

**PONTIFÍCIA UNIVERSIDADE CATÓLICA DE GOIÁS
ESCOLA POLITÉCNICA E DE ARTES
ENGENHARIA DE COMPUTAÇÃO**



PLATAFORMA DE EDIÇÃO/AUTOMAÇÃO PARA TRABALHOS ACADÊMICOS

HIGOR FERREIRA ALVES SANTOS

**GOIÂNIA - GO
2023**

HIGOR FERREIRA ALVES SANTOS

PLATAFORMA DE EDIÇÃO/AUTOMAÇÃO PARA TRABALHOS ACADÊMICOS

Trabalho de Conclusão de Curso apresentado à Escola Politécnica e de Artes, da Pontifícia Universidade Católica de Goiás, como parte dos requisitos para a obtenção do título de Bacharel em Engenharia de Computação.

Orientador:

Prof. M.E.E.Marcelo Antônio Adad

Banca examinadora:

Prof. Dra. Miriam Gusmão

HIGOR FERREIRA ALVES SANTOS

PLATAFORMA DE EDIÇÃO/AUTOMAÇÃO PARA TRABALHOS ACADÊMICOS

Trabalho de Conclusão de Curso aprovado em sua forma final pela Escola Politécnica e de Artes, da Pontifícia Universidade Católica de Goiás, para obtenção do título de Bacharel em Engenharia de Computação, em: ____/ ____/ ____

Orientador: Prof. M.E.E.Marcelo Antônio Adad

Orientador1: Prof. Dra. Miriam Gusmão

Orientador2:

DEDICATÓRIA

AGRADECIMETOS

EPÍGRAFE

RESUMO

ABSTRACT

Lista de ilustrações

Figura 1 – Passo a passo para criar um documento na plataforma	2
Figura 2 – Divisão de blocos em uma imagem	4
Figura 3 – Etapas do processo de Parsing	5
Figura 4 – Pilares da plataforma, (mapa mental)	7
Figura 5 – Logotipo do React	13
Figura 6 – Exemplo de código JSX	15
Figura 7 – Logotipo do NextJs	16
Figura 8 – Logotipo do EditorJs	18
Figura 9 – AntDesign	19
Figura 10 – Exemplo de componente em AntDesign	20
Figura 11 – Logotipo do NodeJs	21
Figura 12 – Correspondências do Regex abc	23
Figura 13 – Correspondências do Regex ab^*c , com metacaractere	24
Figura 14 – Não correspondências do Regex ab^*c , com metacaractere	24
Figura 15 – Estrutura de pastas e arquivos básica do projeto	28
Figura 16 – Estrutura da pasta de roteamento (App Router)	30
Figura 17 – Estrutura da pasta de roteamento (App Router)	31
Figura 18 – Exemplo de roteamento	32
Figura 19 – Sub árvore de renderização do layout principal	33
Figura 20 – Sub árvore de renderização da página principal	34
Figura 21 – Estrutura de pastas do componente Editor	35
Figura 22 – Sub árvore de renderização do componente Editor	35
Figura 23 – Componente React - Editor	35
Figura 24 – Classe - BasePlugin	36
Figura 25 – Estrutura de pastas dos plugins	36
Figura 26 – Parsing de um bloco Header	37
Figura 27 – Parsing de um bloco Paragraph	38
Figura 28 – Estrutura de pastas do módulo de parse	40
Figura 29 – Um ciclo de processamento de texto	41
Figura 30 – Conversão de plugins em código latex	43
Figura 31 – Caminho da pasta types dos plugins no parser	46

Lista de tabelas

Tabela 1 – Tecnologias do ambiente de desenvolvimento	6
Tabela 2 – Diferenças entre o JavaScript e o JSON	12
Tabela 3 – Métodos de string que podem ser usados com regex	25
Tabela 4 – Configurações do servidor NextJs	27
Tabela 5 – Estrutura de pastas básica do projeto e suas atribuições	29
Tabela 6 – Mapeamento de tags em código LaTeX	39
Tabela 7 – Mapeamento de escape de caracteres para código LaTeX	42
Tabela 8 – Mapeamento de pugins para código latex	44
Tabela 9 – Relação do level do título com o código latex	47

LISTA DE SIGLAS

ABNT Associação Brasileira de Normas Técnicas. 1, 2, 22

API Application Programming Interface. 17, 18, 21, 25

CSS Cascading Style Sheet. 10, 17, 20

DOM Document Object Model. 15

ECMA European Computer Manufacturers Association. 11

IES Instituição de Ensino Superior. 1, 2

NBR Norma Brasileira Regulamentadora. 1

PDF Portable Document Format. 2–5, 11, 22

PUC-GO Pontifícia Universidade Católica de Goiás. 22, 48

SSR Server Side Rendering. 14

SVG Scalable Vector Graphics. 10

TCC Trabalho de Conclusão de Curso. 2

UFPB Universidade Federal da Paraíba. 2

UI User Interface. 19, 44

UX User eXperience. 19

XML eXtensible Markup Language. 10

LISTA DE ABREVIATURAS

App Application. 29, 30

bash Bourne Again Shell. 26

CERN Conseil Européen pour la Recherche Nucléaire. 9

HTML HyperText Markup Language. 9, 10, 15, 18, 20, 22, 34, 38, 40, 43–45

JS JavaScript. 11, 12, 15, 17, 20, 25, 39, 46

JSON JavaScript Object Notation. 5, 12, 18

JSX JavaScript XML. 15, 30

LaTeX Lamport TeX. 5, 6, 22, 37–39, 41–45, 47, 49–51

MathML Mathematical Markup Language. 10

PHP Hypertext Preprocessor. 13, 14

RegExp Regular Expression. 22, 25

RegExp Regular Expression. 22

TS TypeScript. 11, 12

TSX TypeScript XML. 30, 31, 34

UUID Universally Unique Identifier. 49

uuid Universally Unique Identifier. 49, 50

W3C World Wide Web Consortium. 9, 10

Web World Wide Web. 8–11, 13, 16, 21, 26

XHTML eXtensible HyperText Markup Language. 10

XSS Cross-Site Scripting. 13

Sumário

Lista de ilustrações	i
Lista de tabelas	ii
Sumário	v
1 INTRODUÇÃO	1
1.1 Objetivo	2
1.2 Fluxo do documento	3
1.2.1 Escrita em blocos	3
1.2.2 Bloco	3
1.2.3 Parsing	4
1.3 Ambiente de desenvolvimento	5
1.3.1 Tecnologias do ambiente de desenvolvimento	6
2 FUNDAMENTAÇÃO TEÓRICA	7
2.1 Do Front-End	8
2.1.1 Tecnologias Web	8
2.1.1.1 <u>Linguagem de Marcação de Hipertexto, HTML</u>	9
2.1.1.2 <u>Funcionamento do HTML</u>	9
2.1.1.3 <u>HTML versão 5</u>	10
2.1.1.4 <u>Folhas de Estilo em Cascata, (CSS)</u>	10
2.1.1.5 <u>JavaScript</u>	11
2.1.1.6 <u>TypeScript</u>	11
2.1.1.7 <u>JavaScript Object Notation, JSON</u>	12
2.1.2 Bibliotecas e Frameworks	13
2.1.2.1 <u>ReactJs</u>	13
2.1.2.1.1 JavaScript XML	15
2.1.2.2 <u>NextJs</u>	16
2.1.2.3 <u>EditorJs</u>	18
2.1.2.4 <u>AntDesign</u>	19
2.2 Do Back-End	20
2.2.1 NodeJs	21
2.3 O processo de Parsing	22
2.3.1 Expressões regulares	22
2.3.1.1 <u>Expressão regular simples</u>	22

2.3.1.2	<u>Expressão regular com metacaractere</u>	23
2.3.1.3	<u>Expressões regulares em JavaScript</u>	24
2.3.2	Lamport Tex, LaTeX	25
2.3.3	Cheerio	25
3	DESENVOLVIMENTO	26
3.1	O servidor NextJs	26
3.1.1	Roteamento (App Router)	29
3.1.1.1	<u>Exemplo de criação de uma rota</u>	30
3.1.1.2	<u>Executando o exemplo</u>	32
3.1.2	Página principal	32
3.1.2.1	<u>Layout</u>	33
3.1.2.2	<u>Page</u>	34
3.2	EditorJS	35
3.2.1	Classe Editor	36
3.2.2	Provider	36
3.2.3	BasePlugin	36
3.2.4	InlineBasePlugin	36
3.2.5	Plugins	36
3.2.5.1	<u>Image</u>	37
3.2.5.2	<u>Header</u>	37
3.2.5.3	<u>FootNote</u>	37
3.3	Parsing, (parser)	37
3.3.1	Visão geral	39
3.3.2	Etapas de processamento	40
3.3.2.1	<u>Escape de caracteres</u>	41
3.3.2.1.1	Código do escape.ts	42
3.3.2.2	<u>Processamento de HTML</u>	43
3.3.2.2.1	Código do processamento HTML	44
3.3.2.3	<u>Pós processamento</u>	45
3.3.3	Plugins	45
3.3.3.1	<u>Tipagem</u>	46
3.3.3.2	<u>Paragraph, (parágrafo)</u>	47
3.3.3.3	<u>Header, (cabeçalhos)</u>	48
3.3.3.4	<u>Image, (imagens)</u>	49
3.3.3.5	<u>List, (listas)</u>	50
3.3.3.6	<u>Page Break, (quebra de página)</u>	51
3.3.3.7	<u>Table, (tabelas)</u>	51
3.3.4	Montagem	53
3.3.4.1	<u>Glossário</u>	53

3.3.4.2	<u>Referências</u>	53
	REFERÊNCIAS	54

1 INTRODUÇÃO

Escrever um trabalho científico pode ser uma tarefa desafiadora. (SEVERINO, 2017) destaca a complexidade e o rigor necessários na elaboração de trabalhos científicos, que não apenas envolvem o domínio do conteúdo específico, mas também a aderência às normas técnicas para apresentação formal e formatação correta.

A Associação Brasileira de Normas Técnicas (ABNT) , é a entidade responsável por, dentre outras, fornecer as normas que regulam o processo de criação de trabalhos acadêmicos. A Norma Brasileira Regulamentadora (NBR) Nº 14724, por exemplo: Especifica os princípios gerais para a elaboração de (teses, dissertações e outros), visando sua apresentação à instituição (banca, comissão examinadora de professores, especialistas designados e/ou outros) (ABNT, 2021).

Ademais, ainda com respeito aos trabalhos acadêmicos, não somente a regulamentação da NBR 14724 deve ser observada. Há ainda a NBR 6023 que trata a respeito da elaboração de referências e a NBR 10520, que diz respeito às citações em documentos.

(CASTRO, 2011), adverte que: "Em ciência, não pode haver uma separação entre forma e conteúdo. Trata-se de uma separação fictícia, pois fica se conhecendo o conteúdo pela forma." Ou seja: A forma do trabalho, sua apresentação, sua formatação e todo o seu arranjo gráfico é tão importante quanto seu conteúdo. (MEDEIROS, 2012) vai complementar essa visão, afirmando que a apresentação gráfica "[...] contribui para a consecução de um trabalho capaz de atingir seu objetivo. Monografia realizada sem a preocupação gráfica, em geral, acaba malsucedida."

Em seu artigo, (SILVA; VITORIA, 2014) vão analisar as percepções e dificuldades dos alunos de um curso superior em Tecnologia de Gestão em Recursos Humanos. Dentre suas dificuldades, (dos alunos em questão), é destacada a questão da formatação do trabalho acadêmico. Há também o fato de que as bancas avaliam os trabalhos baseadas em critérios da própria Instituição de Ensino Superior (IES), critérios estes que não estão necessariamente presentes nas normas da ABNT, ou seja, há uma subjetividade presente que não é comum a todas às IES quanto a questão da formatação. Essa subjetividade contribui para a confusão dos alunos, pois a IES avaliará de acordo com aquilo que julga apropriado, o que muitas vezes pode obscurecer o direcionamento do aluno ao redigir/formatar seu trabalho."

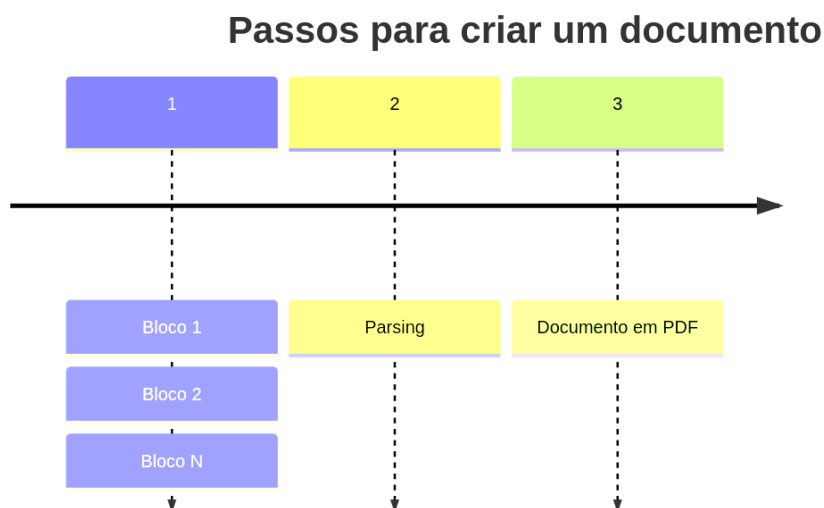
(SANTOS, 2020) em seu Trabalho de Conclusão de Curso (TCC) , também analisa as dificuldades encontradas por egressos, desta vez do curso de Ciências Contábeis da Universidade Federal da Paraíba (UFPB). Em sua pesquisa é destacado que "Quanto a formatação do trabalho com as normas da ABNT, [...], 60% teve alguma dificuldade, inclusive 32% teve muita dificuldade.", ou seja, a formatação do trabalho é um grande desafio presente na vida de boa parte dos estudantes em processo de escrita.

1.1 Objetivo

Levando em consideração os problemas que os alunos de diversas instituições de ensino enfrentam ao elaborar seus respectivos trabalhos (conforme apresentado acima), o objetivo deste instrumento é desenvolver uma plataforma web de alta interatividade¹ e inteligibilidade², de modo que o discente possa se preocupar apenas com o conteúdo. Os detalhes de formatação, de acordo com os padrões da ABNT e da IES, ficarão a cargo da própria plataforma.

A criação de um trabalho de TCC se dará basicamente por três passos básicos: Escrita em blocos; *Parsing*³ e Documento em PDF . A Figura1 ilustra esse fluxo na linha do tempo.

Figura 1 – Passo a passo para criar um documento na plataforma



Fonte: Autoria própria

- ¹ Refere-se à capacidade de um sistema, aplicação ou interface de responder às ações do usuário de maneira eficaz e intuitiva
- ² Refere-se à clareza e compreensibilidade da interface, documentação e feedback fornecidos pelo sistema. Um software inteligível facilita o entendimento do usuário sobre como utilizá-lo e quais são os resultados de suas ações.
- ³ O termo Parsing, (do inglês: análise), será utilizado no sentido de analisar e transformar algo em outra coisa.

O usuário interagirá com a aplicação escrevendo blocos que serão transformados no documento final em PDF. A este processo denominar-se-á Parsing. Após este, bastará enviar o download do PDF ao usuário com todo o padrão de formatação. Os trabalhos desenvolvidos nesta plataforma terão então duas versões: A versão de blocos, (sem formatação e interativa); e a versão final já formatada em PDF.

1.2 Fluxo do documento

1.2.1 Escrita em blocos

A escrita se dará de modo em que tudo será considerado um bloco. A escrita em blocos consiste numa abordagem em que o texto vai sendo escrito em seu fluxo natural, porém blocos podem ser adicionados à escrita. Um bloco é um elemento adicionado ao fluxo de trabalho que desempenha um papel que o diferencia dos demais blocos. Por exemplo: Uma imagem pode ser considerada um bloco nesta abordagem, uma vez que não é um texto mas tem o objetivo de fornecer informações visuais. O próprio corpo do texto em si será considerado um bloco, denominado parágrafo. Um título será um bloco textual cujo objetivo será separar sessões do texto coesas. Uma lista será um bloco para enumerar itens e assim por diante. O documento será basicamente uma composição de diversos blocos dispostos de forma a formar uma unidade coesa final, que será o trabalho propriamente dito.

1.2.2 Bloco

Um bloco é uma unidade lógica no documento que desempenha um papel especializado que nenhum outro bloco o faz. Por exemplo: O bloco mais importante da plataforma⁴ será o de texto, (denominado bloco parágrafo), pois sem texto, não há trabalho. Sem texto não há tão pouca comunicação que transmita informação de caráter acadêmico-científico.

Semelhante ao bloco de texto, diversos outros blocos adjacentes auxiliarão na construção do documento acadêmico. O bloco de imagem, por exemplo, ajuda a exibir informações de forma ilustrativa e auxilia bastante em exemplos que estão sendo dados em determinado contexto do texto.

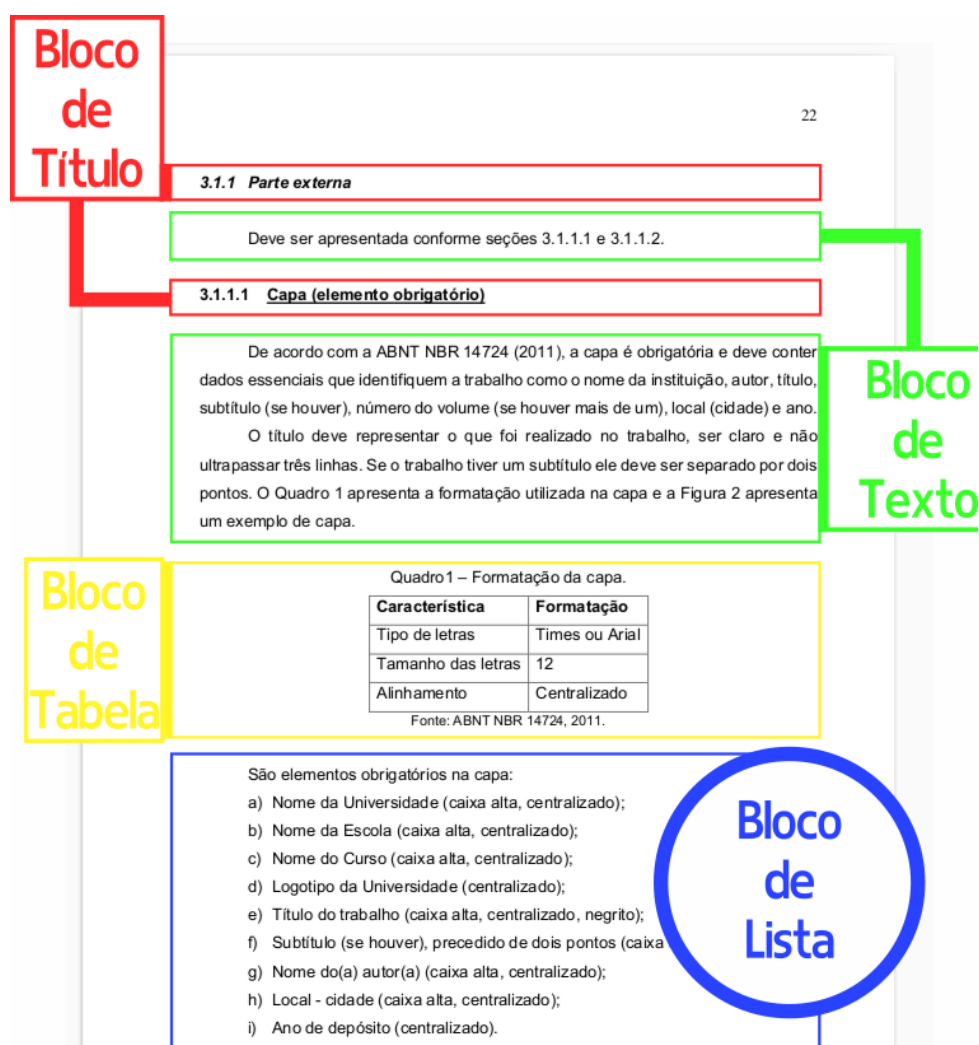
A maior parte dos blocos contará com uma espécie de submenu, (em termos de aplicação), que os permita personalizar. A personalização de blocos é importante para editar configurações e dar autonomia ao usuário em determinar mais precisamente o papel daquele bloco no texto. Por exemplo: Um bloco de título ajuda a separar o texto em unidades coesas. Porém, existem diversos tipos de títulos: Existe o título, o subtítulo, e até o subtítulo do subtítulo.

⁴ O termo plataforma será utilizado de forma intercambiável e como sinônimo de aplicativo; sistema web; ou aplicativo da web

O submenu será a configuração que o usuário fará no bloco após escolhê-lo. No caso do título, por exemplo: Após o usuário escolher este bloco, poderá configurar o nível de título desejado. Nível este que varia do 1 ao 4, sendo 2; 3 e 4 espécies de subtítulos. No caso de uma imagem, o submenu funcionará para que possa ser definida a imagem, bem como seu título de sua descrição.

A imagem abaixo ilustra a composição de um trabalho com seus respectivos blocos:

Figura 2 – Divisão de blocos em uma imagem



Fonte: Adaptado de (PUC-GO, 2022)

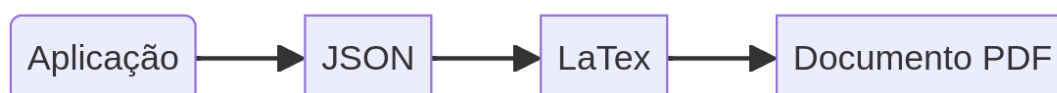
1.2.3 Parsing

O processo de Parsing é o processo que acontecerá sempre que o usuário desejar ver o *layout*⁵ da versão final de seu trabalho. Ele usa o código intermediário gerado pelos blocos para montar o PDF final.

⁵ Do inglês: Disposição, ou esboço. Esta palavra geralmente está associada ao desenho ou visual de algo.

Este processo é, em termos simples, uma espécie de análise a ser aplicada no código gerado pelos blocos da aplicação. A plataforma gerará um código JSON⁶ como resultado das interações do usuário, que posteriormente serão convertidos em código LaTeX⁷. Só então, finalmente será utilizado um utilitário que converterá o código LaTeX em um documento PDF. A Figura 3 ilustra esse processo:

Figura 3 – Etapas do processo de Parsing



Fonte: Autoria própria.

1.3 Ambiente de desenvolvimento

O ambiente de desenvolvimento é de extrema importância para que todas as ferramentas utilizadas possam funcionar em perfeita harmonia em suas respectivas integrações e colaborações. Muitas vezes, problemas de compatibilidade podem afetar o funcionamento das mesmas e impedir que o programa final seja executado corretamente, causando *bugs*⁸ e outros imprevistos impeditivos tanto para a correta execução, quanto para a experiência de desenvolvimento. A lista abaixo diz respeito às ferramentas e ao ambiente onde este *software* foi desenvolvido, bem como todas as suas respectivas versões:

⁶ Ver (sessão que trata do JSON)

⁷ Ver (sessão que trata do

⁸ Do inglês: Inseto. Esta palavra é muito utilizada no contexto de desenvolvimento de aplicativos para se referir a problemas que afetam o funcionamento dos mesmos

1.3.1 Tecnologias do ambiente de desenvolvimento

Atender aos requisitos mínimos de hardware e software é fundamental para garantir uma experiência de usuário satisfatória e evitar problemas de desempenho ou compatibilidade com o aplicativo da plataforma. A seguir enumera-se o ambiente mínimo com seus respectivos softwares necessários para rodar o aplicativo da plataforma:

Tabela 1 – Tecnologias do ambiente de desenvolvimento

Tipo	Tecnologia
Interpretador	NodeJs 20.10.0
Sistema Operacional	Ubuntu 20.04
Gerenciador de pacotes	Npm 10.2.3
Gerenciador de pacotes	Yarn 1.22.19
Utilitário	kpathsea version 6.3.4/dev
Linguagem de Programação	TypeScript 5.3.3
Utilitário	BibTeX 0.99d (TeX Live 2022/dev/Debian)
Navegador de Internet	Google Chrome 119.0.6045.199
Compilador	pdfTeX 3.141592653-2.6-1.40.22 (TeX Live 2022/dev/Debian)
Utilitário	makeglossaries (Utilitário LaTeX)

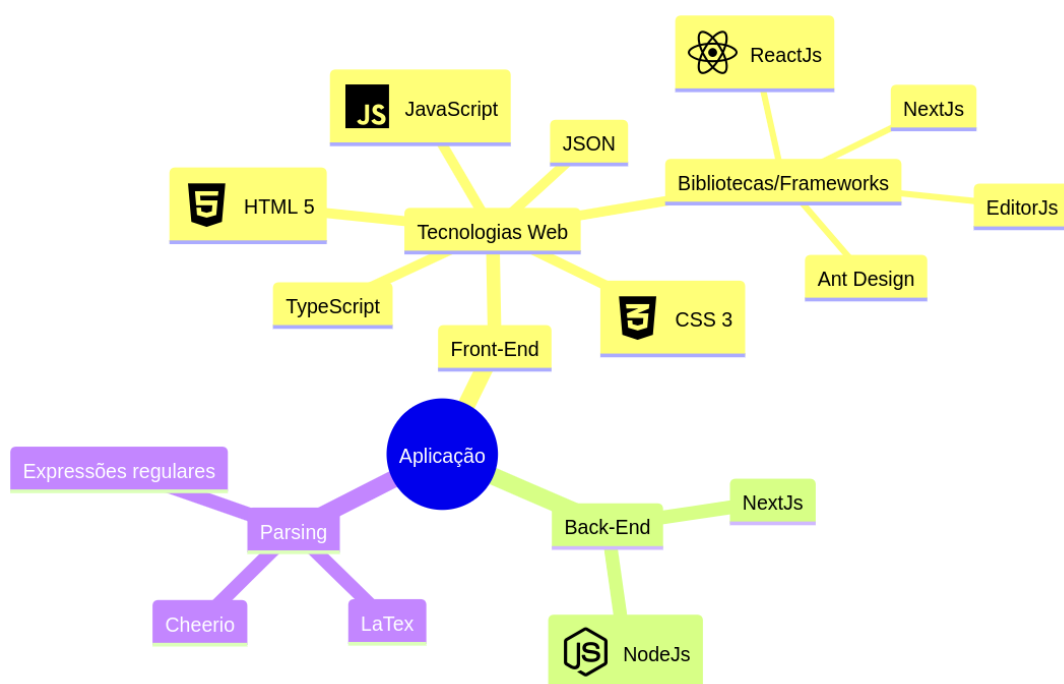
Fonte: Autoria própria

2 FUNDAMENTAÇÃO TEÓRICA

A plataforma será construída sob alguns pilares fundamentais indispensáveis a seu funcionamento. São estes pilares que garantirão o sucesso e o correto funcionamento da aplicação, afim de que todo o objetivo discutido até o presente momento seja atingido.

A Figura4 mostra em forma de mapa mental todos os principais pilares sobre os quais o aplicativo será contruído. Estes pilares são formados por diversas tecnologias, bibliotecas, *frameworks*¹ e conceitos que deverão trabalhar de forma integrada.

Figura 4 – Pilares da plataforma, (mapa mental)



Fonte: Autoria própria.

Estes pilares estão subdivididos em três grandes subcategorias, a saber: Front-End; Back-End e Parsing. Cada qual com seus respectivos conceitos e tecnologias.

¹ Uma framework é como um kit de ferramentas pré-pronto que fornece uma gama de funcionalidades pré-construídas e testadas afim de facilitar o processo de desenvolvimento. (AMAZON; AWS,)

2.1 Do Front-End

O Front-End é, basicamente, a "linha de frente". É a parte da aplicação que interagirá diretamente com o usuário. Ao profissional que codifica e desenvolve esta parte do projeto, denomina-se Desenvolvedor Front-End. A interface do usuário, que é onde o mesmo realiza suas interações com o sistema, normalmente é desenhada por um *designer*², ficando a cargo do desenvolvedor o papel de adaptar o *design*³ ao código afim de obter os efeitos desejados. (TOTVS, 2021)

2.1.1 Tecnologias Web

As tecnologias Web desempenham um papel crucial na criação de experiências digitais interativas, permitindo que os usuários se envolvam com o conteúdo de maneira mais dinâmica e significativa. A incorporação da *Internet*⁴ na vida diária resultou em mudanças significativas, marcada por um ritmo de evolução e aprimoramento sem precedentes, além da distribuição de conteúdo em massa. Juntamente com essas mudanças, surgiram novas tecnologias, variando de *softwares*⁵ a *hardwares*⁶, aprimorando a experiência de navegação na Web (MOLGADO, 2016).

A Internet, que teve origem nos Estados Unidos em 1969, foi inicialmente utilizada por universidades, governos e instituições financeiras antes de se expandir globalmente. No início, a internet era uma via de mão única onde os usuários consumiam informações e se comunicavam de maneira privada. A evolução começou com a introdução de sistemas de busca avançados, destacando-se o lançamento do Google em 1998, que democratizou o acesso à informação. (VITORIANO, 2019).

A grande reviravolta na internet aconteceu em 1999, com o surgimento do *Blogger*, marcando o início da Web 2.0, onde a comunicação tornou-se bidirecional. Os usuários passaram a gerar conteúdo e se relacionar publicamente com marcas, empresas e pessoas por meio de comentários, além de consumir informação. A evolução da tecnologia móvel, em conjunto com o surgimento de redes sociais como *Fotolog*, *MySpace*, *Orkut*, *Facebook*, *YouTube* e *Twitter*, ampliou o conceito de Web 2.0, permitindo o compartilhamento de fotos, vídeos e textos em uma escala maior. (VITORIANO, 2019).

A forma como se interage com a internet também evoluiu ao longo do tempo. Passou-se de sites estáticos para interativos e animados, chegando até aos sites total-

² Profissional que atua com design.

³ Do inglês: Desenho.

⁴ Rede mundial de computadores, (BRASIL, 2014).

⁵ O software é o conjunto de instruções dadas a um computador, de modo que ele execute determinada tarefa. Pode-se dizer que o software é a parte lógica do sistema computacional. (MACHADO,).

⁶ Com hardware, compreende-se o equipamento físico de um sistema computacional. Suas unidades Lógicas de Processamento, memórias e unidades de armazenamento são hardware. (MACHADO,).

mente responsivos⁷ e adaptáveis de hoje. Isso foi possível devido ao desenvolvimento de novos gadgets e ao surgimento de novas linguagens de programação. Atualmente, a Web Moderna é composta por várias técnicas, metodologias, linguagens e ferramentas que permitem o desenvolvimento de aplicações conectadas e interativas, oferecendo diversas formas de interação com interfaces digitais. (VITORIANO, 2019).

2.1.1.1 Linguagem de Marcação de Hipertexto, HTML

A Linguagem de Marcação de Hipertexto, do inglês: *HyperText Markup Language* (HTML) foi criada por Tim Berners-Lee enquanto trabalhava na Organização Europeia para a Pesquisa Nuclear (CERN), o laboratório de física de partículas na Suíça, no final dos anos 1980 e início dos anos 1990. O objetivo era criar uma maneira de compartilhar documentos e informações em um ambiente de rede. A primeira versão do HTML tinha apenas 18 elementos de marcação, permitindo a formatação básica de texto e a inclusão de *links*⁸, imagens e listas. (W3C, 2023).

O HTML rapidamente ganhou popularidade e passou por várias iterações, cada uma adicionando novos elementos e funcionalidades. O HTML4, lançado em 1997, trouxe uma série de melhorias, incluindo mais controle sobre a aparência das páginas web, a introdução de folhas de estilo em cascata (HTML) e melhor suporte a *scripts*⁹. (W3C, 2023).

Finalmente, o HTML5, lançado oficialmente em 2014 pelo *World Wide Web Consortium* (W3C), trouxe uma série de novas funcionalidades, incluindo suporte nativo para vídeo e áudio; novos elementos semânticos; gráficos e animações; geolocalização; armazenamento local e muito mais. (W3C, 2023).

2.1.1.2 Funcionamento do HTML

O HTML funciona como uma linguagem de marcação, o que significa que ele usa "*tags*"¹⁰ para definir diferentes partes de um documento. Essas tags informam ao navegador como exibir o conteúdo da página. Por exemplo, a tag <p> é usada para definir um parágrafo, enquanto a tag <h1> é usada para definir um cabeçalho de primeiro nível. (W3C, 2023).

As páginas HTML são estruturadas usando uma combinação de elementos de bloco, (que formam a estrutura principal da página), e elementos *inline*¹¹, (que forma-

⁷ A responsividade é a capacidade de uma página da Web em se adaptar a diferentes dispositivos e tamanhos de tela. (MDN, 2023b)

⁸ Do inglês: Ligação. Também chamado de hiperlink, é uma referência a um documento eletrônico que, quando clicado, leva o usuário para outro recurso ou documento.

⁹ Do inglês: Roteiro. Aqui usado no sentido de código fonte, que nada mais são do que um conjunto de instruções que o computador seguirá de modo interpretativo.

¹⁰ Do inglês: Marcação.

¹¹ Do inglês: Dentro da Linha. São elementos que podem ser escritos sem quebra de linha.

tam o conteúdo dentro desses blocos). Os elementos são aninhados dentro de outros elementos para criar a estrutura hierárquica da página. (W3C, 2023).

2.1.1.3 HTML versão 5

Com o HTML5, os desenvolvedores podem criar jogos online, reproduzir vídeos e áudios diretamente no navegador, tudo isso sem a necessidade de instalação de plugins externos. Isso resultou em uma melhor experiência geral do usuário, com carregamento mais rápido e maior compatibilidade entre os navegadores. (W3C, 2023).

Além disso, o HTML5 também trouxe recursos avançados de armazenamento local, como o *Web Storage*¹² e o *IndexedDB*¹³. Esses recursos permitem que os sites armazenem dados localmente no navegador do usuário, possibilitando a criação de aplicativos web *offline* e sincronização de dados em tempo real. (W3C, 2023).

Outra contribuição importante do HTML5 é o suporte a tecnologias de geolocalização e acesso aos recursos do dispositivo. Isso permite que os desenvolvedores acessem informações de localização do usuário, câmera, microfone e acelerômetro, abrindo possibilidades para o desenvolvimento de aplicativos web que utilizam esses recursos de forma integrada. (W3C, 2023).

Ao longo de sua história, o HTML tem evoluído constantemente para acompanhar as demandas e os avanços tecnológicos da web. O HTML5 é um marco significativo nessa evolução, trazendo recursos semânticos, multimídia e interativos para a criação de páginas da web modernas. (W3C, 2023).

2.1.1.4 Folhas de Estilo em Cascata, (CSS)

Folhas de Estilo em Cascata, ou *Cascading Style Sheets* (CSS), em tradução livre para o português, é uma linguagem de estilo altamente eficaz e amplamente utilizada. Sua principal função é definir a apresentação de documentos escritos em HTML ou XML. Isso inclui uma série de linguagens baseadas em XML, como SVG, MathML e XHTML. O CSS é responsável por descrever a forma como os elementos são apresentados em diferentes mídias, seja na tela do computador, em papel impresso, por meio de dispositivos de fala ou em outras formas de mídia. (MDN, 2023a).

Considerado uma das principais linguagens da Open¹⁴ Web, o CSS tem uma grande importância na padronização dos navegadores Web. Essa padronização é feita de acordo com as especificações estabelecidas pela W3C, a organização que lidera a Web mundial. O desenvolvimento do CSS é feito em níveis distintos: o CSS1, que hoje é considerado obsoleto; o CSS2.1, que atualmente é uma recomendação; e o CSS3, que está sendo dividido em pequenos módulos e caminha para a sua padronização. (MDN, 2023a).

¹² O termo Web Storage pode ser entendido, em tradução livre, como: Armazém da Web

¹³ Termo abreviado de "Indexed DataBase", (Base de Dados Indexada).

¹⁴ Do inglês: Aberto. Neste contexto, refere-se ao padrão "Aberto" da Web

2.1.1.5 JavaScript

JavaScript é uma linguagem de programação notavelmente versátil que, apesar de ser comumente conhecida pela sua utilização em páginas Web, vai muito além disso. Frequentemente abreviada para JS, essa linguagem é leve, interpretada e orientada a objetos com funções de primeira classe. Graças à sua flexibilidade, o JavaScript se expandiu para uma variedade de ambientes que não são navegadores, incluindo Node.js¹⁵, Apache CouchDB¹⁶ e Adobe Acrobat¹⁷, demonstrando sua adaptabilidade e eficácia em diversos contextos. (MDN, 2023c).

Com sua estrutura baseada em protótipos, o JavaScript é uma linguagem dinâmica que suporta múltiplos paradigmas de programação. Isso significa que, além de ser orientada a objetos, ela também suporta estilos de programação imperativos e declarativos, como a programação funcional. Essa capacidade de suportar diferentes estilos de programação torna o JavaScript uma ferramenta poderosa e flexível para os desenvolvedores. (MDN, 2023c).

O padrão para JavaScript é o ECMAScript. Desde 2012, todos os navegadores modernos oferecem suporte completo ao ECMAScript 5.1. Mesmo os navegadores mais antigos fornecem suporte, pelo menos, ao ECMAScript 3. A sexta versão do ECMAScript, oficialmente chamada de ECMAScript 2015 e inicialmente conhecida como ECMAScript 6 ou ES6, foi publicada pela ECMA International em 17 de junho de 2015. Desde então, as especificações do ECMAScript são lançadas anualmente, demonstrando o desenvolvimento contínuo e o avanço dessa linguagem padrão. (MDN, 2023c).

2.1.1.6 TypeScript

O TypeScript, as vezes abreviado como TS, é uma linguagem fortemente tipada construída em cima do JavaScript, (TYPESCRIPT,). Typescript traz uma sintaxe adicional para o JavaScript de modo que o mesmo possa suportar checagem de tipos estática. Sem o TS, fica difícil saber com quais tipos de dados estar-se a trabalhar durante o processo de desenvolvimento, pois o JavaScript é uma linguagem fracamente tipada. Os parâmetros das funções e variáveis não possuem nenhuma informação, forçando os desenvolvedores a recorrerem a todo momento à documentação ou intuir sobre as tipagens. Typescript resolve esse problema, permitindo tipar o código de modo que erros possam ser reportados quando a tipagem estiver incorreta, por exemplo: ao tentar-se passar uma string¹⁸ para uma função que espera um número,

¹⁵ Ver sessão que trata do Node.js

¹⁶ Base de dados que utiliza o JSON nativamente. Veja mais em:
<https://couchdb.apache.org/#about>

¹⁷ Software que lê e converte arquivos em formato PDF. Veja mais em:
<https://www.adobe.com/br/acrobat.html>

¹⁸ Do inglês: Corda, barbante ou fio. No contexto de programação, é usado como termo para cadeia de caracteres. O caractere é, na maioria das linguagens de programação, um tipo de dado. E textos

TypeScript lançará um erro. O JavaScript, por outro lado, permitirá a execução deste código podendo gerar erros de tempo de execução. (W3C, 2024).

O TypeScript possui um compilador, que nada mais é do que um transpilador. Este transpilador é responsável por transformar o código TS em JS. Desta forma, o código JavaScript resultante da transpilação pode ser rodado em praticamente qualquer navegador ou ambiente que suporte o JavaScript.

2.1.1.7 JavaScript Object Notation, JSON

JavaScript Object Notation, (Notação de Objeto JavaScript), popularmente chamado de JSON. É uma sintaxe para a serialização de objetos do javascript. Com objetos do javascript, compreende-se seus tipos de dados e valores, como: objetos; matrizes; números; strings; booleanos; *null*¹⁹ e *undefined*²⁰. Apesar de baseado na sintaxe do JavaScript, distingue-se desta no sentido da forma de escrita. A serialização é o processo de converter dados estruturados, (ou objetos), em um formato que pode facilmente ser armazenado e transmitido pela rede. O JSON basicamente converte os objetos JavaScript em strings. Uma característica deste, é que JSON é legível tanto por humanos, quanto por máquinas, (na maioria dos casos). (MDN, 2024a).

A Tabela2 mostra as principais diferenças entre o JavaScript e o JSON.

Tabela 2 – Diferenças entre o JavaScript e o JSON

Tipos e valores JavaScript	Diferença para o JSON
Objetos e Arrays	Os nomes das propriedades devem ser strings com aspas duplas; as vírgulas à direita são proibidas.
Números	Zeros à esquerda são proibidos; um ponto decimal deve ser seguido por pelo menos um dígito.
Strings	Apenas um conjunto limitado de caracteres pode ser escapado; certos caracteres de controle são proibidos; o separador de linha Unicode (U+2028) e o separador de parágrafo (U+2029) são permitidos; strings devem ter aspas duplas.

Fonte: (MDN, 2024a).

são formados por estas cadeias denominadas strings,
¹⁹ Do inglês: Nulo. Neste contexto é um valor especial do JavaScript para representar a nulidade de um objeto/variável.
²⁰ Do inglês: Indefinido. Neste contexto é um tipo de dado do JavaScript para variáveis indefinidas.

2.1.2 Bibliotecas e Frameworks

2.1.2.1 ReactJs

Figura 5 – Logotipo do React



Fonte: React Dev, disponível em: <https://react.dev/>

"Biblioteca para interfaces de usuário Web e nativas". O React é uma biblioteca de JavaScript criada pelo Facebook para solucionar desafios de manutenção e escalabilidade em suas aplicações. No início de 2011, a equipe de desenvolvedores do Facebook enfrentava dificuldades em lidar com o crescimento da aplicação de anúncios, que estava se tornando cada vez mais complexa e difícil de ser mantida. O aumento no número de membros da equipe e de funcionalidades estava afetando negativamente os processos da empresa. Com tantas atualizações em cascata, a aplicação estava se tornando lenta e difícil de ser atualizada sem falhas. (MORAIS, 2021).

Para resolver esses problemas, Jordan Walke, engenheiro do Facebook, propôs uma solução inovadora. Ele sugeriu levar o XHP, uma versão do PHP, para o navegador usando JavaScript. O XHP era uma tecnologia desenvolvida para minimizar ataques de Cross-Site Scripting (XSS) em aplicações Web dinâmicas. No entanto, ele não era capaz de lidar com o grande número de requisições necessárias para esse tipo de aplicação. Com o apoio de sua equipe de gerenciamento, Jordan Walke conduziu um experimento de seis meses para explorar essa ideia. O resultado desse experimento foi o surgimento do ReactJS. (MORAIS, 2021).

O ReactJS revolucionou o desenvolvimento de interfaces de usuário ao introduzir o conceito de componentes reutilizáveis e a abordagem de renderização virtual. Com a utilização de componentes, os desenvolvedores podiam criar e reutilizar peças de interface independentes e isoladas, o que simplificava o desenvolvimento e manutenção do código. Além disso, a renderização virtual permitia atualizações de interface eficientes, otimizando o desempenho da aplicação. O ReactJS foi lançado como um

software de código aberto em 2013, permitindo que desenvolvedores de todo o mundo o utilizassem em seus projetos. (MORAIS, 2021).

Desde então, o React ganhou uma imensa popularidade e se tornou uma das principais ferramentas para o desenvolvimento de interfaces de usuário em aplicações web. Sua abordagem declarativa, que permite descrever como a interface deve ser exibida com base no estado da aplicação, simplifica a construção de interfaces complexas. Além disso, a capacidade de reutilização de componentes economiza tempo e esforço durante o desenvolvimento. O React também influenciou o desenvolvimento do React Native, uma versão da biblioteca voltada para a criação de aplicativos móveis multiplataforma. Com a ajuda de uma grande comunidade de desenvolvedores e empresas, o ecossistema do React continua a evoluir e fornecer soluções inovadoras para o desenvolvimento de interfaces de usuário modernas e eficientes. (MORAIS, 2021).

O React teve seus primeiros sinais em 2010, quando o Facebook introduziu o XHP na sua stack de PHP, permitindo a criação de componentes compostos. Em 2011, Jordan Walke criou o FaxJS, protótipo inicial do React, que foi desenvolvido para resolver os desafios de suporte aos anúncios do Facebook. Em 2012, o Instagram foi adquirido pelo Facebook e expressou interesse em adotar o React. Isso levou o Facebook a dissociar o React da empresa e torná-lo open source. Em 2013, ocorreu o lançamento oficial do React, mas inicialmente enfrentou resistência da comunidade de desenvolvedores. No entanto, uma "turnê do React" foi realizada para conquistar os não adeptos. (MORAIS, 2021).

No ano seguinte, o React começou a ganhar reputação e confiança. O *React Developer Tools*²¹ e o *React Hot Reloader*²² foram lançados, trazendo melhorias no desenvolvimento e na experiência do usuário. Em 2015, o React se estabeleceu como uma tecnologia estável, com empresas como Netflix e Airbnb adotando-o. O Redux, responsável pelo gerenciamento de estado, foi lançado, e o React Native expandiu-se para o desenvolvimento de aplicativos móveis para Android. (MORAIS, 2021).

Atualmente, o React continua evoluindo, com o lançamento de novas funcionalidades e recursos para melhorar o desenvolvimento de aplicações. Iniciativas de SSR (*Server Side Rendering*)²³, e o foco em componentes funcionais são algumas das áreas de desenvolvimento. O React permanece como uma biblioteca consolidada no mercado de Front- End, sendo amplamente adotado por grandes empresas em todo o mundo. (MORAIS, 2021).

²¹ Ferramentas de Desenvolvimento React

²² Carregamento e recarregamento ultra rápido React

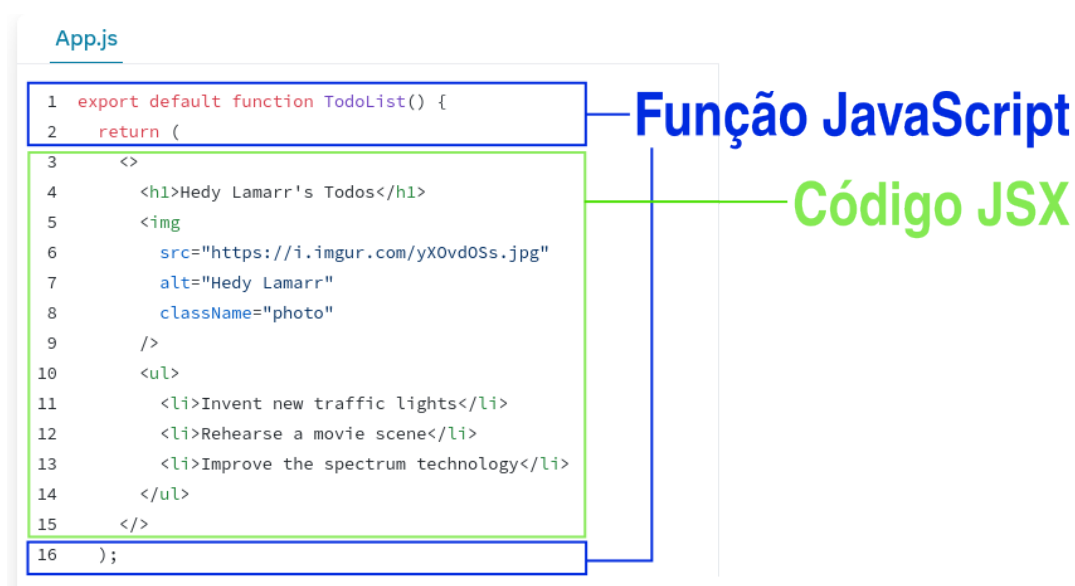
²³ Do inglês: Rederização do Lado do Servidor

2.1.2.1.1 JavaScript XML

O JavaScript XML, (JSX), é uma extensão de sintaxe para o JavaScript que permite escrever código de marcação, (como o HTML), dentro do JavaScript. Os componentes React nada mais são do que funções ou classes chamadas a partir do código JS que atualizam o HTML através da *Document Object Model*²⁴, (DOM). O JSX facilita esse trabalho. Pois ao invés de chamar funções ou instanciar classes no código JavaScript, pode-se escrever a marcação diretamente no mesmo, como se o HTML estivesse dentro do JS. (REACTJS, 2024).

A Figura6 mostra um exemplo de código em JSX. É exportada uma função denominada *ToDoList*, que nada mais é do que um componente React. Esta função retorna um determinado valor, que é o código JSX propriamente dito. Este código será renderizado no lugar onde esta função for chamada.

Figura 6 – Exemplo de código JSX

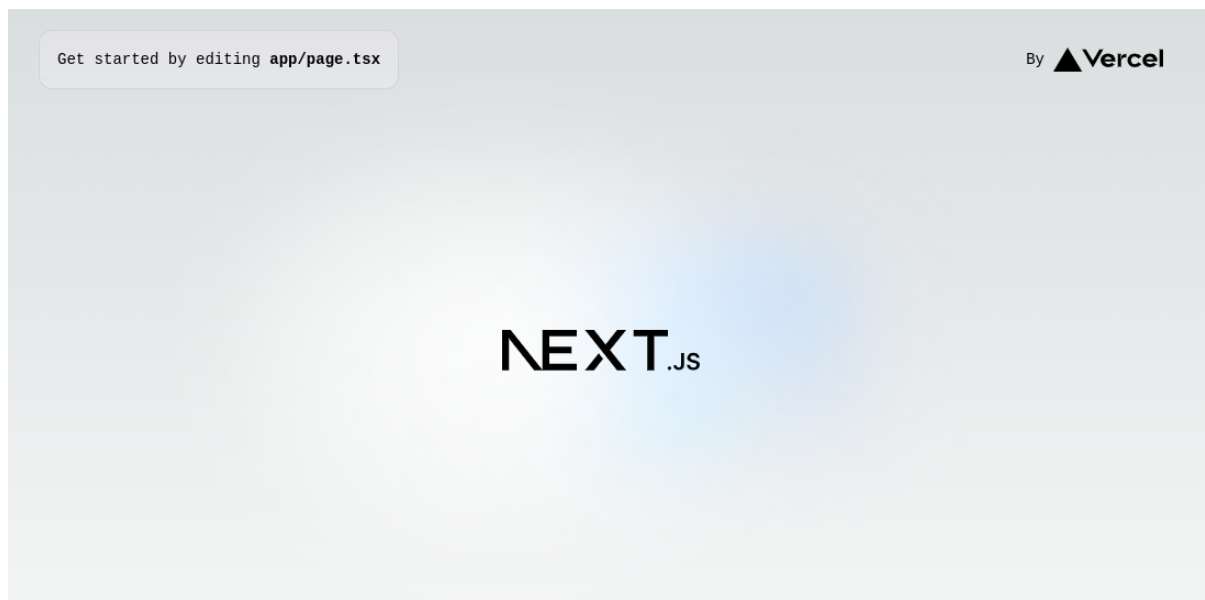


Fonte: Adaptado de: ReactJs. Disponível em:
<<https://pt-br.react.dev/learn/writing-markup-with-jsx>>

²⁴ Do inglês: Modelo de Documento de Objeto. O DOM é utilizado pelo JavaScript para manipular o documento HTML exibido em tela. (PIMENTEL, 2022)

2.1.2.2 NextJs

Figura 7 – Logotipo do NextJs



Fonte: Next.Js, disponível em: <https://nextjs-template.vercel.app/>

O NextJs é uma framework em ReactJs voltada à construção de aplicações Web tando na parte do Front-End quanto no Back-End. Com NextJs, utiliza-se os componentes em React para construir as interfaces de usuário, com o NextJs provendo recursos adicionais e otimizações. (NEXTJS, 2024).

NextJs também se encarrega de todas as configurações necessárias do React, como o processo de enpacotamento²⁵, compilação e etc... Permitindo ao desenvolvedor apenas focar no desenvolvimento da aplicação em si. (NEXTJS, 2024).

Devido à natureza desta framework, O NextJs é um pilar que aparece tanto no Back-End quanto no Front-End. Estas duas frentes serão abordadas com a utilização desta ferramenta, aproveitando ao máximo os recursos fornecidos pela mesma. Os principais recursos oferecidos pelo NextJs são:

- Roteamento: O NextJs provê um roteamento, (que é basicamente a navegação por páginas dentro do app), baseado no sistema de arquivos do Sistema Operacional. Os arquivos em pastas do projeto são mapeados para links, que fornecem os componentes de servidor com suporte a layouts²⁶, rotas aninhadas, estados de carregamento, manipulação de erros, entre outros...

²⁵ Em inglês: Bundling. Um Bundle para a Web, por exemplo, junta todos os códigos e recursos em um pacote otimizado para ser distribuído.

²⁶ Layouts são como templates que são comuns às páginas roteadas. Ajudam no processo de reaproveitamento de componentes pois eles podem ser estendidos às páginas, que herdaram características destes layouts.

- Renderização: NextJs fornece renderização do lado do cliente e do lado do servidor com componentes de cliente, e componentes de servidor.
- Busca de dados: Há um processo de busca de dados simplificado com o uso de *async/await*²⁷ nos componentes de servidor, além de uma API²⁸ expandida para a memorização das requisições, *caching*²⁹ de dados e revalidação.
- Estilização: Suporte para os métodos preferidos de estilização. Com inclusão de: Módulos CSS, *Tailwind* CSS, e CSS-in-JS.
- Otimizações: Otimizações de scripts, imagens e fontes são também fornecidos para aprimorar o núcleo do aplicativo e a experiência de usuário.
- TypeScript: Suporte total ao TypeScript, com uma melhor checagem de tipos e compilação eficiente.

(NEXTJS, 2024).

²⁷ Recurso do JavaScript para lidar com a execução de código assíncrono.

²⁸ Do inglês: Interface de Programação de Aplicações. É uma forma na qual dois ou mais aplicativos ou componentes de de computador se comunicam entre si. É uma interface de software que oferece um serviço para outras partes do mesmo ou de outros softwares. (REDDY, 2011).

²⁹ O processo de caching é o ato de armazenar informações que são acessadas frequentemente de maneira que seu acesso se torne mais rápido. Neste contexto, o resultado de uma requisição pode ser armazenado em cache para que não seja necessário consultar o servidor novamente quando a mesma informação for requisitada.

2.1.2.3 EditorJs

Figura 8 – Logotipo do EditorJs



Fonte: Editor.Js, disponível em: <https://editorjs.io/>

"Editor livre em blocos com saída universal em JSON". O EditorJs é um rico editor de texto em blocos que oferece uma experiência de edição intuitiva e versátil. Tudo o que é feito no EditorJs no fim é transformado em um arquivo JSON ao invés de um documento de marcação em HTML. Essa abordagem deixa o processo mais simples para os desenvolvedores no sentido de projetarem suas próprias integrações. Assim, o EditorJs pode ser aplicado a diversas plataformas. (EDITORJS,).

São recursos do EditorJs:

- Dados de saída limpos
- API baseada em plugins
- Código aberto

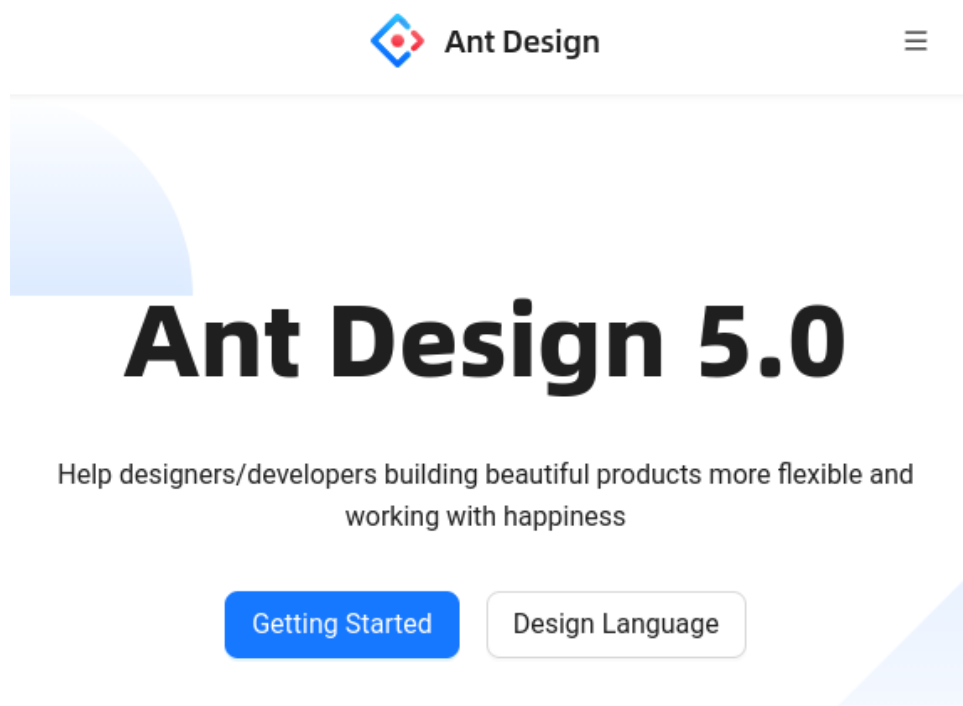
O espaço de trabalho do EditorJs consiste em blocos separados, como: Parágrafos; títulos; listas; etc... Cada um deles independentes entre si, com muitos outros recursos como: Copiar e colar; seleção de vários blocos; e entre outros que funcionam de forma familiar a outras ferramentas. (EDITORJS,).

O conceito chave do EditorJS é sua API, na qual todas as unides funcionais do editor são providas através de plugins externos que fazem uso da mesma. Assim, o núcleo do EditorJs fica sendo mais abstrato e poderoso, de modo que o desenvolvedor

possa implementar diversos desafios com a criação de seus próprios plugins. (EDITORS,).

2.1.2.4 AntDesign

Figura 9 – AntDesign



Fonte: AntDesign, Disponível em: <<https://ant.design>>.

"Ajudando designers e desenvolvedores a construir belos produtos de forma flexível, trabalhando com alegria.". O AntDesign é uma biblioteca de UI, *User Interface*³⁰. Esta biblioteca é uma grande auxiliadora na hora de produzir aplicativos com um design agradável sem que o desenvolvedor gaste muito tempo estilizando os componentes da aplicação.

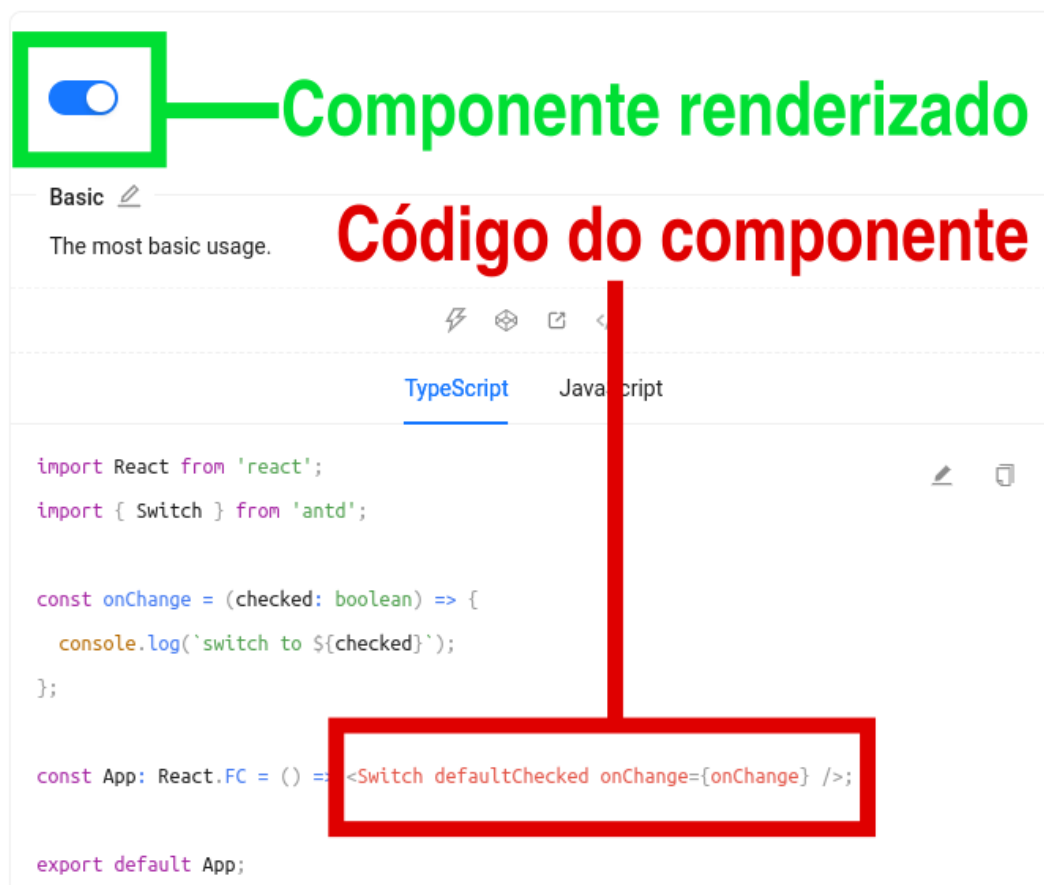
Totalmente integrado ao ReactJs, o AntDesign já traz consigo uma vasta gama de componentes reutilizáveis que podem ser "encaixados" à construção das telas do app. O AntDesign não impacta somente a UI, mas também a *User Experience*³¹, UX, uma vez que seus componentes são construídos em cima do React. Deixando-os altamente reativos e já com seus devidos comportamentos padrão. Ficado a cargo do desenvolvedor apenas personalizar o que se deseja fazer. A Figura10 mostra um exemplo de implementação de um componente em AntDesign.

³⁰ Do inglês: Interface de Usuário

³¹ Do inglês: Experiência de Usuário

Figura 10 – Exemplo de componente em AntDesign

Examples



Fonte: Adaptado de: AndDesign. Disponível em:
<<https://ant.design/components/switch>>

Nota-se que com apenas uma linha tem-se um componente completo com um visual totalmente moderno. Algo que com apenas HTML, CSS e JS, gastaria algumas dezenas de linhas de código. Tanto ReactJs quanto AntDesign significam economia em termos de tempo, código e simplicidade de projeto.

2.2 Do Back-End

Back-end se relaciona com o que está por trás das aplicações desenvolvidas na programação. Ou seja, tudo que dá estrutura e apoio às ações do usuário da máquina é chamado de back-end. Quando se acessa um site, por exemplo, por trás de toda sua apresentação amigável esteticamente, há uma comunicação das informações trocadas entre banco de dados e navegador. Portanto, por trás da interface gráfica do realizador, o back-end está sempre agindo. (TOTVS, 2020).

2.2.1 NodeJs

Figura 11 – Logotipo do NodeJs



Fonte: Adaptado de: NodeJs. Disponível em: <<https://nodejs.org/en>>

"Rode JavaScript em todo lugar". O NodeJs é um ambiente de *runtime*³² que permite ao desenvolvedor criar servidores, aplicativos Web, ferramentas e linha de comando, entre outros... É a principal tecnologia deste projeto e o que permitirá a execução de todas as outras na sequência.

Antes do NodeJs, o JavaScript era uma linguagem que rodava puramente em *Browsers*³³ como uma forma de adicionar interações às páginas da internet. Com o NodeJs, o JavaScript passou do ambiente dos Browsers ao ambiente dos Sistemas Operacionais. Abstraindo APIs dos mesmos. Hoje, com NojeJs, por exemplo, pode-se acessar a API do sistema de arquivos do Sistema Operacional. Algo que há um tempo atrás só se fazia com linguagens de baixo nível como o C++.

A plataforma desenvolvida neste trabalho será um servidor que distribuirá o Sistema Web que é a plataforma em si. O NextJs, tecnologia escolhida para este fim, é um servidor capaz de processar uma parte das interfaces do sistema estaticamente, e entregá-las ao cliente juntamente com os scripts de suas partes dinâmicas. Desta forma, ganha-se segurança no processamento back-end ao mesmo tempo que não perde-se em termos de interatividade com o usuário. Há também diversos ganhos

³² Do inglês: Tempo de execução. Um runtime é basicamente um interpretador capaz de executar um script.

³³ Do inglês: Navegadores. Aqui usado no sentido de navegador da Web, ou seja, o aplicativo no qual acessa-se páginas na internet.

de performance e simplificação de código, uma vez que back-end e front-end serão concentrados no mesmo lugar, não precisando separá-los em projetos diferentes.

2.3 O processo de Parsing

O processo de Parsing é uma das partes mais vitais deste projeto. Sem ele, não é possível obter o documento final formatado de acordo com as normas postas da ABNT e da PUC-GO. Pode-se dizer que o Parsing é o código núcleo da aplicação, pois todas as outras partes, como edição em blocos e navegação, por exemplo, serão feitas com o auxílio de bibliotecas e frameworks. O Parsing, por sua vez, será escrito puramente em TypeScript para processar as saídas do EditorJs.

Seu tratamento utilizar-se-á de uma combinação de expressões regulares para tratar os caracteres especiais de código LaTeX, manipulação de código HTML produzido pelos plugins do editor utilizando-se o Cheerio. E por fim, a compilação de código LaTeX utilizando-se o utilitário pdflatex para gerar o PDF.

2.3.1 Expressões regulares

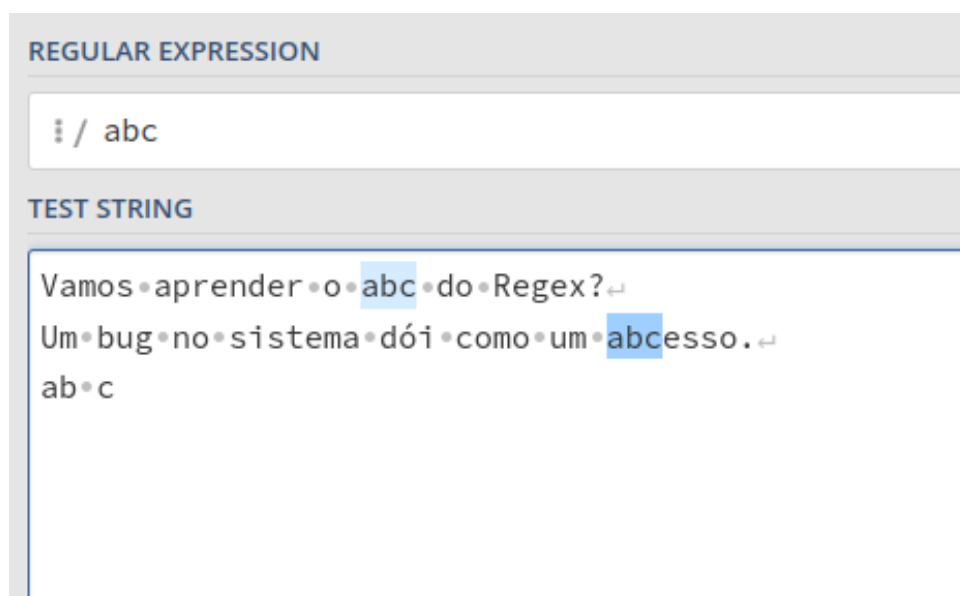
As expressões regulares, (as vezes denominadas RegEx ou RegExp), podem ser descritas basicamente como uma sequência de caracteres que descrevem um padrão de busca em um texto. (MAGALHÃES, 2022).

Uma expressão regular pode ser composta por uma cadeia simples de caracteres, como **abc**; ou uma cadeia composta pela combinação de caracteres simples e caracteres especiais. Aos caracteres especiais, denominam-se metacaracteres. Os metacaracteres são usados quando a busca no texto requer algo mais do que uma simples correspondência direta. (MDN, 2024b).

2.3.1.1 Expressão regular simples

No exemplo anterior, a expressão simples **abc** irá corresponder a uma letra **"a"**, seguida de um **"b"**, seguida de um **"c"** no texto ao qual se está avaliando. Estas letras terão de estar juntas e nesta exata ordem. Esta expressão irá encontrar correspondências nas strings "Vamos aprender o abc do Regex?" e "Um bug no sistema dói como um abcesso.". Nestes dois casos, houve uma correspondência da substring **abc**, mas a string "ab c" não irá obter correspondência, pois aqui a exata substring **abc** não está contida. Observe as correspondências na Figura12. (MDN, 2024b).

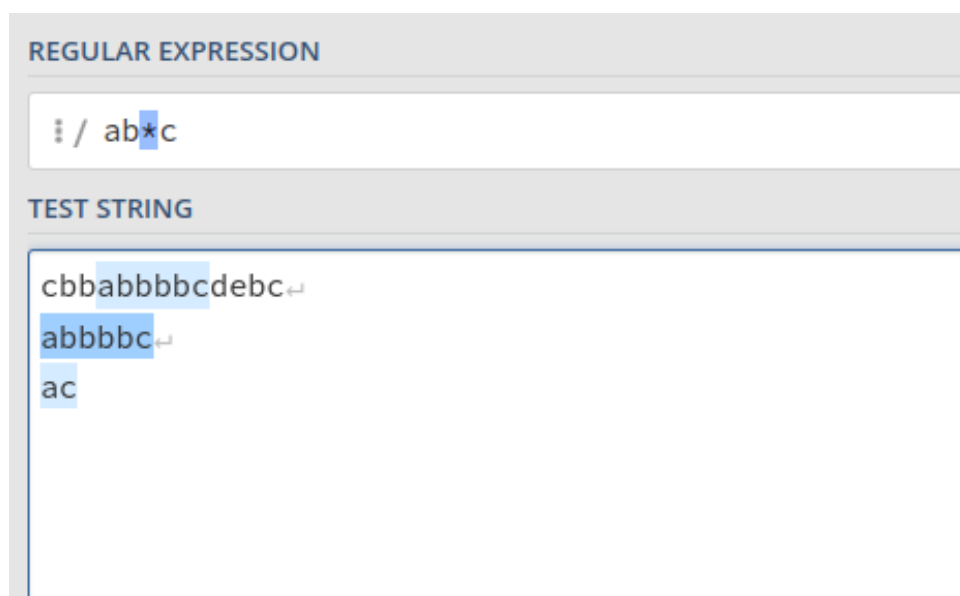
Figura 12 – Correspondências do Regex abc



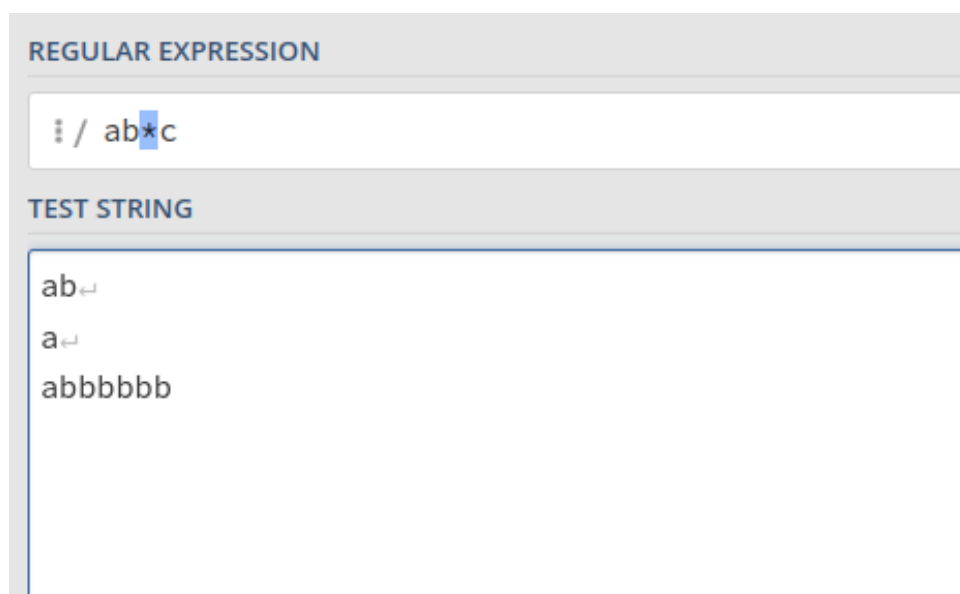
Fonte: Autoria própria

2.3.1.2 Expressão regular com metacaractere

Como mencionado anteriormente, quando se quer algo que é mais do que uma simples correspondência direta, utilizam-se metacaracteres. O caractere `*` é um metacaractere de expressão regular. Por exemplo: **ab*c** utiliza o `*` como metacaractere. A expressão **ab*c** deve ser lida como: Corresponda ao caractere **a**, seguido por zero ou mais caracteres **b**s, seguido por um caractere **c**. Usar `*` após o **b** significa: zero ou mais ocorrências do item anterior, no caso **b**. Como observa-se na Figura13 as strings "cbbabbbbbcdebc", "abbbbc" e "ac" encontrarão correspondências. As strings "ab", "a" e "abbbbbbb", por outro lado, não encontrarão correspondências, conforme a Figura14. (MDN, 2024b).

Figura 13 – Correspondências do Regex `ab*c`, com metacaractere

Fonte: Autoria própria

Figura 14 – Não correspondências do Regex `ab*c`, com metacaractere

Fonte: Autoria própria

Uma expressão regular é uma combinação de caracteres, alguns deles, como o metacaractere quantificador `*`, visto anteriormente, são caracteres especiais. Estes metacaracteres são símbolos especiais que definem como a é interpretada. (MAGALHÃES, 2022).

2.3.1.3 Expressões regulares em JavaScript

Em JavaScript, as expressões regulares podem ser escritas diretamente dentro do código, desde que postas entre barras `//`. Por exemplo: `/abc/` é uma expressão

regular válida em JS. O tipo de dados das expressões regulares é o *object*³⁴, assim como *Array*³⁵ ou *Set*³⁶. As expressões regulares no JS são usadas com dois objetos principais, a saber RegEx e String. A API de strings do JS fornece uma gama de métodos nos quais se pode utilizar juntamente com RegEx. A Tabela3 fornece uma descrição destes métodos:

Tabela 3 – Métodos de string que podem ser usados com regex

Método	Descrição
match	O método match retorna o resultado da correspondência de um dado (MDN, 2024b) à string ao qual está sendo aplicado. Aplicação em código: "Meu abc".match(/abc/);
matchAll	Faz o mesmo que match. Porém, ao contrário de match, que traz apenas a primeira correspondência, matchAll traz todas as correspondências encontradas.
replace	O método replace de uma dada string, retorna uma nova string que substitui a correspondência de uma expressão regular por alguma string de substituição.
replaceAll	Um caminho diferente para fazer o mesmo que replace, porém para todas as ocorrências.
search	Faz uma busca pelo padrão na string e retorna o índice da primeira ocorrência.
split	Divide a string em uma lista ordenada de substrings em que o critério de divisão é a expressão regular fornecida.

Fonte: (MDN, 2024b).

O objeto RegEx também fornece dois métodos em que se usam as expressões regulares, a saber: `exec()` e `test()`. (MDN, 2024b).

2.3.2 *Lamport Tex, LaTeX*

2.3.3 *Cheerio*

³⁴ Do inglês: Objeto

³⁵ Do inglês: Variedade ou matriz. No contexto de JavaScript é um dado estruturado, composto de uma sequência de outros objetos ou dados puros.

³⁶ Do inglês: Conjunto. Assim como Array, é um dado estruturado que agrupa outros objetos, porém de forma não ordenada e não indexada.

3 DESENVOLVIMENTO

A base da aplicação se dará por meio de um servidor onde serão realizadas todas as operações. Este distribuirá o conteúdo estático e dinâmico para a interação com o usuário, além também de fornecer rotas e funções para o processamento do conteúdo gerado pelo usuário. Por se tratar de uma aplicação em seu estágio inicial de concepção, este projeto não se preocupará com autenticação e autorização de usuários. Consistirá apenas de um servidor simples a ser rodado localmente a fim de se desenvolver suas funcionalidades principais.

3.1 O servidor NextJs

O servidor da aplicação será construído em cima do NextJs, a framework React para a Web. Iniciar o projeto é uma tarefa simples. Basta apenas navegar para algum diretório onde se deseja criar o projeto e digitar o seguinte comando em bash¹:

```
1 npx create-next-app@latest
```

Vale ressaltar que é necessário ter o NodeJs na versão 18.17 ou superior para iniciar o projeto em NextJs. Após rodar este comando, o prompt fará uma série de perguntas para a configuração do mesmo, observe o exemplo abaixo:

```
1 What is your project named? my-app
2 Would you like to use TypeScript? No / Yes
3 Would you like to use ESLint? No / Yes
4 Would you like to use Tailwind CSS? No / Yes
5 Would you like to use 'src/' directory? No / Yes
6 Would you like to use App Router? (recommended) No / Yes
7 Would you like to customize the default import alias (@/*)? No / Yes
8 What import alias would you like configured? @/*
```

Os itens a serem configurados são:

1. Nome do projeto

¹ Bash (*Bourne Again Shell*): Interface de linha de comando do linux

2. Será feito o uso de TypeScript?
3. Será feito o uso de ESLint?
4. Será feito o uso de Tailwind CSS?
5. Será usado o diretório 'src' como diretório padrão de código?
6. Será utilizado o App Router (Roteador de App)?
7. Deseja customizar o apelido padrão de importação (@/*)?
8. Qual apelido de importação gostaria de configurar?

A tabela Tabela4 mostra as opções escolhidas para este projeto:

Tabela 4 – Configurações do servidor NextJs

Pergunta	Resposta
What is your project named? my-app	editor2
Would you like to use TypeScript? No / Yes	Yes
Would you like to use ESLint? No / Yes	No
Would you like to use Tailwind CSS? No / Yes	No
Would you like to use 'src/' directory? No / Yes	Yes
Would you like to use App Router? (recommended) No / Yes	Yes
Would you like to customize the default import alias (@/*)? No / Yes	Yes
What import alias would you like configured? @/*	@/*

Os comandos a seguir instalam todas as dependências necessárias do projeto:

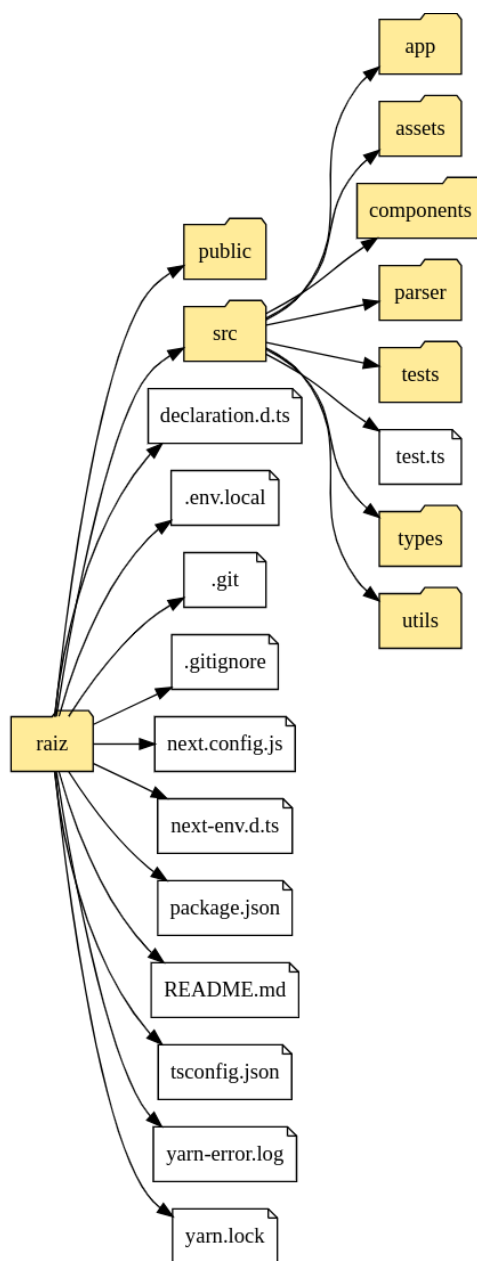
```

1 yarn add @editorjs/editorjs @editorjs/header @editorjs/list
2 yarn add -D @types/editorjs__header @types/node @types/react
3 yarn add @emotion/react @emotion/styled antd cheerio
4 yarn add -D @types/react-dom @types/uuid typescript
5 yarn add node-latex react-icons uuid

```

Após isso o utilitário de criação da aplicação NextJs irá criar uma estrutura de pastas e arquivos de projeto totalmente configurado e pronto para ser programado. A Figura15 mostra a estrutura de pastas base do projeto, com todos os seus respectivos arquivos de configuração, roteamento e componentes básicos do react.

Figura 15 – Estrutura de pastas e arquivos básica do projeto



Fonte: Autoria própria

Tabela 5 – Estrutura de pastas básica do projeto e suas atribuições

Arquivo/Pasta	Descrição
public	Pasta pública para distribuição de arquivos estáticos.
src	Esta pasta é onde praticamente todas as coisas na aplicação vão acontecer, aqui está as páginas e seus devidos roteamentos, o editor e seus plugins, o parser e tudo mais.
declaration.d.ts	Arquivo de declarações. Aqui é anotado algumas tipagens para serem usadas globalmente durante o processo de desenvolvimento.
.env.local	Arquivo de variáveis de ambiente para serem usadas como teste durante o tempo de desenvolvimento
.git	Pasta de controle do git.
.gitignore	Arquivos a serem ignorados pela ferramenta de versionamento.
next.config.js	Arquivo de configurações do NextJs
next-env.d.ts	
package.json	Arquivo que define que o projeto é um projeto NodeJs. Aqui está toda a informação sobre dependências do projeto, que são todos os pacotes usados de terceiros. Também possui definição de scripts úteis para serem utilizados no processo de desenvolvimento.
README.md	Documentação de apresentação do projeto.
tsconfig.json	Configurações do TypeScript.
yarn-error.log	Erros do gerenciador de pacotes yarn
yarn.lock	Arquivo de lock de dependências do gerenciador de pacotes yarn.

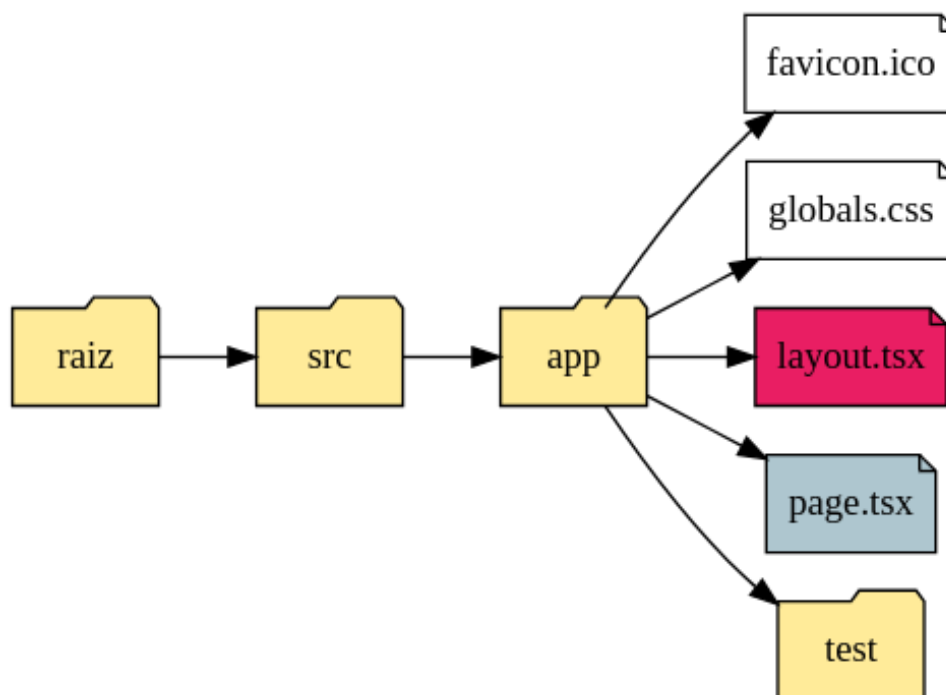
Fonte: Autoria própria

3.1.1 Roteamento (App Router)

No ato de configuração do servidor, conforme mostrado na Tabela4, na pergunta: *"Would you like to use App Router? (recommended) No / Yes"*, foi escolhida a opção sim, "yes". Isto significa que no código fonte do projeto haverá uma pasta chamada App onde serão arquivados os roteamentos, as páginas, layouts, páginas de erro, loading, entre outros.

A Figura16 mostra a estrutura da pasta App.

Figura 16 – Estrutura da pasta de roteamento (App Router)



Fonte: Autoria própria

Observe os arquivos destacados em azul e rosa, respectivamente `page.tsx` e `layout.tsx`. A extensão TSX² significa o mesmo que JSX, com a diferença de que é um arquivo em TypeScript ao invés de JavaScript.

Os arquivos `page` e `layout` são a primeira página da aplicação. Por estarem no nível do diretório App (dentro da pasta `app`) esta página será mapeada para a rota `/` na navegação. Isto significa que quando o servidor estiver rodando e o usuário acessar o endereço do serviço, o código contido em `page.tsx` renderizará o conteúdo da página para o usuário.

O arquivo `layout.tsx` serve como um template para a página. Pode-se pensá-lo como uma espécie de casca da aplicação. A página herda o que está em `layout`, de modo que o que será renderizado no Browser será o template, acrescido da `page`.

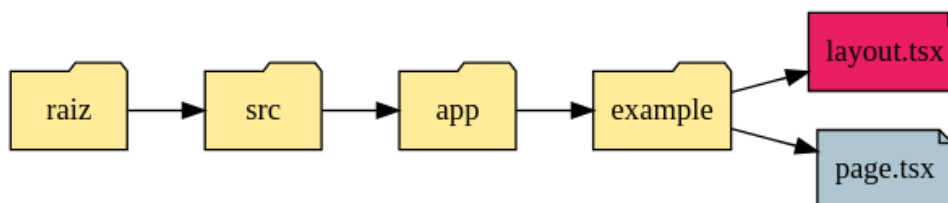
Por estar no diretório raiz de `app`, todas as rotas de páginas da aplicação herdarão o que está em `layout.tsx`. Arquivos de `layout` também podem ser escritos em rotas que não sejam a raiz.

3.1.1.1 Exemplo de criação de uma rota

Para criar uma rota para uma nova página no NextJs, basta que uma pasta no diretório `app` seja criada. Esta pasta deve conter pelo menos um arquivo `page.tsx` ou `route.ts`. Observe a Figura 17.

² TypeScript XML

Figura 17 – Estrutura da pasta de roteamento (App Router)



Fonte: Autoria própria

Ao criar a pasta `example` com o arquivo `page.tsx`. O servidor mapeará para a rota `/example` no Browser, renderizando o componente exportado em `page`. Observe a seguir o código em `layout.tsx`:

```
1 import { PropsWithChildren } from "react";
2
3 export default function Layout(
4   { children }: PropsWithChildren
5 ){
6   return <section>
7     <h1>Eu sou o Layout</h1>
8     { children }
9   </section>
10 }
```

Observe que o arquivo exporta uma função denominada `Layout`, que retorna um código TSX a ser renderizado na página. Observe na linha 4 que a função recebe um objeto como parâmetro que é anotado por `PropsWithChildren`. Dentro deste objeto há a propriedade `children`³, que contém outros componentes React a serem renderizados. No caso, o NextJs injetará dentro de `children` o código que é exportado por `page.tsx`. A linha 8 é a posição onde este componente será renderizado.

Observe abaixo o código de `page.tsx`:

```
1 export default function Page(){
2   return <div>
3     Olá mundo!
4     Eu sou uma page.
5   </div>
6 }
```

No código `page` apenas é exportada uma função (componente) `Page` que retorna uma frase simples: "Olá mundo! Eu sou uma page".

³ Do inglês: Filhos.

3.1.1.2 Executando o exemplo

Para executar o exemplo e ver como ele se comporta no navegador do usuário, basta a partir do diretório do projeto rodar o seguinte comando:

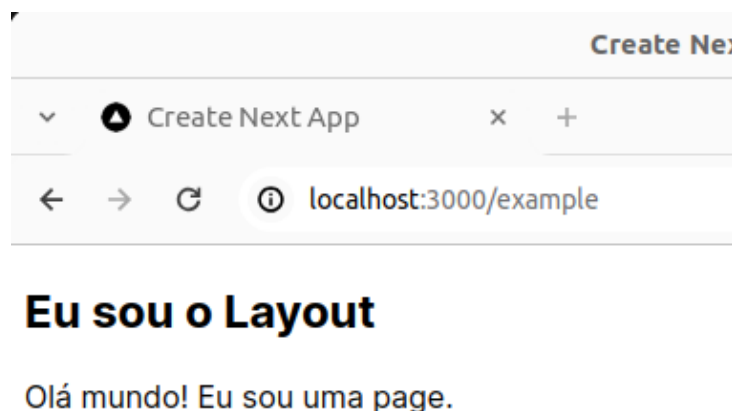
```
1 yarn dev
```

Após rodar o comando o servidor iniciará localmente. A saída do terminal mostrará a porta do serviço a ser acessada afim de renderizar a página:

```
1 yarn run v1.22.19
2 $ next dev
3   Next.js 14.1.1
4   - Local:      http://localhost:3000
5   - Environments: .env.local
6
7   Ready in 2.7s
```

Ao acessar o endereço <http://localhost:3000> na rota `/example`, a página é renderizada conforme a Figura18:

Figura 18 – Exemplo de roteamento



Fonte: Autoria própria

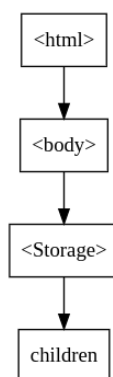
3.1.2 Página principal

Neste primeiro momento, o editor em que o usuário interagirá será renderizado na página principal. Conforme dito anteriormente e ilustrado na Figura16, a página principal está contida ao nível da pasta `app`, juntamente com seu layout.

3.1.2.1 Layout

O arquivo de layout é o template base no qual todas as outras páginas herdarão. Observe na Figura19 a sub árvore de renderização do layout:

Figura 19 – Sub árvore de renderização do layout principal



Fonte: Autoria própria

Tem-se uma estrutura básica com a tag html, (a tag raiz do documento), logo após o body, que diz respeito a área de renderização do documento. Storage é um componente personalizado em React que será discutido mais adiante, ele serve basicamente para armazenar conteúdos no navegador do usuário. Logo em seguida há o último componente, (ou nó folha), que consiste no children. Neste contexto, children pode ser qualquer coisa a depender da rota de página ao qual se está acessando.

Observe abaixo o código de layout:

```
1 import type { Metadata } from 'next'
2 import { Inter } from 'next/font/google'
3
4 import Storage from '@/components/Storage';
5
6 const inter = Inter({ subsets: ['latin'] })
7
8 export const metadata: Metadata = {
9   title: 'Create Next App',
10  description: 'Generated by create next app',
11 }
12
13 export default function RootLayout({
14   children,
15 }: {
16   children: React.ReactNode
17 }) {
```

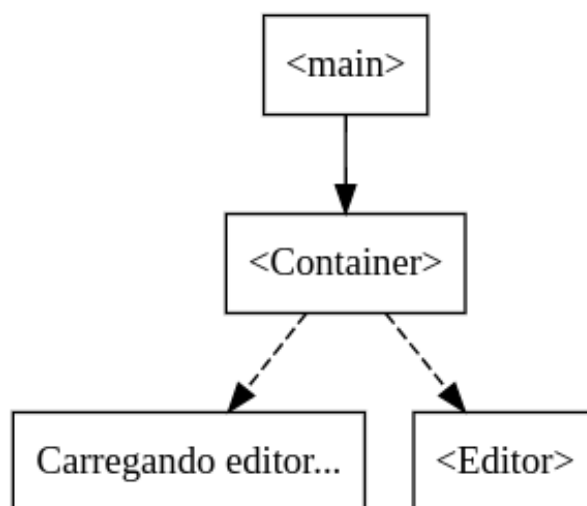


```
18     return (  
19         <html lang="pt-BR">  
20             <body className={inter.className}>  
21                 <Storage>  
22                     {children}  
23                 </Storage>  
24             </body>  
25         </html>  
26     )  
27 }
```

Na linha 13 há a exportação do componente em si. O código TSX é retornado a partir da linha 18. Note que aqui não há o uso da tag head, padrão comum do HTML. Isso se dá pois o gerenciamento das configurações desta tag fica a cargo do NextJs. Observe que da linha 8 à 11 há a exportação de uma constante denominada metadata. Nela há a chave title que o Next utilizará para renderizar o título da página, (que em um html normal seria configurado dentro de head).

3.1.2.2 Page

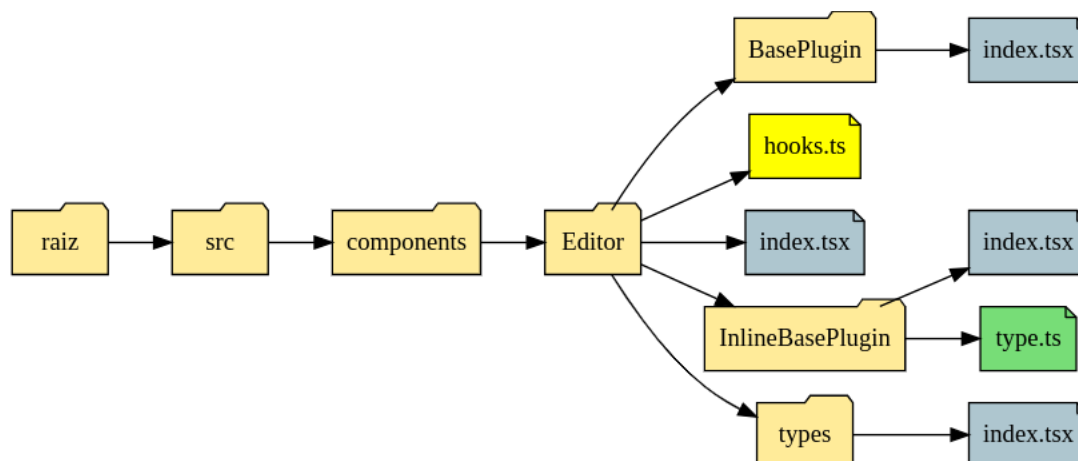
Figura 20 – Sub árvore de renderização da página principal



Fonte: Autoria própria

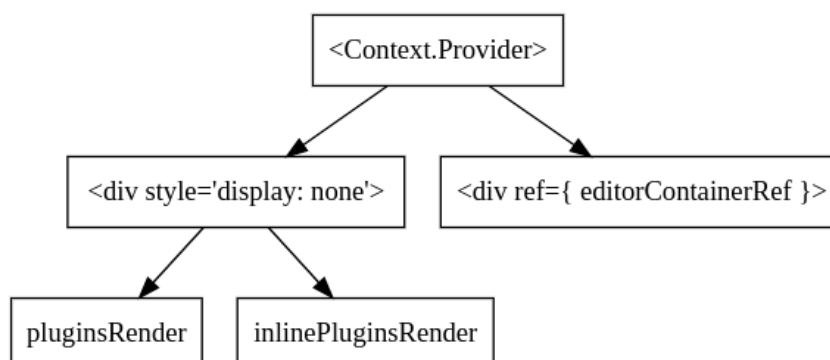
3.2 EditorJS

Figura 21 – Estrutura de pastas do componente Editor



Fonte: Autoria própria

Figura 22 – Sub árvore de renderização do componente Editor



Fonte: Autoria própria

Figura 23 – Componente React - Editor

«React Component» Editor
+editor: Ref +editorContainerRef: Ref +ready: State<boolean> +pluginsList: State<PluginListItemType[]> +inlinePluginsList: State<InlineBasePlugin[]> +pluginsRender: JSX +inlinePluginsRender: JSX
+render() : : JSX +inlineToolRenderHandler(p: RenderCustomEvent) : : void +inlineToolUnmountHandler(p: UnmountCustomEvent) : : void

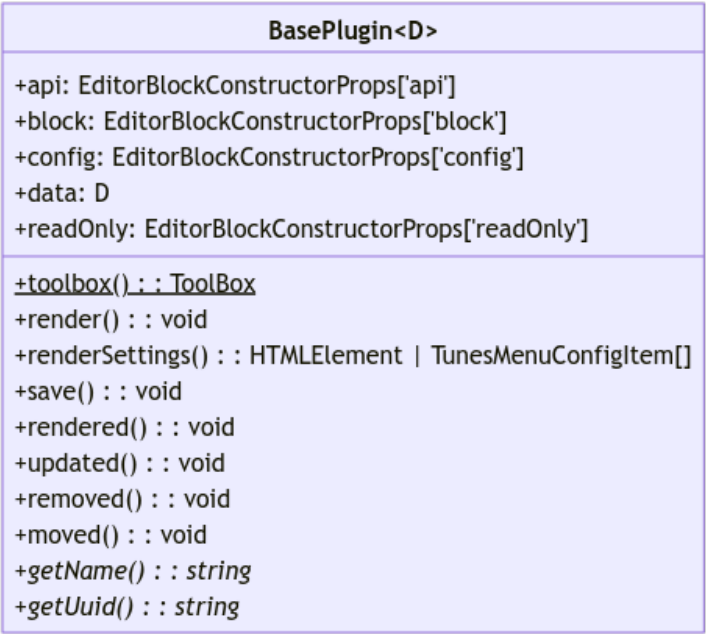
Fonte: Autoria própria

3.2.1 Classe Editor

3.2.2 Provider

3.2.3 BasePlugin

Figura 24 – Classe - BasePlugin



3.2.5.1 Image

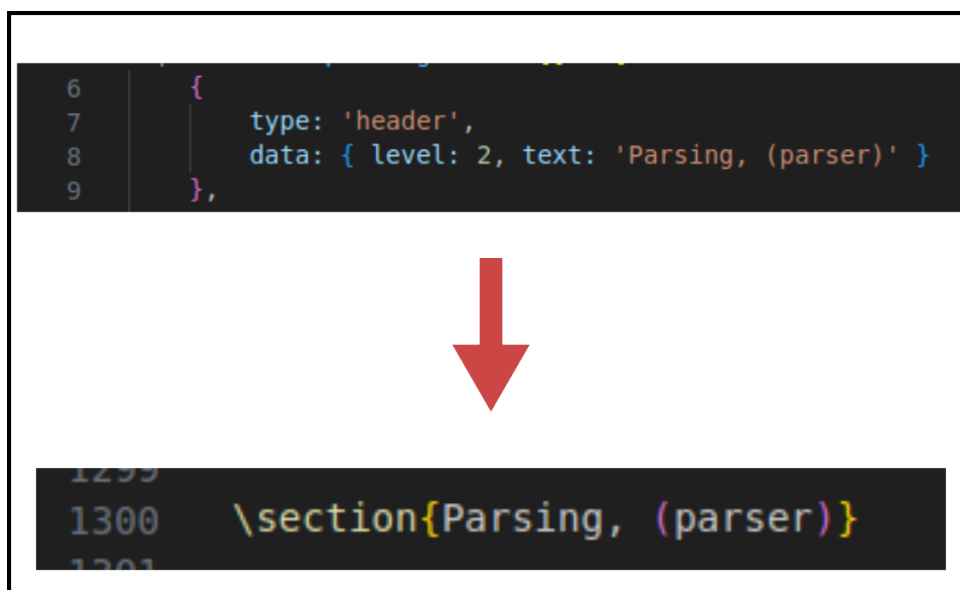
3.2.5.2 Header

3.2.5.3 FootNote

3.3 Parsing, (parser)

O processo de Parsing transformará cada bloco provido pela saída do EditorJs em um trecho de código em LaTeX. Observe na Figura26 e Figura27 os exemplos de parsing aplicados a um objeto de *Header*⁴ e *Paragraph*⁵, respectivamente.

Figura 26 – Parsing de um bloco Header

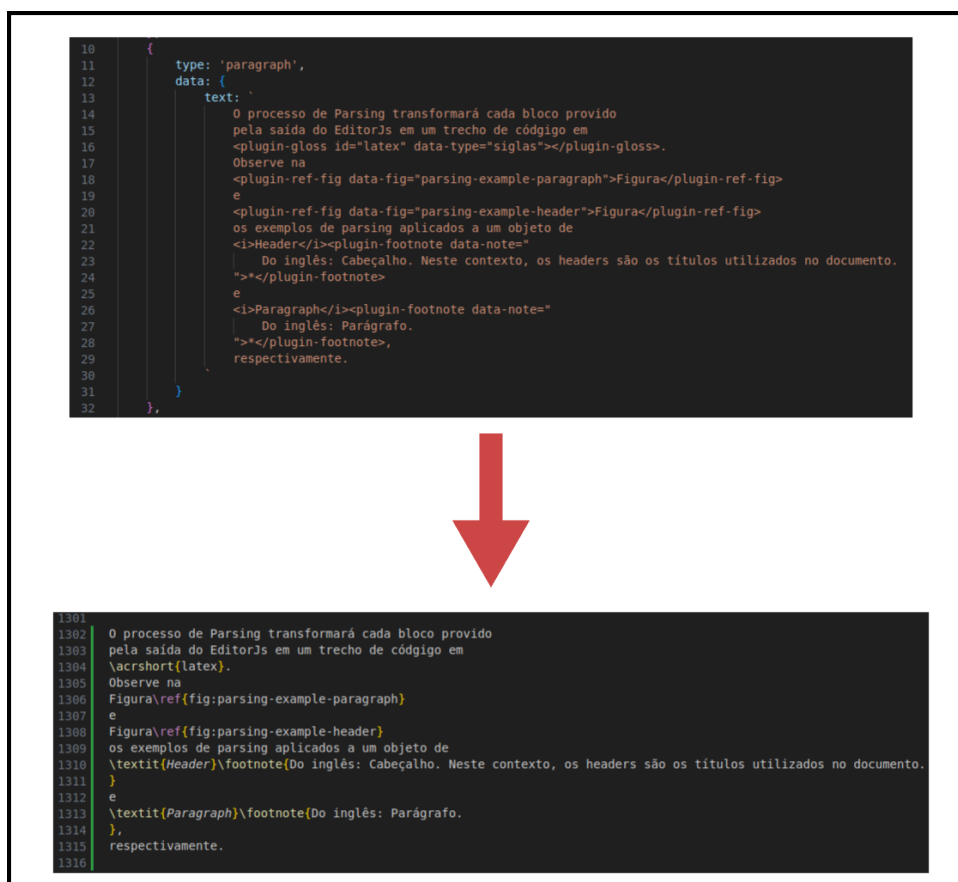


Fonte: Autoria própria

⁴ Do inglês: Cabeçalho. Neste contexto, os headers são os títulos utilizados no documento.

⁵ Do inglês: Parágrafo.

Figura 27 – Parsing de um bloco Paragraph



Fonte: Autoria própria

No exemplo do bloco paragraph, é notável o processo de processamento de HTML acontecendo. Note que o conteúdo textual do bloco não é apenas texto simples. Há quatro tags de marcação personalizadas que se refletem em comandos especiais LaTeX, a saber:

1. plugin-gloss
2. plugin-ref-fig
3. i
4. plugin-footnote

A tabela Tabela6 mostra o mapeamento destas tags, que são plugins *in-line*⁶ do EditorJs. As Tags possuem conteúdo e atributo, cada qual se refletindo a uma particularidade do código LaTeX a depender de sua natureza.

⁶ Do inglês: Em linha.

Tabela 6 – Mapeamento de tags em código LaTeX

Tag	Conteúdo	Atributo	Valor do Atributo	LaTeX
plugin-gloss	Não	id	latex	\acrshort{latex}
plugin-ref-fig	Figura	data-fig	parsing-example-paragraph	\ref{fig:parsing-example-paragraph}
plugin-ref-fig	Figura	data-fig	parsing-example-header	\ref{fig:parsing-example-header}
i	Header	Não	Não	\textit{Header}
plugin-footnote	Não	data-note	<texto da nota>	\footnote{<texto da nota>}
i	Paragraph	Não	Não	\textit{Paragraph}
plugin-footnote	Não	data-note	Do inglês: Parágrafo.	\footnote{Do inglês: Parágrafo.}

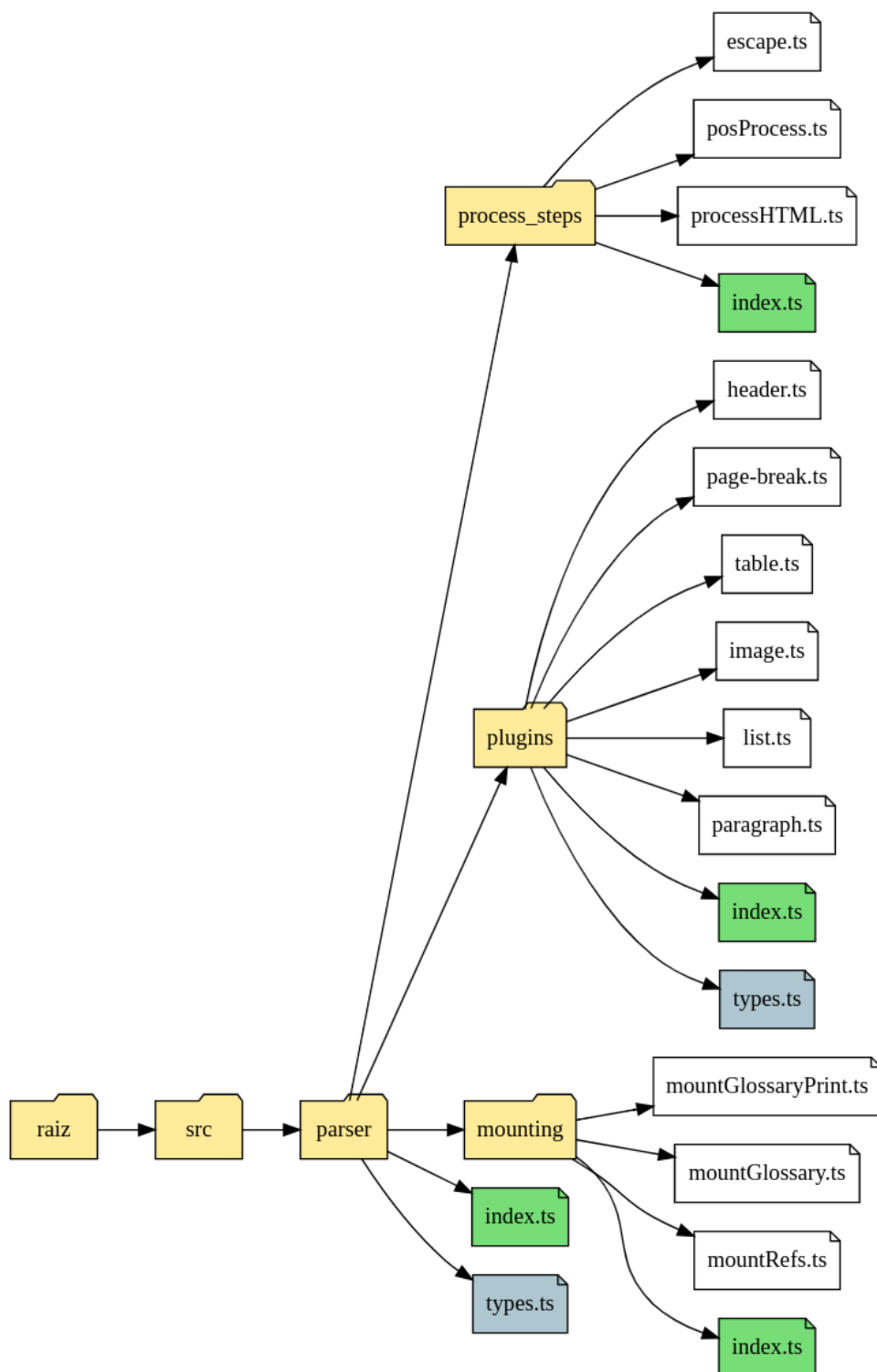
3.3.1 Visão geral

O parser da aplicação não depende de nenhuma biblioteca ou framework de terceiros, (com exceção do processHTML que utiliza a biblioteca cheerio). O processo de parsing é feito apenas utilizando-se de recursos nativos da linguagem TypeScript, (consequentemente JS). A Figura28 mostra a estrutura de pastas do módulo de parse, com três grandes grupos:

- process_steps
- plugins
- mounting

O process_steps diz respeito a funções utilitárias do processamento de texto. É uma das mais importantes partes pois fornece funções que serão utilizadas a todo momento em outras partes do parsing. Provê segurança pois nele reside as funções de escape de caracteres especiais e entre outros. O plugins fornece o processamento respectivo de cada plugin da aplicação, bem como o modo em que cada um deverá ser convertido em código LaTeX. O mounting diz respeito a montagens de partes dinâmicas do documento LaTeX.

Figura 28 – Estrutura de pastas do módulo de parse



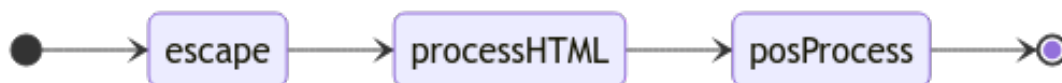
Fonte: Autoria própria

3.3.2 Estapas de processamento

O principal plugin da aplicação a ser processado será o de paragraph. A informação principal deste plugin é tão somente o texto de marcação HTML. Devido a este fato, este plugin será o que mais utilizará as funções providas pelo process_steps.

A Figura29 ilustra um ciclo de processamento de texto de marcação:

Figura 29 – Um ciclo de processamento de texto



Fonte: Autoria própria

3.3.2.1 Escape de caracteres

O escape de processamento é a primeira etapa do processamento de texto. Necessária para evitar o primeiro problema de transformar o texto de marcação em código LaTeX, que é o problema dos caracteres especiais.

O código LaTeX, (como dito na fundamentação teórica), é um código legível tanto para seres humanos quanto por máquinas. Devido a isto, existem alguns caracteres especiais de controle que ajudam a definir como determinado conteúdo aparecerá no documento. Por exemplo: O caractere `\` é um dos mais importantes caracteres do LaTeX, pois ele define uma gama de comandos, como por exemplo o `\chapter{texto}`, que define **texto** como sendo um capítulo no documento. Mas e se o usuário digitar no documento um caractere `\`? Quais problemas isso poderá gerar?

Caso o usuário digite um `\` no documento e este caractere simplesmente seja transcrito no documento LaTeX, gerará uma série de efeitos indesejados, desde comandos inexistentes como: `\esse-comando-não-existe`, como o uso acidental de comandos do próprio LaTeX, como: `\chapter`; `\section` ou `\subsection`. Para evitar isto qualquer ocorrência de `\` no texto será substituída por `\textbackslash`, que é um comando LaTeX responsável por imprimir uma `\` no corpo do texto. A Tabela7 mostra as ocorrências de todos os caracteres que devem ser substituídos por algum comando especial no LaTeX que possua o mesmo significado textual:

Tabela 7 – Mapeamento de escape de caracteres para código LaTeX

Caractere	Substituição
\	\textbackslash
#	\#
\$	\\$
%	\%
^	\textasciicircum
&	\&
_	_
{	\{
}	\}
[{}
]	{}
~	\textasciitilde

Fonte: Adaptado de: (MADEIRA, 2020)

Todas as ocorrências destes caracteres devem ser devidamente substituídas afim de evitar qualquer problema de interpretação do compilador LaTeX.

3.3.2.1.1 Código do escape.ts

Abaixo tem-se a aplicação do processamento de escape de caracteres em typescript. a função `escapeCharacters` recebe uma string na linha 1, e das linhas 3 a 14 faz uma sucessão de novas atribuições desta nova string. As atribuições consistem de uma nova string que, através da função `replace`, substituem as expressões regulares pela nova string, que seguem de acordo com a Tabela7. Ao final, na linha 16, a nova string é retornada.

```

1 export function escapeCharacters(str: string){
2     let newStr = str;
3     newStr = newStr.replace(/\\/gm, '\\textbackslash ');
4     newStr = newStr.replace(/#/gm, '\\#');
5     newStr = newStr.replace(/\$/gm, '\\$');
6     newStr = newStr.replace(/%/gm, '\\%');
7     newStr = newStr.replace(/\~/gm, '\\textasciicircum ');
8     newStr = newStr.replace(/&/gm, '\\&');

```

```

9      newStr = newStr.replace(/_/gm, '\\_');
10     newStr = newStr.replace(/\\{/gm, '\\{');
11     newStr = newStr.replace(/\\}/gm, '\\}');
12     newStr = newStr.replace(/\\[/gm, '\\[');
13     newStr = newStr.replace(/\\]/gm, '\\]');
14     newStr = newStr.replace(/~/gm, '\\textasciitilde ');
15
16     return newStr;
17 }

```

3.3.2.2 Processamento de HTML

O processamento de HTML é a segunda etapa do processo de parsing. É nele que os plugins in-line customizados são transformados em comandos LaTeX. A figura Figura30 ilustra como os plugins de glossário e referência de figuras se comportam no código LaTeX através dos quadrados coloridos:

Figura 30 – Conversão de plugins em código latex

```

text:
0 processamento de
<plugin-gloss id="html"></plugin-gloss>
é a segunda etapa do processo de parsing. É
nele que os plugins in-line customizados são transformados
comandos
<plugin-gloss id="latex"></plugin-gloss>.
A figura
<plugin-ref-fig data-fig="html-processing-example">Figura</plugin-ref-fig>
ilustra como os plugins de glossário e referência de figuras
se comportam no código
<plugin-gloss id="latex"></plugin-gloss>:

1533
1534 0 processamento de
1535 \acrshort{html}
1536 é a segunda etapa do processo de parsing. É
1537 nele que os plugins in-line customizados são transformados
1538 comandos
1539 \acrshort{latex}.
1540 A figura
1541 Figura\ref{fig:html-processing-example}
1542 ilustra como os plugins de glossário e referência de figuras
1543 se comportam no código
1544 \acrshort{latex}:
1545

```

Fonte: Autoria própria

No caso do plugin-gloss, há também um objeto de glossário que contém todas as definições de siglas e abreviaturas. Este objeto é transformado e montado em declarações LaTeX para serem usadas ao longo do documento. Para mais detalhes deste processo, verifique a sessão de montagem

A Tabela8 mapeia como cada plugin se comportará no código LaTeX:

Tabela 8 – Mapeamento de pugins para código latex

Tag plugin	Atributo(s)	LaTeX
plugin-gloss	id	\acrshort{<id>}
plugin-ref	id	\cite{<id>}
plugin-footnote	data-note	\footnote{<data-note>}
plugin-ref-fig	data-fig	Figura\ref{fig:<data-fig>}
plugin-ref-table	data-table	Tabela\ref{fig:<data-table>}
br	Não	\\
strong	Não	\textbf{<conteúdo>}
i	Não	\textit{<conteúdo>}

Observe que no processamento de HTML o conteúdo da maioria das tags é irrelevante, com exceção das tags strong e i. Isso se dá pois na maioria das vezes a informação de interesse é guardada nos atributos, ficando o conteúdo servindo apenas para que se exiba o plugin de forma confortável no texto que o usuário está digitando na UI. As tags strong e i, são respectivamente, o negrito e itálico. São padrões do HTML e aqui seus conteúdos são aproveitados para dentro dos comandos equivalentes em LaTeX.

3.3.2.2.1 Código do processamento HTML

Para processar os plugins está sendo utilizada a biblioteca cheerio. A string a ser analisada, recebida na linha 3, passa por uma checagem de tags presentes. Por exemplo: Observe a linha 5. Sempre que houver qualquer ocorrência de plugin-ref, esta é substituída através do comando replaceWith pelo texto: \cite{\${\$(node).attr('id')}}. Observe que a parte \${\$(node).attr('id')} é o que recupera o atributo "id" da tag.

```

1 import * as cheerio from 'cheerio';
2
3 export function processHTML(text: string): string{
4     const $ = cheerio.load(text);
5     $('plugin-ref').replaceWith((_, node) => {
6         return '\\cite{${$(node).attr('id')}}';
7     }); [...]
```

Por fim, na linha 40 é retornado o resultado do processamento em HTML. Note que ele é recuperado a partir do nó `body` do documento virtual, e então, a função `text()` é chamada para recuperar a representação textual do que foi processado.

```
39 [...]
40     return $('body').text();
41 }
```

O código completo do `processHTML.ts` juntamente com todos os plugins pode ser encontrado no repositório do projeto.

3.3.2.3 Pós processamento

Após feito o processo de escape e processamento de HTML, ainda podem sobrar alguns caracteres remanescentes problemáticos à correta interpretação do código LaTeX. Por exemplo: Os caracteres `<` e `>` são escapados no HTML como as entidades `lt` e `gt`, respectivamente. Eles por sua vez, ao passar pelo escape de caracteres, possuem o caractere `&` escapado como `\&`. Seu código final acaba resultando em `\<` e `\>`, que tem de novamente ser convertido em `<` e `>`. Observe o código abaixo:

```
1 export function posProcess(str: string): string{
2     let new_str = str.replace(/\\</gm, '<');
3     new_str = new_str.replace(/\\>/gm, '>');
4     new_str = new_str.replace(/\\"/gm, '"');
5     new_str = new_str.replace(/"(?!\\w|\\}|\\|\\|\\|~|\\.|,|\\{|\\(|\\[)/gm, '"~');
6     new_str = new_str.replace(/(\\textbackslash)(\\s?\\n|\\s{2,})/gm, '$1~');
7     return new_str;
8 }
```

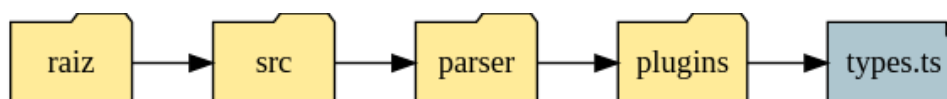
Escrever a conclusão.

3.3.3 Plugins

O parsing dos plugins é algo essencial na montagem do documento ao qual estar-se a escrever. São os plugins que formam cada pequena unidade de bloco que comporá um documento que muitas vezes se tornará gigantesco.

3.3.3.1 Tipagem

Figura 31 – Caminho da pasta types dos plugins no parser



Fonte: Autoria própria

O arquivo de tipagem é onde é concentrada todas as anotações de tipo de cada bloco. O documento todo é um array de objetos, no qual cada um desses objetos corresponde a um bloco do documento. É através de cada tipo de bloco que pode-se por exemplo, discriminar qual tipo de bloco está sendo tratado no momento do processamento. Observe por exemplo a anotação de tipo do ParagraphBlock:

```
37 [...]
38 export interface ParagraphBlock {
39     type: 'paragraph'
40     id: string
41     data: {
42         text: string
43     }
44 }
45 [...]
```

Todo bloco é composto basicamente por estas três propriedades: type; id e data. Sendo que a propriedade data variará sua forma a depender do bloco que está sendo definido. Este código nunca é executado ou incluído na versão de execução final em JS, (por se tratar de uma anotação de tipo). Mas é de extrema utilidade quando se trata de guiar o desenvolvedor na hora de instanciar um objeto do tipo que é definido pelo interface⁷ ParagraphBlock.

Observe como o interface do HeaderComponent se diferencia do HeaderComponent em sua anotação de tipo:

```
5 [...]
6 export interface HeaderComponent {
7     type: 'header'
8     id: string
9     data: {
```

⁷ Neste contexto, está-se utilizando o termo interface em seu original inglês, conforme a documentação do TypeScript. Devido a este fato, interface (não confundir com interface de usuário) será tratado com pronome masculino.

```

10         level: 1 | 2 | 3 | 4 | 5
11         text: string
12     }
13 }
14 [...]

```

No HeaderBlock, a propriedade data é composta por outras duas propriedades: text e level. Ao refletir sobre isso, percebe-se que essas duas propriedades são essenciais para definir o que é um título em um documento. O level (nível) indica o nível hierárquico do título, enquanto o text (texto) representa o conteúdo exibido como título. Perceba que há uma limitação na definição do nível do título, com valores variando apenas de 1 a 5. Cada um desses valores corresponde respectivamente às sessões: primária, secundária, terciária, quaternária e quinária.

Observe na Tabela9 a relação do level do título com o código LaTeX correspondente:

Tabela 9 – Relação do level do título com o código latex

Level	Código LaTeX
1	<code>\chapter{<conteúdo>}</code>
2	<code>\section{<conteúdo>}</code>
3	<code>\subsection{<conteúdo>}</code>
4	<code>\subsubsection{<conteúdo>}</code>
5	<code>\subsubsubsection{<conteúdo>}</code>

3.3.3.2 Paragraph, (parágrafo)

O bloco de parágrafo parece de longe um dos blocos mais simples de todos. Ele simplesmente retorna o texto do bloco recebido, fazendo-o passar pelas etapas de processamento conforme já descrito na Figura29. Observe na linha 7 como o parâmetro block é anotado com o tipo ParagraphBlock. É neste caso que entra a utilidade do TypeScript. Pois além de prevenir erros não deixando que se passe um bloco incorreto para ser processado por essa função, tem-se a condição de saber exatamente o que tem dentro da propriedade data na linha 11. Neste caso, apenas passa-se o texto para o processamento.

```

1 import { escapeCharacters } from '@/parser/process_steps/escape';
2 import { processHTML } from '@/parser/process_steps/processHTML';
3 import { ParagraphBlock } from '@/parser/types';
4 import { posProcess } from '@/parser/process_steps/posProcess';

```

```

5
6
7 export function getParagraph(block: ParagraphBlock){
8     return posProcess(
9         processHTML(
10             escapeCharacters(
11                 block.data.text
12             )
13         )
14     );
15 }

```

3.3.3.3 Header, (cabecçalhos)

O parse do bloco de Header já distoa-se um pouco do bloco de paragraph. Aqui é notável a aplicação do exposto na Tabela 9 onde o cada level se traduz em um comando diferente no latex. Através da clausula switch, o level é testado afim de retornar as particularidades de cada tipo de título. Observe que o level 4, implementado na linha 13, possui um comando a mais de `\underline` para que o título apareça como sublinhado. Este comando a mais foi implementado para atender as particularidades exidas pelo documento da PUC-GO em: (PUC-GO, 2022).

```

1 import { escapeCharacters } from '@/parser/process_steps/escape';
2 import { HeaderBlock } from '@/parser/types';
3
4 export function getHeader(block: HeaderBlock){
5     switch(block.data.level){
6         case 1:
7             return '\\chapter{${escapeCharacters(block.data.text)}}';
8         case 2:
9             return '\\section{${escapeCharacters(block.data.text)}}';
10        case 3:
11            return '\\subsection{${escapeCharacters(block.data.text)}}';
12        case 4:
13            return '\\subsubsection{\\underline{${
14                escapeCharacters(block.data.text)
15            }}}';
16        case 5:
17            return '\\subsubsubsection{${
18                escapeCharacters(block.data.text)

```

```

19         }}‘
20     }
21 }

```

3.3.3.4 Image, (imagens)

O parse do plugin de imagem é um dos mais complexos em termos de comandos LaTeX. Ele retorna toda uma estrutura de modo que o compilador de documentos LaTeX possa renderizar a imagem corretamente e seja capaz de referenciá-la ao longo do texto.

Nas linhas 8 a 12 tem-se a desconstrução das propriedades da imagem que estão presentes na propriedade data do bloco. Estas propriedades são: uuid⁸; title; width; description e fileType.

```

1 import { escapeCharacters } from '@parser/process_steps/escape';
2 import { posProcess } from '@parser/process_steps/posProcess';
3 import { processHTML } from '@parser/process_steps/processHTML';
4 import { ImageBlock } from '@parser/types';
5
6 export function getImage(block: ImageBlock){
7     const {
8         uuid,
9         title,
10        width,
11        description,
12        fileType,
13    } = block.data;
14    [...]

```

Na linha 20 tem-se a inclusão do título da imagem. Note que o título somente precisa passar pelo processo de escape de caracteres, uma vez que no mesmo não haverá plugins no corpo do texto. Nas linhas 25 à 33 há a inclusão da descrição da imagem. Note que diferente do título, a descrição passa por todo o ciclo de processamento, pois na descrição o usuário pode incluir referências e outras coisas que resultam na presença de tags de plugin no corpo do texto.

As linhas 21 a 23 tratam do tamanho que a imagem vai ocupar na tela, bem como a referência ao arquivo de imagem que será renderizado. As propriedades width, uuid, e fileType são utilizadas neste processo. Cada imagem é guardada na pasta

⁸ UUID, *Universally Unique Identifier* (Identificador Universalmente Único).

interna images com o nome do uuid que o software define, bem como sua extensão de arquivo.

A linha 24 utiliza o uuid para definir a label da imagem. Desta forma, a imagem a que se trata o bloco torna-se referenciável no texto do documento.

```
13 [...]
14
15
16 // H Prevents figure to be placed incorrectly
17 return '
18     \begin{figure}[H]
19         \centering
20         \caption{{\escapeCharacters(title)}}
21         \includegraphics[width=${
22             width.toFixed(1)
23         }\\textwidth]{./images/${uuid}.${fileType}}
24         \label{fig:${uuid}} \\\
25         \textnormal{\fontsize{10pt}{12pt}}{
26             posProcess(
27                 processHTML(
28                     escapeCharacters(
29                         description
30                     )
31                 )
32             )
33         }}
34     \end{figure}
35     '.trim().replace(/^\s{8}/gm, '');
36 }
```

3.3.3.5 List, (listas)

Existem dois tipos de lista, as enumeradas e as não enumeradas, respectivamente `enumerate` e `itemize` no código LaTeX. A linha 9 utiliza a propriedade `type` para definir qual tipo de lista será utilizada. Por fim, a lista é composta de uma iteração sobre as strings do array presente na propriedade `list`. Este array, após ajustado com a função `map`, é transformado em uma string unindo seus itens por uma quebra de linha com o auxílio da função `join` na linha 11. Observe também o processamento textual presente na linha 11. Isso se dá pela liberdade do usuário em utilizar plugins nos campos de lista.

```

1 import { escapeCharacters } from "@parser/process_steps/escape";
2 import { processHTML } from "@parser/process_steps/processHTML";
3 import { ListBlock } from "@parser/types";
4 import { posProcess } from "@parser/process_steps";
5
6 export function getList(list: ListBlock){
7     const { data: { type, list: __list } } = list;
8     return '
9     \r\\begin{${ type === 'bullet' ? 'itemize' : 'enumerate' }}
10         ${__list.map(el => '\r\t\\item ${
11             posProcess(processHTML(escapeCharacters(el))) }').join('')}
12     }
13     \r\\end{${ type === 'bullet' ? 'itemize' : 'enumerate' }}
14     '.trim()
15 }

```

3.3.3.6 Page Break, (quebra de página)

O bloco de PageBreak é o bloco de parser mais simples de todos. Embora receba como parâmetro um bloco do tipo PageBreakBlock, nem uso do mesmo é feito. O PageBreak apenas retorna um comando em LaTeX que limpa a página e faz com que o conteúdo após o mesmo seja escrito na próxima página.

```

1 import { PageBreakBlock } from '@parser/types';
2
3 export function pageBreak(block: PageBreakBlock){
4     return '\\clearpage';
5 }

```

3.3.3.7 Table, (tabelas)

Explicar o complexo parser do plugin de tabela.

```

1 import { escapeCharacters } from '@parser/process_steps/escape';
2 import { processHTML } from '@parser/process_steps/processHTML';
3 import { posProcess } from '@parser/process_steps/posProcess';
4 import { TableBlock } from '@parser/types';
5
6
7 export function getTable(block: TableBlock){

```

```

8      const {
9          id,
10         title,
11         description,
12         header,
13         items,
14         width,
15         column_sizes,
16     } = block.data;
17
18     const MAX_WIDTH = 16*width;
19
20     return '
21         \\begin{table}[H]
22             \\centering
23             \\caption{{escapeCharacters(title)}}
24             \\label{tbl:${id}}
25             \\renewcommand{\\arraystretch}{1.5}
26             \\begin{tabular}{{${
27                 column_sizes.map(col => {
28                     return 'p{{(MAX_WIDTH*col).toFixed(4)}}cm}'
29                 }).join(' ')}
30             }}
31             \\hline
32             ${
33                 header.map(
34                     h => '\\textbf{{${
35                         posProcess(
36                             processHTML(
37                                 escapeCharacters(h)
38                             )
39                         )
40                     }}'
41                 ).join(' & ')
42             } \\\\
43             \\hline
44             ${
45                 items.map(row => {
46                     return row.map(cell => {

```

```
47         return posProcess(  
48             processHTML(  
49                 escapeCharacters(cell)  
50             )  
51         )  
52     }).join(' & ')  
53 }).join(' \\\n\t\t')  
54 } \\\n  
55 \\\hline  
56 ${  
57     description &&  
58     '\\\\multicolumn${  
59         header.length  
60     }}>{c}{\\fontsize{10pt}{12pt}}${  
61         posProcess(  
62             processHTML(  
63                 escapeCharacters(description)  
64             )  
65         )  
66     }},'  
67 }  
68 \\\end{tabular}  
69 \\\end{table}  
70 '.trim().replace(/~\s{8}/gm, ' ');  
71 }
```

3.3.4 Montagem

3.3.4.1 Glossário

3.3.4.2 Referências

REFERÊNCIAS

ABNT. *Quem somos? ABNT - Associação Brasileira de Normas Técnicas*. 2021. [Online; acessado em 15-Maio-2023]. Disponível em: <<http://www.abnt.org.br/institucional/sobre>>.

AMAZON; AWS. *O que é uma framework em programação e engenharia?* [Online; acessado em 25-Maio-2024]. Disponível em: <<https://aws.amazon.com/pt/what-is/framework/>>.

BRASIL. Lei nº 12.965, de 23 de abril de 2014. dispõe sobre os direitos civis na internet. *Diário Oficial da União*, 2014.

CASTRO, C. M. *Como redigir e apresentar um trabalho científico*. [S.l.]: São Paulo: Pearson Prentice Hall, 2011.

EDITORJS. Base concepts. [Online; acessado em 30-Maio-2024]. Disponível em: <<https://editorjs.io/base-concepts/>>.

MACHADO, E. Hardware e software. Enciclopédia Significados. Disponível em: <<https://www.significados.com.br/hardware-e-software/s>>.

MADEIRA, D. *Tutorial básico de LATEX*. [S.l.: s.n.], 2020.

MAGALHÃES, C. Regex: o guia essencial das expressões regulares. *Blog DP6*, Medium, 2022. [Online; acessado em 01-Junho-2024]. Disponível em: <<https://blog.dp6.com.br/regex-o-guia-essencial-das-express%C3%B5es-regulares-2fc1df38a481>>.

MDN. Css. Mozilla Developer Network, 2023. [Online; acessado em 16-Maio-2023]. Disponível em: <<https://developer.mozilla.org/pt-BR/docs/Web/CSS>>.

MDN. Design responsivo. Mozilla Developer Network, 2023. [Online; acessado em 28-Maio-2024]. Disponível em: <https://developer.mozilla.org/pt-BR/docs/Learn/CSS/CSS_layout/Responsive_Design>.

MDN. Javascript. Mozilla Developer Network, 2023. [Online; acessado em 16-Maio-2023]. Disponível em: <<https://developer.mozilla.org/pt-BR/docs/Web/JavaScript>>.

MDN. Json. Mozilla Developer Network, 2024. [Online; acessado em 30-Maio-2024]. Disponível em: <https://developer.mozilla.org/pt-BR/docs/Web/JavaScript/Reference/Global_Objects/JSON>.

MDN. Regular expressions. *Web Docs*, Mozilla Developer Network, 2024. [Online; acessado em 01-Junho-2024]. Disponível em: <https://developer.mozilla.org/en-US/docs/Web/JavaScript/Guide/Regular_expressions>.

MEDEIROS, J. B. *Redação científica: a prática de fichamentos, resumos, resenhas*. 11st. ed. [S.l.]: São Paulo: Atlas, 2012.

- MOLGADO, V. *A Evolução da Web: linha do tempo interativa da História da internet*. 2016. [Online: acessado em 15-Maio-2023]. Disponível em: <<https://labvis.eba.ufrj.br/a-evolucao-da-web-linha-do-tempo-interativa-da-historia-da-internet/>>.
- MORAIS, P. A história do react! Medium, 2021. [Online: acessado em 29-Maio-2024]. Disponível em: <<https://medium.com/@ppternunes/a-hist%C3%B3ria-do-react-ba346c416fe1>>.
- NEXTJS. Introduction. 2024. [Online: acessado em 29-Maio-2024]. Disponível em: <<https://nextjs.org/docs>>.
- PIMENTEL, E. O que é dom? Alura, 2022. [Online: acessado em 01-Junho-2024]. Disponível em: <<https://www.alura.com.br/artigos/o-que-e-o-dom>>.
- PUC-GO. Coordenação de tcc. manual para elaboração de trabalho de conclusão de curso. Pontifícia Universidade Católica de Goiás, 2022.
- REACTJS. Writing markup with jsx. 2024. [Online: acessado em 01-Junho-2024]. Disponível em: <<https://pt-br.react.dev/learn/writing-markup-with-jsx>>.
- REDDY, M. *API Design for C++*. [S.l.]: Elsevier Science, 2011. ISSN 9780123850041.
- SANTOS, I. R. As dificuldades na construção do trabalho de conclusão de curso: Percepção de estudantes egressos do curso de ciências contábeis. 2020. 69. trabalho de conclusão de curso (graduação) - ciências contábeis. Universidade Federal da Paraíba, João Pessoa, 2020.
- SEVERINO, A. J. *Metodologia do trabalho científico*. 24st. ed. [S.l.]: Cortez, 2017.
- SILVA, M. O. da; VITORIA, M. I. C. A experiência de escrita no trabalho de conclusão de curso: Percepções de alunos de um curso superior de tecnologia em gestão em recursos humanos. CAMINE: Caminhos da Educação, Franca, 2014. ISSN 2175-4217.
- TOTVS. *O que é back-end e qual seu papel na programação?* 2020. [Online: acessado em 01-Junho-2024]. Disponível em: <<https://www.totvs.com/blog/developers/back-end/>>.
- TOTVS. *Front end: O que é, como funciona e qual a importância*. 2021. [Online: acessado em 25-Maio-2024]. Disponível em: <<https://www.totvs.com/blog/developers/front-end/>>.
- TYPESCRIPT. *TypeScript*. [Online: acessado em 29-Maio-2024]. Disponível em: <<https://www.typescriptlang.org/>>.
- VITORIANO, D. *O que é a Web Moderna*. 2019. Medium. [Online: acessado em 15-Maio-2023]. Disponível em: <<https://blog.danvitoriano.com.br/o-que-%C3%A9-a-web-moderna-b01e4df9a565>>.
- W3C. Html. World Wide Web Consortium, 2023. [Online: acessado em 16-Maio-2023]. Disponível em: <<https://html.spec.whatwg.org/multipage>>.
- W3C. Typescript introduction. W3SCHOOLS, 2024. [Online: acessado em 29-Maio-2024]. Disponível em: <https://www.w3schools.com/typescript/typescript_intro.php>.