Tendências da Pesquisa Brasileira em Ciência da Informação

FERRAMENTAS DE APOIO À INTEGRAÇÃO DE DADOS ENTRE SISTEMAS DE INFORMAÇÃO NO CONTEXTO DA AVALIAÇÃO DA PÓS-GRADUAÇÃO STRICTO SENSU PELA CAPES[1]

TOOLS TO SUPPORT DATA INTEGRATION BETWEEN INFORMATION SYSTEMS IN THE CONTEXT OF STRICTO SENSU POSTGRADUATE EVALUATION BY CAPES

Harrysson Gilgamesh Medeiros Nobrega[2]
Emanuelle Torino[3]
Silvana Aparecida Borsetti Gregorio Vidotti[4]
Jaqueline Pereira Carvalho Halicki[5]
Talita Moreira de Oliveira[6]

Resumo: No processo de avaliação da pós-graduação *stricto sensu* brasileira, é necessário que as instituições informem à Capes um conjunto de dados que subsidiam o processo. Tais dados estão armazenados nas instituições em diferentes fontes, compostas por sistemas informatizados ou mesmo em planilhas, e, em geral são informados à Capes em um processo manual de preenchimento, gerando retrabalho. Considerando a necessidade de integrar os diferentes sistemas de informação das instituições que compõem o ecossistema brasileiro da Educação, Ciência, Tecnologia e Inovação, o presente estudo objetivou estudar e desenvolver uma solução tecnológica para integrar os dados armazenados nos sistemas de informação utilizados pelas instituições que compõem o ecossistema em questão, bem como das fontes de publicação e disseminação de resultados de pesquisa, possibilitando integrá-los às plataformas do governo brasileiro. Com isso, foram desenvolvidas ferramentas tecnológicas customizáveis, capazes de se adaptarem e serem utilizadas em qualquer ecossistema, independentemente das diferentes faixas de dados e maturidade tecnológica. As Ferramentas de Apoio à Integração de Sistemas desenvolvidas e apresentadas neste estudo se constituem como uma solução para auxiliar instituições, fundações

[1] Texto ampliado a partir de artigo submetido, avaliado, aprovado, apresentado e premiado no XXII ENANCIB.

[2] Especialista em Gestão de Tecnologia da Informação. Consultor e gerente de projetos na Rede Nacional de Ensino e Pesquisa (RNP). E-mail: harrysson.nobrega@consultores.rnp.br. Orcid: https://orcid.org/0000-0002-0378-4782.
[3] Doutora em Ciência da Informação (UNESP/Marília). Bibliotecária Documentalista na Universidade Tecnológica Federal do Paraná (UTFPR). E-mail: emanuelle@utfpr.edu.br. Orcid: https://orcid.org/0000-0002-3791-9884.
[4] Doutora em Educação (UNESP/Marília). Docente da Universidade Estadual Paulista Júlio Mesquita Filho. E-mail: silvana.vidotti@unesp.br. Orcid: https://orcid.org/0000-0002-4216-0374.
[5] Mestra em Ciência da Informação (UNESP). Assistente Administrativa da Universidade Estadual do Rio de Janeiro (UERJ) Cedida à Pró-Reitoria de Pós-Graduação da Universidade Estadual Paulista (PROPG/UNESP). E-mail: jaqueline.halicki@unesp.br. Orcid: https://orcid.org/0009-0003-3009-3818.
[6] Doutora em Educação em Ciências Químicas da Vida e Saúde (UFRGS). Analista em Ciência e Tecnologia na Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). E-mail: talita.oliveira@capes.gov.br. Orcid: https://orcid.org/0000-0001-5740-4207.

de apoio e de governo, na comunicação de dados entre sistemas de informação, trazendo ganhos significativos para todos os envolvidos.

Palavras-Chave: Integração de dados. Integração entre sistemas. Ferramentas de Apoio à Integração de Sistemas. Ecossistema de ECTI.

Abstract: *In the process of evaluating Brazilian stricto sensu postgraduate courses, institutions must provide Capes with a set of data that supports the process. Such data are stored in institutions in different sources, made up of computerized systems or even in spreadsheets, and are generally reported to Capes in a manual filling process, generating rework. Considering the need to integrate the different information systems of the institutions that make up the Brazilian Education, Science, Technology and Innovation ecosystem, the present study aimed to study and develop a technological solution to integrate the data stored in the information systems used by the institutions that make up the Brazilian Education, Science, Technology and Innovation ecosystem, as well as sources for publishing and disseminating research results, making it possible to integrate them into Brazilian government platforms. As a result, customizable technological tools were developed, capable of adapting and being used in any ecosystem, regardless of different data ranges and technological maturity. The Systems Integration Support Tools developed and presented in this study are a solution to assist institutions, support foundations and government in communicating data between information systems, bringing significant gains to everyone involved.*

Keywords: *Data integration. Integration between systems. Systems Integration Support Tools. ECTI ecosystem.*

1 INTRODUÇÃO

No ecossistema brasileiro de Educação, Ciência, Tecnologia e Inovação (ECTI), a coleta de dados armazenados em diferentes sistemas de informação é complexa em função das estruturas e formas de armazenamento dos dados. Esse contexto possui vários atores, como educadores, pesquisadores, coordenadores de cursos de graduação e pós-graduação, gestores de instituições de ensino, órgãos governamentais e agências de fomento[7].

Vivenciamos um cenário de mudança na forma como são destinados e utilizados os recursos públicos na ECTI brasileira, sejam eles financeiros, físicos, tecnológicos e/ou humanos. Com isso, há a necessidade de dispor de ferramentas

[7] Quando nos referimos às Agências de Fomento, incluímos as Fundações de Amparo à Pesquisa (FAP).

tecnológicas eficientes para mitigar o retrabalho dos atores supramencionados na alimentação de diferentes sistemas de informação, com finalidades específicas e que necessitam de dados armazenados em outros sistemas para fins institucionais, atualização de currículo, avaliação pelo Ministério da Educação (MEC) ou pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), entre outros. Ainda, é perceptível um cenário heterogêneo em termos de quantidade, qualidade e confiabilidade de dados disponíveis nos sistemas de informação das diversas instituições que compõem o ecossistema de ECTI.

Dietrich *et al.* ([2022?]) afirmam que tecnologias emergentes podem ser utilizadas para fornecer dados aos cidadãos de forma automática, por meio dos preceitos de Dados Abertos, uma vez que grande parte deles são gerados por órgãos públicos. Por outro lado, apontam que tais dados, em geral, não estão disponíveis de modo a favorecer o (re)uso.

Diante do exposto, este trabalho tem como problema de pesquisa: como integrar os dados necessários à avaliação da pós-graduação *stricto sensu* brasileira, disponíveis em sistemas de informação que armazenam dados nas instituições e nas fontes de publicação e disponibilização de resultados de pesquisa, sistemas esses que compõem o ecossistema brasileiro de Educação, Ciência, Tecnologia e Inovação e as plataformas do governo?

Para tanto, objetivou-se estudar e desenvolver uma solução tecnológica para integrar os dados armazenados nos sistemas de informação utilizados pelas instituições que compõem o ecossistema brasileiro de ECTI, em especial no contexto da pós-graduação *stricto sensu* brasileira, bem como dados das fontes de publicação e disseminação de resultados de pesquisa, possibilitando integrá-los às plataformas do governo brasileiro.

Neste sentido, cumpre destacar que

> A CAPES é responsável por avaliar, acompanhar, fomentar e induzir cursos de pós-graduação *stricto sensu*. Assim, para o funcionamento regular dos cursos de mestrado e doutorado no Brasil e para a expedição de diploma válido em todo território nacional, são necessários a avaliação positiva da CAPES, o reconhecimento do Conselho Nacional de Educação e a homologação pelo Ministério da Educação para cada curso pleiteado (COORDENAÇÃO DE APERFEIÇOAMENTO DE PESSOAL DE NÍVEL SUPERIOR, 2023, online).

O estudo abordou a solução intitulada Ferramentas de Apoio à Integração de Sistemas (FAIS), composta pelo FAIS Expositor, FAIS Integrador, FAIS Inventário, Pilha de Desambiguação e Rede Neural – desenvolvida e disponibilizada pela CAPES para favorecer a integração de dados entre sistemas de informação relacionados ao processo de avaliação da pós-graduação *stricto sensu* brasileira pela CAPES.

2 METODOLOGIA

Na busca de um método para orientar a descoberta, o acesso, a integração e a reutilização da vasta quantidade de dados gerados e armazenados no contexto da ECTI, foram utilizados os preceitos estabelecidos pelo movimento de dados abertos (Torino; Trevisan; Vidotti, 2019; Torino; Vidotti, 2021; Dietrich *et al.*, [2022?]; Open Knowledge Foundation, 2022), pela transparência pública (Brasil, 2011) e pelos princípios FAIR (FORCE11, [2020?]; GOFAIR, [2020?]).

O estudo de natureza empírica investigou um fenômeno em um contexto com pouca literatura disponível, o que reforça sua relevância para o cenário do ecossistema brasileiro de ECTI e para a Ciência da Informação.

Os antecedentes que marcaram a necessidade de realização desta pesquisa são identificados na própria CAPES que, ao longo dos anos, realizou diferentes iniciativas visando auxiliar a coleta de dados dos Programas de Pós-Graduação (PPGs). Para

tanto, foram utilizados diferentes recursos tecnológicos, dentre eles o INFOCAPES, o DATACAPES, o Coleta e, a partir de 2012, a Plataforma Sucupira.

Em 2019, com o objetivo de identificar pontos passíveis de melhorias na Plataforma Sucupira, para torná-la uma plataforma de gestão da informação das Instituições de Ensino Superior Brasileiras (IES), foram selecionadas universidades, com base em requisitos tecnológicos e corpo técnico disponível, para participar de testes de integração. Isso visa a cocriação de soluções desejáveis, econômicas e tecnologicamente viáveis para desenvolvimento e disponibilização para uso das instituições interessadas.

A coleta das informações incluiu uma pesquisa realizada no Seminário de Meio Termo de 2019, promovido pela CAPES, referente ao quadriênio 2014-2020. Posteriormente, no segundo semestre de 2019, foram realizadas visitas às IES selecionadas. As visitas objetivaram conhecer as formas de coleta, tratamento e armazenamento dos dados, bem como as tecnologias computacionais existentes nas instituições. Para tanto, os representantes das IES participantes foram entrevistados, utilizando como técnicas de coleta de dados: entrevista em profundidade; Canvas de proposta de valor; teste de usabilidade da Plataforma Sucupira e facilitação de *design thinking*.

No primeiro semestre de 2020, a pesquisa foi ampliada para outras IES privadas, públicas federais e estaduais, além de agências de fomento, utilizando a mesma metodologia empregada anteriormente, e foram obtidos resultados semelhantes. Ao todo, foram visitadas 26 agências de fomento, 8 instituições de ensino superior, 2 agências de fomento federais, 1 empresa privada, 2 órgãos de governo e 2 empresas sem fins lucrativos.

A pesquisa de campo exploratória visou identificar o nível de maturidade no tratamento e armazenamento dos dados, os padrões tecnológicos, bem como as

dificuldades que as instituições encontram nesse processo. Ainda, possibilitou identificar os modelos de dados adotados e mapear as similaridades, necessidades e expectativas.

Dentre as dificuldades encontradas, destacam-se a necessidade de obtenção de dados certificados e o retrabalho nas prestações de contas para as agências de fomento.

O mapeamento subsidiou o desenvolvimento das Ferramentas de Apoio à Integração de Sistemas (FAIS) realizado pela equipe do Hub Nacional de Integração de Dados da CAPES em parceria com uma equipe da Universidade Estadual Paulista "Júlio de Mesquita Filho" (Unesp), que foi designada para atuar no projeto. Esse grupo de trabalho iniciou as atividades em março de 2020, quando começaram a ser realizados experimentos em caráter de testes no ambiente digital da Unesp. O objetivo consistia em identificar as principais necessidades para a criação de um conjunto de ferramentas de integração de dados com capacidade de uso por qualquer instituição brasileira.

3 RESULTADOS E DISCUSSÕES

A pesquisa de campo exploratória possibilitou identificar, nas IES participantes, diversos cenários tecnológicos. Algumas IES possuem setores de Tecnologia da Informação (TI) bem definidos e estruturados, com sistemas acadêmicos implantados e equipe suficiente para colaborar com novos projetos, enquanto outras IES não possuem equipe e setor de TI definido, orçamento para investimento em TI e sistema de gerenciamento implantado.

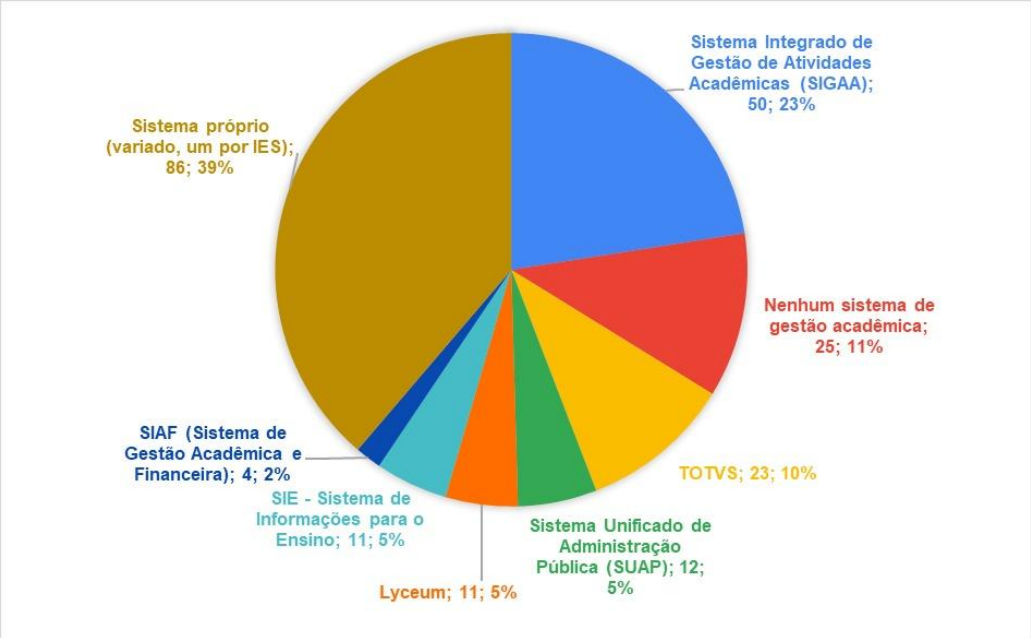
Poucas agências de fomento tinham estrutura necessária para iniciar a colaboração em um projeto de integração de dados. Das 26 agências de fomento

visitadas, apenas 4 possuíam maturidade tecnológica e pessoal disponível para a participação no projeto.

Não foram identificados padrões de dados e ontologias nas IES e tampouco conceitos de semântica e ontologias nas agências de fomento. Em muitos casos, as dificuldades se expandem para os conceitos como, por exemplo, chegar a um consenso sobre o que é bolsa de pesquisa e auxílio.

Posteriormente, no ano de 2023, foi realizada pesquisa junto aos programas de pós-graduação objetivando revisar e atualizar dados, visando identificar se houve alguma modificação em relação à pesquisa anterior, sobretudo no que tange ao amadurecimento tecnológico e à adoção de sistemas de controle acadêmico. A pesquisa foi realizada em forma de formulário disponibilizado na Plataforma Sucupira entre os meses de janeiro a julho de 2023 e contou com 220 respostas. O resultado desta pesquisa é apresentado na Figura 1.

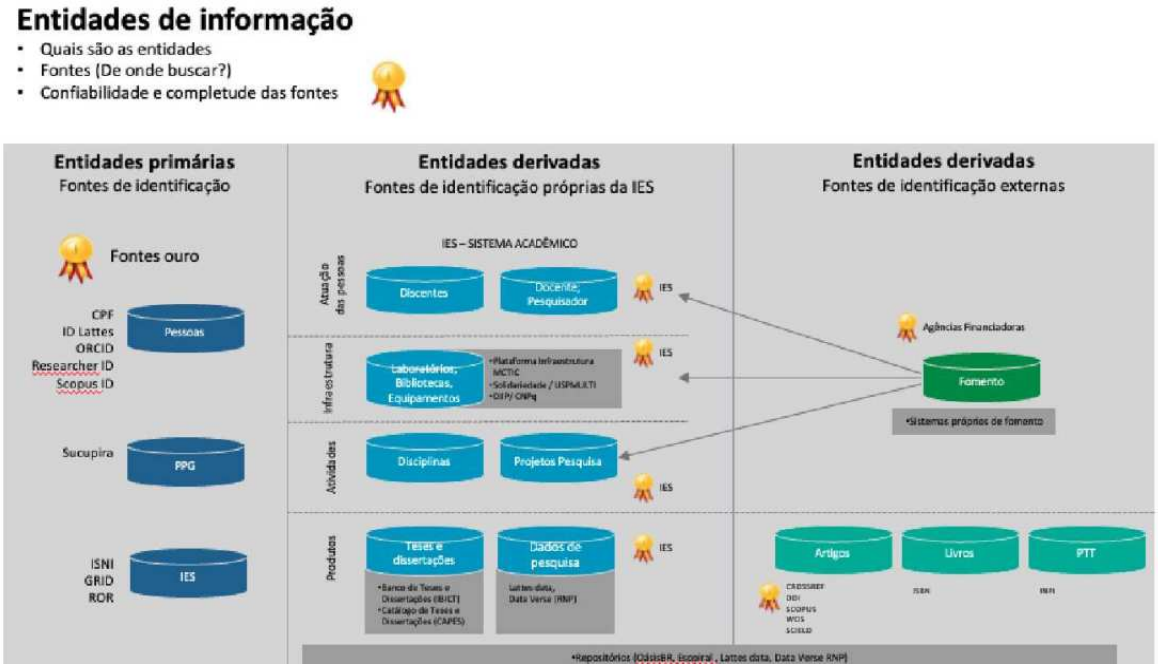
Figura 1: Resultado da pesquisa sobre sistemas acadêmicos - 2023



Fonte: Diretoria de Avaliação (DAV) da CAPES.

Os dados coletados nas pesquisas de campo possibilitaram mapear as fontes de armazenamento de dados existentes em cada instituição, sua origem e a possibilidade de integração, conforme apresentado na Figura 2.

Figura 2: Mapeamento das Fontes de Armazenamento de Dados



Fonte: Nobrega, Torino, Vidotti, Halicki e Oliveira (2022).

Considerando a diversidade de sistemas de informação, modelos de dados e tecnologias utilizadas pelas instituições brasileiras que compõem o ecossistema da ECTI, a solução Ferramentas de Apoio à Integração de Sistemas (FAIS) foi desenvolvida como um produto customizável, capaz de se adaptar e ser utilizado em qualquer instituição, independentemente das tecnologias disponíveis, bem como das diferentes faixas de dados e maturidade no tratamento e armazenamento de dados.

Os testes preliminares mostraram que a solução FAIS é robusta e satisfatória, o que culminou na inclusão das FAIS como um serviço do Hub Nacional de Integração de Dados da CAPES. Com isso, o Hub Nacional de Integração de Dados oferecerá

serviços para as demais instituições que compõem o ecossistema brasileiro de ECTI, conforme explicitado na Figura 3.

Figura 3: Hub Nacional de Integração de Dados



Fonte: Diretoria de Avaliação (DAV) da CAPES.

Os princípios FAIR tornam-se fundamentais para o sucesso da integração dos sistemas de informação que compõem esse ecossistema, tendo em vista a diversidade e o volume de dados que podem ser interoperados entre as diversas instituições do ecossistema brasileiro de ECTI e a necessidade de serem utilizados padrões pré-definidos para a comunicação de dados entre as instituições participantes do Hub de Integração de Dados. Assim, a adoção dos princípios FAIR (FORCE11, [2020?]; GOFAIR, [2020?]) viabilizam a interoperabilidade entre as diferentes fontes de dados, objetivo principal deste trabalho.

Segundo Torino, Coneglian e Vidotti (2020, p. 16),

> Os princípios FAIR possibilitam identificar pontos de convergência no tratamento dos dados, metadados e infraestruturas visando maximizar sua localização, acesso e uso. Dessa forma, é necessário que todos os envolvidos – desde o planejamento da pesquisa, a coleta, o tratamento e o armazenamento dos dados – tenham em mente a adoção dos princípios, considerando as especificidades dos dados e do domínio.

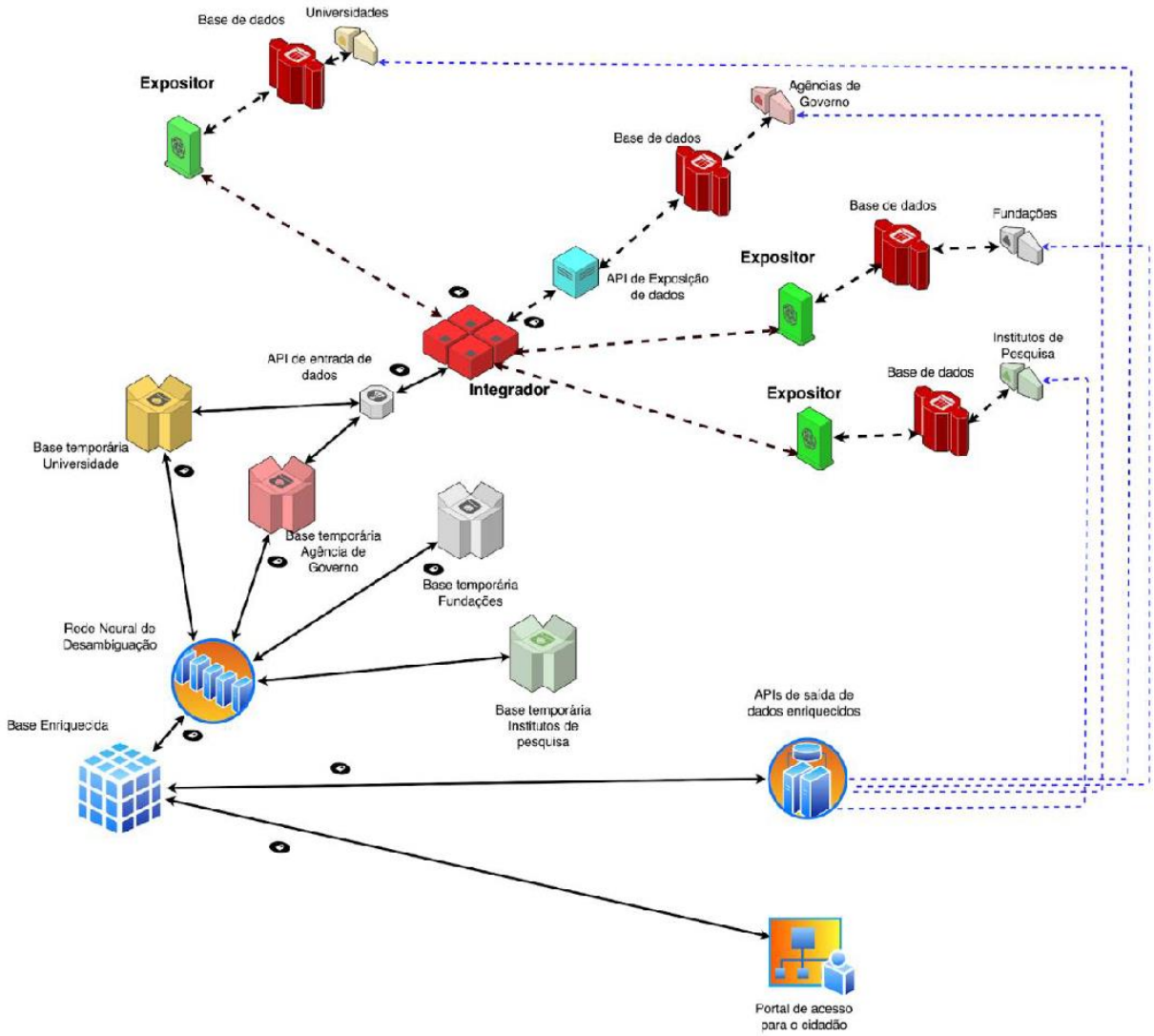
A arquitetura das Ferramentas de Apoio à Integração de Sistemas (FAIS) foi desenvolvida após análise dos dados coletados na pesquisa de campo exploratória. A primeira solução desenvolvida foi um *data display*, denominado FAIS Expositor, que pode ser implementado nos parques tecnológicos de cada instituição participante. A segunda solução desenvolvida foi uma plataforma de consumo de *Application Programming Interface* (API), denominada FAIS Integrador, responsável pelo consumo os dados compartilhados pela FAIS Expositor ou por outra ferramenta tecnológica – a exemplo de protocolos e *web services* –, e alimentação do ambiente informacional digital do Hub Nacional de Integração de Dados (Figura 4).

A FAIS Integrador é uma ferramenta projetada para apoiar a integração de sistemas, com um foco específico em operações de extração, transformação e carga (ETL). Essa ferramenta foi criada com o objetivo de oferecer uma abordagem descomplicada para a movimentação e transformação de dados entre diferentes sistemas de informação. Uma das principais características que torna a FAIS Integrador tão atrativa é possibilidade de integração com a ferramenta do Apache NiFi, proporcionando um desempenho confiável e escalável.

Uma das principais vantagens da FAIS Integrador é sua interface administrativa intuitiva. A tela administrativa foi projetada com foco na facilidade de uso, permitindo que os consumidores realizem transformações desejadas nos dados de maneira eficiente e sem a necessidade de conhecimentos técnicos avançados. Isso possibilita que profissionais de diversas áreas possam utilizar a ferramenta para modelar fluxos

de dados de acordo com suas necessidades específicas, eliminando a barreira entre os departamentos e acelerando os processos de integração.

Figura 4: Arquitetura das Ferramentas de Apoio à Integração de Sistemas (FAIS)



Fonte: Nobrega, Torino, Vidotti, Halicki e Oliveira (2022).

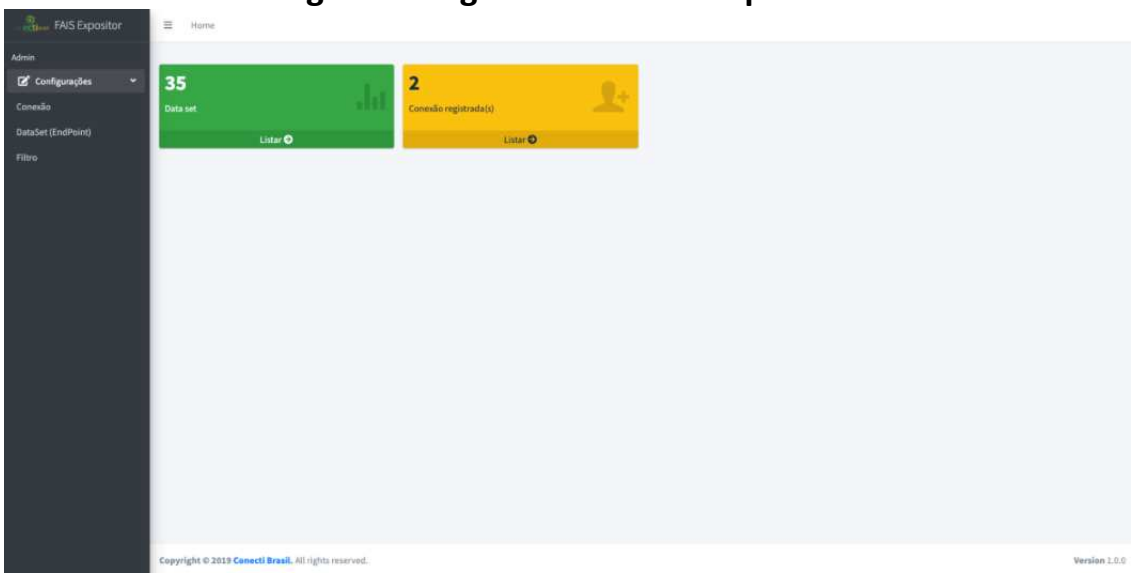
Além da simplicidade da interface administrativa, a FAIS Integrador também oferece uma área exclusiva para receber *templates* do Apache NiFi. Essa funcionalidade é especialmente valiosa para usuários que já estão familiarizados com o NiFi, possibilitando a importação de *templates* predefinidos para acelerar ainda

mais o processo de configuração e transformação de dados. Isso não apenas otimiza tempo, mas também favorece a consistência e a confiabilidade nos fluxos de dados criados.

Além disso, na FAIS Integrador é possível traduzir os dados recebidos para o padrão semântico adotado na base de dados de destino, seguindo todas as regras adotadas pela equipe responsável pela coleta dos dados. Essa tradução é configurada apenas uma vez pelas equipes de tecnologia da informação (TI) de ambos os envolvidos (fornecedor e consumidor dos dados).

A FAIS Expositor foi desenvolvida como uma solução capaz de se conectar em uma fonte de dados com modelos relacionais para atender a heterogeneidade dos cenários encontrados e, rapidamente, ser capaz de expor as tabelas de dados dessa fonte em forma de API Rest. A FAIS Expositor (Figura 5) é alinhada aos modernos padrões existentes e aderente aos padrões de segurança atuais.

Figura 5: Página inicial FAIS Expositor



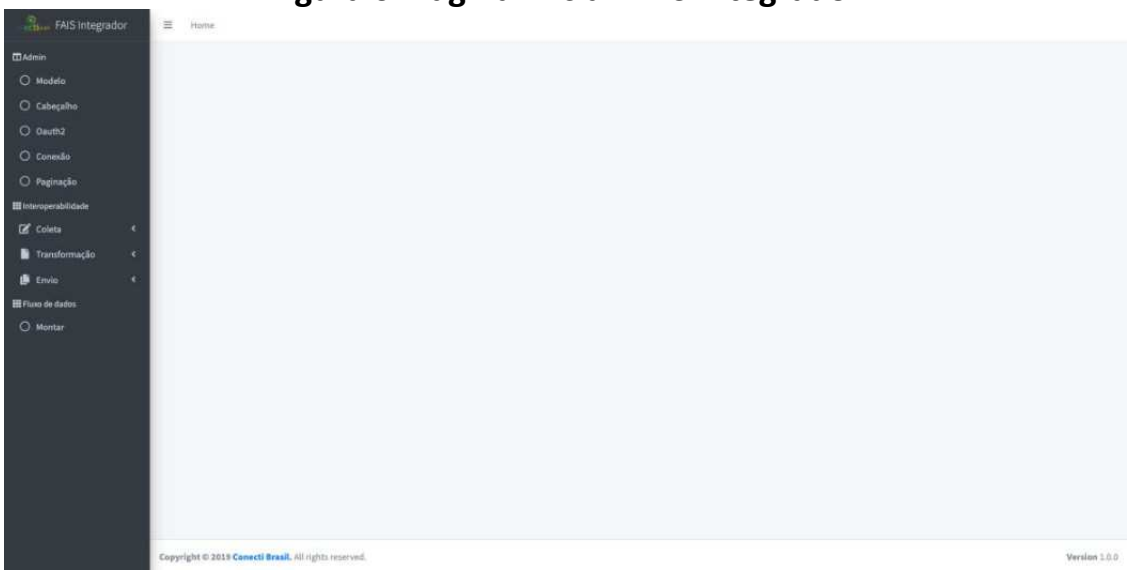
Fonte: autoria própria.

De acordo com Dietrich *et al*. ([2022?]), no contexto dos dados abertos, os dados podem ser disponibilizados por API, que permitem estabelecer quais dados serão disponibilizados a partir de um banco de dados atualizado em tempo real.

Havia, ainda, a necessidade de capturar dados de instituições que não possuem fonte de dados estruturados, mas que possuem dados relevantes a serem integrados e que também não possuem estrutura computacional que permita uma integração. Também foi necessário encontrar uma forma de capturar dados em páginas web, cuja fonte de dados está apenas em portais e sem forma de exposição, como, por exemplo, a base da Biblioteca Nacional Brasileira.

Desse modo, foi desenvolvida a solução FAIS Integrador (Figura 6) para coleta de dados por meio da importação de planilhas e de arquivos de texto. Essa solução é capaz de capturar informações de páginas web, o que possibilita a realização de buscas e coletas de forma automatizada e periódica. Além disso, a solução precisava ser capaz de traduzir os modelos de dados semânticos coletados para o modelo de dados adotados pelo consumidor.

Figura 6: Página inicial FAIS Integrador

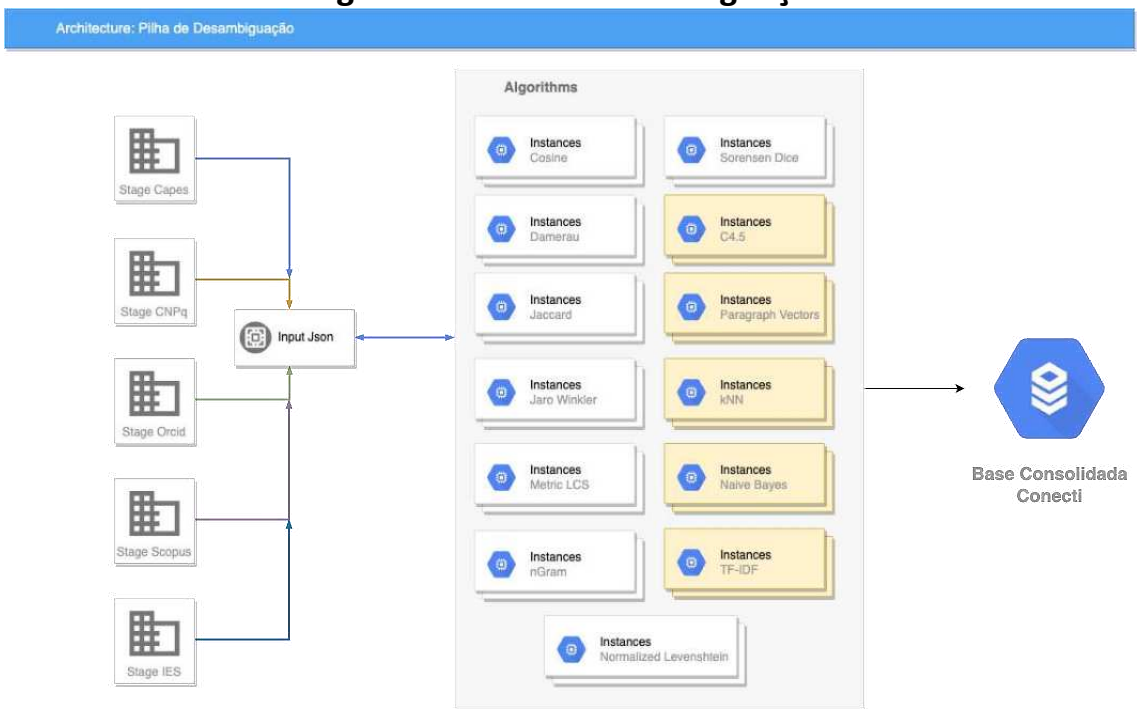


Fonte: autoria própria.

Cada conjunto de metadados coletado utilizando a FAIS Integrador é gravado em um banco de dados relacional temporário, ao qual apenas o fornecedor dos dados tem acesso por meio de uma chave de acesso única para alterar os dados armazenados. Com isso, tem-se bancos de dados separados para cada parceiro, o que se configura como uma fonte de dados.

Ao observar uma amostra desses dados coletados, percebeu-se que seria necessária a higienização e o cruzamento de dados entre essas bases, objetivando atender às políticas públicas vigentes, como, por exemplo, a necessidade de identificar pessoas de forma única entre as bases coletadas.

Figura 7: Pilha de Desambiguação



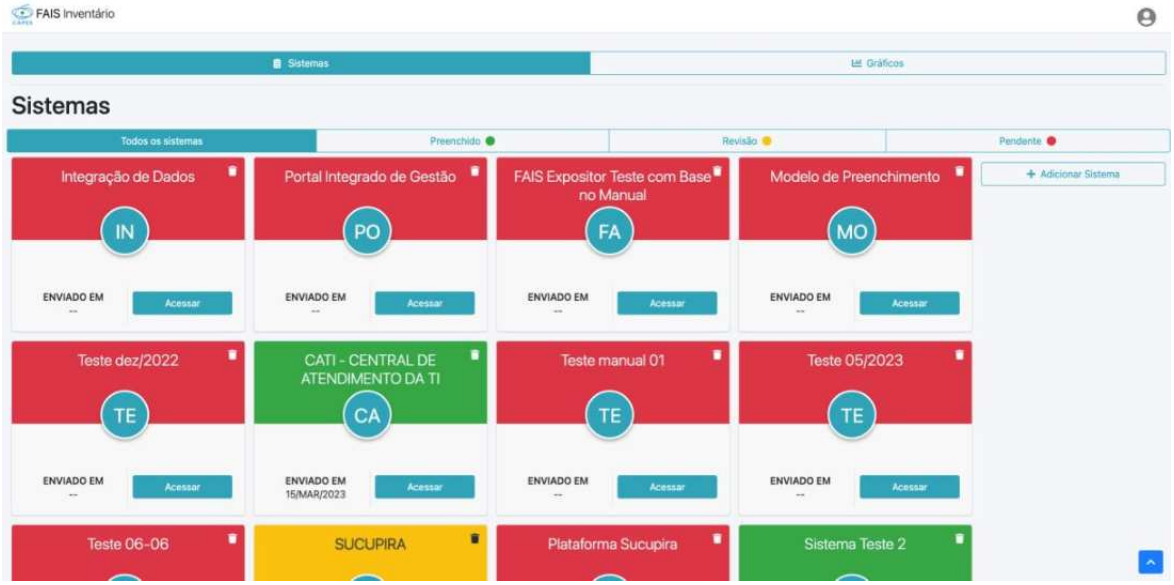
Fonte: autoria própria.

Para solucionar, esse problema, foi desenvolvida uma solução de inteligência artificial composta por 13 algoritmos rodando em pilha – designada como Pilha de Desambiguação (Figura 7) –, capaz de identificar o tipo de metadado que está sendo

trafegado e, de forma autônoma, decidir o melhor algoritmo a ser utilizado em cada conjunto de metadados. O conjunto de dados resultantes dessa desambiguação é salvo em uma base de dados enriquecida e, dessa base, são disponibilizados API's de consulta para que os parceiros possam obter dados mais limpos e completos para seus sistemas.

Também nessa base final, foi planejada uma forma única para que os usuários finais tenham acesso aos dados pessoais trafegados entre essas instituições. Desse modo, as pessoas passam a ter controle sobre a disponibilização ou não dos próprios dados nesse banco, em cumprimento à Lei Geral de Proteção de Dados Pessoais (LGPD) (Brasil, 2019), (Figura 8).

Figura 8: FAIS Inventário



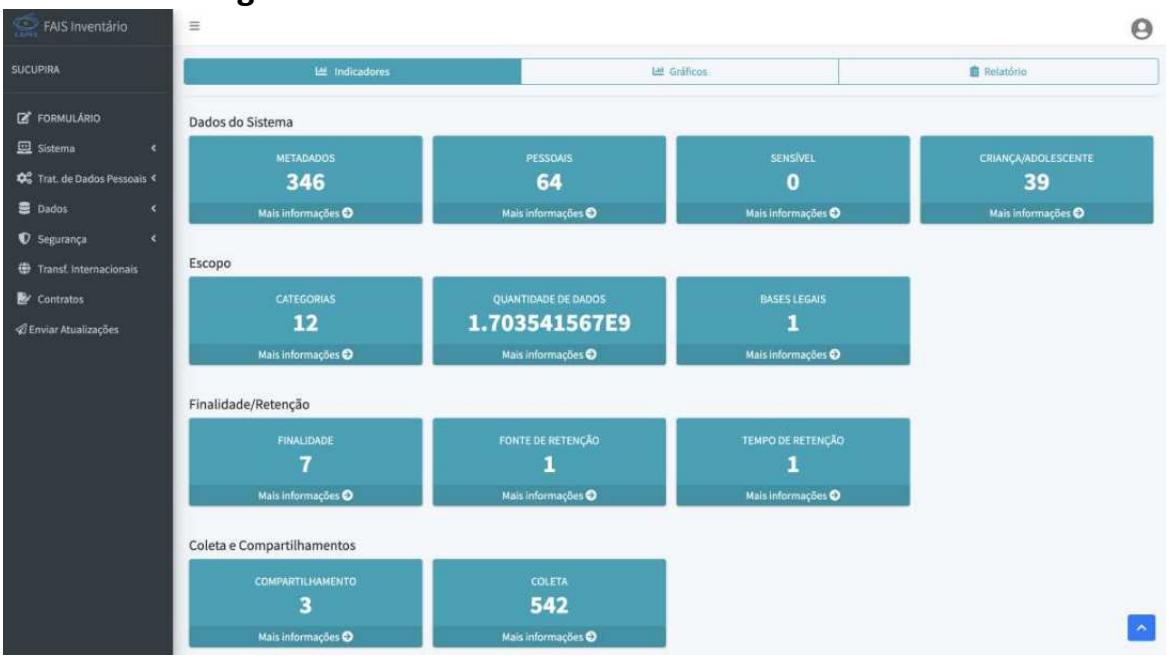
Fonte: autoria própria.

A FAIS Inventário foi construída objetivando facilitar a governança dos dados capturados das diversas fontes mapeadas. A ferramenta foi projetada para se conectar nas bases existentes e, com a ajudada de uma Inteligência Artificial (IA),

identificar de forma autônoma quais metadados foram armazenados e a natureza dos mesmos à luz da LGPD (Figura 9).

A FAIS Inventário também foi projetada para gerar o mapa de relacionamento dos dados e o Relatório de Impacto à Proteção de Dados Pessoais (RIPD), automaticamente após do completo preenchimento dos formulários do sistema.

Figura 9: FAIS Inventário – Visão dos Metadados

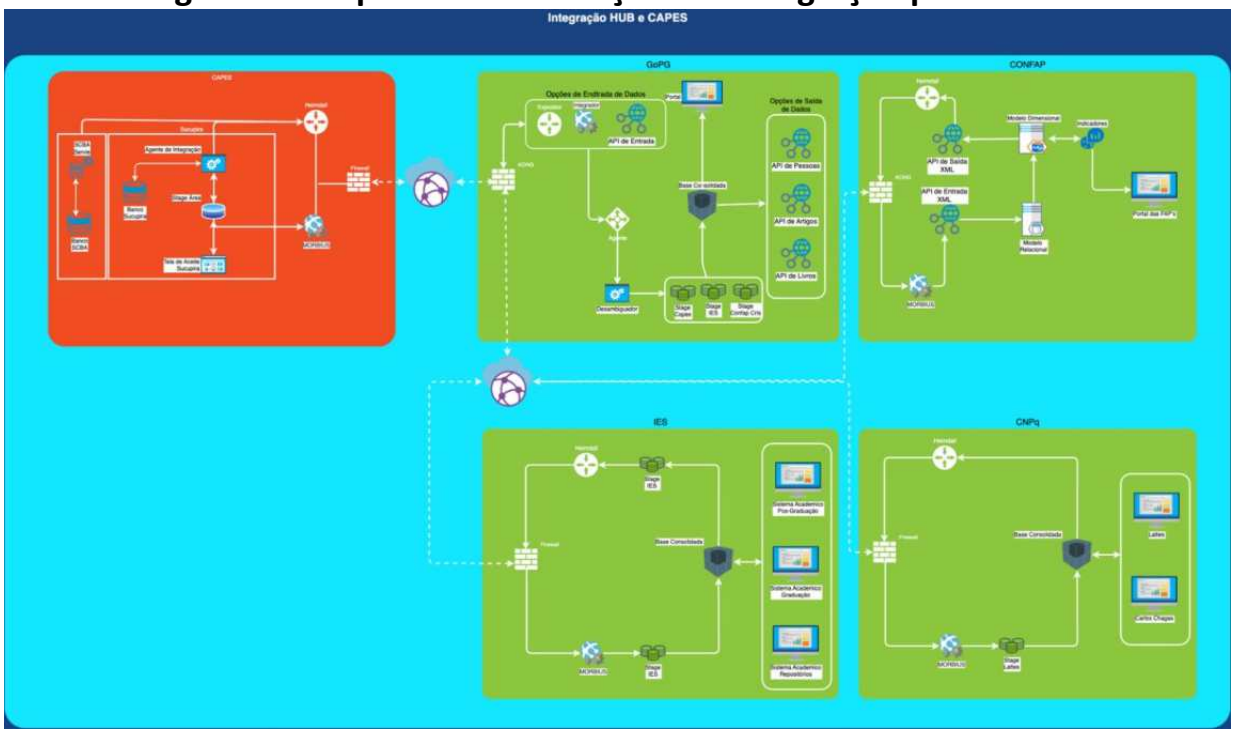


Fonte: autoria própria.

Em suma, foram construídos uma solução capaz de expor dados de instituições sem tecnologia para tal; outra solução para captura de dados em qualquer tipo de base ou arquivo estruturado; um conversor de modelo de dados e semântica; uma rede neural com aprendizado supervisionado destinada a higienização e desambiguação de dados; e uma forma de visualização centralizada para o cidadão averiguar e autorizar o uso de dados pessoais entre as instituições.

Essas soluções foram desenvolvidas de forma dinâmica o suficiente para serem customizadas de acordo com as necessidades de cada instituição. É possível gerar a documentação das API's que podem ser escaláveis para aderir a qualquer protocolo de segurança utilizado. Destaca-se, ainda, que as soluções são independentes e podem ou não ser implementadas em conjunto (Figura 10).

Figura 10: Arquitetura de soluções de Integração para CAPES



Fonte: autoria própria.

As soluções foram construídas para usar poucos recursos computacionais, e também podem ser encapsuladas em *docker*, *kubernetes* ou *openshift*, construído em Java 11 - LTS, usando o framework Spring Boot 2.2.1.RELEASE, Spring cloud 2.2.1.RELEASE, Microservices with validation FK, AOP (programação orientada a aspectos), Service Mesh, Hateoas, GraphQL, Query Search, Rancher v2.3.2, MongoDB,

PostgreSQL, Oracle, Kong, Jfrog, Gitlab Pipeline, Harbor, Graylog 2, Jwt and Jwe, Grafana and Prometheus.

Inicialmente, as soluções FAIS Expositor e FAIS Integrador foram testadas para coletar dados das bases de dados da Capes, Biblioteca Nacional, Google Scholar, Biblioteca Aberta, BD ISBN, Scopus, ORCID, Google Livros, Plataforma Lattes e uma universidade parceira para obter dados do sistema acadêmico. As soluções foram preparadas para suportar 60 mil requisições simultâneas por segundo, com a possibilidade de dobrar a capacidade escalando mais computadores no cluster.

As soluções FAIS Expositor e FAIS Integrador foram projetadas para se adequar a qualquer fonte de dados existente nos participantes, sejam os dados estruturados ou não estruturados, que podem ser recebidos, por exemplo, via API ou upload de arquivos. Toda a configuração para receber e exibir dados é simplificada objetivando atender a qualquer cenário existente no ecossistema brasileiro de Educação, Ciência, Tecnologia e Inovação (ECTI).

4 CONSIDERAÇÕES FINAIS

As Ferramentas de Apoio a Integração de Sistemas (FAIS) idealizadas visam abstrair a necessidade de se definir modelos de dados a serem adotados pelos diversos atores do ecossistema brasileiro de Educação, Ciência, Tecnologia e Inovação (ECTI), possibilitando a conversão dos modelos existentes para o modelo de dados adotado pelo Hub Nacional de Integração de Dados, bem como para o modelo de dados existente na organização na qual as soluções estiverem instaladas. Entretanto, a necessidade de os atores definirem a semântica de dados no ecossistema será o ponto diferencial para a integração entre as plataformas.

A não utilização de técnicas eficientes capazes de promover a colaboração entre organizações que desejam trocar informações, mantendo diferentes estruturas internas e processos de negócios variados, acarretará a inviabilização do projeto de integração proposto para a utilização dessas soluções e a construção do Hub Nacional de Integração de Dados.

O objetivo do serviço ou do desenvolvimento destas tecnologias é otimizar a obtenção dos dados, gerando mais agilidade e eficiência para a instituição e para os próprios envolvidos, além de simplificar o trabalho, seja para a instituição, o governo ou o cidadão.

As ferramentas desenvolvidas e apresentadas neste estudo auxiliam instituições, fundações de apoio e de governo, como solução para a comunicação de dados entre sistemas de informação. É sabido que há a necessidade de avanços nas infraestruturas de governança de dados em todas as instâncias envolvidas, o que requer a compreensão dos dados por humanos e máquinas e a adoção de estrutura semântica para os dados. Dessa forma, tem-se nas FAIS uma solução importante para a melhoria dos processos de comunicação de dados, sobretudo no contexto da pesquisa e da pós-graduação *stricto sensu* brasileira, podendo se constituir em um impulsionador de uma infraestrutura de dados acionáveis por máquinas.

REFERÊNCIAS

BRASIL. **Lei nº 12.527, de 18 de novembro de 2011**. Regula o acesso a informações previsto no inciso XXXIII do art. 5º, no inciso II do § 3º do art. 37 e no § 2º do art. 216 da Constituição Federal; altera a Lei nº 8.112, de 11 de dezembro de 1990; revoga a Lei nº 11.111, de 5 de maio de 2005, e dispositivos da Lei nº 8.159, de 8 de janeiro de 1991; e dá outras providências. Brasília, DF: Casa Civil, 2011. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/l12527.htm. Acesso em: 1 maio 2018.

BRASIL. **Lei nº 13.853, de 8 de julho de 2019**. Altera a Lei nº 13.709, de 14 de agosto de 2018, para dispor sobre a proteção de dados pessoais e para criar a Autoridade Nacional de Proteção de Dados; e dá outras providências. Brasília, DF: Secretaria Geral, 2019. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm. Acesso em: 1 maio 2022.

COORDENAÇÃO DE APERFEIÇOAMENTO DE PESSOAL DE NÍVEL SUPERIOR. **Sobre a CAPES**. [Brasília]: Ministério da Educação, 2023. Disponível em: https://www.gov.br/capes/pt-br/acesso-a-informacao/perguntas-frequentes/sobre-a-cap. Acesso em: 20 nov. 2023.

DIETRICH, Daniel; GRAY, Jonathan; McNAMARA, Tim; POIKOLA, Antti; POLLOCK, Rufus; TAIT, Julian; ZIJLSTRA, Ton. **Open data handbook**: guia de dados abertos. *[S. l.]*: Open Knowledge, [2022?]. Disponível em: http://opendatahandbook.org/guide/pt_BR/. Acesso em: 20 jul. 2022.

FORCE11. **Guiding principles for findable, accessible, interoperable and re-usable data publishing version b1.0**. *[S. l.]*: FORCE11, [2020?]. Disponível em: https://force11.org/info/guiding-principles-for-findable-accessible-interoperable-and-re-usable-data-publishing-version-b1-0/. Acesso em: 7 abr. 2020.

GOFAIR. **FAIR principles**. *[S. l.]*: GOFAIR, [2020?]. Disponível em: https://www.go-fair.org/fair-principles/. Acesso em: 7 abr. 2020.

NOBREGA, Harrysson Gilgamesh Medeiros; TORINO, Emanuelle; VIDOTTI, Silvana Aparecida Borsetti Gregorio; HALICKI, Jaqueline Pereira Carvalho; OLIVEIRA, Talita Moreira de. Integração de dados de ECTI entre diferentes sistemas de informação: proposta de solução. *In*: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 22., 2022, Porto Alegre. **Anais** [...]. Porto Alegre: Universidade Federal do Rio Grande do Sul, 2022. p. 1-11. Disponível em: https://enancib.ancib.org/index.php/enancib/xxiienancib/paper/view/1241. Acesso em: 9 fev. 2023.

OPEN KNOWLEDGE FOUNDATION. **Fair Principles**. Disponível em: https://www.go-fair.org/fair-principles/. Acesso em: 7 abr. 2022.

TORINO, Emanuelle; CONEGLIAN, Caio Saraiva; VIDOTTI, Silvana Aparecida Borsetti Gregorio. Estruturas de representação para reuso de dados no contexto da ecologia

de pesquisa: CRIS Institucional. **Informação & Informação**, Londrina, v. 25, n. 3, p. 1-27, 2020. Disponível em: http://www.uel.br/revistas/uel/index.php/informacao/article/view/41946. Acesso em: 24 nov. 2020.

TORINO, Emanuelle; TREVISAN, Gustavo Lunardelli; VIDOTTI, Silvana Aparecida Borsetti Gregorio. Dados abertos CAPES: um olhar à luz dos desafios para publicação de dados na web. **Ciência da Informação**, Brasília, v. 48, n. 3, p. 38-46, 2019. Disponível em: https://revista.ibict.br/ciinf/article/view/4866. Acesso em: 10 jun. 2022.

TORINO, Emanuelle; VIDOTTI, Silvana Aparecida Borsetti Gregorio. Boas práticas para dados na web: análise do portal Dados Abertos Capes. **Informação & Sociedade**: Estudos, João Pessoa, v. 31, n. 1, p. 1-25, 2021. Disponível em: https://periodicos.ufpb.br/index.php/ies/article/view/50790. Acesso em: 10 jun. 2022.

AGRADECIMENTOS

Agradecemos à Coordenadoria de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo financiamento desta pesquisa.

Copyright: Esta obra está licenciada com uma Licença Creative Commons Atribuição 4.0 Internacional.

tpbci@ancib.org @anciboficial @ancib_brasil