

# Proposta de um Repositório Digital para Transparência de Dados Pessoais

Larissa Amadeu Teixeira<sup>1</sup>, Maria Luisa Moreira de Souza<sup>1</sup>, Thiago Adriano Coleti<sup>1</sup>, André Menolli<sup>1</sup>, Marcelo Morandini<sup>2</sup>, Renato Balancieri<sup>3</sup>

<sup>1</sup>Universidade Estadual do Norte do Paraná (UENP) – Bandeirantes – PR – Brasil

<sup>2</sup>Universidade de São Paulo (EACH-USP) – São Paulo – SP - Brasil

<sup>3</sup>Universidade Estadual do Paraná (UNESPAR) – Apucarana – PR – Brasil

lamadeu2012@hotmail.com, malumsouza2104@gmail.com, thiago.coleti@uenp.eu.br  
menolli@uenp.edu.br, m.morandini@usp.br, renato.balancieri@unespar.edu.br

**Abstract.** *This paper presents a repository of information proposal that concerns the usage of personal data. It aims to support Transparency forced by Brazilian General Data Protection Regulation (LGPD) providing resources to simplify tasks for data controllers who need publishing information about their software projects that involve the processing of personal data. It also aims to assist data subjects in their searches for information regarding the agents and events associated with the usage of their personal data. With the repository, we assumed that is possible to avoid individual and/or complex searches within privacy policies and/or hidden interfaces within the tools. This upcoming release, in the form of a web tool, will be assessed and validated with controllers and data subjects in future works.*

**Resumo.** *Este artigo apresenta a proposta de um repositório de dados sobre a manipulação de dados pessoais. O objetivo é apoiar a Transparência no uso dos dados pessoais, que é uma exigência das regulamentações como a Lei Geral de Proteção de Dados (LGPD) com recursos para simplificar as ações dos controladores que precisam divulgarem informações sobre seus projetos de software que manipulam dados pessoais. Também visa apoiar os titulares dos dados no processo de busca de informações sobre agentes e eventos envolvidos na manipulação de seus dados. Espera-se, com o repositório, evitar uma busca individualizada em documentos ou recursos pouco acessíveis na aplicação. A versão inicial, uma plataforma web, será avaliada e validada com controladores e titulares dos dados.*

## 1. Introdução

A Transparência de Dados Pessoais (Transparência) é um requisito de qualidade que indica o grau no qual uma aplicação de software fornece informações sobre agentes e eventos envolvidos na manipulação dos dados pessoais de seus usuários [Filgueiras et al. 2019]. A Transparência também é uma exigência na Lei Geral de Proteção de Dados<sup>1</sup> (LGPD), do Brasil, em seu Artigo 6º, inciso VI, que descreve a

---

<sup>1</sup>[https://www.planalto.gov.br/ccivil.03/\\_ato20152018/2018/lei/L13709compilado.htm](https://www.planalto.gov.br/ccivil.03/_ato20152018/2018/lei/L13709compilado.htm)

Transparência como *garantia, aos titulares, de informações claras, precisas e facilmente acessíveis sobre a realização do tratamento e os respectivos agentes de tratamento, observados os segredos comercial e industrial*. Na *General Data Protection Regulation (GDPR)*<sup>2</sup>, em vigor na União Europeia, as diretrizes descritas nos Artigos 12, 13 e 14, que destacam a necessidade dos controladores<sup>3</sup> disponibilizarem meios apropriados para prover informações sobre manipulação dos dados pessoais.

Um dos principais desafios da Transparência está na disponibilidade das informações coerentes e sumarizadas aos titulares dos dados<sup>4</sup>, pois as mesmas costumam integrar as Políticas de Privacidade e Segurança (PPS), com textos longos e complexos. Estratégias como ícones para privacidade e transparência [Holtz et al. 2011]; melhorias na legibilidade das PPS [Audich et al. 2021]; *plugins* e *frameworks* para navegadores [Murmam and Fischer-Hübner 2017]; e modelos conceituais e diretrizes para Transparência podem ser encontrados na literatura [Coleti et al. 2020]. Entretanto, para uso dessas estratégias, os titulares dos dados devem acessar e procurar as informações diretamente nas aplicações de software, ocasionando um processo ineficiente.

Com o objetivo de melhorar a divulgação das informações sobre a manipulação dos dados pessoais pelos controladores e o acesso às mesmas pelos titulares dos dados, este artigo apresenta a proposta da criação de um repositório digital para Transparência de Dados Pessoais. A proposta ampara-se nas necessidades das empresas serem transparentes quanto aos seus processos, atendendo assim a LGPD (ou GDPR); e no direito das pessoas acessarem informações sobre agentes e eventos envolvidos na manipulação de seus dados pessoais, de forma simples, objetiva e perceptível, para fiscalizar a manipulação de seus dados. O desenvolvimento do repositório se dará por um processo que contempla as ações de definição das estratégias de uso, seleção do modelo de metadados, definição de requisitos funcionais da aplicação e desenvolvimento de uma plataforma web.

O artigo está estruturado da seguinte forma: a Seção 2 apresenta uma visão geral sobre repositório digital de dados. A Seção 3, descreve as etapas do processo de concepção e desenvolvimento do repositório. Por fim, a Seção 4 apresenta as considerações finais da proposta e os trabalhos futuros.

## 2. Repositório Digital de Dados

Um repositório digital é um ambiente computacional no qual as informações são inseridas, armazenadas, descritas e distribuídas para usuários que desejam utilizá-las [Casagrande et al. 2015]. Os repositórios permitem a construção de metadados sobre o domínio de aplicação que descrevem a informação apresentada e, quando necessário, como utilizá-la [Freund et al. 2019].

No trabalho de [Sayão et al. 2018] são apresentadas diversas características que um repositório digital deve ter e as preocupações com sua implementação, que contemplam: escolha da plataforma de software, a construção de mecanismos de busca e acesso às informações, aspectos de usabilidade e experiência do usuário, definição de um fluxo de trabalho para a entrada, gerenciamento e acesso às informações. [Sayão et al. 2018]

---

<sup>2</sup><https://gdpr-info.eu/>

<sup>3</sup>É o responsável por determinar as finalidades e os meios de tratamento de dados pessoais.

<sup>4</sup>Pessoa física a quem se referem os dados pessoais.

destacam também a necessidade de considerar aspectos de domínio, contexto e usuário e não somente a construção de um banco de dados.

Os repositórios têm ganhado importância em diversas áreas do conhecimento em virtude das possibilidades de divulgação de informações e conhecimentos sobre domínios de aplicação. Na ciência, a construção e a utilização de repositórios públicos e abertos para divulgação de pesquisas e seus materiais complementares é foco de discussão em uma área conhecida como Ciência Aberta [Azzam and Jacobson 2015]. Já em áreas como a programação, o uso de repositórios de códigos fonte e artefatos de software como o GitHub<sup>5</sup> é uma constante entre desenvolvedores, e o conhecimento deste tipo de recurso costuma ser exigência em processos para contratação de desenvolvedores. Há também a criação de repositórios para domínios específicos a fim de prover conhecimento e/ou compartilhar informações, como é o caso do projeto de [Briganó et al. 2021], que propõe um repositório de conhecimentos em governança em Tecnologia da Informação (TI) para empresários da região Norte do Paraná.

### 3. Concepção e Desenvolvimento do Repositório

O escopo do repositório foi definido como um ambiente para dar suporte à Transparência de Dados Pessoais por meio do armazenamento, gerenciamento, divulgação e acesso às informações sobre a manipulação de dados pessoais. Na sequência serão apresentadas as etapas do processo de concepção e desenvolvimento do repositório.

**Definição das estratégias de uso do repositório:** foi definido como cada perfil de usuário faria o uso da aplicação, sendo elas:

- **Controladores** - utilizarão a ferramenta como mecanismo para divulgar como seus projetos de software manipularão dados pessoais. Essa ação ocorrerá por meio do preenchimento de um formulário, a ser discutido posteriormente neste texto, com informações relacionadas aos métodos de uso e pessoas/empresas que poderão acessar os dados pessoais;
- **Titular dos dados** - utilizarão a ferramenta como mecanismo de busca e acesso facilitado, uma vez que poderão consultar como aplicações de software distintas realizam a manipulação de seus dados pessoais sem a necessidade de buscas individuais e complexas.

**Seleção do modelo de metadados:** a existência de um repositório está vinculada à um modelo de metadados. Para este projeto, foi escolhido o o TR-Model [Coleti et al. 2020]<sup>6</sup>. Esse modelo apresenta um conjunto de metadados que indica as informações mínimas a serem apresentadas para os titulares dos dados. Os metadados são organizados em cinco classes: (a) **Atores** - pessoas que realizam alguma ação na manipulação de dados pessoais, como controladores, destinatários de dados em casos de transferência e compartilhamento, operadores que executam ações computacionais nos dados e agências de controle que fiscalizam as ações de manipulação; (b) **Dados pessoais** - informações sobre os dados pessoais manipulados tais como origem, descrição, autorização de uso e período de utilização; (c) **Transferência e compartilhamento** - informações sobre ações de transferência e compartilhamento entre controladores e operadores tais como garantias legais, justificativa e destinatários; (d) **Propósito de uso** -

---

<sup>5</sup><https://github.com/>

<sup>6</sup>[http://each.usp.br/cond\\_met\\_pand/trmodel/trmodel.php](http://each.usp.br/cond_met_pand/trmodel/trmodel.php)

informações sobre os objetivos da manipulação de dados pelo controladores tais como descrição, base legal, período de execução do propósito; e (e) **Ação e Negociação** - informações sobre como o titular dos dados deve agir para reclamar, questionar ou cancelar as permissões ou ações de manipulação de dados.

**Definição de requisitos funcionais:** nesta etapa, foram definidas as principais funcionalidades para o repositório, inspiradas em repositórios já existentes como o *DataONE* [Allard 2012]; nas exigências de Transparência existentes na LGPD e GDPR; e nas demandas de pessoas por Transparência, conforme identificado durante o desenvolvimento do TR-Model [Coleti et al. 2020]. Os requisitos funcionais definidos foram:

- RQ1 - manutenção (inclusão, edição e exclusão) das informações sobre a manipulação dos dados pessoais pelos controladores;
- RQ2 - acesso e visualização das informações de manipulação pelos titulares dos dados; e
- RQ3 - recurso para personalização das opções de busca de projetos por parte dos titulares dos dados.

Também foi definido o requisito RQ4 - recurso para os titulares dos dados inserirem informações, críticas, dúvidas e denúncias sobre a manipulação dos dados pessoais. Entretanto, o RQ4 será implementado em trabalhos futuros, pois assumiu-se que envolverá aspectos de colaboração, gestão da informação e proveniência de dados.

**Desenvolvimento da plataforma web:** a versão inicial do repositório, que tem seu código disponível no GitHub<sup>7</sup> para fins de colaboração, foi desenvolvida para plataforma Web-Responsiva usando o *framework* Laravel, Linguagem de Programação PHP e Banco de Dados PostgreSQL.

Foi proposto um formulário para procurar levantar mais aspectos acerca dos requisitos RQ1 e RQ2. Tal formulário é apresentado em contextos de uso diferentes e com funcionalidades para personalizar seus comportamentos de acordo com o perfil do usuário. No perfil do controlador, o formulário será utilizado para inclusão, alteração e exclusão de informações, uma vez que o objetivo é permitir que o controlador informe como suas aplicações de software manipularão os dados pessoais. Já no perfil dos titulares dos dados, o mesmo formulário preenchido pelos controladores será exibido, mas em modo leitura para permitir somente a visualização da informação. No formulário, as informações foram organizadas utilizando o padrão de *design* de abas, no qual cada aba contempla uma classe do TR-Model com seus respectivos atributos. Para os atributos, buscou-se utilizar componentes de interface que melhor atendiam as especificações dos metadados no TR-Model. Na Figura 1 é mostrada a interface do formulário com alguns atributos referentes ao compartilhamento de dados pessoais.

Para o RQ3, buscou-se propor uma interface de busca avançada a fim de cobrir o maior número de possibilidades de demandas por Transparência dos titulares de dados, inspirada nas buscas avançadas existentes em bases de dados científicas como a IEEE<sup>8</sup>, ACM<sup>9</sup> e SBC Open Library<sup>10</sup>, com opções e busca textual, filtro de informações e montagem de *strings* de busca.

<sup>7</sup><https://github.com/Larissa-Amadeu/ProjetoLaravel>

<sup>8</sup><https://ieeexplore.ieee.org/search/advanced>

<sup>9</sup><https://dl.acm.org/search/advanced>

<sup>10</sup><https://sol.sbc.org.br/busca/>

**Figura 1. Exemplo de formulário com alguns atributos de compartilhamento.**

Os resultados da busca são mostrados em uma interface com o padrão de lista e, a partir dessa interface, o titular dos dados poderá acessar as informações do projeto, conforme já descrito no RQ2. Um Exemplo da interface de busca é mostrada na Figura 2.

**Figura 2. Formulário de busca.**

A seguir são apresentadas algumas considerações finais deste artigo, bem como as atividades que podem ser conduzidas no futuro para melhor adequação e utilização deste repositório.

#### **4. Considerações Finais e Trabalhos Futuros**

Esse artigo apresentou o processo de criação e desenvolvimento de um repositório de informações sobre a manipulação de dados pessoais. A proposta inicial está disponível no GitHub, conforme apresentado anteriormente na Seção 3. Tal disponibilização visa permitir colaboração de demais profissionais. Espera-se que uma versão seja disponibilizada para avaliação e validação com controladores e titulares dos dados em breve.

Espera-se com o repositório, criar um ambiente favorável para empresas disponibilizarem informações sobre a manipulação dos dados pessoais de maneira prática e simplificada sem comprometer segredos comerciais e, ao mesmo tempo, proporcionar uma ferramenta prática para titulares dos dados buscarem e consultarem essas informações sem a necessidade de vasculhar, de forma individual, políticas de privacidade e segurança ou procurar em mecanismos escondidos dentro das aplicações.

Como trabalhos futuros, pretende-se disponibilizar o repositório para atividades de avaliação e validação em duas frentes: (1) controladores - para verificar o interesse e

a aderência do repositório para apoiar as empresas no fato de tornarem-se mais transparentes para seus usuários; (2) titulares dos dados - nessa avaliação pretende-se verificar a experiência do usuário na busca, visualização e interpretação da informação.

## 5. Agradecimentos

Agradecemos à Universidade Estadual do Norte do Paraná pelo apoio financeiro no Programa Institucional de Bolsas de Desenvolvimento Tecnológico e Inovação (PIBITI) e à Fundação Araucária de Apoio ao Desenvolvimento Científico e Tecnológico do Estado do Paraná pelo apoio financeiros pelo Programa Institucional de Bolsas de Iniciação Científica.

## Referências

- Allard, S. (2012). DataONE: Facilitating eScience through Collaboration. *Journal of eScience Librarianship*, 1(1):4–17.
- Audich, D. A., Dara, R., and Nonnecke, B. (2021). Improving readability of online privacy policies through doop: A domain ontology for online privacy. *Digital*, 1(4):198–215.
- Azzam, T. and Jacobson, M. R. (2015). Reflections on the Future of Research on Evaluation. *New Directions for Evaluation*, 29(148):103–116.
- Briganó, G. U., Brancher, J. D., and de Barros, R. M. (2021). Proposta de um Repositório de Conhecimentos em Governança de TIC. pages 767–772.
- Casagrande, M. F. R., Kozima, G., and Willrich, R. (2015). Técnica de Recomendação Baseada em Metadados para Repositórios Digitais Voltados ao Ensino. *Revista Brasileira de Informática na Educação*, 23(02):70.
- Coleti, T. A., Corrêa, P. L. P., Filgueiras, L. V. L., and Morandini, M. (2020). TR-Model. A Metadata Profile Application for Personal Data Transparency. *IEEE Access*, 8(1):75184–75209.
- Filgueiras, L. V. L., Leal, A. S. F., Coleti, T. A., Morandini, M., Correa, P. L., and Alves-Souza, S. N. (2019). Keep System Status Visible: Impact of Notifications on the Perception of Personal Data Transparency. *Human-Computer Interaction. Perspectives on Design*, 1:513–530.
- Freund, G. P., Sembay, M. J., and De Macedo, D. D. J. (2019). Proveniência de Dados e Segurança da Informação: relações interdisciplinares no domínio da Ciência da Informação. *Revista Ibero-Americana de Ciência da Informação*, 12(3):807–825.
- Holtz, L. E., Nocun, K., and Hansen, M. (2011). Towards displaying privacy information with icons. *IFIP Advances in Information and Communication Technology*, 352 AICT:338–348.
- Murmann, P. and Fischer-Hübner, S. (2017). Tools for achieving usable ex post transparency: a survey. *IEEE Access*, 5:22965–22991.
- Sayão, L., Toutain, B. L., Rosa, G. F., and Marcondes, H. (2018). *Implantação e gestão de repositórios institucionais: políticas, memória, livre acesso e preservação*, volume 12.