

UNIVERSIDADE TUIUTI DO PARANÁ

**ALEXANDRE JUNIOR GARCIA DOS SANTOS
ÍCARO DE ASSIS HOSTERT
LEONARDO DE PAULA TEIXEIRA
VITOR ROBERTO BATISTA
SCHIRMER**

EXTRACTIFY

APLICATIVO VOLTADO À EXTRAÇÃO DE DADOS EM MASSA DE ARQUIVOS

Relatório apresentado ao Curso de Análise e Desenvolvimento de Sistemas da
Universidade Tuiuti do Paraná como requisito avaliativo do 2º bimestre da disciplina de
Interface Humano-Computador.

Professores: Patrícia Rucker de Bassi, Angela Zatti e Sergio Luiz Marques Filho

**CURITIBA
2025**

**ALEXANDRE JUNIOR GARCIA DOS SANTOS
ÍCARO DE ASSIS HOSTERT
LEONARDO DE PAULA TEIXEIRA
VITOR ROBERTO BATISTA
SCHIRMER**

EXTRACTIFY

APLICATIVO VOLTADO À EXTRAÇÃO DE DADOS EM MASSA DE ARQUIVOS

Relatório apresentado ao Curso de Análise e Desenvolvimento de Sistemas da
Universidade Tuiuti do Paraná como requisito avaliativo do 2º bimestre da disciplina de
Interface Humano-Computador.

Professores: Patrícia Rucker de Bassi, Angela Zatti e Sergio Luiz Marques Filho

**CURITIBA
2025**

Resumo

O trabalho tem como objetivo a documentação do projeto Extractify, uma plataforma web desenvolvida para automatizar a extração de dados de arquivos PDF e imagens. A plataforma visa oferecer uma interface intuitiva e personalizável, atendendo às necessidades de diferentes perfis de usuários, como profissionais financeiros e administrativos. As principais funcionalidades incluem autenticação de usuários, upload de arquivos, criação de extrações personalizadas, visualização de resultados, exportação de dados, estatísticas de uso, configurações de aparência (temas claro/escuro) e seleção de idioma. O projeto incorpora recursos de acessibilidade, como suporte à Língua Brasileira de Sinais (Libras) por meio do VLibras Widget e alteração de temas para usuários com fotossensibilidade. A avaliação heurística da interface, baseada nos princípios de Nielsen, destaca a visibilidade do status do sistema, consistência, controle do usuário e flexibilidade. Apesar dos avanços, melhorias como a reexecução de extrações e otimização da visualização de dados ainda estão em desenvolvimento. O projeto reflete a aplicação de conhecimentos adquiridos no curso, oferecendo uma solução prática para automação de processos.

Palavras-chave: extração de dados, interface web, acessibilidade, expressões regulares, automação.

Sumário

1 Dados do Projeto	5
1.1 Nome do Projeto	5
1.2 Tema do Projeto	5
1.3 Objetivo do Projeto	5
1.4 Equipe do Projeto	5
2 Funcionalidades do Projeto	5
3 Storytelling	6
3.1 Cenário Envolvendo 3 Perfis de Usuários	6
3.2 Contexto de Uso para Cada Perfil	6
3.3 Mapa de Personas	7
4 Acessibilidade	7
4.1 Mudança de Tema	7
4.2 Tradução do Conteúdo para Libras	9
5 Protótipo	10
5.1 Telas Desenvolvidas (Alta Fidelidade)	10
6 Avaliação Heurística da Interface	14
7 Conclusão	15
Referências	16

1 Dados do Projeto

1.1 Nome do Projeto

Extractify

1.2 Tema do Projeto

Extração de dados automatizada a partir de arquivos enviados pelo usuário.

1.3 Objetivo do Projeto

O objetivo do projeto Extractify é oferecer uma plataforma web intuitiva que permita aos usuários realizar extrações de dados utilizando expressões regulares (regex) em arquivos PDF ou imagens, promovendo praticidade e automação no tratamento de informações.

1.4 Equipe do Projeto

- Alexandre Junior Garcia dos Santos
- Ícaro de Assis Hostert
- Leonardo de Paula Teixeira
- Vitor Roberto Batista Schirmer

Funcionalidades do Projeto

As principais funcionalidades desenvolvidas foram:

- Login e autenticação de usuários
- Criação de extrações com campos personalizados (nome + regex)
- Upload de arquivos (PDF ou imagens)
- Visualização dos resultados extraídos por arquivo
- Listagem de extrações realizadas
- Exportação de dados
- Visualização de estatísticas de uso
- Configurações de aparência (claro/escuro)
- Seleção de idioma

Storytelling

1.5 Cenário Envolvendo 3 Perfis de Usuários

1. **Profissional Financeiro:** Precisa extrair dados de documentos financeiros de forma rápida.
2. **Profissional Administrativo:** Quer automatizar a coleta de e-mails de faturas digitalizadas.
3. **RH:** Extrair automaticamente dados de formulários e comprovantes de ponto digitalizados para alimentar o sistema de folha de pagamento.

1.6 Contexto de Uso para Cada Perfil

- O profissional financeiro acessa a plataforma para extrair valores e preços de notas fiscais.
- O profissional administrativo utiliza para extrair e-mails de documentos escaneados.
- O profissional do RH utiliza para extrair os dados dos holerites dos funcionários.

1.7 Mapa de Personas

Persona	Dados Demográficos	Objetivos	Frustrações	Citação Representativa
Ana, 40 anos	Responsável pelo departamento financeiro em uma startup	Extrair valores de notas fiscais para calcular gastos mensais e semanais	Perder tempo copiando e colando manualmente trechos de PDF	Eu só quero focar no conteúdo, não em ficar garimpando texto.
Marcos, 34 anos	Assistente administrativo, trabalha em consultório médico	Coletar e organizar e-mails de pacientes a partir de documentos escaneados; gerar lista de contatos limpa	Ler OCR impreciso em PDF escaneado; formatação inconsistente que quebra planilha	Preciso de algo confiável, sem ter que revisar linha a linha.
Beatriz, 28 anos.	Analista de RH em empresa de médio porte	Extrair dados de comprovantes de ponto para folha de pagamento e relatórios de horas extras	Revisão manual de PDFs; erros de OCR em horários que comprometem a folha	Quero enviar aquilo já para o sistema de RH, sem perder horas corrigindo planilhas.

Acessibilidade

1.8 Mudança de Tema

- **Recurso aplicado:** Alteração de tema (claro/escuro) para atender usuários com fotossensibilidade ou preferências visuais.
- **Justificativa:** Melhorar a experiência visual e reduzir a fadiga ocular em ambientes com diferentes iluminações.
- **Ícone de acessibilidade:** Presente na aba "Configurações" > "Aparência"

FIGURA 1 - MUDANÇA DE TEMA (CLARO)

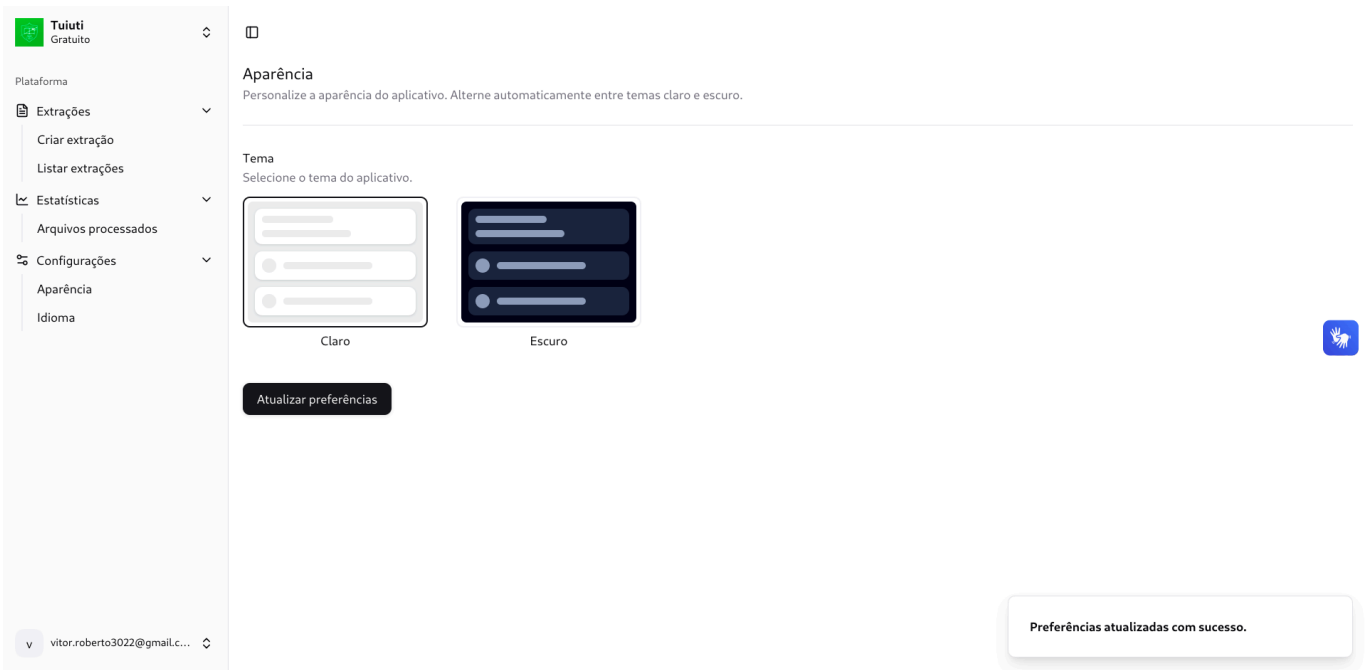
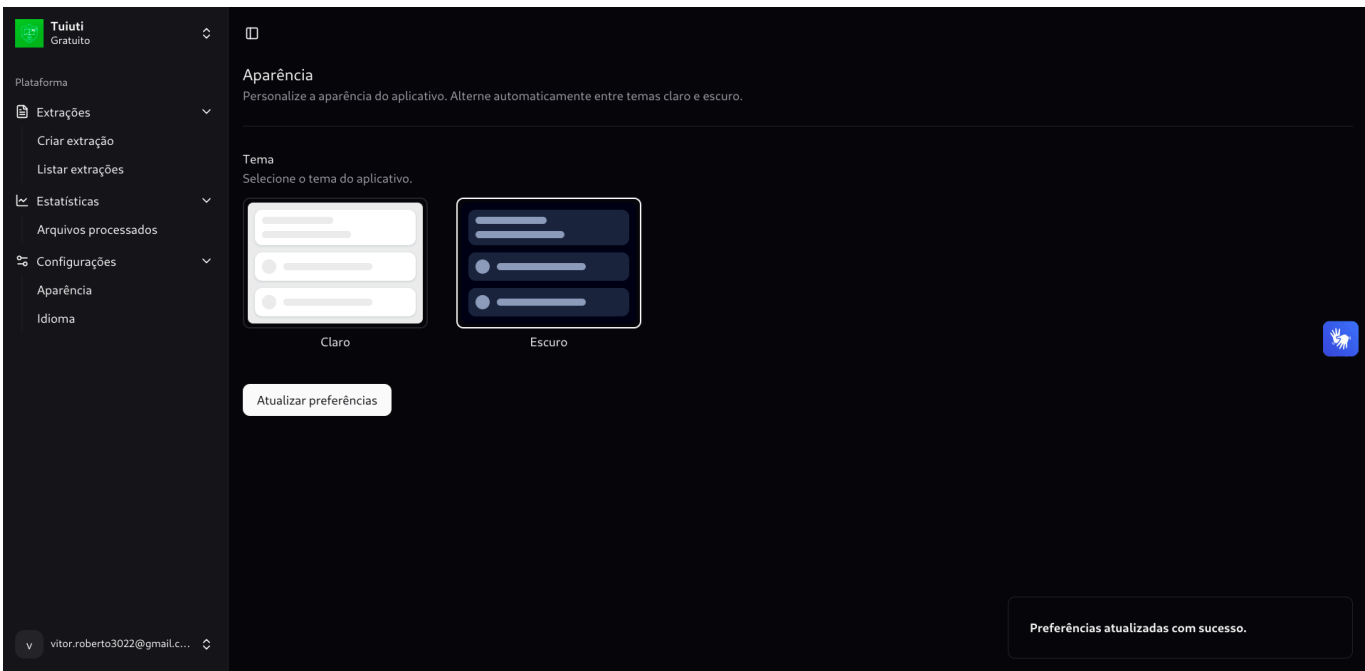


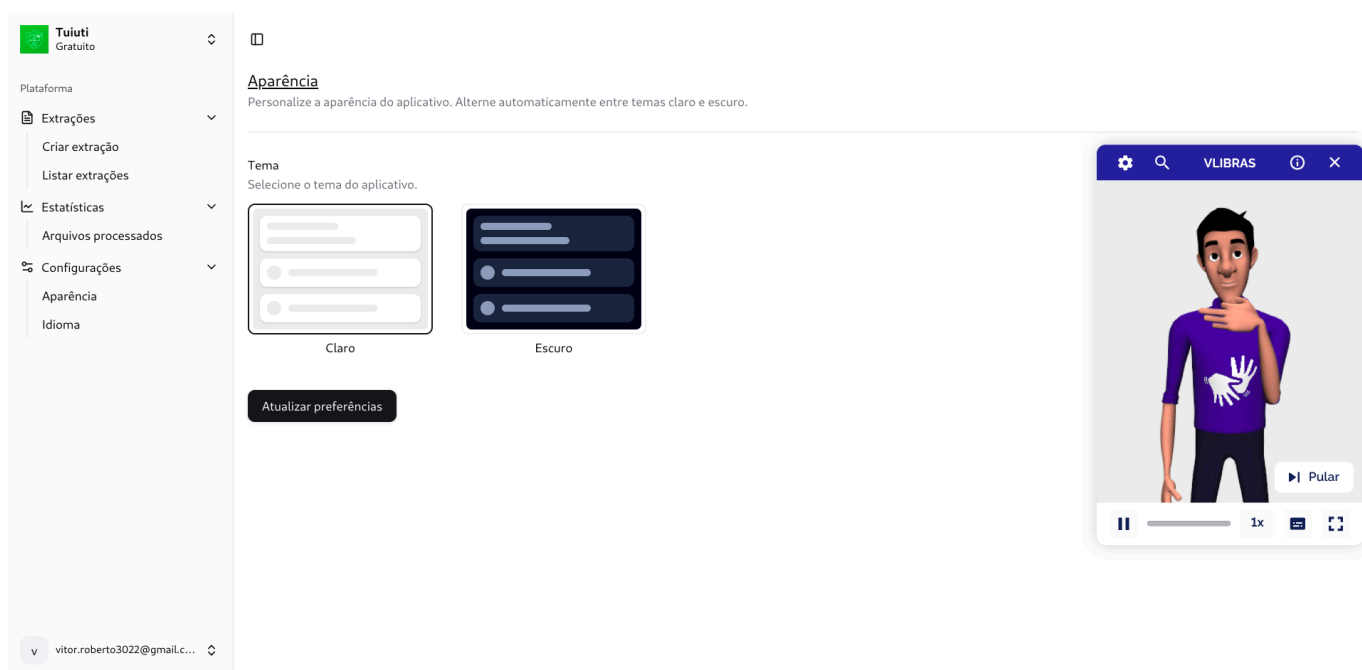
FIGURA 2 - MUDANÇA DE TEMA (ESCURO)



1.9 Tradução do Conteúdo para Libras

- **Recurso aplicado:** Tradução do conteúdo do site para Libras, utilizando o VLibras Widget.
- **Justificativa:** Os surdos enfrentam dificuldades para ler, escrever e se comunicar na língua oral. O VLibras Widget traduz conteúdos do Português Brasileiro para a Língua Brasileira de Sinais (Libras), tornando o site acessível a pessoas surdas.
- **Ícone de acessibilidade:** Presente em todas as telas, centralizado verticalmente no lado direito da página.

FIGURA 3 - TRADUÇÃO PARA LIBRAS

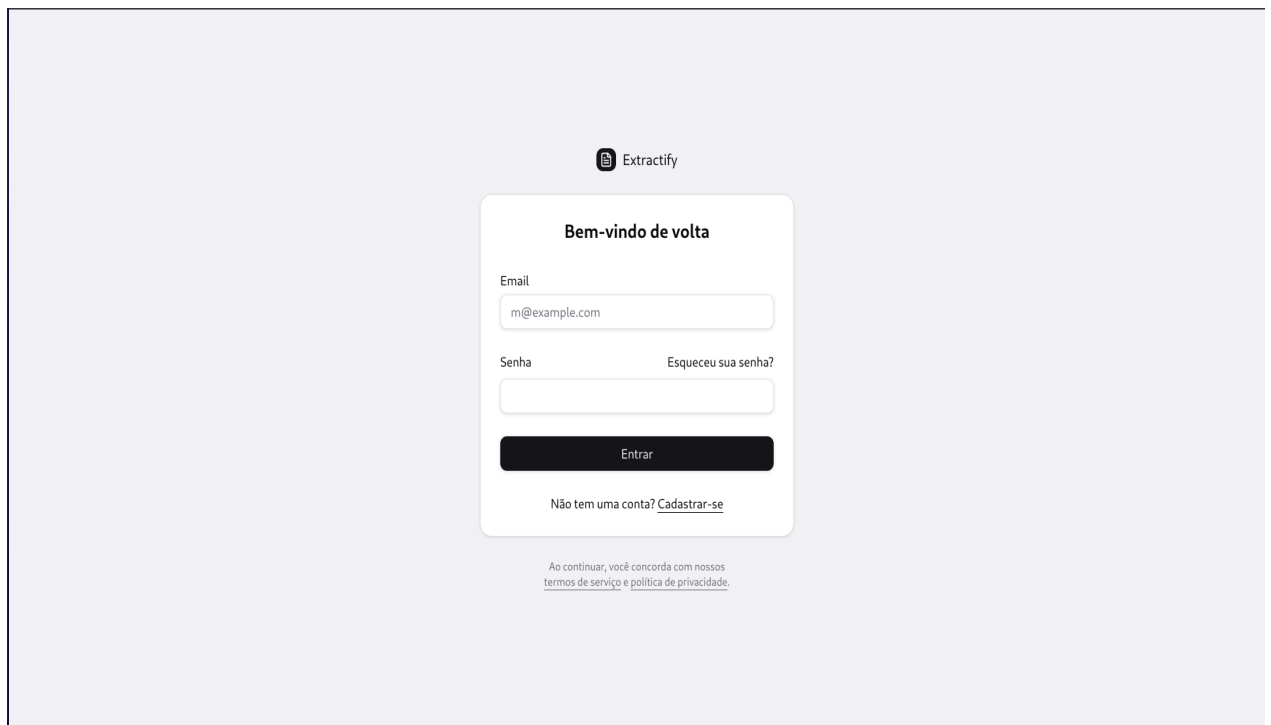


5 Protótipo

1.10 Telas Desenvolvidas (Alta Fidelidade)

- **Login:** Autenticação do usuário.

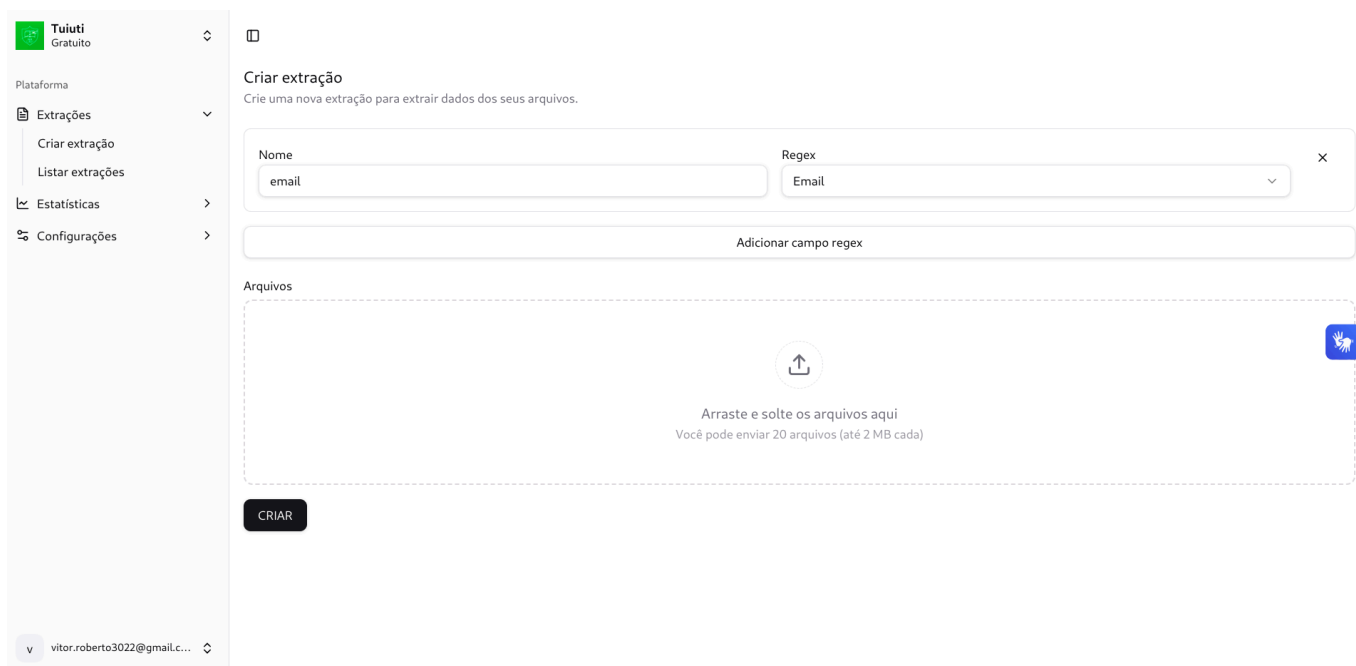
FIGURA 4 - TELA AUTENTICAÇÃO



The login screen for 'Extractify' features a light purple background. At the top center is the 'Extractify' logo. Below it, a white card contains the heading 'Bem-vindo de volta'. The card has two input fields: 'Email' with the placeholder 'm@example.com' and 'Senha'. To the right of the password field is a link 'Esqueceu sua senha?'. Below the inputs is a dark blue 'Entrar' button. At the bottom of the card is a link 'Não tem uma conta? [Cadastrar-se](#)'. Below the card, centered, is a small disclaimer: 'Ao continuar, você concorda com nossos termos de serviço e política de privacidade.'

- **Criar Extração:** Definição de campos e upload de arquivos.

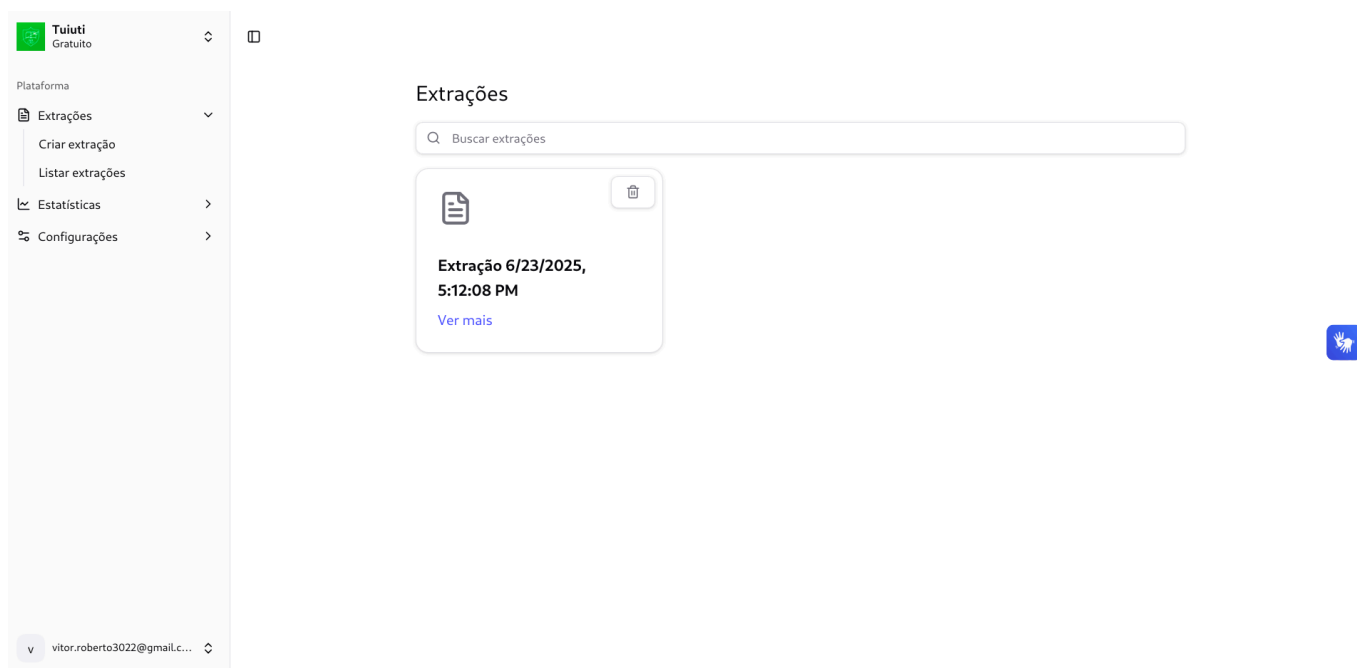
FIGURA 5 - TELA CRIAR EXTRAÇÃO



The 'Criar extração' screen has a light gray background. On the left is a sidebar with the 'Tuiuti' logo and 'Gratuito' text, followed by a 'Plataforma' section containing 'Extrações' (with sub-items 'Criar extração' and 'Listar extrações'), 'Estatísticas', and 'Configurações'. The main area has a title 'Criar extração' and a subtitle 'Crie uma nova extração para extrair dados dos seus arquivos.'. Below this are two input fields: 'Nome' (containing 'email') and 'Regex' (containing 'Email' with a dropdown arrow). A button 'Adicionar campo regex' is below these. The 'Arquivos' section features a large dashed box for file upload, with an upload icon and text: 'Arraste e solte os arquivos aqui' and 'Você pode enviar 20 arquivos (até 2 MB cada)'. A 'CRIAR' button is at the bottom left. The bottom of the sidebar shows a user profile for 'vitor.roberto3022@gmail.c...'.

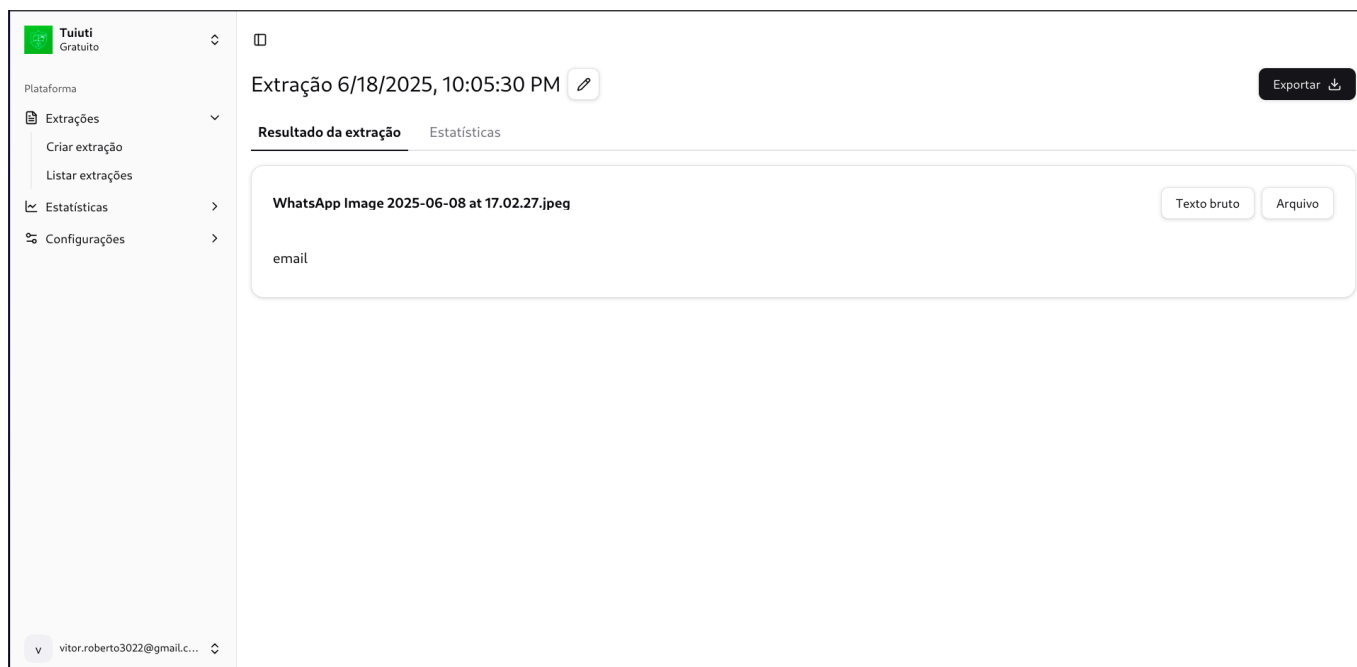
- **Listar Extrações:** Visualização das extrações criadas.

FIGURA 6 - TELA LISTAR EXTRAÇÕES



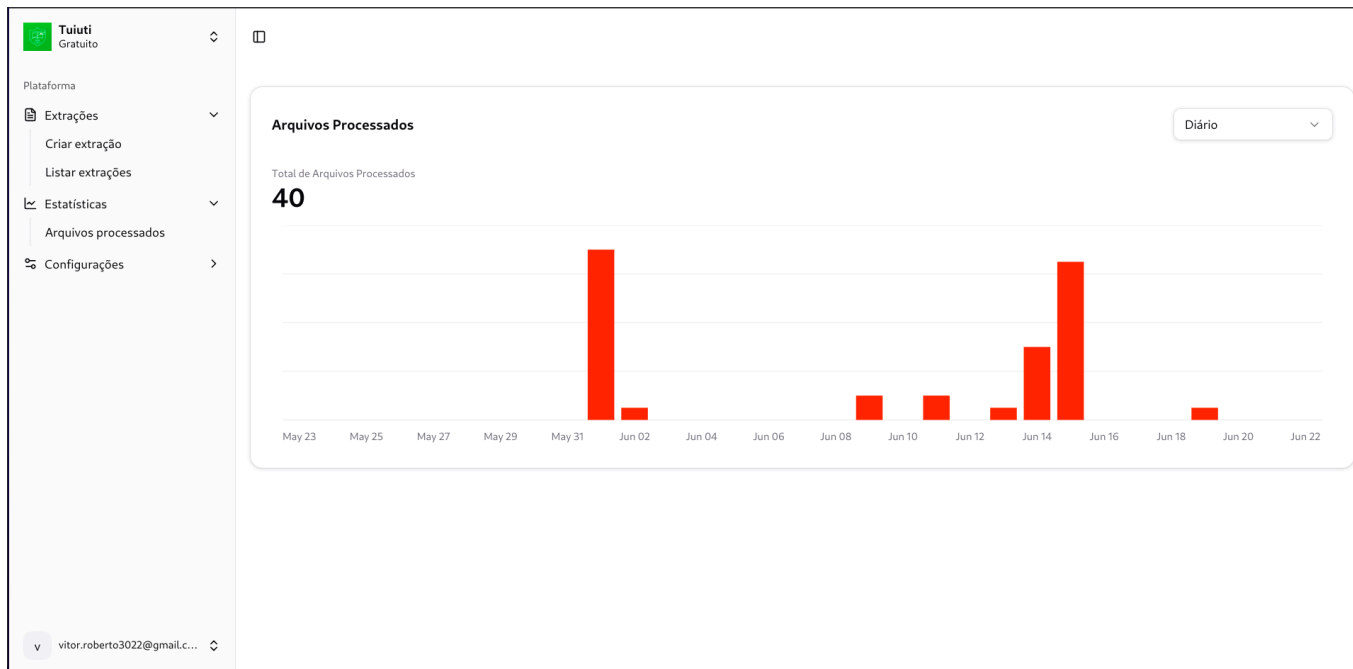
- **Resultado da Extração:** Dados extraídos por arquivo.

FIGURA 7 - TELA RESULTADO DA EXTRAÇÃO



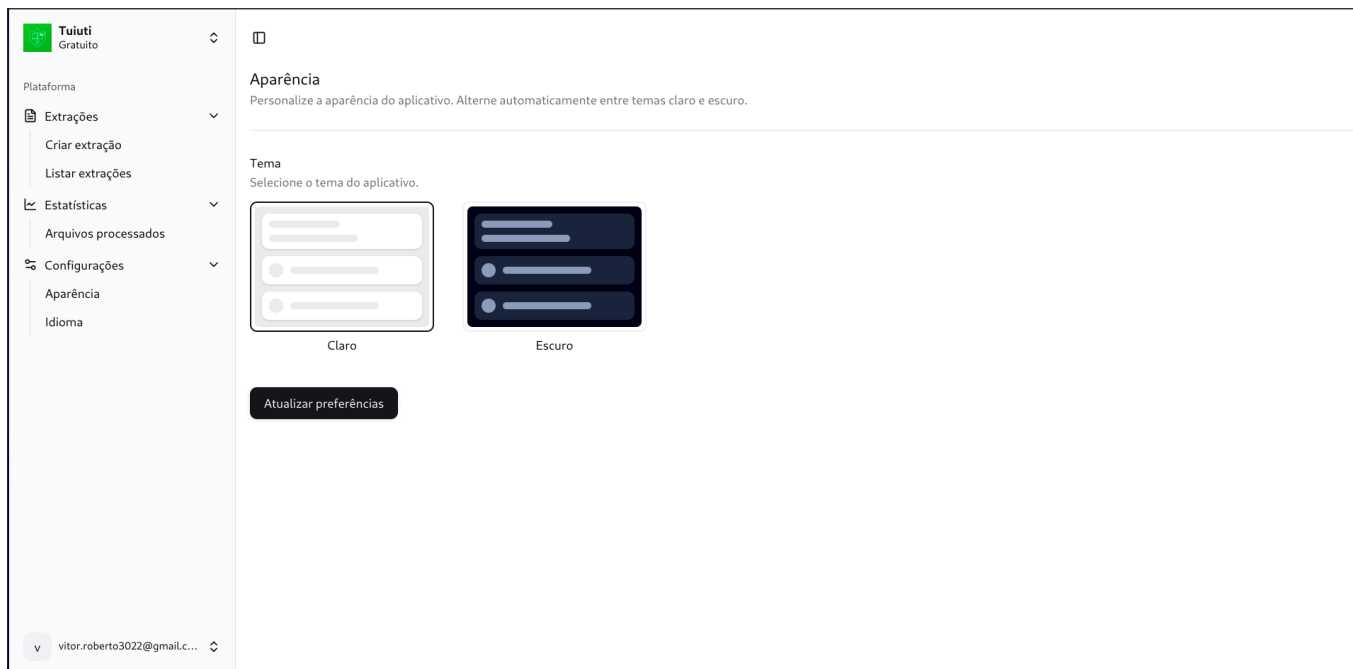
- **Estatísticas:** Visualização de número de arquivos processados por dia.

FIGURA 8 - TELA ESTATÍSTICAS



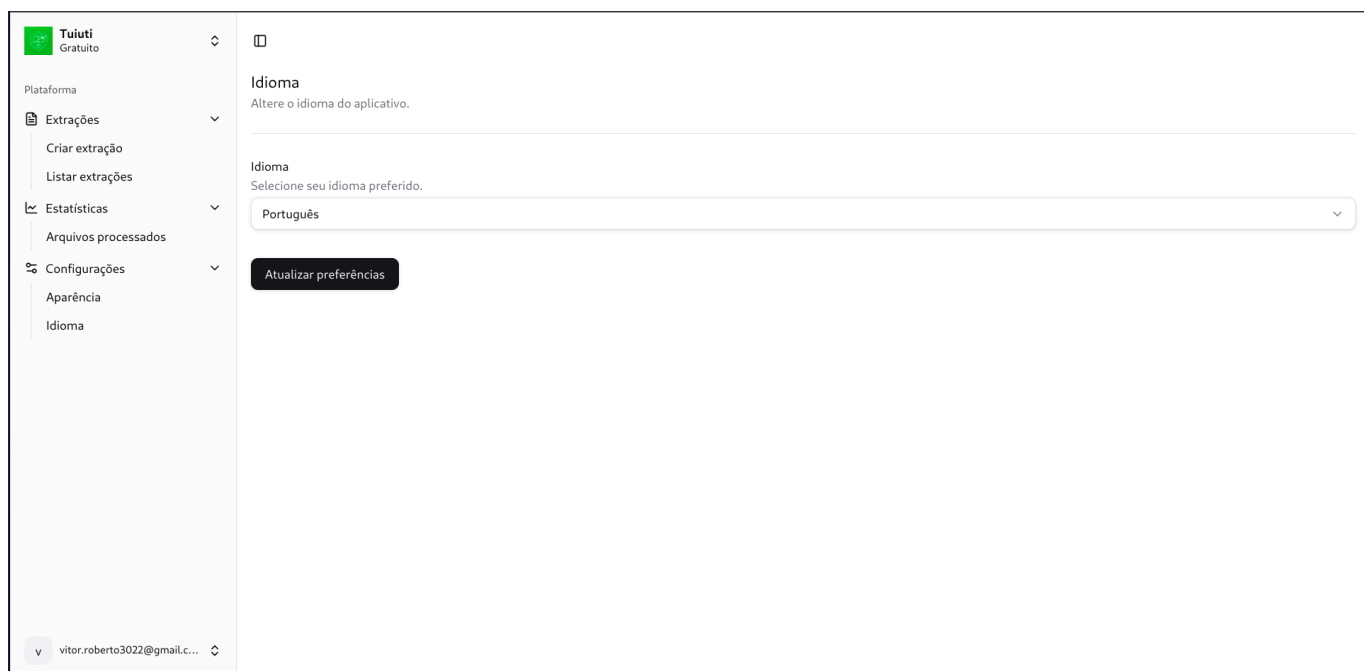
- **Configurações > Aparência:** Seleção de tema claro ou escuro.

FIGURA 9 - TELA APARÊNCIA



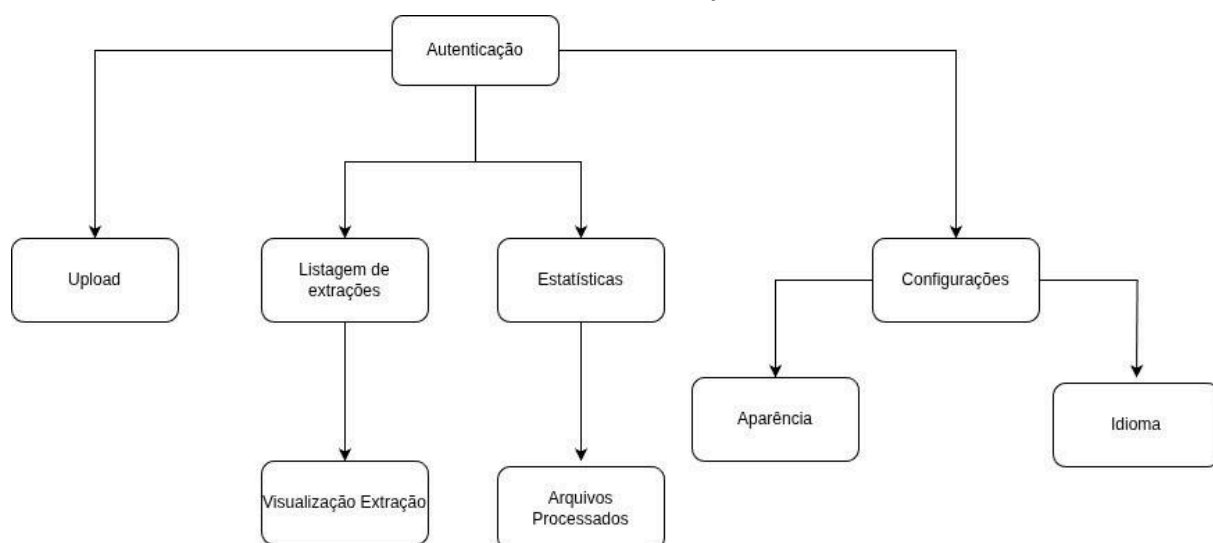
- **Configurações > Idioma:** Seleção do idioma do sistema.

FIGURA 10 - TELA IDIOMA



1.11 Mapa de Navegação

FIGURA 11 - MAPA DE NAVEGAÇÃO



6 Avaliação Heurística da Interface

A equipe realizou uma avaliação heurística da interface, identificando erros de usabilidade no site Extractify. O Quadro 1 apresenta o Framework de avaliação heurística da Interface do Extractify. Para os erros identificados atribuiu-se um Grau de Severidade, que serão tratados no próximo release do site.

QUADRO 1 - AVALIAÇÃO HEURÍSTICA DA INTERFACE

Inspeção Heurística				
Heurística	Itens de Usabilidade	Comentários	Grau de severidade	Grau de severidade da heurística
Visibilidade do status do sistema (feedback)	Facilidade de aprendizagem	Sistema fornece feedback claro durante processos de upload e extração		3
	Eficiência	Interface permite acompanhar o progresso das operações em tempo real		
	Facilidade de lembrar	Status das extrações anteriores são mantidos na listagem		
	Erros	Não há feedback para registros excluídos do sistema ou usuários.	3	
	Satisfação subjetiva	Usuário se sente informado sobre o estado das operações		
Compatibilidade do sistema com o mundo real	Facilidade de aprendizagem	Interface usa terminologia familiar para a maioria dos usuários		1
	Eficiência	Organização das funcionalidades segue lógica intuitiva		
	Facilidade de lembrar	Ícones e elementos visuais são reconhecíveis		
	Erros	Na tela de criar extração(figura 5): o campo a ser extraído é chamado de regex.	1	
	Satisfação subjetiva	Interface se alinha com expectativas dos usuários		
Controle do usuário e liberdade	Facilidade de aprendizagem	Usuário pode navegar livremente entre as funcionalidades		5
	Eficiência	Possibilidade de excluir extrações e alterar configurações		
	Facilidade de	Opções de controle são consistentes		

	relembrar	em toda a interface		
	Erros	Na tela de upload: O usuário não pode restaurar um registro apagado. Não aponta popup de dupla escolha.	5	
	Satisfação subjetiva	Usuário se sente no controle das suas ações		
Consistência e padrões	Facilidade de aprendizagem	Interface mantém padrões visuais consistentes		1
	Eficiência	Elementos similares funcionam de forma padronizada		
	Facilidade de relembrar	Cores, ícones e layout segue padrão estabelecido		
	Erros	Na tela 10 a palavra idioma aparece duas vezes.	1	
	Satisfação subjetiva	Consistência visual transmite profissionalismo		
Prevenção de erros	Facilidade de aprendizagem	Sistema possui validação que evitam erros básicos		1
	Eficiência	Validação de tipos de arquivo e campos obrigatórios		
	Facilidade de relembrar	Regras de validação são claras e consistentes		
	Erros	A equipe não identificou erros para esta heurística	1	
	Satisfação subjetiva	Usuário se sente seguro ao usar o sistema		
Reconhecimento ao invés de relembração	Facilidade de aprendizagem	Menu e navegação sempre visíveis facilitam o uso		1
	Eficiência	Histórico de extrações reduz necessidade de relembrar		
	Facilidade de relembrar	Interface intuitiva minimiza carga cognitiva		
	Erros	A equipe não identificou erros para esta heurística	1	
	Satisfação subjetiva	Usuário não precisa memorizar informações complexas		
Flexibilidade e eficiência de uso	Facilidade de aprendizagem	Sistema permite personalizações básicas (tema, idioma)		3
	Eficiência	Criação de extrações personalizadas aumenta produtividade		
	Facilidade de relembrar	Configurações salvas facilitam uso futuro		
	Erros	Não possui teclas de atalhos para acessar seções do sistema	3	
	Satisfação subjetiva	Usuário pode adaptar o sistema às suas necessidades		
Estética e design minimalista	Facilidade de aprendizagem	Interface limpa facilita o entendimento		1
	Eficiência	Ausência de elementos desnecessários acelera o uso		

	Facilidade de lembrar	Layout organizado torna elementos fáceis de encontrar		
	Erros	-	1	
	Satisfação subjetiva	Design agradável e profissional		
Ajudar os usuários a reconhecer, diagnosticar e corrigir erros	Facilidade de aprendizagem	Mensagens de erro são compreensíveis quando presentes		1
	Eficiência	Validações em tempo real ajudam a corrigir problemas rapidamente		
	Facilidade de lembrar	Padrões de erro são consistentes		
	Erros	A equipe não identificou erros para esta heurística	1	
	Satisfação subjetiva	Usuário consegue resolver problemas sem frustração		
Help e documentação	Facilidade de aprendizagem	Interface intuitiva reduz dependência de documentação		5
	Eficiência	Falta de ajuda pode reduzir eficiência de novos usuários		
	Facilidade de lembrar	Ausência de referência para funcionalidades complexas		
	Erros	Não possui ajuda para usuários via manual do usuário.	5	
	Satisfação subjetiva	Usuário pode se sentir perdido em situações complexas		

A avaliação heurística da interface do Extractify, revelou que o sistema oferece feedback claro e imediato ao usuário em quase todas as etapas do processo, como durante o upload de arquivos e após a conclusão das extrações, informando sobre o progresso e os resultados. Isso permite que o usuário acompanhe o que está acontecendo e entenda o status da sua ação.

A interface do Extractify mantém uma linguagem visual e funcional consistente. Ícones são padronizados em toda a plataforma, o que contribui para a familiaridade e facilita o aprendizado do usuário, reduzindo a necessidade de memorização.

O site concede ao usuário controle sobre suas ações. É possível excluir extrações e alterar preferências de aparência, como o tema claro/escuro. Essa flexibilidade aumenta a sensação de controle e permite que o usuário personalize sua experiência.

Embora o sistema ofereça a possibilidade de o usuário criar extrações personalizadas com campos definidos, a informação fornecida indica que a escolha dos campos de extração pode ser limitada ou pré-definida pelos administradores do site, o que impacta a experiência de extração e pode limitar a flexibilidade para certos casos de uso específicos. A reexecução de extrações e a otimização da visualização de dados ainda estão em desenvolvimento.

Além disso, o projeto incorpora recursos de acessibilidade que melhoram a usabilidade para diversos perfis de usuários, como o suporte à Língua Brasileira de Sinais (Libras) através do VLibras Widget e a alteração de temas para usuários com fotossensibilidade ou para reduzir a fadiga ocular.

Conforme exposto no Quadro 1, o próximo release do Extractify contém um manual do usuário com documentação para as possibilidades de extrações, bem como as limitações do site. A equipe precisa criar uma funcionalidade que permita restaurar extrações excluídas em um intervalo de tempo informado ao usuário.

Além disso, a equipe desenvolverá teclas de atalho para usuários avançados, direcionando estes para telas de criar extrações e consultar extrações, além de feedback constante para as ações no sistema, sobretudo na tela de excluir extrações.

Conclusão

O projeto Extractify representa uma solução inovadora para a automação da extração de dados de arquivos PDF e imagens. Com uma interface web intuitiva e recursos de acessibilidade, como suporte à Libras e temas claro/escuro, a plataforma promove inclusão e usabilidade. A avaliação heurística baseada em Nielsen confirmou a eficácia da interface, destacando sua consistência e flexibilidade. Contudo, melhorias como a reexecução de extrações, a melhoria na visualização de dados e na precisão na extração de dados são necessárias para aprimorar a experiência do usuário. O *Extractify* reflete o aprendizado adquirido no curso, oferecendo uma ferramenta robusta e escalável para automação de processos.

Referências

GOVERNO DO BRASIL. VLibras. Disponível em: <https://www.gov.br/governodigital/pt-br/acessibilidade-e-usuario/vlibras>. Acesso em: 4 maio 2025.

VLibras. Instalação do Widget VLibras em páginas web. Disponível em: <https://vlibras.gov.br/doc/widget/installation/webpageintegration.html?ga=2.205222480.1595640842.1682445746-816840059.1655413110>. Acesso em : 4 maio 2025.