**Especificação funcional do Music Bot**

Sumário

[Objetivo 1](#_Toc42286382)

[Origem das informações 1](#_Toc42286383)

[Fluxo das informações 1](#_Toc42286384)

[Detalhamento das atividades 2](#_Toc42286385)

[Ambiente de desenvolvimento 2](#_Toc42286386)

[Função para limpeza de textos 2](#_Toc42286387)

# Objetivo

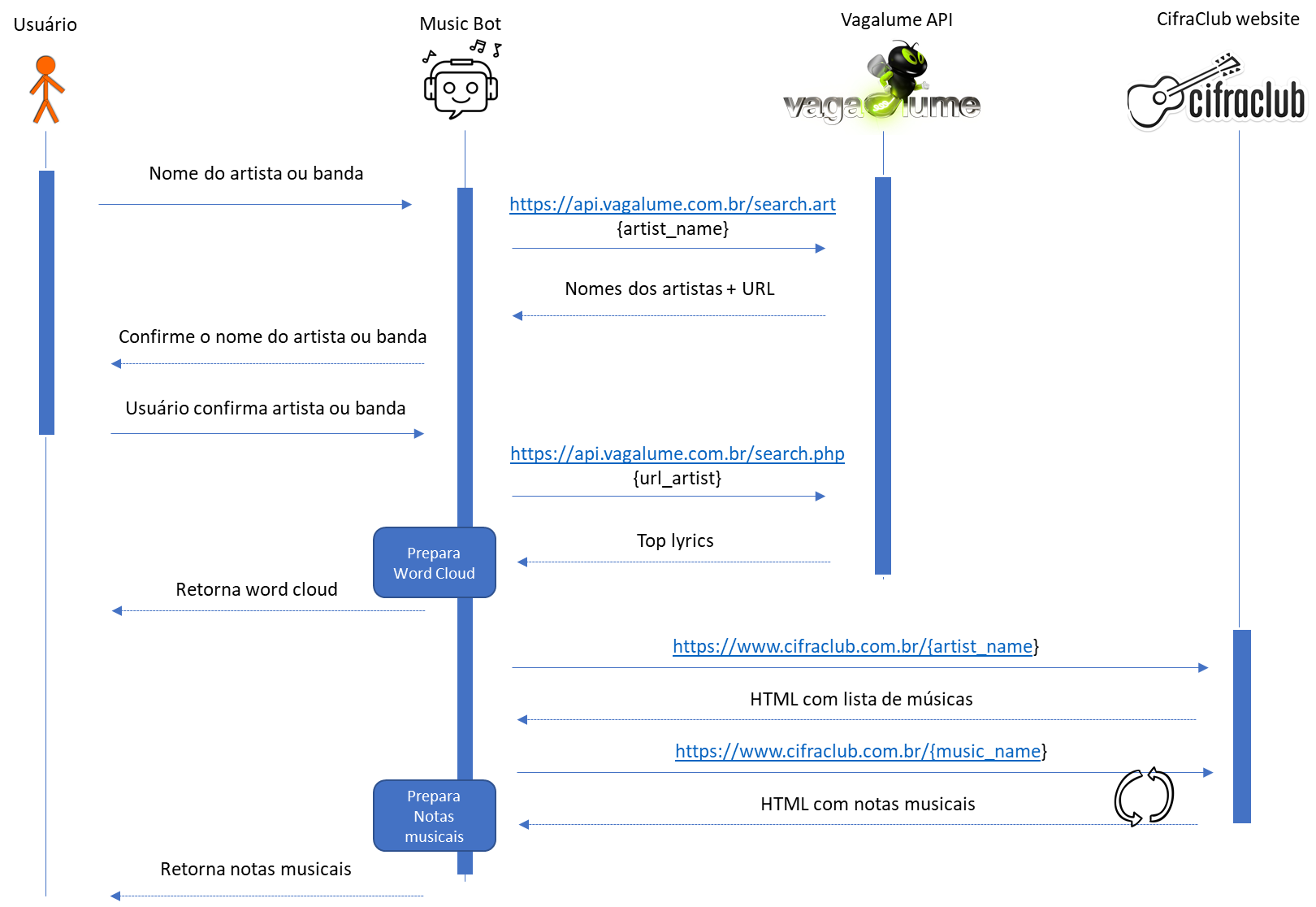
Criar um notebook interativo, capaz de capturar um nome de artista ou banda **digitado pelo usuário** e apresentar uma **nuvem de palavras** com os termos mais utilizados nas principais letras de músicas deste artista. Adicionalmente, o notebook também deve apresentar as **10 notas musicais mais utilizadas** nos acordes musicais do artista ou banda.

## Origem das informações

As letras de músicas do artista ou banda serão buscadas na **API do Vagalume** ([www.vagalume.com.br](http://www.vagalume.com.br/)). As informações sobre os endpoints específicos serão fornecidas no decorrer da documentação. Já as informações sobre notas musicais, serão capturadas por meio de web scraping no **site do CifraClub** (<https://www.cifraclub.com.br/>).

## Fluxo das informações

Com o objeto de facilitar a especificação técnica de cada atividade do projeto, o diagrama abaixo fornece uma visão geral da troca de informações entre usuário, notebook, API do Vagalume e site do Cifra Club. Use este diagrama apenas para entender, de forma geral, *como* e *quando* cada tipo de informação será capturada.



Perceba que teremos que fazer web scraping de várias páginas no site do CifraClub. Isso porque precisamos acessar URLs com o seguinte formato [*https://www.cifraclub.com.br/{music\_name}*](https://www.cifraclub.com.br/%7bmusic_name%7d)*.*

# Detalhamento das atividades

A seguir, serão detalhadas todas as atividades de desenvolvimento. Note que as atividades seguem uma sequência cronológica de acordo com a lógica de todo o processo.

## Ambiente de desenvolvimento

Siga as instruções do link a seguir para configurar o ambiente de desenvolvimento e cadastrar uma API Key: <https://github.com/weslleymoura/music-bot/blob/master/README.md>

## Função para limpeza de textos

Crie uma função que recebe um texto e um código de idioma. Esta função deve retornar o texto após realizar as seguintes etapas de limpeza:

* Remover qualquer pontuação existente
* Remover tokens não alfabéticos
* Remover palavras com apenas um caractere
* Remover stop words

Atenção: remova stop words de acordo com o idioma informado no parâmetro de entrada da função. Dentro da função, use o seguinte mapeamento de linguagens *{1: 'portuguese', 2: 'english', 3: 'spanish'}*

## Solicite o nome do artista

Solicite o nome do artista ou banda ao usuário e grave a informação em uma variável.

## Pesquise o nome do artista na API do Vagalume

Faça uma chamada ao endpoint <https://api.vagalume.com.br/search.art> e informe a seguinte query string *{"q":name\_artist, "limit": "10"}, onde nome\_artist* é o nomedo artista ou banda capturado na etapa anterior.

## Preparação dos nomes dos artistas para confirmação do usuário

A pesquisa anterior retorna, entre outras informações, o nome e a URL das bandas que fazem match com o termo da pesquisa. Use estas informações para preparar os dados de confirmação ao usuário. O objetivo desta tarefa é possibilitar que o usuário nos informe a URL que devemos utilizar deste ponto em diante. Use a função *ipywidgets.interact* para apresentar a informação ao usuário e solicitar a confirmação da pesquisa.

Pesquisando a li