

INTEGRAÇÃO E REUSO DE DADOS PARA O POVOAMENTO SEMIAUTOMÁTICO DE DISSERTAÇÕES E TESES NO REPOSITÓRIO INSTITUCIONAL DA UTFPR

DATA INTEGRATION AND REUSE FOR THE SEMI-AUTOMATIC ARCHIVING OF DISSERTATIONS AND THESIS IN THE UTFPR INSTITUTIONAL REPOSITORY

Emanuelle Torino^a

Nestor Cortez Saavedra Filho^b

RESUMO

Objetivo: Descreve o processo utilizado para integração e reuso de dados para o povoamento semiautomático de dissertações e teses no Repositório Institucional da UTFPR (RIUT), a partir dos Sistemas Corporativos da UTFPR. **Metodologia:** Realizou-se um estudo de caso na Universidade Tecnológica Federal do Paraná (UTFPR), e a observação participante como método de investigação qualitativa para análise dos dados disponíveis nos sistemas de informação corporativos a serem integrados. **Resultados:** Apresenta, com base na Arquitetura da Informação, a estruturação do fluxo de entrega de dissertações e teses e o estudo da camada de dados dos Sistemas Corporativos da UTFPR, bem como a compatibilização e a conversão para o perfil de aplicação do RIUT, visando ao povoamento semiautomático de dissertações e teses em repositório institucional. Tais resultados possibilitaram desenvolver o processo de integração e reuso de dados entre os Sistemas Corporativos da UTFPR e o Repositório Institucional. **Conclusões:** O reuso de dados já existentes nos Sistemas Corporativos para o povoamento semiautomático de dissertações e teses no Repositório Institucional gera benefícios múltiplos, dentre os quais destacam-se: agilidade e precisão, veracidade dos dados, desoneração do aluno e dos bibliotecários, preservação da memória e ampliação da visibilidade da produção acadêmica, científica e tecnológica, com vistas a fomentar o impacto e inserção social desses resultados, além de reduzir os esforços humanos anteriormente despendidos para a atividade.

Descritores: Repositórios institucionais. Dissertações e teses. Integração de dados. Reuso de dados.

^a Doutoranda pelo Programa de Pós-Graduação em Ciência da Informação (PPGCI-Unesp). Bibliotecária da Universidade Tecnológica Federal do Paraná (UTFPR). E-mail: emanuelle@utfpr.edu.br

^b Doutor em Física pela Universidade de São Paulo. Professor Associado no Departamento Acadêmico de Física da Universidade Tecnológica Federal do Paraná (UTFPR). E-mail: nestorsf@utfpr.edu.br

1 INTRODUÇÃO

As universidades, ao longo dos séculos, produzem uma infinidade de dados e informações registradas em diversos suportes apoiados pelas tecnologias existentes. No cenário atual, utilizam tecnologias informáticas para a disponibilização de sistemas de informação que auxiliam na gestão de dados e informações relacionados às diferentes áreas de atuação.

Nesse contexto, a Ciência da Informação pode fornecer suporte para o planejamento de uma ecologia informacional na qual os sistemas possam ser desenvolvidos adequadamente para atender as suas finalidades, bem como para a estruturação dos dados neles armazenados, visando ao reuso em benefício institucional.

A estruturação dessa ecologia deve estar embasada nos objetivos institucionais e em políticas que preconizem as formas de atuação no que tange à disponibilização, ao acesso, à segurança, ao *backup*, à preservação, à integração e à interoperabilidade dos dados, informações e objetos digitais gerenciados pelos sistemas.

A dinâmica da produção de objetos resultantes da atividade de pesquisa apoia-se nos processos de comunicação científica, por meio de seus canais formais e informais. O avanço das tecnologias trouxe a esse contexto ferramentas computacionais para a gestão desses conteúdos, das quais destacam-se, neste estudo, os repositórios digitais, como ambientes para a gestão e a visibilidade de objetos que resultam das atividades fim.

Nesse contexto é inegável que o processo de gestão da informação precisa ser planejado e estruturado de forma a otimizar processos, reduzir retrabalho e utilizar tecnologias para que a redundância de dados, quando necessária, seja realizada por aplicações computacionais. Assim as tecnologias de informática podem automatizar tarefas e melhorar processos, otimizando tempo e esforços dispendidos por humanos.

Dessa forma, a Arquitetura da Informação considerada “[...] uma prática profissional e um campo de estudos focado na solução dos problemas básicos

de acesso e uso das vastas informações disponíveis atualmente.” (RESMINI; ROSATI, 2011) e que objetiva planejar ambientes de modo a atender contexto, conteúdo e usuário, pode oferecer aporte para a estruturação da camada de dados dos sistemas de informação visando a compatibilizá-los de modo a automatizar os processos realizados por humanos.

No tocante à pesquisa científica e tecnológica produzida em centros de pesquisa ou programas de pós-graduação que são, no modelo brasileiro, financiados quase em sua totalidade pelo Poder Público, a visibilidade e a integridade dos dados potencializadas por essas iniciativas vão ao encontro dos anseios da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes) nas avaliações quadrienais dos Programas de Pós-Graduação (PPG), onde os quesitos impacto local/regional/nacional e potencial de internacionalização dos mesmos desempenham um papel central (BRASIL, 2020).

O presente estudo descreve o processo utilizado para integração e reuso de dados para o povoamento semiautomático de dissertações e teses no Repositório Institucional da UTFPR (RIUT), a partir dos Sistemas Corporativos da UTFPR.

Para tanto foi realizado um estudo de caso na Universidade Tecnológica Federal do Paraná (UTFPR), e a observação participante como método de investigação qualitativa para análise dos dados disponíveis nos sistemas de informação institucionais a serem integrados.

Como resultados, apresenta, com base na Arquitetura da Informação, a estruturação do fluxo de entrega da versão final de dissertação e tese e o estudo da camada de dados dos Sistemas Corporativos da UTFPR, a compatibilização e a conversão para o perfil de aplicação do RIUT, visando ao povoamento semiautomático de dissertações e teses em repositório institucional. Tais resultados possibilitaram desenvolver o processo de integração e reuso de dados entre os Sistemas Corporativos da UTFPR e o Repositório Institucional da UTFPR.

2 ARQUITETURA DA INFORMAÇÃO

A Arquitetura da Informação (AI) possui base multidisciplinar e quatro abordagens: arquitetural, sistêmica, informacional e pervasiva, conforme Oliveira (2014).

A abordagem arquitetural, preconizada por Wurman (1996, 2005), utiliza-se da arquitetura e do *design* para o tratamento de informações visando à organização em benefício da orientação e da compreensão do sujeito.

Com o advento dos ambientes digitais na web, Morville e Rosenfeld (2006) embasam-se na teoria geral dos sistemas para propor a abordagem sistêmica, apoiada em sistemas de organização, rotulagem, busca e navegação, que se articulam para que as informações possam ser acessadas e utilizadas em um ambiente.

A estas abordagens, Vidotti, Cusin e Corradi (2008) e Camargo e Vidotti (2011) agregam os elementos de representação, armazenamento, preservação, recuperação, acesso e uso da informação, como contribuições da Ciência da Informação para a chamada abordagem informacional.

E mais recentemente, Oliveira (2014) baseia-se na computação ubíqua, no *design* de serviços e na Ciência da Informação para propor a abordagem pervasiva, que altera o foco do ambiente informacional e sujeitos para as ecologias informacionais complexas, e desloca para o centro do processo a informação, buscando manter a localização do usuário na ecologia considerando ambientes analógicos e/ou digitais, espaços, tecnologias e comportamento dos sujeitos, que interagem e exercem influências múltiplas entre as partes (OLIVEIRA, 2014). Busca ainda a integração de dados e a interoperabilidade entre ambientes e ecologias para favorecer a Encontrabilidade da Informação (VECHIATO; VIDOTTI, 2014).

As demandas sociais e tecnológicas em cada uma das abordagens apresentam uma alteração no cenário da AI, que considera os relacionamentos entre sujeitos, processos e ambientes informacionais, deslocando-os para ecologias informacionais complexas, referindo-se como o “[...] conjunto de espaços, ambientes, canais, mídias, tecnologias e sujeitos com seus

comportamentos, todos interligados e conectados de maneira holística pela informação.” (OLIVEIRA; VIDOTTI, 2016, p. 96).

No contexto deste trabalho, enfoca-se a Arquitetura da Informação com abordagens sistêmica e informacional, uma vez que o foco principal é a estruturação da camada de dados dos sistemas de informação da UTFPR, visando compatibilizá-los e reutilizá-los de modo a tornar o processo de povoamento das dissertações e teses no Repositório Institucional semiautomático.

3 POVOAMENTO DE DISSERTAÇÕES E TESES NO REPOSITÓRIO INSTITUCIONAL DA UTFPR: DIAGNÓSTICO

A Universidade Tecnológica Federal do Paraná (UTFPR) é uma instituição pública federal de ensino, que oferece à comunidade cursos de graduação e pós-graduação em diferentes áreas do conhecimento. Com a transformação em Universidade, no ano de 2005, a expansão e a interiorização no Estado do Paraná, houve uma ampliação na oferta de Programas de Pós-Graduação (PPG), dos então 5 para os atuais 62, e conseqüentemente da produção acadêmica, científica e tecnológica.

Contudo, mesmo diante da Portaria nº 13 da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes), que “Institui a divulgação digital das teses e dissertações produzidas pelos programas de doutorado e mestrado reconhecidos.” (BRASIL, 2006), a UTFPR não disponibilizou um ambiente digital para os referidos materiais, a exemplo de algumas universidades brasileiras que implantaram Bibliotecas Digitais de Teses e Dissertações (BDTD), acarretando uma forma alternativa de atendimento à referida portaria: a disponibilização desses objetos digitais nas páginas web dos PPG, mantendo a versão impressa no acervo físico das Bibliotecas da Universidade.

No ano de 2010, a partir do fomento da agência Financiadora de Estudos e Projetos (FINEP), por meio do edital FINEP/PCAL/XBDB nº 1/2009, a UTFPR disponibilizou o Repositório Institucional da UTFPR (RIUT) para a gestão da

produção acadêmica e científica da Universidade, no qual passaram a ser disponibilizadas, no ano de 2011, versões digitais das dissertações e teses, sem alteração no processo de disponibilização da versão impressa. Apesar disso, os PPG mantiveram a disponibilização paralela nas suas páginas web, que, pela sua natureza, não garantem plena disseminação e recuperação dos dados e resultados de pesquisa, diluindo, portanto, o potencial de apropriação dos mesmos pela sociedade.

Quando da implantação do RIUT, foi estabelecido o trâmite de entrega da versão final de dissertações e teses realizada pelo discente ao do PPG que, após análise de adequação de conteúdo pelo orientador e verificação do cumprimento dos requisitos para a entrega pela coordenação, encaminhava de forma física o trabalho impresso e a versão digital utilizando mídias de armazenamento (CD/DVD) à Biblioteca do câmpus para depósito no RIUT (UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ, 2011).

No que tange a dissertações e teses impressas, no ano de 2018, na Portaria nº 182 da Capes, que “Dispõe sobre processos avaliativos das propostas de cursos novos e dos programas de pós-graduação *stricto sensu* em funcionamento.” (BRASIL, 2018) e revoga a Portaria nº 13 (BRASIL, 2006), não houve qualquer menção à obrigatoriedade da entrega de dissertação e tese, quer seja impressa ou digital, e, ao referir-se aos documentos de área, delega a esses documentos tal questão.

Dessa forma a UTFPR estabeleceu, na Instrução Normativa Conjunta nº 01/2018 - PROPPG/PROGRAD, que “Dispõe sobre o formato das versões de trabalhos finais enviadas às bancas de avaliação e ao depósito no Sistema de Bibliotecas da UTFPR” (UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ, 2018), o depósito exclusivo das versões finais de dissertações e teses em formato digital no Portal de Informação em Acesso Aberto da UTFPR (PIAA), vinculado ao Departamento de Bibliotecas (DEPBIB), e do qual faz parte o RIUT. Em paralelo, a alteração do Portal Institucional da UTFPR trouxe novas definições quanto aos conteúdos a serem disponibilizados, o que reforçou que as dissertações e teses estariam armazenadas apenas no RIUT e acessíveis nas páginas web dos PPG a partir de um *link* para a área do curso no RIUT.

Tal medida auxiliou na consolidação do RIUT como ambiente institucional para a gestão, preservação e ampliação da visibilidade das dissertações e teses defendidas nos PPG da UTFPR. Desde a implantação do RIUT, definiu-se institucionalmente que o depósito dos referidos materiais seria realizado por bibliotecários membros do Comitê Gestor do RIUT, constituído por um representante de cada um dos 13 (treze) câmpus da Universidade, de forma manual, considerando, para tanto, a Política de Submissão do RIUT que estabelece o perfil de aplicação para as diferentes tipologias gerenciadas e rege a representação da informação e a padronização dos metadados.

Por outro lado, a UTFPR possui uma infraestrutura de sistemas de informação mantida pela Diretoria de Gestão de Tecnologia de Informação (DIRGTI), da qual fazem parte os Sistemas Corporativos da UTFPR, constituídos por 13 sistemas, dentre os quais um destinado à Pós-Graduação *Stricto Sensu*, por meio do qual são realizados processos de interesse do presente estudo, como a matrícula e o agendamento de bancas de defesa.

Evidenciou-se assim que o sistema destinado à Pós-Graduação *Stricto Sensu* armazena dados relativos à defesa das dissertações e teses que são necessários à representação dos seus conteúdos no RIUT.

Dessa forma procedeu-se à análise da camada de dados disponíveis nos Sistemas Corporativos da UTFPR visando ao reuso e à integração de dados dos sistemas para estabelecer um fluxo para a entrega de versões finais de dissertações e teses e o reuso dos dados para o povoamento do RIUT, por meio da conversão para o perfil de aplicação de metadados, de modo a otimizar o processo de entrada de dados e torná-lo semiautomático.

Para tanto, considerando a afirmação de Macedo (2005, p. 132), “A finalidade da Arquitetura da informação é, portanto, viabilizar o fluxo efetivo de informações por meio do desenho de ambientes informacionais.”, valeu-se do arcabouço teórico-conceitual da Arquitetura da Informação para a integração dos dados dos Sistemas Corporativos da UTFPR ao RIUT, aplicando-a à camada dos dados.

Nesta perspectiva, aplica-se a Arquitetura da Informação com abordagem sistêmica, considerando os ambientes digitais um conjunto de sistemas que

constituem um todo completo e a abordagem informacional no tratamento computacional e humano necessários à representação, visando à recuperação, ao acesso e ao uso da informação.

4 INTEGRAÇÃO DE DADOS DOS SISTEMAS CORPORATIVOS COM O REPOSITÓRIO INSTITUCIONAL DA UTFPR

O contexto apresentado descreve o diagnóstico, a partir do qual a integração dos sistemas, visando ao reuso de dados para o povoamento de repositório digital, foi planejada. Para tanto, inicialmente foi necessário mapear as instâncias organizacionais envolvidas, para que o objetivo de estabelecer um fluxo de informação e integrar os dados disponíveis nos Sistemas Corporativos da UTFPR com o RIUT pudesse ser atingido. O processo de planejamento envolveu a Pró-Reitoria de Pesquisa e Pós-Graduação (PROPPG), o Departamento de Sistemas de Informação (DESI), vinculado à Diretoria de Gestão de Tecnologia da Informação (DIRGTI), e o Repositório Institucional da UTFPR (RIUT), vinculado ao Departamento de Bibliotecas (DEPBIB).

A partir do planejamento, foram definidos os requisitos necessários, agrupando-os em duas grandes áreas:

- a) gestão: responsável pela definição do fluxo de entrega das dissertações e teses, atividade que envolve discentes, docentes, coordenadores e bibliotecários. O desenvolvimento desta área ocorreu sob responsabilidade da PROPPG e da Coordenação do RIUT;
- b) tecnologia: responsável pela análise, compatibilização e estruturação dos dados para a integração dos sistemas, de forma a atender as definições determinadas pela área de gestão, embasada nos preceitos da Ciência da Informação e da Arquitetura da Informação. A atividade foi realizada pelo RIUT e operacionalizada pelo DESI.

Ressalta-se que as áreas supramencionadas atuaram em todas as etapas do processo, visando operacionalizá-lo de forma adequada ao cenário organizacional. Contudo, para fins de organização do presente estudo, serão

abordadas sequencialmente.

4.1 GESTÃO

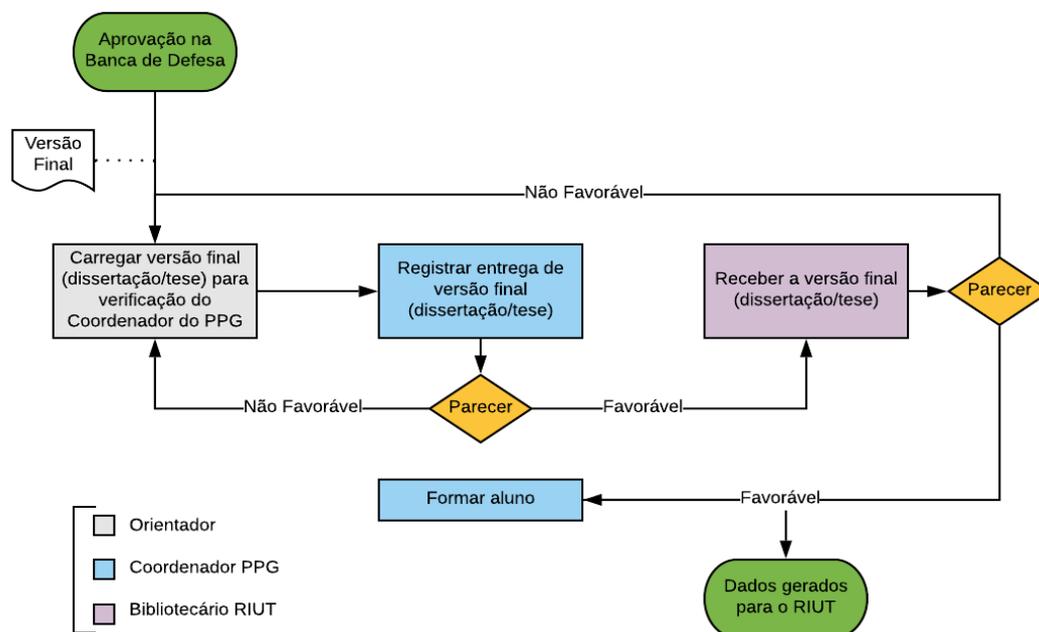
Conforme mencionado, a entrega das versões finais de dissertações e teses ocorria de forma manual, em suporte físico. Contudo, no sistema Pós-Graduação *Stricto Sensu*, estava disponível uma funcionalidade na qual o coordenador do PPG, após o envio manual da versão final do trabalho à Biblioteca, declarava individualmente a entrega, um dos requisitos necessários à solicitação da emissão do diploma do aluno, e registrava no referido sistema o nome do bibliotecário responsável pelo recebimento.

Esse processo, além de moroso, trazia em si um problema, visto que quem gravava a declaração de recebimento no sistema Pós-Graduação *Stricto Sensu* não era o bibliotecário que recebia o trabalho e sim o coordenador que o enviava. Além disso, havia uma lacuna entre os dados dos discentes formados pelos PPG e as versões finais disponíveis nas Bibliotecas da UTFPR, sobretudo no RIUT, que impossibilitava à gestão da PROPPG e do DEPBIB assegurar que a totalidade das dissertações e teses defendidas na Universidade estavam disponíveis no RIUT, de extrema relevância para o processo de avaliação dos programas de pós-graduação e para a disponibilização pública e ampliação de visibilidade dos conteúdos.

Tal situação suscitou a necessidade de estabelecer um fluxo de entrega de versões finais que possibilitasse a cada um dos atores envolvidos no processo a responsabilização pela sua competência, bem como assegurar que a totalidade dos trabalhos defendidos na UTFPR pudessem ser disponibilizados no Repositório Institucional, visando ganhos a todos os envolvidos na elaboração e aprovação dos trabalhos, bem como à Universidade e à sociedade geral e científica.

Este processo foi alterado conforme fluxograma apresentado na Figura 1.

Figura 1 - Fluxo de entrega de versão final de dissertação e tese na UTFPR



Fonte: Autoria própria.

A partir da definição do fluxo de entrega da versão final da dissertação e tese (Figura 1), após a aprovação pela banca de defesa, e atendidas as questões regulamentares, o aluno entrega a versão final ao orientador, responsável pela homologação do conteúdo, pelo *upload* do(s) objeto(s) digital(is) no(s) formato(s) estabelecidos pela IN 10/2020 PROPPG (UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ, 2020b) no sistema Pós-Graduação *Stricto Sensu*, pelo registro do termo de ciência do cumprimento dos documentos Institucionais vigentes e pelo envio ao coordenador do PPG. Destaca-se que os PPG que têm como produção, além da dissertação ou tese, produtos educacionais ou tecnológicos devem carregá-los no sistema Pós-Graduação *Stricto Sensu* e enviá-los ao RIUT.

Neste sentido, compete ao orientador indicar, no momento do *upload* da versão final do trabalho, duas outras informações: a solicitação de proteção da propriedade intelectual do trabalho ou produto elaborado, por meio de patente ou registro, e o uso de informações de empresas/organizações/instituições públicas ou privadas nominadas ou identificáveis no trabalho e/ou produto.

No primeiro caso, quando há perfil de patente ou registro, o trabalho passa a ser entendido como sigiloso e com isso é enviado ao RIUT com metadados

disponíveis e restrição de acesso ao texto completo, que é embargado pelo período de 18 meses a contar da data de defesa. Nesse caso, a Agência de Inovação da UTFPR (AGINT) e o Núcleo de Inovação Tecnológica (NIT) do câmpus da Universidade no qual o curso é ofertado é notificado por e-mail automático de que o trabalho defendido pretende solicitar patente ou registro, para que possam apoiar o autor e o orientador do trabalho no processo.

No segundo caso, quando há uso de informações de empresas/organizações/instituições públicas ou privadas nominadas ou identificáveis no trabalho e/ou produto, é necessário que o orientador faça o *upload* do termo de autorização para a divulgação emitido pela empresa, conforme modelo disponível na IN 10/2020 PROPPG.

Compete ainda ao orientador verificar se o trabalho e o produto educacional ou tecnológico (quando se aplica) apresenta a licença *Creative Commons* regulamentada pelo PPG, conforme determina a Resolução Conjunta nº 01/2020 COGEP-COPPG que “ Dispõe sobre a Política de licenciamento das versões finais dos Trabalhos de Conclusão de Curso da Graduação (TCC) e dos Programas de Pós-Graduação Stricto Sensu (dissertações e teses), bem como dos produtos educacionais e tecnológicos a elas vinculados, produzidas no âmbito da Universidade Tecnológica Federal do Paraná.” (UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ, 2020a).

Posteriormente, o coordenador do Programa de Pós-Graduação verifica a conformidade do processo anterior, realizado pelo orientador, bem como o atendimento aos documentos institucionais e ao documento de área de avaliação Capes vigentes e emite parecer: caso favorável, o fluxo segue para a Biblioteca e permite que o aluno seja formado, ação que possibilita o requerimento do diploma, e caso não favorável, retorna para os ajustes indicados a serem realizados sob a responsabilidade do orientador.

Ao bibliotecário membro do Comitê Gestor do RIUT no câmpus com o qual o Programa de Pós-Graduação possui vínculo compete analisar a versão final do trabalho e do produto educacional a ela vinculado, considerando os documentos institucionais vigentes, o atendimento às normas para a apresentação de trabalho acadêmico vigentes na UTFPR e a presença da licença *Creative Commons*

regulamentada pelo PPG, para a emissão do parecer: caso favorável, os requisitos necessários à conclusão do curso são cumpridos e a situação do aluno no sistema Pós-Graduação *Stricto Sensu* é alterada para formado, o que permite o requerimento de emissão de diploma do aluno; e caso não favorável, o fluxo retorna para o orientador, responsável pelos ajustes indicados.

Após a emissão do parecer favorável pelo bibliotecário membro do Comitê Gestor do RIUT, os dados da defesa são gerados, atendendo aos requisitos definidos pela área tecnológica, descritos a seguir, para serem integrados ao RIUT. Destaca-se que, com eles, os objetos digitais das versões finais e dos produtos educacionais ou tecnológicos, quando se aplica, são migrados do sistema Pós-Graduação *Stricto Sensu* para o RIUT, o que possibilita o tratamento dos dados e a disponibilização no RIUT.

Destaca-se que, em todas as etapas, em caso de parecer não favorável, há um campo para o registro da ocorrência, o que possibilita ao autor e ao orientador da dissertação ou tese identificar a necessidade de adequação. Além disso, em todas as etapas são enviadas notificações por e-mail para todos os envolvidos nas etapas anteriores do fluxo de entrega da versão final da dissertação e tese. Todo o histórico deste fluxo é armazenado no sistema Pós-Graduação *Stricto Sensu*.

4.2 TECNOLOGIA

No que tange à área tecnologia, havia duas necessidades: a análise dos dados disponíveis nos 13 sistemas que compõem os Sistemas Corporativos da UTFPR para compatibilização, conversão para o perfil de aplicação de metadados e integração com o RIUT; e o desenvolvimento de funcionalidades para operacionalizar o fluxo descrito na Figura 1.

A integração foi alicerçada em 3 processos, para os quais foram consideradas as estruturas de dados e especificidades de cada sistema: análise, processamento e depósito, todas elas permeadas por tratamento, quer seja humano ou computacional.

Os Sistemas Corporativos da UTFPRⁱ são desenvolvidos localmente e constituídos por 13 sistemas: Acadêmico, Almoxarifado, Avaliação Institucional,

Corporativo, Identificação, Estágio Interno, Gestão de Pessoas, Orçamento e Gestão, Patrimônio, Portal do Aluno, Pós-Graduação, Sistema de Apoio ao Usuário e Sistema de Acesso, que compartilham da mesma infraestrutura. Esses sistemas são utilizados para os dados das 14 unidades da UTFPR: a Reitoria e seus 13 câmpus.

Por outro lado, o Repositório Institucional da UTFPR (RIUT)ⁱⁱ utiliza o DSpace, software de código aberto desenvolvido e distribuído pela Lyris, visando à implementação de repositórios digitais com a função de gestão, preservação e ampliação de visibilidade da produção intelectual.

Durante a análise, os Sistemas Corporativos da UTFPR foram definidos como provedores de dados, por isso suas estruturas e os dados armazenados foram analisados para determinar os elementos de compatibilização, bem como a possibilidade de adição de dados ainda não existentes. Nesta etapa o perfil de aplicação de metadados do RIUT para dissertações (D) e teses (T) foi de extrema relevância para estruturar a integração.

Destaca-se que os Sistemas Corporativos da UTFPR não adotam vocabulário padronizado de dados, enquanto o RIUT adota o *Dublin Core* (DC), o que gera a necessidade de processamento para a conversão dos dados a serem integrados em DC no formato *Extensible Markup Language* (XML), considerando o perfil de aplicação de metadados do RIUT.

A partir do perfil de aplicação supracitado foram verificados quais dados necessários à representação de dissertações e teses estavam armazenados nos sistemas que compõem os Sistemas Corporativos da UTFPR e aqueles que poderiam ser implantados por sua relevância aos processos institucionais, além da forma de tratamento computacional mais adequada para a conversão e integração ao RIUT. A análise permitiu identificar que os dados necessários são provenientes de três dos sistemas disponíveis nos Sistemas Corporativos da UTFPR (provedores de dados): Acadêmico, Corporativo e Pós-Graduação *Stricto Sensu*, conforme apresentado no Quadro 1.

No Quadro 1 é apresentado o perfil de aplicação do RIUT para dissertações e teses, cuja adaptação consiste na inclusão dos dados explicitados nas colunas Campo nos Sistemas Corporativos da UTFPR, Sistema de origem e Responsável pela alimentação, que indicam a compatibilização dos

dados realizada para a integração.

Na coluna Elemento *Dublin Core* são apresentados os atributos utilizados para a representação das dissertações e teses no RIUT; na coluna Rótulo, como o elemento é exibido na interface do RIUT; a coluna Campo nos Sistemas Corporativos da UTFPR exibe o rótulo do campo na interface dos Sistemas Corporativos da UTFPR; na coluna Sistema de origem são indicados os sistemas disponíveis nos Sistemas Corporativos da UTFPR, dos quais os dados são coletados; a coluna Responsável pela alimentação informa o perfil de quem alimenta o campo de valor daquele metadado no sistema correspondente (Sistemas Corporativos da UTFPR ou RIUT); a coluna Aplicação se refere ao RIUT e determina quais dados são obrigatórios (O) ou obrigatórios quando presentes no documento (P); e, finalmente, a coluna R/NR, também relacionada ao RIUT, esclarece quais metadados são repetíveis e não repetíveis no RIUT.

Quadro 1 - Perfil de aplicação D/T do RIUT com a compatibilização dos dados

Elemento Dublin Core	Rótulo no RIUT	Campo nos Sistemas Corporativos da UTFPR	Sistema de origem	Responsável pela alimentação	Aplicação	R/NR
dc.type	Tipo de documento	Curso ou Tipo de Curso (rotina para virar tipologia)	Acadêmico – Cadastro do Aluno	DESYS	O	NR
dc.title	Título do documento	Título	Pós <i>Stricto Sensu</i> – Marcação de Defesa	Aluno/ Secretaria	O	NR
dc.title.alternative	Título alternativo	Em implementação	Pós <i>Stricto Sensu</i> – Marcação de Defesa	Aluno/ Orientador	O	R
dc.rights	Permissão de acesso	Processamento realizado na conversão dos dados. Em implementação	Pós <i>Stricto Sensu</i> – Upload Versão Final	Orientador/ DESYS	O	NR
dc.rights.uri	Link para a Política ou Licença de Direito Autoral	Em implementação	Pós <i>Stricto Sensu</i> – Cadastro de Coleções RIUT	RIUT	P	NR
dc.date.issued	Data de publicação ou defesa do documento descrito	Data da banca	Pós <i>Stricto Sensu</i> – Marcação Defesa	Aluno/ Secretaria	O	NR
dc.creator	Nome do(s) autor(es) do trabalho	Aluno	Corporativo – Cadastro do Aluno	Aluno/ Secretaria	O	R
dc.creator.Lattes	URL do Currículo Lattes do autor	Implementado	Acadêmico – Cadastro de Lattes e ORCID	Aluno/ Secretaria	O	NR
dc.creator.ID	ORCID do autor	Implementado	Acadêmico – Marcação da Defesa	Aluno/ Secretaria	P	NR

Elemento Dublin Core	Rótulo no RIUT	Campo nos Sistemas Corporativos da UTFPR	Sistema de origem	Responsável pela alimentação	Aplicação	R/NR
dc.contributor.advisor1	Orientador	Orientador	Corporativo – Cadastro de Pessoa	Secretaria	O	NR
dc.contributor.advisor1Lattes	URL do Currículo Lattes do orientador	Implementado	Acadêmico – Cadastro de Lattes e ORCID	Professor/Secretaria	O	NR
dc.contributor.advisor1ID	ORCID do orientador	Implementado	Acadêmico – Cadastro de Lattes e ORCID	Professor/Secretaria	O	NR
dc.contributor.advisor-co1	Coorientador	Coorientador	Corporativo – Cadastro de Pessoa	Secretaria	P	R
dc.contributor.advisor-co1Lattes	URL do Currículo Lattes do coorientador	Implementado	Acadêmico – Cadastro de Lattes e ORCID	Professor/Secretaria	P	R
dc.contributor.advisor-co1ID	ORCID do coorientador	Implementado	Acadêmico – Cadastro de Lattes e ORCID	Professor/Secretaria	P	R
dc.contributor.referee1	Primeiro membro da banca	Membros da banca	Pós <i>Stricto Sensu</i> – Marcação da Defesa	Secretaria	O	NR
dc.contributor.referee1Lattes	URL do Currículo Lattes do primeiro membro da banca	Implementado	Acadêmico – Cadastro de Lattes e ORCID	Professor/Secretaria	O	NR
dc.contributor.referee1ID	ORCID do primeiro membro da banca	Implementado	Acadêmico – Cadastro de Lattes e ORCID	Professor/Secretaria	O	NR
dc.contributor.referee2	Segundo membro da banca	Membros da banca	Pós <i>Stricto Sensu</i> – Marcação da Defesa	Secretaria	O	NR
dc.contributor.referee2Lattes	URL do Currículo Lattes do segundo membro da banca	Implementado	Acadêmico – Cadastro de Lattes e ORCID	Professor/Secretaria	O	NR
dc.contributor.referee2ID	ORCID do segundo membro da banca	Implementado	Acadêmico – Cadastro de Lattes e ORCID	Professor/Secretaria	O	NR
dc.contributor.referee3	Terceiro membro da banca	Membros da banca	Pós <i>Stricto Sensu</i> – Marcação da Defesa	Secretaria	O	NR
dc.contributor.referee3Lattes	URL do Currículo Lattes do terceiro membro da banca	Implementado	Acadêmico – Cadastro de Lattes e ORCID	Professor/Secretaria	O	NR
dc.contributor.referee3ID	ORCID do terceiro membro da banca	Implementado	Acadêmico – Cadastro de Lattes e ORCID	Professor/Secretaria	O	NR
dc.contributor.referee4	Quarto membro da banca	Membros da banca	Pós <i>Stricto Sensu</i> – Marcação da Defesa	Secretaria	P	NR
dc.contributor.referee4Lattes	URL do Currículo Lattes do quarto membro da banca	Implementado	Acadêmico – Cadastro de Lattes e ORCID	Professor/Secretaria	P	NR
dc.contributor.referee4ID	ORCID do quarto membro da banca	Implementado	Acadêmico – Cadastro de Lattes e ORCID	Professor/Secretaria	P	NR
dc.contributor.referee5	Quarto membro da banca	Membros da banca	Pós <i>Stricto Sensu</i> – Marcação da Defesa	Secretaria	P	NR
dc.contributor.referee5Lattes	URL do Currículo Lattes do quinto membro da banca	Implementado	Acadêmico – Cadastro de Lattes e ORCID	Professor/Secretaria	P	NR

Elemento Dublin Core	Rótulo no RIUT	Campo nos Sistemas Corporativos da UTFPR	Sistema de origem	Responsável pela alimentação	Aplicação	R/NR
dc.contributor.refereeSID	ORCID do quinto membro da banca	Implementado	Acadêmico – Cadastro de Lattes e ORCID	Professor/Secretaria	P	NR
dc.description	Notas			Bibliotecário	P	R
dc.publisher	Nome da instituição por extenso	Padrão	Corporativo – Cadastro da Instituição	DESI	O	NR
dc.publisher.initials	Sigla da instituição	Padrão	Corporativo – Cadastro da Instituição	DESI	O	NR
utfpr.publisher.local	Câmpus	Padrão	Corporativo – Cadastro da Instituição	PROPPG	O	R
dc.publisher.program	Programa de Pós-Graduação ou Curso	Programa	Pós <i>Stricto Sensu</i> – Cadastro de Programa	PROPPG	O	NR
dc.publisher.country	País da Publicação	Padrão	Corporativo – Cadastro da Instituição	DESI	O	NR
utfpr.degree.local	Local de Defesa	Câmpus	Pós <i>Stricto Sensu</i> – Marcação da Defesa	Secretaria	O	NR
dc.language	Idioma			Bibliotecário	O	R
dc.subject	Palavras-chave			Bibliotecário	O	R
dc.subject.cnpq	Áreas de conhecimento de acordo com tabela do CNPq			Bibliotecário	O	R
utfpr.subject.capes	Áreas de conhecimento de acordo com tabela da Capes	Código Capes	Pós <i>Stricto Sensu</i> – Cadastro de Programa	PROPPG	O	R
dc.description.resumo	Resumo	Em implementação	Pós <i>Stricto Sensu</i> – Marcação da Defesa	Aluno	P	NR
dc.description.abstract	Abstract	Em implementação	Pós <i>Stricto Sensu</i> – Marcação da Defesa	Aluno	P	NR
dc.description.sponsorship	Agência de fomento			Bibliotecário	P	NR
dc.relation	Relação			Bibliotecário	P	R
dc.identifier.citation	Referência	Processamento realizado na conversão dos dados	Variáveis provenientes de diferentes campos processados para a integração	DESI	O	NR

Nota: O - obrigatório; P - obrigatório quando presente no documento; R - repetível; NR - não repetível.

Fonte: Adaptado da Política de Submissão do RIUT.

Destaca-se que os metadados que não possuem informações nas colunas Campo nos Sistemas Corporativos da UTFPR e Sistema de origem não estão disponíveis nos Sistemas Corporativos da UTFPR e, portanto, são de alimentação manual realizada pelos bibliotecários membros do Comitê Gestor do RIUT, a eles também compete a análise e adequação dos dados provenientes dos Sistemas Corporativos da UTFPR e integrados de forma semiautomática ao

RIUT após conversão e processamento computacional.

Considerando o perfil de aplicação disponível no Quadro 1, para a representação das versões finais no RIUT, são utilizados 46 campos de metadados para a representação de dissertações e teses, dos quais 30 são obrigatórios (O) e 16 são obrigatórios quando presentes no documento (P).

Considerando apenas os campos obrigatórios (30), 15 deles, que correspondem a 50% dos dados obrigatórios e a 32,61% do total dos dados necessários para a representação de dissertações e teses no RIUT (46), conforme o perfil de aplicação de dissertações e teses, já estavam disponíveis nos Sistemas Corporativos da UTFPR e 10 (21,74%) foram implementados. E, considerando a totalidade dos metadados utilizados para a representação (46), 18 (39,13%) já estavam disponíveis e 17 (36,96%) foram implementados.

Assim, considerando os dados já existentes ou implementados, a partir da integração, do total de 46 campos de metadados presentes no perfil de aplicação, 35 (76,09%) são alimentados de forma semiautomática no RIUT, tornando necessária a entrada manual de dados de apenas 11 (23,91%) campos.

Ademais, ressalta-se que, conforme consta no Quadro 1, há 5 campos de metadados necessários à representação de dissertações e teses no RIUT, que estão em implantação nos Sistemas Corporativos da UTFPR. Desta forma, a partir desta disponibilização, do total de 46 campos utilizados para a representação, 40 (86,96%) terão alimentação semiautomática, restando apenas 6 (13,04%) para representação manual, sendo estes: notas, idioma, palavras-chave, áreas de conhecimento de acordo com tabela do CNPq, agência de fomento e relação.

Vale destacar que dos 46 campos de metadados utilizados para a representação de dissertações e teses apresentados no perfil de aplicação do RIUT, 12 (26,09%) são repetíveis e 34 (73,91%) são não repetíveis.

São considerados repetíveis aqueles metadados que, por suas características e considerando as especificidades do objeto digital representado, podem ser utilizados mais de uma vez. Desta forma, na representação das dissertações e teses no RIUT, por exemplo, o campo assunto (dc.subject) é repetível e, conforme a Política de Submissão do RIUT, deve ser utilizado a quantidade mínima de 3 assuntos em português e seus correspondentes em

inglês. Portanto a quantidade de campos utilizada na representação destes materiais é de no mínimo 35 campos, caso sejam utilizados apenas os obrigatórios (O), e no mínimo 51 campos, quando utilizados também os obrigatórios quando presentes no documento (P).

Como dados já disponíveis nos Sistemas Corporativos da UTFPR tem-se: tipo de documento, autor, orientador, coorientador, membros da banca de defesa, título, data de defesa, nome do PPG, instituição de defesa, câmpus de vínculo, país, local de defesa, área de conhecimento Capes. Alguns desses campos foram inseridos como variáveis para que a referência do documento seja elaborada no padrão da Associação Brasileira de Normas Técnicas (ABNT). Ainda estão em implantação os campos: título alternativo, permissão de acesso, link para a Política ou Licença de Direito Autoral, resumo e abstract.

Como necessidades de tratamento computacional tem-se: a adoção de maiúsculas e minúsculas para os campos preenchidos inadequadamente; o tratamento dos nomes de pessoas para o formato de referência, considerando inclusive os agnomes e os identificativos de geração; o tratamento para que o tipo de curso reflita a tipologia do objeto; e a permissão de acesso, cuja padronização está atrelada à solicitação de restrição de acesso ao trabalho.

A partir do estudo do perfil de aplicação, verificou-se que as *Uniform Resource Locator* (URL) do Currículo Lattes e do *Open Researcher and Contributor ID* (ORCID ID) dos pesquisadores poderiam ser implementadas nos Sistemas Corporativos da UTFPR, considerando sua relevância para outros processos Institucionais.

Nesse sentido, houve a alimentação automática das URL do Currículo Lattes da maioria dos docentes da Instituição a partir de dados já disponíveis na PROPPG, de forma que apenas aqueles que não tenham esse dado armazenado devem incluí-lo manualmente nos Sistemas Corporativos da UTFPR, considerando que esse passou a ser um dado obrigatório na vinculação do docente a uma banca de defesa. Quanto ao ORCID ID, o campo foi incluído como não obrigatório, dada a necessidade de ampliar a conscientização de uso e as políticas de adoção pela comunidade da UTFPR.

Além disso, foi implementado um campo, no cadastro do PPG, para registrar a URL da coleção do curso no RIUT, visto que, no processamento, o

registro deve ser direcionado a uma coleção específica no repositório. Inicialmente houve uma carga automática dos dados disponíveis no RIUT e, quando da criação de novos cursos/PPG, o sistema Pós-Graduação *Stricto Sensu* notifica o RIUT por e-mail automático para que a coleção seja criada e o campo alimentado nos Sistemas Corporativos da UTFPR. Além deste, foi implementado um campo para indicar o link de acesso à licença *Creative Commons* regulamentada pelo PPG, que possibilita a gestão dessa informação e o envio desse dado para o RIUT.

O processamento dos dados tem como variável de consulta o nome do aluno e seu vínculo com um PPG específico, tendo em vista ser essa a forma utilizada pelos Sistemas Corporativos da UTFPR.

Considerando o fluxo apresentado na Figura 1, após o parecer favorável da Biblioteca, os dados são gerados pelos Sistemas Corporativos da UTFPR, de acordo com o tratamento computacional adequado, e armazenados no servidor do RIUT, formando uma coleção diária, na qual há uma subpasta para cada registro, contendo: objeto digital, metadados em *Dublin Core XML*, arquivo em texto puro contendo a coleção à qual o registro deverá ser armazenado no RIUT e texto puro contendo o nome do arquivo do objeto digital, para vinculá-lo ao registro. Esse processo, embora gere os dados nos Sistemas Corporativos e os armazene no servidor do RIUT, ainda não os armazena na base de dados do RIUT.

O povoamento consiste na inserção da coleção diária no RIUT por linha de comando, com processamento por *script* como rotina do servidor, com fluxo diário de integração dos dados. O depósito é monitorado por meio de *logs* disponibilizados nos Sistemas Corporativos da UTFPR, para que seja assegurada a inserção dos registros e a credibilidade ao processo.

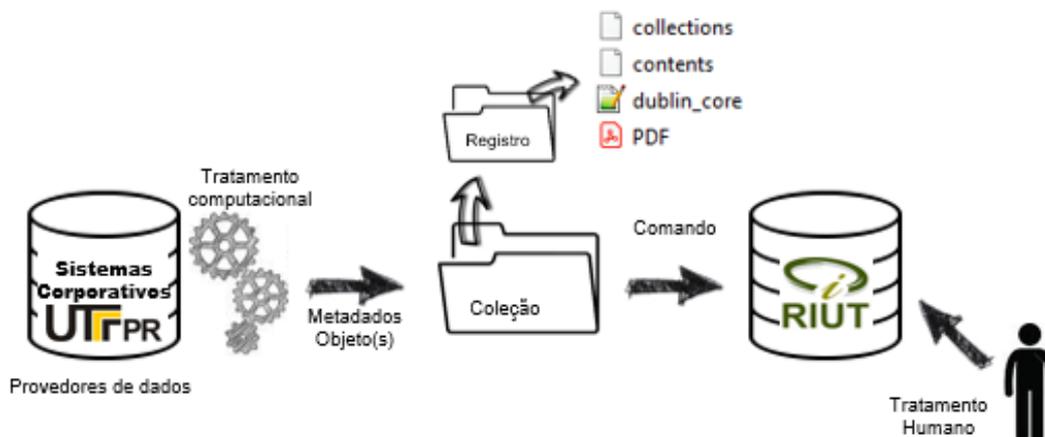
Vale elucidar que, como visualizado na Figura 1, além do processamento para o reuso dos dados, foi realizada uma implementação que permite que o orientador do trabalho realize o *upload* do arquivo da versão final da dissertação ou tese, bem como de arquivos adicionais que porventura a acompanhem, como, por exemplo, os produtos educacionais e tecnológicos requeridos por alguns PPG profissionais ou por áreas específicas. Embora o *upload* seja realizado no sistema Pós-Graduação *Stricto Sensu*, quando do processamento, estes objetos

digitais são parte da coleção criada e, após integrados, são mantidos apenas no RIUT, evitando a duplicação de armazenamento.

A Figura 2 representa graficamente o processo de integração entre os Sistemas Corporativos da UTFPR (provedores de dados) e o RIUT e nele aparecem duas etapas de tratamento, computacional e humano.

O tratamento computacional é realizado nos Sistemas Corporativos da UTFPR no momento em que o registro é enviado para a pasta do servidor do RIUT e consiste nas adequações anteriormente descritas, bem como na conversão da tabela de dados em formato XML, utilizando o vocabulário *Dublin Core* e considerando o perfil de aplicação do RIUT. Já o tratamento humano é realizado pelos bibliotecários membros do Comitê Gestor do RIUT, por meio da interface do RIUT e considerando a Política de Submissão do RIUT para a verificação da adequação dos dados ao(s) objeto(s) digital(is) e a entrada manual dos campos não alimentados de forma semiautomática, que correspondem atualmente a 23,91% e em breve será de apenas 13,04%.

Figura 2 - Processo de integração entre os sistemas



Fonte: Autoria própria.

Como descrito anteriormente, para a integração com o RIUT, os Sistemas Corporativos da UTFPR são os provedores de dados, constituídos por 13 sistemas e suas bases de dados que interagem entre si, por isso consideradas na representação da Figura 2 apenas como Provedores de dados. Os Sistemas Corporativos da UTFPR e o RIUT estão alocados em diferentes servidores, sendo que o primeiro sistema processa os dados e forma a coleção em uma

pasta compartilhada no servidor do segundo, contudo o povoamento é realizado na rotina diária, cujo processamento ocorre por *script*.

Para auxiliar na implantação do processo, conforme fluxo explicitado na Figura 1, foram elaborados tutoriais para apoiar os atores dos diferentes perfis atrelados à integração: orientador, coordenador e bibliotecário. Além disso, as áreas envolvidas: PROPPG, RIUT e DESIS forneceram orientações específicas para cada perfil, considerando a rotina realizada.

A partir da integração dos Sistemas Corporativos da UTFPR com o RIUT, tendo em vista que a entrega da versão final da dissertação ou tese é um requisito para a emissão do diploma, é possível assegurar que a totalidade dos alunos diplomados em cursos de Pós-Graduação *Stricto Sensu* da UTFPR terão a versão final de sua dissertação e/ou tese disponível no Repositório Institucional. Por sua vez, o RIUT atua como ambiente informacional de gestão, disseminação, ampliação da visibilidade desses trabalhos em mecanismos de busca web, bases de dados e diretórios nacionais e internacionais, preservação da memória da pesquisa realizada na Universidade, bem como utilizado para a preservação digital. Além disso é possível disponibilizar os conhecimentos gerados pelos PPG por meio das versões finais, para acesso público e gratuito à sociedade e à comunidade científica.

Diante do exposto, vale ressaltar que o processo descrito anteriormente teve aporte da Arquitetura da Informação, com o objetivo de reuso dos dados visando ao povoamento semiautomático de dissertações e teses no RIUT.

Como entregáveis da Arquitetura da Informação destacam-se o fluxograma de entrega de dissertações e teses, o Perfil de Aplicação D/T do RIUT com a compatibilização dos dados advindos de diferentes camadas de dados dos Sistemas Corporativos da UTFPR e o modelo do processo de integração entre os sistemas.

5 CONSIDERAÇÕES FINAIS

O povoamento de dissertações e teses em repositórios institucionais geralmente ocorre por autoarquivamento ou povoamento mediado por

bibliotecários. Contudo o reuso de dados já existentes nos sistemas disponíveis na Instituição gera benefícios múltiplos, dentre os quais destacam-se: agilidade, dados preenchidos automaticamente, veracidade dos dados, desoneração do aluno e dos bibliotecários, preservação da memória e ampliação da visibilidade da produção acadêmico-científica.

Além destes e considerando os aportes da Arquitetura da Informação é possível a análise dos benefícios para a tríade conteúdo (dissertações e teses, bem como os demais objetos digitais que as acompanhem), contexto (integração entre Sistemas Corporativos da UTFPR e Repositório Institucional da UTFPR) e usuário (aluno, orientador, coordenador de PPG, bibliotecário, comunidade científica e geral e a própria Capes).

A análise das camadas de dados dos Sistemas Corporativos da UTFPR possibilitou identificar a presença de 39,13% dos dados utilizados para a representação de dissertação ou tese no Repositório Institucional (RIUT) e a análise do perfil de aplicação do RIUT possibilitou implementar mais 36,96%, que consistem em dados de interesse institucional pela PROPPG, totalizando 76,09% dos dados utilizados para a representação dos trabalhos provenientes de 3 sistemas disponíveis nos Sistemas Corporativos da UTFPR: Acadêmico, Corporativo e Pós-Graduação *Stricto Sensu*.

Destaca-se que estão em implantação nos Sistemas Corporativos mais 5 campos, o que possibilitará ampliar o reuso de dados, de 76,09% para 86,96% dos campos alimentados de forma semiautomática, restando 13,04% para entrada manual.

Desta forma, por meio das tecnologias de informática e do reuso de dados, tornou-se possível automatizar tarefas e melhorar processos, o que permitiu agilidade e precisão na disponibilização das dissertações e teses no RIUT, além da redução do tempo e dos esforços humanos anteriormente despendidos.

A partir da integração entre os Sistemas Corporativos da UTFPR e o RIUT é possível assegurar que a totalidade dos alunos diplomados em cursos de Pós-Graduação *Stricto Sensu* da UTFPR terão as versões finais de suas dissertações e teses disponibilizadas no Repositório Institucional.

Esta experiência possibilita um novo olhar para a análise de dados

disponíveis nas diferentes fontes já existentes na Universidade. Esses dados podem ser reutilizados em diferentes finalidades, como: utilização em ferramentas de gestão, disponibilização de dados abertos, alimentação automática de conteúdos em *websites*, integração com outros ambientes internos ou externos à Instituição, constituição de sistemas de informação de pesquisa (RIM/CRIS).

Além disso e considerando que a UTFPR disponibiliza a versão final de trabalhos de conclusão de cursos de graduação (TCCs) em Repositório Institucional, já está em curso a análise da camada de dados dos Sistemas Corporativos da UTFPR, a fim de possibilitar a integração para o povoamento semiautomático desses conteúdos, o que possibilitará, de igual maneira, inúmeros benefícios ao conteúdo, contexto e usuários.

Com o avanço das políticas de gestão, a Instrução Normativa Conjunta nº 01/2018 foi revogada por meio da IN 10/2020 PROPPG (UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ, 2020b) e da IN nº 1/2021 PROGRAD (UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ, 2021), que estabelecem a entrega das versões finais de trabalhos de conclusão de cursos de graduação, dissertações e teses apenas em formato digital pelos Sistemas Corporativos da UTFPR; reforçam o RIUT como ambiente institucional exclusivo para a disponibilização desses objetos digitais e estabelecem o fluxo de entrega e depósito no RIUT, considerando a arquitetura apresentada no presente estudo.

Estas implementações possibilitam a utilização do Repositório Institucional como uma ferramenta de gestão e publicização das versões finais dos trabalhos defendidos na UTFPR, tornando-o relevante no processo de gestão da pesquisa e um ambiente centralizado capaz de prover dados dessa natureza para mecanismos de busca gerais e acadêmicos, diretórios de coleta de dados e na integração futura vislumbrada em cenário nacional com ambientes como ORCID, Plataformas Lattes e Sucupira.

Além disso, a disponibilização pública dos trabalhos defendidos permite a aproximação das pesquisas geradas na UTFPR com a comunidade científica e com a sociedade, fator importante para a avaliação dos PPG.

Espera-se que este estudo auxilie pesquisadores e profissionais no desenvolvimento de um novo olhar para a análise de dados disponíveis em

fontes internas e externas às instituições, visando ao reuso para diferentes finalidades, dentre elas, destaca-se neste estudo o povoamento semiautomático de repositórios digitais.

AGRADECIMENTOS

Agradecemos aos servidores da UTFPR vinculados ao DESIS e à PROPPG envolvidos diretamente nas diferentes etapas de implementação do povoamento semiautomático de dissertações e teses no RIUT. Agradecemos também aos bibliotecários membros do Comitê Gestor do RIUT e aos professores orientadores e coordenadores de PPG que nos forneceram *feedbacks* que possibilitaram consolidar o processo.

REFERÊNCIAS

BRASIL. Ministério da Educação. Fundação Coordenação de Aperfeiçoamento de Pessoal de Nível Superior. Portaria nº 182, de 14 de agosto de 2018. **Diário Oficial da União**: seção 1, Brasília, DF, n. 158, p. 11, 16 ago. 2018. Disponível em:

<http://pesquisa.in.gov.br/imprensa/jsp/visualiza/index.jsp?data=16/08/2018&jornal=515&pagina=11>. Acesso em: 06 fev. 2019.

BRASIL. Ministério da Educação. Fundação Coordenação de Aperfeiçoamento de Pessoal de Nível Superior. Portaria nº 13, de 15 de fevereiro de 2006.

Diário Oficial da União: seção 1, Brasília, DF, n. 35, p. 15, 17 fev. 2006.

Disponível em:

https://abmes.org.br/arquivos/legislacoes/Port_CAPES_13_20060215.pdf.

Acesso em: 06 fev. 2019.

BRASIL. Ministério da Educação. Fundação Coordenação de Aperfeiçoamento de Pessoal de Nível Superior. **Ficha de Avaliação – Programas Acadêmicos e Profissionais – Área 46/Ensino**. Dimensão 3, “Impacto na Sociedade”, 22/07/2020. Disponível em: <https://www.gov.br/capes/pt-br/aceso-a-informacao/acoes-e-programas/avaliacao/sobre-a-avaliacao/areas-avaliacao/sobre-as-areas-de-avaliacao/colégio-de-ciencias-exatas-tecnologicas-e-multidisciplinar/multidisciplinar/ensino> Acesso em: 11 mar. 2021.

CAMARGO, L. S. A.; VIDOTTI, S. A. B. G. **Arquitetura da informação**: uma abordagem prática para o tratamento de conteúdo e interface em ambientes informacionais digitais. Rio de Janeiro: LTC, 2011.

MACEDO, F. L. O. de. **Arquitetura da informação**: aspectos epistemológicos, científicos e práticos. 2005. Dissertação (Mestrado em Ciência da Informação) – Universidade de Brasília, Brasília, 2005. Disponível em: https://www.academia.edu/2504826/Arquitetura_da_Informa%C3%A7%C3%A3o_o_aspectos_epistemol%C3%B3gicos_cient%C3%ADficos_e_pr%C3%A1ticos. Acesso em: 06 fev. 2019.

MORVILLE, P.; ROSENFELD, L. **Information architecture for the word wide web**. Sebastopol, CA: O'Reilly, 2006.

OLIVEIRA, H. P. C. de. **Arquitetura da informação pervasiva**: contribuições conceituais. 2014. Tese (Doutorado em Ciência da Informação) – Universidade Estadual Paulista, Marília, 2014. Disponível em: <https://repositorio.unesp.br/handle/11449/110387>. Acesso em: 06 fev. 2019.

OLIVEIRA, H. P. C. de; VIDOTTI, S. A. B. G. Dos ambientes informacionais às ecologias informacionais complexas. **Informação & Sociedade: Estudos**, João Pessoa, v. 26, n. 1, p. 91-101, jan./abr. 2016. Disponível em: <http://www.periodicos.ufpb.br/index.php/ies/article/view/29438/16509>. Acesso em: 06 fev. 2019.

RESMINI, A.; ROSATI, L. A brief history of information architecture. **Journal of Information Architecture**, v. 3, n. 2, p. 33-45, 2011. Disponível em: <https://journalofia.org/volume3/issue2/03-resmini/jofia-0302-03-resmini.pdf>. Acesso em: 16 fev. 2019.

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ. Conselho de Graduação e Educação Profissional; Conselho de Pesquisa e Pós-Graduação. **Resolução Conjunta nº 01/2020 – COGEP-COPPG**. 2020a. Disponível em: https://sei.utfpr.edu.br/sei/publicacoes/controlador_publicacoes.php?acao=publicacao_visualizar&id_documento=1811618&id_orgao_publicacao=0. Acesso em: 22 out. 2020.

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ. Pró-Reitoria de Graduação e Educação Profissional. **Instrução Normativa PROGRAD nº 1/2021**. 2021. Disponível em: https://sei.utfpr.edu.br/sei/publicacoes/controlador_publicacoes.php?acao=publicacao_visualizar&id_documento=2042165&id_orgao_publicacao=0. Acesso em: 13 fev. 2021.

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ. Pró-Reitoria de Pesquisa e Pós-Graduação. **Instrução Normativa Conjunta nº 01/2018 – PROPPG/PROGRAD**. 2018. Disponível em: <http://portal.utfpr.edu.br/documentos/pesquisa-e-pos-graduacao/proppg/instrucoes-normativas-conjuntas/in-conjunta-proppg-prograd-01-2018.pdf/view>. Acesso em: 06 fev. 2019.

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ. Pró-Reitoria de Pesquisa e Pós-Graduação. **Instrução Normativa Conjunta nº 01/2011 – PROGRAD/ PROPPG**. 2011. Disponível em:

<http://portal.utfpr.edu.br/documentos/pesquisa-e-pos-graduacao/proppg/instrucoes-normativas-conjuntas/instrucao-normativa-conjunta-01-2011-prograd-proppg/view>. Acesso em: 06 fev. 2020.

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ. Pró-Reitoria de Pesquisa e Pós-Graduação. **Instrução Normativa nº 10/2020 – PROPPG**. 2020b. Disponível em:

https://sei.utfpr.edu.br/sei/publicacoes/controlador_publicacoes.php?acao=publicacao_visualizar&id_documento=1840028&id_orgao_publicacao=0. Acesso em: 22 out. 2020.

VECHIATO, F. L.; VIDOTTI, S. A. B. G. **Encontrabilidade da informação**. São Paulo: Cultura Acadêmica, 2014. Disponível em:

<http://hdl.handle.net/11449/126218>. Acesso em: 06 fev. 2019.

VIDOTTI, S. A. B. G.; CUSIN, C. A.; CORRADI, J. A. M. Acessibilidade digital sob o prisma da arquitetura da informação. In: GUIMARÃES, J. A. C.; FUJITA, M. S. L. **Ensino e pesquisa em Biblioteconomia no Brasil: a emergência de um novo olhar**. São Paulo: Cultura Acadêmica, 2008.

WURMAN, R. S. **Ansiedade de informação**. São Paulo: Editora de Cultura, 2005.

WURMAN, R. S. **Information architects**. Zurich: Graphis Press Corp, 1996.

DATA INTEGRATION Y REUSE FOR THE SEMI-AUTOMATIC ARCHIVING OF DISSERTATIONS AND THESIS IN THE UTFPR INSTITUTIONAL REPOSITORY

ABSTRACT

Objective: Describes the process used for data integration and reuse for the semi-automatic archiving of dissertations and thesis in the UTFPR Institutional Repository (RIUT), from the UTFPR Corporate Systems. **Methodology:** Conducts a case study at the Universidade Tecnológica Federal do Paraná (UTFPR), and the participant observation as a qualitative research method for analysis of data available in the corporate information systems to be integrated. **Results:** Based on the Information Architecture, it presents the structuring of the dissertation and thesis delivery flow and the study of the data layer of the UTFPR Corporate Systems, the compatibility and conversion to the RIUT application profile, aiming at the archiving semia-automatic of dissertation and thesis in institutional repository. These results made it possible to develop the process of data integration and reuse between UTFPR Corporate Systems and the Institutional Repository. **Conclusions:** The reuse of data already existing in the Corporate Systems for the semi-automatic archiving of dissertations and thesis in the

Institutional Repository generates multiple benefits, among which stand out: agility and accuracy, truthfulness of data, exemption from the student and librarians, preserving the memory and increasing the visibility of academic, scientific and technological production, with a view to promoting the impact and social insertion of these results, in addition to reducing the human efforts previously expended for the activity.

Descriptors: Institutional repositories. Dissertations and thesis. Data integration. Data reuse.

INTEGRACIÓN Y REUTILIZACIÓN DE DATOS PARA POBLAMIENTO SEMIAUTOMÁTICO DE DISERTACIONES Y TESIS EN EL REPOSITORIO INSTITUCIONAL DE UTFPR

RESUMEN

Objetivo: Describe el proceso utilizado para la integración y reutilización de datos para la población semiautomática de disertaciones y tesis en el Repositorio Institucional de la UTFPR (RIUT), desde los Sistemas Corporativos de la UTFPR. **Metodología:** Realiza un estudio de caso en la Universidade Tecnológica Federal de Paraná (UTFPR), y la observación participante como método de investigación cualitativa para el análisis de los datos disponibles en los sistemas de información corporativos a integrar. **Resultados:** Presenta, a partir de la Arquitectura de la Información, la estructuración del flujo de entrega de disertaciones y tesis y, el estudio de la capa de datos de los Sistemas Corporativos de UTFPR, la compatibilidad y conversión al perfil de aplicación del RIUT, con el objetivo de la población semiautomática de disertaciones y tesis en el repositorio institucional. Estos resultados permitieron desarrollar el proceso de integración y reutilización de datos entre los Sistemas Corporativos de la UTFPR y el Repositorio Institucional. **Conclusiones:** La reutilización de los datos ya existentes en los Sistemas Corporativos para la población semiautomática de disertaciones y tesis en el Repositorio Institucional genera múltiples beneficios, entre los que destacamos: agilidad y precisión, veracidad de los datos, aliviar al estudiante y bibliotecarios, preservación de la memoria y ampliación de la visibilidad de la producción académica, científica y tecnológica, con vistas a promover el impacto y la inserción social de estos resultados, además de reducir los esfuerzos humanos que antes se dedicaban a la actividad.

Descriptor: Repositorios institucionales. Disertaciones y tesis. Integración de datos. Reutilización de datos.

i <https://sistemas.utfpr.edu.br/>

ii <http://repositorio.utfpr.edu.br/jspui/>

Recebido em: 06.05.2020

Aceito em: 22.03.2021