

Ferramenta para construção de datasets colaborativos

Definição de Projeto

Objetivo

Construir uma ferramenta *crowd based* para geração de datasets rotulados, é, em que para cada entrada X há um (ou mais) rótulo(s) Y. A idéia principal é fornecer ao usuário uma maneira prática de:

1. Poder criar um dataset para um determinado problema
2. Poder compartilhá-lo com amigos ou com o público geral
3. Poder colaborar com outras pessoas para popular este dataset

O que será feito

1. Backend → Rotas para comunicação com banco de dados (CRUD app basicamente) + processamento de dados (possivelmente algo com estatística/IA)
2. Frontend Web → Interação com o Backend
3. Frontend Mobile → Interação com o Backend (imagino que isso possa ser feito depois dos dois primeiros)

Usuários

A ferramenta conta com 2 tipos de usuários: **autenticados e não autenticados**.

Requisitos Essenciais

Nome	Descrição	Prioridade
Criar Dataset	O usuário autenticado pode criar um dataset	Alta
Definir a privacidade de um dataset	<p>O usuário autenticado pode definir o status de privacidade do seu dataset de acordo com 2 aspectos: visualização e contribuição.</p> <p>Visualização</p> <ol style="list-style-type: none">1. Pública: todos podem ver o dataset2. Não-Listado: qualquer pessoa com link pode ver o dataset3. Privado: apenas pessoas convidadas podem ver o dataset <p>Contribuição</p> <ol style="list-style-type: none">1. Pública: todos podem contribuir com o dataset2. Privado: apenas pessoas convidadas podem	Média

	contribuir com o dataset	
Compartilhar dataset	O usuário, possuindo permissão (haja visto a política acima), pode compartilhar o dataset.	Média
Inserir Entrada no Dataset	<p>O usuário, possuindo permissão (haja visto a política acima), pode inserir uma entrada no Dataset.</p> <p>No caso de imagens, elas podem ser enviadas através de upload ou tiradas com a câmera do dispositivo (caso haja).</p> <p>No caso de texto, eles podem ser enviados através de upload de arquivos (.txt, .csv) ou digitados manualmente.</p> <p>Para ambos casos de upload, deve ser possível o envio de múltiplos arquivos simultaneamente.</p> <p>Os rótulos podem constar nos arquivos enviados ou serem digitados manualmente.</p>	Alta
Remover Entrada no Dataset	O usuário, possuindo permissão (haja visto a política em “ <i>Definir a privacidade de um dataset</i> ”), pode remover uma entrada do dataset	Baixa
Exportar Dataset	O usuário, possuindo permissão (haja visto a política em “ <i>Definir a privacidade de um dataset</i> ”), pode baixar o dataset (.csv e/ou imagens)	Média

Requisitos Extra

Esses podem ser diferenciais do projeto...

Nome	Descrição	Prioridade
Análise Estatística dos Datasets	<p>Seria possível fazer alguma análise estatística dos dados gerados e informá-los ao usuário?</p> <p>Ex.: Média de alguma feature, gráficos, análise de quantis, outliers, etc</p>	?
Análise IA	<p>Seria possível fazer alguma análise com IA dos dados gerados e informá-los ao usuário?</p> <p>Ex.: Verificar se a imagem inserida realmente é do contexto proposto pelo dataset</p>	?

Aplicativos Referência

- [Memento: Smart Database & AI - Apps on Google Play](#)

- [Dataset Creator - Apps on Google Play](#)