



Sistema Progredir: analisando fluxos informacionais para a migração do banco de dados do Cadastro Único

Marcos Pereira de Novais

Mestre em Inovação Tecnológica, Universidade Nacional de Seoul (SNU), Seoul, Coreia do Sul
Tecnologista, Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict), Brasília, DF, Brasil.

<http://lattes.cnpq.br/8317465361448862>

Email: marcosnovais@ibict.br



Tainá Regly

Doutoranda em Ciência da Informação, Universidade Federal do Rio de Janeiro em convênio com o Instituto Brasileiro de Informação em Ciência e Tecnologia (UFRJ-Ibict), Rio de Janeiro, RJ, Brasil.

Pesquisadora, Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict), Brasília, DF, Brasil.

<http://lattes.cnpq.br/5572586959935763>

Email: tainasouza@ibict.br

Jaqueline Gutierri Coelho

Mestranda em Bioinformática, Universidade de Brasília (UnB), Brasília, DF, Brasil.
Assistente de Pesquisa, Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict), Brasília, DF, Brasil.

<http://lattes.cnpq.br/8034101020207172>

Email: jaquelinecoelho@ibict.br

André Daniel Fernandes Causanilhas

Pós-graduado, Fundação Getúlio Vargas (FGV), Brasília, DF, Brasil.

Pesquisador, Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict), Brasília, DF, Brasil.

<http://lattes.cnpq.br/3096598045450907>

Email: andrecausanilhas@ibict.br

João Victor Santos Aguiar

Pós-graduado, Centro Universitário do Instituto de Educação Superior De Brasília (IESB), Brasília, DF, Brasil.

Pesquisador, Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict), Brasília, DF, Brasil.

<http://lattes.cnpq.br/6219402504681132>

Email: joaoaguiar@ibict.br

Joyce Paola Teixeira Sobrinho

Bacharel, Instituto de Educação Superior de Brasília (IESB), Brasília, DF, Brasil.

Pesquisadora, Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict), Brasília, DF, Brasil.

<http://lattes.cnpq.br/1599646488165006>

Email: joycesobrinho@ibict.br

Lisandra Guerrero Pérez

Doutora em Ciência da Informação, Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, Minas Gerais, Brasil.

Pesquisadora, Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict), Brasília, DF, Brasil.

<http://lattes.cnpq.br/1761103654180287>

Email: lisandraperez@ibict.br

Tiago Emmanuel Nunes Braga

Doutor em Ciência da Informação, Universidade de Brasília (UnB), Brasília, DF, Brasil.
Pesquisador, Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict), Brasília, DF, Brasil.

<http://lattes.cnpq.br/8376134230259399>

Email: tiagobraga@ibict.br

Submetido em: 11/09/2023. **Aprovado em:** 30/04/2024. **Publicado em:** 18/07/2024.

RESUMO

Introdução: Este trabalho apresenta os resultados obtidos da 'Análise de fluxos informacionais entre os usuários do Cadastro Único e as instituições da rede de parceiros do desenvolvimento social', correspondente à Meta 2 do projeto de pesquisa Plataforma Informacional para o Sistema Progredir, uma parceria entre o IBICT e o Ministério do Desenvolvimento e Assistência Social, Família e Combate à Fome. Metodologia: Relato descritivo sobre a análise e a migração de bancos de dados do Cadastro Único para o sistema Progredir. Em cada caso, foram descritas as metodologias para o fluxo de dados e as ferramentas utilizadas. Resultados: Foram mapeados os envolvidos na produção e consumo da informação do Plano Progredir. Também foram analisados os fluxos de dados, denominados agentes, quando foi possível fazer o mapeamento de parceiros como o Cadastro Único e o Ministério do Trabalho. A estrutura tecnológica e o ambiente informacional foram mapeados e apresentados por meio de diagramas de extração de dados, com o uso de técnicas de inteligência artificial. Conclusão: O estudo da ciência de dados envolvendo as bases originais do Progredir (com o banco de dados do Cadastro Único para se tornarem interligadas) constituiu um grande desafio que permitiu a extração, transformação e migração dos dados visando auxiliar o usuário final.

Palavras-chave: fluxo de dados; banco de dados; Cadastro Único; Progredir.

INTRODUÇÃO

O programa Progredir foi instituído pelo Decreto nº 9.160, de 26 de setembro de 2017 (BRASIL, 2017), e pertence ao Ministério do Desenvolvimento e Assistência Social, Família e Combate à Fome (MDS). Define-se como “[...] um conjunto articulado de ações de inclusão produtiva para pessoas inscritas no Cadastro Único para Programas Sociais do Governo Federal” (Brasil, 2017).

- I. Estimular e ampliar o acesso de pessoas incluídas no Cadastro Único ao mundo do trabalho e propiciar melhores condições de emprego e renda;
- II. Articular e coordenar a oferta de serviços para inclusão produtiva, de forma a aproximar os trabalhadores e os empreendedores de baixa renda e o mundo do trabalho, por meio de ações de intermediação de mão de obra, qualificação profissional e empreendedorismo;
- III. Incentivar ações municipais e estaduais de inclusão produtiva; e
- IV. Incentivar ações de órgãos e entidades públicas e de instituições privadas que promovam a inclusão do público-alvo do Plano Progredir no mundo do trabalho (Brasil, 2017, *online*).

As ações do programa Progredir assentam-se em 3 (três) eixos principais: i) empreendedorismo; ii) mão de obra; iii) qualificação profissional (Brasil, 2018). O tratamento e organização de tais informações de modo estratégico mostra potencial para o acompanhamento de demandas de inclusão produtiva, características do mercado de trabalho e necessidades de oferta de capacitação regional. De forma geral, no âmbito do Progredir, as informações constituem um insumo fundamental para a promoção de avanços, melhorias e alcance de objetivos e metas.

Assim, foi estabelecido o projeto de pesquisa Plataforma Informacional para o Sistema Progredir a partir da parceria entre o MDS e o Instituto Brasileiro de Informação Científica e Técnica (IBICT). O intuito deste projeto reside na apresentação de resultados expressivos junto a seu público-alvo e na ampliação progressiva de suas ferramentas e rede de parceiros. A partir dessas premissas espera-se promover o acesso e o direcionamento de oportunidades de inclusão produtiva, otimizar programas, planos/projetos em andamento, coordenar e articular ações visando potencializar a efetividade da política social, facilitar resultados e ampliar seu alcance (Brasil, 2018). O projeto objetiva a proposição de um novo sistema Progredir para o desenvolvimento de políticas públicas voltadas à inclusão produtiva de pessoas em situação de vulnerabilidade social (Brasil, 2021).

Desta forma, o objetivo do presente artigo consiste em descrever os procedimentos realizados na Meta 2 do projeto referente à análise de fluxos informacionais entre os usuários do Cadastro Único e as instituições da rede de parceiros do desenvolvimento social. Para isso, foi realizado o mapeamento dos agentes envolvidos na produção e consumo de insumos informacionais relacionados ao projeto, a análise dos fluxos estabelecidos entre os agentes mapeados e a estruturação do fluxo analisado. Doravante, será apresentada a metodologia e os resultados alcançados nas ações voltadas para análise dos fluxos informacionais.

METODOLOGIA

A segunda meta está focada na análise dos fluxos informacionais. Nesta etapa, foram mapeados os envolvidos na produção e consumo da informação do Plano Progredir. Aqui também foram analisados os fluxos de dados, denominados agentes, quando foi possível fazer o mapeamento de parceiros como o Cadastro Único e o Ministério do Trabalho. Nesse sentido, foram efetuadas reuniões para alinhar, solicitar e efetuar acordos para o recebimento da estrutura e dos dados dos produtores socioeconômicos, como o Ministério do Trabalho, o Ministério da Educação, o Ministério do Desenvolvimento Regional, o Ministério da Mulher, da Família e dos Direitos Humanos, bem como o Ministério da Economia, formadores do Grupo Gestor do Plano Progredir (GGPP). Os resultados das reuniões permitiram a realização de negociações e bases na fase de entrega.

Foi realizada a carga da base do Progredir com tamanho de 45 Gigabytes, disponibilizada pela equipe do MDS na plataforma Microsoft Sharepoint. Posteriormente, recebeu-se a carga do sistema do Cadastro Único de aproximadamente 100 Gigabytes de informações. Criou-se um ambiente de banco de dados utilizando a ferramenta PostgreSQL¹ para carregamento de arquivos de dump (backup do banco de dados).

Com as bases internalizadas, houve progresso no andamento do processo científico de análise dos dados no qual foi identificada a necessidade de realizar a documentação e a melhoria de padrões de informação, tanto na modelagem dos dados, quanto no tratamento da qualidade dos dados.

Foi iniciado o processo de documentação do novo banco de dados, seguindo os padrões das formas normais. Com isso, criaram-se indicadores que descrevem o impacto identificado da desnormalização existente com as informações originais e a melhoria significativa após adotados padrões científicos de dados. Na próxima seção, são apresentados os resultados obtidos a partir da execução dos procedimentos metodológicos descritos.

Mapeamento de agentes envolvidos na produção e consumo de insumos informacionais relacionados ao projeto

O mapeamento dos agentes envolvidos na produção e consumo de insumos informacionais relacionados ao projeto foi realizado por meio da delimitação estabelecida em acordo com os técnicos do Ministério da Cidadania e a equipe do projeto no IBICT. O conjunto dos processos foi desenvolvido por meio de reuniões e ações que elucidaram os fluxos de dados a serem coletados.

Os agentes produzem informações diversas que devem ser compreendidas e ajustadas para a realidade do projeto Progredir. Nesse sentido, os Ministérios atuantes no GGPP serviriam de base para a tomada de decisões e o entendimento do que é produzido, assim como para entender quais seriam relevantes nesse mapeamento.

¹ PostgreSQL é um sistema de gerenciamento de bancos de dados (SGBD) de código aberto, que utiliza o padrão SQL para a interação entre dados e informações contidas em seu repositório. É usado principalmente para gestão dos dados de aplicativos móveis, geoespaciais, analíticos e de *websites* (Drake; Worsley, 2002).

Em uma fase inicial, foi realizada uma reunião com o Ministério do Trabalho e Previdência, onde foram mostrados os conjuntos de dados e a interoperabilidade com o projeto Progredir. Nessa reunião delimitaram-se os dados dos empregados e o seu dimensionamento correspondente à base dos usuários do Progredir. Isto permitiu definir os principais agentes envolvidos, assim como sua forma de produção e consumo de informações.

Foram identificados produtores de dados de uma rede que presta serviço a trabalhadores ao nível nacional por meio de 3 (três) artefatos:

- Rede de atendimento física;
- Gov.br;
- App Sine Fácil (que está se preparando para a migração ao app da carreira digital), pois o foco é unir emprego, trabalho e renda.

No caso do Sistema Nacional de Emprego (SINE), podem ser exportados diversos dados segundo segue: experiência profissional, escolaridade, Cadastro de Pessoas Físicas (CPF), número do Cadastro Nacional de Informações Sociais (CNIS) – caso o trabalhador não tenha sido localizado no CNIS, o sistema gera automaticamente o número para o trabalhador e passa para o Sistema Nacional de Emprego (SINE). Comenta-se que já estão sendo realizadas colaborações para poder existir um marcador com o Cadastro Único. Atualmente existe um campo auto declaratório da bolsa família. A maioria dos dados do SINE são extraídos do CNIS, e para que os dados possuam padronização, esses dados não podem ser alterados pelo sistema do SINE.

As exportações dos dados do SINE seriam informadas pelo *Dataprev*², ainda que atualmente a ênfase reside na apresentação do SINE, suas funcionalidades e seus dados. Foi sugerido pelo Ministério do Trabalho que o usuário acessasse o SINE por meio de um link colocado no Progredir, porém, a solução baseada apenas no link não permite assegurar a rastreabilidade da informação.

O cruzamento das informações para o rastreamento da necessidade dos perfis do público-alvo é importante para se estabelecer um projeto a longo prazo. Uma estratégia de trabalho em conjunto (a qual integre o Cadastro Único e o *Data Lake* do poder executivo) poderia apoiar na identificação e estruturação dos indicadores adequados. Essa estratégia também contribuiria para aprimorar possíveis soluções de pesquisa relacionadas à mão de obra. Durante o projeto, vislumbraram-se possíveis aplicações futuras de inteligência artificial dentro do SINE e do Progredir que se alinham na mesma junção da necessidade das informações. Também, nesses debates, foi analisada a formalização da parceria do Ministério do Trabalho com o MDS e o IBICT.

Para estabelecer uma estratégia de acompanhamento do usuário frente à demanda levantada anteriormente junto aos usuários do SINE, entendeu-se que uma possível solução passaria pela utilização de um marcador. O marcador permitiria estabelecer a rastreabilidade parcial com acompanhamento do sistema. Para uma correta efetividade dessa proposta, é

2 Empresa pública cuja missão consiste em prover soluções digitais para o exercício da cidadania, fornecendo soluções de Tecnologia da Informação e Comunicação para o aprimoramento e a execução de políticas sociais do Estado brasileiro (Brasil, 2023).

necessária a sua aplicação em todos os que já possuem o cadastro no Progredir, direcionados a partir da base do SINE. Essa solução poderia ser, posteriormente, inativada, caso no futuro ocorra o acesso direto ao banco de dados, o que garantiria uma maior rastreabilidade dos usuários.

Análise dos fluxos estabelecidos entre os agentes mapeados

O plano Progredir, foco desse debate, está concebido como um conjunto de ações do Governo Federal para gerar emprego, renda e promover a construção da autonomia das pessoas inscritas no Cadastro Único para Programas Sociais do Governo Federal. Foram realizados questionamentos ao então Ministério da Mulher, Família e Direitos Humanos (MMFDH) a respeito de quais dados específicos eram necessários para a disponibilização e o andamento do projeto. Nesse sentido, são necessários os dados dos cidadãos, pois este projeto de construção de uma Plataforma Informacional trabalha com políticas públicas (ex: Cadastro Único).

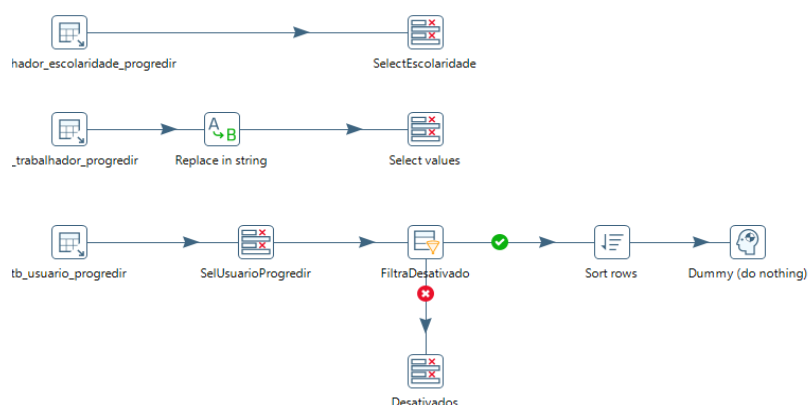
Foi informado por meio do então MMFDH que a área de TI daquele Ministério não possuía essa informação de forma granular, pois os dados não eram tratados pela TI e sim por uma secretaria. Após a classificação e a arguição do acesso aos dados, a autorização de acesso partiria do responsável pelo dado. O grupo responsável pela pesquisa não conseguiu acesso ao banco de dados, e sim a uma cópia não atualizável de um conjunto reduzido de informações. Foi identificado que uma parte desses dados é utilizada para alimentar o repositório, e para isso é necessário o cruzamento do CPF do cidadão para a API por meio de conexões seguras. A arquitetura atual desses dados já comporta essa ação.

Identificou-se que a maioria dos dados não está sistematizado, pois muitos não foram coletados via formulários. Isso gera uma dificuldade para identificar quais pessoas utilizam as políticas públicas que estão vinculadas ao Cadastro Único. As instituições identificadas não conseguiram informar quais pessoas estão vinculadas ao Cadastro Único. O MMFDH informou que uma possível perda de dados e o não controle da informação deveriam ser priorizados nas próximas políticas públicas visando controlar esses dados para cenários futuros.

Estrutura do fluxo analisado

A estrutura tecnológica e o ambiente informacional foram mapeados e apresentados por meio de diagramas de extração de dados, com o uso de técnicas de inteligência artificial, especificamente as técnicas *Extract Transform and Load* (ETL). A **FIGURA 1** e a **FIGURA 2** mostram essas transformações e análises dos dados no processo de construção de uma base inter operável.

FIGURA 1 – Estrutura tecnológica de transformação (Cadastro Único – Progridir) em ETL

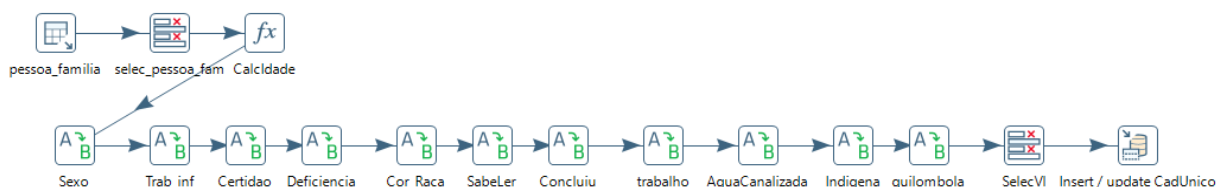


Fonte: Documentação do projeto de pesquisa (2023).

Devido à necessidade de uma melhoria na qualidade dos dados, foram utilizadas ferramentas e metodologias para fazer a transformação e a adequação dos dados, visando integrar as bases do Cadastro Único e a plataforma Progridir.

A **FIGURA 1** apresenta a transformação dos dados originais de 3 (três) tabelas existentes no Progridir utilizando a ferramenta *Pentaho*³. Aqui, o objetivo é converter os dados desorganizados em informações filtradas e de alta qualidade.

FIGURA 2 – Estrutura tecnológica ETL Cadastro Único



Fonte: Documentação do projeto de pesquisa (2023).

A **FIGURA 2** exemplifica uma das tabelas do Cadastro Único usando a mesma ferramenta utilizada na **FIGURA 1**, com o propósito de tratar os dados para que os bancos do Progridir e do Cadastro Único se relacionem.

Com a realização da migração, tem-se atuado continuamente no processo científico de estudo das informações legadas para o novo modelo. Além disso, no final do ano de 2022, foi realizado um *workshop* voltado para a apresentação e disseminação dos resultados parciais do projeto (Workshop, 2022). Nesse evento foram explicadas outras entregas realizadas

3 Pentaho é uma plataforma de inteligência de negócios orientada para a solução e é focada no processo, que inclui todos os principais componentes necessários para implementar soluções baseadas em processos (Gorman, 2009).

juntamente com a migração do banco de dados do Cadastro Único, tal como a criação de um modelo de mala direta para o novo Progredir e de um observatório informacional apto a representar -de forma gráfica- os dados estatísticos relativos à nova plataforma.

CONSIDERAÇÕES FINAIS

Os resultados apresentados neste trabalho baseiam-se nas necessidades específicas e demandadas pelo MDS para o design e construção da Plataforma Informacional do Sistema Progredir.

Considerando os 3 (três) eixos principais do Plano Progredir: I) empreendedorismo; II) mão de obra; e III) qualificação profissional, a equipe do IBICT tem desenvolvido um conjunto de tarefas visando criar uma plataforma capaz de fornecer serviços gratuitos vinculados ao provimento e divulgação de oportunidades de emprego, cursos presenciais e à distância, e possibilidades de microcrédito para financiamento de negócios individuais e em grupo para usuários inscritos no Cadastro Único. Sob a perspectiva da ciência de dados, constitui um grande desafio, ainda, o processo de interligação das bases originais do Progredir com a do Cadastro Único, tarefa que se mantém como prioridade do projeto.

REFERÊNCIAS

BRASIL. **Decreto n. 9.160, de 26 de setembro de 2017**. Institui o Plano Progridir. Diário Oficial da União: seção 1, Brasília, DF, ano 154, n. 186, 27 set. 2017. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2017/decreto/d9160.htm#:~:text=DECRETA%3A,do%20Governo%20Federal%20%2D%20Cadastro%20%C3%9Anico. Acesso em: 30 maio 2023.

BRASIL. **Empresa de Tecnologia e Informações da Previdência**: Dataprev. Gov.br. Brasília, 2023. Disponível em: Acesso em: <https://www.gov.br/gestao/pt-br/aceso-a-informacao/institucional/competencias/dataprev>. 12 jun. 2023.

BRASIL. Ministério do Desenvolvimento Social e Combate à Fome. **Plano Progridir**: enfrentando a desigualdade pela inclusão produtiva. *In*: Cadernos de estudos de desenvolvimento social em debate. – n. 33, 2018 . Brasília: Ministério do Desenvolvimento Social; Secretaria de Avaliação e Gestão da Informação, 2018. ISSN 1808-0758. Disponível em: https://aplicacoes.mds.gov.br/sagirms/ferramentas/docs/caderno_de_estudos_33_com_capa.pdf. Acesso em: 23 maio 2023.

BRASIL. **Plano de Trabalho do Termo de Execução Descentralizada nº 11/2021**. 2021. Ministério da Cidadania Departamento de Inclusão Produtiva Urbana. Coordenação-Geral de Qualificação. Disponível em: <https://www.gov.br/mds/pt-br/aceso-a-informacao/licitacoes-e-contratos/termos-de-cooperacao-e-termos-de-execucao-descentralizada/PlanodeTrabalhoTEDn112021.pdf>. Acesso em: 30 maio 2023.

DRAKE, J. D.; WORSLEY, J. C. **Practical PostgreSQL**. Califórnia: O'Reilly Media, Inc., 2002.

GORMAN, W. **Pentaho Reporting 3.5 for Java Developers**. Birmingham: Packt Publishing Ltd, 2009.

WORKSHOP: divulgação do observatório e nova marca. [S. l.: s. n.], 2022. 1 vídeo (2:13:45 min). Publicado pelo canal MDS. Disponível em: <https://youtu.be/mtRnviGTJaQ>. Acesso em: 30 maio 2023.