

Universidade do Minho Departamento de Informática



Processamento de Linguagens ${\it Grupo~24}$

22 de maio, 2022



Beatriz Rodrigues (a93230)



Francisco Neves (a93202)



Guilherme Fernandes (a93216)

Índice

1	Introdução					
2	Analisador Léxico Gramática					
3						
4	$Parser \ { m e} \ AST$					
5	Dicionários	7				
	5.1 JSON $parser$	7				
	5.1.1 Analisador Léxico	7				
	5.1.2 Gramática	7				
	5.1.3 <i>Parser</i>	8				
6	Pipes					
7	Templates Implementados	10				
	7.1 HTML	10				
	7.2 Markdown	11				
	7.3 Asciidoc	12				
	7.4 Plain Text	13				
	7.5 Bash Tree	14				
	7.6 Manual de Receitas	15				
	7.7 Pokédex	16				
	7.8 Carta de Motivação	18				
8	Extras Estéticos	20				
	8.1 Tratamento de Erros	20				
	8.2 Syntax Highlighting	21				
9	Conclusões e Trabalho Futuro	22				
10	0 Bibliografia					

1. Introdução

koal@ foi desenvolvido no contexto da unidade curricular de Processamento de Linguagens. Retirando inspiração da linguagem de *templates* Pandoc, foi desenvolvida uma ferramenta que permite, de forma semelhante, recorrer a um dicionário (*JSON* ou *YAML*) e a um *template* seguindo a *syntax* definida para dar origem a um resultado final no formato correspondente ao *template* utilizado.

No entanto, ao longo do processo, foi decidido que seria interessante adicionar novas funcionalidades vistas como potencialmente úteis, a descrever em mais detalhe no presente relatório.

Para além disso, como possível valorização, foi implementado ainda um *parser* de ficheiros *JSON*, a utilizar em conjunto com a linguagem desenvolvida.

2. Analisador Léxico

Em primeiro lugar, foram definidos os tokens desta linguagem:

- Texto → Todos os símbolos contidos entre aspas
- $VAR \rightarrow S$ ímbolos iniciados por , que representam possíveis entradas do dicionário, ou seja, as variáveis da template;
- TMPVAR → Símbolos iniciados por #, que podem ser definidas no template, representando assim variáveis temporárias;
- NEWLINE → Símbolo representativo do newline que permite efetuar o controlo de linhas;
- ALIASNAME → Representam funções o nome de funções que podem ser posteriormente utilizadas;
- IF \rightarrow Palavra reservada (if) que indica uma condição do tipo if;
- ELIF → Palavra reservada (elif) que indica uma condição do tipo elif;
- ELSE → Palavra reservada (else) que indica uma condição do tipo else;
- $FOR \rightarrow Palavra reservada (for) que indica um ciclo for;$
- ALIAS → Palavra (alias) que indica a definição de um alias;
- INCLUDE → Palavra reservada (include) que indica a inclusão dos alias de um outro ficheiro indicado em seguida;
- PIPE \rightarrow Caractere '/' que indica que se deverá seguir um *pipe* do programa.

Foram ainda definidos os seguintes caracteres como *literals* da linguagem em questão: '{', '}', ':', '(', ')' e ','.

Além disto, foi ainda definida a regra t_ignore para ignorar caracteres não pretendidos e a regra t_ignore_comment de forma a ignorar os comentários (linhas iniciadas por '//').

Por fim, de forma a efetuar um correto controlo de *newlines* foi criada uma *flag* denominada *ignore_newline* que, tal como o nome indica, tem como propósito indicar em que circunstâncias é pretendido que se ignore um *newline* e vice-versa.

3. Gramática

Tendo isto, passou-se então para a definição da gramática necessária para a linguagem de *templates* pretendida.

Assim, foi decidido que a linguagem (lang) seria definida por statements, sendo que statements seria uma lista de statement com caso de paragem em vazio. Por sua vez, um statement poderia ser as diversas possibilidades oferecidas pela linguagem, desde expressões condicionais, até ciclos for, variáveis, aliases ou texto corrido.

Foi então definida a seguinte gramática:

```
Rule 0
           S' -> lang
Rule 1
           lang -> statements
Rule 2
           statements -> statements statement
Rule 3
           statements -> <empty>
Rule 4
           statement -> variable
Rule 5
           statement -> NEWLINE
Rule 6
           statement -> if elifs else
Rule 7
           if -> IF variable block
Rule 8
           elifs -> elifs elif
Rule 9
           elif -> ELIF variable block
Rule 10
           elifs -> <empty>
Rule 11
           else -> ELSE block
Rule 12
           else -> <empty>
           statement -> FOR TMPVAR : variable block
Rule 13
Rule 14
           statement -> ALIAS ALIASNAME tmpvars block
Rule 15
           tmpvars -> tmpvars TMPVAR
Rule 16
           tmpvars -> <empty>
Rule 17
           statement -> ALIASNAME ( args )
Rule 18
           args -> variables
Rule 19
           args -> <empty>
Rule 20
           variables -> variables , variable
Rule 21
           variables -> variable
Rule 22
           statement -> INCLUDE TEXT NEWLINE
Rule 23
           block -> { statements }
Rule 24
           variable -> VAR
Rule 25
           variable -> TMPVAR
Rule 26
           variable -> TEXT
Rule 27
           variable -> variable FIELD
Rule 28
           variable -> variable PIPE
```

4. Parser e AST

Tendo em conta a biblioteca utilizada para efetuar o parsing da linguagem (módulo yacc do ply), foi observado que se estava perante uma estratégia de parsing bottomup, pelo que, foi prioritária a utilização de recursividade à esquerda.

Além disto, reparou-se ainda que, devido a esta estratégia de *parsing*, por diversas ocasiões como em avaliações de expressões condicionais ou de ciclos era, primeiramente efetuada a verificação da expressão no corpo da funcionalidade, sendo que, apenas após isso se validava a condição da expressão, algo que, não era de todo o pretendido.

Assim, foi decidido que seria benéfica a criação de uma AST (árvore de sintaxe abstrata) que permite inverter o processo de avaliação acima referido e confere uma maior abstração à avaliação da linguagem desenvolvida. Para isto, foi criado o módulo $koala_ast$ que possui diversas classes abstratas que são utilizadas nas classes representativas dos diversos tipos da linguagem criada, quando adequadas.

5. Dicionários

Como referido anteriormente, a linguagem desenvolvida é capaz de interpretar dicionários do tipo JSON ou YAML. Sendo que, de forma a valorizar o trabalho efetuado, para os dicionários do tipo JSON foi também criada uma ferramenta de parsing que permite a sua transformação em dicionários de Python. Por outro lado, para a utilização de dicionários do tipo YAML foi utilizada a biblioteca yaml recorrendo ao seu método load, utilizando o Loader yaml. FullLoader.

5.1 JSON parser

De forma a efetuar o parsing de dicionários JSON para dicionários de Python, foi criado um parser para este tipo de ficheiros.

5.1.1 Analisador Léxico

Inicialmente, foram decididos os *tokens* que o *parser* deveria ser capaz de reconhecer, foram eles:

- STRING → Texto colocado entre aspas;
- TRUE \rightarrow Palavra reservada true;
- **FALSE** \rightarrow Palavra reservada *false*;
- NULL \rightarrow Palavra reservada *null*;
- **NUMBER** Número real que poderá encontrar-se em notação científica ou não.

Além destes, foram ainda utilizados os seguintes literals: '{', '}', ',', ':', '[' e ']'.

5.1.2 Gramática

Tendo isto, foi então criada a seguinte gramática:

```
S' -> json
Rule 0
Rule 1
           json -> element
Rule 2
           value -> object
Rule 3
           value -> array
Rule 4
           value -> STRING
Rule 5
           value -> NUMBER
           value -> TRUE
Rule 6
Rule 7
           value -> FALSE
```

```
Rule 8
           value -> NULL
Rule 9
           object -> { }
Rule 10
           object -> { members }
           members -> member
Rule 11
           members -> members , member
Rule 12
Rule 13
           member -> STRING : element
Rule 14
           array -> <empty>
Rule 15
           array -> [ elements ]
Rule 16
           elements -> element
Rule 17
           elements -> elements , element
Rule 18
           element -> value
```

5.1.3 Parser

Tendo em conta o objetivo deste trabalho prático, o parser foi definido tendo em mente apenas a definição de um dicionário em Python passível de utilização pelo koal@. Assim, recorrendo ao módulo yacc do ply e dando prioridade à recursividade à esquerda na definição da gramática, foi definido o parser para os ficheiros pretendidos.

6. Pipes

A implementação de *pipes* surge de forma a permitir que o utilizador possa efetuar diferentes operações consoante a posição relativa de um elemento numa dada lista ou, até mesmo, converter *Strings* para maiúsculas ou minúsculas e inverter a ordem de uma lista ou de uma *String*. Com a sua utilização, permitimos que, por exemplo, todos os elementos da lista sejam terminados com o caractere ',' exceto o último, conforme uma típica enumeração.

Assim, com isto em mente, foram implementados os seguintes pipes:

- First: Utiliza-se com /first e permite obter o primeiro elemento da lista;
- Last: Utiliza-se com /last e permite obter o último elemento da lista;
- **Head**: Utiliza-se com /head e permite obter todos os elementos de uma lista exceto o último;
- *Tail*: Utiliza-se com /tail e permite obter todos os elemento de uma lista exceto o primeiro;
- *Upper*: Utiliza-se com /upper e permite colocar uma *String* em maiúsculas;
- Lower: Utiliza-se com /lower e permite colocar uma String em minúsculas;
- **Reverse**: Utiliza-se com /reverse e permite inverter a ordem de uma *String* ou de uma lista.

7. Templates Implementados

De forma a serem efetuados alguns testes na linguagem criada, foram desenvolvidos diversos *templates* de linguagens *markup* conhecidas ou outras utilidades consideradas interessantes para a linguagem em questão. Entre eles, podem-se considerar os seguintes:

7.1 HTML

Figura 7.1: Excerto do template de HTML

7.2 Markdown

Figura 7.2: Alias definidos para o template de Markdown

```
examples > markdown > \bar{\textbf{b}} markdown/mdhelp.koa"

include "examples/markdown/mdhelp.koa"

if \textit{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleblock}}{\textit{atitleb
```

Figura 7.3: Template de Markdown

7.3 Asciidoc

```
examples) asclidoc \( \text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\t
```

Figura 7.4: Template de Asciidoc

7.4 Plain Text

```
if @titleblock{
    @titleblock
}
for #item : @headerincludes{
    #item
}
for #item : @includebefore{
    #item
}
if @tableofcontents{
    @tableofcontents
}
@body
for #item : @includeafter{
    #item
}
```

Figura 7.5: Template de Plain text

7.5 Bash Tree

```
alias cd #dir {
    "cd " #dir
}
alias mkdir #dir {
    "mkdir " #dir
}
alias cp #source #dest {
    "cp " #source " " #dest
}
alias touch #file {
    "touch " #file
}
```

Figura 7.6: Comandos bash

Figura 7.7: Template de Bash Tree

7.6 Manual de Receitas

Figura 7.8: Template do manual de receitas

```
recipes:

- name: Bacalhau com Natas
- inage: https://static.clubedaanamariabraga.com.br/wp-content/uploads/2019/11/bacalhau-com-natas-1024×576.jpg
site: https://snamariabraga.globo.com/receita/bacalhau-com-natas/
difficulty: fācil
veggie: Não
doses: 4
time: lh00min
- name: Arroz de Pato
inage: https://wm.tudoreceitas.com/receita-de-arroz-de-pato-a-portuguesa-219_600.jpg
site: https://wm.tudoreceitas.com/receita-de-arroz-de-pato-a-portuguesa-219.html
veggie: Não
doses: 6
time: lh30min
```

Figura 7.9: Fragmento do dicionário exemplificativo

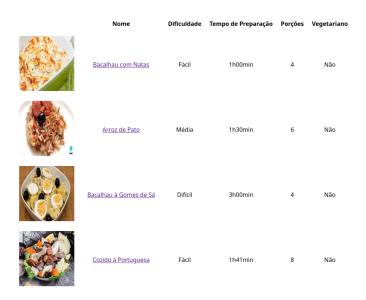


Figura 7.10: Fragmento do resultado do manual de receitas

7.7 Pokédex

```
include "examples/std.koa"
"<!DOCTYPE html>

    <thead>
             Number
             Name
             Type
         </thead>"
    "
for #poke : @pokemon {
" "
            <img src="#poke.img">"
"#poke.num""
"#poke.name""
"join(#poke.type, "<br>""
"
    if #poke.next_evolution {
    for #evo : #poke.next_evolution/head {
          "#evo.name
                  <br>"
                  "#poke.next_evolution/last.name
             "
      "
"<∕table>"
```

Figura 7.11: Template da Pokédex

```
examples > [6] std.koa | alias join #list #delim { for #elem : #list/head {#elem #delim} #list/last }
```

Figura 7.12: Ficheiro stdalias

```
"pokemon": [{
    "id": 1,
    "num": "001",
    "name": "Bulbasaur",
    "img": "http://www.serebii.net/pokemongo/pokemon/001.png",
    "type": [
        "Grass",
        "Poison"
],
    "height": "0.71 m",
    "weight": "6.9 kg",
    "candy': "Bulbasaur Candy",
    "candy-count": 25,
    "egg": "2 km",
    "spawn_chance": 0.69,
    "avg_spawns": 69,
    "spawn_time": "20:00",
    "multipliers": [1.58],
    "weaknesses": [
        "Fire",
        "Ice",
        "Ire",
        "Flying",
        "Psychic"
],
    "next_evolution": [{
        "num": "002",
        "name": "Ivysaur"
}, {
        "num": "003",
        "name": "Venusaur"
}]
```

Figura 7.13: Fragmento do dicionário da Pokédex

	Number	Name	Туре	Evolutions
300	001	Bulbasaur	Grass Poison	Ivysaur Venusaur
	002	Ivysaur	Grass Poison	Venusaur
	003	Venusaur	Grass Poison	
*	004	Charmander	Fire	Charmeleon Charizard
8	005	Charmeleon	Fire	Charizard
X	006	Charizard	Fire Flying	

Figura 7.14: Fragmento do resultado da *Pokédex*

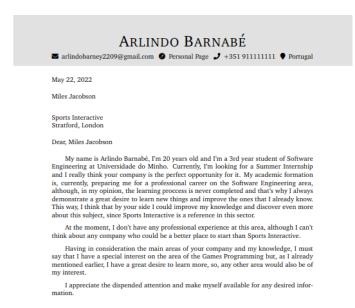
7.8 Carta de Motivação

```
\\newcommand{\\myname}{"@name"}
\\newcommand{\\myname!{"@name"}
\\newcommand{\\myname!!f" @ittle {@title} else {"Applicant"}"]
\\newcommand{\\mynage}{"if @page {@page} elif @git {@git} else { @linkedin }"}
\\newcommand{\\mynpage}{"if @page {@page} elif @git {@git} else { @linkedin }"}
\\newcommand{\\mynpage}{"if @page {@page} elif @git {@git} else { @linkedin }"}
\\newcommand{\\mynpage}{"alocation"}"
if @recipient{
        "\newcommand{\\recipient}{"@recipient"}"
}
\\newcommand{\\coser}{"if @greetingf@greeting", "} else {"Greetings, "}"}
\\newcommand{\\coser}{"if @closerf@closer", "} else{"Best wishes and regards, "}"}
\\newcommand{\\company}{"@company"}"
if @company_city {
        "\newcommand{\\street}{"@company_street"}"
}
if @company_city {
        "\newcommand{\\street}{"@company_state"}"
}
if @company_ztate {
        "\newcommand{\\state}{"@company_state"}"
}
if @company_zip {
        "\newcommand{\\state}{"@company_zip"}"
}
\\begin{document}
\\AddToShipoutPrictureB6{\\colorigr}
\\AddToShipoutPrictureB6{\\colorigr}
\\AddToShipoutPrictureB6{\\colorigr}
\\AtPageUpperLeft{\\rule[-1.3in]{\\paperwidth}{1.3in}}
}
\\begin{document}
\\href{mailto:\\mynemail}{\\faEnvelope\\enspace \\mynemail}\\hfill"
if @page{
        "\href{\\mynage}{\\faEnvelope\\enspace My Website}\\hfill"
}
elif @git{
        "\href{\\mynage}{\\faEit*\enspace Git Page}\\hfill"
}
*\href{\\mynage}{\\faGit*\enspace LinkedIn}\\hfill"
}
*\href{\\mynage}{\\faGit*\enspace \\mynhone}\\hfill
\\faMapMarker\\enspace \\mylocation
\\end{end(center)}
```

Figura 7.15: Excerto do *Template* de Cartas de Motivação

```
name: Arlindo Barnabé
age: 20
job: Student
title: Software Engineering Student
company: Sports Interactive
company_state: London
company_city: Stratford
company_country: England
company_sector: Games Programming
university: Universidade do Minho
course: Software Engineering
year: 3rd
area: Software Engineering
recipient: Miles Jacobson
greeting: Dear
page: https://beasrodrigues24.github.io/yaPP/
email: arlindobarney2209@gmail.com
phone: +351 911111111
location: Portugal
application_for: Summer Internship
```

Figura 7.16: Dicionário exemplificativo para carta de motivação



Best wishes and regards, Arlindo Barnabé Software Engineering Student

Figura 7.17: Ficheiro LaTeX gerado

8. Extras Estéticos

8.1 Tratamento de Erros

Considerando a perspetiva do utilizador, foi considerada relevante a implementação de um tratamento de erros elegante e apelativo. Assim, foi desenvolvida uma classe denominada *PrettyPrint* que recebendo uma *String*, possui métodos que permitem a exibição de erros ou *warnings* ao utilizador de forma colorida.

```
[tgvp@tgvp-mKx recipes]$ python3 ../../src/koala.py -dt yaml recipes.koa recipes.yaml recipes.html Error: recipes.koa:15: Syntax error. Token 'recipe' not expected.
```

Figura 8.1: Exemplo de Mensagem de Erro

[tgvp@tgvp-mKx recipes]\$ python3 ../../src/koala.py -dt yaml recipes.koa recipes.yaml recipes.html Warning: recipes.koa:16: Invalid token: ';'. Skipping...

Figura 8.2: Exemplo de Mensagem de Warning

8.2 Syntax Highlighting

De forma a permitir uma utilização mais simples e prazerosa do programa, foi ainda criada uma extensão para *Microsoft Visual Studio Code* que permite adicionar *syntax highlight* a ficheiros com a extensão .koa.

Para isso, o utilizador deverá dirigir-se ao marketplace do VS Code e a partir de lá efetuar a instalação de uma extensão externa através do ficheiro .vsix fornecido.

Podem ser vistos exemplos desta utilidade através das imagens de templates de exemplo anteriormente fornecidas.

9. Conclusões e Trabalho Futuro

Este projeto permitiu o reforço de conhecimentos obtidos na unidade curricular de Processamento de Linguagens, permitindo ainda a compreensão do poder de uma linguagem de *templates* e a verificação do vasto leque de utilizações possíveis. Considera-se que foram implementadas bastantes funcionalidades criativas e convenientes, que tornam esta ferramenta interessante e útil.

É possível afirmar que foi produzido um resultado positivo, que explorou as potencialidades que este programa poderia ter para além das mecânicas inspiradas pelo Pandoc.

Por fim, considera-se ainda que o koal@ é um programa passível de evolução podendo oferecer melhores e mais vantagens aos seus utilizadores como, por exemplo, os diversos *pipes* que o Pandoc oferece que ainda não foram implementados ou a inserção de novos *templates* pré-definidos que permitiriam que o utilizador pudesse apenas alterar o conteúdo dos dicionários.

10. Bibliografia

- 1. Documentação do ply . Consultado pela última vez a 22 de maio de 2022.
- 2. **Documentação do Python** . Consultado pela última vez a 22 de maio de 2022.
- 3. **Documentação do JSON** . Consultado pela última vez a 22 de maio de 2022.
- 4. **Documentação do YAML** . Consultado pela última vez a 22 de maio de 2022
- 5. **Documentação do Pandoc** . Consultado pela última vez a 22 de maio de 2022.
- 6. **Documentação do AsciiDoc** . Consultado pela última vez a 22 de maio de 2022.
- 7. W3Schools . Consultado pela última vez a 22 de maio de 2022.
- 8. ${\it Dataset}$ de Pokémon GO . Consultado pela última vez a 22 de maio de 2022
- 9. Template da Carta de Recomendação em LaTex . Consultado pela última vez a 22 de maio de 2022.