiro. Lattesminer: uma linguagem de domínio específico para extração automática de informações da plataforma lattes. **XII Workshop de Computação Aplicada - WORCAP**, 2012.

BALANCIERI, Renato *et al.* A análise de redes de colaboração científica sob as novas tecnologias de informação e comunicação: um estudo na plataforma lattes. **Ciência da informação**, IBICT, v. 34, n. 1, 2005.

CAREGNATO, Sonia Elisa. Google acadêmico como ferramenta para os estudos de citações: avaliação da precisão das buscas por autor. **Pontodeacesso**, v. 5, n. 3, p. 72–86, 2011.

CARPENTER, Christopher R; CONE, David C; SARLI, Cathy C. Using publication metrics to highlight academic productivity and research impact. **Academic emergency medicine**, Wiley Online Library, v. 21, n. 10, p. 1160–1172, 2014.

CESSO, Marcus Vinícius da Silva; FERRAZ, Renato Ribeiro Nogueira. Uso de ferramentas computacionais para gestão estratégica da produtividade em pesquisa de um programa de pós-graduação stricto sensu: Uma revisão sistemática da literatura. **Proceedings of the 14th CONTECSI International Conference on Information Systems and Technology Management**, v. 14, 2017.

COSTA, Helder Gomes. Modelo para webibliomining: proposta e caso de aplicação. **Revista da FAE**, v. 13, n. 1, p. 115–126, 2010.

DALMORO, Marlon; VIEIRA, Kelmara Mendes. Dilemas na construção de escalas tipo likert: o número de itens e a disposição influenciam nos resultados? **Revista gestão organizacional**, v. 6, n. 3, 2013.

DIAS, Thiago MR; MOREIRA, Tales HJ; DIAS, Patrícia M. Caracterização e análise das redes de colaboração científica dos bolsistas de produtividade em pesquisa do cnpq. In: SBC. **Anais do VII Brazilian Workshop on Social Network Analysis and Mining**. [S.l.], 2018.

DIGIAMPIETRI, Luciano A *et al.* Minerando e caracterizando dados de currículos lattes. In: SBC. **Anais do I Brazilian Workshop on Social Network Analysis and Mining**. [S.l.], 2012.

DRESCH, Aline; LACERDA, Daniel Pacheco; JÚNIOR, José Antonio Valle Antunes. **Design science research: método de pesquisa para avanço da ciência e tecnologia**. [S.l.]: Bookman Editora, 2015.

FILARDI, Ana Lúcia; TRAINA, Agma Juci Machado. Montando questionários para medir a satisfação do usuário: avaliação de interface de um sistema que utiliza técnica de recuperação de imagens por conteúdo. In: IN: SIMPÓSIO BRASILEIRO DE FATORES HUMANOS EM SISTEMAS COMPUTACIONAIS, 8 . . . . **Embrapa Territorial-Artigo em anais de congresso (ALICE)**. [S.l.], 2008.

FRANÇA, Antônio MB. A inovação no setor público. **Cadernos**, v. 1, n. 1, p. 10–14, 2017.

GIL, Antonio Carlos. **Métodos e técnicas de pesquisa social**. [S.l.]: 6. ed. Editora Atlas SA, 2008.

- HAYASHI, Maria Cristina Piumbato Innocentini. Afinidades eletivas entre a cientometria e os estudos sociais da ciência. Universidade Estadual de Campinas, 2013.
- HICKS, Diana *et al.* Bibliometrics: the leiden manifesto for research metrics. **Nature News**, v. 520, n. 7548, p. 429, 2015.
- JACSÓ, Péter. Metadata mega mess in google scholar. **Online Information Review**, Emerald Group Publishing Limited, 2010.
- KNAFLIC, Cole Nussbaumer. **Storytelling com Dados: Um guia sobre visualização de dados para profissionais de negócios**. [S.l.]: Alta Books, 2019.
- LANE, Julia. Let's make science metrics more scientific. **Nature**, Nature Publishing Group, v. 464, n. 7288, p. 488–489, 2010.
- LEITE, Paula; MUGNAINI, Rogério; LETA, Jacqueline. A new indicator for international visibility: exploring brazilian scientific community. **Scientometrics**, Akadémiai Kiadó, co-published with Springer Science+ Business Media BV ..., v. 88, n. 1, p. 311–319, 2011.
- LOPES, Giseli Rabello. Avaliação e recomendação de colaborações em redes sociais acadêmicas. 2012.
- MACCARI, Emerson Antonio; KNISS, Claudia Terezinha; FERRAZ, Renato Ribeiro Nogueira. Organização e gestão de pós-graduação por meio da ferramenta computacional scriptlattes. **Multi-Science Research (MSR)**, v. 2, n. 1, p. 113–142, 2019.
- MARTINS, Cibele B *et al.* The influence of co-authorship networks on the performance of graduate programs in the brazilian evaluation system. In: IEEE. **Proceedings of PICMET'14 Conference: Portland International Center for Management of Engineering and Technology; Infrastructure and Service Integration**. [S.l.], 2014. p. 1569–1584.
- MEGLIORANSI, Murilo Dutra. Um sistema para coleta e apresentação de indicadores bibliométricos da plataforma lattes. Universidade Federal do Pampa, 2019.
- MENA-CHALCO, Jesús Pascual; JÚNIOR, C. Prospecção de dados acadêmicos de currículos lattes através de scriptlattes. **Bibliometria e Cientometria: reflexões teóricas e interfaces**. São Carlos: Pedro & João, p. 109–128, 2013.
- MINGERS, John; LEYDESDORFF, Loet. A review of theory and practice in scientometrics. **European Journal of Operational Research**, v. 246, n. 1, p. 1–19, 2015. ISSN 0377-2217. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S037722171500274X>>.
- MUGNAINI, Rogério; LEITE, Paula; LETA, Jacqueline. Fontes de informação para análise de internacionalização da produção científica brasileira. **PontodeAcesso**, v. 5, n. 3, p. 87–102, 2011.
- MUKHERJEE, Bhaskar. Do open-access journals in library and information science have any scholarly impact? a bibliometric study of selected open-access journals using google scholar. **Journal of the American Society for Information Science and Technology**, Wiley Online Library, v. 60, n. 3, p. 581–594, 2009.
- NORUZI, Alireza. Google scholar: The new generation of citation indexes. Walter de Gruyter GmbH & Co. KG, 2005.

- PARRA, Maurício Rodrigues; COUTINHO, Renato Xavier; PESSANO, Edward Frederico Castro. Um breve olhar sobre a cienciometria: origem, evolução, tendências e sua contribuição para o ensino de ciências. **Revista Contexto & Educação**, v. 34, n. 107, p. 126–141, 2019.
- PEFFERS, Ken *et al.* A design science research methodology for information systems research. **Journal of management information systems**, Taylor & Francis, v. 24, n. 3, p. 45–77, 2007.
- PRESSMAN, Roger; MAXIM, Bruce. **Engenharia de Software-8ª Edição**. [S.l.]: McGraw Hill Brasil, 2016.
- PRICE, Derek J de Solla. **Little science, big science**. [S.l.]: Columbia University Press, 1963.
- SANTOS, Raimundo Nonato Macedo dos; KOBASHI, Nair Yumiko. Bibliometria, cientometria, infometria: conceitos e aplicações. *Tendências da Pesquisa Brasileira em Ciência da Informação*, 2009.
- SCHNEIDER, Jesper; LARSEN, Birger; INGWERSEN, Peter. A comparative study of first and all-author co-citation counting, and two different matrix generation approaches applied for author co-citation analyses. **Scientometrics**, Akadémiai Kiadó, co-published with Springer Science+ Business Media BV ..., v. 80, n. 1, p. 103–130, 2009.
- SILVA, José Aparecido da; BIANCHI, Maria de Lourdes Pires. Cientometria: a métrica da ciência. **Paidéia (Ribeirão Preto)**, SciELO Brasil, v. 11, n. 21, p. 5–10, 2001.
- SILVEIRA, Diego Santana; NUNES, Maria Augusta Silveira Netto. Avaliação de usabilidade de interfaces computacionais: uma prospecção tecnológica. **Revista GEINTEC-Gestão, Inovação e Tecnologias**, v. 3, n. 3, p. 111–119, 2013.
- SOUZA, Elaine Calasans; OLIVEIRA, Marcus Rogério de. Comparativo entre os bancos de dados mysql e mongodb: quando o mongodb é indicado para o desenvolvimento de uma aplicação. **Revista Interface Tecnológica**, v. 16, n. 2, p. 38–48, 2019.
- SPINAK, Ernesto. Indicadores cienciométricos. **Ciência da informação**, SciELO Brasil, v. 27, n. 2, p. nd–nd, 1998.
- SU, Hsin-Ning; LEE, Pei-Chun. Mapping knowledge structure by keyword co-occurrence: a first look at journal papers in technology foresight. **Scientometrics**, Akadémiai Kiadó, co-published with Springer Science+ Business Media BV ..., v. 85, n. 1, p. 65–79, 2010.
- VAUGHAN, Liwen; SHAW, Debora. A new look at evidence of scholarly citation in citation indexes and from web sources. **Scientometrics**, Springer, v. 74, n. 2, p. 317–330, 2008.
- ZENG, An *et al.* The science of science: From the perspective of complex systems. **Physics Reports**, Elsevier, v. 714, p. 1–73, 2017.



## APÊNDICES

**APÊNDICE A – ARTIGO PUBLICADO NA REVISTA *E-TECH: TECNOLOGIAS PARA COMPETITIVIDADE INDUSTRIAL* EM 2021**

# PROSPECÇÃO DE FERRAMENTAS PARA GESTÃO DE PESQUISA E DESENVOLVIMENTO EM INSTITUIÇÕES DE CIÊNCIA E TECNOLOGIA

Fabiano Paulo Nhoatto<sup>1</sup>  
Adrielson Ferreira Justino<sup>2</sup>  
Éfren Lopes de Souza<sup>3</sup>  
Fábio Manoel França Lobato<sup>4</sup>

## RESUMO

A análise de grupos de pesquisadores e de suas respectivas produções são atividades relevantes e complexas. Na medida em que novos conhecimentos são gerados, é necessário que seja conduzida uma avaliação de todas essas produções, sobretudo quanto à qualidade e ao impacto. No Brasil, os pesquisadores alimentam a base de dados Lattes com todas as suas produções técnico-científicas pertinentes. Sendo à luz deste cenário, o presente estudo objetiva apresentar o cenário das principais ferramentas de extração e análise da Plataforma Lattes, por meio da realização de uma prospecção tecnológica. E ainda, identificar como as Instituições de Ensino Superior publicizam a produção técnico-científica de seus pesquisadores. Este estudo é caracterizado como pesquisa exploratória, descritiva e de abordagem quanti-qualitativa. Os resultados mostram que nenhuma das ferramentas de extração analisadas atendeu todas as funcionalidades necessárias para avaliação cienciométrica de pesquisadores. Também, identificou-se que as Instituições de Ensino Superior possuem maior tendência a organizar seus indicadores e divulgar dados em conjuntos de páginas estáticas. O estudo conclui sugerindo a necessidade de uma nova solução tecnológica, que atenda funcionalidades indispensáveis para o auxílio em tomadas de decisão na gestão de pesquisas, principalmente no que tange à visualização da informação de forma interativa.

**PALAVRAS-CHAVE:** Prospecção Tecnológica. Cienciométrica. Produção Técnico-Científica. Gestão de Pesquisa.

\*\*\*

1. Mestrando, email: fabiano.nhoatto@ufopa.edu.br
2. Graduando, email: adrielferreira28@gmail.com
3. Doutor, email: efren.lopez@ufopa.edu.br
4. Doutor, email: fabio.lobato@ufopa.edu.br

\*\*\*

# 1 INTRODUÇÃO

Aproximadamente 2,5 milhões de novos artigos científicos são publicados a cada ano (JINHA, 2010). Em uma análise preparada pelo Grupo *Web of Science* para a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), com o intuito de fornecer uma visão geral do desempenho da pesquisa brasileira, constatou-se que o Brasil ficou em 13º lugar no mundo em termos de produção de trabalhos técnico-científicos entre 2013 - 2018 (ANALYTIC, 2019). Na medida em que novos conhecimentos são gerados, é fundamental que seja realizada uma avaliação de todas essas produções, sobretudo quanto a sua qualidade e impacto. Devido à constante produção de conhecimentos gerados pela ciência, é essencial a adoção de práticas de uso de técnicas quantitativas e de mensuração, com o intuito de compreender os processos de consolidação de determinadas disciplinas e como esse mecanismo é estruturado (PARRA; COUTINHO; PESSANO, 2019).

A busca pela excelência em áreas de pesquisa e, também, a competição pelos recursos

financeiros das agências de fomento são uns dos principais fatores para realização desses estudos e adoção de métricas bibliométricas (LOPES, 2012). Nesse sentido, a extração de dados de produção científica, análise e identificação de padrões bibliométricos são tópicos relevantes na área de Bibliometria e Cientometria (MENA-CHALCO; CESAR-JR, 2013).

A exemplo disso, editais de financiamento de projetos de pesquisas por agências de fomento consideram dados registrados nos currículos cadastrados na Plataforma Lattes (PL) dos proponentes como uma das formas de avaliação das propostas (DIAS; MOREIRA; DIAS, 2018). A PL tornou-se uma importante base de dados da produção acadêmica brasileira, por armazenar essencialmente informações dos currículos dos pesquisadores (BUFREM, 2005; MEGLIORANSI, 2019). O currículo Lattes inclui identificação, formação acadêmica, prêmios científicos/artísticos, produções bibliográficas, orientações, entre outras informações (AMORIN, 2003).

Figura 1: Página inicial da Plataforma Lattes



Fonte: Dos autores<sup>1</sup> (2021)

1. Print Screen retirado da página inicial do site da Plataforma Lattes. Disponível em: <http://lattes.cnpq.br/>

---

Análises de grandes grupos e de docentes de uma instituição são inviáveis de serem feitas manualmente. Assim, tal procedimento torna-se uma tarefa demorada e altamente suscetível a erros (MENA-CHALCO; CESAR-JR, 2013). Para contornar essa situação, estudos apresentam a utilização de ferramentas computacionais para auxiliar a coleta de dados da PL, oferecer indicadores quantitativos de produção e visualizar graficamente o desempenho de instituições e pesquisadores (ALVES; YANASSE; SOMA, 2012; CESSO; FERRAZ, 2017). A utilização efetiva de *softwares* proporciona inúmeros benefícios para os usuários, que vão desde a automatização, praticidade e maior velocidade na execução de tarefas até a redução de custos (PRESSMAN; MAXIM, 2016; SILVEIRA; AUGUSTA; NETTO, 2013).

No entanto, algumas ferramentas não podem ser acessadas ou distribuídas livremente, outras carecem de alguma funcionalidade, como a geração de redes de colaboração, perfil do pesquisador ou indicadores de uma instituição (MEGLIORANSI, 2019). Além disso, percebe-se uma escassez de iniciativas práticas quanto a sistemas para visualização de dados cientiométricos, abalizando tomadas de decisões orientadas a dados, compreensíveis

e interativas (SANTA ANNA; COURA MACULAN, 2019).

Estudos de prospecção tecnológica são considerados uma boa ferramenta para a realização de pesquisas que contribuem para o desenvolvimento de novas tecnologias (MARIA *et al.*, 2020). No setor de desenvolvimento tecnológico, indispensável para qualquer projeto de desenvolvimento do país, a necessidade de estudos prospectivos é evidente (REIS; SPINOLA; AMARAL, 2017). Sendo assim, a partir dos conceitos supracitados, o presente trabalho visa trazer à tona, por meio da realização de uma prospecção tecnológica, o cenário das principais ferramentas de extração e produção de indicadores bibliométricos, que utilizam dados provenientes da PL. O estudo também questiona a necessidade de reconfiguração nas Instituições de Ensino Superior (IES), acerca dos tradicionais serviços de indexação, organização, consumo e meios de divulgação de informação das produções técnico-científicas.

O restante deste artigo está organizado como segue. Na Seção 2 é discutida a metodologia utilizada para o desenvolvimento da pesquisa. Em sequência, na Seção 3 é realizada a análise dos dados, bem como discussão dos resultados obtidos na prospecção tecnológica. Por fim, na Seção 4 as considerações finais serão apresentadas.

## 2 METODOLOGIA

---

Este estudo caracteriza-se como uma pesquisa exploratória, descritiva e de abordagem quanti-qualitativa. A prospecção tecnológica pode ser considerada como uma metodologia capaz de mapear desenvolvimentos científicos e tecnológicos futuros com potencial de influenciar significativamente os rumos de uma economia, indústria ou sociedade (FERNANDES *et al.*,

2018; SILVEIRA; AUGUSTA; NETTO, 2013). Para este estudo, a metodologia de pesquisa empregada foi adaptada de estudos de prospecção tecnológica, como principal ferramenta de busca científica, na qual foram considerados registros em repositórios de *softwares* e de pedidos de patente em bases de dados (DOS SANTOS AMPARO; DO

---

RIBEIRO; GUARIEIRO, 2014; SILVA; AUGUSTA; PROCC, 2014; SILVEIRA; AUGUSTA; NETTO, 2013).

Dessa forma, foram realizados dois levantamentos, o primeiro com o intuito de rastrear soluções tecnológicas que realizam extração e análise de dados da PL, e consistiu em algumas determinações de busca. O segundo verifica os serviços das IES acerca da organização e comunicação de dados de pesquisa para a sociedade. As buscas ocorreram no período de março a novembro de 2020.

Para o primeiro levantamento, a prospecção tecnológica foi conduzida por meio de buscas em repositórios de *softwares* e bases de dados que envolveu o *GitHub*, *SourceForge* e base do Instituto Nacional de Propriedade Intelectual (INPI), além de artigos científicos que reportam soluções tecnológicas de extração da PL. Em seguida, foram determinadas as seguintes palavras-chave, para serem utilizadas nos processos de busca: (1) “extração lattes”, (2) “extraction lattes”, (3) “extract lattes”, (4) “lattes extractor”, (5) “lattes extrator”. Os termos (4) e (5) são semelhantes, porém ambos podem ser encontrados na literatura. Ademais, foram consideradas algumas ferramentas citadas nos trabalhos relacionados, para incrementar o maior número possível de softwares na análise comparativa.

Mediante os resultados das buscas, foi realizada a seleção dos *softwares* de interesse por meio da análise de seus resumos, na qual, aqueles que eram foco da busca foram analisados integralmente. Para a análise comparativa foram definidas funcionalidades com base nas informações mais relevantes disponibilizadas

nos currículos Lattes na construção de indicadores bibliométricos/cientométricos, além de meios de visualização da informação. Sendo elas: (1) Produções Bibliográficas por grupo, (2) Indicadores do membro do grupo, (3) Propriedade intelectual, (4) Orientações, (5) Prêmios científicos/ Artísticos, (6) Gráficos quantitativos de produção, (7) Grafos de Coautoria, (8) Gráficos interativos. Além disso, também foi definida uma nova funcionalidade, (9) Acesso Livre, para consultar a existência de restrições de acesso às ferramentas.

Para o segundo levantamento, primeiramente foi criada uma lista com 39 IES brasileiras considerando ranqueamento de produção técnico-científica das bases Nature Index<sup>2</sup> e Microsoft Academic<sup>3</sup>. Além disso, buscando uma melhor distribuição da análise, considerou-se também a representatividade regional. Incluíram-se algumas estaduais baseando-se nos mesmos critérios. Em seguida, foram definidos os termos a serem procurados por meio dos menus e por pesquisa de texto nos sites institucionais, pró-reitorias de pesquisa e Google. As palavras-chave buscadas foram: (1) anuário, (2) indicadores, (3) indicadores de pesquisa e (4) indicadores de produção.

De acordo com a Lei nº 12.527, de 18 de novembro de 2011, denominada Lei de Acesso à Informação (LAI), em seu Art. 8º: “É dever dos órgãos e entidades públicas promover, independentemente de requerimentos, a divulgação em local de fácil acesso, no âmbito de suas competências, de informações de interesse coletivo ou geral por eles produzidas ou custodiadas”. Esse fato reflete na transparência ativa (SILVA; HOCH; RIGHI, 2013), pois ela regulamenta o direito constitucional de

---

2. <https://www.natureindex.com/country-outputs/brazil>

3. <https://academic.microsoft.com/institutions>

acesso às informações públicas. Essa forma de transparência pode ser definida como a divulgação, por parte do Estado, das informações de interesse geral da população, independente de solicitação, principalmente por meios eletrônicos na internet, como sites e portais.

Dessa forma, foi verificado quais das IES utilizam como meio de divulgação: (1) Portal Institucional; (2) Relatório, dos quais podem

ter atualizações bimestral, semestral ou anual; (3) Portal Interativo, para IES que possuem sites que permitem filtrar ou interagir com dados quantitativos de pesquisas. Além disso, foi consultado o tipo de informação disponibilizada por meio de indicadores bibliométricos relevantes como: (4) Produções, (5) Projetos, (6) Pesquisadores, (7) Grupos de pesquisa, (8) Bolsas e (9) Propriedade Intelectual (AMORIN, 2003).

### 3 RESULTADOS E DISCUSSÃO

No Quadro 1 são apresentados os números dos registros encontrados com as suas respectivas palavras-chaves. Foram encontrados três registros de *softwares* na base do INPI, 44 ocorrências no GitHub e apenas um no SourceForge.

**Quadro 1:** Levantamento das ferramentas de extração da Plataforma Lattes

	Palavras-Chave	INPI	GitHub	SourceForge
(1)	Extração <i>lattes</i>	0	12	1
(2)	<i>Extraction latt</i>	0	1	0
(3)	<i>Extract latt</i>	1	15	0
(4)	<i>Lattes extractor</i>	1	10	0
(5)	<i>Lattes extrator</i>	1	6	0

Fonte: Dos autores (2021)

O Quadro 2 mostra as diferenças entre os sistemas selecionados. Constatou-se que nenhuma das seis ferramentas selecionadas atendem todas as funcionalidades em questão. O indicador bibliométrico de produção por grupo (e.g. uma instituição, instituto, grupo de pesquisa) e orientações é gerado em todas as ferramentas. Além de produções bibliográficas, o ScriptLattes, LucyLattes e Latticles disponibilizam separadamente produções artísticas, produções técnicas e projetos, as orientações, prêmios e eventos (DA SILVA *et al.*, 2019).

**Quadro 2:** Análise comparativa dos sistemas de extração da Plataforma Lattes

	Funcionalidades	SL	LL	LT	eL	SS	SE
(1)	Produções Bibliográficas por grupo	✓	✓	✓	✓	✓	✓
(2)	Indicadores do membro do grupo	✗	✗	✓	✗	✓	✓
(3)	Propriedade intelectual	✓	✓	✓	✗	✓	✓
(4)	Orientações	✓	✓	✓	✓	✓	✓
(5)	Prêmios científicos/ Artísticos	✓	✓	✓	✗	✗	✓
(6)	Gráficos quantitativos de produção	✓	✓	✗	✓	✓	✓
(7)	Grafos de Coautoria	✓	✓	✓	✓	✗	✗
(8)	Gráficos interativos	✗	✗	✗	✗	✓	✓
(9)	Livre Acesso	✓	✓	✓	✓	✓	✗

SL - ScriptLattes; LL LucyLattes; LT - Lattices; eL e-Lattes; SS- Sistema Somos; SE - Stela Experta.

Fonte: Dos autores (2021)

Adicionalmente, as ferramentas Somos e Stela Experta também fornecem informações de propriedade intelectual. No entanto, a e-Lattes e Somos são as únicas sem avaliação de prêmios científicos ou artísticos. Também pode ser observado que a avaliação de indicadores de forma individual, por membro do grupo, é fornecida somente pelas ferramentas Lattices, Somos e Stela Experta. A elevada quantidade de dados e informações dificulta a interpretação das informações científicas (REIS; SPINOLA; AMARAL, 2017). Para tal problemática, as ferramentas de extração e análise podem contar com recursos de técnicas de visualização, possibilitando ao tomador de decisão (e.g. pesquisador, gestor e outros) utilizar sua percepção visual para melhor analisar e compreender as informações presentes nesses indicadores (MACCARI; KNISS; FERRAZ, 2019; REIS; SPINOLA;

AMARAL, 2017). Isso possibilita garantir vantagens para organizações, visto que o ato de inovar é comumente uma decisão que perpassa a alta gestão, é indispensável que existam informações completas e precisas que subsidiem uma tomada de decisão inteligente (ZSCHORNACK *et al.*, 2019).

Ainda de acordo com o Quadro 2, é possível constatar que somente o Lattices não possui gráficos quantitativos de produção. Com exceção das plataformas Somos e Stela Experta todas possuem gráficos de colaboração científica, visto que, por ausência de um processo automático de identificação de colaborações pela própria Plataforma Lattes, a relação de colaboração não é constatada de forma automática. Isto reforça a necessidade de ferramentas de extração automática e técnicas para a identificar colaborações em grande volume de dados (DIAS; MOITA; DIAS, 2016).



Estudos sobre as redes de colaboração científica vêm recebendo atenção de analistas de diversas áreas do conhecimento, devido seu potencial de identificar e melhor compreender como as pesquisas em uma determinada área têm sido realizadas, ou como grupos de pesquisadores têm colaborado no desenvolvimento de suas pesquisas (DOS SANTOS AMPARO; DO RIBEIRO; GUARIEIRO, 2014; LOPES, 2012). Somente as ferramentas Somos e Stela Experta são as que dispõem de gráficos de produção que podem ser alterados usando a sua interface. Destaca-se que Stela Experta é a única que possui restrição de acesso.

A PL disponibiliza livremente os currículos Lattes, estes são visualizados por interface de consulta disponibilizada pelo CNPq. No entanto, ela apresenta currículos de forma individual, devido não ser projetada para apresentar análises e comparações com outros currículos (DIAS; MOITA; DIAS, 2016).

O Quadro 3 mostra as 39 IES estudadas no segundo levantamento, em que foram verificadas quais possuem os meios de divulgação em questão e quais os principais indicadores bibliométricos disponíveis nos seus repositórios de produção.

**Quadro 3:** Levantamento dos meios de comunicação de dados de pesquisa das IES

	UNIVERSIDADE	POR	REL	PO_I	PRD	PRJ	PES	GRU	BOL	PI
(1)	UFSCAR	X	✓	X	✓	✓	✓	✓	X	X
(2)	UFPA	X	✓	X	✓	✓	✓	✓	✓	X
(3)	UFAM	X	X	X	✓	X	✓	X	X	X
(4)	UFC	X	✓	X	✓	X	✓	✓	X	X
(5)	UFRN	✓	X	X	✓	✓	✓	✓	X	X
(6)	UFSC	✓	X	X	✓	✓	X	✓	✓	X
(7)	UFPE	✓	X	X	✓	X	✓	X	X	✓
(8)	UFF	X	✓	X	✓	X	✓	X	X	X
(9)	UFRGS	✓	X	X	X	✓	✓	X	✓	X
(10)	UFMG	✓	✓	✓	✓	X	✓	✓	X	✓
(11)	UFABC	X	✓	X	✓	X	X	X	X	X
(12)	UFOP	✓	✓	✓	✓	✓	✓	✓	✓	✓
(13)	UFRJ	✓	✓	X	✓	✓	✓	X	X	✓
(14)	UFES	✓	✓	X	✓	X	✓	X	X	X
(15)	UNB	✓	✓	✓	✓	✓	✓	✓	✓	✓
(16)	UFRRJ	✓	X	X	✓	✓	✓	✓	X	X



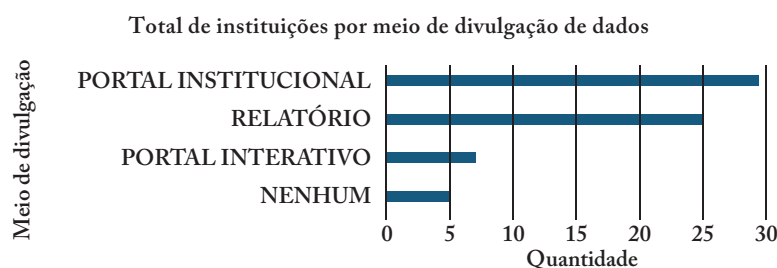
	UNIVERSIDADE	POR	REL	PO_I	PRD	PRJ	PES	GRU	BOL	PI
(17)	UFMS	X	X	X	✓	X	X	X	X	X
(18)	UNIFESP	✓	✓	✓	✓	✓	✓	✓	X	X
(19)	UFGD	✓	X	X	✓	X	✓	X	X	X
(20)	UFG	✓	X	✓	✓	✓	✓	✓	✓	X
(21)	UFMA	✓	✓	X	✓	✓	✓	X	X	X
(22)	UFPI	✓	✓	X	✓	✓	✓	✓	✓	X
(23)	UNIR	✓	✓	X	✓	✓	✓	✓	X	X
(24)	UFRR	✓	X	X	✓	✓	✓	✓	✓	X
(25)	UFAC	✓	✓	X	✓	✓	✓	X	✓	X
(26)	UNIFAP	✓	✓	X	✓	✓	✓	✓	✓	X
(27)	UFT	✓	✓	X	✓	X	✓	✓	X	X
(28)	UFV	✓	✓	✓	✓	✓	✓	X	✓	X
(29)	UNIRIO	✓	✓	X	✓	✓	✓	✓	✓	X
(30)	UTFPR	✓	✓	X	✓	X	✓	✓	✓	X
(31)	UFFS	X	✓	X	✓	✓	X	✓	✓	✓
(32)	UFPR	✓	X	✓	✓	X	✓	✓	X	✓
(33)	UFOPA	X	X	X	✓	✓	✓	✓	✓	X
(34)	UERJ	X	✓	X	✓	✓	X	✓	X	X
(35)	UEA	✓	X	X	✓	X	✓	X	X	X
(36)	UEMA	✓	✓	X	✓	✓	✓	✓	✓	X
(37)	UNICAMP	✓	✓	X	✓	✓	✓	✓	✓	✓
(38)	USP	✓	✓	X	✓	✓	✓	X	✓	✓
(39)	UNESP	✓	X	X	✓	X	✓	X	X	X

POR - Portal Institucional, REL - Relatório, PO\_I - Portal Interativo, PRD - Produções, PRJ - Projetos, PES - Pesquisadores, GRU - Grupos, BOL - Bolsas, PI - Propriedade Intelectual.

Fonte: Dos autores (2021)

De acordo com a Figura 2, apenas sete IES possuem portal interativo, 25 utilizam relatórios e 29 usam portais institucionais, caracterizando como o meio mais utilizado para divulgação de informações técnico-científicas. Apenas três não possuem nenhum meio de divulgação entre os pesquisados. Dessa forma, a maioria das IES estudadas realizam a divulgação científica por meio de algum repositório de produção. Entretanto, possuem maior tendência a organizar seus indicadores e divulgar dados em conjuntos de páginas estáticas, o que dificulta sua utilização e eventuais manutenções (MEGLIORANSI, 2019).

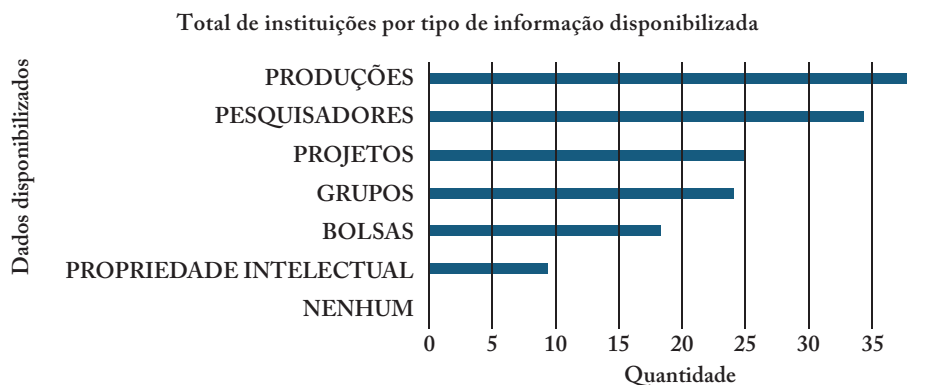
**Figura 2:** Total de instituições por Meio de Divulgação de Dados



Fonte: Dos autores (2021)

Por meio da Figura 3, observa-se que produções bibliográficas possuem maior frequência entre todos os tipos de informações disponibilizadas, que é realizada por 38 IES. Por outro lado, no que tange às produções de propriedade intelectual (e.g. produção técnica, produtos tecnológicos, softwares e outros), somente nove disponibilizam. Informações individuais de pesquisadores e de grupos (e.g. instituição, institutos, grupos de pesquisas e outros) ocorrem em mais da metade das IES, com o total de 34 e 24 ocorrências, respectivamente.

**Figura 3:** Total de instituições por tipo de informação disponibilizada



Fonte: Dos autores (2021)

---

A comunicação de dados de projetos de pesquisa possui uma boa frequência, um total de 25 instituições adotam essa prática. No entanto, o número de ocorrências de informações sobre bolsas é feito somente por 18, ou seja, menos da metade. Foi constatado que em relação ao quantitativo de IES que não divulgaram nenhum dado entre os levantados é nulo. Além disso, por meio das buscas realizadas no segundo levantamento verificou-se que informações sobre internacionalização, parcerias interinstitucionais e agências financiadoras não estão presentes ou não estão de fácil acesso. Essas informações são extremamente relevantes para a gestão da pesquisa e também para a comunidade.

## 4 CONCLUSÃO

---

É inegável a relevância que a tomada de decisão orientada a dados tem ganhado nos últimos anos (ZSCHORNACK *et al.*, 2019). A gestão eficiente de programas de pesquisa tem se beneficiado de forma crescente do grande volume de dados produzidos, que fez se consolidar uma área conhecida como cienciométrica. No Brasil, a maior parte dos estudos cienciométricos é conduzido baseando-se nos dados disponibilizados na Plataforma Lattes.

Visando conhecer o cenário das principais ferramentas de extração e análise da Plataforma Lattes, conduziu-se uma prospecção tecnológica em bases de artefatos de *software* como Instituto Nacional de Propriedade Intelectual, Github e SourceForge. Além disso, investigou-se como as Instituições de Ensino Superior publicizam a produção técnico-científica de seus pesquisadores, diminuindo a distância entre a academia e a sociedade.

Constatou-se uma carência quanto a ferramentas de código-livre ou gratuitas para obtenção e análise de dados advindos da Plataforma

Lattes. E ainda, as ferramentas proprietárias analisadas não contemplavam algumas produções importantes, como prêmios, títulos e grafos de coautorias. De forma complementar, analisou-se também a forma com a qual as IES apresentam índices cienciométricos. A maior parte da divulgação ocorre por meio de anuários estatísticos em pdf, longos e de difícil entendimento para a sociedade em geral. Dessa forma, conclui-se salutar o desenvolvimento de uma ferramenta interativa, capaz de auxiliar na avaliação da pesquisa institucional, abalizando o processo de tomada de decisão, permitindo ainda o acesso simples e flexível à produção técnico-científica institucional

O método de pesquisa utilizado não permite generalizações das observações, mas disponibiliza uma sistemática objetiva das ferramentas de extração da PL e o papel exercido pelos serviços de informação pelas IES. Futuras pesquisas poderiam ampliar a amostra analisada e comparar os resultados aqui alcançados.



# PROSPECTING TOOLS FOR RESEARCH AND DEVELOPMENT MANAGEMENT IN SCIENCE AND TECHNOLOGY INSTITUTIONS

## ABSTRACT

*The analysis of groups of researchers and their respective productions are relevant and complex activities. As new knowledge is generated, it is necessary to carry out an evaluation of all these productions and especially their quality. This study is characterized as exploratory, descriptive and quantitative-qualitative research. It aims to present the scenario of the main extraction and analysis tools of the Lattes Platform by conducting a technological prospecting. As well, to approach the methods about the indexing, organization and dissemination services of technical and scientific productions in Higher Education Institutions. It was found that none of the analyzed extraction tools met all the researched functionalities. It was identified that Higher Education Institutions have a greater tendency to organize their indicators and disseminate data in sets of static pages. The study concludes by suggesting the need for a new technological solution, which meets essential features to aid decision making in research management, especially with regard to the interactive display of information.*

\*\*\*

**KEYWORDS:** *Technological  
Prospecting. Extraction Tools.  
Scientific Communication.  
Technical-Scientific Production.  
Research Management.*

\*\*\*

## REFERÊNCIAS

- ALVES, A. D.; YANASSE, H. H.; SOMA, N. Y. LattesMiner: uma linguagem de domínio específico para extração automática de informações da Plataforma Lattes. **XII Workshop de Computação Aplicada - WORCAP**, p. 6, 2012.
- AMORIN, C. V. Organização do currículo: plataforma Lattes. **Pesquisa Odontológica Brasileira**, v. 17, n. suppl 1, p. 18–22, 2003.
- ANALYTIC, C. A **Pesquisa no Brasil:** Promovendo a excelência Análise preparada para a CAPES pelo Grupo Web of Science. [S. l.: s. n.], 2019, p. 42.
- BUFREM, L. O saber científico registrado e as práticas de mensuração da informação. **Ciência da Informação**, v. 34, n. 2, p. 9–25, 2005.
- CESSO, M. V.; FERRAZ, R. R. N. Uso De Ferramentas Computacionais Para Gestão Estratégica Da Produtividade Em Pesquisa De Um Programa De Pós-Graduação Stricto Sensu: Uma Revisão Sistemática Da Literatura. **Proceedings of the 14th CONTECSI International Conference on Information Systems and Technology Management**, v. 14, p. 977–995, 2017.

DIAS, T. M. R.; MOITA, G. F.; DIAS, P. M. Adoção da plataforma lattes como fonte de dados para caracterização de redes científicas. **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, v. 21, n. 47, p. 16, 2016.

DIAS, T. M. R.; MOREIRA, T. H. J.; DIAS, P. M. Caracterização e Análise das Redes de Colaboração Científica dos Bolsistas de Produtividade em Pesquisa do CNPq. Anais do VII Brazilian Workshop on Social Network Analysis and Mining - BraSNAM 2018. **Anais...**2018.

DOS SANTOS AMPARO, K. K.; DO RIBEIRO, M. C. O.; GUARIEIRO, L. L. N. Estudo de caso utilizando mapeamento de prospecção tecnológica como principal ferramenta de busca científica. **Perspectivas em Ciencia da Informacao**, v. 17, n. 4, p. 195–209, 2014.

FERNANDES, D. *et al.* Prospecção tecnológica de programas (softwares) para indústria petroquímica technological forecasting of softwares for petrochemical industry. **Revista GEINTEC-Gestão, Inovação e Tecnologias**, v. 8, p. 4231–4239, 2018.

JINHA, A. Article 50 million: An estimate of the number of scholarly articles in existence. **Learned Publishing**, v. 23, n. 3, p. 258–263, 2010.

LOPES, G. R. **Avaliação e Recomendação de Colaborações em Redes Sociais Acadêmicas**. [S. l.: s. n.], 2012.

MACCARI, E. A.; KNISS, C. T.; FERRAZ, R. R. N. **Organização e gestão de pós-graduação por meio da ferramenta computacional scriptlattes**. [S. l.: s. n.], 2019, p. 113–142.

MARIA, L. *et al.* Prospecção Científica e Tecnológica de Nanopartículas Poliméricas de PLGA contendo Antimicrobianos para o Combate às Biofilmes. **Revista GEINTEC-Gestão, Inovação e Tecnologias**, v. 10, p. 5676–5685, 2020.

MEGLIORANSI, M. D. **Um sistema para coleta e apresentação de indicadores bibliométricos da Plataforma Lattes**. [S. l.: s. n.], 2019.

MENA-CHALCO, J. P.; CESAR-JR, R. M. Prospecção de dados acadêmicos de currículos Lattes através de scriptLattes. **"Bibliometria e Cientometria: reflexões teóricas e interfaces"**. [S. l.: s. n.], 2013, p. 109–128.

PARRA, M. R.; COUTINHO, R. X.; PESSANO, E. F. C. Um Breve Olhar Sobre a Cienciometria: Origem, Evolução, Tendências e sua Contribuição para o Ensino de Ciências. **Revista Contexto & Educação**, v. 34, n. 107, p. 126–141, 28 mar. 2019.

PRESSMAN, R.; MAXIM, B. **Engenharia de Software**. 8. ed. [S. l.: s. n.], 2016, p. 968.

REIS, J. E. DOS; SPINOLA, A. T. P.; AMARAL, R. M. DO. Incipiência da visualização de indicadores bibliométricos e altmétricos nos Repositórios Institucionais brasileiros. **Em Questão**, v. 23, p. 213, 2017.

SANTA ANNA, J.; COURA MACULAN, B. A Gestão dos Dados de Pesquisa Nas Universidades e o Papel dos Serviços Informacionais Oferecidos Nas Bibliotecas: Uma Revisão Narrativa Resumo. **Múltiplos Olhares Em Ciência Da Informação**, v. 9, p. 1–16, 2019.

SILVA, R. L. DA; HOCH, P. A.; RIGHI, L. M. Transparência pública e a atuação normativa do CNJ. **Revista Direito GV**, v. 9, n. 2, p. 489–514, 2013.

SILVA, Í. D.; AUGUSTA, M.; PROCC, S. N. N. Prospecção de software para auxílio em tratamento de câncer exploration of software to aid cancer treatment. **Revista GEINTEC-Gestão, Inovação e Tecnologias**, v. 4, p. 1402–1413, 2014.

SILVEIRA, D. S.; AUGUSTA, M.; NETTO, S. Avaliação de Usabilidade de Interfaces Computacionais: Uma Prospecção Tecnológica. **Revista GEINTEC-Gestão, Inovação e Tecnologias**, v. 3, p. 111–119, 2013.

ZSCHORNACK, T. *et al.* Observatório de inovação como ferramenta de subsídio à inteligência nas organizações. **Revista E-Tech: Tecnologias para Competitividade Industrial-ISSN-1983-1838**, v. 10, n. 1, p. 7–22, 2019.

---

## SOBRE OS AUTORES



### **Fabiano Paulo Nhoatto**

Mestrando em Propriedade Intelectual e Transferência de Tecnologia para Inovação (PROFNIT) pelo Instituto de Engenharia e Geociências (IEG) da Universidade Federal do Oeste do Pará (UFOPA), em Santarém/PA (2019-2021). Graduado (2014) em Sistemas de Informação pela Universidade Federal do Pará (UFPA). Servidor público no cargo de Analista de TI da UFOPA, atuando principalmente em desenvolvimento web e mobile, mineração de dados e ciência de dados.



### **Éfren Lopes de Souza**

Possui Doutorado (2014) e Mestrado (2010) em Informática pela Universidade Federal do Amazonas (UFAM) e Graduação (2006) em Processamento de Dados pela Universidade Federal do Pará (UFPA). Atualmente é Professor Adjunto da Universidade Federal do Oeste do Pará (UFOPA). Tem experiência na área de Ciência da Computação, com ênfase em Redes de Sensores Sem Fio, atuando principalmente nos seguintes temas: algoritmos distribuídos, fusão de dados, localização e rastreamento de alvos.



### **Adrielson Ferreira Justino**

Graduando em Bacharelado em Ciência da Computação pelo Instituto de Engenharia e Geociências da Universidade Federal do Oeste do Pará (UFOPA), em Santarém/PA (2017-2022). Bolsista do Programa de Educação Tutorial (PET) do Ministério da Educação (MEC). É membro do Grupo de Estudo e Pesquisa do Laboratório de Computação Aplicada (LACA-UFOPA), desenvolvendo projetos de análise de redes sociais.



### **Fábio Manoel França Lobato**

É Professor de Computação na Universidade Federal do Oeste do Pará e lidera o Grupo de Pesquisa em Computação Aplicada. É bolsista de Produtividade Desenvolvimento Tecnológico e Extensão Inovadora do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), ciclo 2021-2023. Seus interesses de pesquisa são ciência de dados para sistemas de apoio à decisão, análise de mídia social e mercados eletrônicos.

\*\*\*

## APÊNDICE B – FORMULÁRIO DE AVALIAÇÃO

### Avaliação do protótipo - Sciensh

Este formulário refere-se a etapa de avaliação do projeto. As respostas serão usadas para mensurar a qualidade do protótipo desenvolvido.



fabiano.nhoatto@gmail.com (não compartilhado)

[Alternar conta](#)



**\*Obrigatório**

Nome do servidor \*

Sua resposta

Setor \*

Sua resposta

Cargo \*

Sua resposta





O protótipo se mostra promissor em:

	Discordo Plenamente	Discordo Parcialmente	Não concordo Nem discordo	Concordo Parcialmente	Concordo Plenamente
Calcular pontuações personalizadas de indicadores cientométricos.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Permitir maior agilidade na gestão da pesquisa e inovação na instituição.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Identificar os perfis e tendências de pesquisa da instituição.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Gerar relatórios dinâmicos das produções dos pesquisadores e suas relações de cooperação.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Prover informações e auxiliar no planejamento e a avaliação de políticas científicas.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Disponibilizar dados das atividades de pesquisa à comunidade.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Permitir a difusão das produções da	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>



instituição,  
principalmente  
para a  
sociedade.

Permitir maior  
eficiência no  
processo de  
análise da  
produtividade  
dos  
pesquisadores.

☐☐☐☐☐

Identificar a  
conexão de  
pesquisadores  
na instituição.

☐☐☐☐☐

Mensurar itens  
relevantes para  
a Gestão e a  
comunidade.

☐☐☐☐☐

Sugestões ou comentários

Sua resposta

Enviar

Limpar formulário

Nunca envie senhas pelo Formulários Google.

Este conteúdo não foi criado nem aprovado pelo Google. [Denunciar abuso](#) - [Termos de Serviço](#) - [Política de Privacidade](#)

Google Formulários



ANEXOS

## ANEXO A – EDITAL PESQUISA 2020



**UNIVERSIDADE FEDERAL DO OESTE DO PARÁ**  
**PRÓ-REITORIA DE PESQUISA, PÓS-GRADUAÇÃO E INOVAÇÃO TECNOLÓGICA**

## ANEXO I – FICHA DE PONTUAÇÃO

PRODUÇÕES DOCENTES	PONTOS	LIMITE MÁXIMO	Nº DE PRODUÇÕES
<b>A) ARTIGOS PUBLICADOS EM PERIÓDICOS DE ACORDO COM O MAIOR QUALIS DA REVISTA</b>		<b>500</b>	
Artigo completo em periódico A1	50		0
Artigo completo em periódico A2	45		0
Artigo completo em periódico B1	35		0
Artigo completo em periódico B2	25		0
Artigo completo em periódico B3	20		0
Artigo completo em periódico B4	15		0
Artigo completo em periódico B5	10		0
Artigo completo em periódico C ou sem extrato	5		0
<b>B) LIVROS E CAPÍTULOS DE LIVROS</b>		<b>200</b>	
Autoria de livro especializado (Edição internacional)	60		0
Autoria de livro especializado (Edição nacional/local)	50		0
Autoria de capítulo de livro especializado (Edição internacional)	40		0
Autoria de capítulo de livro especializado (Edição nacional/local)	30		0
Organização de livro especializado (Edição internacional)	20		0
Organização de livro especializado (Edição nacional/local)	10		0
<b>C) TRABALHOS EM EVENTOS*</b>		<b>300</b>	
Trabalho completo em evento Internacional	10		0
Trabalho completo em evento nacional	7		0
Trabalho completo em evento regional/local	5		0
Resumo ou resumo expandido em evento Internacional	7		0
Resumo ou resumo expandido em evento nacional	5		0
Resumo ou resumo expandido em evento regional/local	3		0
<b>D) PROPRIEDADE INTELECTUAL (COM REGISTRO OU PEDIDO DE PATENTE)</b>		<b>200</b>	
Processo ou técnica	40		0
Produto tecnológico	40		0
Produto de design	20		0
Software	20		0
<b>E) ORIENTAÇÕES NA PÓS-GRADUAÇÃO</b>		<b>200</b>	
Tese de doutorado concluída (orientador principal)	40		0
Tese de doutorado concluída (coorientador)	20		0
Orientação de doutorado em andamento (orientador principal)	10		0
Dissertação de mestrado concluída (orientador principal)	20		0
Dissertação de mestrado concluída (coorientador)	10		0
Orientação de mestrado em andamento (orientador principal)	5		0
Monografia de especialização concluída	5		0
<b>F) ORIENTAÇÕES NA GRADUAÇÃO</b>		<b>300</b>	
Trabalhos de conclusão de curso ou orientações de outra natureza concluídos	10		0



**UNIVERSIDADE FEDERAL DO OESTE DO PARÁ**  
**PRÓ-REITORIA DE PESQUISA, PÓS-GRADUAÇÃO E INOVAÇÃO TECNOLÓGICA**

**Planos anuais de iniciação científica, de extensão ou tecnológica concluídos	15		0
**Planos anuais de iniciação científica, de extensão ou tecnológica em andamento	5		0
Programa de Educação Tutorial concluído	15		0
<b>G) OUTRAS PRODUÇÕES</b>		<b>100</b>	
Coordenação de projeto de pesquisa	25		0
Participação em projeto de pesquisa	5		0
Prêmios científicos/ Artísticos	10		0
<b>TOTAL MÁXIMO</b>		<b>1800</b>	
** Incluindo Iniciação Jr.			<b>Nota final=</b>

## ANEXO B – QUESTIONÁRIO DE SATISFAÇÃO

21/07/2021

Avaliação do Edital 02/2020-PROPPIT-UFOPA- PIBIC, PIBITI e PIVIC

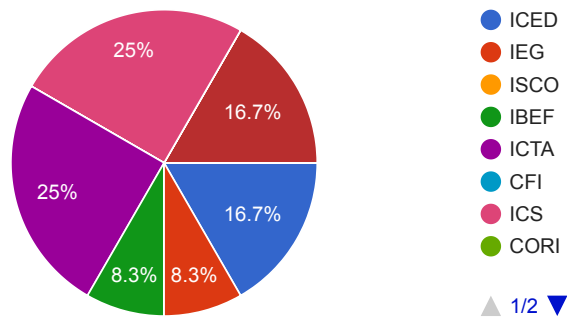
### Avaliação do Edital 02/2020-PROPPIT-UFOPA- PIBIC, PIBITI e PIVIC

12 responses

[Publish analytics](#)

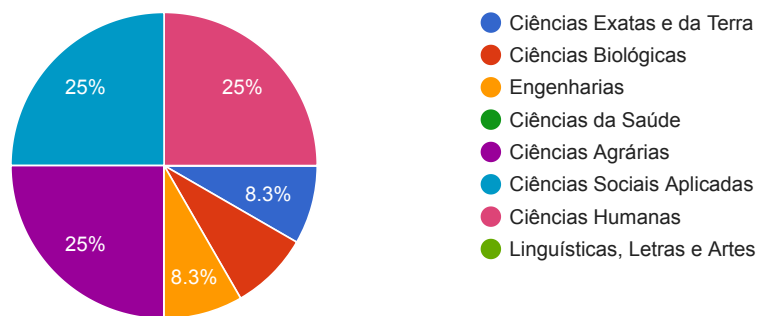
#### Qual a sua lotação?

12 responses



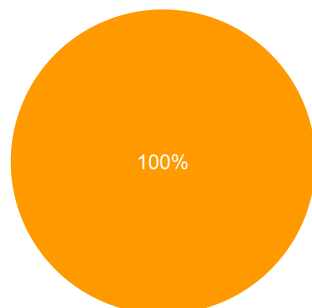
#### Em qual área de conhecimento você atua?

12 responses



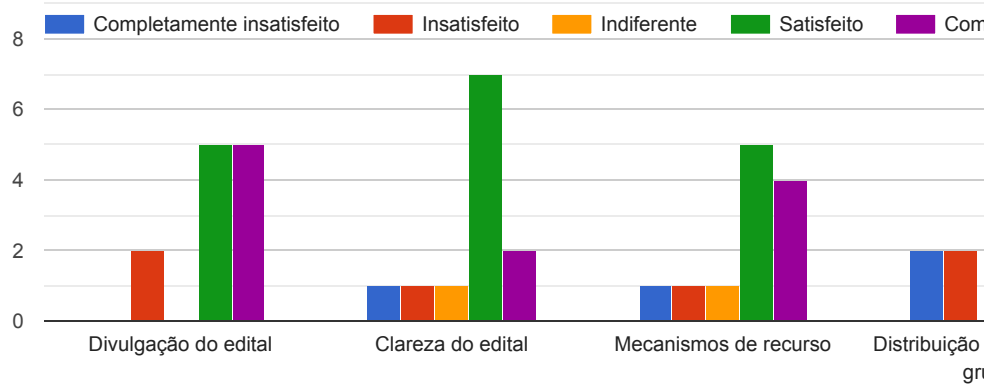
## Qual sua titulação?

12 respostas



- Especialização
- Mestrado
- Doutorado

## Como você avalia os seguintes itens?







## ANEXO C – REGISTRO DE SOFTWARE



REPÚBLICA FEDERATIVA DO BRASIL  
MINISTÉRIO DA ECONOMIA  
INSTITUTO NACIONAL DA PROPRIEDADE INDUSTRIAL  
DIRETORIA DE PATENTES, PROGRAMAS DE COMPUTADOR E TOPOGRAFIAS DE CIRCUITOS INTEGRADOS

## Certificado de Registro de Programa de Computador

Processo Nº: **BR512021003141-5**

O Instituto Nacional da Propriedade Industrial expede o presente certificado de registro de programa de computador, válido por 50 anos a partir de 1º de janeiro subsequente à data de 19/11/2021, em conformidade com o §2º, art. 2º da Lei 9.609, de 19 de Fevereiro de 1998.

**Título:** Sciensh: sistema de apoio à gestão da pesquisa e inovação baseado em cientometria

**Data de publicação:** 19/11/2021

**Data de criação:** 19/11/2021

**Titular(es):** UNIVERSIDADE FEDERAL DO OESTE DO PARÁ

**Autor(es):** FÁBIO MANOEL FRANÇA LOBATO; ÉFREN LOPES DE SOUZA; FABIANO PAULO NHOATTO

**Linguagem:** JAVA SCRIPT; PYTHON

**Campo de aplicação:** AD-02; AD-04; IF-07; IF-10; IN-02

**Tipo de programa:** AP-02; AT-02; AV-01; FA-01; GI-01

**Algoritmo hash:** SHA-256

**Resumo digital hash:** ac7e2030c384f55b5f1239f7e3f4a390e50eef3838ef01a4bc6d6a3c280fcc9d

**Expedido em:** 28/12/2021

**Aprovado por:**  
Carlos Alexandre Fernandes Silva  
Chefe da DIPTO