

UNIVERSIDADE FEDERAL DE ITAJUBÁ
INSTITUTO DE ENGENHARIA DE PRODUÇÃO E GESTÃO
MESTRADO PROFISSIONAL EM ADMINISTRAÇÃO (MPA)

EUNICE CRISTINA DA SILVA

PROPOSTA DE UM AMBIENTE DE *BUSINESS INTELLIGENCE* PARA
ACOMPANHAMENTO DOS EGRESSOS DOS PROGRAMAS DE INICIAÇÃO
CIENTÍFICA DO IFSULDEMINAS

Itajubá/MG

2023

EUNICE CRISTINA DA SILVA

**PROPOSTA DE UM AMBIENTE DE *BUSINESS INTELLIGENCE* PARA
ACOMPANHAMENTO DOS EGRESSOS DOS PROGRAMAS DE INICIAÇÃO
CIENTÍFICA DO IFSULDEMINAS**

Dissertação submetida ao Programa de Pós-Graduação em Administração da Universidade Federal de Itajubá, como requisito parcial para a obtenção do grau de Mestre em Administração.

Linha de pesquisa: Sistemas de Informação para Apoio à Decisão.

Orientador: Prof. Dr. Fábio Favaretto

Itajubá/MG

2023

Ficha Catalográfica

Ficha catalográfica elaborada automaticamente de acordo com os dados fornecidos pelo(a) autor(a).

EUNICE CRISTINA DA SILVA

PROPOSTA DE UM AMBIENTE DE *BUSINESS INTELLIGENCE* PARA
ACOMPANHAMENTO DOS EGRESSOS DOS PROGRAMAS DE INICIAÇÃO
CIENTÍFICA DO IFSULDEMINAS

Dissertação submetida ao Programa de Pós-Graduação em Administração da Universidade Federal de Itajubá, como requisito parcial para a obtenção do grau de Mestre em Administração.

Aprovado em 05/06/2023.

BANCA EXAMINADORA

Prof. Dr. Fábio Favaretto – presidente e orientador
Universidade Federal de Itajubá – UNIFEI

Prof. Dr. Everton Drohomeretski – membro
FAE Centro Universitário – FAE

Prof. Dr. Alexandre Ferreira de Pinho – membro
Universidade Federal de Itajubá – UNIFEI

Itajubá/MG

2023

AGRADECIMENTOS

Dedico este trecho para prestar meus agradecimentos a todos envolvidos direta ou indiretamente ao longo desta etapa da minha vida. Em primeiro lugar, presto minha gratidão ao meu orientador Prof. Fábio Favaretto, que esteve comigo ao longo do período, fornecendo conhecimentos e orientações indispensáveis para o desenvolvimento e sucesso deste trabalho, sempre se dispondo com muita paciência, gentileza e generosidade.

Aos professores Pinho e Everton, pelo aceite em fazer parte da banca examinadora, avaliando a dissertação e fornecendo contribuições indispensáveis para o resultado; muito obrigada pela forma como abordaram as sugestões.

A todos os docentes e técnicos administrativos do PPGMPA da UNIFEI.

A todos os colegas da turma e em especial aqueles que estão diariamente ao meu lado na reitoria. Aos amigos da Pró-Reitoria de Pesquisa, Pós-Graduação e Inovação (PPPI) do Instituto Federal de Educação, Ciência e Tecnologia do Sul de Minas Gerais (IFSULDEMINAS), que forneceram suporte profissional quando precisei me ausentar para o desenvolvimento da dissertação e suporte emocional quando achei que não seria possível terminar (obrigada, pessoal).

Meu Obrigada especial a profa. Sindynara Ferreira, ex-pró-reitora de Pesquisa, PG e Inovação, que me incentivou a buscar novos conhecimentos, nunca aceitando a minha negativa “esse momento é complicado, não tenho condições de me dedicar”, ela é minha principal incentivadora, sempre fornecendo conselhos valiosos como só uma amiga pode fazer. Agradecimento especial também a Kélica, Regiane, Renata, Rosimeire, Maria Rita e Gissélida, que com amizade ímpar, me incentivaram e não me deixaram desistir, por vezes acreditando em mim mais do que eu mesma, demonstrando prazer com meu amadurecimento; aos meus amigos que além de me ouvirem e apoiarem, propiciaram necessários momentos que ajudaram a aliviar o peso e cansaço nas fases difíceis dessa jornada.

Aos meus pais (*in memoriam*) que me ensinaram sobre a autonomia e a responsabilidade por meus atos; às minhas irmãs que sempre entenderam os momentos de estresse e torceram por mim, demonstrando paciência e muito amor. Ao Caio por tudo, principalmente por suportar minhas crises e estar ao meu lado sempre, você é minha pessoa favorita.

Por fim, ao IFSULDEMINAS que me proporcionou a oportunidade de cursar o mestrado por meio do convênio com a UNIFEI e pela liberação das horas para a ação de desenvolvimento em serviço.

A Deus.

*Você pode ter dados sem informação,
mas você não pode ter informação sem dados.*

(Daniel Keys Moran)

RESUMO

O Programa Institucional de Bolsas de Iniciação Científica tem como objetivos principais: contribuir para a formação de recursos humanos para a pesquisa, qualificar os estudantes para os programas de pós-graduação e contribuir para diminuir o tempo médio de permanência na pós-graduação. Logo, um sistema para acompanhar os egressos da Iniciação Científica (IC) deve fazer parte da gestão das IES, ao possibilitar, por meio de informações, manter o acompanhamento da trajetória acadêmica e profissional, a absorção desses egressos nos cursos de mestrado e doutorado e o impacto na sociedade, além de fornecer subsídios para o planejamento de ações institucionais. Para esse tipo de acompanhamento, a utilização de ferramentas de análise de dados é fundamental, dentre elas destaca-se o *Business Intelligence* (BI). O principal componente de um BI é o *Data Warehouse* (DW), que é um repositório de dados que tem por finalidade alimentar o BI, relatórios e funções analíticas gerando informações. Diante disso, este estudo teve como objetivo principal propor um ambiente de *Business Intelligence* para o acompanhamento dos egressos da iniciação científica do Instituto Federal de Educação, Ciência e Tecnologia do Sul de Minas (IFSULDEMINAS) com informações gerenciais que auxiliem os gestores na tomada de decisão no âmbito da iniciação científica, pesquisa e pós-graduação. Para atingir este objetivo o estudo foi conduzido pelo método de pesquisa-ação, descrevendo os passos desde a identificação do problema e requisitos informacionais até a proposta de solução, a qual é a proposição de um ambiente de BI para a Gestão dos Egressos da IC. Para identificar os requisitos foram realizadas reuniões com roteiro semiestruturado, pesquisa documental e bibliográfica e observação direta. A análise dos dados foi qualitativa e de cunho descritivo. O desenvolvimento do DW do BI foi baseado na metodologia de Ralph Kimball, e usado o software Power BI para a apresentação das informações em forma de dashboards. A ferramenta foi avaliada e validada pelos Gestores da IC do IFSULDEMINAS. Os resultados alcançados atingiram os objetivos propostos, apresentando informações relevantes para o acompanhamento de egressos da IC. O modelo dimensional de DW desenvolvido conseguiu transformar os dados em informações importantes, e o relatório de dashboard possibilitou demonstrar o impacto dos egressos da IC na formação de recursos humanos para pesquisa (mestres e doutores), bem como a trajetória profissional e acadêmica e o impacto social. A base de dados final construída, diante da sua magnitude de dimensões e dados, possibilita novas análises com as dimensões presentes, permitindo estabelecer cruzamentos não oportunizados pela direção do estudo. Como contribuição, além da social/econômica (formação de recursos humanos para a pesquisa), da prática/pessoal (desenvolvimento de uma ferramenta de BI como solução ao problema contextualizado), esta pesquisa contribui também no nível acadêmico/científico, pois o resultado vai além de um novo estudo sobre BI, contribui para preencher a lacuna da literatura sobre ferramentas de BI na Gestão de Egressos e estudos sobre BI em instituições públicas, contribuindo para a evolução da teoria.

Palavras-chave: Iniciação científica, *Data Warehouse*, Análise de dados, Acompanhamento de egressos, Pós-Graduação.

ABSTRACT

The Institutional Scientific Initiation Scholarship Program has the following main objectives: to contribute to the formation of human resources for research, to qualify students for graduate programs and to contribute to reducing the average length of stay in graduate school. Therefore, a system to follow up Scientific Initiation (CI) graduates should be part of the management of the HEIs, as it makes it possible, through information, to keep track of the academic and professional trajectory, the absorption of these graduates in master's and doctoral courses and the impact on society, in addition to providing subsidies for the planning of institutional actions. For this type of follow-up, the use of data analysis tools is fundamental, among which Business Intelligence (BI) stands out. The main component of a BI is the Data Warehouse (DW), which is a data repository whose purpose is to feed the BI, reports and analytical functions generating information. Therefore, this study had as main objective to propose a Data Warehouse environment for the monitoring of graduates of the scientific initiation of the Federal Institute of Education, Science and Technology of the South of Minas (IFSULDEMINAS) with managerial information that can help managers in taking decision in the scope of scientific initiation, research and graduate studies. To achieve this objective, the study was conducted using the action-research method, describing the steps from identifying the problem and informational requirements to proposing a solution, which is the proposition of a DW (dimensional model) for the Management of Graduates of IC, and the visualization of information through Dashboards. To identify the requirements, semi-structured interviews, documentary and bibliographical research and direct observation were carried out. Data analysis was qualitative and descriptive. The development of the BI DW was based on Ralph Kimball's methodology, and Power BI software was used to present data in the form of dashboards. The tool was evaluated and validated by CI Managers at IFSULDEMINAS. The results reached achieved the proposed objectives, presenting relevant information for the follow-up of IC graduates. The developed DW dimensional model was able to transform the data into important information, and the developed dashboard report allowed the analysis of the results, demonstrating the impact of CI graduates in the training of human resources for research (masters and doctors), as well as their trajectory professional and academic and social impact. The final database built, in view of its magnitude of dimensions and data, enables new analyzes with the present dimensions, allowing to establish intersections not provided by the direction of the study. As a contribution, in addition to the social/economic (training of human resources for research), practical/personal (development of a BI tool as a solution to the contextualized problem), this research also contributes at the academic/scientific level, as the result will in addition to a new study on BI, it contributes to filling the gap in the literature on BI tools in Graduate Management and studies on BI in public institutions, which contributes to the evolution of the theory.

Key words: Scientific initiation, Data Warehouse, Data analysis, Monitoring of graduates, Postgraduate studies.

LISTA DE FIGURAS

Figura 1 – Abordagem proposta por Bill Inmon	28
Figura 2 – Abordagem proposta por Ralph Kimball	29
Figura 3 – Modelo Dimensional <i>Star Schema</i>	31
Figura 4 – Modelo Dimensional <i>Snowflake Schema</i>	31
Figura 5 – Processo Extração, Transformação e Carregamento dos dados (ETL)	32
Figura 6 – Áreas com mais estudos na literatura sobre <i>Business Intelligence</i> na <i>Scopus</i> e <i>Web Of Science</i> até 2022	35
Figura 7 – Áreas com mais estudos na literatura sobre <i>Business Intelligence</i> no CTD/CAPES entre 2018 e 2022.	36
Figura 8 – Evolução dos estudos sobre BI - literatura internacional e nacional	36
Figura 9 – Unidades do Instituto Federal do Sul de Minas	42
Figura 10 – Fases da pesquisa-ação	44
Figura 11 – Ciclo de vida do <i>Data Warehouse</i>	47
Figura 12 – Abordagem do planejamento da pesquisa-ação	48
Figura 13 – Etapas do planejamento da pesquisa-ação (abordagem por meio da identificação de um problema)	49
Figura 14 – Modelo Conceitual do <i>Data Warehouse</i>	64
Figura 15 – Plano de ação para o desenvolvimento da proposta de solução	64
Figura 16 – Ciclo de vida do <i>Data Warehouse</i>	65
Figura 17 – Modelo Dimensional do <i>Data Warehouse</i>	66
Figura 18 – <i>Dashboard</i> : Trajetória Acadêmica	67
Figura 19 – <i>Dashboard</i> : Impacto do PIBIC na Pós-Graduação	68
Figura 20 – <i>Dashboard</i> : Trajetória Profissional	68
Figura 21 – <i>Dashboard</i> : Produção Científica	69
Figura 22 – <i>Dashboard</i> : Desenvolvimento Tecnológico e Regional	69
Figura 23 – <i>Dashboard</i> : Formação do Pesquisador	70
Figura 24 – <i>Dashboard</i> : Meninas na Ciência	70

LISTA DE QUADROS

Quadro 1	– Apresentação do resumo das técnicas de coleta de dados	51
Quadro 2	– Modelo de controle dos bolsistas de IC - DPPG/IFSULDEMINAS	52
Quadro 3	– Relação das decisões para nortear o Acompanhamento de Egressos	54
Quadro 4	– Requisitos para o acompanhamento de Egressos do IFSULDEMINAS	56
Quadro 5	– Requisitos e indicadores – Dados abertos da CAPES	60
Quadro 6	– Requisito e indicador – Dados da Plataforma Lattes	61
Quadro 7	– Requisitos e indicadores – Dados da CB/DPPG	62
Quadro 8	– Requisitos e indicadores – Dados do WebGiz/SUAP	62
Quadro 9	– Requisitos e indicadores – Dados da RAIS/MEI/LinkedIn	63
Quadro 10	– Requisitos e indicadores – Dados de outras bases	63
Quadro 11	– Relação das fontes de dados	65
Quadro 12	– Requisitos do DW de Acompanhamento de Egressos da IC	76

LISTA DE TABELAS

Tabela 1	– Síntese do resultado de seleção dos estudos	37
Tabela 2	– Distribuição dos Egressos da IC no IFSULDEMINAS	56

LISTA DE SIGLAS E ABREVIATURAS

BI	<i>Business Intelligence</i>
CAPES	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
CGEE	Centro de Gestão e Estudos Estratégicos
CNPq	Conselho Nacional de Desenvolvimento Científico e Tecnológico
C&T	Ciência e Tecnologia
CT&I	Ciência, Tecnologia e Inovação
CB	Coordenação de Bolsas
CNE	Conselho Nacional de Educação
DDI	Diretoria de Desenvolvimento Institucional
DTIC	Diretoria de Tecnologia da Informação
DM	<i>Data Mart</i>
DW	<i>Data Warehouse</i>
DPPG	Diretoria de Pesquisa e Pós-Graduação
ETL	Processo de extração, transformação e carregamento
FAP	Fundação de Apoio a Pesquisa
FAPEMIG	Fundação de Amparo à Pesquisa do Estado de Minas Gerais
IC	Iniciação Científica
IES	Instituições de Ensino Superior
IFES	Instituições Federais de Ensino Superior
IFSULDEMINAS	Instituto Federal de Educação, Ciência e Tecnologia do Sul de Minas Gerais
MCTI	Ministério da Ciência, Tecnologia e Inovações
OLAP	Processamento Analítico On-Line
PIBIC	Programa Institucional de Bolsas de Iniciação Científica
PG	Pós-Graduação
PNE	Plano Nacional de Educação
PNPG	Plano Nacional de Pós-Graduação
PPG	Programa de Pós-Graduação
PROAD	Pró-Reitoria de Administração
PROEN	Pró-Reitoria de Ensino
PROEX	Pró-Reitoria de Extensão
PROGEP	Pró-Reitoria de Gestão de Pessoas

PPPI	Pró-Reitoria de Pesquisa, Pós-Graduação e Inovação
RH	Recursos Humanos
RN	Norma Reguladora
SI	Sistema de Informação
TI	Tecnologia da Informação
UNESCO	Organização das Nações Unidas para a Educação, Ciência e Cultura
WOS	<i>Web of Science</i>

SUMÁRIO

1 INTRODUÇÃO	15
2 REFERENCIAL TEÓRICO	19
2.1 Iniciação Científica	19
2.1.1 O Programa Institucional de Bolsas de Iniciação Científica (PIBIC) do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) na política de iniciação científica no Brasil	20
2.1.2 Contribuição da iniciação científica na graduação	22
2.1.3 Contribuições da Iniciação Científica para a pós-graduação: Plano Nacional de Educação e Plano Nacional de Pós-Graduação	23
2.2 Sistemas de Informação / <i>Business Intelligence</i>	24
2.2.1 <i>Data Warehouse</i> (DW)	26
2.2.1.1 Modelagem Dimensional	30
2.2.1.2 Processo <i>Extract, Transform and Load</i> (ETL)	32
2.2.1.3 Ferramentas OLAP para visualização e análise dos dados	32
2.3 Trabalhos Relacionados	34
2.3.1 Desenvolvimento de ferramentas de Business Intelligence em Universidades	34
2.3.2 Sistemas de Informação e o Acompanhamento de Egressos	39
3 METODOLOGIA	40
3.1 Objeto de estudo	41
3.2 Classificação/Tipo da pesquisa	43
3.3 Pesquisa-ação	44
3.4 Proposta metodológica: Fases do desenvolvimento da pesquisa-ação e técnicas de coleta e análise dos dados	45
4 DESENVOLVIMENTO	48
4.1 Fase 1 – Planejamento da pesquisa-ação	48
4.1.1 Definir contexto e propósito	49
4.1.2 Definir estrutura conceitual-teórica	50
4.1.3 Selecionar a unidade de análise e técnicas de coleta de dados	51
4.2 Fase 2 – Coleta dos dados	52
4.3 Fase 3 – Análise dos dados e proposta de plano de ação	59
4.4 Fase 4 – Implementação do plano de ação – Desenvolvimento do <i>Data Warehouse</i>	64
4.5 Fase 5 – Avaliação da ferramenta de <i>Business Intelligence</i>	71
5 DISCUSSÃO	73
6 CONSIDERAÇÕES FINAIS	80
REFERÊNCIAS	84
APÊNDICE A – Roteiro semiestruturado	91

APÊNDICE B – Estudos selecionados nas revisões integrativas	92
APÊNDICE C – Protocolo de coleta de dados	94
ANEXO A – Anuência da Instituição para a realização da pesquisa e a coleta dos dados	96

1 INTRODUÇÃO

Acompanhar e monitorar é parte da rotina de administrar nas organizações em geral e sobretudo nas Instituições de Ensino Superior (IES). As IES possuem um importante papel no desenvolvimento da sociedade por meio da formação profissional e devem se preocupar com a necessidade de avaliar e acompanhar o resultado da formação e o impacto na sociedade, sendo importante entender e analisar o destino dos seus egressos, estabelecer e manter um banco de dados, sobretudo no que tange à informação sobre a colocação profissional dos titulados. A educação de qualidade tornou-se uma questão crucial para os formuladores de políticas públicas; no âmbito de políticas públicas e científicas, no Brasil, destacam-se os Programas Institucionais de Bolsas de Iniciação Científica (PIBIC) implementados pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). O PIBIC é “voltado para o desenvolvimento do pensamento científico e iniciação à pesquisa de estudantes de graduação do ensino superior” (CNPq, 2006). Entre os seus objetivos está o de contribuir para a formação de recursos humanos e encaminhar os ex-bolsistas aos programas de Pós-Graduação (PG) esperando que a integralização do curso seja em menor tempo.

Araújo e Andriola (2020) relataram que os resultados de programa ligados à Iniciação Científica (IC) devem ser analisados de forma periódica, para que se possa identificar a evolução e a concretização dos objetivos destes programas, como a formação de recursos humanos para a pesquisa.

Houve progresso nos estudos sobre IC e inúmeras pesquisas sobre egressos da pós-graduação (PG) no mundo do trabalho, mas poucas pesquisas que cruzassem os descritores PIBIC, Pós-Graduação e formação de recursos humanos para a pesquisa (DE ALBUQUERQUE *et al.*, 2022; SANTOS, 2021; SANCHEZ, 2019). Nas pesquisas que analisam o impacto do PIBIC encontram-se sugestões de implantação de Sistemas de Informação, bem como a importância da organização e análise dos dados dos egressos (SILVA, 2021; DANTAS, 2019; CANUTO, 2018), demonstrando lacuna na literatura científica nessa área.

O uso de um Sistema de Informação (SI) para acompanhar os egressos da IC deve ser parte da gestão das IES de forma permanente, pois, por meio do registro de informações, é possível manter o acompanhamento da trajetória dos egressos e oferecer subsídios para o planejamento de ações institucionais. Para isso, é indispensável utilizar ferramentas de análise massiva de dados. Com o desenvolvimento da tecnologia e a quantidade de dados produzidos torna-se necessário o desenvolvimento de SIs e ferramentas que possam extrair e analisar as informações, de forma rápida, para auxiliar na tomada de decisões. A divulgação da informação

é o elo entre a IES e a sociedade, e sua viabilidade depende da capacidade de seu gerenciamento em bancos de dados que possam reuni-la e divulgá-la. Dentre as ferramentas de análise de dados está o *Business Intelligence* (BI); assim, a implantação de uma ferramenta de BI possibilita uma análise mais ampla e atualizada sobre o acompanhamento de egressos da IC.

No setor público, a implantação do BI ainda é muito incipiente, poucos órgãos da Administração Pública possuem essa ferramenta de auxílio à gestão e tomada de decisão implementada. As (IES) estão investindo, gradativamente, em BI, para atender à crescente demanda por informações (STEWART, 2022). O recente relatório da Educause (2020) revelou as 10 principais tecnologias estratégicas nas quais as IES dos EUA estão investindo, e a maioria dessas tendências tecnológicas depende de soluções de BI e análise de dados.

O Grupo Gartner conceitua BI como um termo abrangente que inclui aplicações, infraestruturas, ferramentas e práticas que permitem o acesso e análise de informação para melhorar e otimizar decisões e o desempenho (GARTNER, 2019). O principal componente na arquitetura de um BI é um repositório de armazenagem de dados, que organiza, centraliza e consolida dados e informações coletados de diferentes sistemas transacionais e é otimizado para consultas (LAUDON; LAUDON, 2022); esse repositório é denominado de *Data Warehouse* (DW).

Entre as literaturas que desenvolvem ferramentas de BI no acompanhamento de egressos destacam-se os estudos de Ismail *et al.* (2022), Blanco Rojas *et al.* (2016), Hamdane *et al.* (2022) e Niño (2020). Blanco Rojas *et al.* (2016) construíram uma base de dados atualizada com informações da população em estudo (egressos) provenientes das redes sociais (*LinkedIn, Facebook e Twitter*). Devido à ausência de um sistema da informação (SI) para fornecer indicadores e estatísticas confiáveis e centralizados sobre a empregabilidade dos graduados universitários marroquinos, Hamdane *et al.* (2022) propuseram um *Data Warehouse* acadêmico como uma ferramenta poderosa e confiável para conectar os egressos da universidade com o mundo profissional usando a mineração de dados. No entanto, não houve um estudo que contemplasse a análise de dados secundários sobre a empregabilidade formal e informal, o impacto do PIBIC e o desenvolvimento regional.

O PIBIC do Instituto Federal de Educação, Ciência e Tecnologia do Sul de Minas Gerais (IFSULDEMINAS) precisa de ferramentas sistêmicas que apresentem o acompanhamento do PIBIC sobre o encaminhamento dos egressos à PG e seu destino profissional, e para essa operacionalização é necessário que se tenha, primeiramente, uma fonte geradora de dados e que esses dados estejam disponíveis para consulta pelos gestores; assim é necessário o desenvolvimento de uma solução capaz de extrair dados de diferentes fontes na IES e fora dela,

usando como estratégia a integração de dados (secundários) que permita obter informações de qualidade, em tempo hábil e que responda aos requisitos e necessidades da IES.

A pesquisa surge, então, com a seguinte pergunta: quais os requisitos necessários para subsidiar a criação de um ambiente de *Business Intelligence* (BI) para o acompanhamento de egressos dos programas de iniciação científica?

Diante da problemática apresentada, seguem, em seguida, os objetivos gerais e específicos deste estudo.

O objetivo geral é propor um ambiente de *Business Intelligence* (BI) para o acompanhamento dos egressos da iniciação científica do IFSULDEMINAS com informações gerenciais que auxiliem os gestores na tomada de decisão no âmbito da iniciação científica, pesquisa e pós-graduação.

Para cumprir o objetivo geral foram elencados os seguintes objetivos específicos: i) identificar os requisitos informacionais necessários para subsidiar a criação do ambiente de BI para o acompanhamento dos egressos da IC; ii) elencar os indicadores necessários para o acompanhamento dos egressos da IC, principalmente aqueles relacionados ao impacto do PIBIC na pós-graduação *Stricto sensu*, trajetória profissional e desenvolvimento tecnológico e regional; iii) identificar as fontes de informações necessárias ao processo de tomada de decisão; iv) mapear e levantar os dados dos egressos do PIBIC/CNPq do IFSULDEMINAS, e v) desenvolver o DW do BI e, por meio do software Power BI, um painel (*Dashboard*) para exploração das informações e apresentação para os usuários finais.

A relevância deste estudo se justifica: i) diante da importância do ensino superior para o desenvolvimento do país e à necessidade das IES responderem à sociedade sobre os resultados de seus processos formativos; ii) da escassez de estudos sobre egressos da IC e sobre a utilização de BI no acompanhamento de egressos e em instituições de ensino superior; e iii) pela ausência de acompanhamento dos resultados do desempenho do PIBIC no IFSULDEMINAS.

A premissa é que as informações apresentadas pelo BI podem se tornar um importante instrumento na tomada de decisão da Gestão em relação ao investimento/captação de novas bolsas, e um diferencial na prestação de contas com o CNPq e agências de fomento, o que pode acarretar aumento na quantidade de bolsas da cota institucional. Justifica-se ainda que a implantação de um SI para divulgação e acesso aos dados sobre a evolução dos egressos da IC na pós-graduação significa a materialização desta evolução, que, ao centralizar as informações em um sistema, torna atrativo o universo da pesquisa científica, apresentando a carreira científica (em IES ou empresas) como uma possibilidade de atuação profissional real e próxima.

Como contribuição prática/profissional, a ferramenta de BI, para acompanhamento de egressos da IC, permitirá um maior conhecimento sobre a formação de recursos humanos para a pesquisa científica. A contribuição acadêmica/científica vai muito além de um novo estudo sobre BI, ao contribuir na lacuna da literatura sobre ferramentas de BI para a Gestão de Egressos, como destacado na revisão bibliométrica apresentada, contribuindo para a evolução da teoria; e por fim, a contribuição econômica e social demonstra o impacto de políticas públicas em educação na região foco do estudo.

A pesquisa se dará em um universo restrito aos egressos do PIBIC do IFSULDEMINAS e será tratado como enfoque o PIBIC promovido pelo CNPq, devido à quantidade e qualidade dos dados disponíveis e à sua abrangência nacional, tendo como população os egressos do PIBIC que foram bolsistas entre os anos de 2011¹ e 2021.

Esta dissertação foi estruturada em seis seções. A primeira seção, a introdução do trabalho, apresenta o tema ao leitor com seus objetivos, além da justificativa para a realização da pesquisa. A segunda seção apresenta a fundamentação teórica, com a pesquisa bibliográfica realizada acerca dos conceitos que edificaram a pesquisa: Iniciação Científica, Sistemas de Informação, *Business Intelligence* e *Data Warehouse* (que embasaram as etapas do desenvolvimento do DW) e Trabalhos Relacionados. Na terceira seção foi descrito o desenho da pesquisa, a metodologia adotada ao longo do estudo, desde a descrição do objeto de estudo, sua caracterização, as técnicas de coleta e análise dos dados e as fases da metodologia usada no desenvolvimento da pesquisa. A quarta seção apresenta o desenvolvimento da proposta de solução para o problema apresentado, ou seja, descreve a identificação dos requisitos, o desenvolvimento do DW do BI, a visualização das informações por meio do software Power BI e sua validação pela Gestão. Na quinta seção foram discutidos os resultados encontrados, e a sexta apresenta as considerações finais sobre o modelo desenvolvido, as limitações do estudo, as contribuições e as sugestões de pesquisas futuras; na sequência, são listados as referências utilizadas, os apêndices e os anexos.

¹ Aprovação na chamada de cotas de bolsas do PIBIC/CNPq no IFSULDEMINAS.

2 REFERENCIAL TEÓRICO

Nesta seção apresentam-se os conceitos e o estado da arte para evidenciar o processo de formação de recursos humanos para a pesquisa, bem como a relevância do PIBIC do CNPq nesse processo.

Em seguida, são apresentados os conceitos relacionados a sistemas de informação, e *Business Intelligence*, destacando a importância da informação para a tomada de decisão; após são apresentados os conceitos de um DW e seu processo de desenvolvimento para embasar o método escolhido na intervenção prática da pesquisa.

Para finalizar o referencial teórico, são apresentadas: a revisão bibliométrica sobre o desenvolvimento de ferramentas de *Business Intelligence* em IES, e uma revisão narrativa sobre Sistemas de Informação no Acompanhamento de Egressos.

2.1 INICIAÇÃO CIENTÍFICA

As IES, que historicamente direcionam seus esforços na formação de cidadãos e profissionais das diversas áreas do conhecimento, vem se destacando há alguns anos como um ambiente de iniciação para jovens estudantes candidatos a pesquisadores (carreira científica). A partir de 1950 observou-se o início da inserção dos discentes de graduação em atividades relacionadas à pesquisa científica, com o intuito de estimular e preparar novos talentos para a pesquisa.

Para Pelinski (2021) a IC disponibiliza aos estudantes um conjunto de conhecimentos necessários para iniciar e desenvolver pesquisas acadêmicas; para Massi e Queiroz (2015), pode ser vista sob dois pontos de vista no ensino superior: no primeiro, é entendida como um processo das vivências do estudante durante o período da graduação (programas de treinamento, fichamentos, desenvolvimento de metodologia científica, visitas programadas, etc.); no segundo, e o que é adotado nesta pesquisa, a IC é definida como a integração do estudante em um ambiente de pesquisa, com ou sem o recebimento de contrapartida financeira (bolsa), por meio do desenvolvimento de atividades científicas em projetos de pesquisa, e sempre sob a orientação de um pesquisador/docente da IES.

Ao longo dos anos, os objetivos da IC se diversificaram sendo ampliados, evoluindo para um programa financiado pelo governo federal. A promoção e financiamento pelo governo federal de uma política institucional nacional de IC é um importante diferencial do Brasil perante outros países, onde essa atividade, na maioria das vezes, decorre de iniciativas individuais de docentes. A política institucional de IC no Brasil teve avanço sistemático e continuidade ao longo de mais de três décadas, sendo um ótimo exemplo.

Para que a ciência se desenvolva constantemente é necessário que o conhecimento seja ampliado, corrigido ou aprimorado, o que ocorre por meio dos resultados de novas pesquisas (MUELLER, 1995). A pesquisa científica é uma das principais bases para resolver problemas e conflitos. Os resultados das pesquisas científicas contribuem tanto na busca pelo bem-estar coletivo e de melhorias nas condições de vida, como para o progresso da humanidade. Logo, todo investimento em ciência deve ser considerado prioridade entre as políticas dos governos.

2.1.1 O Programa Institucional de Bolsas de Iniciação Científica (PIBIC) do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) na política de iniciação científica no Brasil

O CNPq, criado pela Lei 1.310, em 15 de janeiro de 1951, tem como objetivo incentivar o desenvolvimento da pesquisa científica e tecnológica em qualquer área do conhecimento (CNPq, 2023a). É o principal órgão brasileiro de fomento à IC.

Segundo Carvalho (2002, p. 145),

O surgimento da iniciação científica nas universidades fundamentou-se em três fatores, a saber: o surgimento do CNPq, a expansão do sistema de ensino superior e a consolidação da Pós-Graduação. Esses três fatores, então, [...] criaram as condições para que a pesquisa científica crescesse no âmbito das instituições de ensino superior, possibilitando, a partir daí o surgimento dos programas de iniciação científica no cenário nacional.

A política científica no Brasil, por meio do estímulo para o desenvolvimento da pesquisa científica, acabou sendo prerrogativa, principalmente, do CNPq. Esse estímulo ocorre por meio da concessão de recursos para a pesquisa (financiamento de bolsas para a formação de pesquisadores), e desenvolvimento de pesquisas (financiamento de despesas de custeio e capital).

O CNPq foi fundamental na orientação de um modelo nacional de formação de pesquisadores (recursos humanos). Desde 1951, houve investimentos em bolsas de IC visando preparar os então discentes de graduação para a pós-graduação. Segundo Oliveira e Bianchetti (2019), nos 12 primeiros anos do CNPq houve uma instabilidade nos investimentos em bolsas de pesquisa no país. Esta instabilidade ocorreu devido a uma política articulada para a área da Ciência e Tecnologia (C&T), e, ao mesmo tempo, à instabilidade política, uma característica do período.

O CNPq sofreu alterações estruturais ao longo dos anos e, hoje, possui natureza jurídica de fundação vinculada ao Ministério da Ciência, Tecnologia e Inovação (MCTI). A sua atuação contribuiu e contribui para o progresso nacional e para o reconhecimento dos pesquisadores e das instituições de pesquisa brasileiras pela comunidade científica internacional (CNPq,

2023b). É importante mencionar que, além do CNPq, outras agências de fomento, inclusive estaduais, atuam no financiamento de pesquisas científicas e/ou tecnológicas e incentivam e contribuem para a formação de pesquisadores brasileiros, como, por exemplo, as Fundações de Amparo à Pesquisa (FAPs). Contudo, no presente trabalho, optou-se por limitar o objeto de estudo ao PIBIC financiado pelo CNPq, cujos resultados exitosos foram salientados em estudo realizado pelo CGEE (CGEE, 2017).

No dia 20 de julho de 1988, o Conselho Deliberativo do CNPq oficializou a criação do Programa Institucional de Bolsas de Iniciação Científica (PIBIC). Alguns estudiosos do tema indicam que esse programa teve inspiração em ações de apoio à graduação implementadas nos Estados Unidos, em especial no “Research Experience for Undergraduates – REU”, um programa da National Science Foundation (NSF) que, desde 1986, contribui para o envolvimento dos estudantes de graduação em programas de pesquisa e que possui algumas características semelhantes ao PIBIC (DANTAS, 2019, p. 38).

Após a análise da evolução das normas reguladoras do PIBIC, identificou-se um esforço na elaboração de políticas públicas para a formação científica de base (recursos humanos para a pesquisa) e que essas políticas impactassem diretamente na pós-graduação (PG), ficando isso claro, a partir de 1996, com a atualização das normas e na qual um dos principais objetivos do PIBIC é o de contribuir para a redução do tempo médio de permanência do ex-bolsista de IC na pós-graduação (CNPq, 2006).

De acordo com Trombelli (2013), o PIBIC do CNPq foi criado com a intenção de proporcionar benefícios a longo prazo, esperando uma possível economia de recursos financeiros na qualificação dos pesquisadores na PG, sobretudo no ingresso dos ex-bolsistas de IC na pós-graduação, que, na teoria, deveriam concluir o curso em menor tempo.

Hoje, orientado pela RN-017/2006-CNPq², o PIBIC é um programa voltado para o desenvolvimento do pensamento científico e a iniciação à pesquisa de estudantes de graduação do ensino superior. O programa tem como objetivos gerais: a) contribuir para a formação de recursos humanos para a pesquisa; b) contribuir para a formação científica de recursos humanos que se dedicarão a qualquer atividade profissional; e c) contribuir para reduzir o tempo médio de permanência dos estudantes na pós-graduação (CNPq, 2006).

² Fonte: CNPq. Anexo III - RN-017/2006 - Bolsas por Quota no País. Disponível em: http://memoria2.cnpq.br/web/guest/view/-/journal_content/56_INSTANCE_0oED/10157/100352?COMPANY_ID=10132. Acesso em 22 abr. 2023.

2.1.2 Contribuição da iniciação científica na graduação

Na literatura sobre IC, há diversos estudos que analisaram as contribuições na vida acadêmica do estudante (PELINSKI, 2021; MOURA *et al.*, 2020; TOZATO *et al.*, 2020; PINHO, 2017; CGEE, 2017).

Para Pelinski (2021) a IC é um programa que disponibiliza aos estudantes um conjunto de conhecimentos indispensáveis para iniciar e desenvolver pesquisas acadêmicas. Moura *et al.* (2020) analisaram se o PIBIC/CNPq teve influência sobre os bolsistas egressos que permaneceram na área científica. Os resultados indicaram que os egressos adquiriram capital científico, o que favoreceu a entrada e a permanência no campo científico. Para Pinho (2017, p. 665), a atividade de pesquisa na graduação “[...] é considerada, por professores e acadêmicos, uma atividade motivadora e enriquecedora, que torna a formação mais abrangente e dota, os graduandos, com maior capacidade de observação e crítica”.

Entre os resultados das pesquisas as principais contribuições são: desenvolvimento ao nível pessoal dos estudantes (habilidades), como a autonomia, criatividade, maturidade e pensamento crítico e responsabilidade. Além disso, no que diz respeito à formação de pesquisadores e envolvimento com pesquisas, foi ressaltado o papel dos congressos e seminários de iniciação científica, onde é possível obter visibilidade do trabalho em desenvolvimento, socializar o conhecimento com outros pesquisadores e conseguir feedback para melhorias nos projetos (PELINSKI, 2021; MOURA *et al.*, 2020; TOZATO *et al.*, 2020; PINHO, 2017; CGEE, 2021).

Divulgado por meio do relatório “A Formação de novos quadros para CT&I: avaliação do programa institucional de bolsas de iniciação científica (PIBIC)”, estudo realizado pelo CGEE em 2017, em âmbito nacional, permitiu conhecer um panorama da IC, apontando-a como uma vivência de sucesso na formação acadêmica e pessoal do discente de graduação e no seu encaminhamento para a formação de recursos humanos para a pesquisa. O relatório também apontou diferenças estatisticamente significativas nos índices de desempenho acadêmico dos discentes de graduação que participaram do PIBIC e nos que não participaram do programa (CGEE, 2017).

Diante do exposto, pode-se afirmar que o PIBIC é o principal programa de formação de pesquisadores durante o período de graduação dos estudantes.

2.1.3 Contribuição da Iniciação Científica para a pós-graduação: Plano Nacional de Educação e Plano Nacional de Pós-Graduação

Com o objetivo de melhorar a qualidade da PG, o Plano Nacional de Educação (PNE) e o Plano Nacional de Pós-Graduação (PNPG) têm atuado para oportunizar a educação científica nas IES. O PNE, aprovado em 2014 para o período de 2014/2024, por meio da Lei nº 13.005, de 25 de junho de 2014 (BRASIL, 2014), traçou estratégias e 20 metas para orientar a melhoria da qualidade da educação no Brasil. Dentre as metas, diretamente ligadas a esse estudo, destaca-se a meta 14 “Elevar gradualmente o número de matrículas na pós-graduação, de modo a atingir a titulação anual de 60 mil mestres e 25 mil doutores” (BRASIL, 2014).

Aprovada pela CAPES, em 10 de outubro de 2018, a Proposta de Aprimoramento do Modelo de Avaliação da PG, elaborada pela Comissão Nacional de Acompanhamento do PNPG, identificou o PIBIC do CNPq como uma política pública relevante e de elevado destaque no Brasil, e que o PIBIC constitui a base (egressos) de alimentação da PG brasileira (CAPES, 2018).

A universidade é o local de produção e divulgação do conhecimento científico. No Brasil, a geração de novos conhecimentos por meio de atividades de pesquisa ocorre, sobretudo, no âmbito da pós-graduação, que é, atualmente, responsável pela maior parte das pesquisas realizadas no Brasil (BRAMBILLA, 2011). No Brasil, os cursos de PG são fundamentais para o desenvolvimento da pesquisa científica. Além disso, em geral, os pesquisadores brasileiros são, tradicionalmente, egressos dos cursos de mestrado e doutorado. Portanto, a pós-graduação tem um papel estratégico e se baseia na formação e preparação de recursos humanos de alto nível (CABRERO, 2007).

De Albuquerque *et al.* (2022) demonstraram a contribuição do PIBIC/ICMBio na formação de pesquisadores na área de conservação da biodiversidade, tanto na área acadêmica quanto para o mercado de trabalho em geral. Santos (2021) apresentou o impacto do PIBIC na inserção dos beneficiários em programas de pós-graduação no Brasil.

Moura *et al.* (2020) analisaram se o PIBIC/CNPq teve influência sobre os bolsistas egressos a permanecerem na área científica. A tese de Sanchez (2019) teve por objetivo conhecer os percursos (profissionais e acadêmicas) dos egressos (mestres/doutores) do PPG em Educação da Universidade Estadual de Campinas (UNICAMP). A pesquisa identificou que a carreira acadêmica é o destino preferencial dos egressos do mestrado/doutorado do PPG e também padrões de deslocamento geográficos diferentes para os mestres e doutores: enquanto os percursos formativos e a trajetória profissional dos mestres apresentam um caráter local/regional, os dos doutores se caracterizam por uma maior dispersão desde a formação no

nível da graduação até o ingresso na carreira docente após a obtenção do título (SANCHEZ, 2019, p. 113).

Para Souza (2018) os estudantes IC registraram taxas melhores em todas as etapas da seleção dos processos de mestrado, devido à formação diferenciada que os capacita para que se sobressaiam nas provas escritas, entrevistas, análises de projetos e avaliação curricular, culminando na aprovação geral. Costa (2016) analisou, em sua pesquisa, o destino profissional dos egressos do PIBIC da Universidade Federal de Santa Catarina (UFSC), entre 1990 a 2012, bem como a contribuição do programa no processo de formação de pesquisadores. Galvão *et al.* (2017), analisaram a inserção profissional dos titulados mestres e doutores no mercado de trabalho, sem a pretensão de considerar questões que dizem respeito a motivações pessoais, expectativas, avaliação dos cursos, entre outros. Lima (2016) investigou as diferenças entre titulados dos PPGs na UFRGS em 2012 que fizeram IC e aqueles que não tiveram experiência em IC (NIC) na graduação. Foram investigados aspectos como tempo de integralização, idade de titulação e diferenças na produção científica. Em geral, o grupo IC se titulou mais jovem que o NIC. Em Pires (2012), foi discutido os aspectos relacionados ao desenvolvimento e formação do egresso do PIBIC/UNEB como professor pesquisador do ensino superior.

Apesar da recomendação constante do CNPq para que as IES realizem pesquisas que permitam acompanhar a eficiência do PIBIC, na maioria das IES não se percebe, em forma de relatórios/portais, a análise do impacto da IC na trajetória formativa de seus egressos.

2.2 SISTEMAS DE INFORMAÇÃO / BUSINESS INTELLIGENCE

Antes do surgimento dos computadores, as instituições já utilizavam o conceito de Sistemas de Informação (SI) para o gerenciamento e análise da produtividade, mas os administradores gastavam muito tempo e esforço para organizar e juntar as informações devido ao excesso de dados armazenados (STAIR; REYNOLDS, 2020). Com o tempo e a revolução do conhecimento, da informação e da tecnologia, os processos foram sendo minimizados e as instituições passaram a usar sistemas computadorizados, tendo acesso à manipulação, armazenamento e distribuição da informação de forma rápida e sem transtornos (LAUDON; LAUDON, 2022).

De acordo com O'Brien (2011), os SI são fundamentais para o desempenho da organização, para auxiliar na tomada de decisões e em estratégias para alcançar vantagem competitiva. Neste sentido, os SI são uma ferramenta de apoio nos processos de decisão, tanto na gestão de dados quanto no acompanhamento e formulação dos resultados das organizações (O'BRIEN, 2011). Segundo Laudon e Laudon (2022, p. 34), "um sistema de informação (SI)

pode ser definido tecnicamente como um conjunto de componentes inter-relacionados que coletam (ou recuperam), processam, armazenam e distribuem informações destinadas a apoiar a tomada de decisões, a coordenação e o controle em uma organização”. Para entender esse conceito é necessário compreender as definições e diferenças entre dados e informação.

Segundo os mesmos autores, dados são compreendidos como uma sequência de fatos que ocorrem nas instituições que ainda não foram organizados nem analisados; informação são os dados já modelados, isto é, a informação é obtida quando os dados são tratados e organizados para que as pessoas possam entender, analisar e utilizá-los. Um dado é uma representação simbólica, numérica, algorítmica e outros tipos de um atributo ou variável, que pode ser quantitativa ou qualitativa. Os dados descrevem fatos empíricos, eventos ou entidades isoladas. É importante notar que dados isolados podem não conter informações relevantes (EGAS, 2015).

Para as informações serem úteis, é necessário ter conhecimento, que representa a compreensão de informações e de como elas podem ser úteis para apoiar uma tarefa ou decisão (STAIR; REYNOLDS, 2020). Para auxiliar na tomada de decisão, para torná-la o mais assertivo possível, é fundamental o uso de SI capazes de apresentar soluções baseadas em fatores de decisão. Nesse contexto e com a quantidade e disponibilidade de dados, as instituições veem um ativo na informação, e saber manipular e extrair valor desses dados é essencial. A análise de dados e os sistemas de informação são a base para insights acionáveis e facilitam decisões baseadas em dados (KUSHWAHA *et al.*, 2021).

Nos anos 1980, surgiu o conceito de Sistemas de Informações Executivas, ampliando o suporte computadorizado a outros níveis nas organizações, proporcionando, mesmo que de forma incipiente, acesso a informações importantes como: prognósticos e previsões, análise de tendências, detalhamento e acesso a status permanecendo até meados da década de 1990. Assim, todas as informações necessárias à administração podem estar em um sistema baseado em *Business Intelligence* (BI), conceito evoluído dos Sistemas de Informações Executivas (TURBAN *et al.*, 2009).

O conceito de BI e seu uso como solução técnica para fornecer informações ocorreu pela primeira vez em 1958 em um artigo publicado pela IBM *Journal of research and development* (DRAKE, 2018). O artigo descreve o BI como um sistema automático para disseminar informações para as várias divisões de qualquer organização industrial, científica ou governamental (DRAKE, 2018; LUHN, 1958),

Em 1989, Howner Dresner, definiu BI como um termo “guarda-chuva”, que retrata conceitos e métodos, para auxílio à tomada de decisão e com suporte de sistemas dotados de informações retroalimentadas rotineiramente (GARTNER, 2013).

Para Trieu (2017) BI é um termo genérico que abrange as tecnologias, processos e conceitos para melhorar a tomada de decisões. Para Drake (2018) a definição de BI inclui os aspectos de armazenamento de dados, o processo de extrair, transformar e carregar (ETL) e os métodos de apresentação de dados. Neste sentido, não se trata apenas de obter dados, mas sim de implementar processos para assegurar que eles estejam sempre disponíveis e acessíveis para a tomada de decisões (TRIEU, 2017).

Para Stewart (2022) BI é uma área de pesquisa nova e em desenvolvimento com oportunidades de replicação e validação de estudos, incluindo estudos de casos em contextos não europeus.

O BI possui componentes principais: um DW, uma coleção de ferramentas para manipular e analisar os dados no DW e uma interface de usuário (como o *dashboard*). DW é um banco de dados especial, cujo objetivo é auxiliar na tomada de decisão da organização, e o ambiente de análise conecta os usuários ao sistema de BI por meio de uma interface (TURBAN *et al.*, 2009).

Na revisão sistemática realizada por Gamboa-Cruzado (2023) identificou-se que os EUA foram o país mais produtivo em pesquisas sobre BI; constatou-se ainda que a metodologia de Kimball foi a mais utilizada no desenvolvimento do DW, a parte mais relevante no desenvolvimento de ambientes de BI.

No tópico seguinte são apresentados conceitos para o desenvolvimento de um DW.

2.2.1 Data Warehouse (DW)

Data Warehouses (DW) são depósitos de dados digitais que podem guardar informações de uma instituição, de forma organizada e personalizada, permitindo o seu uso posterior para subsídio na tomada de decisões. De acordo com Inmon (1997, p. 33), um DW é “um conjunto de dados baseado em assuntos, integrado, não volátil e variável em relação ao tempo, de apoio às decisões gerenciais”. Singh (2001) definiu como um processo de junção dos dados de uma instituição em um único local (repositório) onde os gestores e/ou usuários finais podem executar análises, consultas e gerar relatórios. Esse repositório é chamado DW, sendo um processo desenvolvido para armazenar dados históricos e otimizar consultas (LAUDON; LAUDON, 2022; SINGH, 2001). Novos dados são incorporados ao DW e não sobrescrevem os anteriores,

o que permite acompanhar a evolução de métricas no tempo (KIMBALL *et al.*, 2008; SINGH, 2001).

Segundo Primak (2008), para facilitar o entendimento do conceito de DW é importante comparar um DW com um banco de dados (BD) tradicional, onde um BD pode ser entendido como um conjunto de dados operacionais armazenados e utilizados por um sistema de aplicação; já o DW armazena os dados em formato diferenciado possibilitando rápida agregação desses dados e detalhamento das análises. Primak (2008, p. 37-38) definiu DW como “uma coleção de dados derivados dos dados operacionais para um sistema de suporte à decisão”. Estes dados derivados são, muitas vezes, chamados de dados gerenciais, informacionais ou analíticos. Para Favaretto (2016), um DW apresenta um ambiente flexível, permitindo a análise de qualquer tipo de medida e possibilitando que os próprios usuários criem suas consultas para suporte à decisão.

Um DW é um ambiente que armazena dados de fontes diferentes e os organiza e entrega aos gestores, independente da tecnologia/plataforma que usam. Em síntese, um DW é uma tecnologia de análise e gestão de dados. Para Machado (2010), DW é um armazém de dados históricos cuja finalidade é apresentar informações para identificar indicadores e sua evolução ao longo do tempo, e para Favaretto (2005), por meio do DW é possível a criação de indicadores que permitem consultas analíticas pelo usuário final para o suporte na tomada de decisão.

Kimball e Ross (2013) propuseram uma visão descentralizada do DW, chamada de *Data Mart* (DM). O *Data Mart* é um subconjunto de dados de um DW, organizado por assunto pertencente a uma área ou departamento específico. Inmon (2005) descreveu os DMs como estruturas de dados que incluem informações de acordo com necessidades e interesses de determinado departamento de uma organização, ou seja, as informações são armazenadas em assuntos ou áreas específicas.

Segundo Kimball e Ross (2013) são requisitos de um DW: facilitar o acesso à informação; tornar a informação consistente; ser uma fonte de informações adaptável e resiliente, e manter a informação segura e ser o apoio para a tomada de decisões.

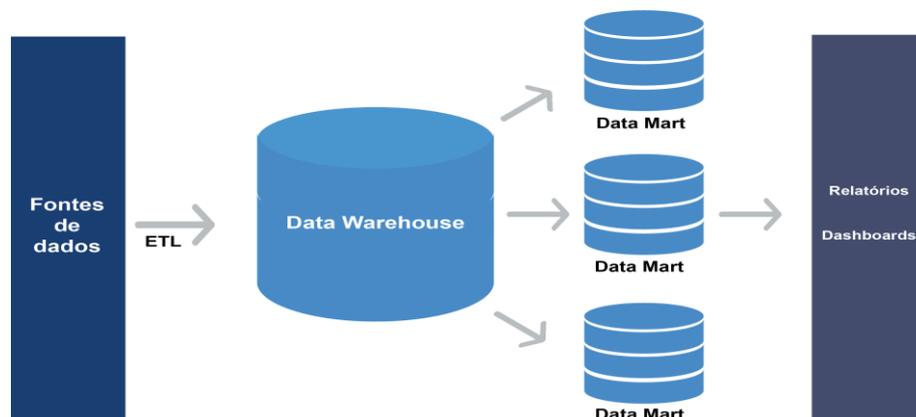
Um DW não é um software/aplicação que possa ser comprado e instalado nos computadores; sua implementação requer a integração de vários processos.

O desenvolvimento/implementação de um DW é influenciado por diversos fatores como a arquitetura escolhida, os recursos disponíveis, o escopo da implementação, a infraestrutura de TI, a forma de acesso aos dados corporativos, bem como o retorno de investimento e a velocidade de realização do projeto (MACHADO, 2010).

Para a arquitetura de um DW existem duas metodologias, uma proposta por Bill Inmon e outra por Ralph Kimball. De acordo com Inmon (2005) um DW é um repositório com as seguintes características: orientado por assunto; integrado (os dados são oriundos de vários sistemas, e por isso precisam ser limpos, filtrados e transformados); não volátil (dados não sofrem alterações ou atualizações); variável no tempo (dados históricos). A metodologia proposta por Bill Inmon, considerado o pai do DW, começa analisando os dados das diversas fontes de dados da instituição. Esses dados são tratados por meio de um processo chamado *Extract Transform and Load* (ETL) e, posteriormente, carregados no DW (INMON, 2005). Após os dados estarem completamente integrados ao DW são gerados vários *data marts* de acordo com a necessidade da instituição. Como a única fonte de dados é o DW, isto garante que a integridade e a consistência sejam mantidas (JUKIC, 2006). Por fim é realizada a análise dos dados, por meio de relatórios e *dashboards*, dentre outros. Esta abordagem é também chamada *top-down*.

Esta abordagem é recomendada se os requisitos de análise não estiverem definidos e as fontes de origem forem estáveis (JUKIC, 2006). Para organizações de grandes dimensões, com a necessidade de obter novas fontes de informação constantemente, esta arquitetura pode não ser a melhor solução, ao exigir um enorme gasto de tempo (JUKIC, 2006). A abordagem de Inmon propõe criar um *Data Warehouse* Institucional – DWI (repositório central) que reúna todas as informações importantes que, posteriormente, podem ser divididas em repositórios menores (ZEA; SAMPEDRO; MORA, 2016). A Figura 1 apresenta graficamente essa abordagem.

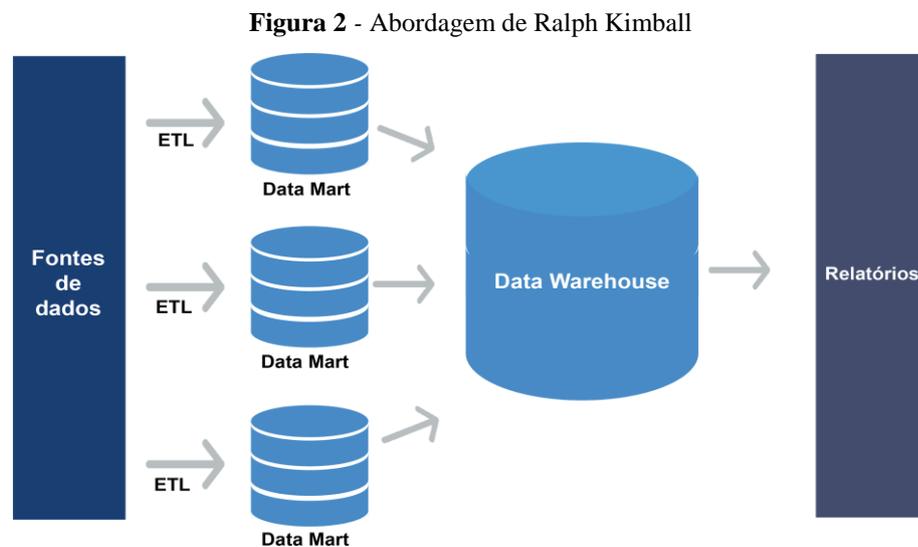
Figura 1 - Abordagem de Bill Inmon



Fonte: Adaptado de Zea, Sampedro e Mora (2016)

Já a abordagem de Kimball propõe que os dados sejam consolidados em *data marts* (DM), versões menores do DW, e podem ser consolidados em um repositório institucional (ZEA; SAMPEDRO; MORA, 2016). A metodologia de Kimball, chamada de *bottom-up*, foi

criada nos anos 90, oferecendo uma nova perspectiva da arquitetura de um DW. As fontes de dados são identificadas e o processo de ETL é realizado. Após estas duas etapas, são criados os DMs de acordo com as necessidades da instituição e, posteriormente, integrados no DW (KIMBALL; ROSS, 2013). Essa abordagem tem como base o modelo dimensional que é desnormalizado por natureza, e divide os dados em duas tabelas. Tabela fato que contém os dados transacionais e tabelas de dimensão, que contém as informações que descrevem os fatos (JUKIC, 2006). Recomenda-se a abordagem *bottom-up* se os requisitos forem definidos e conhecidos. A Figura 2 apresenta graficamente essa abordagem.



Fonte: Adaptado de Zea, Sampedro e Mora (2016)

A abordagem de Ralph Kimball foca nos DM que respondem aos setores das instituições; a junção de vários DM pode resultar em um DW, permitindo que ele seja gradualmente desenvolvido e implementado em um menor período em comparação à abordagem de Inmon (ZEA; SAMPEDRO; MORA, 2016).

Outra divergência entre as abordagens diz respeito ao ciclo de desenvolvimento do DW. Enquanto Kimball considera o primeiro passo para a implementação do DW o levantamento dos requisitos do negócio, dos dados necessários e a capacidade de obtenção desses dados, para Inmon o primeiro passo deve ser a integração dos dados disponíveis em um DW, compreendê-los e, após, determinar quais são os requisitos possíveis de serem obtidos.

Para Kimball (2008), os requisitos de informação são importantes para determinar o nível de detalhamento ou a granularidade dos dados em um DW.

A metodologia de Kimball foi escolhida para o DW do BI para acompanhamento dos egressos da IC.

2.2.1.1 Modelagem Dimensional

A modelagem dimensional é uma técnica usada na construção de um DW e envolve a análise e a concepção de um modelo de dados. Para Machado (2010), trata-se de uma modelagem simples e a visualização do modelo é representada por um cubo de dados que pode ser visto de vários ângulos; o centro é chamado de fato (composto de dados de medidas) e cada lado é uma dimensão, permitindo visualizar a mesma informação de diferentes ângulos. No mesmo sentido, para Kimball e Ross (2013) deve-se projetar, modelar e desenvolver o banco de dados em um formato multidimensional, de modo que os dados possam se relacionar e serem representados como um cubo.

Um modelo de dados que se inicia de forma simples tende a continuar até o final da modelagem, ao contrário de um modelo que já se inicia de forma complexa (KIMBALL; ROSS, 2013).

A tabela fato e tabelas de dimensões formam o modelo dimensional.

- Tabela fato é a tabela principal de um modelo dimensional. O termo “fato” representa uma medida de negócio (KIMBALL; ROSS, 2013).
- Tabela dimensão apresenta contextos descritivos dos negócios, o que a torna a tabela a chave para o modelo ser de fácil entendimento. Kimball e Ross (2013) relacionaram a qualidade do DW com a qualidade e profundidade dos atributos das tabelas de dimensão, ou seja, as dimensões determinam o contexto para os fatos.

Para implementar um projeto de DW, é necessário realizar um levantamento e descobrir as necessidades da informação e indicadores que se deseja obter. O objetivo da modelagem de dados é construir um sistema capaz de atender as necessidades informadas de acordo com a realidade. Para Kimball (2008), a modelagem de dados procura apresentar os dados em um formato que, ao mesmo tempo, seja intuitivo e atenda às consultas com um alto desempenho.

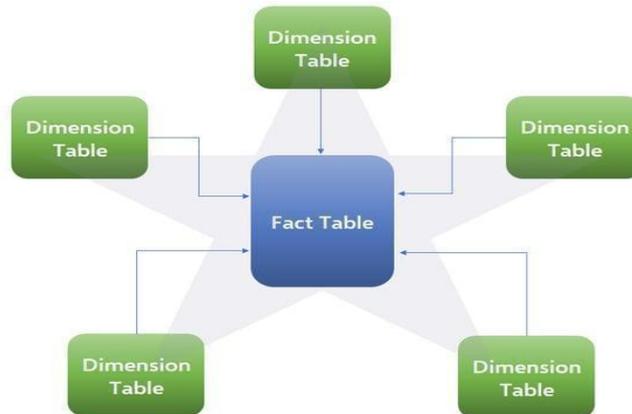
Segundo Machado (2010), a análise multidimensional representa os dados como dimensões. Combinando essas dimensões, o usuário tem uma visão dos dados de um DW podendo, por exemplo, alterar as dimensões e navegar entre os níveis de detalhamento dos dados do DW.

A modelagem dimensional pode ser implementada por meio de dois modelos, *Star Schema* e *Snowflake Schema*. Em ambos os casos, há uma tabela dominante no centro (fato) com múltiplas junções a outras tabelas (dimensões).

No modelo *Star Schema* (Estrela), os fatos são envolvidos por ligações com as dimensões do problema. Sua representação gráfica se assemelha a uma estrela e, por isso, recebe o nome de modelo estrela ou *Star Schema* (KIMBALL, 2008). A Figura 3 apresenta o

esquema estrela, onde a tabela fato é centralizada e as tabelas de dimensão ficam relacionadas somente a ela.

Figura 3 - Modelo Dimensional *Star Schema*

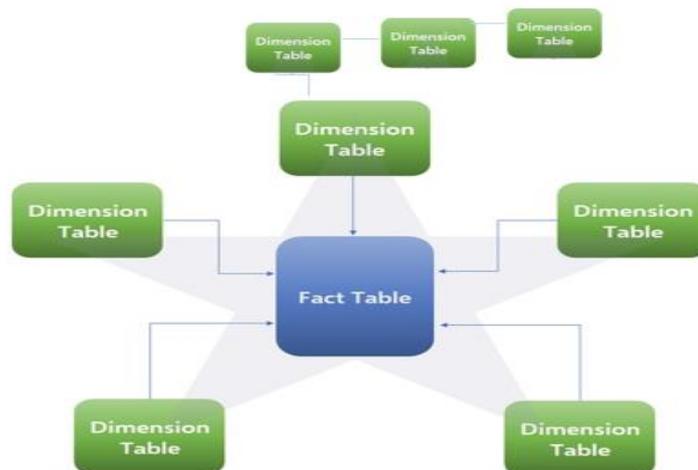


Fonte: Adaptado de Machado (2010, p. 74)

“O modelo *snowflake* é o resultado da decomposição de uma ou mais dimensões que possuem hierarquia entre seus membros” (MACHADO, 2010, p. 75). No mesmo sentido, para Elmasri e Navathe (2021), o modelo floco de neve é resultado da normalização e expansão das tabelas de dimensão do esquema estrela.

A Figura 4 ilustra um exemplo de esquema floco de neve, onde as dimensões produto e vendedor foram normalizadas. Por ser baseada na Figura 3, é possível afirmar que este esquema é uma extensão do esquema estrela.

Figura 4 - Modelo Dimensional *Snowflake Schema*



Fonte: Adaptado de Machado (2010, p. 76)

A modelagem dimensional é fundamental para a implementação do *Data Warehouse*, pois serve de base para o processo de ETL (KIMBALL, 2008).

2.2.1.2 Processo *Extract, Transform and Load* (ETL)

Kimball e Ross (2013), explicaram que a sigla ETL, do inglês *Extract, Transform and Load*, significa Extração, Transformação e Carga. Ela concentra-se na extração de dados das fontes externas e operacionais (sistemas transacionais) e recebem a seguinte descrição:

- Extração: É a primeira etapa do processo e compreende o entendimento das fontes dos dados, copiando os que são necessários (KIMBALL; ROSS, 2013).
- Transformação: Segunda etapa com início após a fase de extração ter sido concluída. Os dados podem sofrer diversas alterações, como correções, conversão de formatos, remoção de campos inúteis, combinação entre dados de diferentes fontes, entre outros (KIMBALL; ROSS, 2013).
- Carga: Última etapa do processo, que começa após a finalização do processo de transformação. Nesta etapa os dados estão prontos para serem carregados no DW, permitindo que os dados apresentem informações (KIMBALL; ROSS, 2013).

Para a realização da ETL, é necessário mapear todas as fontes de dados que serão usadas para compor o repositório (DW). Assim, o processo visa extrair os dados das fontes, tratá-los de acordo com as definições dos modelos lógicos e físicos, e carregá-los no DW (KIMBALL *et al.*, 2008). A Figura 5 apresenta os passos do processo ETL.

Figura 5 – Processo ETL



Fonte: Adaptado de Kimball *et al.* (2008)

2.2.1.3 Ferramentas OLAP para visualização e análise dos dados

A análise dos dados é imprescindível em um processo de tomada de decisão e essa análise depende da capacidade dos SI em oferecer possibilidades para analisar os dados em um DW. Existem muitas ferramentas usadas para manipular e analisar dados, dentre elas, as mais utilizadas são ferramentas de previsão e *Online Analytical Processing* (OLAP).

OLAP é entendido como um conjunto de ferramentas que permitem explorar os dados de um DW em uma arquitetura multidimensional (BARBIERI, 2011).

As ferramentas OLAP possibilitam aos usuários que não tem conhecimento sobre banco de dados manipular e analisar os dados, essas ferramentas são úteis para a tomada de decisão, pois ajudam o analista dos dados a criar gráficos e filtros, enxergando mais rapidamente a informação desejada, identificando problemas, gaps e padrões, sem precisar de uma equipe de TI para a análise (PITON, 2018).

Machado (2010) apresentou os quatro tipos de operações usadas para navegar pelo modelo dimensional usando ferramentas OLAP: *drill down*, *drill up*, *slice e dice*. As duas primeiras são usadas para navegar entre os níveis de detalhamento dos dados do DW, enquanto as outras duas são usadas para mudar as dimensões.

As tabelas dinâmicas do software Microsoft Excel permitem criar relatórios dinâmicos e multidimensionais, permitindo a manipulação dos dados armazenados e diversas análises exploratórias que auxiliam nas decisões. Uma tabela dinâmica consegue resumir quantidades de dados, permitindo decisões conscientes sobre dados importantes, com características como: consultar abundantes quantidades de dados; totalizar, agregar e subtotalizar dados numéricos, resumir dados por categorias e subcategorias, criar cálculos e fórmulas personalizados, expandir e recolher níveis de dados para destacar resultados, mover linhas para colunas ou colunas para linhas, filtrar, classificar, agrupar e formatar condicionalmente o subconjunto de dados mais útil e interessante para se concentrar apenas nas informações desejadas (MICROSOFT, 2022). Ainda por meio desta ferramenta é possível construir *Dashboards*.

O Power BI é uma ferramenta da Microsoft, criada em 2015. O seu objetivo é fornecer visualizações interativas e recursos de BI, em uma interface para os usuários finais criarem relatórios e *dashboards* personalizados. As fontes de dados podem ser importadas do Excel, ou de outros sistemas que armazenam dados. O Power BI pode ser encontrado em diversas versões, a mais comum é o Power BI Desktop para computadores com sistema Windows (MICROSOFT, 2023).

O *Dashboard* é um painel que apresenta informação, métricas e indicadores com o objetivo de representar os números relevantes para a estratégia de negócio e para atingir os objetivos institucionais. Nos *dashboards*, gráficos, mapas, tabelas de dados, barras, indicadores e metas podem ser usados para compor indicadores conforme as necessidades monitoradas (ARAÚJO, 2014).

As ferramentas de previsão (análise preditiva e mineração de dados) tem como objetivo extrair informações e conhecimentos por interpretação e previsão. Podem ser obtidos, por meio da mineração de dados, as seguintes informações: provisões, classificações, associações e

sequências. A mineração de dados consegue fazer análises de alto nível em padrões e tendências, além de fornecer mais detalhes (LAUDON; LAUDON, 2022).

2.3 TRABALHOS RELACIONADOS

A revisão de literatura é essencial para qualquer trabalho acadêmico; por meio deste processo é possível desenvolver e construir uma base de conhecimento referente a teorias e conceitos sobre investigações das diversas áreas. Assim, esta seção apresenta uma revisão bibliométrica da literatura internacional e nacional sobre a utilização de ferramentas de *Business Intelligence* em Universidades. A seguir foi realizada uma revisão narrativa sobre o Acompanhamento de Egressos.

2.3.1 Desenvolvimento de ferramentas de *Business Intelligence* em Universidades

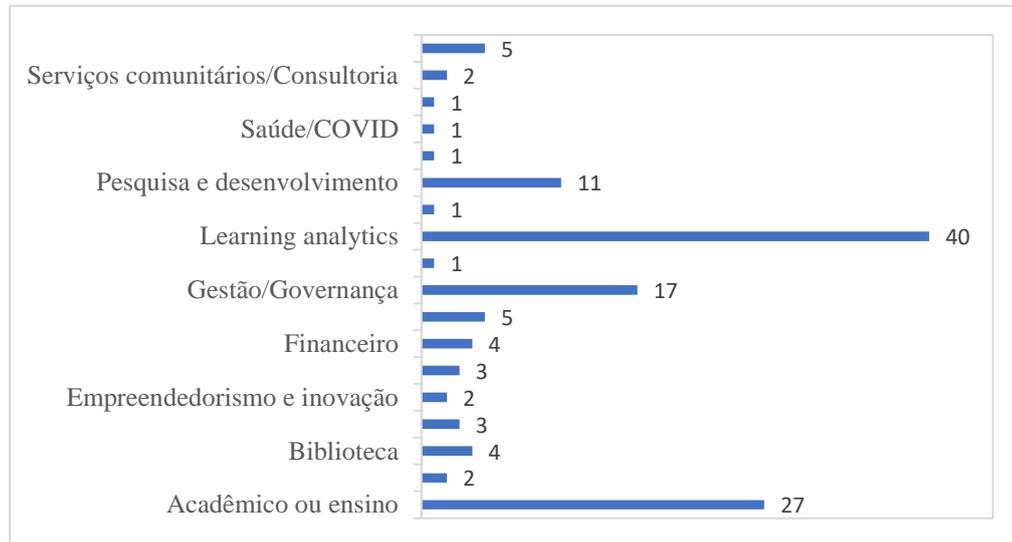
Foi realizada uma pesquisa bibliométrica sobre o uso de ferramentas de BI em Universidades, por meio de buscas sistematizadas em bases internacionais e nacionais. A primeira busca da literatura foi realizada nas bases de dados da *Scopus* e *Web of Science (WoS)*, entre 2 e 6 de janeiro de 2023. Os termos foram pesquisados nos campos títulos e palavras-chave, são eles: “*business intelligence*” ou “*dashboard*” ou “*Data Warehouse*” e “*universities*” ou “*university*” ou “*higher education*”. Foram selecionados somente artigos de periódicos (resultados de estudos revisados por pares) e não foi definido corte temporal para esta primeira busca.

A busca na base de dados *Scopus* retornou trezentos e sessenta e sete (367) artigos e na *WoS* cento e vinte e nove (129), totalizando quatrocentos e noventa e seis (496) artigos. Ao importar o resultado das pesquisas para o software Microsoft Excel, identificou-se cento e dois (102) artigos repetidos; após, foi realizada a leitura do título onde foram desconsiderados cento e trinta (130) artigos; por fim procedeu-se à leitura dos resumos dos demais artigos, onde foram desconsiderados cento e trinta e quatro (134) por não tratarem do tema BI, DW ou *Dashboards* em universidades. A amostra foi finalizada com cento e trinta (130) artigos. Seguiu-se a análise dos artigos para identificar as áreas/tópicos relacionados a pesquisas de desenvolvimento de *BI/DW/Dashboards* em universidades; os dados foram coletados e planilhados.

As áreas/tópicos identificadas nos artigos foram: acadêmico ou ensino, ingresso, biblioteca, egressos, empreendedorismo e inovação, evasão, financeiro, gerenciamento de projeto, gestão/governança, identificação de KPI, *learning analytics*, marketing, pesquisa e desenvolvimento, ranking de universidades, saúde/COVID, segurança cibernética, serviços

comunitários/consultoria e sustentabilidade, conforme distribuição apresentada na Figura 6.

Figura 6 - Áreas com mais estudos na literatura sobre *Business intelligence* na *Scopus* e *Web Of Science* até 2022



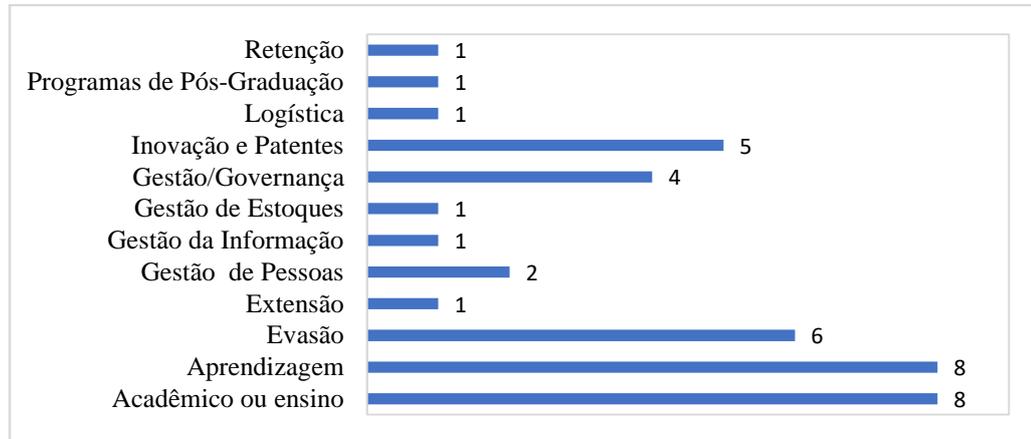
Fonte: Da autora (2023)

A área/tópico com maior número de artigos foi *learning analytics*, com 40 artigos. *Learning analytics* (análise da aprendizagem por meio da tecnologia) é uma área em crescimento, pois as IES estão explorando o potencial para prever retenção de discentes e entender o comportamento de aprendizado, dentre outros.

Na segunda busca sistematizada da literatura optou-se por realizá-la na literatura nacional. A pesquisa foi realizada no Catálogo de Teses e Dissertações da CAPES, entre os dias 10 e 12 de janeiro de 2023; os termos pesquisados foram “*business intelligence*” ou “*dashboard*” ou “*Data Warehouse*” e o corte temporal foram os últimos 5 anos. Foram encontrados 135 estudos (82, 34 e 19 respectivamente). Foi feita a leitura do título e do resumo para verificar se eram estudos relacionados ao desenvolvimento de ferramentas em universidades. Foram descartados 6 estudos que estavam em duplicidade e 90 que não estavam no escopo da revisão, restando 39 estudos. Seguiu-se a análise das teses/dissertações para identificar as áreas/tópicos relacionados a pesquisas de desenvolvimento de ferramentas de BI. Os dados foram coletados e planilhados.

As áreas/tópicos identificadas nas dissertações foram: acadêmico/ensino, aprendizagem, evasão, extensão, gestão de pessoas, gestão da informação, gestão de estoques, gestão/governança, inovação e patentes, logística, Programas de Pós-Graduação e retenção, conforme distribuição apresentada na Figura 7.

Figura 7 - Áreas com mais estudos na literatura sobre *Business Intelligence* no CTD/CAPES entre 2018 e 2022

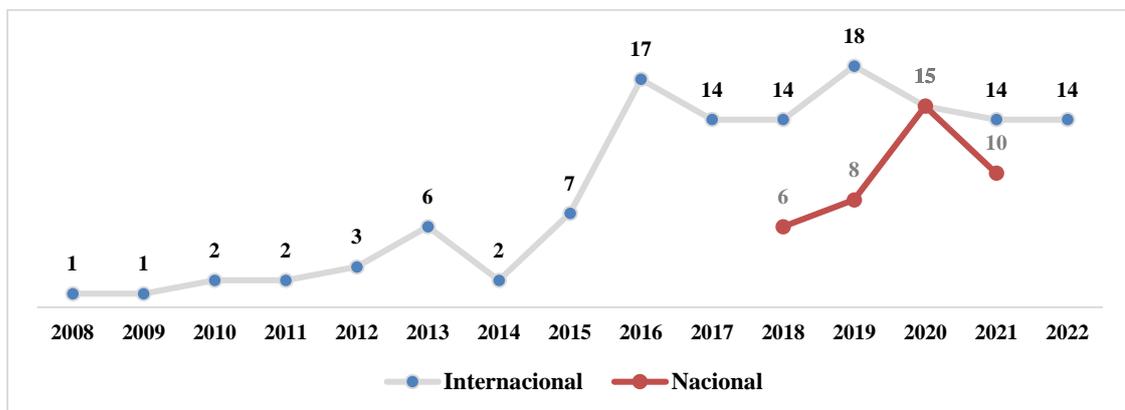


Fonte: Da autora (2023)

A área/tópico com maior número de estudos foi aprendizagem e Gestão Acadêmica/Ensino, com 8 estudos cada.

Analisando o ano de publicação dos artigos, verificou-se um aumento nos últimos sete anos, com pico de crescimento no ano de 2016 e 2019, e manutenção nos últimos três anos. Na literatura nacional houve um salto no ano de depósito das teses/dissertações em 2020 (15 dissertações) em comparação aos anos anteriores; em 2021 foram encontrados 10 estudos e nenhum em 2022 (porém deve-se considerar que no momento desta revisão muitos estudos podem não ter sido lançados na plataforma Sucupira/CAPES ainda). A evolução dos estudos pode ser observada na Figura 8.

Figura 8 - Evolução dos estudos sobre BI na literatura internacional e nacional



Fonte: Da autora (2023)

Assim, foi detalhado na Tabela 1 os resultados dos procedimentos para seleção dos estudos.

Tabela 1 - Síntese do resultado da seleção dos estudos

Repositório	Descritores	Total	Excluídos	Incluídos
Scopus e Web of Science	"business intelligence" ou "dashboard" ou "Data Warehouse" e "universities" ou "university" ou "higher education"	496	366	130
Catálogo de Teses e Dissertações da CAPES	"business intelligence" ou "dashboard" ou "Data Warehouse" (últimos 5 anos)	135	96	39
	TOTAL	631	462	169

Fonte: Da autora (2023)

Após a fase de inclusão e exclusão foi realizada a leitura aprofundada dos estudos que possuíam como foco o contexto e os objetivos desta dissertação. Assim foram selecionados artigos e dissertações mais próximos ao tema para discussão, entre eles: estudos com egressos, pós-graduação, gestão/governança, estudos que usaram a metodologia de Kimball no desenvolvimento do DW do BI e aqueles que trabalharam com a identificação de requisitos informacionais para o DW do BI.

Foram encontrados poucos estudos desenvolvendo ferramentas de BI no acompanhamento de egressos, entre eles destaca-se o estudo de Ismail *et al.* (2022) que teve como objetivo desenvolver um painel exploratório utilizando técnicas de visualização de dados de egressos de uma universidade pública da Malásia - *Universiti Teknologi MARA*. Os resultados foram interpretados por meio de análise descritiva e diagnóstica. Blanco Rojas *et al.* (2016) construíram uma base de dados atualizada com informações da população em estudo (egressos) provenientes das redes sociais (*LinkedIn, Facebook e Twitter*). Segundo os autores, as redes sociais são as principais fontes de dados, além de ser um meio de comunicação inovador, com grande conteúdo informativo. Após a extração dos dados das redes sociais foi realizado o processo ETL deixando o DW pronto para consultas e geração de relatórios que responderam aos requisitos de negócio levantados na fase inicial do projeto.

Ainda sobre estudos com egressos, devido à ausência de um sistema da informação (SI) para fornecer indicadores e estatísticas confiáveis e centralizados sobre a empregabilidade dos graduados universitários marroquinos, Hamdane *et al.* (2022) propuseram um DW acadêmico como uma ferramenta poderosa e confiável para conectar os egressos da universidade com o mundo profissional usando a mineração de dados.

Niño (2020) realizou um diagnóstico para identificar o nível de maturidade em análises de dados, assim foi desenhado um modelo para fortalecer a cultura organizacional, a infraestrutura, a gestão de dados, a análise de dados e a governança. A proposta contempla a

definição de um quadro de governança, princípios orientadores, estratégias, políticas, processos e órgão de decisão.

Silva (2018) propôs uma ferramenta de BI para a análise da produtividade do professor permanente nos itens de produção intelectual/técnica do sistema de avaliação da Capes, no quadriênio 2013/2016, verificando a interdependência entre as variáveis. Lessak (2018) desenvolveu uma ferramenta de BI para apoiar a gestão na sistematização de atividades e competências dos pesquisadores do Polo Embrapii do IFSC e seus relacionamentos com empresas. As informações podem ser utilizadas para melhorar a comunicação com as empresas e outras Instituições.

Para Da Silva Souza (2021, p. 8) “o desenvolvimento do Data Mart utilizando ferramentas de BI, permite ao gestor extrair informações globais e detalhadas sobre a gestão de pessoas da instituição”. A pesquisa de Da Silva Souza (2021) permitiu analisar o quadro de servidores da UFV, apontando possíveis desligamentos e evidenciando questões sobre reposição de vagas para concurso (cargos extintos ou vedados). Vendruscolo (2020) desenvolveu um sistema de BI para a extensão universitária da UFSC. Foram identificados indicadores para a extensão universitária e os dados utilizados para os cálculos dos indicadores nos sistemas da universidade. Os *dashboards* foram desenvolvidos para apresentar a informação de uma forma clara e objetiva e que, simultaneamente, pudesse ser realizada uma análise mais profunda. A ferramenta já foi utilizada pela PROEX da IES para o PDI 2020/2024.

No trabalho de Oraee, Sanatjoo e Ahanchian (2021) foi destacada a importância de identificar as necessidades de informação no processo de inteligência de negócios, pois além de melhorar as decisões é um pré-requisito para as próximas ações e atividades no processo de desenvolvimento do BI.

Mendes (2020) desenvolveu um DW sobre evasão na UNIFEI, explicando os dados necessários para essa informação, quais as fontes de dados e como extraí-los, transformá-los e agregá-los nas métricas de evasão. Para o desenvolvimento do DW foi utilizado a metodologia de Kimball. Para Mendes (2020, p. 82), “o levantamento dos requisitos do DW é a primeira e mais importante etapa do seu desenvolvimento”. Vieira (2021) desenvolveu um ambiente de BI para analisar o rendimento acadêmico dos estudantes de graduação da UNIFEI para auxílio aos gestores. A ferramenta desenvolvida supre parte da deficiência dos Sistemas de informação da IES entregando informações relevantes à Gestão. O autor também utilizou a metodologia de Kimball para o desenvolvimento do DW do BI, a mesma que será utilizada neste estudo.

A revisão permitiu verificar como as ferramentas de BI estão sendo desenvolvidas para apoio a tomada de decisão em IES e para mostrar a escassez de pesquisas relacionadas a

egressos e sua trajetória profissional, revelando a importância do desenvolvimento de ferramentas para apoio a gestão de egressos, um assunto tão atual.

2.3.2 Sistemas de Informação e o Acompanhamento de Egressos

Sistemas de informação no acompanhamento dos egressos representa um relevante instrumento para o aperfeiçoamento das IES, pois gestores necessitam de informações sobre seus egressos para o constante aperfeiçoamento das IES. Para que a utilização dos SI seja efetiva é necessário no seu desenvolvimento a identificação dos processos organizacionais e de qual informação existente é necessária para sua operacionalização.

O acompanhamento se refere ao processo de monitoramento dos egressos com vistas ao estabelecimento e à manutenção de um banco de dados, sobretudo no que tange à colocação profissional dos titulados. Esse acompanhamento pode ser realizado por *surveys* (questionários com egressos), portais de egressos (relacionamento) e por meio da análise de dados secundários; pode-se ainda realizar a junção de mais de uma forma de acompanhamento para se ter um resultado mais completo. Nas IES o acompanhamento pode ser realizado em diversos níveis, ou seja, por curso de graduação, curso de especialização, curso de PG *Stricto sensu* ou com todos os cursos da IES.

Nas pesquisas que analisam o impacto do PIBIC encontramos sugestões de implantação de SI, bem como a importância da organização e análise dos dados. Segundo Canuto (2018) apesar dos órgãos e agências de fomento não estabelecerem um parâmetro ou meta que deve ser alcançada por meio da utilização das cotas de bolsas por elas concedidas, as IES devem realizar o acompanhamento dos resultados dos seus egressos dos Programas de IC, como um instrumento relevante para avaliar sua eficiência e eficácia.

Ângelo (2020) verificou se a IC contribuiu na formação de recursos humanos para a pesquisa por meio de uma perspectiva censitária. Segundo o autor, "a falta de acompanhamento de egressos gera desperdícios ao país, pois possibilita o descaminho do ex-bolsista para outras atividades, como forma emergencial de sobrevivência. Assim, outras iniciativas podem ser realizadas, de modo complementar, tais como o encaminhamento do egresso ao mercado de trabalho, como também a aplicação da pesquisa na sociedade" (ÂNGELO, 2020, p. 106).

Dantas (2019) ressaltou a importância do desenvolvimento de um sistema que possibilite a visibilidade, bem como a análise dos dados referente aos resultados dos programas de IC nas IES para acompanhamento pela Gestão. Ainda segundo a autora, o SI, serviria como contribuição para definição de políticas públicas que aumentem a eficiência desses programas.

O estudo de Cabral (2021) teve como objetivo principal elencar indicadores e identificar diretrizes e estratégias para a gestão de egressos da PG.

Sobre o acompanhamento da trajetória profissional de mestres e doutores, Maccari *et al.* (2014) salientam que uma das dificuldades das IES para o acompanhamento dos egressos (trajetória profissional e acadêmica) é compilar a grande quantidade de dados isolados e desatualizados. Neste sentido, o estudo propôs a identificação de quais devem ser os pontos observados pelo gerente do projeto para subsidiar a implantação de um sistema com o intuito de obter o melhor resultado na abordagem e planejamento.

3 METODOLOGIA

Para Martins e Theóphilo (2016), a metodologia é uma forma de captar olhares significativos da realidade para responder a questionamentos de pesquisa, delineando o percurso para atingir determinado objetivo.

Nesta seção será apresentada a metodologia adotada no desenvolvimento da pesquisa. Inicia-se caracterizando a IES eleita como objeto de estudo e os sujeitos (egressos da IC). Em seguida, será apresentada a caracterização da metodologia, o método de pesquisa escolhido, as suas fases de desenvolvimento e as técnicas usadas na coleta e análise dos dados.

3.1 Objeto de estudo

Esta pesquisa foi desenvolvida na Pró-Reitoria de Pesquisa, Pós-Graduação e Inovação (PPPI), no âmbito da Coordenadoria de Bolsas de Iniciação Científica (CB), setor subordinado à Diretoria de Pesquisa e Pós-Graduação (DPPG), onde são administrados e monitorados os Programas de IC e os projetos de pesquisa desenvolvidos na instituição, incluindo, os processos seletivos, implementação de bolsas, cancelamentos, substituições, controle e acompanhamento de bolsistas e voluntários, além de prestação de informações e instruções tanto a orientandos como orientadores. Além das cotas institucionais de bolsas de IC das agências de fomento (CNPq e FAPEMIG), a PPPI também fomenta a IC via recursos específicos para editais de bolsas.

O IFSULDEMINAS, criado pela Lei nº 11.892, de 29 de dezembro de 2008, teve sua origem na junção das Escolas Agrotécnicas localizadas nos municípios de Muzambinho, Inconfidentes e Machado, constituindo uma nova instituição e assumindo como compromisso o desenvolvimento regional do Sul de Minas Gerais por meio da excelência na Educação Profissional e Tecnológica (IFSULDEMINAS, 2023). Atualmente é composto por oito campi localizados nas cidades de Muzambinho, Machado, Inconfidentes, Passos, Poços de Caldas, Pouso Alegre, Carmo de Minas e Três Corações. Além de núcleos avançados e “polos de rede”, ampliando o acesso ao ensino em 178 municípios de abrangência. As unidades (campis) do IFSULDEMINAS são apresentadas na Figura 9.

Figura 9 - Unidades do Instituto Federal do Sul de Minas Gerais



Fonte: IFSULDEMINAS (2023)

A reitoria do IFSULDEMINAS, localizada na cidade de Pouso Alegre/MG, coordena toda a estrutura educacional e administrativa dos campi. É responsável pela administração geral da instituição e pela execução das políticas de gestão educacional, orçamentária, patrimonial e de pessoal, composta pelo Gabinete, cinco pró-reitorias: pró-reitoria de Pesquisa, Pós-Graduação e Inovação (PPPI), pró-reitoria de Ensino (PROEN), pró-reitoria de Extensão (PROEX), pró-reitoria de Gestão de Pessoas (PROGEP), pró-reitoria de Administração (PROAD), e duas Diretorias Sistêmicas: Diretoria de Desenvolvimento Institucional (DDI) e Diretoria de Tecnologia da Informação (DTIC) (IFSULDEMINAS, 2023).

O IFSULDEMINAS tem como missão “promover a excelência na oferta da educação profissional e tecnológica, em todos os níveis, formando, cidadãos críticos, criativos, competentes e humanistas, articulando ensino, pesquisa e extensão e contribuindo para o desenvolvimento sustentável do Sul de Minas Gerais” (IFSULDEMINAS, 2023).

Os sujeitos desta pesquisa são os egressos do PIBIC/IFSULDEMINAS que foram bolsistas do programa entre 2011 e 2021.

Neste sentido, o acompanhamento dos egressos, no caso desta pesquisa, os egressos da iniciação científica, pode avaliar a formação acadêmica ao nível de pós-graduação (mestrado e doutorado) e a inserção no mercado de trabalho, além de contribuir com o processo de avaliação da instituição no sentido do cumprimento de sua missão, a formação de recursos humanos qualificados. É importante salientar que conhecer o percurso acadêmico e a atuação profissional dos egressos da IC constitui tanto uma informação do impacto das IES na sociedade, quanto

um dado relevante para o planejamento de ações institucionais; é importante saber a trajetória acadêmica, onde estão e o que estão fazendo os estudantes que fizeram parte do PIBIC/IFSULDEMINAS.

3.2 Classificação/Tipo da pesquisa

Para execução de uma pesquisa é natural que se busque a classificação e organização dos fatos para o seu melhor entendimento (GIL, 2010). Segundo Andrade (2010), as pesquisas podem ser classificadas de várias formas, porém é relevante definir a pesquisa quanto: a sua natureza, aos objetivos, a abordagem e ao método de pesquisa.

Quanto à natureza da pesquisa, este estudo foi enquadrado com uma pesquisa aplicada, uma vez que serão desenvolvidos artefatos para se resolver uma situação real, presente em instituições/organizações reais.

Segundo Thiollent (1986, p.108) a pesquisa aplicada:

se concentra em torno dos problemas presentes nas atividades de instituições, organizações, grupos ou atores sociais. Ela está empenhada na elaboração de diagnósticos, identificação de problemas e busca de soluções (THIOLLENT, 1986, p.108).

Quanto ao seu objetivo, a pesquisa caracteriza-se como uma pesquisa descritiva, visto que pretende obter familiaridade com o problema e torná-lo mais explícito. A pesquisa descritiva tem por objetivo descrever as características de uma população definida ou estabelece relações entre variáveis (GIL, 2002).

Quanto à abordagem, a pesquisa fundamentou-se na abordagem qualitativa. A pesquisa qualitativa objetiva responder questões particulares, que não podem ser simplesmente quantificadas, e aprofunda-se no significado das ações e relações humanas que não são perceptíveis em estatística, equações e medidas (MINAYO, 2015).

Quanto ao método de pesquisa, devido à natureza do objeto de pesquisa e os aspectos que o rodeiam, será realizada uma pesquisa-ação. Na pesquisa-ação, a investigação se inicia com a identificação de um problema em um ambiente real, demandando uma descrição tanto do problema quanto do seu contexto; assim o pesquisador deve participar como ator da ação, e focar em uma ação para a solução de um problema, relatando suas impressões e as soluções investigadas. “Pela pesquisa-ação é possível estudar dinamicamente os problemas, decisões, ações, negociações, conflitos e tomadas de consciência que ocorrem entre os agentes durante o processo de transformação da situação” (THIOLLENT, 2022, p. 25).

3.3 Pesquisa-ação

Para Mello *et al.* (2012) além de produzir conhecimento, a pesquisa-ação busca resolver um problema prático onde o pesquisador procura compreender os fenômenos observando-os, descrevendo-os e interpretando-os. Ainda segundo o autor, obrigatoriamente, a primeira etapa do método da pesquisa ação é referente a pesquisa e a segunda referente a ação. As partes interessadas e os pesquisadores se envolvem para atingir o objetivo da pesquisa por meio da definição do problema, busca de soluções e aprofundamento do conhecimento científico disponível (THIOLLENT, 2022).

Na literatura sobre pesquisa-ação são encontradas diferentes formas de apresentar seus ciclos e fases. Nesta pesquisa, adotou-se o ciclo proposto por Mello *et al.* (2012) composto por cinco fases: planejar a ação — subdivida em três etapas: seleção da unidade de análise e técnicas de coleta de dados, definição da estrutura conceitual-teórica e definição do contexto e propósito da pesquisa; coletar os dados — podem ser coletados de diferentes formas e podem ser primários e/ou secundários; analisar dados e projetar ações — os critérios para a análise precisam ter relação ao propósito da pesquisa e ao final desta etapa deve-se elaborar um plano de ação; implementar a ação — desenvolver a solução de um determinado problema, e avaliar resultados — a avaliação dos resultados deve ter relação com os objetivos da pesquisa (científico e técnico) e as proposições estabelecidas no início da pesquisa.

Estas fases são executadas em sequência e de forma cíclica, isto é, na última etapa o resultado é avaliado, e se necessário, é considerada a preparação de um novo ciclo da pesquisa ação até que se chegue no resultado esperado. A Figura 10 ilustra as fases da pesquisa-ação.

Figura 10 – Fases da pesquisa-ação



Fonte: adaptado de Mello *et al.* (2012)

Este ciclo possui um número de etapas que não é nem tão pequeno com sub-etapas embutidas, nem tão grande que fragmente demais procedimentos relacionados. A ação e a pesquisa são os principais resultados de uma pesquisa-ação, assim eles não são apenas soluções para problemas reais, mas o aprendizado dos resultados não esperados e esperados, produzindo contribuição para o conhecimento científico e para a teoria. Esses resultados podem ainda informar a outras organizações como atuar em um problema específico.

3.4 Proposta metodológica: Fases do desenvolvimento da pesquisa-ação e técnicas de coleta e análise dos dados

A representação da pesquisa-ação é de um ciclo iterativo, porém ela é tipicamente realizada em ciclos iterativos que sucessivamente refinam o conhecimento adquirido nos ciclos anteriores. Ele pode ter início de duas formas (abordagens): a) o pesquisador identifica um problema (lacuna) na literatura que possa ser resolvido cientificamente, e b) o pesquisador participa da proposta de solução de um problema identificado em uma organização por um método de pesquisa (MELLO *et al.*, 2012).

Esta pesquisa seguiu o modelo proposto por Mello *et al.* (2012), com o fluxo de execução em cinco fases, e teve seu início com a identificação de um problema institucional.

Fase 1) planejar a pesquisa-ação: definiu-se o contexto e propósito da pesquisa, a estrutura conceitual-teórica, e a seleção da unidade de análise e técnicas de coleta de dados.

A pesquisa nasceu por meio de um problema identificado no ambiente de trabalho da pesquisadora, a necessidade de se acompanhar a trajetória acadêmica e profissional dos egressos da IC do IFSULDEMINAS. Teve como objetivo geral propor um ambiente de BI para o acompanhamento desses egressos. Foi levantada a fundamentação teórica para contextualizar e fundamentar o desenvolvimento do DW do BI e para subsidiar a identificação de requisitos do BI. A unidade de análise foi selecionada, sendo neste caso, os egressos da iniciação científica do IFSULDEMINAS; quanto a definição das técnicas para a coleta de dados, foram combinadas as seguintes: reuniões com roteiro semiestruturado (necessidade dos requisitos), pesquisa documental (análise de documentos e arquivos), pesquisa bibliográfica (referencial teórico para desenvolvimento de um DW e revisão da literatura) e observação participante do pesquisador (ambiente da pesquisa).

Fase 2) coletar os dados: foi iniciado o processo de definição das informações necessárias à ferramenta de *Business Intelligence*. Foram utilizadas fontes primárias e secundárias. A coleta dos dados foi conduzida visando construir um modelo de acompanhamento que permita avaliar as trajetórias de estudos e profissional dos egressos da

iniciação científica na PG *Stricto sensu*, entre outros. Estas trajetórias permitirão analisar de forma sistêmica a contribuição dos egressos da IC/IFSULDEMINAS no sistema de ensino superior brasileiro (na formação do mercado acadêmico) ou como pesquisador em instituições privadas (no mercado privado). Os dados foram tabulados sendo analisados na próxima fase.

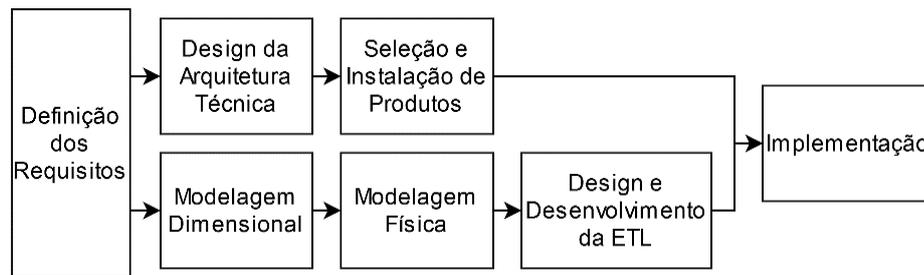
Fase 3) analisar os dados e projetar ações: nesta fase os critérios para a análise precisam ter relação com o propósito da pesquisa e ao final deve-se elaborar um plano de ação. Assim os dados coletados foram analisados; na análise de dados, de cunho descritivo, a comparação entre os dados tabulados e a teoria envolvida no tema pesquisado e na solução foi fundamental. Todos os dados coletados nas reuniões, na pesquisa documental e bibliográfica, bem como o acesso às fontes dos dados e sua disponibilidade foram analisados. Após a análise de viabilidade foi elaborado um plano de ação para a solução do problema identificado.

Um dos objetivos da pesquisa-ação é a resolução de um problema real/prático, portanto o final desta etapa se deu com a elaboração do plano de ação para o desenvolvimento do BI, com todas as recomendações para a solução do problema.

Fase 4) implementar a ação: o plano de ação foi executado. Foi uma fase importante e demorada, que incluiu o processo de modelagem do DW, a extração, transformação e carga dos dados (ETL) e a ferramenta para acesso e divulgação das informações (*Dashboard*). Foi fundamentada no referencial teórico apresentado na seção “2.2.1 (*Data Warehouse*)” que descreveu que para o desenvolvimento de um DW existem dois métodos/abordagens, a de Bill Inmon e a de Ralph Kimball.

Nesta pesquisa, o método de Ralph Kimball foi o escolhido. Uma característica desta abordagem e que contribuiu na escolha é seu foco em repositórios locais (DMs), construídos por setores funcionais, que podem ser, posteriormente, conectados a um DW de nível organizacional (KIMBALL *et al.*, 2008). Além do foco em DMs, outra característica está relacionada com a arquitetura de dados do DW, a modelagem dimensional de Kimball é mais simples para ser desenvolvida quando comparada ao modelo proposto por Inmon. Nesse sentido, essas características tornam o modelo de modelagem de Kimball mais prático e acessível para aqueles que não são da área de TI, permitindo que o esforço e o tempo despendido para o desenvolvimento do DW sejam menores e mais bem aproveitados.

As etapas de desenvolvimento de um DW propostas por Kimball são apresentadas na Figura 11.

Figura 11 - Ciclo de vida do *Data Warehouse*

Fonte: Adaptado de Kimball *et al.* (2008)

Na definição dos requisitos foram usados os resultados da fase 3 da pesquisa-ação, ou seja, os requisitos após a análise dos dados e viabilidade do acesso. Esses mesmos resultados foram utilizados no desenvolvimento do modelo conceitual e a modelagem dimensional do DW. Para o design da arquitetura técnica (*bottom-up*) e seleção dos produtos foi utilizado o Access para as consultas (popular tabela), o Microsoft Excel como área de preparação dos dados (*flat table*) e o Software Microsoft Power BI como ferramenta OLAP para a visualização das análises dos dados.

Fase 5) avaliar os resultados: A avaliação dos resultados deve ter relação com os objetivos (científico e técnico) e as propostas estabelecidas no início da pesquisa. Assim finalizado o desenvolvimento do BI, foi feita a demonstração da ferramenta para avaliação dos Gestores; para isso foi realizada uma reunião não-estruturada para apresentação, análise e testes e validação da ferramenta.

O objetivo da avaliação/validação pelos Gestores e pesquisadores relacionados à área de pesquisa foi o de verificar se o BI cumpre os requisitos informados na fase 2, analisados e aprovados na fase 3 da pesquisa-ação. A apresentação foi dividida em três fases: apresentação da pesquisa (relato com o objetivo inicial), resultados da análise dos requisitos e apresentação das funcionalidades do BI pelo pesquisador e teste com os gestores; e por fim, respostas às perguntas por parte dos mesmos e parecer da ferramenta.

Foram solicitadas pequenas alterações nos gráficos para que os mesmos apresentassem de forma mais clara as análises, porém nada que demandasse um segundo ciclo da pesquisa-ação. A ferramenta foi validada pelos Gestores e pesquisadores convidados.

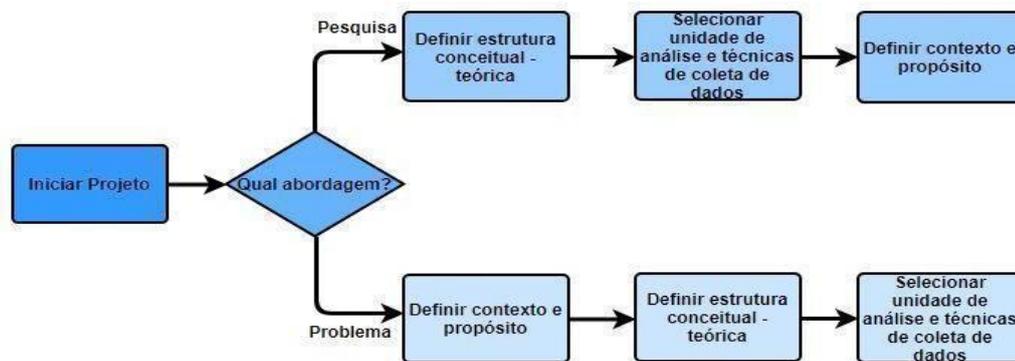
4 DESENVOLVIMENTO

Nesta seção foram descritos os resultados da proposta metodológica apresentada, de acordo com as fases da pesquisa-ação, no desenvolvimento da proposta de um ambiente de BI para o Acompanhamento dos Egressos da IC do IFSULDEMINAS. Assim, a seção descreve desde o planejamento da pesquisa-ação até a avaliação pelos gestores da ferramenta desenvolvida por meio do software Power BI (interfaces para os usuários finais), passando pela coleta e análise dos dados, modelagem dimensional e o processo de extração e tratamento e carregamento dos dados (ETL).

4.1 Fase 1 – Planejamento da pesquisa-ação

De acordo com Mello *et al.* (2012), a primeira fase da pesquisa-ação, o planejamento, possui três etapas: definição do contexto e objetivo da pesquisa, definição da estrutura conceitual-teórica e seleção da unidade de análise e técnicas de coleta de dados; porém, seu início depende da forma de iniciação da pesquisa-ação que pode ocorrer por duas abordagens: a) o pesquisador identifica um problema (lacuna) na literatura que possa ser resolvido cientificamente, e b) o pesquisador participa da proposta de solução de um problema identificado por uma organização por meio de um método de pesquisa (MELLO *et al.*, 2012), conforme demonstrado na Figura 12.

Figura 12 – Abordagens do planejamento da pesquisa-ação



Fonte: Mello *et al.* (2012)

O objetivo desta pesquisa surgiu de um problema identificado no trabalho da pesquisadora, que necessita de uma proposta para solução. Para Mello *et al.* (2012), na abordagem da pesquisa-ação dirigida por um problema, a estrutura conceitual-teórica é realizada após a definição do problema da organização a ser solucionado. Assim, a fundamentação teórica, contextualiza e fundamenta o problema diagnosticado. A questão de pesquisa e seus objetivos são definidos para propor recomendações e resolver o problema,

contribuindo assim para a base de conhecimento. Dessa forma, o planejamento desta pesquisa seguiu as etapas apresentadas na Figura 13.

Figura 13 – Etapas do planejamento da pesquisa-ação (abordagem - identificação de um problema)



Fonte: Adaptado de Mello *et al.* (2012)

4.1.1 Definir contexto e propósito

Nesta etapa da pesquisa foram definidos o contexto e o propósito da pesquisa, cujo objetivo geral é propor um ambiente de BI para o acompanhamento dos egressos da IC do IFSULDEMINAS com informações gerenciais que auxiliem os gestores na tomada de decisão no âmbito da iniciação científica, pesquisa e pós-graduação.

Até o momento, o IFSULDEMINAS não dispõe de um Sistema Integrado de Informações que apresente informações para apoiar as decisões dos gestores. Até 2016, o IFSULDEMINAS utilizava o sistema WebGiz, que era responsável pelo controle da trajetória acadêmica dos discentes.

A Diretoria de Gestão de TI (DGTI) do Instituto Federal do Rio Grande do Norte (IFRN) desenvolveu o Sistema Unificado de Administração Pública (SUAP) para a Gestão dos Processos Administrativos do IFRN. Atualmente, devido ao êxito do projeto, o SUAP já é utilizado em outros Institutos Federais, como o IFSULDEMINAS, que, desde 2015, tem implantado gradualmente os módulos do SUAP. Em 2017, começou a transição do WebGiz para o SUAP na parte acadêmica; porém, mesmo o SUAP sendo um sistema que está em constante atualização dos módulos e novas funcionalidades, ainda não possui no módulo SUAP/PESQUISA uma funcionalidade que nos permita ter um controle sobre os estudantes que participaram do PIBIC e sobre sua carreira após sua saída ou continuidade de estudos no IFSULDEMINAS.

O IFSULDEMINAS, hoje, para controlar os dados do PIBIC, usa planilhas eletrônicas por processo seletivo (vigência das bolsas) que contêm os dados gerais dos bolsistas, campus, orientador, modalidade das bolsas, vigência, etc. Esses dados estão dispersos, havendo pouca integração que forneça informações.

Apesar da recomendação do CNPq, de que as IES realizem pesquisas focais que permitam acompanhar a efetividade do PIBIC, ainda não se observa, na maioria das IES,

relatórios/portais com a apresentação da análise do impacto da IC na trajetória formativa de seus egressos.

Para auxiliar na solução desse problema, este estudo tem como foco principal a proposição de um ambiente de BI para acompanhamento dos egressos da IC que permita, entre outros, o monitoramento de sua formação (pesquisador) em cursos de PG *Stricto sensu*. Salieta-se ainda que conhecer o percurso acadêmico e a atuação profissional dos egressos da IC é tanto uma do impacto das IES na sociedade, quanto um dado relevante para o planejamento de ações institucionais. É importante saber a trajetória acadêmica, onde estão e o que estão fazendo os egressos do PIBIC/IFSULDEMINAS.

4.1.2 Definir estrutura conceitual-teórica

Na etapa dois, buscou-se a fundamentação teórica para contextualizar e fundamentar o problema identificado (Sistema para Acompanhamento dos Egressos da Iniciação Científica).

No item 2.1 foi apresentada a política de iniciação científica no Brasil (PIBIC/CNPq), os objetivos gerais do programa e o processo de formação do pesquisador que contribuíram para identificar os requisitos para o DW do BI. Os objetivos gerais do programa são: a) contribuir para a formação de recursos humanos para a pesquisa; b) contribuir para a formação científica de recursos humanos que se dedicarão a qualquer atividade profissional; e c) diminuir o tempo médio de permanência dos estudantes na pós-graduação (CNPq, 2006); no item 2.2.1, foi apresentada a conceituação do método de desenvolvimento (prático) de um *Data Warehouse* (DW).

Além da fundamentação teórica sobre o desenvolvimento de um DW e da política de iniciação científica no Brasil, foram realizadas duas (2) revisões integrativas da literatura que tratam da temática de Acompanhamento de Egressos. A revisão concentrou-se nos egressos das IES e não nos egressos da iniciação científica, porém os resultados ajudaram consideravelmente na proposta de uma solução para Acompanhar os Egressos da IC.

A 1ª revisão identificou na literatura o uso de sistemas de informação no acompanhamento dos egressos nas IES, o material utilizado nesta revisão (artigos) foi pesquisado em duas bases de dados internacionais: *Web of Science (WOS)* e *Scopus*. Estas bases foram escolhidas por apresentarem os melhores estudos no campo do conhecimento científico; o artigo foi publicado em 2022. A 2ª revisão teve como objetivo analisar as abordagens dos estudos sobre sistemas de acompanhamento de egressos na literatura nacional, especificamente no Catálogo de Teses e Dissertações (CTD) da CAPES, O resultado desta pesquisa foi relevante

para a formulação e identificação de requisitos para o BI, uma vez que alguns dos estudos selecionados para a revisão abordaram o acompanhamento de egressos na pós-graduação.

4.1.3 Selecionar a unidade de análise e técnicas de coleta de dados

Na 3ª etapa do planejamento da pesquisa-ação, a unidade de análise (objeto e sujeitos do estudo) e as técnicas de coleta de dados foram selecionadas. Conforme Mello *et al.* (2012), no caso da abordagem da pesquisa-ação motivada pelo problema (demanda da organização), a definição da unidade de análise (objeto de estudo) já foi realizada, sendo neste caso, os egressos da IC do IFSULDEMINAS; neste estudo foram utilizados os dados dos egressos do PIBIC promovido pelo CNPq, pela quantidade e qualidade dos dados disponíveis e em razão de sua abrangência nacional, tendo como população (sujeitos) os egressos do PIBIC que foram bolsistas entre os anos de 2011³ e 2021. O item 3.1 trata do objeto de estudo com mais informações sobre a instituição, seus campi e a estrutura administrativa do IFSULDEMINAS.

Para a coleta de dados, foram combinadas as seguintes técnicas: reuniões com roteiro semiestruturado e não-estruturadas (levantamento dos requisitos), pesquisa bibliográfica, pesquisa documental (análise de documentos e arquivos) e observação participante do pesquisador (ambiente da pesquisa).

No Quadro 1, apresentam-se as técnicas de coleta de dados usadas de acordo com as fases da pesquisa-ação e em consonância com os objetivos específicos desta pesquisa.

Quadro 1 - Apresentação do resumo das técnicas de coleta e análise dos dados

Objetivos Específicos	Fase da metodologia	Instrumento de coletas de dados
Identificar os requisitos informacionais necessários para subsidiar a criação do BI para o acompanhamento de egressos da IC	Fase 2 e 3	Reunião com roteiro semiestruturado com os Gestores do PIBIC/IFSULDEMINAS e PPPI. Observação direta, pesquisa documental e pesquisa bibliográfica.
Identificar as principais fontes de informações relacionadas ao processo de tomada de decisão.	Fase 2 e 3	
Elencar quais os indicadores necessários para o acompanhamento dos egressos da IC	Fase 2 e 3	
Mapear e levantar os dados dos egressos do PIBIC/CNPq do IFSULDEMINAS.	Fase 2, 3 e 4	
Desenvolver o DW do BI e por meio do software Power BI um painel (<i>dashboard</i>) para exploração das informações.	Fase 4	Pesquisa bibliográfica (conforme referencial teórico na seção 2).
Avaliação da ferramenta pela IES.	Fase 5	Reunião não-estruturada com os Gestores do PIBIC/IFSULDEMINAS e PPPI.

Fonte: Da autora (2023)

³ Neste ano ocorreu a aprovação na chamada do CNPq para as cotas institucionais de bolsas do PIBIC/CNPq no IFSULDEMINAS.

4.2 Fase 2 – Coleta dos dados

Para Mello *et al.* (2012), os dados podem ser coletados de diferentes formas, dependendo do contexto da pesquisa, e podem envolver dados primários (observação e reuniões) e secundários (pesquisa documental e bibliográfica).

Esta etapa é importante para definir a necessidade da instituição com relação às informações necessárias ao ambiente de BI. Assim, analisou-se o processo atual da instituição no acompanhamento dos egressos da IC para entendimento do processo e do problema a ser solucionado. Os especialistas (Diretor e Coordenador da Gestão de Bolsas) foram ouvidos para entendimento da situação problema, por meio de reuniões (semiestruturadas e não-estruturadas), para levantar os requisitos necessários para o acompanhamento dos egressos da IC e para identificar as necessidades de dados e informações para direcionar a coleta nos diversos sistemas de informações.

Para obter informações mais precisas para o levantamento dos requisitos do BI foram realizadas três reuniões. Na primeira reunião, foi possível ter uma ideia geral do setor de Bolsas de Iniciação Científica e teve como foco a estrutura do setor, os processos seletivos de bolsas de iniciação científica, os controles dos bolsistas, como eles são avaliados, dentre outros. O Quadro 2 demonstra como é feito o controle dos dados dos bolsistas no IFSULDEMINAS.

Quadro 2 – Modelo de controle dos bolsistas de IC realizada pelo DPPG do IFSULDEMINAS

Relação de Projetos e Bolsistas PIBIC/CNPq 2011/2012						
Campus	Título		Início	Duração	Término	Área do conhecimento
MUZ	Produção de soja sob diferentes plantas de cobertura em sistema de plantio direto no Sul de Minas Gerais		1/9/11	11	31/7/12	Ciências Agrárias
IFS	Controle Químico <i>Cercospora Zea maydis</i> na cultura do milho		1/9/11	11	31/7/12	Ciências Agrárias
IFS	Estudo da aplicabilidade das plataformas de redes sensores WIRELESS EM AMBIENTES RURAIS		1/9/11	11	31/7/12	Ciências Agrárias

Fonte: Da autora (2023)

Na segunda reunião, foi apresentado aos gestores a estrutura da pesquisa, informações sobre sistemas de apoio à decisão e análise dos dados e a importância da análise de dados e a geração de informações; tratou-se de uma conversa informal e muito proveitosa onde se discutiu questões relacionadas ao problema e preocupações com a disponibilidade de dados/informações para apoiar o processo de tomada de decisão. Estas duas primeiras reuniões tiveram o intuito

de explanação geral sobre a situação problema, levando-os a analisarem a importância da adequada identificação dos requisitos informacionais e acesso aos dados.

Deve-se observar que na pesquisa-ação, entre os participantes da pesquisa está o próprio pesquisador, que também relata suas observações e experiências. As reuniões foram presenciais e duraram aproximadamente uma hora (cada uma).

A terceira reunião com os gestores do setor teve como objetivo identificar quais informações são consideradas importantes pelos tomadores de decisão que devem ser armazenadas no DW (KIMBALL *et al.*, 2008).

Para esta reunião foi necessário a elaboração de um roteiro com a formulação de questões para identificar os requisitos informacionais necessários ao acompanhamento de egressos, como: quais informações são importantes para o acompanhamento de egressos no âmbito da IC (enquanto formação de pesquisador a partir da vivência na IC)? Quais indicadores refletem estas informações? Todas as informações relevantes estão disponíveis? De que forma é o acesso às informações que serão necessárias? Quais as fontes de dados da instituição estão disponíveis para acesso? Os dados disponíveis atendem a todas as necessidades? Os dados são todos da instituição ou serão necessários dados externos? Para os dados externos é possível identificar a fonte desses dados? Quais? Qual a dificuldade para extrair os dados? Essa ação é permitida? Temos acesso? Quais seriam as decisões tomadas com base em um Sistema de Acompanhamento de Egressos? Qual a importância da Gestão de Egressos da IC?

O roteiro da reunião foi elaborado a partir das questões da entrevista semiestruturada desenvolvida por Vidigal (2019), no seu estudo, que tinha como objetivo principal levantar necessidades de informações semelhantes. Houve o cuidado de analisar as questões adequando a linguagem, a estrutura e a sequência das mesmas para se chegar a um roteiro com perguntas básicas, que norteassem a conversa, mas que atingissem e respondessem ao objetivo principal da pesquisa.

Os Gestores elencaram os requisitos mínimos, necessários e importantes para o acompanhamento dos egressos, bem como as fontes dos dados: informação sobre o tempo para conclusão do curso de mestrado e doutorado, percentual de egressos que evoluíram para o mestrado e doutorado, principais áreas de conhecimentos dos egressos mestres e doutores, análise da contribuição na formação de mestres e doutores que trabalham em instituições de ensino e/ou empresas privadas; elencaram também análises de como os egressos de IC ingressaram no IFSULDEMINAS (Ações Afirmativas (AF), Sistema de seleção unificado (SISU), Vestibular), se o egresso da IC é também egresso do ensino médio do IFSULDEMINAS, se foi bolsista no mestrado e doutorado, quantas meninas são mestras e

doutoras, em quais Universidades e estado fizeram a PG, cursar IC impactou a nota durante a graduação, houve influência do orientador na escolha da Universidade de PG, o egresso tem produção científica relevante?

Sobre os indicadores para o acompanhamento dos egressos da IC foram elencados: tempo de titulação do mestrado e doutorado, porcentagem de mestres e doutores por campi, ano e área de conhecimento, principais áreas de titulação no mestrado e doutorado, percentual de Docentes em IES, percentual de pesquisadores em indústrias, forma de ingresso, percentual de egresso do ensino médio (EM), percentual de meninas pesquisadoras.

Sobre a importância da Gestão de acompanhamento dos Egressos da IC, o Gestor respondeu que a informação sobre estas trajetórias permitirá analisar de forma sistêmica a contribuição dos egressos da IC/IFSULDEMINAS no sistema de ensino superior brasileiro (docência) ou no setor privado (pesquisador em instituições privadas) e principalmente na região do Sul de Minas. Salientou ainda que a IES conhecer o destino de seus egressos é de extrema importância.

Sobre as decisões que podem ser tomadas por meio da ferramenta de BI foram elencadas algumas apresentadas no Quadro 3.

Quadro 3 - Relação das decisões para nortear o acompanhamento de Egressos

	Decisões a partir de informações da Gestão de Egressos da IC
1	Analisar se os objetivos do PIBIC/CNPq estão sendo cumpridos (titulação em menor tempo e formação de recursos para a pesquisa).
2	Analisar a possibilidade de investimentos em novas bolsas de IC fomentadas pelo IFSULDEMINAS, conforme o indicador de mestres e doutores.
3	Analisar a partir das Áreas de Conhecimento de formação dos mestres e doutores a necessidade de um planejamento para abertura de cursos de mestrados (áreas de mais procura de nossos egressos).
4	Analisar o percentual de meninas e propor ações de incentivo a meninas na ciência.
5	Analisar a forma de ingresso dos egressos da IC para um melhor planejamento em políticas estudantis.
6	Analisar se o IF contribui na formação de mestres e doutores para as IES ou para empresas - estimular o empreendedorismo.
7	O egresso da IC é também egresso do IFSULDEMINAS no ensino médio integrado? - pensar em políticas para estimular nossos discentes.
8	Analisar se existe um padrão para nossos egressos (universidade ou cursos de PG).
9	Nosso egresso foi bolsista na PG - analisar (em caso de abertura de PPG é um fator importante).
10	Subsídios para reformulação da política de iniciação científica do IFSULDEMINAS.
11	Investimento em Programas de Pós-Graduação <i>Stricto sensu</i> (retorno do egresso)
12	Subsídios para Avaliação Institucional e Política de Acompanhamento de Egressos Institucional.

Fonte: Da autora(2023)

Os dados coletados foram planilhados e analisados (Fase 3 da pesquisa-ação). As reuniões iniciais foram conduzidas visando elaborar um modelo de Gestão de acompanhamento de egressos da IC que permita, dentre outras informações, analisar as trajetórias de estudos na PG *Stricto sensu* dos egressos da IC.

Além das reuniões, foi realizada a coleta de dados por meio de pesquisa documental; primeiro foi analisada a RN-017/2006⁴ que é norma reguladora do PIBIC/CNPq. Os objetivos do PIBIC são: contribuir para a formação de recursos humanos para a pesquisa, e diminuir o tempo médio de permanência dos estudantes na PG (CNPq, 2006). Neste sentido, esses dois requisitos são essenciais no acompanhamento dos egressos da IC.

Ainda na pesquisa documental, foi analisado o estudo realizado pelo Centro de Gestão e Estudos Estratégicos⁵ (CGEE) em 2017 sobre a iniciação científica que apontou que a participação no PIBIC reduz o tempo entre a conclusão da graduação e o ingresso no mestrado, entre outros pontos avaliados. Constatou-se, por exemplo, que a participação em IC configura uma experiência mais rica para o estudante, estimulando o pensamento crítico e a criatividade, reforçando a escolha profissional e despertando o interesse pela pesquisa (CGEE, 2017). Neste relatório também foram observadas experiências objetivas e subjetivas (satisfação com as habilidades desenvolvidas) e as expectativas de trajetória profissional futura na perspectiva dos bolsistas. Entre as experiências foram citados a exposição do bolsista à outra língua, o seu envolvimento com atividades de pesquisa e a experiência de publicação e apresentação dos resultados de pesquisa.

Também foram mapeadas as informações de requisitos considerados importantes por meio da análise das duas revisões integrativas: qualidade do ensino, impacto da formação na PG, inserção e atuação profissional, perfil dos mestres e doutores, vínculo por meio de Grupos de Pesquisa das IES, tempo de conclusão da graduação comparado com tempo curricular e desempenho acadêmico (IRA).

Após a análise documental, foram coletados os dados disponíveis na Diretoria de Pesquisa e Pós-Graduação dos egressos da IC em formato de planilhas eletrônicas. Na Tabela 2 apresenta-se a distribuição dos egressos por ano.

⁴ <https://www.gov.br/observatorio/pt-br/assuntos/programas-academicos/iniciacao-cientifica-e-tecnologica/documentos/regulamento/rn-017-2006-cnpq.pdf>

⁵ <https://www.cgee.org.br/documents/10195/734063/PIBIC-pdf>

Tabela 2 – Distribuição dos Egressos da IC no IFSULDEMINAS

Ano	Quantidade de Bolsista CNPq
2011	5
2012	6
2013	8
2014	9
2015	9
2016	14
2017	14
2018	18
2019	14
2020	15
2021	18
Total Geral	130

Fonte: Da autora (2023)

As respostas da reunião com roteiro semiestruturado, o relatório da pesquisa documental e bibliográfica foram compiladas e planilhadas para análise qualitativa de cunho descritivo. O Quadro 4 apresenta a compilação das necessidades de requisitos (informações e coleta de dados).

Quadro 4 – Requisitos para a Gestão de Acompanhamento de Egressos do IFSULDEMINAS

Requisitos	Fonte dos dados	Endereço das Fontes (Onde e como)	Indicador	Coleta do Requisito
Informação sobre o tempo de conclusão do curso de Mestrado	Dados abertos da CAPES	https://dadosabertos.capes.gov.br/organizacao/diretoria-de-avaliacao	Tempo de titulação (mestrado)	RN-17/2006 – CNPq e DPPG
Informação sobre o tempo de conclusão do curso de Doutorado	Dados abertos da CAPES	https://dadosabertos.capes.gov.br/organizacao/diretoria-de-avaliacao	Tempo titulação (doutorado)	RN-17/2006 – CNPq e DPPG
Quantidade de egressos da iniciação científica evoluíram para o mestrado	Dados abertos da CAPES	https://dadosabertos.capes.gov.br/organizacao/diretoria-de-avaliacao	Quantidades ingressantes e titulados	RN-17/2006 – CNPq e DPPG
Quantidade de egressos da iniciação científica evoluíram para o doutorado	Dados abertos da CAPES	https://dadosabertos.capes.gov.br/organizacao/diretoria-de-avaliacao	Quantidades ingressantes e titulados	RN-17/2006 – CNPq e DPPG
Quais as áreas de conhecimento com mais egressos, mestres e doutores	Dados abertos da CAPES	https://dadosabertos.capes.gov.br/organizacao/diretoria-de-avaliacao	Áreas de conhecimento	Reunião CB e DPPG
Formação de mestres e doutores para as instituições de ensino?	RAIS/LinkedIn	https://www.Linkedin.com/http://rais.gov.br/sitio/consulta_trabalhador_identificacao.jsf	Percentual de Docentes em IES	Reunião CB e DPPG

Requisitos	Fonte dos dados	Endereço das Fontes (Onde e como)	Indicador	Coleta do Requisito
Formação de mestres e doutores para empresas privadas?	RAIS/LinkedIn/MEI	https://www.Linkedin.com/http://rais.gov.br/sitio/consulta_trabalha_dor_identificacao.jsf	Percentual de pesquisadores em indústrias	CGEE-2017, Reunião CB e DPPG
Forma de ingresso dos egressos de IC no IFSULDEMINAS (AF, Enem, Vestibular)	WebGiz e SUAP	www.suap.ifsuldeminas.edu.br e planilhas dos campi (WebGiz)	Forma de ingresso	Reunião CB e DPPG
Egresso da IC foi discente do ensino médio no IFSULDEMINAS	WebGiz e SUAP	www.suap.ifsuldeminas.edu.br e planilhas dos campi (WebGiz)	Percentual de egresso do EM do IF	Reunião CB e DPPG
Foi bolsista no mestrado/doutorado	Dados abertos da CAPES	https://dadosabertos.capes.gov.br/organizacao/diretoria-de-avaliacao	Quantidade de bolsistas	Reunião CB e DPPG
Informação sobre o percentual de meninas pesquisadoras (M/D)	Planilhas CB e DPPG	Planilha de controles da CB/DPPG	Percentual de meninas pesquisadoras	Reunião CB e DPPG
Onde os egressos cursaram a PG/ - Universidade	Dados abertos da CAPES	https://dadosabertos.capes.gov.br/organizacao/diretoria-de-avaliacao	Universidades mais cursadas	Reunião CB e DPPG
Onde nossos egressos cursaram a PG - Distância em relação ao campus de origem	Dados abertos da CAPES	https://dadosabertos.capes.gov.br/organizacao/diretoria-de-avaliacao	Distância	Reunião CB e DPPG
Cursar IC impactou a nota durante a graduação	Questão subjetiva não há dados	Questão subjetiva, não há dados	Questão subjetiva, não há dados	Reunião CB e DPPG
Influência da formação do orientador na escolha da IES de mestrado/doutorado	Questão subjetiva, sem dados	Questão subjetiva, não há dados	Questão subjetiva, não há dados	Reunião CB e DPPG
Produção científica do egresso é relevante	Plataforma <i>Lattes</i>	https://lattes.cnpq.br/	Quantidade de produção científica	Reunião CB e DPPG
Egressos por campus	Planilhas CB e DPPG	Planilha de controles da CB/DPPG	Quantidade de egresso/campus	Reunião CB e DPPG
Egressos por ano/área	Planilhas CB e DPPG	Planilha de controles da CB/DPPG	Quantidade de egressos/ano	Reunião CB e DPPG
Tempo do término da graduação e ingresso no mestrado	WebGiz e SUAP	www.suap.ifsuldeminas.edu.br	Tempo médio - graduação e o mestrado	CGEE-2017
Habilidades decorrentes da experiência com a IC	Questão subjetiva, sem dados	Questão subjetiva, não há dados	Questão subjetiva, não há dados	CGEE-2017
Qualidade do ensino	Questão subjetiva não há dados	Questão subjetiva, não há dados	Questão subjetiva, não há dados	Revisão Integrativa
Desempenho acadêmico na graduação	WebGiz e SUAP	www.suap.ifsuldeminas.edu.br	Média geral (IRA)	Revisão Integrativa
Inserção profissional	RAIS/LinkedIn/MEI	https://www.Linkedin.com/http://rais.gov.br/sitio/consulta_trabalha_dor_identificacao.jsf	Quantidade de egressos trabalhando	Revisão Integrativa
Impacto da IC na pós-graduação	Verificar	Verificar	Verificar	Revisão Integrativa

Requisitos	Fonte dos dados	Endereço das Fontes (Onde e como)	Indicador	Coleta do Requisito
Vínculo com a IES por Grupos de Pesquisa	Verificar	Verificar	Verificar	Revisão Integrativa

Fonte: Da autora (2023)

Os dados disponíveis e relatados nas reuniões foram coletados (baixados em formato de planilha eletrônica) para análise na Fase 3 da pesquisa-ação (análise dos dados). Os dados foram divididos em: Dados Abertos da CAPES, RAIS/MEI/LinkedIn, Plataforma *Lattes*, WebGiz/SUAP, Planilhas de controles de Bolsas da CB/DPPG e “outras bases”.

A CAPES disponibiliza dados abertos à sociedade. O portal Dados Abertos CAPES⁶ (2023), disponibiliza, no ambiente “[...] dados e informações sobre a pós-graduação brasileira, sobre a formação de professores para educação básica e outros temas relacionados à educação”. Foi definido como fonte para a análise os conjuntos de dados do tema “Avaliação da Pós-Graduação *Stricto sensu*” que possui 37 conjuntos de dados. Foram utilizados os seguintes conjuntos de dados: Discentes da Pós-Graduação do Brasil, dos anos de [2013 a 2016], [2017 a 2020] e [2021 a 2024]. Assim, foram baixados dos conjuntos de dados (quadrienal) as planilhas de 2014, 2015, 2016, 2017, 2018, 2019, 2020 e 2021⁷; a coleta das planilhas ocorreu entre dezembro de 2022 e janeiro de 2023.

Para a coleta de informação sobre vínculo empregatício, vigente ou não, foi realizada uma pesquisa na Relação Anual de Informações Sociais (RAIS), do Ministério da Economia. Ela disponibiliza informações do mercado de trabalho, fornecendo dados para a elaboração de estatísticas sobre o mundo do trabalho. A pesquisa foi realizada por meio do portal da RAIS. Salienta-se que a informação se limita ao conhecimento da situação laboral do egresso no ano de 2021, período em que a informação estava disponível no momento da coleta. Para a coleta da informação sobre o cadastro como Microempresário individual (MEI) que permite a emissão do Comprovante de Condição de Microempreendedor Individual (CCMEI) foi realizada pesquisa por meio de acesso pelo Gov.Br; por fim, foram coletadas as informações inseridas na plataforma do LinkedIn pelos egressos da IES; o LinkedIn é uma rede social profissional que facilita o trabalho de recrutadores. O objetivo de coletar os dados nas três plataformas (MEI-LinkedIn e RAIS) foi para comparar as informações coletadas de emprego formal, informal e

⁶ Acesso ao conjunto de dados por meio do link: <https://dadosabertos.capes.gov.br/organization/diretoria-de-avaliacao?q=discentes&sort=>

⁷ Conforme a Portaria CAPES nº 55, de 17 de março 2022, que estabelece o calendário da CAPES para o Coleta ano base 2021, informamos que haverá uma segunda etapa de preenchimento do Coleta desse ano base prevista para encerrar em março de 2023, o que poderá gerar alterações nos dados de 2021.

comprovações do MEI. Os dados foram coletados em dezembro de 2022 e após a extração, foram consolidados em uma planilha eletrônica. A planilha foi elaborada com as opções "Sim" e "Não", representando se o discente esteve associado a algum vínculo empregatício formal, informal ou empreendedor. O protocolo para a coleta de dados, em cada uma das plataformas, foi elaborado e está nos apêndices desta dissertação.

Os dados da produção científica dos egressos foram coletados por meio do acesso ao currículo *Lattes*, baixado em formato .html da Plataforma *Lattes*, e inserido na plataforma *Latticles*⁸, que é um aplicativo web construído com o pacote shiny do software livre R para auxiliar na avaliação de produções declaradas na Plataforma *Lattes* do CNPq. Após a extração, os dados foram consolidados em uma planilha eletrônica.

Os dados acadêmicos dos egressos foram coletados nas plataformas institucionais. O IFSULDEMINAS foi criado em 2008 por meio da junção de 3 escolas técnicas dos municípios de Machado, Inconfidentes e Muzambinho. Os três campi utilizavam um Sistema de Informação Acadêmico chamado WebGIZ. Em meados de 2016 o IF começou a utilizar o SUAP que contém o módulo SUAP/EDU que registra toda a vida acadêmica dos discentes; porém essa mudança de Sistema foi efetuada gradativamente começando pelos campi novos (Passos, Pouso Alegre, Poços de Caldas, Carmo de Minas e Três Corações) seguido pelo campus Machado, Inconfidentes e por último Muzambinho. Neste sentido, para filtrarmos os dados acadêmicos dos egressos da IC foi necessário juntar arquivos do SUAP com os arquivos do WebGIZ de cada campus. O sistema SUAP possui dados mais completos que o sistema WebGIZ; os dados foram exportados para planilha eletrônica. Não foi possível obter a tabela original do WebGiz do Campus Inconfidentes referente aos anos de 2011 a 2016.

Cabe ressaltar que a utilização e a coleta dos dados encontram-se respaldados na Lei Geral de Proteção de Dados Pessoais LGPD, nº 13.709, de 14 de agosto de 2018, que dispõe no artigo 4º, inciso II, alínea "b", e 7º, inciso IV, sobre a inaplicabilidade da lei em dados para fins exclusivamente acadêmicos, e que o estudo precisa estar vinculado a um órgão de pesquisa, e que, sempre que possível, os dados sejam anonimizados, não permitindo a identificação do titular (BRASIL, 2018).

4.3 Fase 3 – Análise dos dados e proposta de plano de ação

Nesta fase da pesquisa-ação foram analisados os requisitos informacionais necessários descritos nas reuniões, e o acesso às bases dos dados. O intuito foi verificar o que eles contam

⁸ Acesso por meio do link: <https://arsilva.shinyapps.io/latticles2/>

conforme as decisões que serão tomadas. Foi necessário entender os dados e verificar como eles devem ser tratados, ou seja, como os valores zerados serão mostrados, quais dados não são necessários e serão excluídos, se os formatos estão corretos, se será necessária fusão de dados e principalmente quais variáveis serão utilizadas na modelagem. Essa análise será útil para o processo de ETL que será realizado na próxima fase da pesquisa-ação.

A planilha com os requisitos foi dividida por acesso às fontes de dados para os requisitos serem analisados de forma mais clara. As fontes de dados foram divididas em: Dados Abertos da CAPES, RAIS/LinkedIn/MEI, Plataforma *Lattes*, WebGiz/SUAP, Planilhas de controle de Bolsas da CB/DPPG e “outras bases”. Nesta fase foi necessário entender e compreender os dados e os objetivos da solução proposta para que se possa fazer bom uso entre a “riqueza dos dados” das diversas bases mencionadas.

Iniciando o processo de análise dos dados, analisaram-se os requisitos da fonte dos Dados abertos da CAPES que atendem a variável relacionada ao Impacto do PIBIC no IFSULDEMINAS e à Formação de pesquisador. Os dados foram extraídos do conjunto de dados “Discentes da Pós-Graduação *Stricto sensu* do Brasil” do Portal de Dados Abertos da CAPES, sendo usadas as planilhas dos anos de 2014 a 2021. Após a análise, foram anotadas as colunas da planilha que podem ser inutilizadas no processo ETL, uma vez que não fornecem dados úteis para gerar informação. Em média, cada planilha tem 400.000 registros (linhas) com 37 colunas. Das variáveis disponíveis foram utilizadas as seguintes: Nome_grande_área_conhecimento, Nome_modalidade_programa, Nome_grau_programa, Nome_programa_IES, Nome_região, Sigla_UF_programa, Nome_município_programa_IES, Faixa_etária, Grau_acadêmico_discente, Situação_ingressante, Nome_situação_discente, Data_matrícula_discente, Data_situação_discente, Qtdade_mês_titulacão, Nome_área_avaliacão, Sigla_entidade_ensino, Grau_acadêmico_discente e Situação_ingressante. O Quadro 5 apresenta o resultado da análise dos requisitos e a viabilidade do uso dos dados. As planilhas da CAPES conseguem responder à maioria dos requisitos; quanto ao requisito para verificar se o egresso foi bolsista no mestrado e doutorado, infelizmente, não há nenhum dado referente ao recebimento de bolsas nas planilhas baixadas das bases da CAPES.

Quadro 5 – Requisitos e indicadores - Dados abertos da CAPES

Requisitos	Indicador	Viabilidade do requisito
Tempo de conclusão do Mestrado	Quantidade de egressos com prazo igual ou inferior a 24 meses	Ok
Tempo de conclusão do Doutorado	Quantidade de egressos com prazo igual ou inferior a 48 meses	Ok

Requisitos	Indicador	Viabilidade do requisito
Quantidade de egressos da IC que ingressaram no mestrado	Quantidade em curso Quantidade Titulados	Ok
Quantidade de egressos da IC que ingressaram no doutorado	Quantidade em curso Quantidade Titulados	Ok
Quais as áreas de conhecimento com mais egressos, M/D	Quantidade de Áreas de conhecimento/Cursos	Ok
Foi bolsista no mestrado/doutorado	Quantidade de bolsistas	Não
Onde os egressos cursaram a PG/ Universidade	Quantidade de Universidades	Ok
Cidade/Estado onde os egressos cursaram a PG - Distância em relação à origem -	Quantidade de Estados	Ok

Fonte: Da autora (2023)

A seguir foi realizada a análise dos requisitos com base nos dados da Plataforma *Lattes* que coletou dados da produção científica dos egressos do PIBIC, uma vez que parte significativa dos egressos ingressaram em cursos de pós-graduação e, assim, teriam seus currículos cadastrados e atualizados na plataforma.

Sanchez (2019) apontou que a Plataforma *Lattes* tem um grande potencial como fonte de dados, mas ainda é pouco explorada pelos pesquisadores. A plataforma pode auxiliar em diversas informações laborais, porém deve-se considerar que os dados podem estar desatualizados. A extração da produção científica foi realizado usando a ferramenta Latticles; após, foram definidas as variáveis que, presentes no Currículo *Lattes*, eram de interesse desta pesquisa, são elas: total_art, total_qualis_A, total_artigos_tec, total_resumos, livros, capítulo de livros, orientação_IC, orient_TCC_graduação, orient_especializacao, patentes, editoração, cursos_minis, mat_didatico, org_eventos, artes_visuais e total_geral_public. O Quadro 6 mostra o requisito e o indicador por meio do acesso aos dados da Plataforma *Lattes*.

Quadro 6 – Requisito e indicador - Plataforma *Lattes*

Requisitos	Indicador	Viabilidade do requisito
Produção científica do egresso é relevante	Quantidade de egressos com produção científica (artigos, qualis A, patentes, entre outros)	Ok

Fonte: Da autora (2023)

O IFSULDEMINAS não possui um Sistema Integrado para controle de bolsistas de IC, neste caso, o controle e manuseio dos dados é baseado somente em planilhas eletrônicas. A análise dos dados da CB e DPPG mostrou que o controle dos dados de egressos de IC não contém todas as informações essenciais para a triangulação com outras bases, sendo necessário, antes da análise, realizar um mapeamento de outros dados para complementar a planilha com a

identificação dos documentos dos bolsistas e conferência dos nomes, modalidade da bolsa e campus de origem. Foram observadas e analisadas as variáveis que podem ser excluídas, pois não são necessárias.

A planilha possui 130 registros de egressos entre os anos de 2011 e 2021. O Quadro 7 apresenta os requisitos, indicadores e sua viabilidade.

Quadro 7 – Requisito e indicador por meio dos dados da CB/DPPG

Requisitos	Indicador	Viabilidade do requisito
Egressos por campus	Quantidade de egresso/campus	Ok
Egressos por Ano/Modalidade	Quantidade de egressos/ano	Ok
Informação sobre meninas pesquisadoras (M/D)	Percentual de meninas nas ciências	Ok

Fonte: Da autora (2023)

A seguir foi realizada a análise dos requisitos com base nos dados das plataformas institucionais com informações acadêmicas e de ingresso no IFSULDEMINAS. A análise dos dados acadêmicos dos egressos nos sistemas do IFSULDEMINAS foi bastante complexa; foi necessário juntar os dados de 2 sistemas diferentes em uma única planilha (Suap, WebGiz do campus Machado e Muzambinho). Os dados das duas planilhas diferem, pois o sistema SUAP tem possibilidade de coletar dados mais completos que o sistema WebGIZ, assim, para filtrar os dados acadêmicos dos egressos da IC foi necessário a junção dos arquivos no mesmo formato. A planilha final, composta pelos três arquivos, contém 74.628 registros e 29 colunas. As variáveis da planilha atendem aos requisitos, sendo elas: Campus, Nome_Discente, Número_documento_Discente, Ano_ingresso, Ano_Conclusão, Descrição_Curso, Cor_raça, Forma_Ingresso, IRA, Município, Naturalidade, Gênero, Tipo_escola_origem, Data de Conclusão do Intercâmbio e Ano_conclusão_ensino_anterior. O Quadro 8 apresenta os resultados da análise dos requisitos.

Quadro 8 – Requisitos e indicadores por meio dos dados do WebGiz/SUAP

Requisito	Indicador	Viabilidade do requisito
Forma de ingresso dos egressos de IC no IFSULDEMINAS (AF, SISU, Vestibular)	Total por forma de ingresso	Ok
Egresso da IC foi discente do ensino médio no IFSULDEMINAS	Percentual de egresso do EM do IF	Não
Prazo de integralização da graduação e prazo entre término da graduação e no M/D	Prazo de integralização	Ok
Desempenho acadêmico na graduação	Total de discentes com IRA maior que 8	Ok

Fonte: Da autora (2023)

Para a análise dos dados da trajetória profissional foram verificados os dados coletados na Plataforma da RAIS, MEI e do LinkedIn. O objetivo desta análise é demonstrar a quantidade de egressos no mercado de trabalho formal, informal, ou atuando como microempreendedor.

Com esses dados foi possível analisar se os egressos são docentes e aqueles que atuam como pesquisadores em empresas/indústrias. O acesso aos dados da RAIS, do Ministério da Economia, é uma forma de verificar se o egresso tem vínculo empregatício vigente ou não. O Quadro 9 demonstra o resultado da análise.

Quadro 9 – Requisitos e indicadores - RAIS/MEI/LinkedIn

Requisitos	Indicador	Viabilidade do requisito
Formamos mestres e doutores para as instituições de ensino?	Percentual de Docentes em IES	Ok
Formamos mestres e doutores para empresas privadas?	Percentual de pesquisadores em indústrias	Ok
Inserção profissional	Percentual de emprego formal/empreendedor e informal	Ok

Fonte: Da autora (2023)

Por fim, foi analisada a tabela com os últimos 6 requisitos coletados, porém, os mesmos não têm dados reais associados, trata-se de questões subjetivas. A maioria desses requisitos requer a coleta de dados diretamente com os egressos, por meio de questionários (survey). No entanto, a parte subjetiva não será abordada nesta pesquisa, pois o objetivo é trabalhar com fonte de dados secundários. No Quadro 10 são apresentados os resultados da análise.

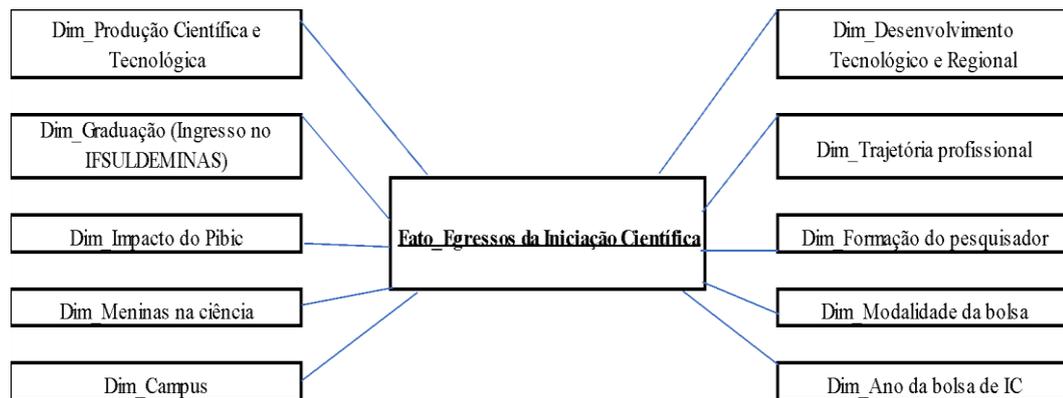
Quadro 10 – Requisitos e indicadores por meio de dados de outras bases

Requisitos	Indicador	Viabilidade do requisito
Cursar IC impactou a nota durante a graduação	Não há	Não temos dados
Influência da formação do orientador na escolha da IES de M/D	Não há	Não temos dados
Habilidades decorrentes da experiência (IC)	Não há	Não temos dados
Qualidade do ensino	Não há	Não temos dados
Impacto da IC na pós-graduação	Quantidade de egressos M/D	Ok
Vínculo com a IES por Grupos de Pesquisa	Não há	Não temos dados

Fonte: Da autora (2023)

Dessa forma, esta etapa de análise e identificação de quais requisitos serão trabalhados e como respondem à questão da pesquisa foi finalizada. Os requisitos para a solução foram identificados, organizados e disponibilizados para a próxima fase da pesquisa-ação. Ao se analisar e avaliar as fontes de dados, definiu-se um processo de atualização anual das planilhas que necessitam da extração de algum sistema. A Figura 14 apresenta a primeira versão do modelo conceitual do DW.

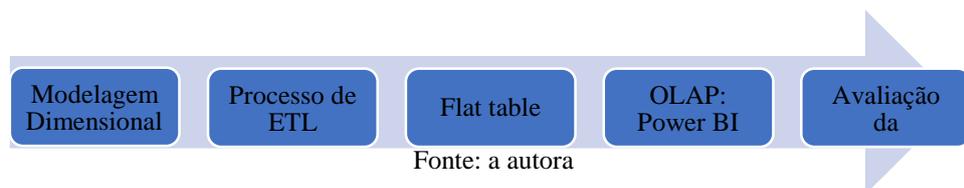
Figura 14 – Modelo Conceitual do *Data Warehouse*



Fonte: Da autora (2023)

De acordo com Mello *et al.* (2012), um dos objetivos da pesquisa-ação é a resolução de um problema real/prático, logo, o final desta etapa se deu com a elaboração do plano de ação para o desenvolvimento do ambiente de BI. Este plano incluiu todas as recomendações para a solução do problema. A Figura 15 apresenta o plano de ação para o acompanhamento dos egressos da IC.

Figura 15 – Plano de ação para o desenvolvimento da proposta de solução



Fonte: a autora

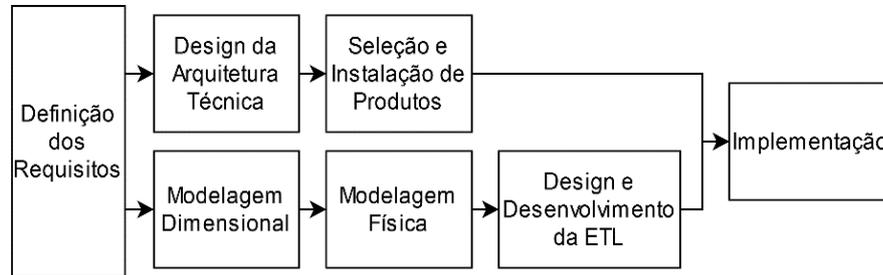
- 1) Realizar a modelagem dimensional do DW.
- 2) Processar o ETL.
- 3) Disponibilizar a *flat table* do DW para a análise dos dados.
- 4) Desenvolver os *Dashboards* por meio do Microsoft Power BI.
- 5) Apresentar a ferramenta para avaliação dos Gestores e realizar um novo ciclo da pesquisa-ação, se necessário.

4.4 Fase 4 – Implementação do plano de ação – Desenvolvimento do *Data Warehouse*

Nessa fase, descreveu-se a execução do plano de ação para o desenvolvimento da ferramenta, incluindo a modelagem dimensional, o processo de extração, transformação/tratamento e carregamento dos dados (ETL), a disponibilização da *flat table* (DW), o desenvolvimento das interfaces orientadas aos usuários finais por meio do software Microsoft Power BI e a avaliação da ferramenta pelos gestores do IFSULDEMINAS. Esta fase fundamentou-se no referencial teórico apresentado na seção 2.2. O método de Ralph Kimball

foi o escolhido para o desenvolvimento do DW do BI. As etapas de desenvolvimento de um DW propostas por Kimball são apresentadas na Figura 16.

Figura 16 - Ciclo de vida do *Data Warehouse*



Fonte: Adaptado de Kimball *et al.* (2008)

Na definição dos requisitos foram utilizados os resultados da fase 3 da pesquisa-ação, mais especificamente a seção 4.3, onde foram definidos os requisitos informacionais, as fontes dos dados e os indicadores para o acompanhamento de egressos da IC.

Para o design da arquitetura técnica: os sistemas de banco de dados que serviram como fonte de dados para o DW foram identificados conforme descrito na seção 4.3. No Quadro 11 eles são apresentados com uma breve descrição.

Quadro 11 – Relação das fontes de dados

Sistemas de origem ou Banco de dados	Descrição
SUAP e WebGiz	Sistemas utilizados para registro do controle acadêmico dos cursos do IFSULDEMINAS.
Dados abertos da CAPES	Disponibiliza dados e informações da Pós-graduação brasileira.
RAIS e MEI (ME) e LinkedIn.	Disponibiliza informações sobre o trabalho formal (registrado) e sobre a comprovação de Microempresário Individual; rede social profissional facilitadora.
Plataforma Lattes/CNPq	Disponibiliza dados e informações do registro da vida pregressa e atual dos estudantes e pesquisadores do país.
Planilhas de controle de bolsistas da IC	Sistema de gestão e administração dos dados dos bolsistas de iniciação científica do IFSULDEMINAS.

Fonte: Da autora (2023)

Foi utilizada a abordagem *Bottom-Up*, pois o objetivo foi desenvolver um *Data Warehouse* específico para acompanhamento dos egressos da iniciação científica. Outro motivo para usar essa abordagem é não haver uma estrutura específica para o desenvolvimento de um DW geral. Este tipo de implementação não requer uma infraestrutura complexa, é mais fácil de ser desenvolvida, apresenta retornos rápidos e pode-se reaproveitar dimensões a cada novo processo, pois nessa arquitetura as dimensões podem ser compartilhadas.

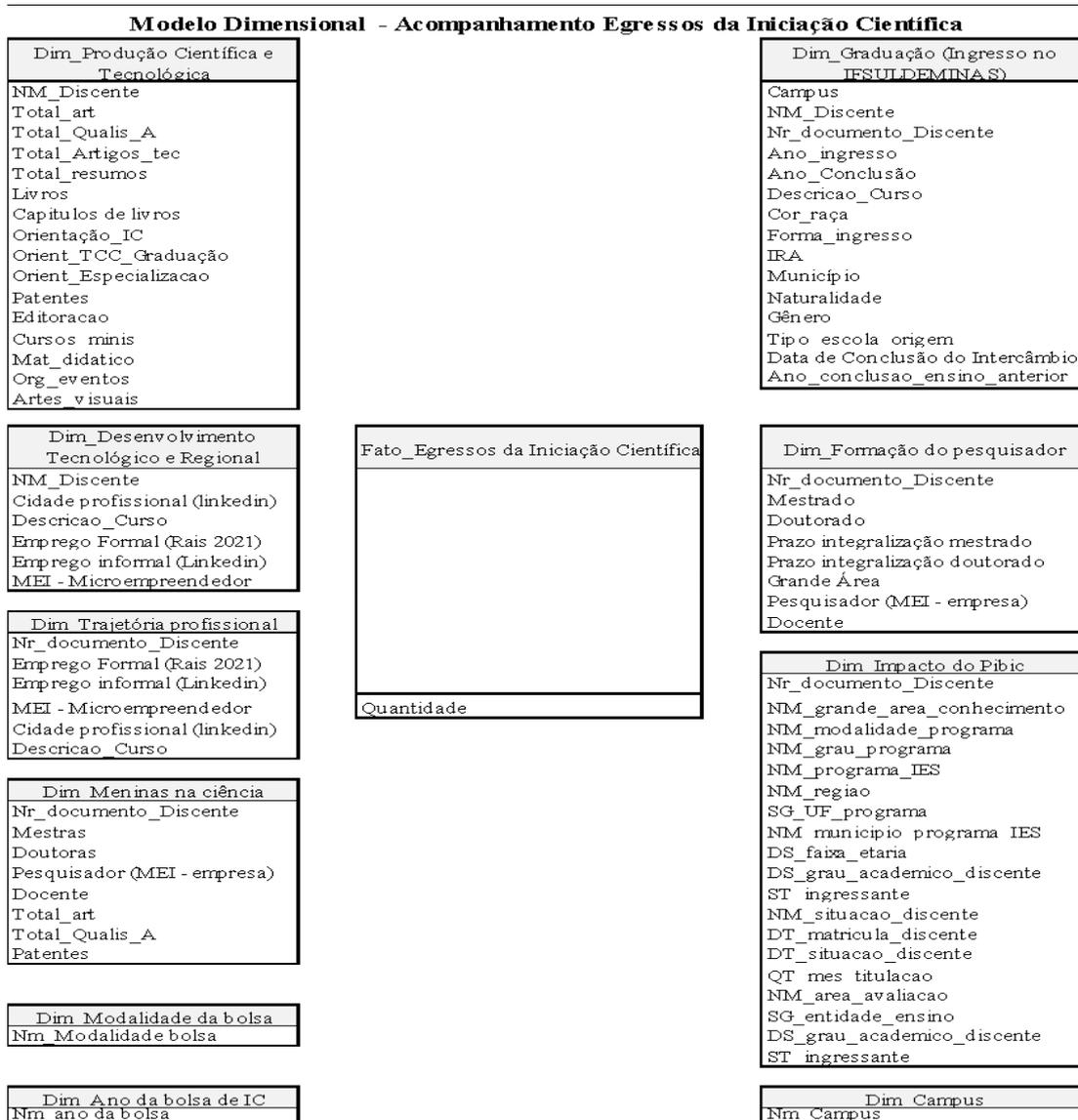
No processo de integração de dados optou-se por utilizar ferramentas de fácil compreensão e manuseio, entre elas o Microsoft Excel e o Microsoft Access; os dados foram copiados dos sistemas pelas ferramentas de ETL, após foram efetuadas as transformações e o

resultado foi carregado no DW. Após o desenvolvido o DW, os dados ficaram disponíveis para utilização, em formato de relatórios de Dashboards, por meio do Microsoft Power BI para acesso pelos usuários finais.

No desenvolvimento da Modelagem Dimensional, é preciso definir alguns conceitos como granularidade, dimensões e fatos. Granularidade é o menor nível de detalhe que se deseja analisar. A definição da granularidade pode afetar a capacidade de armazenamento dos dados e o nível de análise dos dados. Na Tabela de Fato são associadas às informações quantitativas do tema e as tabelas de Dimensão referem-se às variáveis associadas às tabelas de Fatos.

Assim, foi utilizada a técnica de modelagem dimensional de dados, em um modelo estrela, detalhando a tabela de Fato e as tabelas de Dimensão. A Figura 17 apresenta a modelagem dimensional.

Figura 17 – Modelo Dimensional do Data Warehouse



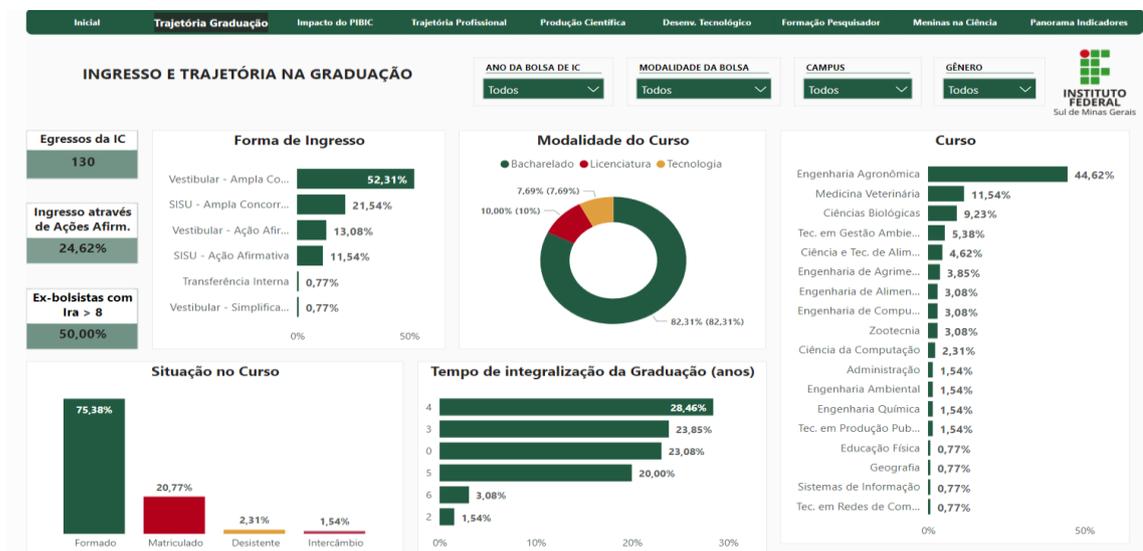
Na extração dos dados da CAPES e sistemas acadêmicos do IFSULDEMINAS foram realizadas diversas consultas utilizando o Microsoft Access para popular as tabelas dimensão relacionadas aos dados que seriam analisados. No processo de transformação para padronização dos dados foram realizados: limpeza de dados, formatação, transformação de campos nulos e dados incompletos, filtros, união, divisão, agregação de dados, extração de linhas e colunas, conversão de valores, dentre outros. Após os tratamentos necessários, a base ficou pronta para ser executada de acordo com a modelagem definida. Informações como nome, RG, CPF foram descartados para que não permitisse a identificação direta de qualquer egresso, mantendo a base e o estudo conforme a LGPD. Assim, a base de dados foi carregada para uma “flat table” que foi importada para o software Power BI desktop para o desenvolvimento das interfaces para os usuários finais.

Houve diversos testes para analisar e avaliar os gráficos e tabelas que melhor apresentariam os dados aos gestores; assim, foram escolhidos alguns layouts que melhor se adequam à análise dos indicadores. Para cada dimensão do Modelo dimensional, criou-se uma página/tela e cada uma delas apresenta gráficos, indicadores, filtros e agrupamentos diferentes. Foi criada uma primeira versão destes *dashboards* (relatórios) para os gestores avaliarem a ferramenta. Foram criados oito *dashboards* principais para apresentar as informações.

A primeira versão teve como objetivo responder aos requisitos informacionais para o BI de acompanhamento dos egressos e as análises apresentadas nas reuniões. Cada página da apresentação dos relatórios de *dashboards* apresenta uma dimensão analisada.

A página 2 do relatório de *Dashboards* responde aos indicadores e análises relacionados aos ex-bolsistas durante a graduação no IFSULDEMINAS, conforme Figura 18.

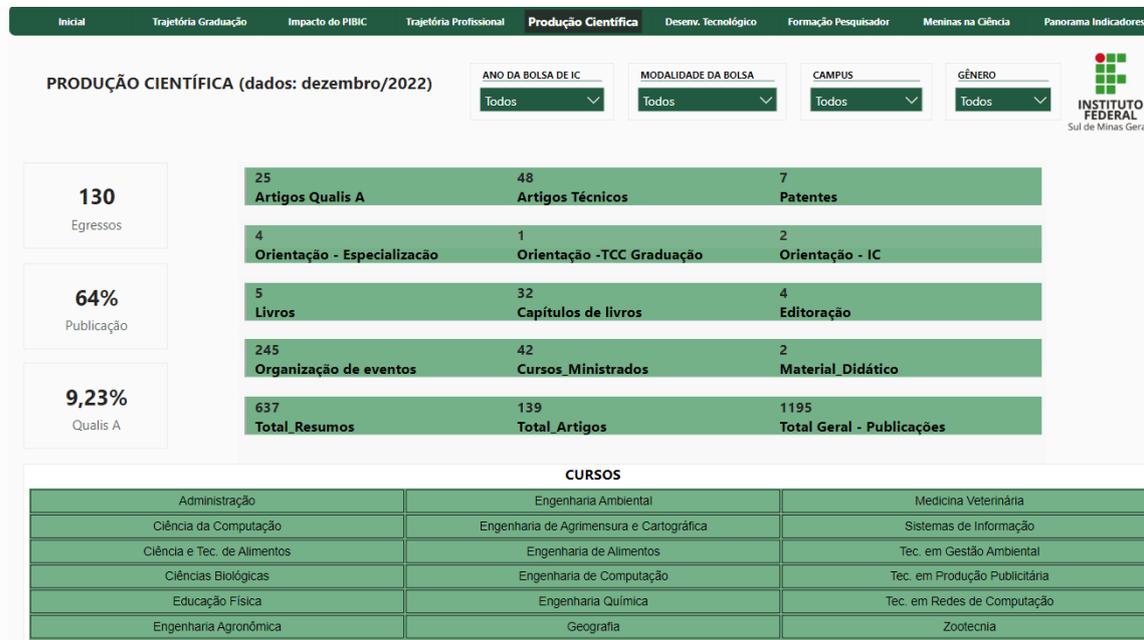
Figura 18 – Dashboard: Trajetória Acadêmica



Fonte: Da autora (2023)

A página 5 do relatório de *Dashboards* responde aos indicadores e análises relacionados à produção científica dos ex-bolsistas do IFSULDEMINAS, conforme Figura 21.

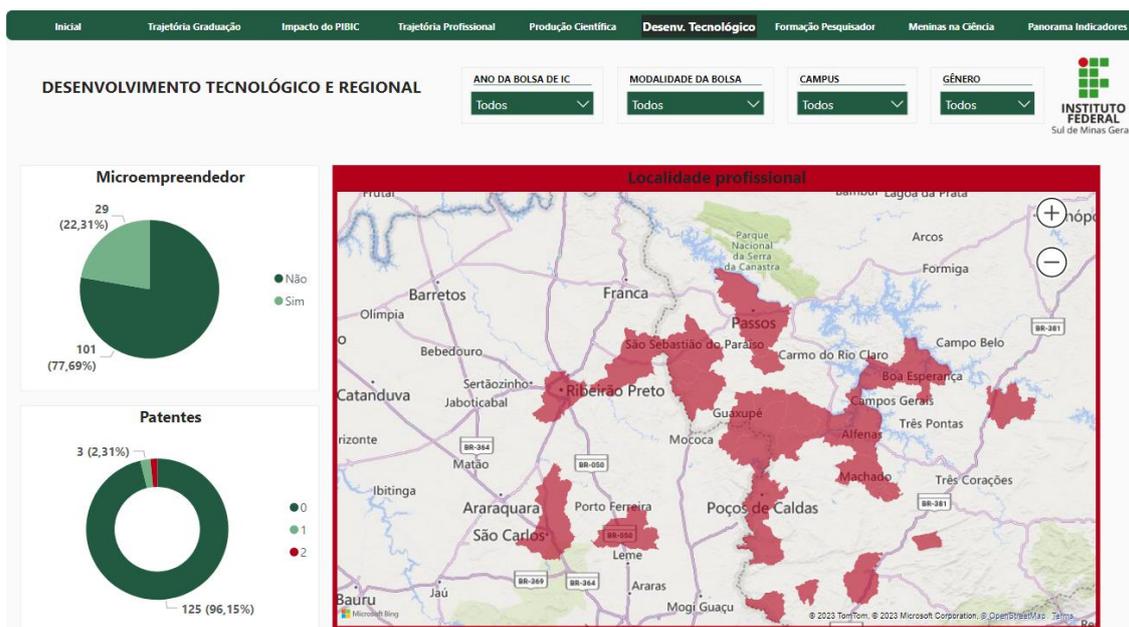
Figura 21 – Dashboard: Produção Científica



Fonte: Da autora (2023)

A página 6 do relatório de *Dashboards* responde aos indicadores e análises relacionados ao desenvolvimento tecnológico e regional, conforme Figura 22.

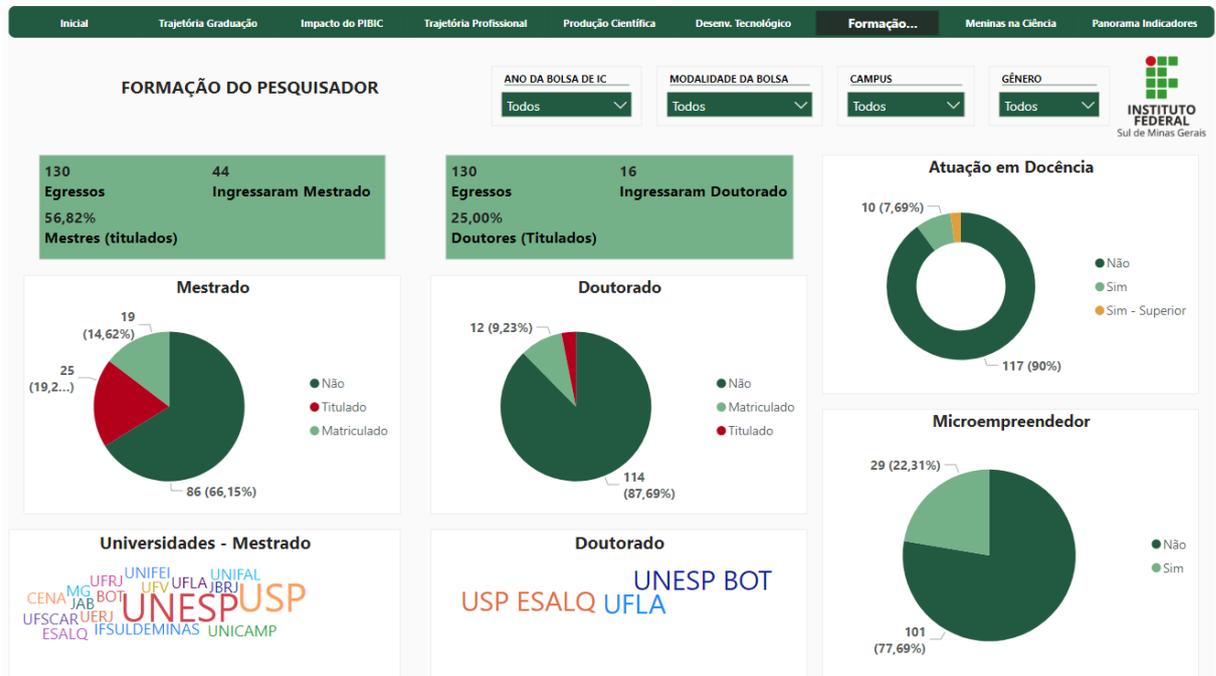
Figura 22 – Dashboard: Desenvolvimento Tecnológico e Regional



Fonte: Da autora (2023)

A página 7 do relatório de *Dashboards* responde aos indicadores e análises relacionados à formação do pesquisador, conforme Figura 23.

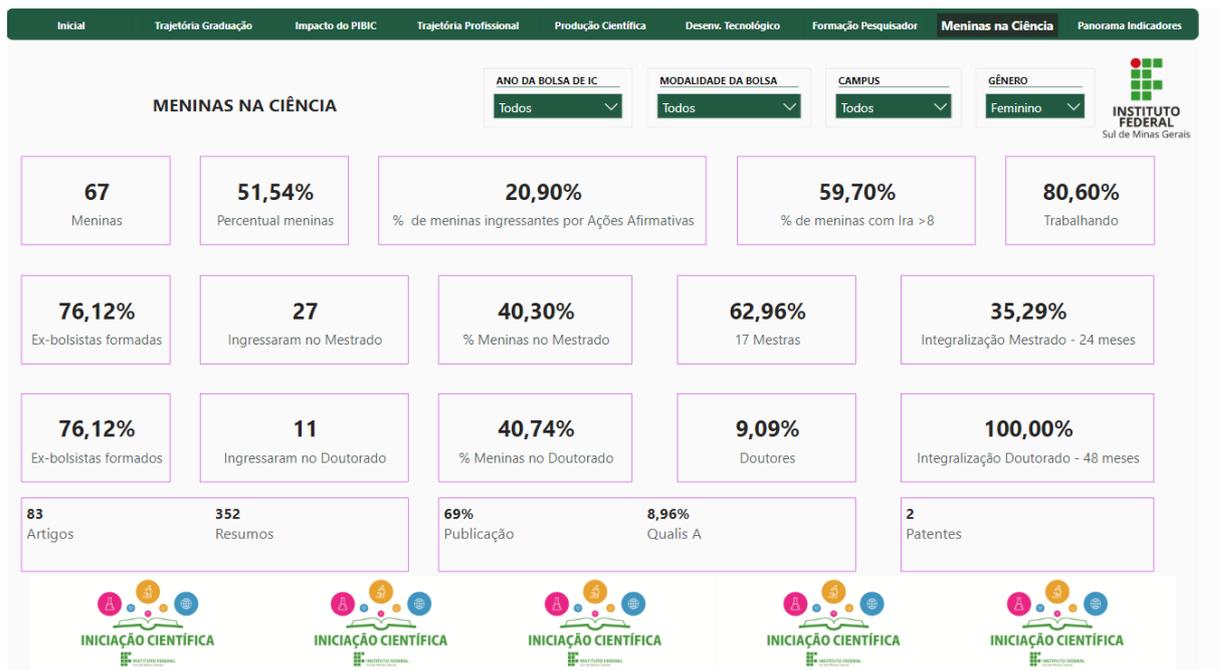
Figura 23 – Dashboard: Formação do Pesquisador. UNIFEI, Itajubá/MG, 2023.



Fonte: Da autora (2023)

A página 8 do relatório de *Dashboards* responde aos indicadores e análises relacionados à relação das meninas na ciência, conforme Figura 24.

Figura 24 – Dashboard: Meninas na Ciência



Fonte: Da autora (2023)

4.5 Fase 5 – Avaliação (validação) da ferramenta

A avaliação da ferramenta pelos gestores da IC no IFSULDEMINAS é descrita nesta fase. A validação é uma fase importante para apresentar a ferramenta desenvolvida com dados reais dos egressos da IC. Um ambiente de homologação foi disponibilizado para testar a ferramenta e validar com mais propriedade. A validação tem como objetivo verificar se a ferramenta responde ao problema da pesquisa. A apresentação foi realizada por meio de reunião realizada em março de 2023 com os mesmos gestores que participaram das reuniões iniciais e que apresentaram as necessidades dos requisitos. Na apresentação foi informado o processo de seleção e análise dos requisitos indispensáveis à Gestão de Egressos da IC, a base teórica da pesquisa (*Business Intelligence, Data Warehouse* e impacto do PIBIC na PG), quais as fontes de dados utilizados, a lógica no processo de construção do modelo e o treinamento para uso da ferramenta. Além da apresentação aos gestores, a ferramenta foi apresentada a alguns docentes que desenvolvem pesquisas em temáticas relacionadas a egressos, IC e análise de dados.

Nestas reuniões foram apontadas algumas inconsistências nos gráficos com solicitação de substituição de alguns modelos, porém, os gestores concordaram que eles são de fácil interpretação, bem organizados e com informações claras e objetivas. Com base nas avaliações a ferramenta foi validada; todos os gestores concordaram que os *dashboards* fornecem as informações necessárias para a gestão dos egressos da IC do IFSULDEMINAS.

Dessa forma, foram realizadas as adequações solicitadas; após a realização dos ajustes solicitados os relatórios de *Dashboards* foram apresentados a um grupo de Coordenadores de Pesquisa e PG dos campi do IFSULDEMINAS, onde surgiram novas solicitações de ajustes, como, por exemplo, o estudo de metas relacionadas aos indicadores para que se possa acompanhar o alcance ou não da meta. Apesar de, no escopo do projeto, ter sido delimitado a definição de requisitos básicos para o acompanhamento dos egressos, acolheu-se a solicitação de implantação de metas para alguns indicadores, como, por exemplo, uma meta de ingressantes nos cursos de mestrado/doutorado; porém essa meta só será implementada após a realização da análise dos egressos da FAPEMIG no mesmo período (2011 a 2021) para que se tenha o panorama e seja possível traçar uma meta

Também foram solicitados alguns ajustes em termos (palavras) usados nos dashboards, como: correções pontuais de texto, inclusão, correção e exclusão de outros termos. Todas as sugestões foram acatadas.

Nas reuniões de apresentação surgiram outras sugestões, como comparação dos resultados com indicadores nacionais e internacionais; foi explicado aos presentes que não se

tratou, nesta pesquisa, analisar e comparar indicadores, mas sim de apresentar uma ferramenta que pudesse demonstrar por meio de dados a trajetória acadêmica e profissional dos egressos da iniciação científica enquanto formação do pesquisador. Porém, todas as sugestões, importantes e interessantes, foram anotadas e serão descritas como sugestão de pesquisas futuras.

A ferramenta foi validada por todos (gestão e convidados) com os devidos ajustes. A resposta positiva dos avaliadores foi gratificante. O orgulho dos mesmos ao notar os resultados positivos das análises foi a melhor parte.

5 DISCUSSÃO

Nos últimos anos, o número de estudos que se interessam pela trajetória acadêmica e profissional de egressos da PG no mundo do trabalho, aumentou, porém, pesquisas que cruzassem os descritores “PIBIC”, “Pós-Graduação” e “formação de recursos humanos para a pesquisa” foram poucas. Neste sentido, a ferramenta de *Business Intelligence* desenvolvida neste estudo contempla a escassez de estudos na temática; o estudo apresenta importantes variáveis como: trajetória acadêmica, trajetória profissional (trabalho formal, informal e microempreendedor), impacto do PIBIC na PG, desenvolvimento tecnológico e regional, produção científica, formação de recursos humanos para a pesquisa e o papel das meninas na ciência) que proporcionam um olhar reflexivo sobre seus resultados.

Cabe ressaltar que as diversas variáveis usadas nas análises do BI desenvolvido foram corroboradas pela literatura. Neste sentido, com a ferramenta de BI é possível acompanhar a trajetória dos ex-bolsistas de IC (antes, durante e após a titulação em cursos de PG), para avaliar a formação recebida quanto à inserção no mercado de trabalho e a sua contribuição regional.

Assim, com o protocolo de desenvolvimento da pesquisa-ação concluído e os resultados descritos, esta parte da dissertação discute os achados da seção anterior e suas contribuições para a teoria. Para isso, foram utilizados estudos presentes nas literaturas nacional e internacional. Selecionou-se aqueles que direcionam de maneira semelhante para as características exploradas durante esta pesquisa. Alguns estudos citados na seção 3 (referencial teórico) são resgatados e auxiliam na discussão dos resultados encontrados.

Para organizar a discussão, esta seção foi dividida em duas partes: a primeira parte discorre sobre a conexão entre os requisitos de informação e o uso da metodologia de Kimball no desenvolvimento do DW/BI, e a segunda parte sobre o resultado da análise por meio da ferramenta de BI para o Acompanhamento de Egressos/Gestão de Egressos.

Requisitos de informação e utilização da Metodologia de Kimball

O processo de definição dos requisitos de informação viáveis e desejáveis do BI para acompanhamento dos egressos da IC foi realizado por triangulação de dados: pesquisa bibliográfica, pesquisa documental e reuniões com os Gestores. A identificação das necessidades de requisitos de informação é um pré-requisito para as próximas ações e atividades no desenvolvimento de ferramentas para apoio à tomada de decisão/BI. Esses requisitos foram identificados e definidos após a análise da viabilidade de acesso às fontes dos dados e após a constatação de que esses forneceriam as informações desejadas. As pesquisas bibliográfica e

documental corroboraram com os requisitos informados nas reuniões com os gestores, sendo, neste caso, necessidades de requisitos já mencionados em estudos anteriores.

Oraee, Sanatjoo e Ahanchian (2021) usaram a entrevista para determinar os requisitos; o estudo destaca três problemas clássicos na definição das necessidades de informação por meio da coleta de dados com gestores: aqueles que não expressam suas necessidades, aqueles que querem saber de tudo e acreditam que saberão o que precisam quando o encontrarem e aqueles que respondem “você me diz o que eu preciso”. O estudo indicou que a correta identificação dos requisitos de informação e acesso aos dados podem ajudar na arquitetura dos sistemas de informações gerenciais para a tomada de decisão. Na mesma linha de raciocínio, a pesquisa de Mendes (2020, p. 82), destacou que “o levantamento dos requisitos do DW é a primeira e mais importante etapa do seu desenvolvimento”; na pesquisa de Mendes (2020), os requisitos de informação surgiram da triangulação de dados secundários (pesquisa documental e bibliográfica) e foi possível conhecer a interseção das informações desejáveis no DW de acordo com o que é possível ser atendido pelos dados disponíveis no sistema da IES estudada. Destaca-se que os requisitos de informação são necessários para determinar a granularidade dos dados no DW ou seu nível de detalhamento (KIMBALL *et al.*, 2008).

O desenvolvimento de DW locais e a utilização da técnica de modelagem dimensional tornam a abordagem de Kimball mais acessível a profissionais sem formação na área de TI. Assim, após a correta definição dos requisitos foi possível aplicar o método de Kimball para o desenvolvimento do DW desta pesquisa, envolvendo a modelagem dimensional (tabela de fato e dimensões), o processo de extração, tratamento e carregamento (ETL) dos dados na estrutura do DW (*flat table*) e desenvolvimento da área de visualização dos dados aos usuários por meio do software Microsoft Power BI. Cabe relatar que uma das dificuldades encontradas nesta pesquisa foi o processo de ETL relacionado as tabelas dos dados oriundos das plataformas institucionais e da CAPES.

Na pesquisa de Da Silva Souza (2021) o DM foi desenvolvido usando a metodologia de Kimball e ferramentas de BI, permitindo ao gestor obter informações globais e detalhadas sobre a gestão de pessoas da instituição. Mendes (2020) usou com êxito, em sua pesquisa, o método de Ralph Kimball para desenvolvimento do DW. Vendrúscolo (2020) desenvolveu um sistema de BI para a extensão universitária utilizando a metodologia de Kimball no desenvolvimento do DW e para a visualização dos *Dashboards* foi utilizado o software Power BI. A ferramenta desenvolvida foi usada pela PROEX da IES para a elaboração do PDI 2020/2024. Silva (2018) salientou que, no desenvolvimento do DW/BI, o maior problema foi a obtenção e manipulação de dados da Plataforma Sucupira, pela estrutura em que eles são disponibilizados, e o processo

ETL, mais especificamente na transformação dos dados, foi o passo mais moroso da criação da estrutura de BI.

Por fim, pode-se afirmar que a conexão entre a correta definição dos requisitos e a utilização da metodologia de Kimball no desenvolvimento do DW/BI foi essencial nos resultados positivos vislumbrados na literatura.

A ferramenta desenvolvida responde aos requisitos informados na fase de coleta de dados e os resultados apresentados por meio dos relatórios de *dashboards* permitem a interação e análises dos gestores e pessoas externas em cada uma das dimensões e conforme as segmentações disponíveis. Por meio dos *dashboards* é possível visualizar a forma de ingresso, prazo de integralização dos cursos, os cursos que mais absorvem discentes na IC, o impacto do PIBIC nos cursos *Stricto sensu*, a trajetória profissional dos egressos, incluindo o emprego formal, MEI e emprego informal; além de apresentar produção científica relevante dos egressos, produção tecnológica, impacto no desenvolvimento regional e o protagonismo das meninas egressas.

Na pesquisa de Ismail (2022), o processo de criação do *dashboard* foi dividido em três etapas principais: concepção, visualização e finalização. O objetivo foi visualizar os dados dos egressos por meio de técnicas de visualização de dados multidimensionais com opções de segmentações no painel para que pessoas externas interajam com os dados apresentados. Handane (2022) propôs o desenvolvimento de um framework para ligar egressos a recrutadores por meio de uma arquitetura de um banco de dados real; o projeto está em fase de aperfeiçoamento para processamento de consultas para aplicações de algoritmos complexos que permitam consultas analíticas mais rápidas. A pesquisa de Niño (2020) elaborou uma proposta de governança de BI em universidades, incentivando iniciativas de projetos de *Business Intelligence*, transformando-se em insumos indispensáveis para a tomada de decisões que gerem valor às IES.

Em seguida, são abordadas as análises dos requisitos por meio do DW/BI desenvolvido para o acompanhamento de egressos da IC comparando-os com os resultados apresentados na literatura.

BI para Acompanhamento de Egressos de IC/Gestão Universitária

Na identificação dos requisitos para a gestão do acompanhamento dos egressos da IC, foram relacionadas às dimensões/variáveis e as fontes de dados para a gestão do acompanhamento do egresso, conforme apresentado no Quadro 12.

Quadro 12 – Requisitos para o DW de Acompanhamento de Egressos da IC

Dimensões/Variáveis	Fonte dos dados
Produção Científica e Tecnológica	CV <i>Lattes</i> /CNPq
Desenvolvimento Tecnológico e Regional	CV <i>Lattes</i> /CNPq – RAIS, LinkedIn e MEI
Graduação (Ingresso no IFSULDEMINAS)	SUAP e GIZ
Trajatória profissional	RAIS, LinkedIn e MEI
Impacto do Pibic	CAPES – tabelas da SUCUPIRA
Meninas na ciência	Cv <i>Lattes</i> /CNPq - RAIS, LinkedIn e MEI - CAPES
Formação do pesquisador	RAIS, LinkedIn e MEI - CAPES

Fonte: Da autora (2023)

Ao analisar, na literatura, os resultados que tratam da variável produção científica, é importante mencionar a utilização da plataforma *Lattes* do CNPq como fonte de dados para a análise. Para a extração dos dados desta pesquisa foi utilizada a ferramenta Latticles na extração da produção científica.

Para Sanchez (2019, p. 23) “embora tenha um grande potencial como fonte de dados, a Plataforma *Lattes* é relativamente pouco explorada por pesquisadores brasileiros”. Altran *et al.* (2015) analisaram a produção científica de docentes permanentes e colaboradores dos PPGs em Ciência da Informação, entre os anos de 2008 a 2012. A metodologia incluiu a extração automática das informações pelo ScriptLattes23.

Alguns estudos nacionais podem ajudar a compreender os resultados encontrados nesta pesquisa sobre o impacto do PIBIC na PG e sobre a formação do Pesquisador. Na população analisada (130 egressos) neste estudo por meio da ferramenta desenvolvida, o impacto dos egressos na PG *Stricto sensu* foi de 33,85% para ingressantes no mestrado e 36,36% para ingressantes nos programas de doutorado. Não foi intenção dos pesquisadores confrontar este percentual com outras IES ou compará-los com metas ou indicadores de outras instituições, ou com agências de fomento; a intenção foi demonstrar, por meio das análises apresentadas, os resultados do impacto da IC dos ex-bolsistas analisados relacionados a sua absorção pelos PPGs.

Oliveira *et al.* (2020) conduziram um estudo sobre os resultados do PIBIC realizado com 3.459 discentes de graduação; foi identificado que a taxa de encaminhamento aos cursos de PG *Stricto sensu* foi de 44% para os discentes que participaram de iniciação científica na graduação. O estudo apontou que 19,3% dos que concluíram o mestrado ingressaram no doutorado. No estudo de Ângelo (2020) verificou-se a eficácia e impacto do PIBIC na área da Ciência da Informação, contribuindo para a formação de recursos humanos para a pesquisa.

Souza (2018) evidenciou, entre seus achados, que discentes ex-bolsistas de IC têm maiores taxas de aprovação em processos seletivos de cursos *Stricto sensu*.

A formação de pesquisadores é a base para o desenvolvimento científico e tecnológico. Neste sentido, ser bolsista de IC, além de possibilitar a vocação para a ciência, é o primeiro passo na formação de mestres e doutores que contribuem para a produção científica, tecnológica e regional.

Para Santos (2021) além de apresentar o impacto do PIBIC em PPGs o estudo contribuiu nas disparidades de gênero; as mulheres são a maioria enquanto bolsistas de IC e tituladas em cursos *Stricto sensu*. É comum encontrar estudos com análise de população de bolsistas de IC, que demonstram maior quantidade de discentes do sexo feminino (SILVA, *et al.*, 2020; OLIVEIRA *et al.*, 2020; ARAÚJO; ANDRIOLA, 2020). Assim, o resultado da análise pelo BI, mostra que 51,54% da população analisada são meninas e a absorção destas meninas pelos PPGs foi relevante, dado importante para demonstrar o sucesso das meninas na ciência.

As discentes eram 54% de todos os estudantes *Stricto sensu*, matriculados no ano de 2021, no Brasil (CAPES, 2023). No entanto, apesar de as discentes do sexo feminino serem maioria nas pesquisas relacionadas a IC e a maioria de matrículas nos cursos *Stricto sensu*, uma pesquisa recente da Unesco (2023) constatou que apenas 35% dos estudantes das áreas de ciência, tecnologia, engenharia e matemática, as quais são consideradas fundamentais para as economias nacionais, são mulheres (UNESCO, 2023).

A literatura carece de estudos que abordem aspectos profissionais (trajetória profissional/mercado de trabalho) junto aos discentes egressos da IC. Para a coleta de dados dos egressos desta pesquisa foi utilizado a RAIS, a plataforma LinkedIn e o cadastro como MEI; a intenção foi analisar a trajetória profissional de emprego formal, informal ou empreendedorismo.

Para Cabral (2021, p. 142), “as mídias sociais são ferramentas fundamentais, sobretudo a rede social LinkedIn, que podem evidenciar a trajetória profissional dos egressos”. O autor salientou que a plataforma pode auxiliar em informações laborais, contudo, fica restrita aos dados informados e que podem não estar atualizados (CABRAL, 2021). Para Blanco Rojas *et al.* (2016) há algum tempo as redes sociais tomaram muita força nas organizações e instituições, elas são hoje as fontes mais utilizadas para encontrar informações mais completas e verdadeiras relacionadas ao campo profissional. Segundo os autores, as redes sociais são as principais fontes de dados, além de serem um meio de comunicação inovador, com grande conteúdo informativo (BLANCO ROJAS *et al.*, 2016).

A pesquisa de Galvão *et al.* (2017) apresentou a trajetória laboral dos mestres e doutores entre 2009 e 2014, analisando as condições de emprego e a relação que os egressos mantêm com sua área de formação pós-graduada, usando para isso informações da CAPES e da RAIS. Oliveira *et al.* (2020) identificaram que 51% dos egressos do PIBIC, de universidades da Bahia, atuam em atividade de docência, considerando aqui os níveis de educação básica. Em Novais (2020), os egressos da IC exibiram maior chance de empregos quando o assunto foi a área de educação de maneira geral.

A influência da IC no período de graduação dos ex-bolsistas demanda muita reflexão. Nem todas as variáveis associadas à graduação apresentam variedade ou profundidade de investigações que auxiliem a descrever, discutir e embasar resultados descobertos, pois a IC possui como característica principal o desenvolvimento desses estudantes relacionados principalmente a aspectos científicos.

A literatura demonstra a possibilidade de melhor desempenho em sala de aula devido às habilidades geradas pela participação como bolsista em projetos de pesquisa, pois os mesmos favorecem aspectos como a escrita, pensamento crítico, expressão oral, organização de tarefas, entre outros (NARDINI *et al.*, 2019). No mesmo sentido, o estudo de Araújo e Andriola (2020) realizado com bolsistas do PIBIC, evidenciou que 88% concordam que a participação em IC despertou o interesse pela pesquisa científica e a seguirem em cursos de PG. Nardini *et al.* (2019) compararam o rendimento acadêmico entre discentes vinculados a IC e outros que não participaram, da Faculdade São Leopoldo Mandic, nos anos de 2013 a 2016; os resultados indicaram que os estudantes vinculados a IC exibiram coeficiente de rendimento acadêmico maior do que os estudantes que não participaram de IC.

Nas análises proporcionadas pela ferramenta de BI constatou-se que 50% dos ex-bolsistas têm IRA (índice de rendimento acadêmico) maior que 8, o que é um percentual bastante positivo, sendo que a nota mínima para aprovação é 6. Além disso, verificou-se que a área de Ciências Agrárias é a que concentra a maior quantidade de ex-bolsistas, e a área na qual a grande maioria dos egressos foi absorvida pelos PPGs. Entre os cursos da área de Ciências Agrárias, o de Engenharia Agrônoma tem 44,62% dos egressos da IC, além disso, 80% dos egressos que seguiram na PG são dessa área. Isto se justifica, pois o IFSULDEMINAS foi formado pela união de três Escolas Agrotécnicas do sul de Minas Gerais.

Na análise de resultados sobre a integralização do curso analisando a trajetória dos egressos do PIBIC na graduação, no IFSULDEMINAS, podemos citar algo que chamou a atenção, na população analisada (130 egressos) apenas 3 discentes desistiram do curso (evadidos). Não se pode afirmar que a participação em IC contribui para diminuir a evasão,

para isso é necessário um estudo mais robusto e aprofundado, porém nesta pesquisa, a análise mostrou que apenas 2,31% dos ex-bolsistas de IC não concluíram a graduação na IES.

Não foi intenção apresentar profundidade de estudos à discussão, mas sim, apresentar uma nova contribuição para a área de sistemas de apoio à decisão (BI) e para o acompanhamento de egressos, esperando assim despertar interesse de pesquisadores nesta direção.

As limitações encontradas estão relacionadas a aspectos de desenvolvimento de sistemas, área onde a autora não tem domínio. Neste sentido, o uso de ferramentas mais complexas no desenvolvimento do processo de ETL, por exemplo, tornaria o processo menos moroso e mais fácil de atualização por meio de scripts e APIs.

Concluindo esta seção, ressaltam-se as implicações práticas desta pesquisa para o acompanhamento de egressos da IC por uma ferramenta de *Business Intelligence*. Plataforma essa que pode ser desenvolvida por outras IES ou, com base nos requisitos identificados, possa-se viabilizar o desenvolvimento de um sistema de informações para esse acompanhamento.

6 CONSIDERAÇÕES FINAIS

Inicialmente, para as considerações finais, é importante relembrar a questão que norteou este estudo: “quais os requisitos necessários à criação de um ambiente de *Business Intelligence* para o acompanhamento de egressos dos programas de iniciação científica?”; complementando a questão que norteou a pesquisa, é necessário listar e discorrer sobre os objetivos específicos propostos: i) Identificar os requisitos informacionais necessários para subsidiar a criação do BI para o acompanhamento dos egressos da IC; ii) Elencar quais os indicadores necessários para o acompanhamento dos egressos da IC, principalmente aqueles relacionados ao impacto do PIBIC na pós-graduação *Stricto sensu*, produção científica e desenvolvimento tecnológico e regional; iii) Identificar as fontes de informações necessárias ao processo de tomada de decisão; iv) Mapear e levantar os dados dos egressos do PIBIC/CNPq do IFSULDEMINAS, e v) Desenvolver o DW do BI e por meio do software Microsoft Power BI um painel (*Dashboard*) para exploração das informações e apresentação aos usuários finais.

Os três primeiros objetivos foram atingidos nas Seções 4.2 e 4.3 deste estudo, onde foram descritos os resultados da coleta e análise dos requisitos, indicadores e fontes de dados disponíveis para acesso aos dados. A coleta dos requisitos foi feita por triangulação de dados: pesquisa bibliográfica, documental, observação participante e reuniões com Gestores. Neste ponto a investigação sistematizada em bases científicas foi imprescindível. Assim, o estudo revelou resultados que merecem ser discutidos sucintamente: os requisitos para o acompanhamento de egressos da IC podem ser resumidos nas seguintes variáveis/dimensões: Graduação (ingresso no IFSULDEMINAS - trajetória Acadêmica), Trajetória Profissional, Impacto do PIBIC na PG, Desenvolvimento Tecnológico e Regional, Produção Científica, Formação do Pesquisador e Meninas na Ciência.

Os dois últimos objetivos específicos foram atingidos nas seções 4.4 e 4.5 onde foram descritos o desenvolvimento do DW do BI (processo de modelagem, ETL e visualização dos dados) e sua validação pela Gestão do IFSULDEMINAS. Nesta etapa da pesquisa, foi possível reunir as informações de diferentes bases de dados num único repositório para análises posteriores por meio do software Power BI. Foram realizados diversos tratamentos e consultas, gerando um repositório único (DW), contendo os dados da população de egressos da IC/CNPq entre os anos de 2011 a 2021. O DW permitiu diversas opções de análises para investigar as variáveis descritas na identificação dos requisitos para acompanhamento dos egressos. A ferramenta de BI foi validada pela Gestão do IFSULDEMINAS por meio da apresentação aos Gestores e Coordenadores de Pesquisa e PG da Instituição.

Esta pesquisa utilizou como procedimento técnico a pesquisa-ação, cumpriu rigorosamente todas as etapas propostas por Mello *et al.* (2012) e contou com a interação entre os pesquisadores e a Gestão do PIBIC na IES estudada para resolver um problema identificado e gerar conhecimento.

É importante salientar que o resultado da análise, por meio da ferramenta de BI, demonstrou o impacto promovido pela IC na trajetória acadêmica e profissional dos egressos da IC. Foi possível identificar o desempenho acadêmico durante a graduação, a maioria dos egressos de IC apresentam índice de rendimento acadêmico considerado muito bom. Foi possível identificar também que o curso de Engenharia Agrônoma foi o curso com mais egressos da IC, e com melhores indicadores de progressão à PG. Isto se justifica, pois o IFSULDEMINAS foi formado pela união de três Escolas Agrotécnicas do sul de Minas Gerais. Os egressos da IC apresentaram um índice considerado ótimo de encaminhamento aos cursos de PG, demonstrando o impacto do PIBIC neste encaminhamento. A trajetória profissional demonstrou que a IC está contribuindo tanto para a formação de pesquisadores quanto para profissionais (emprego formal) ou empreendedores. Ainda no aspecto científico, foi possível visualizar o desenvolvimento de produtos oriundos da prática da pesquisa como apresentações em eventos, patentes, artigos em anais e periódicos, demonstrando que o crescimento científico é a consequência natural da vivência dos mesmos em ambientes de pesquisa.

Os resultados apresentados foram consistentes e atenderam ao objetivo geral, que era a criação de um ambiente de *Business Intelligence* para o acompanhamento dos egressos da IC do IFSULDEMINAS com informações gerenciais que auxiliem os gestores na tomada de decisão no âmbito da iniciação científica, pesquisa e pós-graduação. A base final, dada a sua grande variedade de dimensões e dados, possibilita novas análises com as dimensões presentes. É possível estabelecer cruzamentos não oportunizados pela direção do estudo com o problema de pesquisa. Os cruzamentos podem ser realizados por diversas combinações, e poderão proporcionar novos achados não só para a IC, mas também para outros escopos.

Os objetivos foram atingidos; porém, este estudo não se encontra livre de limitações; entre elas pode-se citar o acesso às fontes de dados públicos e governamentais, em virtude da complexidade tanto no acesso como na manipulação dos dados, entretanto, mesmo com as limitações foi possível desenvolver uma ferramenta que atenda a necessidade do acompanhamento dos egressos da IC com a possibilidade de atualização anual dos dados. A população analisada foi apenas dos ex-bolsistas de IC do CNPq. É relevante, no caso do IFSULDEMINAS, que sejam analisados os outros ex-bolsistas de outras agências de fomento,

de fomento interno e voluntários para que se tenha um panorama do impacto do PIBIC e do acompanhamento destes egressos.

A implantação efetiva da ferramenta desenvolvida no IFSULDEMINAS é fundamental para seu aprimoramento, assim como a automatização e a atualização dos dados. O estudo mostra que é possível desenvolver DW/BI em IES para auxiliar na tomada de decisão, neste caso, no acompanhamento de egressos da IC; o estudo também demonstrou que a IC é uma política pública relevante na formação de novos recursos humanos para a pesquisa científica no país. Torna-se imprescindível, porém, que haja o incentivo por parte do IFSULDEMINAS na implementação de mecanismos periódicos e eficazes no acompanhamento dos resultados obtidos. Para que a IC continue sendo uma referência na formação de recursos humanos para a pesquisa, é necessário que o incentivo e os recursos financeiros para a IC sejam constantes.

Ressaltam-se as implicações práticas desta pesquisa para o acompanhamento de egressos da IC por uma ferramenta de *Business Intelligence*. A plataforma de BI pode ser desenvolvida por outras IES, e permite que os gestores obtenham informações relevantes sobre o acompanhamento dos egressos da IC, permitindo um maior conhecimento sobre a formação de recursos humanos para a pesquisa científica. Embora esta pesquisa descreva um modelo de Gestão de Egressos da IC, é necessário considerar que é impossível que um modelo seja ideal para todas as IES, devido a particularidades locais de algumas regiões do país e outras características da Gestão, sendo necessárias alterações e adequações no modelo desenvolvido de acordo com essas particularidades.

Este estudo contribuiu para preencher a lacuna da literatura sobre ferramentas de BI para a Gestão de Egressos, conforme destacado na revisão bibliométrica apresentada; assim, esse estudo aumentou o corpo de conhecimento em uma área pouco explorada, porém com grande impacto para a sociedade, uma vez que o acompanhamento de egressos é uma temática muito atual. No nível científico, a contribuição vai além de um novo estudo sobre BI; o estudo pode ser o pontapé inicial para o surgimento de mais estudos no desenvolvimento de ferramentas para a gestão de egressos em geral e até para o desenvolvimento de um sistema de informação baseado nos requisitos informacionais coletados.

Como contribuição acadêmica, este estudo adotou a pesquisa-ação como procedimento técnico; esse procedimento pode ser replicado por outras IES e os resultados podem ser discutidos e comparados. Assim, o estudo contribuiu para o progresso da teoria que evolui, mesmo que devagar, do particular para o geral.

Como sugestão de estudos futuros e com base nos requisitos aqui identificados, sugere-se o desenvolvimento de um sistema de informações para o acompanhamento de egressos (geral

ou para um público específico); também sugerem-se estudos que desenvolvam uma relação entre os resultados da análise do BI com metas nacionais e internacionais, como, por exemplo, indicadores de formação de recursos humanos, meta para absorção de ex-bolsistas da IC em cursos de PG, meta para análise do tempo de integralização dos cursos de PG, meta para inserção de recursos humanos (pesquisadores) nas IES e em empresas, entre outras.

É importante desenvolver estudos que contemplem a inserção de egressos das IES, absorvidos pelos PPGs *Stricto Sensu*, comparando os estudantes que tiveram a experiência de IC com aqueles que não tiveram e avaliar esse impacto. E não menos importante e muito necessário, sugerem-se estudos relacionados à análise da inserção de meninas ex-bolsistas de IC, ao nível local, estadual e nacional, na formação de recursos humanos para a pesquisa, por meio, é claro, da análise de dados de diversas bases.

Como agenda de estudos futuros e a partir dos resultados encontrados nesta dissertação, identifica-se a necessidade de investigações sobre estudos dos egressos/as das políticas de ação afirmativa na pós-graduação; estudos dos egressos/as das cotas para determinados grupos sociais (diversidade) na pós-graduação, de trajetórias de estudantes e docentes negros/as, quilombolas e indígenas na pós-graduação; estudos de inclusão de determinados grupos sociais na pós-graduação e estudos sobre a influência da iniciação científica no desenvolvimento de habilidades que contribuíssem no direcionamento profissional.

REFERÊNCIAS

- ANDRADE, Maria Margarida de. **Introdução à metodologia do trabalho científico**. 10ª ed. São Paulo: Atlas, 2010.
- ANGELO, Edna da Silva et al. **Estudo dos egressos do programa de iniciação científica PIBIC/CNPq na área da ciência da informação**. 2020.
- ARAÚJO, Adriana Muniz; ANDRIOLA, Wagner Bandeira. Avaliação da eficácia do Programa de Iniciação Científica (PIBIC): Estudo de caso no Instituto Federal do Ceará (UFC). **Revista Eletrônica Acta Sapientia**, v. 7, n. 1, p. 53-82, 2020.
- ARAÚJO, Luciano Vieira de. **Administração da informação e banco de dados**. In: Edmir P. V. Prado, Cesar Alexandre de Souza. (Org.). Fundamentos de Sistemas de Informação. 1ed. São Paulo: Campus Elsevier, 2014, v. 1, p. 171-193.
- BARBIERI, Carlos. **BI2--Business intelligence: Modelagem & Qualidade**. Editora Elsevier, 2011.
- BLANCO ROJAS, Tatiana; ARCHILA CORDOBA, Diana Milena; ANTONIO BALLESTEROS-RICAURTE, Javier. Management of the Data Obtained from Social Networks by Applying Business Intelligence Engineering Process (BIEP). **Revista Virtual Universidad Católica Del Norte**, [s. l.], v. 49, p. 72–91, 2016.
- BRAMBILLA, Sônia Domingues Santos. **Produção científica da Universidade Federal do Rio Grande do Sul representada na Web of Science (2000-2009)**. 2011. 218 f. Tese (Doutorado em Comunicação e Informação). Universidade Federal do Rio Grande do Sul.
- BRASIL. **Plano Nacional de Educação**. Lei nº 13.005, de 25 de junho de 2014 - Aprova o Plano Nacional de Educação - PNE e dá outras providências. Brasília, 2014. Disponível em: http://www.planalto.gov.br/ccivil_03/ato2011-2014/2014/lei/113005.htm. Acesso em 14 dez. 2021.
- CABRAL, Thiago Luiz de Oliveira. **Gestão de egressos da pós-graduação *Stricto sensu*: concepção de um modelo para programas de administração**. 2021.
- CABRERO, Rodrigo de Castro. **Formação de pesquisadores da UFSCar e na área de educação especial: impactos do programa de iniciação científica do CNPq**. São Carlos: UFSCar, 2007, 276 p. Tese (Doutorado em Educação Especial) - Programa de Pós-graduação em Educação Especial, Centro de Educação e Ciências Humanas, Universidade Federal de São Carlos.
- CANUTO, Jussara Gomes. **Análise dos processos de gerenciamento dos programas de Iniciação Científica da Universidade Federal de Itajubá: Um estudo de caso**. (2018). Dissertação (Mestrado Profissional em Administração) - Universidade Federal de Itajubá.
- CAPES. **Dados e Estatísticas: Geocapes**. Coordenação de Aperfeiçoamento de Pessoal de Nível Superior. Brasília, DF, 2021. Disponível em: <https://geocapes.capes.gov.br/geocapes/>. Acesso em: 10 jan. 2023.
- CAPES. **Plano Nacional de Pós-Graduação (PNPG) 2011-2020**. Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Brasília, DF, 2010. Disponível em: <https://www.gov.br/capes/pt-br/centrais-de-conteudo/livros-pnpg-volume-i-mont-pdf>. Acesso em: 12 de março de 2022.

CAPES. **Comissão Nacional de Acompanhamento do PNPg 2011-2020**. Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Brasília, DF, 2018. Disponível em: <https://www.gov.br/capes/pt-br/centrais-de-conteudo/2018-pnpg-cs-avaliacao-final-10-10-18-cs-final-17-55-pdf>. Acesso em: 12 de março de 2022.

CARDOSO, Elsa; SU, Xiaomeng. Projetando um Modelo de Maturidade de Business Intelligence e Analytics para o Ensino Superior: Uma Abordagem da Ciência do Design. **Ciências Aplicadas**, v. 12, n. 9, pág. 4625, 2022.

CARVALHO, Adalberto Grassi. **O PIBIC e a difusão da carreira científica na Universidade de Brasília**. 2002. 159 f. Dissertação (Mestrado em Sociologia) – Universidade de Brasília.

CENTRO DE GESTÃO E ESTUDOS ESTRATÉGICOS - CGEE. **A Formação de novos quadros para CT & I: avaliação do programa institucional de bolsas de iniciação científica (PIBIC)**. Brasília, DF: 2017. 44p. Disponível em: <https://www.cgee.org.br/documents/10195/734063/PIBIC-pdf>. Acesso em: 19 de outubro de 2021.

CENTRO DE GESTÃO E ESTUDOS ESTRATÉGICOS - CGEE. **Panorama da ciência brasileira: 2015-2020**. Boletim Anual OCTI, Brasília, v.1, jun. 2021. Disponível em: https://www.cgee.org.br/documents/10195/734063/CGEE_Pan_Cie_Bra_2015-20.pdf/7e762635-eaee-4daf-bfc9-814c785300c1?version=1.9. Acesso em 19 out. 2021.

CNPq. **Apresentação institucional**. 2023a. Disponível em: <https://www.gov.br/cnpq/pt-br/acesso-a-informacao/institucional/historico>. Acesso em: 12 de março de 2023.

CNPq. **Apresentação institucional**. 2023b. Disponível em: <https://www.gov.br/cnpq/pt-br/acesso-a-informacao/institucional/institucional>. Acesso em: 12 de março de 2023.

CNPq. RN-017/2006: **Estabelece as normas gerais e específicas para bolsas por quota no País. Brasília, 2006**. Disponível em: http://memoria2.cnpq.br/web/guest/view/-/journal_content/56_INSTANCE_0oED/10157/5483901. Acesso em: 12 de março de 2023.

COSTA, Airton; PINTO, Adilson Luiz. **De bolsista a cientista: a experiência da UFSC com o Programa de Iniciação Científica no processo de formação de pesquisadores (1990 a 2012)**. 2016. Dissertação de Mestrado.

DANTAS, Eloisa Helena Fonseca. **Formação de recursos humanos para pesquisa: avaliação do impacto do programa institucional de bolsas de iniciação científica na pós-graduação da UFRN (2019)**. Dissertação de Mestrado – Universidade Federal do Rio Grande do Norte.

DA SILVA SOUZA, Edivaldo; ABRANTES, Luiz Antônio; LISBOA-FILHO, Jugurta. O uso de Data Mart para apoio à tomada de decisão na Gestão de Pessoas em uma Instituição Federal de Ensino. In: **Anais do IX Workshop de Computação Aplicada em Governo Eletrônico**. SBC, 2021. p. 203-214.

DE ALBUQUERQUE, Elizabeth Maria Maia et al. Contribuições da Iniciação Científica para a Formação Acadêmica e Profissional de Estudantes do PIBIC/ICMBio: Percepções dos Egressos. **Biodiversidade Brasileira-BioBrasil**, v. 12, n. 1, p. 159-183, 2022.

DRAKE, Brent M.; WALZ, Aaron. Evoluindo a inteligência de negócios e a análise de dados no ensino superior. **New Directions for Institutional Research**, v. 2018, n. 178, pág. 39-52, 2018.

EGAS SÁNCHEZ, Pablo Giovanni. (2015). **Análise de viabilidade para implementação de uma plataforma de gestão de informação não estruturada**. Estudo de caso: empresa de serviços petrolíferos sediada no Equador.

ELMASRI, Ramez; NAVATHE, Shamkant B. **FUNDAMENTALS OF Database Systems**. 2021.

FAVARETTO, Fábio. Estudo de caso de implantação de *Data Warehouse* para gerenciamento de refugos. **Anais do II Seminário de Gestão de Negócios – FAE**, Curitiba, 2005.

FAVARETTO, Fábio. Ambiente de Data Warehouse para análise de algumas medidas utilizadas na administração da produção. **Revista da FAE**, v. 8, n. 2, 2016.

FIORI, Felipe Correa; DE SOUZA, Márcio Rogério; BEZERRA, Cicero Aparecido. Publicações científicas e acesso a mestrados de alunos de Iniciação Científica: um estudo nos cursos do Setor de Saúde da Universidade Federal. **Em Questão**, p. 204-227, 2021.

GALVÃO, Antonio Carlos Filgueira et al. **O quadro recente de emprego dos mestres e doutores titulados no Brasil**. *Parcerias Estratégicas*, v. 21, n. 43, p. 147-172, 2017.

GAMBOA-CRUZADO, Javier et al. Business Intelligence as Decision Support in Organizations: A Systematic Review of the Itinerary. **Journal of Positive Psychology and Wellbeing**, p. 464-480, 2023.

GARTNER. (2013). **Business Intelligence (BI)**. **Gartner IT Glossary**. Retrieved June 27, 2013, from <http://www.gartner.com/it-glossary/business-intelligence-bi/>

GARTNER. (2019). **Analytics and Business Intelligence (ABI)**. Retrieved October 12, 2019, from <https://www.gartner.com/en/information-technology/glossary/businessintelligence-bi>.

GIL, Antonio Carlos. **Como classificar as pesquisas**. Como elaborar projetos de pesquisa, v. 4, n. 1, pág. 44-45, 2002.

GIL, Antonio Carlos. **Como elaborar projetos de pesquisa**. 5.ed. São Paulo: Atlas, 2010.

HAMDANE, Abdemounaime; *et al.* Big data based architecture to bringing together graduates and recruiters: case of Moroccan university(Article)(Open Access). **Indonesian Journal of Electrical Engineering and Computer Science**. Volume 26, Issue 3, June 2022, Pages 1701-1709. <https://ijeecs.iaescore.com/index.php/IJEECS/article/view/28155>

INMON, William H. **Como construir o Data Warehouse**. Rio de Janeiro: Campus, 1997. 388 p.

INMON, William H. **Building the Data Warehouse**. John Wiley & Filhos, 2005.

INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DO SUL DE MINAS GERAIS. **O IFSULDEMINAS**, 2023. Disponível em: <https://portal.ifsuldeminas.edu.br/index.php/o-instituto>. Acesso em: 17 de março de 2023.

ISMAIL, Nurhanani Izzati; ABDALLAH, Nur Atiqah Sia; OMAR, Abdullah. “Exploring Alumni Data using Data Visualization Techniques”. **International Journal of Advanced Computer Science and Applications (IJACSA)**, 13(9), 2022. <http://dx.doi.org/10.14569/IJACSA.2022.0130922>.

JUKIC, Nenad. Modelagem de estratégias e alternativas para projetos de *data warehouse*. **Comunicações da ACM**, v. 49, n. 4, pág. 83-88, 2006.

KIMBALL, Ralph *et al.* **The Data Warehouse lifecycle toolkit: expert methods for designing, developing and deploying Data Warehouses**, 2nd ed., John Wiley & Sons, 2008.

KIMBALL, Ralph; ROSS, Margy. **O Data Warehouse Toolkit: O Guia Definitivo de Modelagem Dimensional**. 2013.

KUSHWAHA, Amit Kumar; KAR, Arpan Kumar; DWIVEDI, Yogesh K. Applications of big data in emerging management disciplines: A literature review using text mining. **International Journal of Information Management Data Insights**, v. 1, n. 2, p. 100017, 2021.

LAUDON, Kenneth C.; LAUDON, Jane P. **Sistemas de Informações Gerenciais**. 11. ed. São Paulo: Pearson Education do Brasil, 2014.

LAUDON, Kenneth C.; LAUDON, Jane P. **Sistemas de informação gerenciais: administrando a empresa digital**. Bookman editora, 2022.

LESSAK, Alisson Luiz et al. **Desenvolvimento de uma ferramenta de apoio à gestão para o polo de inovação do Instituto Federal de Santa Catarina utilizando business intelligence**. 2018. Dissertação de Mestrado. Universidade Federal de Santa Catarina.

LUHN, Hans Peter. Um sistema de inteligência de negócios. **IBM Journal of research and development**, v. 2, n. 4, pág. 314-319, 1958.

MACCARI, Emerson Antonio et al. Proposta de um modelo de gestão de programas de pós-graduação na área de Administração a partir dos sistemas de avaliação do Brasil (CAPES) e dos Estados Unidos (AACSB). **Revista de Administração**, v. 49, n. 2, pág. 369-383, 2014.

MACHADO, Felipe Nery Rodrigues. **Tecnologia e projeto de Data Warehouse**. Saraiva Educação SA, 2010.

MARTINS, Gilberto de Andrade; THEÓPHILO, Carlos Renato. **Metodologia da Investigação Científica para Ciências Sociais Aplicadas**. 3. ed. São Paulo: Atlas, 2016. 2009

MASSI, Luciana; QUEIROZ, Saete Linhares. (Org.). **Iniciação científica: aspectos históricos, organizacionais e formativos da atividade no ensino superior brasileiro [online]**. São Paulo: Editora UNESP, 2015. 160 p. ISBN 978-85-68334-57-7. Disponível em: <http://books.scielo.org/id/s3ny4/pdf/massi-9788568334577.pdf>. Acesso em: 12 dez. 2021.

McCORMACK, Mark; BROOKS, D. Christopher; SHULMAN, Ben. *Observação de tendências de 2020 do ensino superior e as 10 principais tecnologias estratégicas as 10 principais tecnologias estratégicas para 2020*; **Educase: Boulder, CO, EUA**, 2020; Disponível online: <https://www.educase.edu/ecar/research-publications/higher-education-trend-watch-and-top-10-strategic-technologies/2020/the-top-10-strategic-technologies-for-2020#TheTop10StrategicTechnologiesfor2020> (acessado em 29 de março de 2023).

MELLO, Carlos Henrique Pereira *et al.* Pesquisa-ação na engenharia de produção: proposta de estruturação para sua condução. **Production [online]**. 2012, v. 22, n. 1, pp 1-13. Disponível em: <https://doi.org/10.1590/S0103-65132011005000056>. Acesso em: 26 ago. 2022.

MENDES, Felipe Rodrigo de Moraes et al. **Proposta de um data warehouse para apoio à tomada de decisão sobre evasão institucional em uma instituição federal de ensino superior**. 2020.

MOURA, Daiana De Nez; CECCHETTI, Elcio; DOS SANTOS BERNARDI, Luci Teresinha Marchiori. Contribuições do PIBIC/CNPq para a constituição do habitus de pesquisador. **Revista Eletrônica de Educação**, v. 14, p. e3257096-e3257096, 2020.

MICROSOFT. **Visão geral de Tabela Dinâmica e Gráfico Dinâmico**. Suporte do Office. Disponível em: <https://support.microsoft.com/pt-br/office/vis%C3%A3o-geral-de-tabela-din%C3%A2mica-e-gr%C3%A1fico-din%C3%A2mico-527c8fa3-02c0-445a-a2db-7794676bce96>. Acesso em: 17 ago. 2022.

MICROSOFT. **O que é o Power BI desktop?** Disponível em: <https://learn.microsoft.com/pt-pt/power-bi/fundamentals/desktop-what-is-desktop>

MINAYO, Maria Cecília de Souza; DESLANDES, Suely Ferreira; GOMES, Romeu. **Pesquisa social: teoria, método e criatividade**. 30a ed. Petrópolis: Vozes, 2015.

MUELLER, Suzana Pinheiro Machado. O crescimento da ciência, o comportamento científico e a comunicação científica: algumas reflexões. **Revista da Escola de Biblioteconomia da UFMG**, v. 24, n. 1, 1995.

NARDINI, Elisa Fonseca et al. Política de estímulo à iniciação científica: impacto no coeficiente de rendimento de graduandos em Odontologia. **Revista da ABENO**, [S. l.], v. 19, n. 1, p. 33–39, 2019. DOI: 10.30979/rev.abeno.v19i1.619. Disponível em: <https://revabeno.emnuvens.com.br/revabeno/article/view/619>. Acesso em: 23 mar. 2023.

NIÑO, Harold Arturo Combata; NIÑO, Johana Patrícia Cómbita; ORTEGA Roberto Morales. Estrutura de governança de inteligência de negócios em uma universidade: estudo de caso da Universidad de la costa. **International Journal of Information Management**, v. 50, p. 405-412, 2020.

NOVAIS, Geisa Rodrigues. **Programa de iniciação científica da Universidade de Brasília: uma análise da concentração, do diferencial para ingresso no mercado de trabalho e na pós-graduação, no período de 2004 a 2018**. Dissertação (Mestrado em Gestão Pública), Universidade de Brasília: Brasília, DF, Brasil, 2020.

O'BRIEN, James Aloysius. **Sistemas de informação e as decisões gerenciais na era da internet**. 3. ed. São Paulo: Saraiva, 2011.

OLIVEIRA, Adriano de; BIANCHETTI, Lucídio. Estudantes do ensino médio e do ensino superior: explicitando o modus operandi dos bolsistas do Programa Institucional de Bolsas de Iniciação Científica para o Ensino Médio. **Revista Brasileira de Estudos Pedagógicos**, Brasília, v. 100, n. 255, pág. 464-480, maio de 2019. Acesso em 15 de agosto de 2022. <https://doi.org/10.24109/2176-6681.rbep.100i255.3809>.

OLIVEIRA, Rosângela Moreira de et al. A eficácia dos programas de fomento à iniciação científica: Uma análise sobre o PIBIC da Fundação de Amparo à Pesquisa do Estado da Bahia. **Pesquisa, Sociedade e Desenvolvimento**, v. 9, n. 7, pág. e596974432-e596974432, 2020.

ORAE, Narges; SANATJOO, Azam; AHANCHIAN, Mohamad Reza. An exploratory study on competitive intelligence: Managers' information needs in higher education sector. **Malaysian Journal of Library & Information Science**, [S. l.], v. 26, n. 2, p. 125–142, 2021. DOI: 10.22452/mjlis.vol26no2.7. Disponível em: <https://mjlis.um.edu.my/index.php/MJLIS/article/view/25149>. Acesso em: 11 jan. 2023.

PELINSKI, Joana D. et al. **Efeito da iniciação científica no rendimento acadêmico e na inserção em Programas de Pós-Graduação “Stricto Sensu”**: uma análise com Propensity Score Matching. 2021. Dissertação de mestrado.

PINHO, Maria José de. Ciência e ensino: contribuições da iniciação científica na educação superior. Avaliação: **Revista da Avaliação da Educação Superior** (Campinas) [online]. 2017, v. 22, n. 3 [acessado em 15 janeiro 2023], pp. 658-675. Disponível em: <https://doi.org/10.1590/S1414-40772017000300005>.

PIRES, Regina Celi Machado. O vir a ser professor/pesquisador da educação superior: o caso dos egressos PIBIC/CNPq/UNEB. **Revista de Educação Pública**, v. 21, n. 45, p. 67-84, 2012. Disponível em: <https://periodicoscientificos.ufmt.br/ojs/index.php/educacaopublica/article/view/333>. Acesso em: 29 dez. 2022.

PITON, Rafael. **Data Warehouse Passo a Passo: O Guia Prático de Como Construir um Data Warehouse do Zero**. Porto Alegre, 2018.

PRIMAK, Fábio Vinícius. **Decisões com B.I. (Business intelligence)**. Rio de Janeiro: Ciência Moderna, 2008.

SANCHEZ, Ilara. **Trajetórias acadêmica e profissional dos egressos do Programa de Pós-graduação em Educação da Unicamp**. 2019. Tese de Doutorado. Faculdade de Educação, Universidade Estadual de Campinas.

SANTOS, Lisandra Helena Barros et al. **A contribuição do CNPq para a formação de recursos humanos para pesquisa: uma avaliação de impacto do Programa Institucional de Bolsas de Iniciação Científica-PIBIC**. 2021. Dissertação de mestrado.

SILVA, Natália Rodrigues Araújo da et al. Integridade científica entre alunos de enfermagem que participam do Programa de Iniciação Científica: Estudo exploratório. **Revista da Escola de Enfermagem da USP**, v. 54, 2020.

SILVA, Vinícius Tolentino Oliveira e. **A produtividade do professor no sistema de avaliação dos Programas de Pós-Graduação em Administração da CAPES: uma análise por meio da Business Intelligence**. Dissertação de mestrado. UFMG. Belo Horizonte, 2018.

SINGH, Harry S. **Data Warehouse: Conceitos, Tecnologias, Implementação e Gerenciamento**. 1ª ed. São Paulo: Makron Books, 2001.

SOUZA, Sarlene Gomes de. **A Formação Diferenciada em Egressos de Iniciação Científica: implicações na pós-graduação**. 2018. Tese de Doutorado. Programa de Pós-Graduação em Educação, Universidade Estadual do Ceará, Fortaleza.

STAIR, Ralph; REYNOLDS, George. **Princípios de sistemas de informação**. Cengage Learning, 2020.

STEWART, Christopher Lee; DEWAN, M. Ali Akber. A Systemic Mapping Study of Business Intelligence Maturity Models for Higher Education Institutions. **Computers**, v. 11, n. 11, p. 153, 2022.

TURBAN, Efraim et al. **Business Intelligence - Um Enfoque Gerencial para a Inteligência do Negócio**. Bookman, 2009. ISBN 9788577803347.

THIOLLENT, Michel. **Metodologia da pesquisa-ação**. In: Metodologia da pesquisa-ação. 1986 - 2ª Edição.

THIOLLENT, Michel. **Metodologia da pesquisa-ação**. Cortez editora, 2022.

TOZATO, Heloisa de Camargo *et al.* Avaliação de impacto de políticas públicas: o estudo de caso do PIBIC/ICMBio no Brasil. **Avaliação (Campinas)**, Sorocaba, v. 25, n. 3, p. 676-700, dez. 2020. Disponível em http://old.scielo.br/scielo.php?script=sci_arttext&pid=S1414-40772020000300676&lng=pt&nrm=iso. Acessos em 15 ago. 2022. Epub 18-Nov-2020. <https://doi.org/10.1590/s1414-40772020000300009>.

TRIEU, Van-Hau. Obtendo valor de sistemas de Business Intelligence: Uma agenda de revisão e pesquisa. **Sistemas de Apoio à Decisão**, v. 93, p. 111-124, 2017.

TROMBELLI, Renata Oliveira. **PIBIC/CNPq no divã: um olhar para a efetividade do processo de iniciação científica na formação de pesquisadores em contabilidade**. 2013. Dissertação (Mestrado em Contabilidade) — Universidade Federal do Paraná.

United Nations Educational, Scientific and Cultural Organization (Unesco). (2019). **Women in science (UIS Fact Sheet, 55)**. <http://uis.unesco.org/sites/default/files/documents/fs55-women-in-science-2019-en.pdf>

VENDRÚSCOLO, Juliana de Bona Garcia et al. **Um sistema de *business intelligence* para a extensão universitária**. Dissertação de Mestrado. Universidade Federal de Santa Catarina. 2020.

VIEIRA, Benedito Gomes et al. **Criação de um ambiente de inteligência de negócios para auxílio à tomada de decisão baseado no desempenho acadêmico dos discentes de graduação da Universidade Federal de Itajubá**. 2021. Dissertação de Mestrado. Universidade Federal de Itajubá.

ZEA, Oswaldo Moscoso; SAMPEDRO, Andrés Paz; MORA, Sérgio Luján. *Data Warehouse design for educational data mining*. **15th International Conference on Information Technology Based Higher Education and Training (ITHET)**. 2016, pp. 1-6, doi:10.1109/ITHET.2016.7760754. <https://ieeexplore.ieee.org/document/7760754>

APÊNDICE A – ROTEIRO DA REUNIÃO SEMIESTRUTURADA

Roteiro para reunião com roteiro semiestruturado com os gestores para entendimento do problema a ser solucionado e as informações sobre requisitos mínimos para o Sistema de Apoio à Decisão na Gestão do acompanhamento de Egressos da Iniciação Científica. Instrumento validado por Vidigal (2019), cuja pesquisa tinha proposta de levantamento de necessidades de informações semelhantes.

Objetivo: Levantamento das necessidades de dados, indicadores e informações para a gestão do acompanhamento de egressos da IC do IFSULDEMINAS, segundo a percepção dos gestores da área.

Questões que nortearam a conversa, pois como a reunião foi semiestruturada cada questão deu margem e abertura para novos questionamentos.

- 1) Quais informações são importantes para a Gestão do acompanhamento de Egressos no âmbito da IC (enquanto formação de pesquisador a partir da vivência na IC)?
- 2) Quais indicadores refletem estas informações?
- 3) Todas as informações relevantes estão disponíveis? De que forma é o acesso às informações que serão utilizadas?
- 4) Quais as fontes de dados da instituição estão disponíveis para acesso? Os dados disponíveis atendem a todas as necessidades?
- 5) Os dados são todos da instituição ou serão necessários dados externos?
- 6) Para os dados externos é possível identificar a fonte desses dados? Quais?
- 7) Qual a dificuldade para extrair os dados? Essa ação é permitida? Temos acesso?
- 8) Quais seriam as decisões tomadas com base em um Sistema de Acompanhamento de Egressos?
- 9) Qual a importância da Gestão de Egressos da IC?

APÊNDICE B – ESTUDOS SELECIONADOS NAS REVISÕES INTEGRATIVAS

Artigo 1 – Artigos selecionados para a revisão

	Artigos selecionados	Citações	Ano
1	Mahaney, R. C. & Fisher, J. H. (2009). Leveraging Alumni and Business Community Relations to Assess the Information Systems Curriculum. <i>Journal of Education for Business</i> , 84(3), 142-150.	13	2009
2	Mijic, D. & Jankovic, D. (2012). <i>Towards Improvement of the Study Programme Quality: Alumni Tracking Information System</i> . ICT Innovations, 291–300.	4	2012
3	Pratt J. A., Keys A., & Wirkus T. (2014). Preparing information systems graduates for a complex society: Aligning is curricula with liberal education learning outcomes. <i>Journal of Information Systems Education</i> . 25(1), 35-44.	6	2014
4	Hordosy, R. (2014). Who knows what school leavers and graduates are doing? Comparing information systems within Europe. <i>Comparative Education</i> . 50(4), 448-73.	4	2014
5	Sipos, N. (2017). Graduate Career Tracking System Across the World - as Information Systems in higher Education Decision-making Process. <i>Strategic Management</i> , 22(4), 24-31.	1	2017
6	Hordosy, R. (2016). How do different stakeholders utilise the same data? The case of school leavers' and graduates' information systems in three European countries. <i>International Journal of Research & Method in Education</i> , 40(4), 403-420.	1	2017
7	Morillo, J. P., Cante, C. E. P., Daza, N. J. P. (2018). Evaluation of the professional profile in information systems, library science and archival science at the university of la salle, Colombia, from the perspective of graduates. <i>Meta: Avaliacao</i> 10(29), 387-416.	7	2018
8	Azis, A., Sugiarti, Y., Kumaladewi, N., & Huda, M. Q. Designing and Building an Information System of Career Development and Alumni Based on Android (Case Study: Information Systems Department, Syarif Hidayatullah State Islamic University Jakarta). 2018. <i>6th International Conference on Cyber and IT Service Management (CITSM)</i> . Inna Parapat Hotel – Medan, August 7-9.	1	2018
9	Mahaney, R. C., Fisher, J. H. (2019). A Survey of CIS Alumni to Assess Curriculum: Findings, Implications, and Future Directions. <i>Journal of Information Technology Education: Research (JITE:Research)</i> , 18, 571-587.	1	2019
10	Alban, C. S. G., Fernandez, A. R., & Galarza, F. P. C. (2019). Information System for monitoring graduates in Ecuadorian higher education institutions. <i>Dilemas Contemporaneos-Educacion Política y Valores</i> , 7(1), 1-22.	0	2019

Artigo 2 – Dissertações selecionadas para a revisão

Estudos Selecionados	Autor/Ano Publicação
Acompanhamento de egressos como instrumento de gestão: Proposta de Inovação nas Políticas Acadêmicas da Universidade Federal do Triângulo Mineiro.	Almeida (2021)
Diretrizes de Gestão de egressos para Faculdades de Tecnologia.	Arnoni (2019)
A Gestão do relacionamento com egressos: Uma proposta de diretrizes para o Programa De Pós-Graduação em Administração da UFSC.	Cabral (2017)
Trajetórias, Repercussões e Inserção Profissional dos egressos do Programa de Pós-Graduação em Desenvolvimento Regional – UEPB.	Cavalcanti (2021)
Mapeamento do perfil e destino profissional dos egressos de doutorado da Pós-Graduação em Estudos da Tradução da UFSC: 2010 – 2017.	Christmann (2018)
Política Institucional de integração e acompanhamento de egressos nas Universidades Federais: Uma proposta para a UFGD.	Freire (2018)
Radiografia dos egressos PPGE/UFOPA (2015-2019): Subsídios para elaboração de proposta de modelo de sistema de acompanhamento Institucional.	Gonçalves (2021)
O acompanhamento dos alunos egressos do Mestrado Profissional em Gestão e Avaliação da Educação Pública.	Junior (2017)
Sistemas de acompanhamento de egressos do Ensino Superior: Uma análise sob a ótica de Atores Institucionais.	Leopoldo (2019)
Proposta de diretrizes para a Gestão de egressos de Programas de Graduação.	Pereira (2018)
O acompanhamento de egressos da Universidade Federal do Estado do Rio de Janeiro: Avaliação do Programa Joia Rara.	Santos (2021)
A Gestão de egressos como fonte de informações estratégicas para a UFFS.	Simon (2017)
Instituição de um sistema eficaz para o acompanhamento de egressos de Programas de Pós-Graduação: O egresso como referência para a autoavaliação do PPGACL.	Soares (2019)
Processos de acompanhamento de egressos dos cursos superiores no Instituto Federal do Piauí.	Souza (2021)
O feedback do egresso como fonte de desenvolvimento organizacional: Subsídios para a implantação de um instrumento aplicado.	Volpato (2019)

APÊNDICE C – PROTOCOLO DE COLETA DE DADOS

A utilização de um protocolo de coleta de dados busca orientar e regular a estratégia de pesquisa, para garantir que os resultados possam ser replicados.

1) RAIS

A Relação Anual de Informações Sociais (RAIS) tem por objetivo o suprimento às necessidades de controle da atividade trabalhista no país, para identificação dos trabalhadores com direito ao recebimento do Abono Salarial. Outras funções são o provimento de dados para a elaboração de estatísticas do trabalho e a disponibilização de informações do mercado de trabalho às entidades governamentais.

- Acessar o site: <http://rais.gov.br/sitio/index.jsf>
- No lado direito clique em declaração já entregue
- Consultar trabalhador
- Entrar com os dados para acessar a informação.

2) MEI

Este serviço permite que seja emitido o Comprovante de Condição de Microempreendedor Individual (CCMEI). Este documento certifica a condição atual do MEI. Qualquer interessado em validar as informações do MEI pode acessar os dados. O CCMEI comprova: a inscrição do MEI no CNPJ e na Junta Comercial do Estado, além da dispensa de alvará e licença de funcionamento; a situação cadastral atual do MEI (ativo, baixado, suspenso, inapto, etc.).

- Acessar o site: <https://mei.receita.economia.gov.br/certificado/consulta>
- Entrar com o login do GOV.BR
- Abrirá o portal do empreendedor
- Consultar empreendedor – entrar com os dados para acessar as informações (um de cada vez) - <https://www.gov.br/empresas-e-negocios/pt-br/empreendedor/servicos-para-mei/emissao-de-comprovante-ccmei>

3) LinkedIn

A plataforma pode auxiliar em informações laborais, contudo, fica restrita aos dados informados, e podem não estar atualizados.

- Acessar o site: <https://br.linkedin.com/>

- Clicar em entrar e logar com seus dados
- Na sua página, localize a instituição que deseja identificar pessoas
- No caso de IES, clique na IES, no meu caso estava em PAGES clique na IES escolhida
- Abrirá a página da IES, clique em ex-alunos na barra principal
- Abrirá uma caixa de diálogo que pode ser consultado pelo nome e filtrado por período
- Efetuar as consultas desejadas.

4) Latticles

O Latticles é um aplicativo web construído com o pacote shiny do software livre R para auxiliar na avaliação de produções científicas, técnicas e tecnológicas, artísticas e culturais declaradas na Plataforma *Lattes* do CNPq, em especial as publicações em periódicos científicos com Qualis-CAPES. Ele foi desenvolvido pelo Prof. Dr. Anderson Rodrigo da Silva do Instituto Federal Goiano - Campus Urutaí, para a autoavaliação de Programas de Pós-Graduação, com o objetivo primário de promover o crescimento e fortalecimento desses. Registro INPI: BR512019001166-0.

O Latticles contabiliza produções técnicas, tecnológicas, tudo de forma automatizada a partir de arquivos XML do currículo *Lattes*.

- Acessar o site: <https://arsilva.shinyapps.io/latticles2/>
- Do lado direito tem a opção de selecionar os arquivos dos lattes em .XML (até 50 arquivos) e o período para a extração de dados
- Após selecionar os arquivos e o período clicar em executar, o resultado é exibido na ferramenta com a possibilidade de exportar os dados em arquivos com extensão .csv.
- A extração é apresentada em tópicos: artigos em periódicos, produtos técnicos e tecnológicos, livros e seções, artística e cultural, orientações concluídas e participações em bancas, todos podem ser exportados em formatos .csv.

ANEXO A – ANUÊNCIA DA INSTITUIÇÃO PARA A REALIZAÇÃO DA PESQUISA E A COLETA DOS DADOS

Termo de Anuência Institucional

Eu, Carlos Henrique Rodrigues Reinato, ocupante do cargo de Pró-Reitor de Pesquisa, Pós-Graduação e Inovação, do Instituto Federal de Educação, Ciência e Tecnologia do Sul de Minas Gerais - IFSULDEMINAS, confirmo a anuência da Instituição para realização da pesquisa "**PROPOSTA DE UM AMBIENTE DE *BUSINESS INTELLIGENCE* PARA ACOMPANHAMENTO DOS EGRESSOS DOS PROGRAMAS DE INICIAÇÃO CIENTÍFICA DO IFSULDEMINAS**".

A pesquisa será conduzida pela mestrande Eunice Cristina da Silva, sob a orientação do professor Dr. Fábio Favaretto, no âmbito do Programa de Pós-Graduação Profissional em Administração da Universidade Federal de Itajubá (UNIFEI).

O objetivo da pesquisa é propor um ambiente de *Business Intelligence* para a gestão do acompanhamento de egressos da iniciação científica do IFSULDEMINAS com informações gerenciais que auxiliem os gestores na tomada de decisão no âmbito da iniciação científica.

Foi assegurado pelo pesquisador responsável que a pesquisa envolverá a coleta de dados (pesquisa documental) dos egressos da Iniciação Científica do IFSULDEMINAS nos Sistemas de Informação da Instituição.

Pouso Alegre/MG, 12 de setembro de 2022.