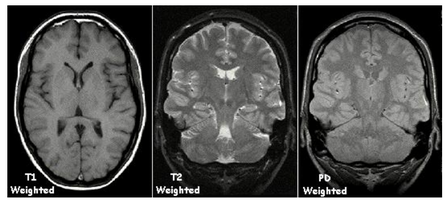
Bancos de dados de imagens
Uma aplicação em Neurociência

Jéssica Yumi Kojima
Orientadora: Kelly Rosa Braghetto

IME-USP

Contato:
Instituto de Matemática e Estatística
Universidade de São Paulo

Email: `yumioh@usp.br`



Introdução

A quantidade de experimentos de neurociência que envolvem a aquisição de imagens, chamados de **experimentos de neuroimagem**, vêm aumentando nos últimos anos a medida que se ampliou a exploração das funções cognitivas pelos cientistas.

Uma tendência forte nessa área são experimentos de **ressonância magnética funcional** (fMRI, do inglês *functional Magnectic Resonance Imaging*), devido a alta qualidade das imagens e por ser uma abordagem não nociva à saúde, como ocorre com as máquinas de raio-X.

Contudo, a complexidade de tais estudos de neuroimagem dificultam o armazenamento completo de todos os dados da pesquisa. Assim, muitas vezes eles não são relatados com um nível detalhes suficientes para a reprodução exata do estudo por outros neurocientistas, ou até para a continuação do mesmo.

Além disso, a maioria dos bancos de dados com boa documentação e uma diversidade de projetos disponíveis são fechados, devido à apreensão da comunidade científica em misturar dados confiáveis com duvidosos. Porém, o tempo gasto para a coleta de dados suficientes para provar a hipótese de um experimento é tão grande, que estudos iriam se beneficiar muito ao **reaproveitar os dados** já armazenados por outrem.

Logo, este projeto propõe um **modelo de banco de dados** relacionados à aquisição de imagens de fMRI, e uma possível solução de código aberto para sua a implementação.

Motivação

1. **Incentivar** projetos de **código aberto**
2. Incentivar **banco de dados abertos** ao público
3. Facilitar o **compartilhamento** e **reuso de dados de neurociência**
4. **Facilitar o armazenamento dos dados** de experimentos associados a imagens de fMRI, visando contribuir com os avanços de conhecimento em neurociência

Análise dos Bancos de Dados Existentes

Foram comparados alguns dos bancos de dados mais populares em neurociência atualmente, a fim de explicitar as vantagens e disvantagens de cada um e entender quais pontos precisam ser melhorados para modelar um banco de forma mais completa possível.

Critério	ADNI	OpenfMRI	OASIS	Carmen	Neurosynth	Neurovault	Brain-Development	Brede Database	Whole Brain Atlas	BrainMap	BiND	MaND	MSU Brain Biodiversity Bank	BrainMuseum
Acesso aberto ao público	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Não precisa fazer login para ver os dados		x	x		x	x	x	x	x		x	x	x	x
Os dados estão disponíveis para download	x	x	x	x	x	x	x	x		x	x	x	x	x
Possui uma licença definida	x	x	x				x		x	x				
Não possui copyright		x	x	?	x	x					x	x		x
Pode fazer upload de dados	x	x		x	x	x				x				
Precisa de registro para fazer upload	x	x		x	x	x	-	-	-	x				
Possui formulário de upload	x	x		x	x	x				x				
Pouca burocracia para se registrar		x	x	x	x	x	-	-	-	x	-	-	-	-
Não precisa instalar software ou fazer download	x	x	x	x	x	x	x	x	x				x	x
É amigável ao usuário	x	x	+-	x	x	x	+-		x	+-	+-	+-	+-	+-
Possui documentação para usuário	x	x		x					x	x	x	x		+-
Possui artigos disponíveis	x	x	x	x	x	x	x	x		x	x	x		
O sistema tem manutenção	x	x		x	x	x		?	?	x				
Possui um sistema de busca adequado	x			x				x		x				
Possui ferramenta para visualizar imagens				?	x	x			x	x				
Código aberto					x	x								
Armazena imagens de MRI	x	x	x	x			x		x				x	
Armazena metadados	x	x	x	x	x	x	x	x		x	x	x		x
Armazena dados brutos	x	x	x	?				x		x				
Armazena dados de seres humanos	x	x	x	x	x	x	x	x	x	x	x	x	x	
Rico em dados armazenados	x	x	x	x						x				
Os documentos estão padronizados	x	x	x		x	x	x	x	x		-	-		x

Tabela 1: Tabela de comparação dos BDs

Ressonância Magnética Funcional

A fMRI é uma técnica que usa as propriedades eletromagnéticas das moléculas para formar imagens, capaz de detectar variações no fluxo sanguíneo em resposta à atividade neuronal [1]. Assim, ela é utilizada para entender como diferentes partes do cérebro reagem a estímulos externos.

Para analisar os dados obtidos por fMRI, são aceitos diversos pacotes de software e modelos estatísticos. Tal plasticidade é tanto uma vantagem quanto desvantagem, pois essa abrangência acaba tornando o processo de análise dos dados ainda mais complexo. Modelagens estatísticas diferentes e até algoritmos diferentes de uma única modelagem podem gerar resultados diferentes para um mesmo conjunto de dados. E a própria ordem de um *workflow* também pode alterar os reslultados, já que cada método estatístico envolve uma probabilidade erro, que aumenta conforme a quantidade de dados. Assim, a análise dos dados de neurociência é bem complexa, visto que uma única sessão de fMRI chega a gerar arquivos de 6MB. Logo, esse trabalho foca apenas no processo de aquisição de dados de fMRI.

Item	Legenda
x	sim
+-	mais ou menos
-	não se aplica
?	não foi possível definir
(em branco)	não

Tabela 2: Legenda dos itens da tabela de BDs

Modelo de Banco de Dados Proposto

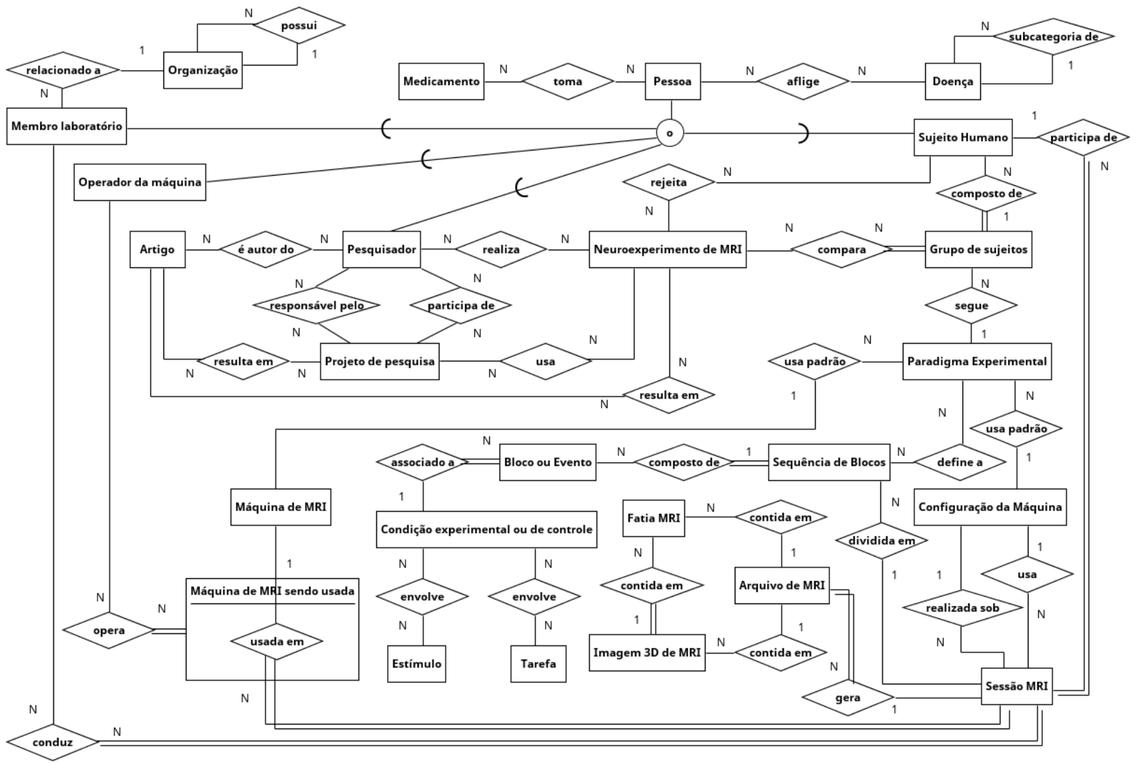


Figura 1: Modelo do banco de dados proposto sem os atributos

Ferramenta XNAT

A ferramenta XNAT (do inglês *Extensible Neuroimaging Archive Toolkit* já é utilizada por alguns bancos de dados, como o OASIS (do inglês *Open Access Series of Imaging Studies*). Como essa ferramenta possui código aberto [2], ela pode ser usada para implementar o modelo de banco de dados proposto.

Considerações Finais

- Para ficar completo, o modelo precisa abranger as modelagens estatísticas usadas nas análises dos dados.
- É importante melhorar a documentação dos bancos de dados abertos para incentivar o seu uso.

Referências

[1] Scott A. Huettel, Allen W. Song, and Gregory McCarthy. *Functional Magnectic Resonance Imaging*. Sinauer, 1st edition, 2004.

[2] Daniel S. Marcus, Timothy R. Olsen, Mohana Ramaratnam, and Randy L. Buckner. The extensible neuroimaging archive toolkit: An informatics platform for managing, exploring, and sharing neuroimaging data. *Neuroinformatics*, 5(1):11–34, March 2007.