

Haskeract

Serviço para extração do texto em imagens feito em Haskell.

Dependências

Para que o projeto seja executado localmente (em um ambiente Linux), certifique-se de que as dependências

- zlib1g-dev
- build-essential
- libffi-dev
- libgmp-dev
- libgmp10
- libncurses-dev
- tesseract-ocr
- ghc v9.4.8
- cabal v3.10.3.0

estejam instaladas no sistema.

Instalando e executando localmente

Para rodar o **ImageTextExtractor** localmente,

- Atualize o repositório de dependências, para que os pacotes do repositório Hackage possam ser instalados, através do seguinte comando:

```
cabal update
```

- Execute o projeto com o comando

```
cabal run
```

Relatório

Dificuldades

Nossa principal dificuldade em desenvolver este projeto foi encontrar uma biblioteca que permitisse a criação de um servidor web para upload e processamento de arquivos. Inicialmente nós tentamos utilizar o Integrated Haskell Platform (IHP), que é um Framework para criar aplicações web com a arquitetura MVC (Model View Controller). Entretanto, diferente do Scotty, ele não oferece a possibilidade de acessar o conteúdo do arquivo submetido ou o seu caminho no sistema de arquivos.

O Scotty, ao receber uma requisição HTTPS com o método POST e conteúdo do tipo “application/x-www-form-urlencoded”, coloca os arquivos contidos nela temporariamente na past “/tmp” do servidor, para que ele possam ser processados. A partir disso nós precisamos pesquisar sobre funções que permitissem a execução de comandos parametrizados no servidor e que retornassem informações sobre os processos, como a saída do comando e se houve sucesso ou falha em sua execução.

Destaques

- O projeto faz intenso uso de Monads para realização de operações sensíveis, como a leitura de arquivos e resultados de processos;
- Tratamento de erros quando nenhum arquivo é enviado ou caso ocorra algum problema na execução do programa tesseract.

Surpresas

Nós só conseguimos utilizar a biblioteca Scotty com o Cabal ao invés do Stack. No Stackage a versão disponível do Scotty é muito antiga, e não contém algumas funções que foram necessárias para o projeto. No Hackage a versão disponível é a mais recente.

Dados da entrega

- Integrantes:
 - Thiago Fernandes Dias, RA 11202130847;
 - William Fernanes Dias, RA 11202020043.
- URL do vídeo: <https://youtu.be/XQOH9VIHU9k>