

**Universidade do Minho**  
Escola de Engenharia

Filipe Bernardino Fernandes

**Arquitetura Distribuída para Análises  
Multimodais de Conectividade Cerebral**



**Universidade do Minho**  
Escola de Engenharia

Filipe Bernardino Fernandes

## **Arquitetura Distribuída para Análises Multimodais de Conectividade Cerebral**

Dissertação de Mestrado  
Mestrado Integrado em Engenharia Biomédica  
Ramo Informática Médica

Trabalho efetuado sob orientação do  
**Professor Doutor Victor Manuel Rodrigues Alves**  
e do  
**Mestre Paulo César Gonçalves Marques**

# Declaração

**Nome:** Filipe Bernardino Fernandes

**Endereço eletrónico:** fernandesbiomed@hotmail.com

**Cartão de Cidadão:** 14088597

**Título da Dissertação:** Arquitetura Distribuída para Análises Multimodais de Conectividade cerebral

**Orientador:** Professor Doutor Victor Manuel Rodrigues Alves

**Supervisor:** Mestre Paulo César Gonçalves Marques

**Ano de conclusão:** 2015

**Designação do Mestrado:** Mestrado Integrado em Engenharia Biomédica  
Ramo de Informática Médica

DE ACORDO COM A LEGISLAÇÃO EM VIGOR, NÃO É PERMITIDA  
A REPRODUÇÃO DE QUALQUER PARTE DESTA DISSERTAÇÃO.

Universidade do Minho, \_\_ / \_\_ / \_\_

Assinatura: \_\_\_\_\_

# Agradecimentos

Gostaria de agradecer em primeiro lugar ao meu orientador, o professor Vítor Alves, pela disponibilidade prestada e pela ajuda na elaboração deste projecto. Sem menor graditão, uma palavra de apreço aos Mestres Paulo Marques e Ricardo Magalhães, pelo tempo e paciência demonstrada. Agradecer todas as suas sugestões, que se revelaram essenciais no desenvolvimento deste trabalho. Queria também demonstrar a meu reconhecimento à Xpand-IT pela flexibilidade concedida e que permitiu conciliar a minha primeira experiência profissional e a conclusão deste projecto.

*Sentes que o tempo acabou, primavera de flores adormecida, qualquer coisa que não volta, que voou..*

Acabou, tinha de acabar, só assim se define como uma etapa. Foram muitos os momentos, muitas as pessoas e uma infindável lista de memórias que resultaram deste feliz percurso académico. Não posso por isso deixar de agradecer aos meus amigos, os de sempre e os mais recentes, que comigo partilharam sorrisos, preocupações, a sua vida. A todos os que me transmitiram valores de solidariedade, união e respeito, o meu sincero agradecimento. A quem através da sua verdadeira amizade conquistou um lugar especial em mim, obrigado. É impossível não ceder à saudade no momento de acabar, mas fica a esperança de que as loucuras vividas ainda estão para começar.

Um agradecimento puro aos meus pais. Torna-se doloroso expressar o carinho e reconhecimento que tenho por vós, pois ultrapassa fortemente a minha capacidade de expressão. Obrigado pelo orgulho feroz que têm por mim, são os maiores responsáveis pelo meu sucesso.

Por último, um agradecimento muito especial aos meus avós. Obrigado pelo carinho, suporte e cumplicidade que sempre demonstraram. São um exemplo de humildade e luta intemporal que guardarei para sempre, e merecem de mim a maior estima e reconhecimento.

Dedico-te esta dissertação. Tu que partiste quando não merecias, não mereces e nunca merecerás. Foste primo, amigo, irmão. Eras força e perseverança e demonstraste a todos os que tiveram o privilégio de te conhecer, que apesar das adversidades que a vida te colocou é possível ser feliz e contagiar o mundo de alegria.

Nunca te esquecerei.

# Resumo

O estudo da conectividade, estrutura, e integração das funções cerebrais é actualmente uma das ferramentas de maior importância na compreensão do cérebro humano. A realização destes estudos via aquisições de Ressonância Magnética exige no entanto acessibilidade e disponibilidade constante de informação. A quantidade de procedimentos e técnicas de análise, associado à produção de grandes volumes de dados e multitude de soluções de software são alguns dos principais entraves à organização, manutenção e partilha de estudos neuroimagiológicos.

Desta forma, o objectivo principal deste trabalho consiste na concepção de um fluxo de processamento, que possa servir de padrão à conjugação de resultados de análises multimodais, através de uma estrutura de estudos definida e nomenclatura de ficheiros própria. Tendo por base este fluxo foi desenvolvida uma aplicação, designada BrainArchive, para automatização do processo de organização e partilha de estudos neuroimagiológicos. Esta permite por sua vez a disponibilização de grandes volumes de informação sem necessidade de caracterização manual de cada ficheiro, ao mesmo tempo que se revela uma ferramenta simples e intuitiva na aquisição de dados.

O protótipo desenvolvido responde às necessidades do seu contexto de utilização e, por conseguinte, é espectável que potencie o processo de investigação, através da simplificação e redução do tempo despendido na organização e partilha de informação.



# Abstract

The study of the connectivity, structure and integration of brain functions is nowadays one of the most relevant tools in the comprehension of the human brain. The realization of these MRI studies requires constant accessibility and availability of information. The amount of procedures and analytical techniques, associated with the production of large volumes of data and multitude of software solutions are among the main obstacles to organization, maintenance and sharing of neuroimaging studies.

Thus, the main objective of this project consists on the conception of a processing flow, which can serve as a standard to the combination of results from multimodal analysis, through a defined study structure and file nomenclature. Based on this flow was developed an application, designated BrainArchive, for automating the process of organization and sharing of neuroimaging studies. This in turn allows the delivery of large volumes of data, without the need for manual characterization of each file, while reveals a simple and intuitive tool for data acquisition.

The developed prototype meets the needs of the context of use and therefore it is expected to enhance the research process, through the simplification and reducing of the time spent organizing and sharing information.



# Conteúdo

<b>Acrónimos</b>	<b>xxiv</b>
<b>1 Introdução</b>	<b>1</b>
1.1 Enquadramento . . . . .	1
1.2 Problema . . . . .	7
1.3 Objectivos . . . . .	11
1.4 Metodologia de Investigação . . . . .	11
1.5 Estrutura do documento . . . . .	12
<b>2 Processamento de dados de Ressonância Magnética</b>	<b>15</b>
2.1 Introdução . . . . .	15
2.2 Ressonância Magnética Estrutural . . . . .	17
2.2.1 Segmentação de imagens estruturais . . . . .	18
2.2.2 Processamento de structural Magnetic Resonance Imaging (sMRI) . . . . .	18
2.3 Ressonância magnética funcional . . . . .	19
2.3.1 Pré-processamento de fMRI e sMRI . . . . .	19
2.3.2 Processamento de fMRI . . . . .	25
2.4 Ressonância Magnética de Difusão . . . . .	26
2.4.1 Pré-processamento de DTI . . . . .	27
2.4.2 Processamento de DTI . . . . .	28
2.5 Análise de Grafos . . . . .	30
2.5.1 Definição de Vértices e Arestas . . . . .	30
2.5.2 Análise . . . . .	32
2.6 Fluxo de processamento . . . . .	32

2.7	Volume de dados . . . . .	44
2.8	Soluções Disponíveis . . . . .	46
2.8.1	ANIMA . . . . .	46
2.8.2	1000 Functional Connectones Project (FCP) . . . . .	47
2.8.3	GAAIN . . . . .	49
2.8.4	XNAT . . . . .	50
<b>3</b>	<b>Tecnologia</b>	<b>55</b>
3.1	Introdução . . . . .	55
3.2	Arquiteturas de Redes . . . . .	55
3.2.1	Cliente-Servidor . . . . .	56
3.2.2	Peer-to-Peer . . . . .	57
3.2.3	Comparação Cliente/Servidor e Peer-to-Peer (P2P) . . . . .	58
3.3	Base de Dados . . . . .	60
3.3.1	Relacionais . . . . .	60
3.3.2	NoSQL . . . . .	63
<b>4</b>	<b>BrainArchive</b>	<b>71</b>
4.1	Introdução . . . . .	71
4.2	Análise de Requisitos . . . . .	72
4.3	Base Tecnológica . . . . .	73
4.4	Hierarquia e Nomenclatura de Estudos . . . . .	83
4.5	Arquitetura e Implementação do Repositório . . . . .	89
4.6	Interface de Utilizador . . . . .	96
4.7	Principais funcionalidades . . . . .	102
<b>5</b>	<b>Resultados e Conclusões</b>	<b>105</b>
5.1	Sinopse . . . . .	105
5.2	Resultados . . . . .	107
5.3	Conclusões . . . . .	111
	<b>Apêndices</b>	<b>129</b>

<b>A</b>	<b>Ressonância Magnética</b>	<b>129</b>
A.1	Princípios de funcionamento . . . . .	129
A.2	Ressonância Magnética estrutural . . . . .	131
A.2.1	Volumetria . . . . .	132
A.2.2	Voxel-based Morphometry . . . . .	133
A.2.3	Cortical thickness analysis . . . . .	134
A.3	Ressonância Magnética funcional . . . . .	135
A.3.1	Princípios básicos . . . . .	135
A.3.2	Task-related fMRI . . . . .	136
A.3.3	Resting-state fMRI . . . . .	137
A.3.4	General Linear Model . . . . .	138
A.3.5	Seed Correlation Analysis . . . . .	140
A.3.6	Independent Component Analysis . . . . .	141
A.3.7	Análise de grafos . . . . .	141
A.4	Ressonância Magnética de difusão . . . . .	143
A.4.1	Aquisição e Tensor de difusão . . . . .	144
A.4.2	Visualização de dados de DTI . . . . .	145
A.4.3	Tratografia . . . . .	146
A.4.4	Análise quantitativa . . . . .	148
<b>B</b>	<b>Nomenclatura e descrição de ficheiros</b>	<b>151</b>
<b>C</b>	<b>BrainArchive - Descrição Funcional</b>	<b>155</b>
C.1	Autenticação e exportação de um ficheiro . . . . .	155
C.2	Importação local assistida . . . . .	161
C.3	Exportação automática de estudo . . . . .	163
C.4	Alteração de permissões em ficheiros presentes no repositório . . . . .	165
C.5	Extração de ficheiros presentes no repositório . . . . .	169
C.6	Outras considerações de utilização . . . . .	171
C.6.1	Exportação de ficheiros que não cumpram com a estruturação hierárquica definida . . . . .	171
C.6.2	Exportação de ficheiros que se encontrem já presentes no repositório . . . . .	173



# Lista de Figuras

1.1	Domínios fundamentais de Informática Médica . . . . .	2
1.2	Padrão de crescimento na utilização de Magnetic Resonance Imaging (MRI) em países da Organização para a Cooperação e Desenvolvimento Económico (OCDE) (retirado de [11]) . . .	4
1.3	Visão geral do fluxo de processamento de dados <i>BrainCat</i> (retirado de [43]) . . . . .	10
2.1	Fluxo de pré-processamento de imagens funcionais e estruturais. As imagens estão representadas a tracejado e os passos de processamento em caixas contínuas. O fluxo das imagens funcionais está representado a vermelho e as estruturais a verde. Passos comuns são representados a azul e quando o resultado de um fluxo é utilizado no outro encontra-se em castanho. . .	20
2.2	Fluxo de processamento global para estudos MRI . . . . .	34
2.3	Fluxo de processamento para análises funcionais de MRI . . .	42
2.4	Fluxo de processamento para análises de difusão de MRI . . .	43
2.5	Fluxo de processamento para análises estruturais de MRI . .	43
2.6	Volume de dados aproximado para um caso completo de estudo de MRI . . . . .	45
2.7	Página de pesquisa da interface online Anima (retirado de [39])	47
2.8	Página de submissão de estudos da interface online ANIMA. (A) <i>Study tab</i> (B) <i>Elements tab</i> (C) <i>Submit tab</i> (retirado de [39])	48
2.9	Visão geral de utilização da base de dados ANIMA, ilustrando o processo de pesquisa, seleção, <i>download</i> e importação de dados na biblioteca <i>Volume Viewer</i> (retirado de [39]) . . . . .	49

2.10	Fluxo de informação presente na arquitetura Global Alzheimer's Association Interactive Network (GAAIN) (retirado de [64]) . . . . .	51
2.11	Esboço da arquitetura cliente/servidor GAAIN (retirado de [64])	52
2.12	Arquitetura Extensible Neuroimaging Archive Toolkit (XNAT) como serviço de captura de dados em múltiplas fontes e distribuição a uma variedade de utilizadores (retirado de [65]) . .	53
2.13	Fluxo de trabalho da solução XNAT (retirado de [65]) . . . . .	53
3.1	Troca de mensagens padrão numa rede Cliente-Servidor . . . .	56
3.2	Arquitetura Cliente-Servidor . . . . .	57
3.3	Arquitetura Peer-to-Peer . . . . .	58
3.4	Evolução do volume e tipo de dados gerados desde os anos 80 até ao presente (retirado de [75]) . . . . .	62
3.5	Teorema de Consistency, Availability, Partition-Tolerance (CAP) (adaptado de [83] ) . . . . .	66
4.1	Eclipse Kepler Service Release 1 na perspectiva PyDev . . . . .	75
4.2	Interface Qt Designer 5.4.1 . . . . .	76
4.3	Modelo de dados documental representativo de uma aplicação de <i>blogging</i> em MongoDB. (retirado de [91]) . . . . .	78
4.4	Representação da divisão em duas coleções com GridFS (retirado de [92]) . . . . .	80
4.5	MongoDB <i>Sharded Collection</i> retirado de [91] . . . . .	80
4.6	MongoDB <i>Sharded Cluster</i> retirado de [91] . . . . .	81
4.7	Esboço de um <i>replica set</i> MongoDB retirado de [91] . . . . .	82
4.8	Hierarquia de estudos definida para análises multimodais . . . .	84
4.9	Exemplo de organização de estudos segundo uma hierarquia de pastas . . . . .	85
4.10	Exemplo de conteúdo da diretoria <i>BedpostX.bedpostx</i> pertencente à hierarquia de estudos. . . . .	88
4.11	Arquitetura geral da aplicação BrainArchive . . . . .	90
4.12	Exemplo de documento de uma base de dados MongoDB . . . .	92
4.13	Exemplo de criação de utilizador numa base de dados MongoDB	93

4.14	Exemplo de criação de instância <i>mongod</i> com suporte Secure Socket Layer (SSL) . . . . .	94
4.15	Esboço da interface de autenticação e mensagem de falha da mesma na aplicação BrainArchive. . . . .	96
4.16	Divisão lógica da interface inicial da aplicação BrainArchive. (A) representação dos estudos presentes localmente para um utilizador. (B) seção de exportação de ficheiros/estudos. (C) estruturação de subpastas descendentes de localização selecionada em (A). . . . .	97
4.17	Interface da aplicação BrainArchive, para visualização de ficheiros presentes no repositório a que um utilizador tem acesso. . . . .	98
4.18	Interface de interrogação e importação da aplicação BrainArchive. (A) seção de filtragem de ficheiros. (B) alteração de permissões de acesso a ficheiros. (C) seção de visualização e importação de ficheiros presentes no repositório. . . . .	99
4.19	Janelas da interface BrainArchive para importação local assistida. (A) definição do estudo e número de casos. (B) seleção de sujeito/caso e diretorias a importar. . . . .	101
4.20	Ícone da aplicação BrainArchive . . . . .	102
5.1	Processo de exportação/importação de ficheiros entre utilizadores da aplicação BrainArchive. Linhas a vermelho representam processos de exportação e a verde de importação. . . . .	109
5.2	Processo de exportação e importação de estudos através da aplicação BrainArchive, baseado em hierarquia e nomenclatura definida. . . . .	111
A.1	Siemens Magnetom Avanto (retirado de [95]) . . . . .	130
A.2	<i>Output</i> do fluxo de processamento Freesurfer: imagens de segmentação cerebral segundo os atlas Destrieux e Desiken. . . . .	133
A.3	Resting state networks obtidas através de análise de componentes independentes. . . . .	139

A.4	Convulsão do sinal HRF com a matriz de respostas ao paradigma experimental que dá origem ao sinal funcional esperado. Adaptado de [123] . . . . .	139
A.5	Vista sagital de um conectoma funcional. Cores dos nós identificam o módulo a que pertencem e cor das ligações reflete a intensidade da ligação. . . . .	143
A.6	Elipsoides que resultam da combinação dos vetores (a) e valores próprios (b) obtidos na análise de tensores de difusão. Representação gráfica da anisotropia fracional. . . . .	145
A.7	Vistas sagital (a) e coronal (b) do resultado da aplicação de algoritmo de tratografia probabilística do Trackvis. . . . .	147
C.1	Autenticação de um utilizador na aplicação BrainArchive . . .	155
C.2	Confirmação de paralelismo entre a diretoria <i>MyBrain</i> e a aplicação BrainArchive. (A) Exemplo de estruturação na diretoria <i>MyBrain</i> . (B) Interface inicial da aplicação BrainArchive . . .	157
C.3	Exportação de um ficheiro representado na interface da aplicação BrainArchive. . . . .	158
C.4	Alteração no estado de exportação (quadrado de estado a verde) representado na interface BrainArchive . . . . .	159
C.5	Interface de interrogação da aplicação BrainArchive ( Confirmação de presença de ficheiro exportado para o repositório) .	160
C.6	Paralelismo evidenciado entre a diretoria <i>MyBrain</i> e interface da aplicação BrainArchive (ambas vazias). (A) diretoria <i>MyBrain</i> . (B) interface da aplicação BrainArchive. . . . .	161
C.7	Janelas da aplicação BrainArchive para importação local assistida de estudos. (A) Definição do estudo e número de sujeitos/casos. (B) Definição de cada sujeito/caso e seleção das opções a importar. . . . .	162
C.8	Confirmação de importação local de Estudo na interface inicial da aplicação BrainArchive. . . . .	163
C.9	Confirmação de exportação de Estudo na interface de interrogação da aplicação BrainArchive. . . . .	164

C.10	Confirmação de inacessibilidade de ficheiros exportados por um utilizador por parte de outro, na interface de interrogação da aplicação BrainArchive. . . . .	166
C.11	Visualização exclusiva de ficheiros exportados por um utilizador, na interface de interrogação da aplicação BrainArchive. . . . .	166
C.12	Esboço de extensão de permissões a um utilizador por parte de outro, na Interface de interrogação da aplicação BrainArchive. . . . .	167
C.13	Filtragem de ficheiros segundo parâmetros desejados, na interface de interrogação da aplicação BrainArchive. . . . .	167
C.14	Confirmação de partilha de ficheiros exportados por um utilizador para outro, na interface de interrogação da aplicação BrainArchive. . . . .	168
C.15	Extração de ficheiros por parte de um utilizador, na interface de interrogação da aplicação BrainArchive . . . . .	169
C.16	Confirmação de extração de ficheiros para a diretoria <i>MyBrain</i> . . . . .	170
C.17	Resultado de extração de ficheiros evidenciado na interface inicial da aplicação BrainArchive . . . . .	170
C.18	Exportação de ficheiro sem estruturação definida. (A) Visualização do ficheiro adicionado à diretoria <i>MyBrain</i> . (B) Processo de exportação manual do ficheiro na interface inicial da aplicação BrainArchive . . . . .	171
C.19	Confirmação do sucesso de exportação de ficheiro desestruturado. (A) realocação do ficheiro segundo a estrutura definida pelo utilizador na interface inicial da aplicação BrainArchive . (B) presença do ficheiro no repositório evidenciado na interface de interrogação da aplicação BrainArchive. . . . .	172



# Lista de Tabelas

B.1	Dicionário de ficheiros para a análise <i>FMRI</i> . . . . .	152
B.2	Dicionário de ficheiros para a análise <i>DTI</i> . . . . .	153
B.3	Dicionário de ficheiros para a análise <i>MRI</i> . . . . .	153
B.4	Dicionário de ficheiros de utilização do software TrackVis, Fressurfer e Bayesian Estimation of Diffuion Parameters using Sampling Techniques (BEDPOSTX), respectivamente . . . . .	154



# Acrónimos

Dado que a maioria dos acrónimos e termos técnicos, normalmente se usam em inglês, optou-se por manter estes termos em inglês.

**AAL** Automated Anatomic Labelling. 31

**ACID** Atomicity, Consistency, Isolation e Durability. 64, 67

**AD** Axial Diffusivity. 39

**ADC** Apparent diffusion coefficient. 28

**AFNI** Analysis of Functional Neuroimages. 22, 24

**BASE** Basically Available, Soft-state, Eventually consistent. 66, 67

**BDR** Bases de Dados Relacionais. 61–64, 69

**BEDPOSTX** Bayesian Estimation of Diffuion Parameters using Sampling Techniques. xix, 29, 35, 40, 84, 87, 154

**BET** Brain Extraction Tool. 23, 28, 38, 39

**BOLD** Blood Oxigen Level Dependent. 5, 19, 22, 25, 36, 135, 136, 138

**BSON** Binary JSON. 68, 77

**CAP** Consistency, Availability, Partition-Tolerance. xiv, 65, 66

**CSV** Comma-separated values. 50

**CT** Computed Tomography. 15, 129

**DICOM** Digital Imaging and Communications in Medicine. 21, 33, 36, 38, 40, 44, 52, 83, 86, 87, 91, 151, 165

**DMN** Default-mode network. 138

**DPARFS** Data Processing Assistant for Resting-State fMRI. 21

**DPC** Data Partner Client. 49, 50

**DTI** Diffusion Tensor Imaging. 5, 6, 8, 16, 17, 26–28, 35, 38, 71, 143–146, 148

**EEG** Electroencephalography. 3, 129

**ETL** Extract Transform Load. 63

**FA** Fractional Anisotropy. 28, 29, 35, 39, 86, 146, 148, 149

**FAST** FMRIB's Automated Segmentation Tool. 25, 41

**fcMap** functional connectivity Map. 140, 141

**FDT** FMRIB Diffusion Toolbox. 29, 35, 39, 86

**FEAT** FMRI Expert Analysis Tool. 25

**FFT** Fast Fourier Transform. 24

**FLIRT** FMRIB's Linear Image Registration Tool. 24, 25, 29, 37, 38, 40

**fMRI** functional Magnetic Resonance Imaging. 5, 6, 8, 16, 17, 19, 21, 22, 25, 28, 31, 36, 71, 135–138, 140–142

**FNIRT** FMRIB's Nonlinear Registration Tool. 19, 37, 40

**FSL** FMRIB'S Software Library. 19, 21–25, 27, 29, 33, 35, 36, 38, 87

**FWHM** Full Width at Half Maximum. 24

**GAAIN** Global Alzheimer's Association Interactive Network. xiv, 49–52, 110

**GB** Gigabyte. 44

**GLM** General Linear Model. 25, 33, 36, 41, 134, 138

**GUI** Graphical User Interface. 75, 76

**HIPAA** Health Insurance Portability and Accountability Act. 49

**HRF** Hemodynamic Response Function. 138–140

**ICA** Independent Component Analysis. 16, 26, 35, 37, 138, 141

**IDE** Integrated Development Environment. 74, 75

**IM** Informática Médica. 1–3

**INDI** International Neuroimaging Data-Sharing Initiative. 47

**JSON** Javascript Option Notation. 68

**Kb** KiloByte. 44

- LCR** Líquido cefalorraquidiano. 25, 41
- MB** Megabyte. 44, 87
- MCFLIRT** Motion Correction FMRIB's Linear Image Registration Tool. 36
- MD** Mean Diffusivity. 39, 145, 148
- MD5** Message-Digest algorithm 5. 173
- MELODIC** Multivariate Exploratory Linear Optimized Decomposition into Independent Components. 26
- MRI** Magnetic Resonance Imaging. xiii, 3, 4, 7, 11, 15–18, 21, 26, 31–34, 39, 40, 42–45, 60, 73, 81, 83, 88, 105–107, 112, 129–132, 134, 135, 137, 144
- NBS** Network Based Statistics. 32, 35
- NiDB** Neuroinformatics Database. 8
- NoSQL** Not Only SQL. 12, 63–67, 70, 77, 79, 106
- OCDE** Organização para a Cooperação e Desenvolvimento Económico. xiii, 4
- OLAP** Online analytical processing. 63
- OLTP** Online transaction processing. 63
- P2P** Peer-to-Peer. x, 56–59, 106
- PCE** Processo Clínico Electrónico. 2
- PET** Positron Emission Tomography. 3, 31, 129, 137
- ProbtrackX** Probabilistic Tracking. 29, 35, 84
- RA** Relative Anisotropy. 146
- RCE** Registo Clínico Electrónico. 1
- RDMS** Relational Database Management System. 63
- ROIs** Regions of Interest. 30, 31, 37, 148, 149
- rsfMRI** resting-state fMRI. 16, 26, 136, 137, 141, 142
- RSN** Resting State Networks. 138, 141
- SB** Substância branca. 25, 41
- SC** Substância cinzenta. 41
- SCA** Seed Correlation Analysis. 140

- sMRI** structural Magnetic Resonance Imaging. ix, 4–6, 8, 16–19, 22, 131, 132
- SPM** Statistical Parametric Mapping. 21, 22, 24, 33
- SQL** Structured Query Language. 61, 63
- SSL** Secure Socket Layer. xv, 82, 93, 94, 102, 109
- TBSS** Tract Based Special Statistics. 16, 29, 35, 39, 148, 149
- TC** Tomografia Computadorizada. 3
- TCP/IP** Transmission Control Protocol/Internet Protocol. 56
- TI** Tecnologias da Informação. 1
- TLS** Transport Layer Security. 93
- TOI** Tracts of Interest. 149
- TPM** Tissue Probability Map. 41
- TR** Tempo de Repetição. 22
- trFMRI** task-related fMRI. 136, 138
- TTL** Time-to-Live. 79
- VBCT** Voxel Based Cortical Thickness. 135
- VBM** Voxel based Morphometry. 4, 17–19, 33, 41, 133–135, 148
- XML** EXtensible Markup Language. 52, 68, 89
- XNAT** Extensible Neuroimaging Archive Toolkit. xiv, 50, 52, 53, 110

# Capítulo 1

## Introdução

### 1.1 Enquadramento

Numa sociedade intimamente relacionada com o desenvolvimento das **Tecnologias da Informação (TI)**, a área de Engenharia Biomédica assume crescente influência. A multidisciplinaridade desta vertente contempla conhecimento de engenharia, biologia e medicina com o intuito de melhorar a qualidade de vida e saúde humana, através da integração de prática clínica com as ciências de engenharia [1] .

As **TI** possuem um efeito muito difundido em praticamente todos os aspectos da atualidade, alterando a forma como vivemos e trabalhamos. Esse mesmo impacto reflete-se diretamente na pluralidade de serviços prestados por profissionais desta área. A capacidade de organização de informação de forma eficiente e compacta, bem como a sua distribuição segura e facilitada, conduz à optimização de processos com melhoria nos serviços prestados [2].

Sendo evidente que a utilização de sistemas informáticos potencia todo o quotidiano do ser humano nas suas mais diversas vertentes, também a sua preponderância se faz notar na saúde. Tal relação dá origem à disciplina de **Informática Médica (IM)**. A **IM** é uma área multidisciplinar, como evidenciado na Figura 1.1, que utiliza tecnologia de informação médica, para melhorar a prestação de cuidados de saúde, como verificável na facilidade de acesso a informação clínica ou migração dos registos em papel para **Registo**

Clínico Electrónico (RCE) [3]. É actualmente inimaginável um mundo onde cirurgias sejam realizadas sem software de apoio ao seu planeamento e realização; hospitais e centros de saúde organizados sem recorrer a ferramentas de Processo Clínico Electrónico (PCE); inexistência de ferramentas de apoio ao diagnóstico como aplicações de imagiologia.

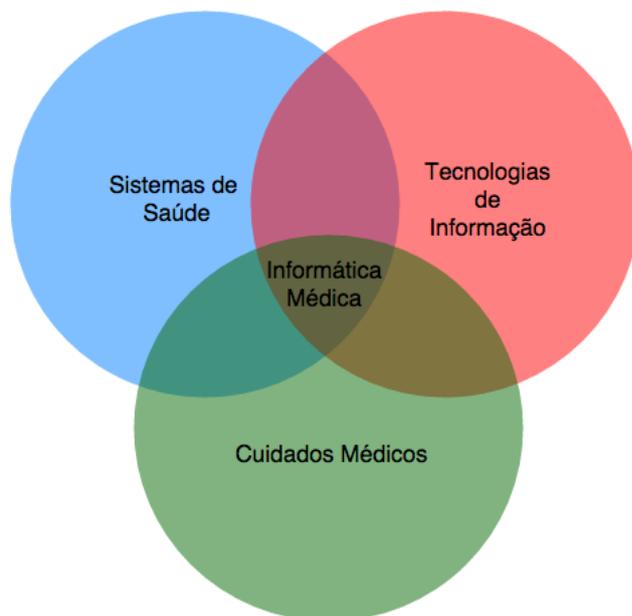


Figura 1.1: Domínios fundamentais de Informática Médica

No que respeita a concepção, desenvolvimento, implementação e manutenção de sistemas de informação médicos, existe um conjunto de funcionalidades e comportamentos específicos a adoptar, pelo que a qualidade dos cuidados de saúde a prestar se desenvolve, por conseguinte, numa relação direta com a qualidade dos sistemas de gestão de informação que aí venham a ser implementados [1,4].

Devido à multidisciplinaridade do conceito de **IM**, a sua definição não é única e consensual. (Greenes & Shortlife, 1990 [5]) definem **IM** como o campo responsável pelo processamento e comunicação de informação relacionada com a prática, pesquisa e educação médica, incluindo todas as tecnologias informáticas necessárias para o apoio destas tarefas. Noutra perspectiva, (Degoulet Fieschi, 1997 [6]) sugerem dividir a área de **IM** em três dimensões,

o auxílio na prestação de cuidados de saúde, a gestão de informação e conhecimento e a influência na sociedade, alterando a forma como a comunidade médica interage com o mundo.

Tal como referido, são várias as vertentes onde a IM possui um forte crescimento e entre elas, a área de Imagem Médica usufrui de grande destaque. A Imagem Médica ou Imagiologia inclui todos os processos e técnicas que permitem adquirir e processar imagens do corpo humano com o objectivo de possibilitar e/ou facilitar o diagnóstico e tratamento de pacientes [7].

A Imagem Médica tornou-se virtualmente essencial para todas as condições clínicas, sendo um modelo de diagnóstico médico para cancro, acidente vascular cerebral, traumatismos, problemas neurológicos, entre outros. É utilizada por um vasto leque de especialistas médicos, desde oncologistas a internistas. Entre as principais técnicas utilizadas destacam-se o raio-x, ultrassons, Tomografia Computadorizada (TC) e Ressonância Magnética (*Magnetic Resonance Imaging (MRI)*). [8].

Uma das tendências atuais da área de Imagem Médica é a disciplina de Neuroimagiologia, responsável pela investigação das estruturas e respectivas funções do cérebro. Para tal são empregues técnicas como MRI, Electroencefalografia (*Electroencephalography (EEG)*), TC ou Tomografia por Emissão de Positões (*Positron Emission Tomography (PET)*) [9].

A MRI é a modalidade predominante em neuroimagem uma vez que apresenta diversas propriedades vantajosas. Além de não envolver radiação ionizante, não existem efeitos secundários conhecidos associados a campos magnéticos e ondas rádio. É uma técnica relativamente barata, não invasiva, com crescente utilização (como evidenciado na Figura 1.2), providenciando excelente resolução espacial e permitindo maior contraste entre tecidos. Desta forma é possível obter imagens com grande qualidade e detalhe sobre diferentes aspectos do funcionamento e estrutura cerebral utilizando o mesmo equipamento, alterando apenas o tipo de aquisição e parametrização [10,11].

O cérebro humano é um órgão essencial à vida e funcionamento de todo o corpo, assumindo-se como um dos mais difíceis de compreender e estudar na sua função e estrutura. Este é um sistema complexo, caracterizado por comunicações dinâmicas em conjuntos funcionalmente especializados e interações

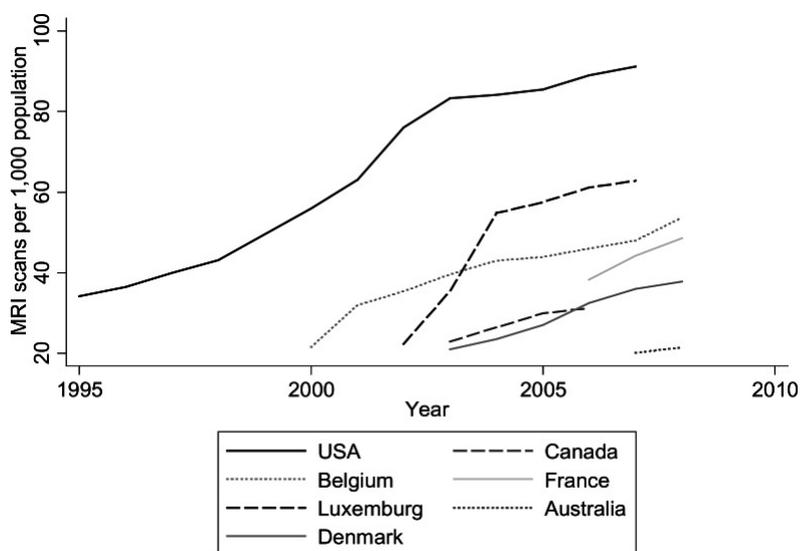


Figura 1.2: Padrão de crescimento na utilização de MRI em países da OCDE (retirado de [11])

mútuas de grande alcance entre estas mesmas estruturas. A dificuldade de acesso e complexa interação entre todos os seus elementos na realização de tarefas cognitivas, sensoriais e motoras transformam o cérebro humano numa das estruturas de maior dificuldade e interesse de estudo [12].

A abrangência de diferentes aquisições de MRI permite a realização de diferentes estudos com propósito de exploração de distintas características cerebrais. De entre estas técnicas destacam-se dois grandes grupos: modalidades de estudo estrutural e funcional.

A Ressonância Magnética Estrutural (*structural Magnetic Resonance Imaging (sMRI)*) é uma poderosa ferramenta utilizada em estudos clínicos e na linha de investigação em neurociências, fornecendo um método não invasivo para obtenção de imagens com alta resolução da estrutura anatômica cerebral bem como de patologias neurológicas [13].

Com estas imagens é possível realizar diferentes estudos da estrutura cerebral: análise visual, no qual são verificadas diferenças entre imagens e análises quantitativas onde se destacam a segmentação, *Voxel based Morphometry (VBM)* e estudos de espessura cortical [14,15]. A sMRI providencia informação qualitativa e quantitativa relativamente à forma, tamanho e integridade

da substância branca e cinzenta cerebral.

A capacidade de segmentação e classificação de diferentes regiões cerebrais a partir de imagens de *sMRI* é extremamente importante em estudos de desenvolvimento cerebral, degeneração neuronal e avaliação de doenças neurológicas e psicológicas [13, 16]. Estudos baseados em análises volumétricas são essencialmente relevantes na exploração de doenças do foro neurológico como a esquizofrenia [17], alzheimer [18] ou autismo [19].

Uma técnica que tem recebido grande destaque em estudos estruturais é a técnica de *Diffusion Tensor Imaging (DTI)*, realizada a partir de aquisições de difusão. Esta utiliza informação sobre a difusão das partículas de água, para determinar a presença e orientação das fibras cerebrais. A partir destes dados é possível construir tratografias, onde se procura reconstruir os diferentes feixes neuronais do cérebro. Por sua vez estes feixes podem ser observados por meio de mapas determinísticos e probabilísticos [20].

São várias as aplicações de *DTI* na área de investigação neurológica e em ambiente clínico como ferramenta de diagnóstico tais como: epilepsia [21], esclerose múltipla [22], esquizofrenia [23] ou tumores cerebrais [24].

Por outro lado, a técnica de Ressonância Magnética Funcional (*functional Magnetic Resonance Imaging (fMRI)*) faz uso da medição da resposta hemodinâmica cerebral para obter um mapeamento da atividade/função cerebral [25, 26]. Esta medida baseia-se no efeito *Blood Oxygen Level Dependent (BOLD)*, assim designado por permitir verificar o nível de oxigenação sanguíneo. Assim, regiões cerebrais com maior actividade neuronal poderão ser identificadas devido ao aumento de irrigação com sangue oxigenado [27].

O estabelecimento de contrastes perante diferentes níveis de oxigenação permite detectar zonas de activação cerebral em diferentes contextos, quer pela realização de tarefas (utilização de paradigmas), quer em repouso.

Este tipo de mapeamento, utilizado essencialmente em investigação na área das neurociências, pode também ser explorado para planeamento cirúrgico e radioterapêutico de tumores [28–30]. Outras aplicações de *fMRI* incluem a avaliação das funções motoras ou da fluência verbal em pacientes com vários níveis patológicos, ou ainda o estudo de como diferentes regiões cerebrais sensíveis ao movimento, se articulam na percepção dos movimentos

globais [31].

O estudo das interdependências estatísticas entre as variações dos níveis de oxigenação sanguíneo de diferentes regiões permite estabelecer medidas de conectividade funcional. [32].

Tal como existem várias formas de se observar o cérebro humano, também surgiram várias abordagens para descrever e quantificar a conectividade cerebral. De uma forma geral, entende-se por conectividade cerebral o conjunto de padrões de ligação entre as diferentes unidades neuronais que compõem o cérebro humano. No entanto, é comum fazer-se uma distinção clara entre conectividade funcional e conectividade estrutural. [33].

A conectividade funcional designa as interações dinâmicas, verificadas por desvios do princípio da independência estatística, entre unidades neuronais espacialmente separadas [34].

Por outro lado, a conectividade estrutural refere-se às ligações físicas (anatômicas), estabelecidas por ligações sinápticas, entre elementos neuronais [35]. Estas ligações podem representar desde circuitos locais de células individuais a redes de grande escala. Estes padrões físicos são relativamente estáticos, podendo ser dinâmicos a escalas temporais na ordem das horas ou dias.

Os dois tipos de conectividade previamente descritos encontram-se relacionados e a forma como se moldam e limitam mutuamente tem-se tornado a questão chave de vários estudos, na forma como as diferentes regiões cerebrais se correlacionam e interligam [36]. Neste sentido é estudada a integração de diferentes estruturas cerebrais nas mais diversas funções e a avaliação da influência destas ligações em casos como o envelhecimento saudável do cérebro ou em doenças como a esquizofrenia ou autismo [37].

Embora modalidades singulares de análise de **fMRI** ou **DTI** tenham permitido inúmeras descobertas na compreensão do cérebro e doenças como o Alzheimer, esquizofrenia ou bipolaridade, a nova tendência de estudo recai na combinação destas técnicas de análise. Esta aglomeração de diferentes modalidades procura ultrapassar limitações anteriormente detectadas fazendo prever uma nova vaga de conhecimento sobre o complexo sistema cerebral [38].

Não obstante da relevância dos estudos de **sMRI**, **fMRI**, bem como na conjugação de modalidades, a sua utilização por profissionais é condicionada

por um conjunto de factores relacionados com o armazenamento, partilha e aquisição da informação gerada nas várias fases de estudo.

## 1.2 Problema

A partilha de informação é uma prática fundamental na sociedade face à vasta quantidade de dados continuamente adquiridos. A aquisição, processamento e análise de imagens de [MRI](#) envolve um conjunto vastíssimo de processos, ferramentas e informação gerada. Este tipo de análise é extremamente importante no âmbito das neurociências, sendo que a validação de resultados num estudo clínico é sempre o passo que antecede a transposição para a prática clínica.

Um estudo neuroimagiológico é geralmente um estudo de grupo, em que é necessário incluir vários sujeitos, tipicamente designados de casos. No exemplo mais típico de avaliação de uma doença neurológica existem dois grupos: um grupo de controlo e um grupo com a condição neurológica em estudo, cada um constituído por vários casos. Além disso num estudo multimodal, para cada caso são necessárias aquisições de diferentes modalidades. Cada aquisição estrutural é constituída por imagens  $3D$  e em aquisições funcionais e difusão até  $4D$ . Do processamento destas imagens resulta não só um grande volume de dados mas também um elevado número de ficheiros. A dimensão destes estudos é apresentado em maior detalhe na [Seção 2.7](#)

Além disso, para que dados à saída do equipamento de [MRI](#) possam conduzir a resultados interpretáveis, é necessário que sejam executadas três etapas distintas. Inicialmente, as imagens são submetidas a um conjunto de etapas de pré-processamento com o objectivo de reduzir o ruído e assim melhorar os resultados a obter. De seguida, é aplicado um modelo com o intuito de extrair dos dados informação relevante. Por fim, é realizado algum tipo de inferência, análise estatística ou ainda pós-processamento para melhorar a apresentação de resultados.

Para a execução destas etapas é necessária a utilização de *software* distinto, muitas vezes em processos computacionalmente pesados e demorados. Destes são obtidos vários ficheiros com potencial interesse de armazenamento,

partilha e reutilização.

Actualmente são várias as soluções que permitem o armazenamento e partilha de dados neuroimagiológicos como e.g. a plataforma *Anima* para partilha de estudos meta-analíticos de **fMRI** [39], a base de dados **Neuroinformatics Database (NiDB)** onde são armazenados grandes volumes de dados neuroimagiológicos provenientes de diversos projectos e modalidades de aquisição [40] ou o projeto **ConnectomeDB** para alojamento e disseminação de dados sobre a estrutura, função e conectividade cerebral [41].

Além da quantidade de informação gerada é comum as imagens serem processadas de forma sequencial, isto é, são estabelecidos fluxos de processamento [42].

Existem neste sentido alguns fluxos de processamento concretamente estabelecidos para o processamento e análise de imagens de **fMRI** e **sMRI**, como o utilizado na aplicação *BrainCat* para análise conjunta de imagens de **fMRI** e **DTI** [43] e evidenciado na Figura 1.3 .

No entanto a capacidade de interligação entre soluções que apresentam fluxos de processamento definidos, e a organização de dados e ficheiros resultantes do mesmo é ainda insuficiente.

A falta de padronização de procedimentos aliada à inexistência de uma plataforma de organização estruturada dos ficheiros gerados, torna o reaproveitamento, partilha e colaboração entre estudos uma tarefa árdua e contraprodutiva. Esta situação, aumenta consideravelmente a probabilidade de perda de informação, além de não permitir um meio seguro e automático para armazenamento e partilha de dados. A possibilidade de se cometerem erros inadvertidos, na combinação entre estudos e técnicas de análise é também maior, podendo comprometer os resultados obtidos. Neste sentido é necessário um esforço suplementar por forma a manter a coerência da informação gerada.

A introdução de uma alternativa que permita a conjugação entre ciclos de processamento e a estruturação de ficheiros num repositório poderá permitir que os problemas apresentados sejam em parte ultrapassados. Além de permitir a correcta organização e partilha de ficheiros, tal solução, pode ser um recurso viável na constituição de um serviço orientado ao cliente. A reutili-

zação de informação é simplificada, permitindo uma colaboração mais eficaz entre instituições e profissionais, reduzindo a necessidade de processamento repetido, muitas vezes computacionalmente dispendioso.

No mesmo sentido a tarefa de pesquisa e obtenção de dados específicos deverá ser transparente e simplificada, através de uma solução escalável, segura e intuitiva. A conjugação destes factores assume particular importância face à crescente produção de informação em neuroimagem, a qual deverá ser facilmente transmitida e utilizada, fomentando a capacidade de investigação e diagnóstico clínico cerebral.

É com base neste panorama que se justifica o desenvolvimento de uma solução inovadora que permita ultrapassar estes problemas, nomeadamente, a falta de armazenamento de informação segundo uma estrutura padronizada e a partilha de dados de forma segura e intuitiva entre entidades.

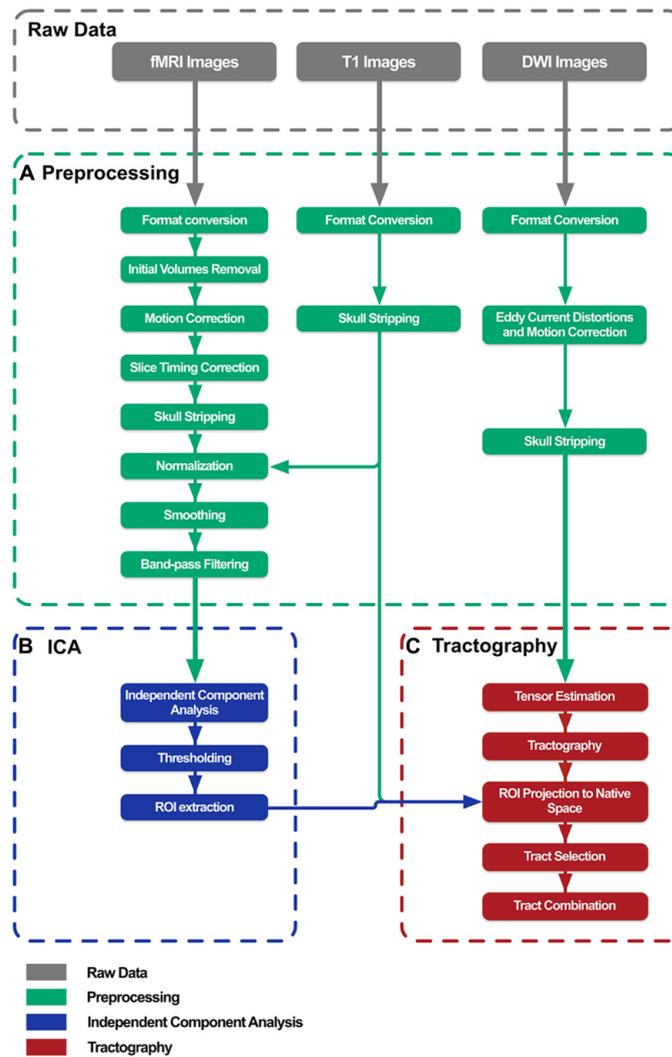


Figura 1.3: Visão geral do fluxo de processamento de dados *BrainCat* (retirado de [43])

## 1.3 Objectivos

Tendo em conta os problemas acima referidos, pretende-se com este projecto a concepção e desenvolvimento de um protótipo de solução para estruturação, arquivo e partilha de estudos neuroimagiológicos de *MRI*, baseado numa arquitectura orientada ao serviço.

- Desenvolver uma aplicação intuitiva e amigável para armazenamento, partilha e disponibilização de estudos neuroimagiológicos de *MRI*;
- Adoptar o armazenamento e estrutura de dados a um fluxo de processamento implementado, permitindo a reconstrução do mesmo de forma distribuída;
- Garantir que todo o tipo de informação esteja sempre disponível e acessível, independentemente do local/entidade em que o profissional se encontra, para que através de ferramentas próprias seja possível manipular essa informação;
- Providenciar a disponibilidade ubíqua de grandes volumes de dados, entre investigadores, entidades e profissionais de saúde;

## 1.4 Metodologia de Investigação

O trabalho da presente dissertação foi realizado durante o quinto ano do Mestrado Integrado em Engenharia Biomédica do ramo de Informática Médica. A realização dos objectivos apresentados seguiu a metodologia Ação-Pesquisa. Segundo esta metodologia o investigador testa uma teoria em situações reais, recebe *feedback*, modifica a teoria em função desse *feedback* e tenta novamente [44].

Esta metodologia inclui um conjunto de passos que devem ser seguidos [45, 46]:

- Identificação do problema e especificação das suas características;
- Actualização interativa do estado da arte dos objectivos do trabalho;

- Modelação e desenvolvimento de um protótipo que satisfaça as especificações;
- Análise de resultados e correção com base nos mesmos;
- Validação do protótipo.

## 1.5 Estrutura do documento

O presente trabalho compreenderá, para além deste capítulo introdutório, mais quatro capítulos distintos estruturados da seguinte forma:

- **Capítulo 2** - Processamento de dados de Ressonância Magnética

Neste capítulo é abordada a temática que envolve a técnica imagiológica de Ressonância Magnética, com destaque para as modalidades de aquisição estrutural, funcional e de difusão. São apresentados princípios básicos de cada modalidade, as análises que podem ser efectuadas, bem como software adequado para a implementação de cada etapa.

É também apresentado um fluxo de processamento para estudos multimodais, com descrição das etapas que o compõem. Para este fluxo são pormenorizados os ficheiros considerados essenciais, resultantes da realização do mesmo. Neste sentido, é por último realizada uma exploração do volume de dados gerados para um estudo típico multimodal e apresentadas soluções de mercado para a manutenção e partilha de informação neurológica.

- **Capítulo 3** - Tecnologia

No segundo capítulo são revistas características técnicas exigidas para a realização deste trabalho. Neste sentido são referidos princípios básicos para arquiteturas em rede, bem como vantagens e desvantagens da utilização das soluções apresentadas. É também realizada uma análise detalhada de diversas soluções de bases de dados, com especial destaque para a comparação da vertente de bases de dados relacionais e [Not Only SQL \(NoSQL\)](#).

- **Capítulo 4** - BrainArchive

No quarto capítulo é apresentada e descrita a aplicação desenvolvida (BrainArchive) que permite a integração, armazenamento e partilha automática de ficheiros, associados ao fluxo de processamento apresentado no capítulo 2. Neste é realizada uma análise de requisitos do sistema a implementar, apresentada a sua arquitetura, implementação, principais componentes e funcionalidades.

- **Capítulo 5** - Resultados e Conclusões

No último capítulo, realiza-se uma exposição sumária do trabalho realizado e principais contribuições, juntamente com sugestões para trabalho futuro.



# Capítulo 2

## Processamento de dados de Ressonância Magnética

### 2.1 Introdução

Nas últimas décadas, a técnica de **MRI** tem-se tornado uma das mais relevantes modalidades de neuroimagem, contribuindo para tal vantagem inerentes à própria técnica. Além de ser uma técnica não invasiva, não apresenta efeitos adversos conhecidos comparativamente a **Computed Tomography (CT)**, que utiliza radiação ionizante [47]. As principais vantagens de **MRI** prendem-se com o facto de proporcionar um bom contraste em tecidos moles e apresentar excelente resolução espacial. Desta forma é capaz de caracterizar e discriminar tecidos através das suas propriedades físicas e bioquímicas [15].

A abrangência de diferentes modalidades de **MRI** permite a realização de diferentes estudos com propósito de exploração de distintas características cerebrais. Estas modalidades dividem-se em dois grandes grupos: modalidades de estudo estrutural e funcional.

A utilização de fluxos de processamento é cada vez mais utilizada no processamento de dados neuroimagingológicos. No entanto, a falta de informação e padronização de processos é um obstáculo à partilha estruturada de estudos nas mais diversas etapas de análise.

Para que dados à saída do equipamento de **MRI** possam conduzir a resultados interpretáveis, é necessário que sejam executadas três etapas distintas. Inicialmente, as imagens são submetidas a um conjunto de etapas de pré-processamento com o objectivo de reduzir o ruído e assim melhorar os resultados a obter. De seguida, é aplicado um modelo com o intuito de extrair dos dados informação relevante. Por fim, é realizado algum tipo de inferência, análise estatística ou ainda pós-processamento para melhorar a apresentação de resultados.

A análise constitui a etapa central de qualquer fluxo de processamento neuroimagiológico. Com efeito, as etapas de pré/pós-processamento podem ser em parte, influenciadas pela análise que se pretende realizar. É assim necessário definir qual a metodologia de análise empregue aos dados de cada modalidade.

No que diz respeito a análises de **fMRI**, destacam-se dois tipos de análise: as análises de correlação e as **Independent Component Analysis (ICA)**. Apesar de permitirem estabelecer uma métrica direta de correlação funcional entre uma região de interesse e os restantes vóxeis do cérebro, os resultados das análises de correlação são altamente dependentes da escolha da região de interesse que constitui a *seed* [48]. Por outro lado, as análises **ICA** são a metodologia mais popular na identificação de padrões de conectividade funcional em **resting-state fMRI (rsfMRI)**, pelo facto de constituírem uma análise puramente exploratória dos dados, não dependendo diretamente da escolha prévia de uma região de interesse.

O crescente interesse no estudo dos feixes neuronais é fruto dos recentes avanços nas técnicas de **DTI**. Neste âmbito, a tratografia tem sido a metodologia empregue em maior escala. Além de permitir uma medida intuitiva do grau de conectividade estrutural pela visualização gráfica das fibras geradas, pode ser facilmente combinada com os resultados de análise **fMRI** [49]. Por outro lado a técnica de **Tract Based Special Statistics (TBSS)** permite melhorar a sensibilidade e capacidade de interpretação de análise em estudos de difusão com múltiplos sujeitos.

Por fim, a utilização de **sMRI** fornece um método para avaliação da estrutura anatómica cerebral. Uma modalidade de grande interesse neuroimagio-

lógico é o método **VBM**. Com recurso a esta técnica é possível a identificação de diferentes locais na densidade de substância cinzenta e branca, através de testes estatísticos que comparam todos os vóxeis de diferentes grupos. Além disso, a capacidade de segmentação e classificação de diferentes regiões utilizando **sMRI** é fundamental na utilização de análise de grafos em imagens de **fMRI** e **DTI**.

Maior detalhe sobre **MRI**, correspondentes modalidades e técnicas de análise é apresentado no **Apêndice A**.

## 2.2 Ressonância Magnética Estrutural

A **sMRI** é uma poderosa ferramenta utilizada na pesquisa cerebral e neurologia clínica. Esta fornece um método não invasivo para obtenção de imagens com alta resolução da estrutura anatômica do cérebro bem como de patologias neurológicas. Muitas sequências de **sMRI** são volumétricas, o que significa que as medições podem ser realizadas em estruturas cerebrais específicas para cálculo de volumes de tecido. A capacidade de segmentar e classificar diferentes regiões a partir de **sMRI** tem grande importância em estudos de desenvolvimento cerebral, degeneração neuronal e na avaliação de doenças neurológicas e psicológicas [13].

Independentemente do método utilizado, a segmentação tridimensional de **MRI** cerebral apresenta-se como um grande desafio. A quantidade de estruturas com diferente forma e aparência, em conjunto com a necessidade de alta precisão e qualidade dos resultados justificam a diversidade de métodos aplicados. A existência de factores como ruído próprio da aquisição, artefactos e a presença de outros tecidos como o crânio dificultam a segmentação de imagens de **sMRI**. Assim sendo, antes de aplicar qualquer método de segmentação é necessário aplicar alguns passos de pré-processamento. Alguns passos de pré-processamento em **sMRI** são semelhantes aos utilizados em **fMRI** pelo que serão explicitados na **Subsecção 2.3.1** deste capítulo.

### 2.2.1 Segmentação de imagens estruturais

A segmentação de imagens cerebrais é um passo essencial em estudos neuroimagiológicos. A segmentação de uma imagem é aplicada para delinear as diferentes regiões anatómicas, classificadas segundo determinado critério. Esta pode ser realizada através da identificação de todos os vóxeis que pertencem a uma região ou estrutura de interesse. Para tal a utilização de atlas cerebrais (apresentados em maior detalhe na [Seção 2.5.1](#)) é prática comum em neuroimagiologia.

A alternativa à utilização de um atlas pré-segmentado para todos os casos passa por realizar a segmentação para cada um dos casos individualmente. Quando a identificação das estruturas cerebrais numa imagem é feita simplesmente pela sobreposição com o atlas, não existe garantia (mesmo estando as duas imagens no mesmo espaço anatómico), da fiabilidade dos resultados. Neste sentido, uma das mais fortes tendências na classificação de imagens é a utilização de ferramentas de classificação anatómica automática, tal como o Freesurfer [50].

O fluxo de processamento do Freesurfer para a segmentação e classificação de imagens de [MRI](#) compreende três grandes passos:

- inicialmente são realizadas tarefas elementares de pré-processamento: correção de movimentos, normalização de intensidade, normalização espacial e remoção do crânio;
- de seguida, são realizados mais passos de pré-processamento e iniciadas operações de segmentação;
- o passo final é responsável pelo registo, parcelação, reconstrução de superfícies e classificação segundo as *labels* existentes.

### 2.2.2 Processamento de [sMRI](#)

Entre as técnicas de processamento de [sMRI](#) destacam-se a Volumetria, [VBM](#) e análises de espessura cortical. Para maior detalhe acerca da implementação destas modalidades e correspondente utilidade, consultar [Seção A.2](#) dos Apêndices.

### Voxel-Based Morphometry

Na utilização de imagens estruturais, é possível investigar diferenças no volume/topografia de substância cinzenta. A abordagem **VBM** é imparcial, na medida em que não requer informação a priori acerca da possível localização de substância cinzenta, não sendo dependente do operador.

Por forma a comparar todas as imagens de substância cinzenta numa base *voxelwise*, é necessária a sua transformação para o espaço padrão MNI, o que por sua vez envolve registo não-linear. Tal é possível e.g. recorrendo ao pacote **FMRIB'S Software Library (FSL)**, mais especificamente à ferramenta **FMRIB's Nonlinear Registration Tool (FNIRT)** [51].

## 2.3 Ressonância magnética funcional

A modalidade de **fMRI** é actualmente uma das técnicas mais utilizadas para o mapeamento do funcionamento do cérebro humano na área das neurociências. Os seus métodos podem medir a perfusão, volume ou oxigenação (neste, a técnica **BOLD** é a mais comum) do sangue no cérebro. São métodos conhecidos por possuírem excelente resolução espacial mas fraca resolução temporal. A **fMRI** tem sido utilizada para investigar e testar hipóteses da topografia funcional do cérebro e investigar mecanismos neurofisiológicos [52, 53], através da medição da resposta hemodinâmica cerebral para obter um mapeamento da atividade/função cerebral.

### 2.3.1 Pré-processamento de fMRI e sMRI

Como referido anteriormente alguns passos de pré-processamento são comuns a análises de **fMRI** e **sMRI**. Nesse sentido é apresentado na Figura 2.6 o fluxo de pré-processamento para imagens funcionais e estruturais. Nesta seção será ainda descrito cada um destes passos bem como as ferramentas utilizadas.

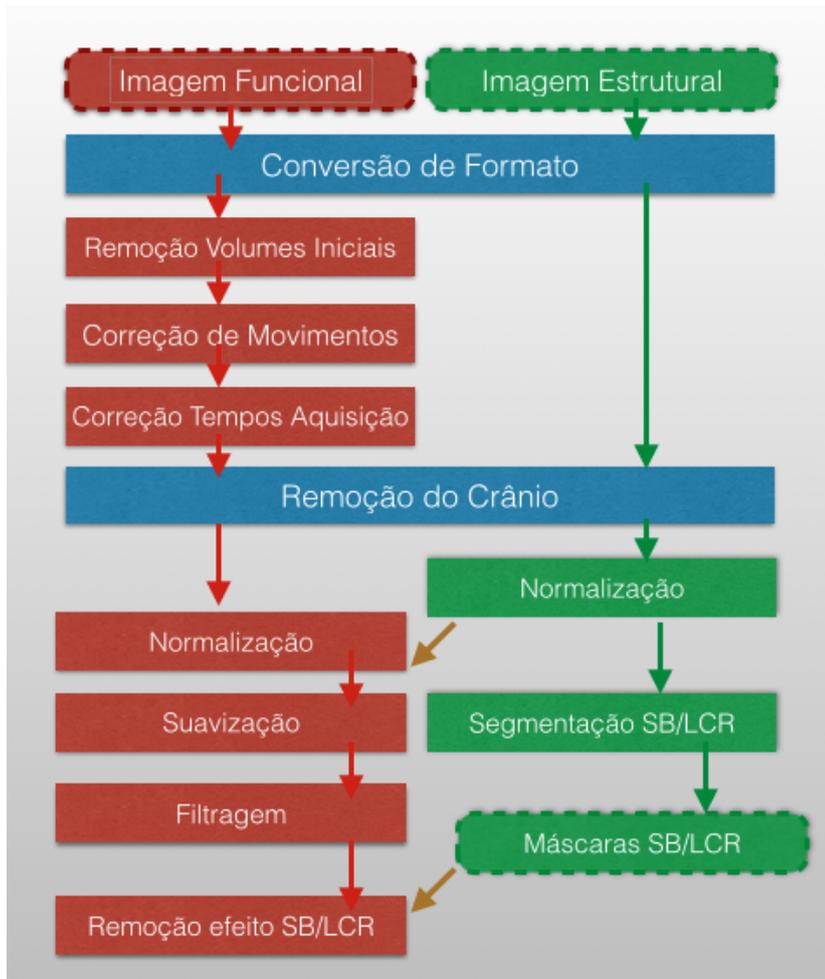


Figura 2.1: Fluxo de pré-processamento de imagens funcionais e estruturais. As imagens estão representadas a tracejado e os passos de processamento em caixas contínuas. O fluxo das imagens funcionais está representado a vermelho e as estruturais a verde. Passos comuns são representados a azul e quando o resultado de um fluxo é utilizado no outro encontra-se em castanho.

### Conversão de Formato

As imagens são normalmente exportadas do equipamento de aquisição em formato [Digital Imaging and Communications in Medicine \(DICOM\)](#), o formato padrão utilizado em ambiente hospitalar e departamentos imagiológicos, enquanto os mais importantes pacotes de software como [FSL](#) e [Statistical Parametric Mapping \(SPM\)](#) [54] usam o formato NIfTI. Nesse sentido a primeira etapa de processamento consiste em converter imagens do formato DICOM para NIfTI. Tratando-se de uma simples conversão de formato, sem grandes possibilidades de erro, é conveniente utilizar um programa igualmente simples. Este é o caso do software *dcm2nii*, disponibilizado em conjunto com o software de visualização MRICron [55].

Estas definições incluem a conversão para o formato NIfTI 4D quando aplicável, em que a quarta dimensão corresponde aos diferentes volumes da aquisição. Além disso, por defeito as imagens são compactadas o que diminui o espaço ocupado pelos ficheiros resultantes.

### Extração de Volumes iniciais

Quando uma aquisição de [MRI](#) é iniciada, o campo magnético demora algum tempo a estabilizar, sendo conveniente remover os volumes iniciais de cada aquisição que poderão possuir informação distorcida.

.Presente na interface de [SPM](#) e [Data Processing Assistant for Resting-State fMRI \(DPARFS\)](#) [56], este passo pode ser realizado pelo comando de [FSL](#) *fslroi*, estando incluído como uma opção por defeito no script *afni\_proc*.

### Correção de Movimentos

A existência de movimentos da cabeça durante uma aquisição funcional pode ter sérios efeitos no sinal [fMRI](#), podendo invalidar possíveis conclusões sobre diferenças na fisiologia cerebral .

Uma vez que todos os volumes da mesma aquisição pertencem ao mesmo sujeito, o tamanho e proporções do cérebro mantêm-se inalterados ao longo desses volumes. Assim, um alinhamento rígido (rotações e translações das

imagens) é suficiente. O volume de referência poderá ser um qualquer volume da mesma aquisição, contudo, a utilização da média da aquisição como referência minimiza a operação de procura do melhor alinhamento.

Os principais pacotes de software apresentam de alguma forma capacidade para corrigir este problema. No **FSL** a correção de movimentos é realizada utilizando a função *McFlirt*.

### Correção dos tempos de aquisição

Cada corte de um volume de **fMRI** é obtido em diferentes instantes temporais. No entanto, como nas análises de **fMRI** se assume que cada volume cerebral representa a resposta **BOLD** de um só instante temporal, é essencial corrigir este desfasamento.

A correção dos tempos de aquisição é levada a cabo através da deslocação no tempo dos sinais de cada corte por uma fracção do **Tempo de Repetição (TR)** equivalente à diferença entre os tempos de aquisição do corte em questão e um corte de referência. Devido à deslocação do sinal, em cada instante de tempo, os valores dos vóxeis do corte em estudo são alterados e devem ser estimados por interpolação.

A ferramenta utilizada nesta operação no pacote **FSL** designa-se *slicetimer*, à qual é necessário fornecer o **TR** da aquisição, a ordem pela qual os cortes são obtidos e realiza a interpolação utilizando uma função seno. No **Analysis of Functional Neuroimages (AFNI)** [57] a função *3dTshift* funciona de forma semelhante, permitindo no entanto escolher entre diferentes tipos de interpolação. É também possível utilizar **SPM** nesta operação.

### Extração do crânio

Devido à própria natureza da aquisição, as imagens de **fMRI** e **sMRI** contêm informação de vários tipos de tecido para além dos tecidos cerebrais. É assim desejável isolar o tecido cerebral dos restantes, pois a exibição de outras estruturas, aumenta desnecessariamente o espaço ocupado pelos ficheiros de imagem, assim como a probabilidade de erros no processo de normalização.

Sendo um dos passos de pré-processamento mais comuns, os principais

pacotes de software apresentam ferramentas para realizar esta operação. No FSL o comando *Brain Extraction Tool (BET)* é responsável por este passo.

A extração do crânio é realizada em dois passos distintos. No primeiro passo, o comando *BET* é utilizado para calcular uma máscara que inclua apenas as estruturas cerebrais a partir do volume da média da aquisição. Neste, as definições padrão deverão conduzir a bons resultados. Contudo, caso os resultados da extração não sejam satisfatórios, o programa deverá ser executado novamente, mas desta vez definindo o limite de intensidade fraccionário a ser utilizado na extração. O valor deste limite que conduz aos melhores resultados só poderá ser alcançado por tentativa erro e poderá diferir de sujeito para sujeito.

Uma vez calculada a máscara, esta deverá ser aplicada a todos os volumes da aquisição através do programa *fslmaths* e da opção correspondente que permite a aplicação de máscaras às imagens fornecidas como entrada.

### Normalização

A normalização é um processo necessário quando são realizados estudos que incluem vários sujeitos e no qual se pretendem estabelecer comparações. Após a normalização, vóxeis posicionados nas mesmas coordenadas em imagens de indivíduos diferentes, corresponderão tanto quanto possível à mesma estrutura anatómica, eliminando factores de variabilidade entre sujeitos.

A normalização é geralmente composta por duas etapas. Inicialmente é calculada a matriz de normalização das imagens estruturais para o espaço anatómico de referência e noutra é calculada a matriz de co-registo dos volumes funcionais com a imagem estrutural. A normalização mais comum regista as imagens para um espaço anatómico standard como o MNI ou Talairach. O co-registo consiste em alinhar as imagens funcionais de um indivíduo com as respectivas imagens estruturais. Os alinhamentos podem ser realizados utilizando transformações rígidas (apenas movimentos de rotação e translação) e transformações não rígidas (onde são alteradas as dimensões da imagem), variando entre o benefício de possuir diferentes imagens num espaço comum, e possíveis distorções na informação da imagem. A combinação das matrizes

obtidas permite normalizar as imagens funcionais para o espaço anatómico de referência.

A função *FMRIB's Linear Image Registration Tool (FLIRT)* do *FSL* pode ser utilizada para calcular e aplicar a matriz de alinhamento que possui todas as operações realizadas para alinhar duas imagens perante uma função de custo.

### Suavização

O processo de suavização funciona calculando para cada vóxel uma média pesada dos seus vóxeis vizinhos, sendo normalmente utilizada a convolução das imagens com uma função Gaussiana tridimensional. Este processo suaviza o sinal de toda a imagem diminuindo diferenças acentuadas entre estruturas anatómicas próximas, diminuindo variações entre sujeitos e melhorando a relação sinal-ruído.

A extensão da função ou curva Gaussiana controla o grau de suavização aplicado, sendo esta geralmente expressa em termos de *Full Width at Half Maximum (FWHM)* que representa a extensão da função quando a sua curva atinge metade do valor máximo. Com o *FSL* é possível realizar esta operação utilizando a função *fslmath*. Tanto o *SPM* como o *AFNI* na função *3dFWHMx* implementam o mesmo tipo de suavização.

### Filtragem

As redes neuronais de repouso apresentam frequências na gama dos 0,01 Hz a 0,08 Hz. Para preservar o sinal destas e remover sinal associado a ruído e sinais fisiológicos é comum aplicar-se um filtro passa-banda que limite o sinal a estas frequências.

Para tal, o programa *fslmaths* é utilizado. Especificando a devida opção, este programa permite combinar a aplicação de um filtro passa-alto não linear e um filtro passa-baixo linear Gaussiano. No *AFNI*, a função *3dBandpass* aplica um filtro baseado na *Fast Fourier Transform (FFT)*.

### Remoção do sinal SB e LCR

A remoção do sinal de Substância branca (SB) e do Líquido cefalorraquidiano (LCR) permite diminuir a influência de variações no sinal BOLD que podem interferir com os resultados finais. Para remover o sinal de SB e LCR são calculadas as respectivas máscaras, segmentando os volumes estruturais. Usando essas máscaras, as séries temporais médias para cada tecido podem ser calculadas. O sinal obtido pode então ser removido utilizando regressão linear.

No FSL ainda não existe uma ferramenta de automização deste processo, no entanto é possível realizar recorrendo a este pacote, seguindo os seguintes passos:

- Utilizar a função [FMRIB's Automated Segmentation Tool \(FAST\)](#) para segmentar as imagens estruturais obtendo máscaras de SB e LCR;
- De seguida, usando o `fslmaths` é possível binarizar as imagens de forma que os valores entre 0.95 e 0.99 sejam 1 e tudo o resto 0, garantindo que as máscaras não estão contaminadas por sinal de substância cinzenta;
- Recorrendo à função [FLIRT](#) é possível converter as máscaras para o espaço funcional. A sua combinação com a imagem funcional permite extrair as séries temporais através da função `meants`;
- Por último, o sinal obtido pode ser utilizado com a função [FMRI Expert Analysis Tool \(FEAT\)](#), para remover a influência da SB e LCR do sinal BOLD.

### 2.3.2 Processamento de fMRI

Actualmente podem ser encontrados na literatura vários métodos para analisar dados de fMRI [58]. Nos estudos típicos de fMRI destaca-se um método conhecido por [General Linear Model \(GLM\)](#). Já para o estudo da conectividade funcional, podem utilizar-se abordagens por modelos, como as análises de correlação, e abordagens independentes de modelos, como o

caso das análises **ICA**. Uma descrição mais detalhada destas modalidades é apresentado na **Seção A.3** dos Apêndices.

### **Independent Component Analysis**

O programa **Multivariate Exploratory Linear Optimized Decomposition into Independent Components (MELODIC)** [59] permite ao utilizador realizar análises de primeiro nível (estudos individuais) e análises **ICA** de segundo nível (estudos de grupo). Nas análises de segundo nível é possível optar pela identificação das componentes que demonstram maior dependência temporal ou espacial entre sujeitos. Em estudos de **rsfMRI**, procuram-se encontrar os padrões espaciais mais comuns ao longo de todo o grupo, não existindo garantias de que a evolução temporal dos sinais seja consistente entre os diferentes elementos do grupo. Assim sendo, e para este caso, justifica-se a aplicação de análise de componentes independentes temporal.

## **2.4 Ressonância Magnética de Difusão**

A ressonância magnética de difusão teve o seu início nos anos 80 com o intuito de permitir estudos funcionais do cérebro, assentando no pressuposto de que poderia ser utilizada para avaliar o fluxo sanguíneo cerebral. É no entanto um padrão de avaliação de alterações na substância branca cerebral, acompanhamento de vítimas de acidente vascular cerebral e criação de mapas de ligação cerebrais [60].

A imagem por Tensor de Difusão (**DTI**) baseia-se no princípio de difusão das moléculas de água, um processo físico que ocorre naturalmente em condições de equilíbrio térmico, caracterizado por movimentos moleculares aleatórios. É o movimento preferencial das moléculas de água ao longo das estruturas fibrosas que torna a **MRI** de difusão uma modalidade imagiológica particularmente útil no estudo destas estruturas.

O estudo da difusão de moléculas de água *in vivo* tornou o **DTI** uma ferramenta singular no estudo de várias patologias incluindo esclerose múltipla, dislexia, derrames cerebrais e esquizofrenia [61].

### 2.4.1 Pré-processamento de DTI

#### Conversão de formato

À semelhança dos ficheiros de imagem de aquisição funcional e estrutural, também os ficheiros da aquisição DTI devem ser convertidos para o formato NIfTI. Contudo, neste caso além do ficheiro NIfTI 4D, é também gerado um ficheiro com os vectores dos gradientes de difusão utilizados (ficheiro com extensão ".bvec") e um outro ficheiro com os valores de b aplicados (ficheiro com extensão ".bvec").

#### Correção de movimentos e distorções

Tal como no caso das aquisições funcionais, as aquisições de difusão devem também ser corrigidas de forma a compensar os movimentos da cabeça. No entanto, a esta correção acresce a necessidade de corrigir também as distorções provocadas por correntes residuais.

A correção deste tipo de distorção pode ser efectuada pelo alinhamento rígido dos volumes de difusão com um volume de referência. Tendo em conta que o volume b0 não resulta da aplicação de um gradiente de difusão, não possui este tipo de distorção e, por isso, é geralmente utilizado como volume de referência. Este alinhamento permite simultaneamente corrigir também os movimentos da cabeça já que os procedimentos são equivalentes, reduzindo as duas correções numa única etapa.

O programa *eddy\_correct* da biblioteca FSL foi especialmente desenvolvido para este efeito.

#### Remoção do crânio

A remoção das estruturas não cerebrais em aquisições de difusão apresenta o mesmo grau de relevância, pelas mesmas razões. Nesse sentido, o processo de extração do crânio é implementado de forma semelhante mas com as seguintes alterações:

- o volume utilizado na criação da máscara das estruturas cerebrais é b0 e não a média dos volumes;

- o limite de intensidade fracionário utilizado pelo programa [BET](#) não deverá ser o valor definido por defeito, uma vez que demonstra remover uma grande extensão do tecido cerebral do volume  $b_0$

### 2.4.2 Processamento de DTI

A visualização da informação contida no tensor de uma forma intuitiva e perceptível é um dos maiores desafios que se colocam à utilização desta modalidade. A elevada dimensionalidade dos dados e as relações complexas no domínio do tensor de difusão são as duas principais razões para essa dificuldade. Consequentemente, várias foram as alternativas propostas ao longo dos últimos anos para facilitar essa visualização, passando muitas dessas alternativas por uma redução da dimensionalidade para extrair informação relevante do tensor. Informação mais detalhada acerca da modalidade de [DTI](#) e correspondentes técnicas de análise, encontra-se descrito na [Seção A.4](#) dos Apêndices.

#### Cálculo dos tensores

As soluções de software de licença gratuita que permitem, além do cálculo dos tensores e tratografia, importar regiões de interesse de estudos de [fMRI](#) são o [MEDInria](#) [62] e [TrackVis](#) [63]. No entanto, o [Diffusion Toolkit](#) disponibilizado em conjunto com o [TrackVis](#), incorpora um conjunto de soluções que lhe conferem maior flexibilidade comparativamente com o [MEDInria](#), e.g. normalizar os resultados da tratografia, facilitando assim a realização de estudos de grupo.

Assim sendo, no cálculo de tensores é de comum utilização o programa *dtifit*, disponibilizado com o [Diffusion Toolkit](#). Este programa permite simultaneamente calcular vários mapas de índices escalares, nomeadamente mapas de [Fractional Anisotropy \(FA\)](#) e [Apparent diffusion coefficient \(ADC\)](#).

#### Tratografia

Uma vez calculados os tensores, torna-se possível realizar a tratografia. Para tal, é utilizado um conjunto de programas disponibilizados pelo

programa do pacote [FSL FMRIB Diffusion Toolbox \(FDT\)](#). Inicialmente é utilizada a ferramenta [Bayesian Estimation of Diffuion Parameters using Sampling Techniques \(BEDPOSTX\)](#). Neste, são obtidas distribuições em parâmetros de difusão de cada vóxel, possibilitando a execução de tratografia probabilística.

De seguida, caso se pretenda armazenar o resultado de tratografia num espaço que não o espaço de difusão, deverá ser implementado registo. O processo de registo é realizado em [FDT](#) recorrendo ao programa [FLIRT](#). Apesar de os resultados de tratografia poderem ser resolvidos em qualquer espaço disponível (e.g. espaço de difusão, estrutural ou MNI), a tratografia em si é sempre executada no espaço de difusão.

Por último, para que a tratografia seja realizada é utilizado o programa [Probabilistic Tracking \(ProbtrackX\)](#) implementado pelo [FDT](#). Este envolve a criação de distribuições de conectividade a partir de *seed voxels* especificados pelo utilizador. O resultado é uma imagem no espaço especificado do *seed*. Todos os vóxeis cerebrais apresentam um valor (apesar de muitos serem 0), representando o valor de conectividade entre o vóxel em causa e o *seed voxel*. Os resultados de tratografia podem em última instância ser analisados visualmente com a ferramenta [TrackVis](#).

### Tract-Based Spatial Statistics

A técnica de [TBSS](#) permite efectuar estudos de grupo com base nos mapas de [FA](#). Esta análise única, baseia-se no alinhamento local dos valores máximos de [FA](#) dos feixes de cada sujeito com um *skeleton* de [FA](#) criado para o efeito. Este *skeleton* é por sua vez criado com base no alinhamento não linear dos mapas de [FA](#) de cada sujeito. Desta forma é possível realizar análises ao nível dos vóxeis, uma vez que os dados originais não sofrem transformações não lineares nem suavizações espaciais.

Esta análise é processada recorrendo à ferramenta que lhe dá o nome, [TBSS](#), disponibilizada pela biblioteca [FSL](#).

## 2.5 Análise de Grafos

A teoria de grafos possui especial interesse uma vez que é aplicável a vários níveis de conectividade estrutural e funcional. Os grafos são compostos por vértices (neurónios ou regiões cerebrais) e arestas (correspondendo a sinapses, "caminhos" ou dependências estatísticas entre elementos neuronais). Na sua forma mais simplificada, um grafo pode ser descrito através de uma matriz de conexão ou adjacência, com elementos binários representando a presença ou ausência de uma aresta direta entre pares de vértices.

A análise de padrões de conectividade estrutural, como e.g. matrizes de conectividade do córtex cerebral, permitem a quantificação de um alargado conjunto de características da rede.

### 2.5.1 Definição de Vértices e Arestas

Terminado o pré-processamento de imagens funcionais e estruturais, é necessário definir a forma como podem ser extraídas medidas de conectividade e como estas se podem traduzir num grafo. Neste sentido, existem dois aspectos essenciais a definir: a tradução das estruturas anatómicas para vértices do grafo e a relação entre essas estruturas, que representará as arestas do mesmo.

As *Regions of Interest (ROIs)* determinarão o espaço onde a informação é extraída, existindo para tal diferentes abordagens:

- utilização de um atlas já existente, com a divisão do cérebro num número razoável de regiões, sendo necessário normalizar as imagens de cada indivíduo para o espaço do atlas, ou vice-versa;
- utilização de técnicas de segmentação aplicadas à aquisição estrutural de cada indivíduo;
- análise de componentes independentes e criação de clusters.

### Atlas cerebrais

A utilização de atlas cerebrais é uma prática comum em neuroimagem. Investigadores que fazem uso de técnicas como PET e fMRI deparam-se com a necessidade de fazer corresponder os focos de actividade cerebral a estruturas anatómicas. Actualmente, existe uma grande variedade de atlas que mapeam diferentes estruturas, de diferente forma e com diferente resolução.

O atlas *Automated Anatomic Labelling (AAL)* resulta da segmentação anatómica de uma imagem espacialmente normalizada para o espaço MNI. Cada hemisfério cerebral é dividido em 45 áreas e define 26 outras áreas dedicadas ao cerebelo e vérmis, perfazendo um total de 116 regiões. Esta segmentação traduz-se numa imagem NIfTI em que cada vóxel possui um valor entre 1 e 116 conforme a estrutura anatómica em que foi classificado.

O atlas Destrieux possui uma maior resolução que o *AAL*, dividindo cada hemisfério em 74 regiões. Na sua construção, a partir de imagens MRI, foi realizada uma segmentação manual assistida por computador que classificou sulcos do cérebro.

Por último, o atlas de Desikan foi desenvolvido tendo em mente a variabilidade entre sujeitos. Para tal, foi construído a partir de imagens de MRI de 40 sujeitos, onde cada aquisição foi segmentada manualmente. As segmentações resultantes foram co-registadas criando um atlas com 68 regiões diferentes.

As segmentações obtidas através do software *Freesurfer* podem ser utilizadas como *ROIs* para definir os nós utilizados em análises de grafos.

### Ligação das redes

A ligação das redes representando as arestas dos grafos são obtidas de forma distinta para análises de conectividade funcional e estrutural. Com as imagens funcionais processadas e com as regiões de interesse definidas, é possível extrair séries temporais. Para cada ponto temporal é calculada para cada região anatómica da máscara de *ROIs*, o valor médio puro da intensidade de activação dos vóxeis correspondentes. Extrai-se desta forma um vector para cada região anatómica com tantos elementos quantos volumes

da imagem funcional. Com estas séries temporais é possível calcular quão relacionadas as diferentes regiões se encontram, obtendo-se efectivamente os valores das arestas dos grafos.

Por outro lado, no que diz respeito à conectividade estrutural, a obtenção das arestas é realizada por tratografia, através dos número de fibras que ligam as ROIs. Em ambos os casos são obtidos ficheiros de texto com entre 90x90 a 160x160 entradas a partir do qual é possível realizar análises distintas.

### 2.5.2 Análise

Com o pré-processamento e a construção de matrizes de adjacência terminados, é possível utilizar estes dados em estudos de conectividade funcional e estrutural. Face à variabilidade e circunstancialidade de cada aquisição de [MRI](#), e por forma a representar de forma fiel as ligações relevantes existentes no cérebro humano, é fundamental a capacidade de distinção entre correlações casuais, ocorrentes das circunstâncias, e aquelas que podem ser encontradas de forma recorrente. Nesse sentido, a utilização de um grupo de estudo permite, recorrendo a ferramentas estatísticas, determinar quais as ligações que são estatisticamente significativas e representativas de algum achado.

Terminada a análise de grupo, é possível proceder à análise e caracterização dos grafos. Para tal, duas abordagens são geralmente implementadas. A primeira designada por [Network Based Statistics \(NBS\)](#) permite testar a significância das ligações da rede e das subsequentes redes encontradas, baseando-se em princípios de *cluster thresholding* tradicionais de mapas paramétricos estatísticos. Uma outra implementação permite realizar uma análise da topologia da rede, com extração de métricas utilizadas na classificação da mesma.

## 2.6 Fluxo de processamento

O processamento de estudos de [MRI](#) em neurociências envolve um grande conjunto de passos, utilização de software distinto e produção de um elevado

volume de dados. A falta de padronização de processos é um grande obstáculo à organização e cooperação entre investigadores, diminuindo a eficiência, rapidez de estudo e aumentando a probabilidade de erros e consequente obtenção de resultados incorretos.

Neste sentido é prática comum a definição de fluxos de processamento, sendo a sua importância de maior relevo em estudos multimodais, ou seja, que combinem diferentes técnicas de análise.

Tendo por base este pressuposto pretende-se nesta seção apresentar um fluxo de processamento para análise estrutural, funcional e de difusão, desde a aquisição até à obtenção de resultados.

Na Figura 2.2 é apresentado um complexo e estruturado fluxo de processamento que pretende demonstrar a variedade de passos presentes na realização de estudos multimodais de MRI. Como observável na Figura 2.2, os diversos processos envolvidos foram numerados, pretendendo-se de seguida explicar cada um destes passos:

1. Realização de aquisições estruturais, funcionais e de difusão MRI;
2. Transporte de imagens em formato IMA num dispositivo (Pen ou disco rígido) da instituição, no qual foram realizadas as aquisições para o centro de investigação;
3. Conversão de formato IMA para DICOM, e.g. através do software Osirix;
4. Estruturação em diretorias das imagens DICOM consoante o tipo de aquisição;
5. Conversão de formato DICOM para NIFTI, e.g. através do software dcm2nii;
6. Utilização do software Freesurfer para pré-processamento, segmentação, registo e reconstrução de superfícies em imagens estruturais;
7. Utilização do software FSL e SPM para análise VBM e GLM de imagens estruturais;

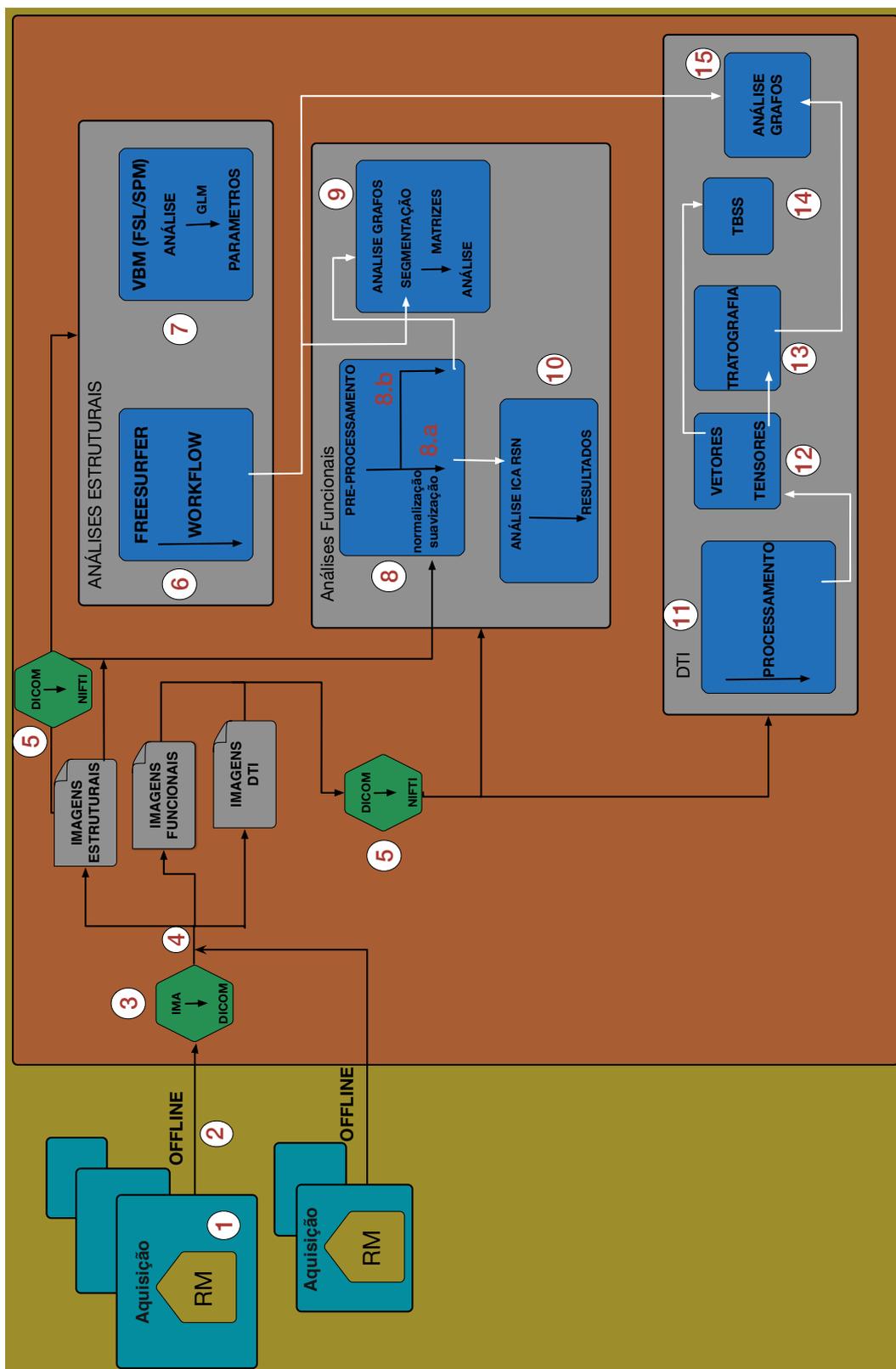


Figura 2.2: Fluxo de processamento global para estudos MRI

8. Utilização do software [FSL](#) para pré-processamento de imagens funcionais;
  - (a) com normalização e suavização;
  - (b) sem normalização e suavização.
9. Utilização de *scripts* Matlab, e [NBS](#) para análise de grafos e construção de matrizes através de imagens funcionais pré-processadas, obtidas no passo 8.b, e segmentação de imagens estruturais, obtida no passo 6;
10. Utilização do software [FSL](#) para análise [ICA](#) e posterior análise estatística de imagens funcionais pré-processadas obtidas no passo 8.a;
11. Utilização do software [FSL](#) para pré-processamento de imagens de difusão;
12. Utilização do software [FSL](#) em imagens [DTI](#) pré-processadas, para obtenção de vetores e cálculo de tensores (e.g. [FA](#));
13. Preparação de dados e utilização do software [FDT](#) e [TrackVis](#) para tratografia determinística e probabilística ([BEDPOSTX](#) e [ProbtrackX](#));
14. Utilização do software [FSL](#) para [TBSS](#);
15. Utilização de *scripts* Matlab, e [NBS](#) para análise de grafos e construção de matrizes através de ficheiro de tratografia, obtido no passo 13, e segmentação de imagens estruturais, obtida no passo 6;

Como evidenciado na lista de passos de processamento descrita, foram consideradas 15 etapas necessárias para um estudo multimodal completo. Da realização destas etapas resulta uma grande quantidade de informação que deverá acessível ao(s) investigador(es) envolvido(s) num determinado estudo.

De seguida pretende-se aprofundar cada momento de análise apresentado na Figura 2.2, referindo a informação gerada que deverá ser mantida organizada, e idealmente partilhável entre investigadores. Desta forma na Figura 2.3, 2.4 e 2.5 é apresentado com maior detalhe o fluxo de processamento efectuado em imagens funcionais, difusão e estruturais, respectivamente. Em

cada uma destas Figuras, são enumerados momentos de produção de ficheiros considerados essenciais para o sucesso de estudos multimodais, bem como o software utilizado.

Inicialmente será considerado o fluxo de processamento para análises funcionais representado na Figura 2.3, seguido do fluxo para imagens de difusão (Figura 2.4), concluindo com os passos utilizados em aquisições estruturais (Figura 2.5):

- 1) Na conversão de formato de imagens **DICOM** funcionais para **NIFTI** é obtido uma imagem 4D, sendo sobre esta que se inicia o fluxo de pré-processamento de análises funcionais. Assim, inicialmente é realizada o passo de extração de volumes com utilização do software **FSL**;
- 2) Após extração de volumes é realizado o passo de correção dos tempos de aquisição, também com utilização de software **FSL**;
- 3) O passo seguinte é a correção de movimentos de imagens funcionais, do qual resultam vários ficheiros e parâmetros:
  - (a) é obtida uma imagem (formato *nii.gz*) resultante da aplicação do software **Motion Correction FMRIB's Linear Image Registration Tool (MCFLIRT)**, com os movimentos corrigidos;
  - (b) além desta imagem são obtidos parâmetros de movimento (formato *par*), bem como 2 imagens (formato *png*) que correspondem ao movimento rotacional e translacional efectuado. Também para estes é utilizado o software **MCFLIRT**;
  - (c) por último, é utilizada a ferramenta *fsl\_motion\_outlier* (desenvolvida para detectar momentos temporais num *dataset fMRI* corrompido por grandes movimentos, calculando métricas de movimento utilizadas em análise **GLM**). Neste sentido são obtidos dois ficheiros (formato *txt*) contendo métricas de *motion outliers* e *DVARs*. *DVARs* indexa a taxa de variação do sinal **BOLD** ao longo do cérebro em cada *frame*, ou por outras palavras, a forma como a intensidade de uma imagem cerebral se altera relativamente a um momento temporal anterior.

- 4) concluído o passo de correção de movimentos, é possível realizar a extração do crânio. Como resultado é obtida uma máscara cerebral, através da utilização do comando *fslmaths* sobre o ficheiro com correção de movimentos.
- 5) após conclusão do passo de extração do crânio podem ser seguidos dois rumos de pré-processamento:
  - (a) no primeiro ramo (representado no lado esquerdo do fluxo de pré-processamento de imagens funcionais (Figura 2.3), após extração do crânio), são realizados os passos de normalização, *denoise*, suavização e filtragem, representados pela numeração 5,6,7 e 8 respectivamente. No passo de normalização são geradas duas matrizes (formato *mat*), que resultam da transformação de imagens funcionais para o espaço MNI e estrutural, através da execução do software **FLIRT**. Cumprido o passo de normalização é então obtida informação normalizada (formato *nii.gz*), através do software **FLIRT** e **FNIRT**, bem como uma imagem da média de normalização (formato *nii.gz*), obtida com recurso ao software *fslmaths*. Após normalização, são efectuados os passos de *denoise*, suavização e filtragem, podendo ser obtidas diferentes combinações de ficheiros (formato *nii.gz*), resultantes da "não realização" de um dos passos de suavização ou *denoise* ou mesmo de ambos. Para estes passos é utilizado o software *fslmaths*, obtendo-se no final imagens pré-processadas (formato *nii.gz*) que poderão ser utilizadas para análise **ICA**;
  - (b) no segundo ramo (representado no lado direito do fluxo de pré-processamento de imagens funcionais (Figura 2.3), após extração do crânio), são realizados apenas os passos de pré-processamento de *denoise* e filtragem, representados no número 9 do fluxo em questão. Na obtenção destes ficheiros é utilizado o software *fslmaths*, obtendo-se no final imagens pré-processadas que poderão ser utilizadas em análises de grafos. Para tal são obtidas séries temporais de cada umas das **ROIs** selecionadas para as imagens

funcionais pré-processadas (número 10 do fluxo representado). Estas séries temporais são obtidas através do comando `FSL fslmeants`, resultando um ficheiro *txt*. Este é utilizado em conjunto com ficheiros de segmentação, obtidos em imagens estruturais, para posterior análise de grafos;

De seguida serão apresentados os passos envolvidos em análises de difusão, cujo fluxo está representado na Figura 2.4. Da conversão de imagens `DICOM` de difusão para o formato `NIFTI` (imagem 4D) resultam ainda ficheiros com informação relativa ao gradiente de pulsos aplicado durante a aquisição. Desta forma, é obtido um ficheiro com extensão *bval* com valores do gradiente de difusão e um ficheiro com extensão *bvec* com os vectores do mesmo. É ainda obtida uma imagem (formato *nii.gz*) designada *b0*, que corresponde a uma imagem base sem qualquer informação de difusão.

- 11) tal como descrito para análises funcionais, também em imagens de difusão são aplicados passos de pré-processamento. Neste sentido o fluxo é iniciado com a correção de movimentos e distorções. Da execução do comando `FSL eddy_correct` é obtido um ficheiro de extensão *ecclog* com informação para a rotação dos vectores.
- 12) após conclusão destes passos é realizada a extração do crânio, através do software `BET`, obtendo-se uma máscara cerebral (formato *nii.gz*);
- 13) concluída a extração do crânio, é executada a rotação de vectores com o comando `FSL rotate_bvecs`, obtendo-se deste processo um ficheiro de extensão *bvec* com informação dos vetores de difusão processados. No final do pré-processamento é obtido um ficheiro de extensão *nii.gz* com os dados pré-processados de `DTI`. Para tal é de novo utilizado o software o programa *eddy\_correct*. Tal como em aquisições funcionais, podem também ser obtidas matrizes de transformação para o espaço de coordenadas `MNI` e estrutural, através da utilização do software `FLIRT`;

Concluído o pré-processamento das imagens de difusão é possível obter um diverso conjunto de informação. Tendo como *input* o ficheiro de dados de difusão pré-processados e o ficheiro de extensão *bvec* com informação dos vetores de difusão processados (gradientes de difusão rodados) são calculados mapas, tensores e vetores. Mais especificamente são obtidos 11 ficheiros de seguida discriminados:

- 14) inicialmente são obtidos os tensores de difusão, com utilização do software [FDT](#);
- 15) de seguida são obtidos o mapa do primeiro *engenvalue*, necessário para cálculo de [Mean Diffusivity \(MD\)](#), [Axial Diffusivity \(AD\)](#) e [FA](#), o mapa do segundo *engenvalue*, para cálculo de [MD](#), e [FA](#) e o terceiro *engenvalue* para cálculo de [MD](#), e [FA](#). São também obtidos o primeiro, segundo e terceiro *eingenvalue*, utilizando para ambos o software [FDT](#);
- 16) uma vez obtidos os *eigenvalues* e *eigenvectors* são obtidos os mapas necessários para análise *voxelwise* [TBSS](#), respectivamente os mapas [FA](#), [MD](#) e ;
- 17) para realização de tratografia determinística é necessário a utilização de um ficheiro de extensão *bvec*, contendo vetores de difusão. Este ficheiro diferencia-se dos restantes *bvec*, uma vez que se encontra no sistema de coordenadas do sistema de [MRI](#), que é por sua vez o utilizado pelo software requerido, designado [TrackVis](#);
- 18) da realização de tratografia determinística, através do software [Track-Vis](#) são obtidos dois ficheiros de extensão *trk*, correspondendo à tratografia determinística e mapa de [FA](#) colorido;
- 19) para realização de tratografia probabilística é necessário preparar uma diretoria de *input* de dados para o software *bedpostx*. Desta forma, deverá estar presente nessa mesma diretoria os seguintes ficheiros: um ficheiro de volumes de dados 4D de difusão, uma máscara cerebral resultante da execução do software [BET](#), ficheiro de texto com os gradientes de difusão aplicados e um outro ficheiro de texto com a lista de

valores de  $b$  aplicados durante a aquisição de difusão. Após utilização de **BEDPOSTX** serão obtidos vários ficheiros, os quais podem ser utilizados pelo software *probtrackx* para tratografia probabilística. Após execução de tratografia probabilística, estes dados podem ser utilizados em conjunto com ficheiros de segmentação de imagens estruturais para análise de grafos;

Por último pretende-se explorar o fluxo existente para análises estruturais, representado na Figura 2.5, no qual é também obtida informação de suporte para análises funcionais e de difusão.

- 20) como evidenciado na seção esquerda da Figura 2.5 é executado o fluxo de processamento freesurfer. Tal como referido, este fluxo é utilizado para segmentação e classificação de imagens de **MRI**. Para tal são utilizadas imagens **DICOM** estruturais ao qual são aplicadas etapas de pré-processamento, seguidas de segmentação, registo, reconstrução de superfícies e classificação. Da execução deste fluxo complexo são gerados vários ficheiros e parâmetros, do qual se destacam os ficheiros representados em 20, 21 e 22. O ficheiro *aparc.a2009s.aseg.mgz* (20) corresponde à segmentação de estruturas subcorticais utilizando o atlas *Destrieux*. O ficheiro *aparc..aseg.mgz* (21) é semelhante ao ficheiro anteriormente referenciado, com a diferença de utilização do atlas *Desikan-Killiany*. Por último o ficheiro *T1.mgz* (22) é, como o nome indica, a representação de imagens *T1-weighted*;
- 23) além da utilização de imagens estruturais em formato **DICOM**, também estas devem ser convertidas para o formato NIfTI (imagem 3D), que será utilizada na seção de processamento direita da Figura 2.5. A estas é também aplicado um fluxo de pré-processamento, iniciado pelo passo de normalização. Para tal é utilizado o software **FLIRT** e **FNIRT**. Da utilização de **FNIRT** obtém-se um ficheiro (formato *nii.gz*) de transformação não linear para MNI, utilizada como apoio a todas as análises. Da mesma forma, da utilização de **FLIRT** são obtidos mais 3 ficheiros: matriz de transformação para MNI (apoio a todas as análises), matriz de transformação para o espaço de difusão (suporte a todas as análises

de difusão) e matriz de transformação para o espaço funcional (apoio a todas as análises funcionais). No final deste passo de pré-processamento são obtidos dados normalizados (formato *nii.gz*);

- 24) Concluída a normalização, é iniciado o processo de segmentação de imagens estruturais. Para tal é utilizado o software `bet`, sendo obtida uma máscara cerebral (formato *nii.gz*).
- 25) uma vez concluído o processo de segmentação é possível obter **Tissue Probability Maps (TPMs)**, essenciais para a realização de análise **VBM** e suporte a análises funcionais. Estes **TPMs** devem representar a probabilidade a priori de cada voxel pertencer a uma classe particular. Estes são obtidos ponderando um número alargado de tecidos espacialmente normalizados de diferentes sujeitos. Para obtenção destes dados é utilizado o software **FAST** resultando 6 ficheiros (formato *nii.gz*): estimativa de volume parcial de **LCR**, **SB** e **Substância cinzenta (SC)**; **TPM** de **LCR**, **SB** e **SC**. Com a obtenção destes ficheiros é possível a realização de **VBM** e posterior análise estatística **GLM**.

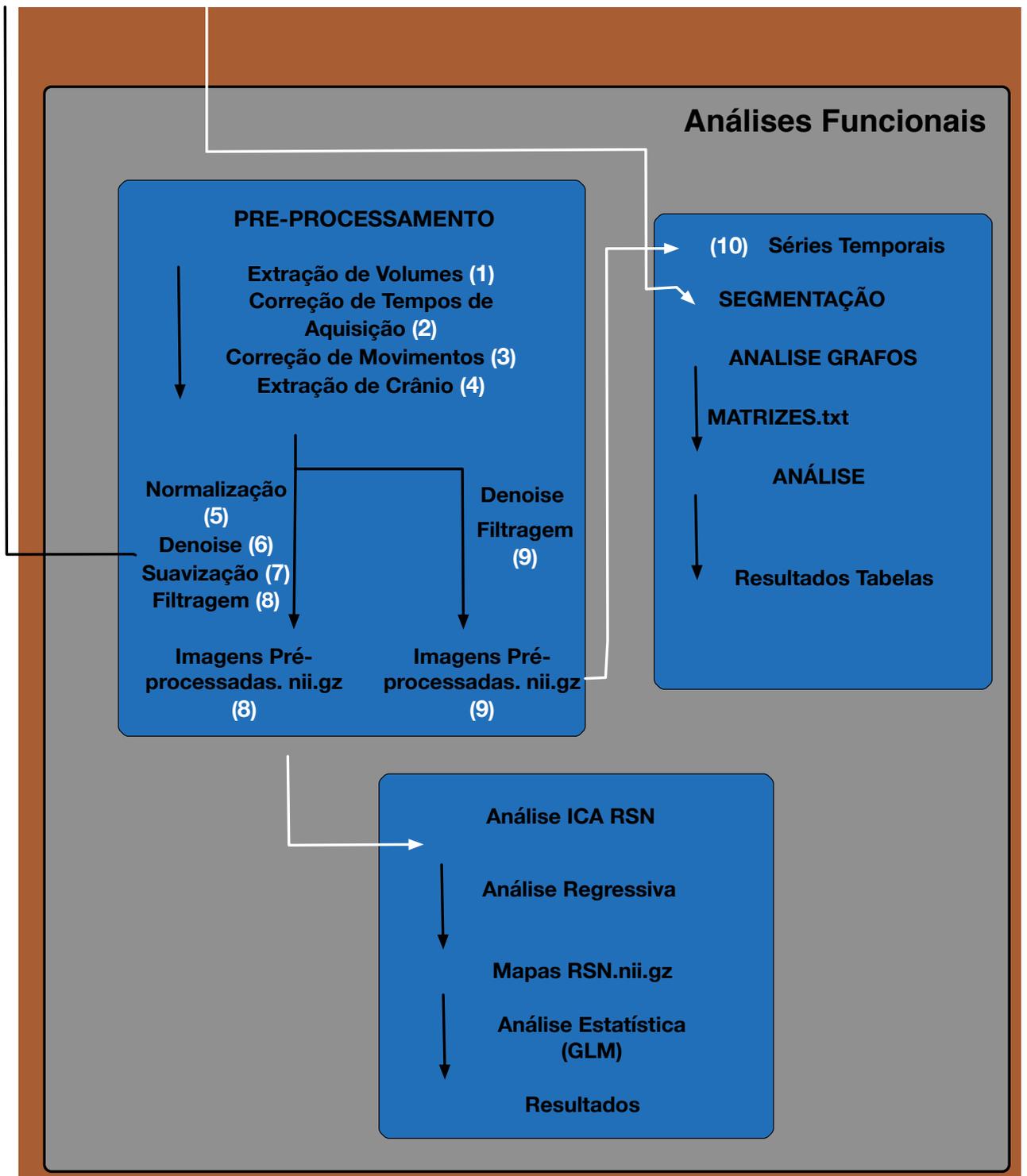


Figura 2.3: Fluxo de processamento para análises funcionais de MRI

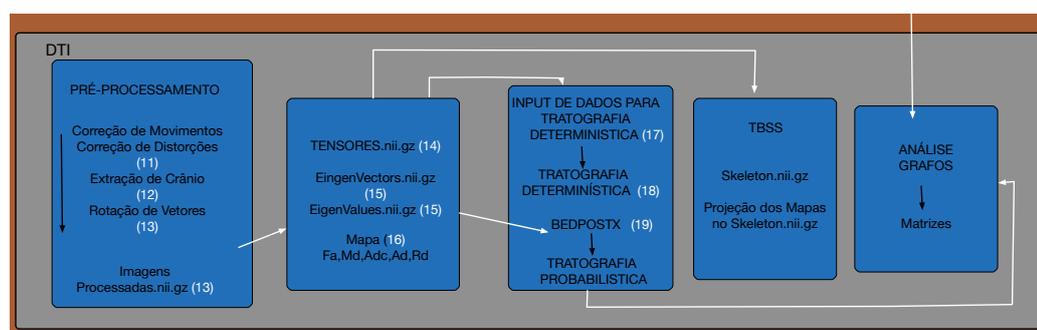


Figura 2.4: Fluxo de processamento para análises de difusão de MRI

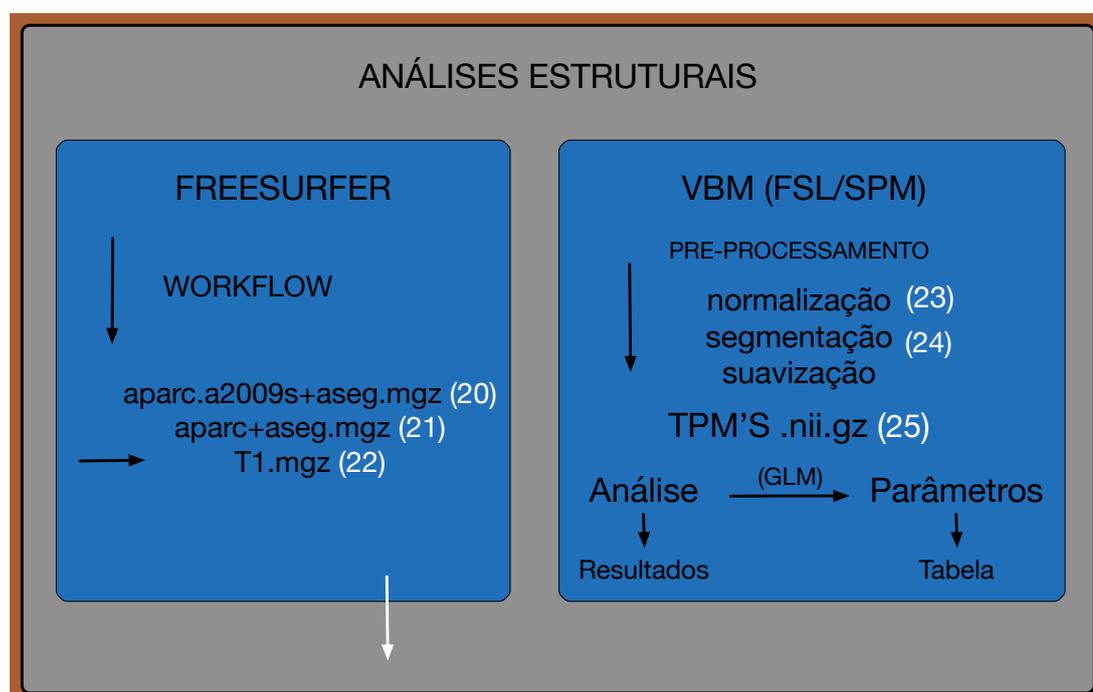


Figura 2.5: Fluxo de processamento para análises estruturais de MRI

## 2.7 Volume de dados

Estudos de **MRI** em neurociências produzem uma grande quantidade de informação, a qual idealmente deverá ser organizada e partilhada de forma segura entre instituições, investigadores e profissionais de saúde. Nesta seção pretende-se descrever com maior detalhe o volume de dados gerados em estudos de **MRI** cerebral.

O número de casos num estudo de aquisições de **MRI** é variável. São tipicamente utilizados entre 50 e 200 casos, dependendo de recursos e necessidades da própria investigação. Cada caso é geralmente composto por 4 a 8 aquisições. Numa aquisição estrutural são obtidas 176 imagens em formato **DICOM**, cada uma com 201 **Kylobytes (Kbs)**. Após conversão para formato **NIfTI** é gerada uma imagem 3D com 11 **Megabytes (MBs)**. Em aquisições de difusão são obtidas tipicamente entre 30 e 260 imagens em formato **DICOM**, com 2.2 **MBs** cada. Após conversão para formato **NIfTI** é gerada uma imagem 4D com 64 **MB**. Por fim, são tipicamente utilizadas 2 a 4 aquisições funcionais. Para cada aquisição funcional são obtidas aproximadamente entre 100 a 1000 imagens **DICOM**, com 400 **Kbs** cada uma. Após conversão de para o formato **NIfTI** é obtida uma imagem 4D com tamanho variável entre aproximadamente 23 e 205 **MB**.

Na realização dos diversos passos de pré-processamento são gerados inúmeros ficheiros. Estes possuem tamanho variável, desde alguns bytes até mais de 300 **MBs**. Desta forma, para um caso típico de análise estrutural, funcional e de difusão é obtida informação num total aproximado de 80 **MBs**, 2.6 **Gigabytes (GBs)** e 800 **MBs**, respectivamente.

Neste sentido, para um caso com aquisições estruturais, funcionais e de difusão são produzidos aproximadamente 1600 ficheiros num total aproximado de 5 **GB**. Assim, considerando um estudo constituído por 100 casos é gerada informação num valor superior a 500 **GB**.

Na Figura 2.6 é apresentado de forma esquemática o volume de dados gerado para um caso típico de estudo.

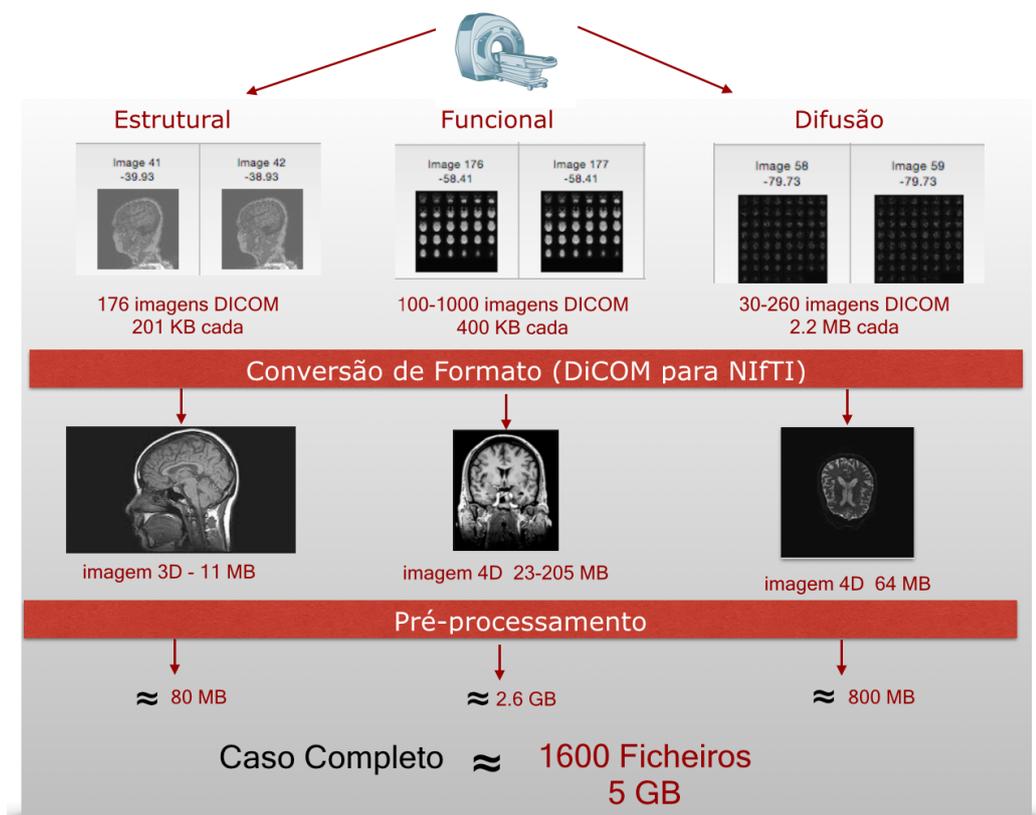


Figura 2.6: Volume de dados aproximado para um caso completo de estudo de MRI

## 2.8 Soluções Disponíveis

A procura de soluções é essencial para a percepção do que já se encontra disponível. Neste sentido, foi realizada uma pesquisa por soluções que permitam a partilha, armazenamento e difusão de estudos de neuroimagem. Ao estudar outros sistemas é possível encontrar funcionalidades genéricas, verificar as funcionalidades que não se encontram implementadas e aquelas que podem ser exploradas de outra forma.

### 2.8.1 ANIMA

A arquitetura ANIMA foi desenvolvida para corresponder a um conjunto de funções. Inicialmente, é um repositório online para resultados de meta-análises de neuroimagem publicados. A informação é organizada em estudos individuais, juntamente com informação essencial descrevendo esses estudos (autor, título, revista, data, resumo, entre outros).

Em cada estudo existe um conjunto de ficheiros para armazenamento de aspectos relevantes do mesmo. Estes são tipicamente ficheiros de volume individuais, armazenados no formato NIfTI. No entanto outro tipo de dados pode ser submetido, incluindo ficheiros de texto contendo descrições úteis, arquivos de imagem ou ficheiros de sessão da aplicação *Volume Viewer*, no qual são armazenadas múltiplas imagens e informação de como devem ser visualizadas.

A cada elemento do estudo está associado um conjunto de metadados descrevendo o seu conteúdo (e.g. título, legenda da Figura, tipo de informação, PDF). Na Figura 2.7 é visível a página de pesquisa da interface online Anima.

Uma segunda função da ANIMA é a possibilidade de submissão do utilizador dos seus próprios estudos à base de dados. Para tal, a interface online inclui uma *Submit page* que permite a utilizadores registados definir os seus estudos e exportar ficheiros.

Por questões de segurança e integridade dos dados, cada estudo é primeiramente verificado pelos administradores da base de dados. Na Figura 2.8 é possível observar a página de submissão de estudos da interface online

Study	Description	Size
<input type="checkbox"/> <b>Kohn et al. 2013</b> Neural network of cognitive emotion regulation—an ALE meta-analysis and MACM analysis. <i>Neuroimage</i> , 87: 345-355.	We here quantitatively summarize the published literature on cognitive emotion regulation using activation likelihood estimation in fMRI and PET (23 studies/479 subjects). According to our model, the superior temporal gyrus, angular gyrus and (pre) supplementary motor area should be involved in execution of regulation initiated by frontal areas. The dorsolateral prefrontal cortex may be related to regulation of cognitive processes such as attention, while the ventrolateral prefrontal cortex may not necessarily reflect the regulatory process per se, but signals salience and therefore the need to regulate.	1.4M
<input type="checkbox"/> <b>Schilbach et al. 2012</b> Introspective Minds: Using ALE Meta-Analyses to Study Commonalities in the Neural Correlates of Emotional Processing, Social & Unconstrained Cognition. <i>PLoS One</i> , 7(2).	We present results of quantitative meta-analyses of neuroimaging studies, which confirm a statistical convergence in the neural correlates of social and resting state cognition. By using conjunction analyses across studies investigating social cognition, resting state, and emotion, we demonstrate significant overlap of task-related signal change in dorso-medial prefrontal and medial parietal cortex, brain regions that have recently been linked to introspective abilities.	4.2M
<input type="checkbox"/> <b>Bzdok et al. 2012</b> Parsing the neural correlates of moral cognition: ALE meta-analysis on morality, theory of mind, and empathy. <i>Brain Structure and Function</i> , 217(4): 783-96.	We tested the hypothesis that moral decisions might be implemented in brain areas engaged in "theory of mind" and empathy, by conducting a large-scale activation likelihood estimation (ALE) meta-analysis of neuroimaging studies, which assessed 2,607 peak coordinates from 247 experiments in 1,790 participants.	506.6K
<input type="checkbox"/> <b>Friebe et al. 2011</b>	We performed a meta-analysis based on a literature search of published functional	

Figura 2.7: Página de pesquisa da interface online Anima (retirado de [39])

## ANIMA

Por último, ANIMA foi desenvolvido com o intuito de permitir que dados importados sejam facilmente organizados e pesquisados a nível local por parte do utilizador. Nesse sentido foi implementada uma aplicação *Desktop* designada *Volume Viewer*, no qual estudos podem ser representados e organizados.

Por conseguinte se um estudo incluir ficheiros de sessão *Volume Viewer*, conjuntos de imagens podem ser visualizadas facilmente por parte do utilizador. Na Figura 2.9 é apresentada a visão geral de funcionamento da arquitetura ANIMA.

### 2.8.2 1000 Functional Connectomes Project (FCP)

Este projecto é parte do [International Neuroimaging Data-Sharing Initiative \(INDI\)](#), apresentado-se como solução de partilha de 1200 imagens 4D de *resting state functional MRI* independentemente colecionada de 33 sites. Estas imagens formam uma ampla comunidade de acesso a datasets funcionais. Dados como idade, sexo e centro de imagiologia são providenciados

**A**

LOGIN STUDY ELEMENTS SUBMIT

Study name

Study ID  Content is saved.

Full text PDF   (uploaded)

Description

Abstract

Reference

Title

---

**B**

LOGIN STUDY ELEMENTS SUBMIT

Enter a new study element

Name

Type

Caption

File

Coordinates

Quantity

Threshold

Name	Type	File	Caption	Actions
Figure 2	Volume file (nii, mnc, mgz)	NbackSternbergVerbalNonverbal_andMore_3Phasen_thr20.nii.gz	4	

---

**C**

LOGIN STUDY ELEMENTS SUBMIT

Almost there! Review the summary below, and go back and make changes if necessary. If you're ready to submit, just click the button...

Study	Description	Size (uncompressed)
Smith et al., 2010	We investigated the neural correlates of aggressive behaviour using an activation likelihood estimation (ALE) analysis.	6.9M
<p><small>Why so aggressive? A meta-analytic analysis of aggressive behaviour. Journal of Excellent Neuroscience. 1(2): 123-130.</small></p>		

Figura 2.8: Página de submissão de estudos da interface online ANIMA. (A) *Study tab* (B) *Elements tab* (C) *Submit tab* (retirado de [39])

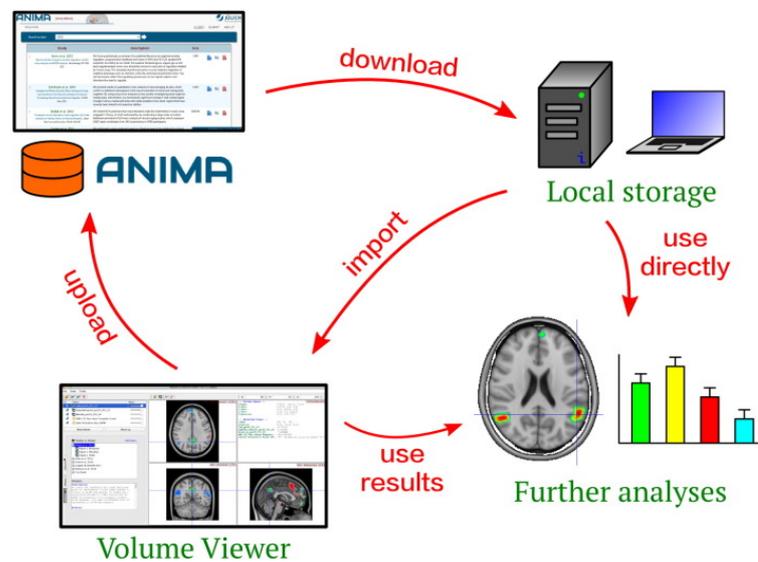


Figura 2.9: Visão geral de utilização da base de dados ANIMA, ilustrando o processo de pesquisa, seleção, *download* e importação de dados na biblioteca *Volume Viewer* (retirado de [39])

para cada imagem. De acordo com as orientações [Health Insurance Portability and Accountability Act \(HIPAA\)](#), todas as imagens são anônimas, não apresentando informação clínica protegida.

### 2.8.3 GAAIN

O [Global Alzheimer’s Association Interactive Network \(GAAIN\)](#) é uma grande comunidade de colaboração e exploração de informação. Nesta é promovida a partilha de dados numa rede global de parceiros que estudam a doença de alzheimer entre outras problemas neurológicos.

Cada parceiro na rede [GAAIN](#) executa um cliente local [Data Partner Client \(DPC\)](#) que, por sua vez, responde a pedidos dos servidores [GAAIN](#). O [DPC](#) é um ficheiro executável Java jar, contendo a informação necessária para a ligação à rede *gaain*. Esta incorpora um servidor web e uma base de dados. A instalação é rápida e simples, sendo a configuração realizada utilizando o *web browser*. Os parceiros em causa possuem sempre controlo dos seus dados, uma vez que cada possui um interruptor ON/OFF, que

permite a ligação segura à rede **GAAIN**, assim como o encerramento da ligação e remoção dos seus dados da rede.

Na Figura 2.10 é explicado com maior detalhe o fluxo de informação decorrente da arquitetura **GAAIN**

Dados clínicos são adicionados à rede **GAAIN** utilizando um ou mais ficheiros **Comma-separated values (CSV)** exportados através de um **DPC**. Este por sua vez não necessita de acesso à base de dados corrente do *Data Partner*, podendo ser executado de um sistema fisicamente separado. Os ficheiros **CSV** são lidos no **DPC** e os dados são mapeados no esquema **GAAIN** seguindo código CDISC (uma organização sem fins lucrativos e multidisciplinar que estabelece padrões de suporte à aquisição, troca, submissão e arquivo de dados clínicos e correspondente metadata).

Através da padronização de códigos e convenções de cada parceiro, os investigadores **GAAIN** não necessitam de compreender individualmente a complexidade de dados de cada estudo. Estes dados mapeados são armazenados localmente no **DPC** e, quando necessário, enviados para os servidores **GAAIN** para responder a solicitações. Na Figura 2.11 é ilustrada a relação cliente servidor da arquitetura **GAAIN**.

Os investigadores **GAAIN** pesquisam por todos os dados disponibilizados por *Data Partners* utilizando o interrogador **GAAIN**. Este exhibe a investigadores a forma como os dados se relacionam, através de dados gráficos interconectadas. Neste sentido, em vez de introduzir números em caixas de texto, os dados são filtrados pela interação com os gráficos, alterando por conseguinte os parâmetros de pesquisa e actualizando automaticamente a visualização dos dados.

Cada resultado de pesquisa ilustra o número de sujeitos envolvidos em cada *Data Partner* e direciona os investigadores na forma de obtenção de dados.

#### 2.8.4 XNAT

O **Extensible Neuroimaging Archive Toolkit (XNAT)** é uma plataforma de software desenvolvida para facilitar a gestão e produtividade de tarefas

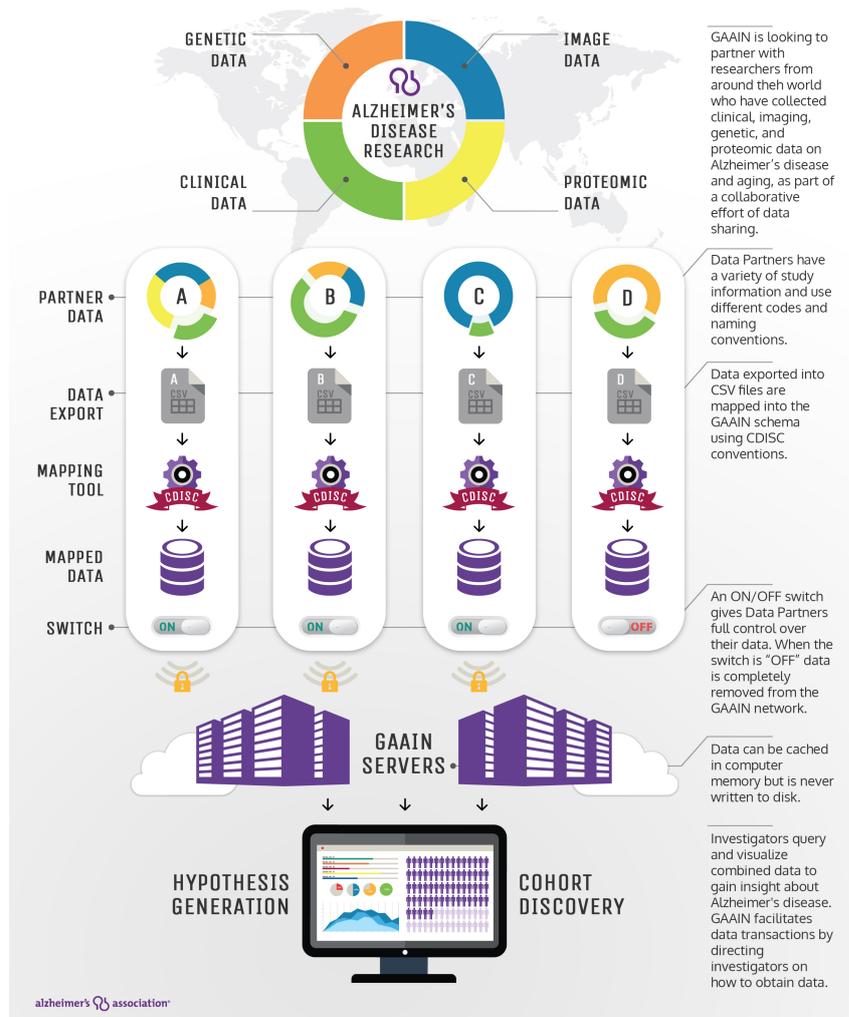


Figura 2.10: Fluxo de informação presente na arquitetura GAAIN (retirado de [64])

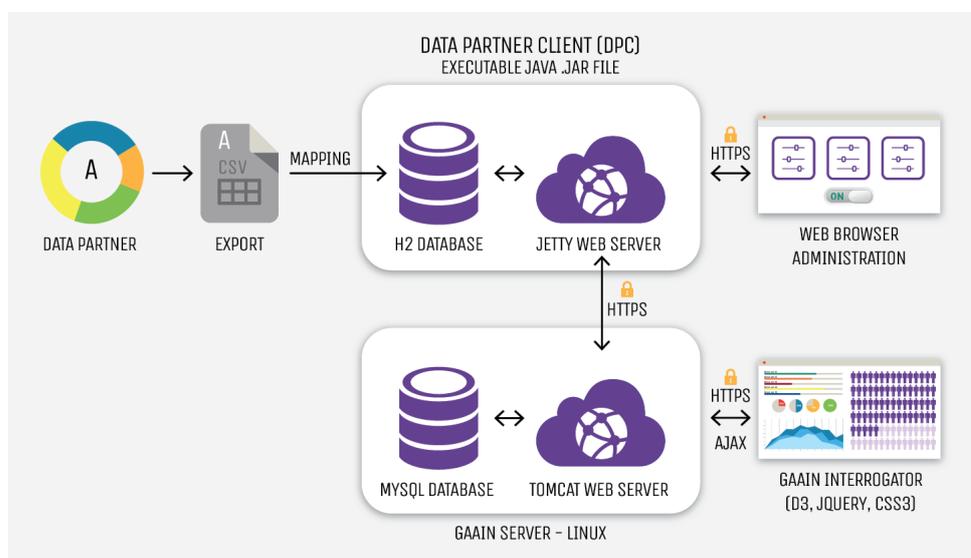


Figura 2.11: Esboço da arquitetura cliente/servidor GAAIN (retirado de [64])

de neuroimagem e dados associados. De forma particular, este possui procedimentos de controlo de qualidade e providencia acesso e armazenamento de dados seguro.

Na Figura 2.12 é possível observar a ideia sobre o qual a arquitetura XNAT foi desenvolvida.

Os dados podem ser introduzidos no sistema como **EXtensible Markup Language (XML)** ou através de formulários. Inicialmente são armazenados numa quarentena virtual até que um utilizador autorizado o valide. Subsequentemente é mantido um histórico de alterações dos dados a gerir.

O acesso ao repositório é realizado através de uma aplicação web segura, providenciando pesquisa por tipo de dados específicos e combinados, relatórios detalhados, listas de dados experimentais, ferramentas de importação/exportação e acesso a fluxos de processamento laboratoriais.

XNAT inclui ainda um visualizador online de imagens suportando formatos de neuroimagem como **DICOM** e **Analyze**.

Na Figura 2.13 é apresentado o fluxo de trabalho realizado através da arquitetura XNAT.

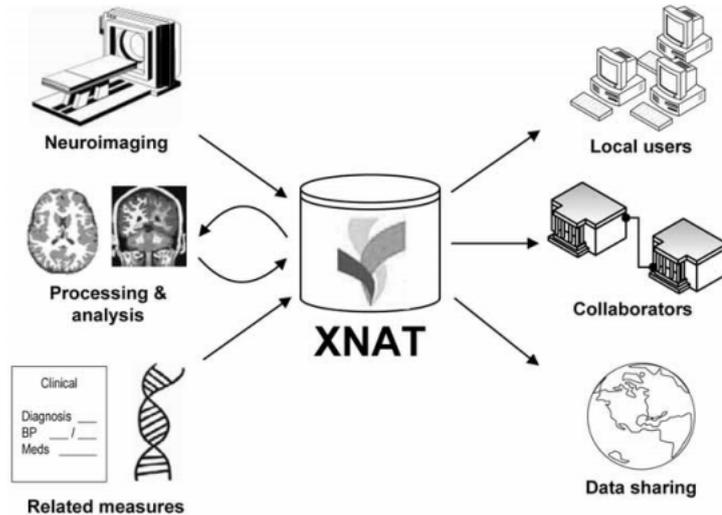


Figura 2.12: Arquitetura XNAT como serviço de captura de dados em múltiplas fontes e distribuição a uma variedade de utilizadores (retirado de [65])

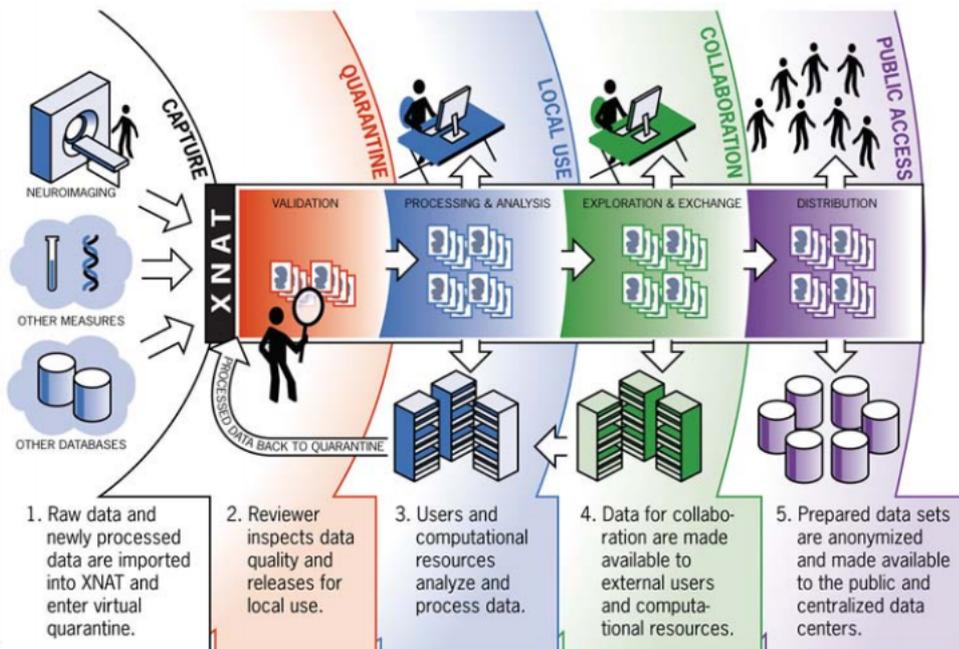


Figura 2.13: Fluxo de trabalho da solução XNAT (retirado de [65])



# Capítulo 3

## Tecnologia

### 3.1 Introdução

No desenvolvimento de uma aplicação é necessário ter em conta uma série de factores. Devem ser escolhidas as tecnologias mais adequadas tendo em conta o objectivo da aplicação e o meio em que a mesma vai ser inserida. Apesar da existência de uma multiplicidade de escolhas para resolver o mesmo problema, pretende-se, no entanto, decidir a solução mais adequada.

Inicialmente é necessário fazer o levantamento de requisitos da aplicação de forma a analisar quais as tecnologia que devem ser utilizadas na elaboração da mesma. Além disso é necessário reconhecer o meio em que esta vai estar integrada e com o que deve interagir, uma vez que pode existir a necessidade de integração com outros sistemas ou usar recursos partilhados com outras aplicações (e.g. base de dados, sistemas de armazenamento).

Deve ainda ser considerado o público-alvo, percebendo as suas necessidades e características, garantindo uma interface do utilizador cuidada, facilmente acessível e de simples interação.

### 3.2 Arquiteturas de Redes

Um dos aspectos de maior relevância no desenvolvimento de uma aplicação distribuída, baseia-se na escolha e implementação de um modelo de

comunicação. Estes são geralmente classificados em duas categorias abrangentes, arquiteturas cliente-servidor e [Peer-to-Peer \(P2P\)](#).

### 3.2.1 Cliente-Servidor

O paradigma de redes cliente-servidor é um sistema no qual um ou mais dispositivos designados clientes estabelecem comunicação com um outro dispositivo designado servidor. O processamento de dados é assim decomposto em duas entidades: solicitador (cliente) e fornecedor (servidor). O cliente envia um ou mais pedidos ao(s) servidor(es) para realização de determinada tarefa e o servidor, por sua vez, providencia serviços aos clientes [66]. Este mecanismo de pedido-resposta é o padrão de comunicação entre clientes e servidores, como evidenciado na Figura 3.1.

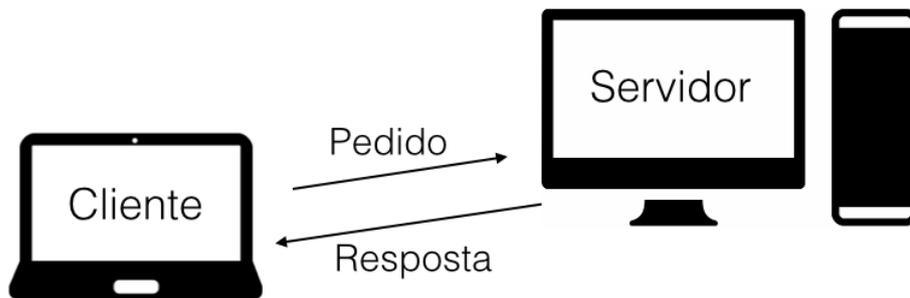


Figura 3.1: Troca de mensagens padrão numa rede Cliente-Servidor

Para comunicar, os dispositivos envolvidos devem utilizar uma linguagem comum e regras específicas definidas num protocolo de comunicação. Todos os protocolos cliente-servidor atuam na designada camada de aplicação, cujo protocolo define os padrões base de diálogo.

Transações computacionais no qual um servidor "satisfaz" o pedido de um cliente são muito comuns, tornando o modelo cliente-servidor numa das ideias centrais de redes em computadores. Assim sendo, várias aplicações fazem uso deste paradigma como e.g. o conjunto de protocolos Internet [Transmission Control Protocol/Internet Protocol \(TCP/IP\)](#) [67].

Cientes e servidores são geralmente parte de um grande programa ou aplicação. Nesse sentido, e uma vez que múltiplos clientes partilham serviços de um mesmo servidor, um servidor específico denominado *daemon* deve estar ativo por forma a aguardar e gerir pedidos dos clientes [68]. Um esboço típico de arquitetura cliente-servidor pode ser observado na Figura 3.2

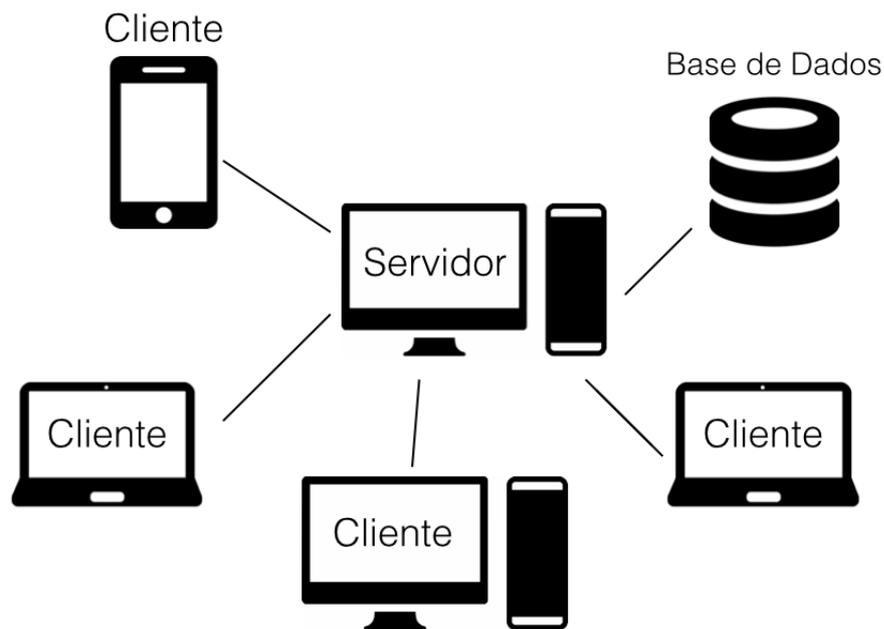


Figura 3.2: Arquitetura Cliente-Servidor

### 3.2.2 Peer-to-Peer

Num modelo P2P, a arquitetura do sistema é decomposta em vários nós computacionais, de capacidade e responsabilidade equivalente. Estes nós (peers) tornam uma porção de recursos como capacidade de processamento, armazenamento em disco ou largura de banda de rede disponível para outros elementos da rede, sem a necessidade de coordenação central por parte de um servidor [69].

Neste sentido, os nós são simultaneamente fornecedores e consumidores de recursos (como evidenciado na Figura 3.3), em contraste como o modelo cliente-servidor no qual o consumo e fornecimento de recursos é dividido.

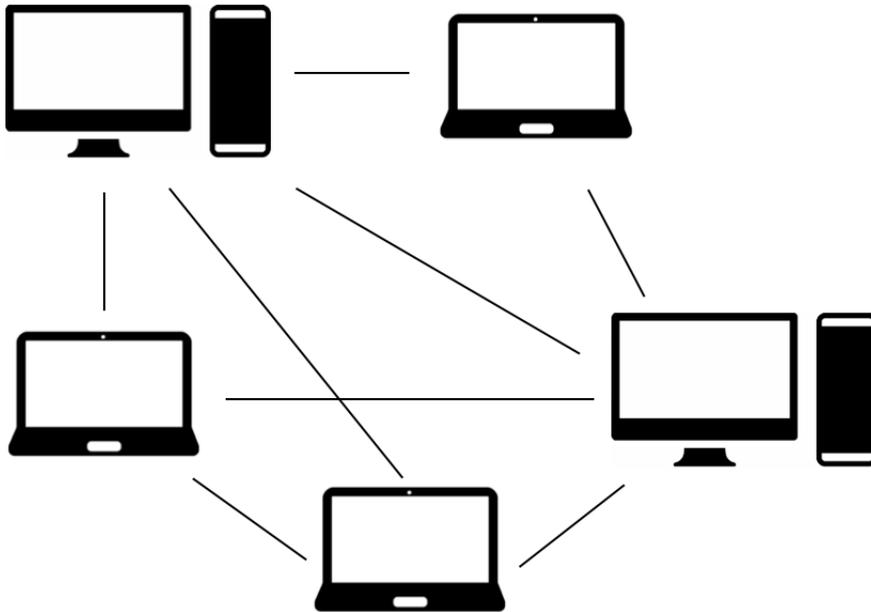


Figura 3.3: Arquitetura Peer-to-Peer

**P2P** atraiu grande atenção da comunidade académica e empresarial. Aplicações **P2P** como Napster, Gnutella, FastTrack, BitTorrent ou Skype apresentaram um grande sucesso entre os seus utilizadores. Contrariamente a um sistema cliente-servidor, os nós incorporam no sistema capacidade de "serviço". Nesse sentido, assim que a exigência do sistema **P2P** aumenta, a capacidade da rede é também acrescida. Isto permite que uma aplicação **P2P** apresente baixo custo e grande escalabilidade [70].

### 3.2.3 Comparação Cliente/Servidor e **P2P**

#### **P2P**

O **P2P** representa a partilha de recursos e serviços pela troca direta entre sistemas. Entre estes recursos inclui-se a partilha de informação, ciclos de processamento, armazenamento de cache ou de ficheiros em disco. Numa arquitetura **P2P**, dispositivos tradicionalmente utilizados como clientes, comunicam directamente entre si, assumindo-se como clientes e servidores, desempenhando o papel (*role*) mais eficiente para a rede [71]. De seguida

pretende-se apresentar vantagens e desvantagens associadas a este tipo de arquitetura:

### Vantagens

- Inexistência de um único ponto de falha, ou seja, se um nó falhar, os restantes continuam a poder comunicar;
- Capacidade de utilização de recursos, e.g poder de processamento ou capacidade de armazenamento;
- Evitar *bottleneck* registado por sobrecarga de tráfego, uma vez que uma arquitetura P2P consegue distribuir informação e equilibrar solicitações sem utilização de um servidor central;
- Escalabilidade facilitada face à inexistência de um controlador central e porque os nós interagem entre si.

### Desvantagens

- Soluções P2P não permitem satisfazer as necessidades de aplicações com padrões de segurança elevados;
- Ligação entre nós não são implementadas para elevadas taxas de transferência;
- Um sistema centralizado mantém as suas funções desde que o servidor as mantenha ativas. Numa arquitetura P2P, se os *peers* começam a abandonar o sistema, os serviços deixam de estar disponíveis;
- Mecanismos de pesquisa funcionam melhor quando o podem fazer a uma base de dados central.

### Cliente/Servidor

Numa arquitetura cliente-servidor, cada dispositivo ou processo na rede é um cliente ou servidor, mas não ambos. Servidores são geralmente poderosos

computadores ou processos dedicados à gestão de drivers de discos, impressoras ou tráfego de rede. Clientes são por sua vez dispositivos pessoais ou estações de trabalho, no qual clientes utilizam aplicações. Estes dependem de servidores na obtenção de recursos como ficheiros, acesso a dispositivos e também poder de processamento [71]. De seguida pretende-se apresentar vantagens e desvantagens associadas a este tipo de arquitetura:

### Vantagens

- Gestão de dados facilitada, uma vez que os ficheiros se encontram presentes num único local. Desta forma é possível realizar de forma eficiente o processo de recuperação de dados e gestão de erros;
- O hardware do servidor é implementado para responder aos pedidos dos clientes de forma rápida. Dados são processados no servidor e apenas o resultado é retornado para o cliente. Desta forma é reduzido o tráfego entre cliente e servidor, melhorando o desempenho em rede;

### Desvantagens

- Sistemas cliente-servidor são geralmente mais caros e necessitam de maior manutenção;
- O servidor constitui um único ponto de falha. Se o servidor encontrar um problema é provável que todo o sistema sofra um grande atraso ou mesmo falhe por completo, afectando clientes a si associados;

## 3.3 Base de Dados

### 3.3.1 Relacionais

A partilha de informação é uma prática fundamental na sociedade face à vasta quantidade de dados continuamente adquiridos. Tal como apresentado anteriormente, a aquisição, processamento e análise de imagens de MRI en-

volve um conjunto vastíssimo de processos, ferramentas e informação gerada, que por sua vez deverá ser de fácil acesso e capacidade de utilização.

Neste sentido, a utilização de uma base de dados é essencial na organização de toda a estrutura de estudo (geralmente distribuída), cuja escala poderá variar entre algumas máquinas num único laboratório, até diferentes departamentos ou mesmo instituições.

No que a base de dados diz respeito, as [Bases de Dados Relacionais \(BDR\)](#) são ainda a tecnologia predominante para armazenamento de informação estruturada na web e em aplicações, e.g. Oracle, MySQL ou Microsoft SQL Servers. Desde o artigo de Codd “A relational model of data for large shared data banks” [72], que bases de dados com suporte de cálculo relacional e fornecendo implementações de pesquisa por [Strucutred Query Language \(SQL\)](#) foram largamente adoptadas e são actualmente consideradas a solução de armazenamento mais consistente por alguns clientes.

No modelo relacional introduzido por Codd, a informação é estruturada em tabelas constituídas por linhas e colunas. A cada linha constitui um único registo de atributos, organizados em colunas contendo elementos “de e só do mesmo tipo”, de acordo com as regras definidas para essa mesma coluna. O sucesso deste tipo de base de dados advém de características próprias, das quais se destacam [73]:

- **Simplicidade:** Informação organizada naturalmente segundo o modelo relacional, simplificando o desenvolvimento e uso da base de dados;
- **Facilidade na recuperação de informação:** Resultados podem ser filtrados segundo o conteúdo de cada coluna, e em qualquer número de colunas, permitindo aos utilizadores recuperar informação de interesse;
- **Integridade:** característica essencial em bases de dados relacionais. Integridade referencial entre tabelas previne registos incompletos ou “órfãos”, assegurando precisão e consistência dos dados;
- **Normalização:** metodologia sistemática existente para assegurar que a base de dados relacional é livre de anomalias que possam afectar a

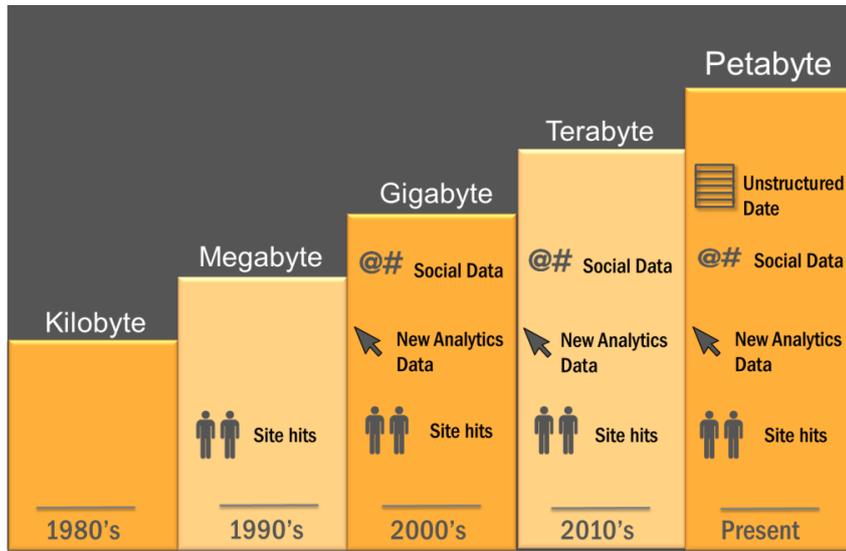


Figura 3.4: Evolução do volume e tipo de dados gerados desde os anos 80 até ao presente (retirado de [75])

integridade e exatidão da mesma;

Apesar das propriedades descritas constituírem vantagens de BDR, desvantagens deste modelo podem também ser apresentadas. O volume de dados gerados actualmente (surgimento do termo Big Data) superam em larga escala os valores obtidos há alguns anos, pelo que a escalabilidade da base de dados é um aspecto de grande importância. Além disso, a informação a gerir é cada vez mais diversa, com inúmeros produtos multimédia (e.g. imagens 3D) a desafiar a capacidade de categorização de bases de dados relacionais [74].

Neste sentido, os principais problemas associados a este tipo de base de dados prendem-se com os seguintes aspectos: esquema fixo e rígido, o que constitui uma menor flexibilidade à base de dados; dificuldade de escalabilidade, com redução de desempenho com o aumento de informação a armazenar; consistência “desnecessária” para determinadas soluções, no qual é necessário processar grande quantidade de informação num curto período de tempo.

Apesar de existirem diferentes abordagens, tais como *Object Databases* ou armazenamento XML (Extensible Markup Language), estas tecnologias

nunca obtiveram a mesma adopção e mercado que as bases de dados relacionais. Ao invés, estas soluções foram “absorvidas” por BDR que e.g. permitem armazenamento XML e utilizam-no com propósito de indexação de texto, ou tornaram-se produtos de nichos como e.g. *Online analytical processing (OLAP)* ou processamento de fluxos [76]. No entanto, o “mundo digital” está a crescer muito rapidamente tornando-se mais complexo em volume (terabyte a petabyte), variedade (estruturado, não-estruturado e híbrido) e velocidade.

Neste sentido, nos anos recentes a hegemonia do modelo relacional foi questionada, tanto pela comunidade científica como pelas companhias web, o que despoletou o aparecimento de diversas soluções alternativas. Este movimento possui a designação de *NoSQL*.

### 3.3.2 NoSQL

*NoSQL* refere-se a um grupo eclético de base de dados não relacionais. Nestes sistemas, a base de dados não é estruturada em tabelas e geralmente não usa *SQL* para manipulação de dados. Arquiteturas *NoSQL* são distribuídas, desenvolvidas para armazenamento de dados em grande escala (estruturados e desestruturados) e para processamento paralelo massivo ao longo de vários *commodity servers*. Utilizam linguagens e mecanismos não *SQL* para interação com dados (embora algumas novas APIs convertam queries *SQL* para a linguagem de consulta nativa do sistema ou ferramenta) [77].

Estas bases de dados encontram-se presentes nas maiores companhias Web como Google, Amazon e Facebook, que encontraram desafios em lidar com grande quantidade de informação recorrendo às soluções de BDR tradicionais. Estas suportam múltiplas actividades incluindo análise exploratória e predicativa, *Extract Transform Load (ETL)* e *Online transaction processing (OLTP)* não crítico. Originalmente motivados por aplicações Web 2.0, estes sistemas são desenvolvidos para escalar milhares ou milhões de utilizadores realizando actualizações (update) e leituras (reads), em oposição a *Relational Database Management System (RDMS)* tradicionais e *data warehouses* [78]

Um exemplo típico de aplicação que pode ser um desafio para BDR tradicionais é um blog web. Este pode ser constituído por diferentes tipos de

atributos: texto, comentários, tags, imagens, vídeos, código fonte, entre outro tipo de informação armazenada em múltiplas tabelas. Uma vez que este tipo de aplicação web é ágil, a correspondente base de dados deverá ser flexível por forma a suportar uma fácil avaliação do esquema [79]. No entanto, a adição ou remoção de uma característica do blog não é possível sem a indisponibilidade do sistema caso uma BDR seja utilizada. Os sistemas NoSQL por outro lado, exibem a capacidade de armazenar e indexar arbitrariamente grande conjunto de dados enquanto permite o acesso concorrente de inúmeros utilizadores [80].

Por forma a garantir a integridade da informação, grande porção das bases de dados clássicas baseiam-se em transações. Desta forma, é assegurada a consistência dos dados em todas as situações de gestão dos mesmos. Estas características transaccionais são designadas por *Atomicity, Consistency, Isolation e Durability (ACID)*, emergindo como conceitos unificadores para computação distribuída [81]:

- **Atomicidade:** Uma transição diz-se atómica, se “parecer” passar de um estado inicial para o correspondente resultado sem nenhum estado intermédio observável, ou se parecer nunca ter saído do estado inicial (princípio do “tudo ou nada”);
- **Consistência:** A transação produz apenas resultados consistentes, caso contrário aborta. Uma vez que é impossível verificar todas as restrições cada vez que uma transação é iniciada, é assumido que os dados são consistentes no estado inicial da base de dados ou se tiverem sido produzidos por uma transação confirmada (committed);
- **Isolamento:** Esta propriedade sugere que um programa em execução num ambiente multi-utilizador, se deve comportar exatamente como num ambiente de utilizador único;
- **Durabilidade:** Requer que os resultados de transações completadas com sucesso não devem ser esquecidos pelo sistema.

No entanto, escalar sistemas complacentes com as propriedades ACID tem demonstrado ser um problema. Os diferentes aspectos de alta disponibilidade

necessária em sistemas distribuídos, não são totalmente solucionáveis, com o Professor Eric Brewer a avançar com o teorema de [Consistency, Availability, Partition-Tolerance \(CAP\)](#) em 2000 [82]:

- **Consistência (*Consistency*)**: todos os clientes vêem a mesma versão dos dados, mesmo em actualizações do *dataset*;
- **Disponibilidade (*Availability*)**: todos os clientes conseguem sempre pelo menos uma cópia da informação requisitada, mesmo que algumas máquinas do cluster estejam em baixo;
- **Tolerância de partição (*Partition Tolerance*)**: o sistema completo mantém as suas propriedades mesmo quando implementado em diferentes servidores, com transparência para o cliente.

A ideia principal deste teorema, visualizado na Figura 3.5 é a de que um sistema distribuído não pode cumprir com os três pressupostos apresentados ao mesmo tempo, mas apenas dois. De acordo com o teorema de [CAP](#) e tendo em conta diferentes abordagens de sistemas [NoSQL](#), uma classificação preliminar destes sistemas pode ser realizada:

- **Orientados a consistência e disponibilidade (CD)**: parte da base de dados não se “preocupa” com tolerância de partição, e geralmente utiliza uma abordagem de replicação para assegurar a consistência e disponibilidade;
- **Orientados a consistência e tolerância de partição (CT)**: tal sistema de base de dados armazena a informação em nós distribuídos, mas assegura inclusivamente a consistência dos mesmos, com perda na capacidade de disponibilidade;
- **Orientados a disponibilidade e tolerância de partição (DP)**: estes sistemas obtêm uma “consistência eventual” através de replicação e verificação.

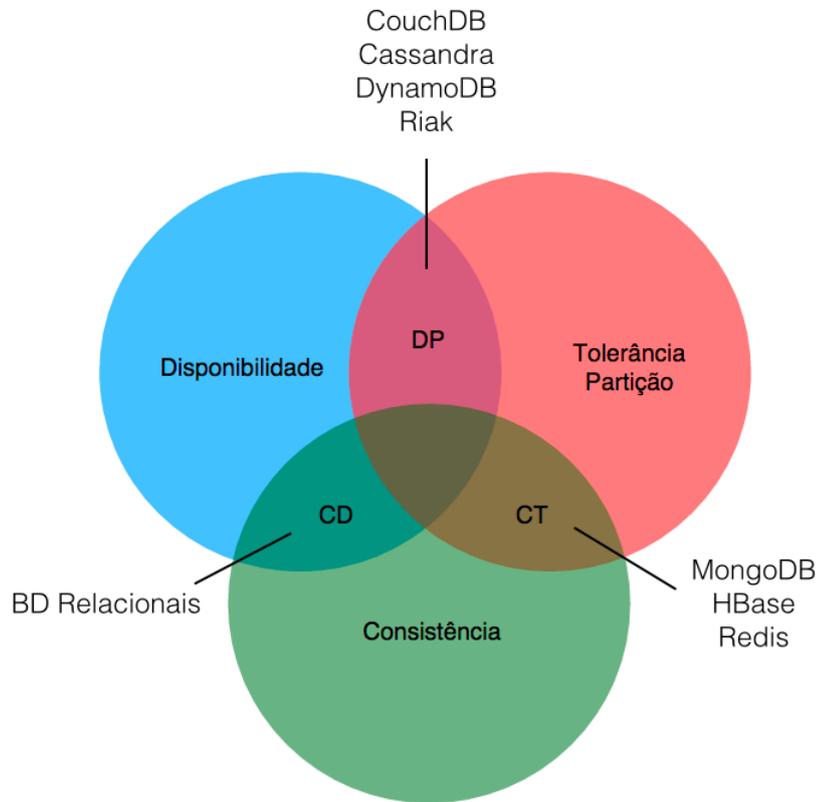


Figura 3.5: Teorema de CAP (adaptado de [83] )

Como resultado das exigências de mercado, inúmeras bases de dados NoSQL desprenderam-se essencialmente de requisitos sobre consistência e isolamento, a fim de conseguir maior disponibilidade, partição de dados, desempenho e “degradação suave”.

Assim sendo, desenvolveu-se um sistema designado por *Basically Available, Soft-state, Eventually consistent (BASE)* [84], cujas principais características são de seguida apresentadas:

- *Essencialmente disponível (Basically Available)*: O sistema garante um certo nível de disponibilidade para os dados, mesmo no que diz respeito a falhas nos nós do sistema. Os dados podem ser obsoletos, mas o sistema irá responder e aceitar respostas;
- *Soft-State*: Os dados encontram-se num constante estado de fluxo, pelo

que enquanto uma resposta pode ser dada, a versão e consistência dos dados não é garantidamente a mais actual;

- Eventualmente Consistente (*Eventually consistent*): Os dados serão eventualmente consistentes ao longo de todos os nós das bases de dados, mas não em todas as transações a cada instante.

Ippolito resume as propriedades **BASE** da seguinte forma [85]: uma aplicação executa basicamente todo o tempo, não precisando de ser consistente em todos os momentos, alcançando eventualmente um estado “conhecido”.

De acordo com as propriedades deste tipo de sistema, podem ser apresentadas as aplicações padrão de bases de dados **NoSQL**: processamento de dados em grande escala (processamento paralelo em sistemas distribuídos); análise exploratória de dados semi estruturados; armazenamento de grande volume de dados (sem estrutura, semi estruturados ou estruturados).

Neste sentido, providenciam soluções relativamente económicas, com alta escalabilidade para armazenamento de grandes volumes de dados e.g. logs, registo de dados de chamadas telefónicas, leitura de medições e armazenamento de dados flexíveis. Os seus frameworks distribuídos tornam estas implementações ideais para processamento massivo de dados (agregação, filtragem ou ordenação). Além disto, são também indicadas para recuperação e substituição de dados máquina a máquina e processamento de grande volume de transações, uma vez que as restrições **ACID** são “relaxadas”, ou pelo menos forçadas ao nível da aplicação ao invés de o fazer no sistema de base de dados.

Em [74] a classificação de bases de dados **NoSQL** foi diferenciada em três tipos: armazenamento Chave-Valor (*Key-value stores*), e.g. SimpleDB; orientado a colunas (*Column-oriented*) e.g. Cassandra, HBase ou a Google Big Table [86]; armazenamento baseado em documentos, e.g. CouchDB ou MongoDB.

Por outro lado, esta divisão pode ser alargada a quatro categorias, cada uma apropriada para diferente tipo de tarefa: (1) armazenamento Chave-Valor; (2) bases de dados documentais; (3) armazenamento por família de colunas; (4) bases de dados de grafos.

- **Armazenamento Chave-Valor** - Neste tipo de sistema de base de dados, os itens são tipicamente armazenados como identificadores alfanuméricos (chaves) e valores associados em tabelas simples (referidas como tabelas de hash). Por sua vez, os valores podem ser simples *string* textuais ou listas e conjuntos (*sets*) mais complexos. A pesquisa de dados é geralmente realizada unicamente nas chaves, não em valores, e é limitada a valores exatos [87].

A simplicidade deste tipo de armazenamento chave-valor, torna-o idealmente apropriado para extração de valores extremamente rápido e escalável, necessário em tarefas como a gestão do perfil de utilizadores, sessões ou obtenção do nome de produtos. Neste sentido, a Amazon faz uso extensivo do seu próprio sistema Chave-Valor, Dynamo, para e.g. gerir o “carrinho de compras”.

O Dynamo é um sistema de armazenamento chave-valor altamente disponível, utilizado por serviços essenciais da Amazon, para providenciar armazenamento de dados escalável, distribuído e acessível [88]. Alguns exemplos deste tipo de sistema são Dynamo (Amazon); Voldemort (LinkedIn); Redis; BerkeleyDB ou Riak.

- **Bases de dados Documentais** - Tal como o próprio nome indica, estas bases de dados são desenvolvidas para gerir e armazenar documentos. Estes documentos são codificados num formato de dados normalizado como e.g. [XML](#), [Javascript Option Notation \(JSON\)](#) ou [Binary JSON \(BSON\)](#). Ao contrário do armazenamento chave-valor acima descrito, a coluna de valor em bases de dados documentais contém dados semi-estruturados, especificamente pares de atributos nome/valor. Uma única coluna pode conter milhares desses atributos, e o número e tipo de atributos registados pode variar de linha para linha. Além disso, tanto a chave como valor são completamente “pesquisáveis” neste tipo de base de dados [89].

Bases de dados documentais são apropriadas para armazenamento e manutenção de coleções de grande tamanho (*Big Data-size*), tais como documentos de texto, imagens, mensagens de email, [XML](#), além de

documentos como a representação desnormalizada de uma entidade da base de dados, e.g. um produto ou consumidor. São também adequadas para armazenar dados "irregulares" (semi-estruturados) que requerem a extensa utilização de valores nulos (espaços reservados cujo valor estaria em falta ou é inexistente) em BDR. Alguns exemplos deste tipo de sistema são: CouchDB, MongoDB ou RavenDB.

- **Armazenamento por Família de Colunas** (implementações *Big-Table*) - Tal como as bases de dados documentais, estas bases de dados empregam uma estrutura de dados distribuída e orientada a colunas que adapta múltiplos atributos por chave. Enquanto alguns destes sistemas possui um "ADN" chave-valor (e.g. a base de dados inspirada em Dynamo, Cassandra), a grande parte é padronizada após a Google's Bigtable, o sistema de armazenamento interno distribuído à escala de petabyte, desenvolvido para o seu índice de pesquisa e outras coleções como o Google Earth e Google Finance [90].

Estes sistemas replicaram não só a estrutura de armazenamento de dados do Google's Bigtable, mas também o sistema distribuído de ficheiros da Google (*GFS*), bem como a framework de processamento paralelo MapReduce, tal como é o caso de Hadoop, que compreende o sistema de ficheiros Hadoop (*HDFS*, aliado a Hbase (sistema de armazenamento de estilo *BigTable*) e MapReduce.

Este tipo de sistema é relevante para: armazenamento distribuído de dados, especialmente dados de versão devido às funções *time-stamping*; processamento de dados batch-oriented em larga escala como e.g. classificação, análise, conversão de dados ou análise exploratória e preditiva realizada por estatísticos e programadores experientes.

Alguns exemplos deste tipo de sistema são: Bigtable (Google); Hyper-table; Cassandra (Facebook); SimpleDB (Amazon).

- **Graph Databases** - As bases de dados de grafos substituem tabelas relacionais por grafos relacionais estruturados de pares chave-valor interconectados. São similares a bases de dados orientadas a objetos

uma vez que os grafos são representados como: uma rede de nós orientada a objetos (objetos conceptuais), relações entre nós e propriedades (atributos dos objetos expressa por pares chave-valor). São por sua vez, o único tipo de base de dados **NoSQL** aqui discutido que se “preocupa” com relações, e o seu foco na representação visual de informação torna-as mais “human-friendly” que os restantes sistemas **NoSQL** [89].

No geral, bases de dados de grafos são úteis quando o interesse recai nas relações entre dados, mais do que nos próprios dados em si. Assim sendo, alguns exemplos de utilização são: representar e percorrer redes sociais, produção de recomendações ou condução de investigações forenses e.g. detecção de padrões. De realçar que estes sistemas são optimizados para *transversing* e não para consulta (querying). Neste sentido, se for necessário explorar relações mas também consultar e analisar os valores incorporados nelas (e/ou ser capaz de utilizar consultas de linguagem natural para analisar relações), então a utilização de um sistema baseado em procura é aconselhado. Alguns exemplos deste tipo de base de dados são: Neo4j; InfoGrid; Sones GraphDB; AllegroGraph; InfiniteGraph

# Capítulo 4

## BrainArchive

### 4.1 Introdução

O cérebro humano é o sistema mais complexo alguma vez criado. Neste sentido, neurocientistas conjuntamente com investigadores de diversas áreas estudam o funcionamento cerebral. Nos anos recentes, inúmeros projetos e parcerias resultaram em descobertas notáveis, desde a compreensão de funções cerebrais e seu relacionamento, bem como de doenças como alzheimer e esquizofrenia.

Técnicas como **fMRI** ou **DTI** representam uma parte essencial na pesquisa actualmente realizada. Estas tecnologias geram enormes quantidades de informação, levando à necessidade de a organizar de forma eficiente e acessível de forma intuitiva e segura. Assim, é possível diminuir o tempo consumido nestas tarefas pelas entidades que dela fazem uso, o que por sua vez acarreta uma melhoria na investigação realizada pois permite ocupar mais tempo com a prática da técnica em si.

A forma como esta informação é gerida e partilhada entre profissionais e entidades é um problema, sendo muitas vezes incoerente e de difícil acesso. A falta de padronização de fluxos de processamento bem como a variedade de ferramentas a utilizar na aplicação destas técnicas é um obstáculo à organização da informação gerada. Neste sentido, a capacidade de armazenamento eficiente de dados nas mais diversas fases de estudo é afectada, impossibi-

litando uma correcta partilha de informação e, por sua vez, a obtenção da mesma de forma intuitiva pelos profissionais envolvidos.

Este cenário justifica a criação de um sistema que permita colmatar os problemas apresentados, através do desenvolvimento de uma plataforma de armazenamento e partilha de informação cerebral, tendo por base uma estrutura e fluxo de processamento padronizado.

## 4.2 Análise de Requisitos

O sistema desenvolvido no ambiente deste trabalho apresenta requisitos necessários de consideração e contexto. Assim é possível compreender a motivação de cada requisito, o que permite encontrar soluções apropriadas para cada característica do sistema, que possam levar à obtenção de resultados concordantes com o pretendido. O principal objetivo visa a concepção e desenvolvimento de um protótipo baseado numa arquitetura orientada ao serviço para armazenamento, estruturação, partilha e obtenção de estudos de Ressonância Magnética Cerebral.

As soluções informáticas permitem melhorar a performance de serviços hospitalares, laboratórios, universidades, entre outros na realização das mais diversas tarefas, sendo prática comum em muitas instituições de prestação de cuidados de saúde e centros de investigação. Actualmente observa-se um crescente avanço tecnológico e resultante proliferação dos sistemas de informação disponíveis, com efeito positivo na qualidade dos cuidados de saúde prestados e pesquisa. Esta relação direta de qualidade resulta num maior investimento no desenvolvimento de soluções informáticas, por forma a colmatar carências específicas nas mais diversas vertentes desde a prestação de cuidados de saúde a investigação como é o caso do complexo sistema cerebral.

Uma vez que este sistema visa ser integrado em ambiente de investigação existem várias especificidades a garantir. Um aspecto essencial é a garantia de segurança de acesso a dados por parte dos utilizadores do sistema. Este deverá permitir a correcta estruturação e visualização de estudos, importação e exportação de ficheiros (desde um único ficheiro a um estudo contendo

inúmeros sujeitos e modalidades de análise), além de permitir configurar permissões de acesso a determinado ficheiro(s) ou estudo. Sendo necessário disponibilizar acesso e partilha de grande quantidade de dados, a partir de diversas localizações e plataformas, é essencial o desenvolvimento de um repositório, seguindo uma abordagem de computação em nuvem. Com o sistema implementado deseja-se otimizar a utilização do fluxo de processamento proposto na [Seção 2.6](#), facilitando a organização e partilha de estudos mediante a apresentação de uma interface gráfica. Tendo em conta estes objectivos, a aplicação deverá:

- ser uma aplicação *plug-and-play*, capaz de ser utilizada em diversos sistemas operativos;
- permitir a estruturação de dados de estudos de [MRI](#) nas diversas etapas de um fluxo de processamento, permitindo localizar e identificar facilmente os resultados de cada etapa;
- possuir uma estrutura ajustada a fluxos de processamento existentes, facilitando assim a sua integração em ambiente real;
- permitir a partilha e acesso a dados de forma segura e transparente;
- apresentar uma interface amigável, limpa e simples, minimizando a necessidade de interação com o utilizador;
- permitir a gestão de permissões de acesso a informação presente no repositório;
- disponibilizar um mecanismo simples e eficiente de obtenção de dados, segundo a especificação do utilizador.

### 4.3 Base Tecnológica

Para o desenvolvimento de uma aplicação informática impõe-se definir os requisitos necessários a satisfazer por parte da mesma. O software utilizado em fluxos de processamento como o apresentado na [Seção 2.6](#), é geralmente

fornecido para as plataformas Linux ou MacOS X, pelo que estas plataformas se tornam as mais relevantes no desenvolvimento da aplicação apresentada. Existem também alternativas para a plataforma Windows que embora utilizadas em menor número continuam a ser interessantes.

Neste sentido, pretende-se que o acesso à aplicação seja transversal a vários sistemas operativos. A escolha da linguagem de programação a utilizar no desenvolvimento de uma aplicação requer o estudo de soluções e tecnologias apropriadas disponíveis. Tendo por base os requisitos apresentados e características do sistema a implementar optou-se por desenvolver a aplicação em causa em Python. Esta linguagem de programação apresenta benefícios face a soluções como Java ou C, entre os quais se destaca:

- desenvolvimento multi-plataforma escalável e estável;
- sintaxe simples - embora o uso de indentação para estabelecimento da estrutura do programa possa ser um entrave inicial para novos utilizadores de Python, a redução de estruturas sintáticas como '{}' permite reduzir o esforço de escrita e aumentar a produtividade de desenvolvimento;
- escrita de alto nível - desprendimento de um *data typing* "pesado" como evidenciado em linguagens como C ou Java, o que conduz simultaneamente ao corte considerável de volume de código de suporte e à escrita simplificada de código flexível e introspectável;
- estruturas de dados poderosas e de fácil implementação - listas ou dicionários implementados em Python podem ser utilizados em conjunto para desenvolvimento de estruturas de dados em curto período de execução;
- biblioteca padrão extensa - Python possui uma vasta biblioteca de suporte desde expressões regulares a *xml parsing*, ferramentas de serviços web a módulos de processamento de imagem e acesso a base de dados.

Uma vez escolhida a linguagem de programação, é também necessário escolher o ambiente de desenvolvimento a utilizar ([Integrated Development](#)

Environment (IDE)), assim como as restantes tecnologias incorporadas. Assim sendo, adoptou-se o IDE Eclipse por ser uma ferramenta livre multi-plataforma, baseada num *workspace* e *plugins* extensíveis para personalização do ambiente de programação. Desenvolvido essencialmente em Java, pode ser utilizado no desenvolvimento de aplicações em linguagens de programação como C++, JavaScript, Python, PHP ou Scala.

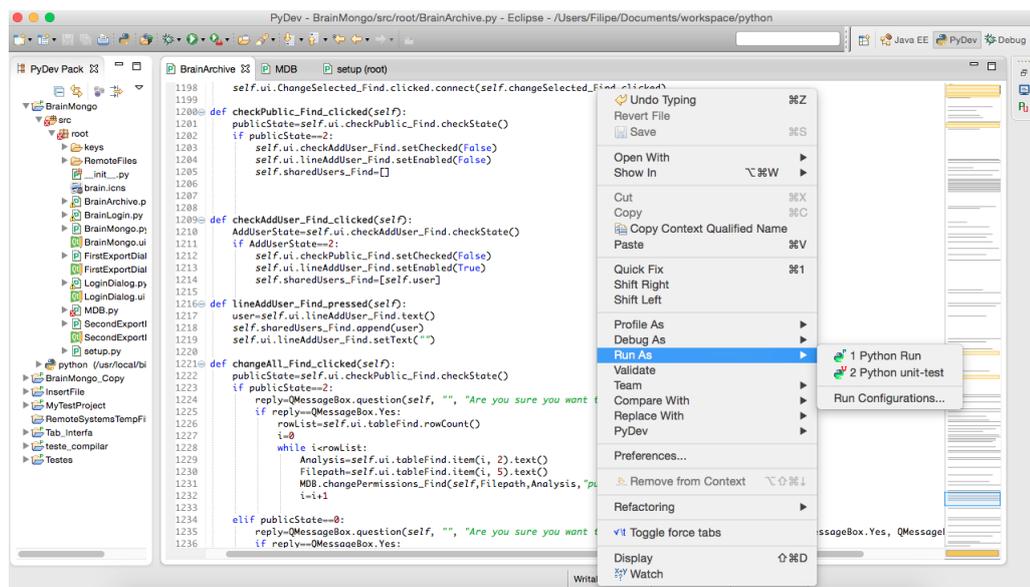


Figura 4.1: Eclipse Kepler Service Release 1 na perspectiva PyDev

No desenvolvimento em Python no IDE Eclipse optou-se por utilizar o *plugin* denominado PyDev. Além de permitir o desenvolvimento de aplicações em Python suporta ainda refaturação do código e *debugging* gráfico. Um exemplo de utilização deste IDE é apresentado na Figura 4.1

Uma vez que se pretende desenvolver uma aplicação com interface gráfica (Graphical User Interface (GUI)) torna-se necessário escolher a ferramenta a utilizar para a sua criação. O desenvolvimento da composição visual e comportamento da GUI é essencial na interação humano-computador. Tipicamente, um utilizador interage com informação através da manipulação de *widgets* visuais que permitem uma correta interação com o tipo de dados que suportam. Para o desenvolvimento da GUI optou-se pela combinação de PyQt e Qt Designer. PyQt é um software livre, implementado como plu-

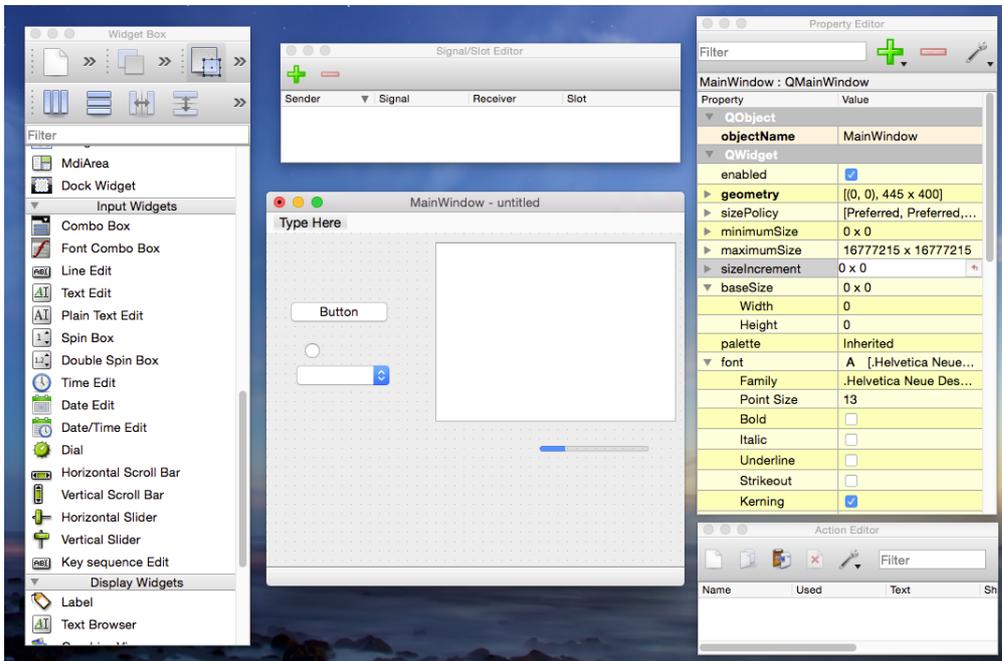


Figura 4.2: Interface Qt Designer 5.4.1

gin de ligação de Python ao conjunto de ferramentas multi-plataforma Qt. O Qt é por sua vez uma framework de desenvolvimento de aplicações, que podem ser executadas em várias plataformas de software e hardware através de alterações mínimas no código subjacente. O PyQt é uma das opções mais utilizadas na programação de GUI, rivalizando com alternativas como PySide, PyGTK, wxPython e TKinter. Na Figura 4.2 é apresentada a interface amigável da ferramenta Qt Designer.

Desta forma, a personalização de janelas e diálogos é possível recorrendo a uma abordagem *drag-and-drop*, com possibilidade de teste em diferentes estilos e resoluções. Os Widgets utilizados são facilmente integrados com código, através de um mecanismo de sinais, que permite a atribuição de comportamentos a elementos gráficos.

Após a escolha de tecnologias de suporte ao desenvolvimento da aplicação em Python, importa definir a infraestrutura responsável pela manutenção do repositório de estudos. Como evidenciado na [Capítulo 3](#), são várias as bases de dados actualmente disponíveis, apresentado entre elas diferenças

profundas de implementação. Por forma a melhor optar por determinada solução, importa definir objetivos essenciais do repositório a desenvolver:

- permitir o armazenamento de grande quantidade de dados;
- facilitar o processo de obtenção de informação através de metadados associados;
- possibilitar a partilha de dados proveniente de diversas fontes, formato e tamanho;
- permitir mecanismos de autenticação e segurança no acesso à mesma;
- utilizar uma solução livre comprovada, confiável e com boa documentação;
- bibliotecas de comunicação poderosas, estáveis e com comunidade e documentação aprovada.

Assim sendo e tendo em conta a análise de tecnologias realizada na [Seção 3.3](#) optou-se por utilizar o sistema de base de dados [NoSQL MongoDB](#). Desenvolvido pelo projeto 10gen, MongoDB é uma base de dados *open-source* orientada a documentos, de alta performance e sem esquema rígido. MongoDB armazena informação como documentos numa representação binária designada [BSON](#). Documentos que partilham uma estrutura similar são geralmente organizados em coleções. Estas podem ser analogamente associadas a tabelas numa base de dados relacional (no qual documentos se assemelham a linhas (rows) e campos (fields) a colunas).

A informação para cada registo tende a ser armazenada num único documento, enquanto que numa base de dados relacional, os dados são geralmente distribuídos em várias tabelas. Um exemplo desta diferenciação no modelo de dados é observável numa aplicação de *blogging*. Numa base de dados relacional, o modelo de dados compreende várias tabelas como Categoria, Tags, Utilizadores, Comentários e Artigos. Por outro lado, utilizando MongoDB, a informação pode ser modelada contendo apenas 2 coleções: Utilizadores e

Artigos. Desta forma cada documento de blog contendo múltiplos comentários, tags e categorias podem ser expressas através de um array incorporado. Um exemplo desta implementação pode ser verificado na Figura 4.3.

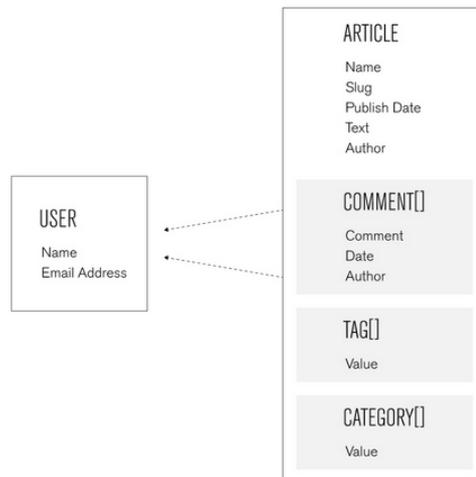


Figura 4.3: Modelo de dados documental representativo de uma aplicação de *blogging* em MongoDB. (retirado de [91])

Como resultado do modelo documental, os dados em MongoDB são mais localizados, reduzindo drasticamente a necessidade de unir (JOIN) tabelas separadas. Desta forma, a performance e escalabilidade da aplicação é aumentada, uma vez que apenas um processo de leitura (read) permite obter todo o documento.

Além disso, documentos MongoDB aproximam-se da estrutura de objetos em linguagens de programação. Por conseguinte, é mais fácil para o programador modelar a forma como os dados na aplicação mapeiam a informação presente na base de dados. MongoDB apresenta um esquema dinâmico, no qual campos podem variar de documento para documento. Desta forma, não existe a necessidade de declaração da estrutura de documentos ao sistema, uma vez que estes são autodescritivos. Caso seja necessário introduzir um novo campo a um documento, este pode ser adicionado sem que afecte outros documentos, sem que obrigue a actualização de um catálogo central e sem tornar o sistema offline.

A variedade de *queries* que MongoDB suporta para operações altamente escaláveis e com potencial analítico é também interessante, podendo estas retornar um documento ou um subconjunto de campos de um documento. Contrariamente a outras abordagens NoSQL, MongoDB não se limita a operações chave-valor. Torna-se assim possível desenvolver poderosas aplicações utilizando *queries* complexas em dados estruturais, semi-estruturais ou desestruturados. Além disso, o mecanismo de indexação existente em MongoDB é essencial na optimização da escalabilidade e desempenho do sistema, existindo vários tipos de índice (e.g. *unique*, *geospatial* ou *Time-to-Live (TTL)*) possíveis de ser aplicados a qualquer campo, incluindo campos em arrays.

Como referido anteriormente, é necessário que a solução escolhida permita integração com linguagens de programação. Neste sentido, MongoDB providencia drivers nativos para populares linguagens de programação, e.g. Java, Ruby, .NET, Python, Perl, Scala, PHP, entre outras.

Documentos em MongoDB apresentam um limite de 16MB. Uma vez que se pretende desenvolver um repositório contendo ficheiros de tamanho e formato variável, esta restrição poderia excluir MongoDB como uma solução viável. No entanto, existe em MongoDB uma especificação denominada GridFS (suportada por todos os drivers oficiais MongoDB) para armazenamento e recuperação de ficheiros que excedam este limite. Para tal, GridFS divide um ficheiro em partes, ou *chunks*, armazenando cada uma destas partes como um documento separado.

Por defeito, GridFS limita o tamanho do *chunk* a 255k, utilizando duas coleções (*.files* e *.chunks*) para o armazenamento dos ficheiros. Desta forma, uma coleção (*chunks*) armazena os diversos chunks e a outra armazena os metadados associados (*files*). Na Figura 4.4 é apresentada a divisão em duas coleções (*files* e *chunks*) efectuada pela API GridFS.

Quando se procura em GridFS por um ficheiro, o driver ou cliente reagrupa quantos chunks for necessário, sendo também possível realizar *range queries*, acedendo a informação em secções arbitrárias de ficheiros. GridFS é também relevante quando se pretende armazenar ficheiros, aos quais se pretende acesso sem carregamento do arquivo inteiro em memória.

MongoDB apresenta características como Replicação e *Sharding* relativas

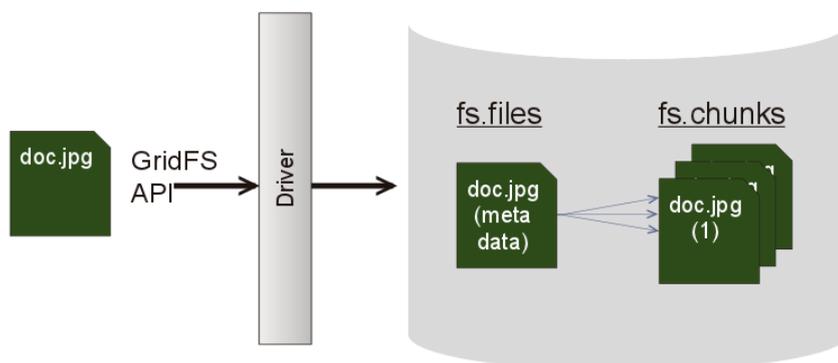


Figura 4.4: Representação da divisão em duas coleções com GridFS (retirado de [92])

à performance e escalabilidade de grandes volumes de dados. *Sharding* é uma técnica que permite escalar a base de dados horizontalmente, dividindo os dados em *shards* que correspondem a partições físicas que se podem encontrar em diferentes localizações. Este processo é ilustrado na Figura 4.5.

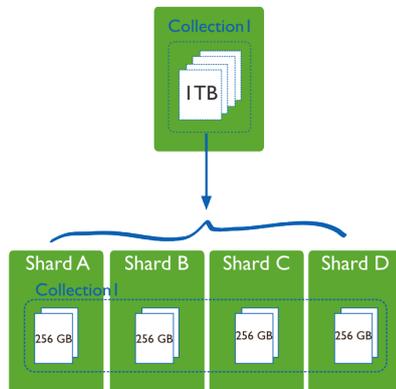


Figura 4.5: MongoDB *Sharded Collection* retirado de [91]

Esta técnica permite acesso transparente a dados distribuídos, sem adicionar complexidade à aplicação. Os sistemas MongoDB utilizam um *Sharded Cluster* constituído por três componentes, como evidenciado na Figura 4.6

Os três componentes designam-se: *Shards*, que são os componentes no qual os dados são armazenados, os *Query Routers* que são as interfaces com aplicações cliente, direcionando as operações para os *shards* respectivos e

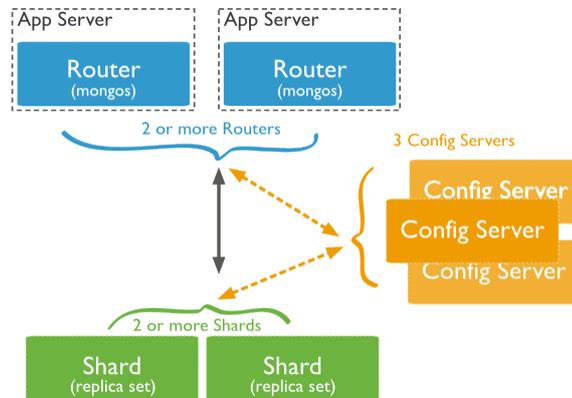


Figura 4.6: MongoDB *Sharded Cluster* retirado de [91]

devolvendo uma resposta ao cliente e os *Config Servers* que guardam metadados do *cluster* para mapeamento dos dados do *cluster* ao respectivo *shard* (*query routers* utilizam esta informação para direcionamento de pedidos). Para *Sharding* de uma coleção é necessário definir uma *shard key*. Uma *shard key* é um campo indexado existente em cada documento na coleção. Desta forma, MongoDB divide os valores da *shard key* em *chunks*, distribuindo-os uniformemente ao longo dos *shards* existentes.

Para providenciar grande disponibilidade, consistência dos dados e tolerância de falhas em *sharded clusters* de produção, cada *shard* é geralmente um *replica set*. Um *replica set* é então um conjunto de réplicas, em cuja configuração existe uma réplica principal e outras secundárias. Cada réplica de um *replica set* contém os mesmos dados, o que permite que caso a réplica principal falhe, uma secundária possa assumir o seu papel. Este mecanismo pode ser observado na Figura 4.7.

A segurança e privacidade de dados é um aspecto crítico, assumindo grande relevância em entidades hospitalares e de investigação. Uma vez que se pretende desenvolver um repositório de dados sensíveis de MRI é necessário assegurar a segurança do sistema. MongoDB apresenta diferentes mecanismos de segurança e acesso a dados e.g. autenticação, autorização ou encriptação.

A base de dados MongoDB foi então o sistema escolhido para desenvolvimento do repositório de ficheiros, uma vez que cumpre com os requi-

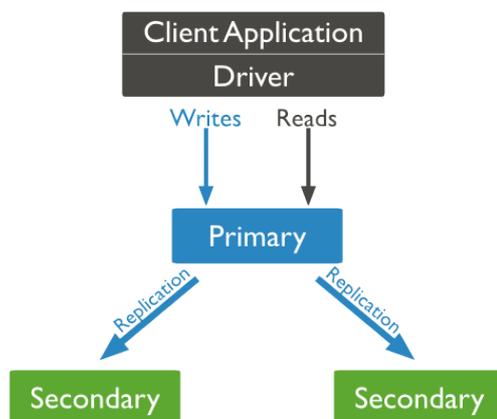


Figura 4.7: Esboço de um *replica set* MongoDB retirado de [91]

sitos apresentados. Esta solução é ideal para armazenamento de ficheiros, apresentando-se como uma solução eficaz na manutenção de grandes volumes de dados. É possível utilizar *queries* avançadas por forma a filtrar ficheiros e a especificação GridFS, permite lidar com os mesmos independentemente do tamanho que apresentem. Simultaneamente a criação de duas coleções (*.chunks* e *.files*) é extremamente útil quando se pretende dotar os ficheiros de metadados associados.

MongoDB permite escalar horizontalmente a base de dados através da técnica designada por *Sharding*. Esta permite distribuir os dados ao longo de partições físicas, por forma a ultrapassar limitações de hardware, equilibrando automaticamente os dados em *clusters*. Suporta também, como referido, a criação de replica sets, um mecanismo de controlo de falhas automaticamente gerido. Se um servidor primário falhar, um servidor secundário tornar-se-á automaticamente primário, sem intervenção humana.

Além disso, MongoDB possui mecanismos de autenticação, implementação de ligações com suporte [Secure Socket Layer \(SSL\)](#) e encriptação de dados. Possui também uma vasta documentação, comunidade e *drivers* para as principais linguagens de programação.

Detalhes de implementação da aplicação e repositório serão descritos e analisados nas próximas seções.

## 4.4 Hierarquia e Nomenclatura de Estudos

A necessidade de padronização no processamento de estudos de **MRI** estrutural, funcional e difusão é essencial por forma a maximizar a organização do grande volume de dados gerados. Neste sentido, é relevante a escolha de uma solução, que permita de alguma forma facilitar todo o fluxo de processamento descrito na [Seção 2.6](#), bem como a organização, cooperação e partilha dos estudos em investigação. Para tal optou-se pela utilização de uma hierarquia de ficheiros semelhante à de fluxos de processamento já existentes.

Uma hierarquia de ficheiros bem estruturada facilita o processo de identificação de ficheiros, a sua organização e possibilidade de partilha. Para tal é necessário o cumprimento com orientações definidas para essa mesma estrutura. Como referido anteriormente, estudos multimodais de **MRI**, envolvem geralmente vários sujeitos, diversos tipos de análise e etapas de pré-processamento, refletindo-se na produção de inúmeros ficheiros os quais importa localizar e reutilizar. Sem a utilização de uma hierarquia de estudos a tarefa de identificação de ficheiros é penosa, contra-produtiva, além de potenciar a perda de ficheiros e incorreta utilização dos mesmos.

Para que a hierarquia a adoptar possa otimizar o processo de organização é essencial compreender os parâmetros que melhor definem um ficheiro. Neste sentido propôs-se a estruturação ilustrada na [Figura 4.8](#).

Como evidenciado na [Figura 4.8](#), pretende-se que o grau de maior abrangência para um ficheiro seja o estudo a que pertence. A partir deste diferenciam-se os sujeitos incluídos no mesmo. A cada sujeito pode pertencer um conjunto de análises/aquisições distintas, cada uma delas envolvendo a utilização de diferente software e/ou ficheiros. Seguindo o fluxo de processamento apresentado na [Seção 2.6](#), para cada sujeito podem existir nove subpastas, contendo ficheiros agrupados por aquisição/análise:

- **FMRIDCM**: imagens **DICOM** resultantes de aquisição funcional;
- **MRIDCM**: imagens **DICOM** resultantes de aquisição estrutural;
- **DTIDCM**: imagens **DICOM** resultantes de aquisição de difusão;

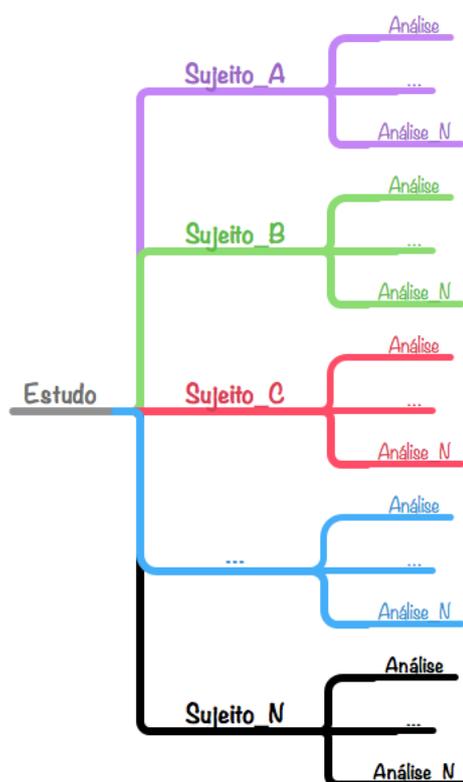


Figura 4.8: Hierarquia de estudos definida para análises multimodais

- **FMRI**: ficheiros resultantes do pré-processamento de imagens funcionais;
- **MRI**: ficheiros resultantes do pré-processamento de imagens estruturais;
- **DTI**: ficheiros resultantes do pré-processamento de imagens de difusão, obtenção de vetores e tensores;
- **TRKVIS**: ficheiros necessários à realização de tratografia determinística com o software TrackVis;
- **BedpostX.bedpostx**: ficheiros necessários à realização de tratografia probabilística com o software [BEDPOSTX](#) e [ProbtrackX](#) ;

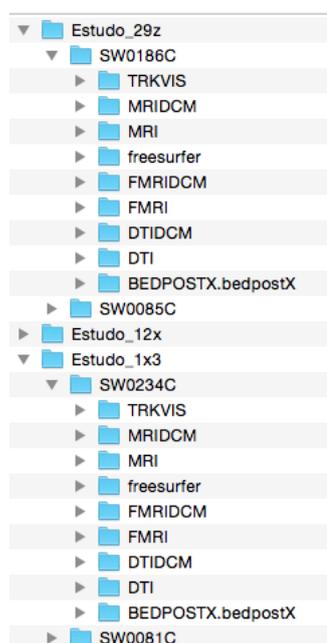


Figura 4.9: Exemplo de organização de estudos segundo uma hierarquia de pastas

- **Freesurfer** : ficheiros resultantes da utilização do software Freesurfer para segmentação de imagens estruturais.

Tendo em conta a estruturação apresentada, um exemplo de hierarquia de pastas pode ser visualizado na Figura 4.9. Neste é perceptível a existência de três estudos (*Estudo\_29z*, *Estudo\_12x* e *Estudo\_1x3*). Além disso é possível observar os sujeitos pertencentes ao estudo *Estudo\_29z* (*SW0085C* e *SW0186C*) e *Estudo\_1x3* (*SW0234C* e *SW0081C*). Da mesma forma é notória a existência das 9 subpastas representadas anteriormente para o sujeito *SW0186C* e *SW0234C*.

Após definir a hierarquia a utilizar, considera-se essencial criar para cada ficheiro, especificado na [Seção 2.6](#), uma nomenclatura específica e descrição própria. Uma vez que se pretende estender a utilização da aplicação BrainArchive a um nível de investigação global, optou-se por utilizar nomenclatura em língua inglesa, assim como será utilizado na aplicação BrainArchive.

A utilização de uma nomenclatura estritamente definida, assume um papel fundamental na organização e partilha de estudos. Através desta é pos-

sível identificar de forma automática o conteúdo de um ficheiro através da relação entre o seu "nome" e aquilo que efectivamente representa. Para tal foi definido um dicionário, apresentado na íntegra no [Apêndice B](#), e mais especificamente nas tabelas [B.1](#), [B.2](#), [B.3](#) e [B.4](#). Este representa todos os ficheiros, especificados na [Seção 2.6](#), caracterizando de forma concisa o mesmos. Assim para cada ficheiro será apresentada a nomenclatura a utilizar, a designação do mesmo e que será utilizada como metadado na aplicação BrainArchive, o software utilizado na sua produção e a *tag* correspondente ao momento de produção do ficheiro no fluxo de processamento apresentado na [Seção 2.6](#). Por forma a melhor compreender a utilidade e implementação deste dicionário, apresentam-se de seguida alguns exemplos:

- **ficheiro** : ...\_diff\_dtifit\_FA.nii.gz
  - **contexto** : mapa de anisotropia fractional (**FA**), obtido pela utilização do software **FDT** em imagens de difusão pré-processadas;
  - **ascendente** : DTI;
  - **descrição** : Fractional anisotropy map.
  
- **ficheiro** : ...\_str\_crop\_.nii.gz
  - **contexto** : ficheiro resultante da conversão das imagens **DICOM** estruturais para formato NIfTI;
  - **ascendente** : FMRI;
  - **descrição** : Raw Data.
  
- **ficheiro** : ...\_fnc\_2mni.mat
  - **contexto** : matriz de transformação para o espaço MNI na normalização de imagens funcionais;
  - **ascendente** : MRI;
  - **descrição** : Transformation matrix to MNI.
  
- **ficheiro** : ...\_diff\_dtk\_track.trk

- **contexto** : ficheiro de tratografia determinística, resultante da aplicação do software TRKVIS
- **ascendente** : TRKVIS
- **descrição** : Deterministic tractography

Na listagem anterior deverá entender-se "...", como o identificador do sujeito/caso a que pertence o ficheiro, e.g. *SW0085*. Existem no entanto algumas exceções no qual não é utilizado o nome do sujeito/caso no início do ficheiro. Estes são os ficheiros pertencentes às opções *MRIDCM*, *FMRIDCM*, *DTIDCM*, *Freesurfer* e *BedpostX.bedpostx*. As primeiras três contêm as imagens em formato **DICOM** do tipo de aquisição correspondente. *MRIDCM* contém 176 imagens **DICOM**, *FMRI* entre 100 a 1000 imagens **DICOM** e *DTIDCM* entre 30 a 260 imagens **DICOM**. Para cada uma destas imagens a designação será `IM-{"aquisição"}-{"imagem"}.dcm`. Um exemplo de imagem será então: *IM-0004-0130.dcm*, corresponderá à imagem **DICOM** número 130 da aquisição 4.

A utilização do software *Freesurfer* gera uma grande quantidade de informação, resultante do complexo processamento de dados que realiza. Desta forma são gerados vários ficheiros com nomenclatura própria, entre os quais os ficheiros considerados no dicionário *aparc.a2009s.aseg.mgz*, *aparc.aseg.mgz* e *T1.mgz*. Desta forma, optou-se por utilizar a nomenclatura automaticamente gerada pelo software *Freesurfer* no dicionário implementado. Por fim, na opção *Bedpostx.bedpostx* encontram-se os ficheiros de *input* necessários para tratografia probabilística, resultantes da aplicação do software **FSL BEDPOSTX**. Desta resultam vários ficheiros e subpastas como é possível observar na Figura 4.10.

Todos os ficheiros apresentados na Figura 4.10, que perfazem um total aproximado de 230 MB são necessários para a realização de tratografia probabilística. Neste sentido não é necessário individualizar cada ficheiro, tornando-se útil a possível partilha de toda a diretoria em causa. Assim sendo, e tal como acontece para as opções *MRIDCM*, *FMRIDCM* e *DTIDCM* é criado um ficheiro *ZIP*, que resulta da compressão de cada uma destas diretorias.

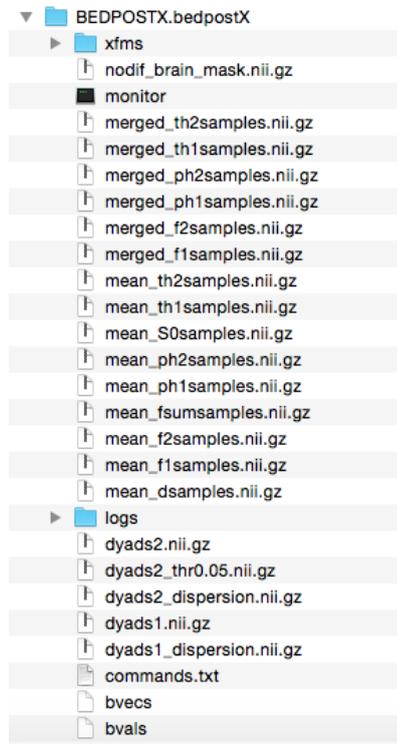


Figura 4.10: Exemplo de conteúdo da diretoria *BedpostX.bedpostx* pertencente à hierarquia de estudos.

Através da definição clara de uma hierarquia de pastas e nomenclatura de ficheiros é possível caracterizar de forma eficaz cada um dos ficheiros obtidos no fluxo de processamento descrito na [Seção 2.6](#). Desta forma a organização de ficheiros e posterior partilha de estudos entre investigadores é facilitada.

Uma vez definido este dicionário de ficheiros e respectiva estrutura de pastas, é possível a sua integração com a aplicação BrainArchive. Neste tipo de estudos/análise é muitas vezes necessário manipular os ficheiros já existentes. Como tal torna-se imperativo possuir uma cópia local dos ficheiros que compõem o estudo a ser realizado. Desta forma optou-se pela criação de uma diretoria designada *MyBrain*. Nesta diretoria, cada utilizador poderá colocar todos os ficheiros de [MRI](#) resultantes do fluxo de processamento descrito, idealmente cumprindo a hierarquia de pastas e nomenclatura descrita. Caso esta seja cumprida, a aplicação BrainArchive reconhecerá automaticamente cada ficheiro, mais especificamente, o estudo, sujeito e análise a que

pertence, bem como o tipo de ficheiro a que diz respeito. Esta diretoria permitirá o duplo propósito de adicionar automaticamente múltiplos ficheiros ao repositório, bem como criar uma cópia local de um estudo presente no repositório que não fora processado localmente, mantendo a estrutura com que foi originalmente incluído no repositório.

Neste sentido o processo de partilha de um estudo com e.g. 50 casos, múltiplas análises e ficheiros pode ser realizado na aplicação BrainArchive através de um único passo, uma vez que não será necessário ao utilizador definir nenhum parâmetro adicional. Maior detalhe acerca do paralelismo e utilização da diretoria *MyBrain* e aplicação BrainArchive será apresentado na [Seção 4.7](#).

## 4.5 Arquitetura e Implementação do Repositório

Tendo em conta a necessidade de desenvolvimento de uma interface de suporte à organização, armazenamento e partilha de estudos multimodais de Ressonância Magnética cerebrais, foi então desenvolvida uma aplicação em Python designada BrainArchive. Esta é constituída por 3 elementos chave que estabelecem correspondência com um modelo de desenvolvimento, no qual a implementação da interface seja independente do comportamento e alterações de estado dos objectos que a compõem.

Neste sentido para a criação da interface, convenientemente analisada numa seção posterior, foi utilizado o software Qt Designer. Através deste foi possível definir a janela principal bem como os restantes diálogos necessários. Da utilização deste resultam ficheiros de extensão *.iu* com estrutura XML. Posteriormente através da utilização do comando *pyuic5* é possível a conversão destes ficheiros para o formato *.py* de forma a serem utilizados pela aplicação.

Outro tipo de classe está encarregue da manipulação dos *widjets* criados na interface, assegurando a correcta utilização dos mesmos por parte do utilizador. Estas classes assumem o papel de "Controlador", estabelecendo

a comunicação entre o utilizador e a interface.

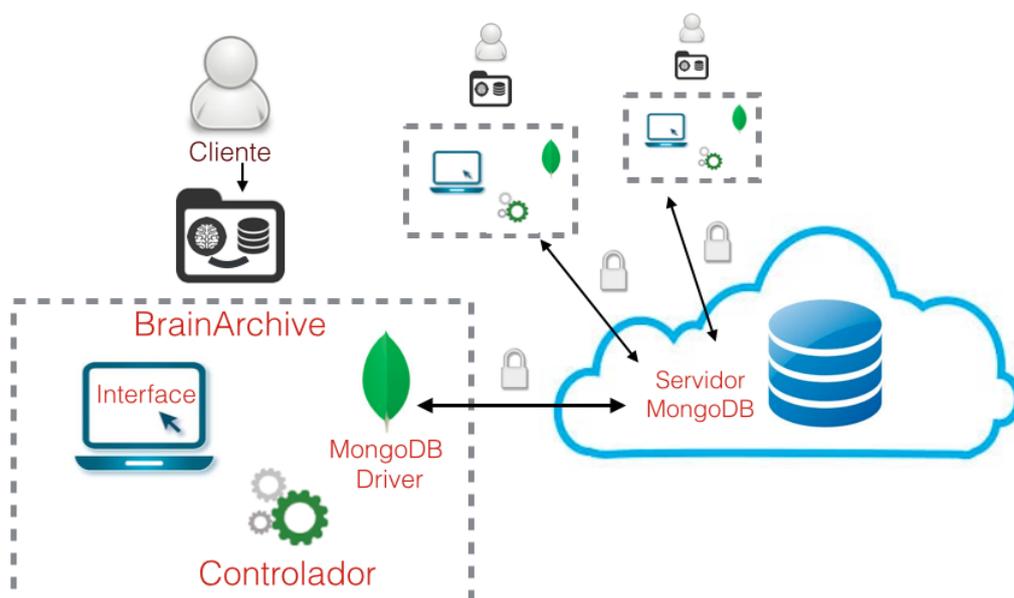


Figura 4.11: Arquitetura geral da aplicação BrainArchive

Por último foi estabelecida uma classe de interação com a base de dados MongoDB, realizando todas as operações de autenticação, leitura e escrita na base de dados de forma transparente. Esta classe faz uso do driver *pymongo*, possibilitando a comunicação entre os clientes da aplicação e o servidor central no qual se encontra o repositório. Detalhes da implementação da base de dados e correspondentes processos e comunicação serão explicados posteriormente. Na Figura 4.11 é observável a interação entre clientes da aplicação BrainArchive o repositório MongoDB. Esta ligação é transparente e resultante da ligação efectuada entre o driver MongoDB presente na aplicação e um servidor responsável pela comunicação com a base de dados.

Tendo por base o fluxo de processamento apresentado na Seção 2.6 e a correspondente divisão e estruturação de cada um dos ficheiros obtidos na Seção 4.4, pretende-se explicar a implementação do repositório utilizando MongoDB.

Como referido anteriormente, os documentos em MongoDB são agrupados

em coleções. Desta forma foram desenvolvidas sete coleções de maneira a estabelecer uma divisão lógica dos ficheiros por análise (tal como estabelecido na [Seção 4.4](#)). Estas sete coleções são então designadas por *FMRI*, *MRI*, *DTI*, *freesurfer*, *DicomZIP*, *Bedpostx* e *TRKVIS*. Tal como é perceptível o número de divisões de análise foi diminuído de nove para sete. Esta redução deve-se ao facto de a coleção *DicomZIP* incluir as diretorias *FMRIDCM*, *MRIDCM* e *DTIDCM*. Esta opção permite assim obter de forma simplificada as aquisições funcionais, estruturais e de difusão num só passo.

Apesar de estas aquisições se incluírem numa só coleção, cada uma encontra-se presente num documento individual. Este documento resulta da compressão das imagens **DICOM** de uma aquisição num arquivo de formato *Zip*. Desta forma pode ser obtido apenas um tipo de aquisição caso seja o pretendido.

O principal objetivo da utilização de MongoDB é a criação de um repositório de ficheiros. Uma vez que alguns destes ficheiros apresentam tamanho superior a 16 MB (tamanho máximo de um documento em MongoDB) foi utilizado o mecanismo GridFS. Como referido este mecanismo "divide" cada documento em diversas partes (*chunks*) de acordo com o tamanho definido para cada parte (*chunksize*). Neste sentido o número real de coleções é quatorze (sete com extensão *.chunks* e outras sete *.files*).

Uma vez enumeradas as coleções da base de dados designada *BrainMongo*, pretende-se agora explicar os campos que as compõem:

- **`_id`** - *ObjectId* desenvolvido no momento de criação do documento em MongoDB. Este é composto por uma combinação complexa, garantindo a unicidade de cada documento;
- **`chunksize`** - tamanho de cada *chunk* que compõe o documento;
- **`length`** - tamanho total do documento/ficheiro;
- **`uploadDate`** - data de criação do documento na base de dados;
- **`md5`** - *hash* md5 utilizado para controlo de versão;
- **`filename`** - caminho (*path*) do ficheiro exportado para a base de dados;

- **file** - caminho (*path*) do ficheiro sem informação da máquina e utilizador;
- **user** - identificação do utilizador responsável pela exportação do ficheiro;
- **study** - nome do estudo do qual o ficheiro faz parte;
- **subject** - identificação do sujeito/caso de cuja aquisição foi produzido o ficheiro;
- **analysis** - tipo de divisão ("análise") a que pertence o ficheiro;
- **filetype** - descrição do ficheiro exportado;
- **permission** - permissão de acesso ao ficheiro definida pelo utilizador no momento de exportação do ficheiro e alterável posteriormente pelo mesmo.

Na Figura 4.12 é apresentado um exemplo de documento da coleção *FMRI.files* da base de dados *BrainMongo*.

```
{
  "_id" : ObjectId("561447d93249bdcb3157f9f5"),
  "filename" : "/Users/Filipe/Desktop/MyBrain/Estudo_1x3/SW0081C/
  FMRI/SW0081C_fnc_vol_stime_mcf_mean_bet_mask.nii.gz",
  "chunkSize" : 261120,
  "uploadDate" : ISODate("2015-10-06T22:14:49.408Z"),
  "md5" : "07cc3cc5dd13a6f494b69cb3ea011a0e",
  "length" : 3039,
  "file" : "Estudo_1x3/SW0081C/
  FMRI/SW0081C_fnc_vol_stime_mcf_mean_bet_mask.nii.gz",
  "user" : "User2",
  "permission" : "User2",
  "analysis" : "FMRI",
  "study" : "Estudo_1x3",
  "filetype" : "Brain Mask",
  "subject" : "SW0081C"
}
```

Figura 4.12: Exemplo de documento de uma base de dados MongoDB

As bases de dados MongoDB podem ser iniciadas, acedidas e terminadas através da linha de comandos ou através de uma biblioteca específica de cada

```
db.createUser(  
  {  
    user: "Filipe Fernandes",  
    pwd: "password",  
    roles: [  
      { role: "readWrite", db: "BrainMongo" }  
    ]  
  }  
)
```

Figura 4.13: Exemplo de criação de utilizador numa base de dados MongoDB

linguagem de programação (*MongoDB Driver*). A base de dados *BrainMongo* aqui especificada é constituída por um servidor, cuja instância em MongoDB se designa *mongod*. Este é o processo (daemon) principal do sistema e tem como funções tratar dos pedidos de dados, gerir o formato dos mesmos e realizar operações de gestão em *background*.

Uma vez que se pretende que a aplicação seja utilizada por diversos utilizadores é necessário fazer a gestão dos mesmos no sistema. Por forma a incluir utilizadores na base de dados é inicialmente necessário a criação de um administrador da base de dados. Este, possui as competências necessárias para criar utilizadores e definir as tarefas garantidas aos mesmos. Na Figura 4.13 é exibido um exemplo de criação de um utilizador na base de dados *BrainMongo*.

Além da necessidade de autenticação no sistema considerou-se necessário proteger o acesso à instância *mongod*. Neste sentido foi utilizada a tecnologia de segurança SSL (actualmente referida como *Transport Layer Security (TLS)*) para estabelecimento de um túnel encriptado entre servidor e cliente. Este protocolo de segurança utiliza certificados *X.509* para autenticação das partes a comunicar e para negociar uma chave de sessão simétrica [93]. Esta chave de sessão é utilizada para encriptar o fluxo de dados entre as partes a comunicar, com intuito de garantir a confidencialidade e integridade dos mesmos.

Como consequência da utilização de certificados *X.509* é necessário recorrer a uma autoridade de certificação, bem como a uma infraestrutura de chave

pública para verificar a relação entre um certificado e o seu proprietário, assim como gerar, assinar e administrar a validade dos mesmos.

Neste sentido foi utilizada a ferramenta *openssl* na criação e validação dos certificados/chaves a utilizar. Para configuração de uma base de dados MongoDB com suporte [SSL](#) é necessário gerar um certificado de chave pública e correspondente chave privada, posteriormente concatenados num ficheiro com extensão *.pem*. Adicionalmente, caso se pretenda desenvolver uma instância *mongod* com encriptação [SSL](#) utilizando um certificado assinado por uma autoridade de certificação é necessária a criação de um ficheiro *.pem* contendo o certificado de raiz da autoridade de certificação.

```
mongod -port <porta> -dbpath -sslMode requireSSL -sslPEMKeyFile <pem>
      -sslCAFile <ac> -fork -logpath [caminho log] -auth
```

Figura 4.14: Exemplo de criação de instância *mongod* com suporte [SSL](#)

Na Figura 4.14 é apresentado um exemplo de criação de uma instância *mongod* com suporte [ssl](#). De seguida pretende-se explicar os parâmetros presentes neste comando:

- **port** - define qual a porta em o *mongod* vai ser executado. Por definição a instância usa a porta 27017;
- **dbpath** - especifica a diretoria onde a instância *mongod* guardará os dados. A localização predefinida é /data/db (em *Windows* )
- **sslMode requireSSL** - permite definir o modo como o processo *mongod* lida com encriptação [SSL](#). Neste caso foi definido "requireSSL", forçando o servidor a aceitar apenas ligações encriptadas. Outras opções são "allowSSL" ou "preferSSL";
- **sslPEMkeyFile** - ficheiro *.pem* contendo o certificado de chave pública e chave privada associada;
- **sslCAFile** - ficheiro *.pem* contendo o certificado de raiz da autoridade de certificação;

- **fork** - permite que o processo *mongod* seja executado em *background*;
- **logpath** - define o caminho para o ficheiro de log onde toda a informação de log será escrita;
- **auth** - força a autenticação para acesso a uma base de dados.

No decorrer deste trabalho serão utilizadas as expressões: exportação para o repositório e importação do repositório, pretendendo-se de seguida clarificar as mesmas. A primeira diz respeito ao processo de introdução de ficheiros locais no repositório desenvolvido. A segunda refere-se ao passo de extração de ficheiros presentes no repositório para a pasta local designada *MyBrain*.

## 4.6 Interface de Utilizador

O principal objectivo de uma interface gráfica é tornar a interação de um programa com o utilizador o mais simples e eficiente possível, melhorando a capacidade do utilizador alcançar determinado objectivo. No caso específico da aplicação BrainArchive foi dada especial atenção à implementação de um design simples e minimalista, especialmente focado nas funções a que se destina.

O objectivo primordial da aplicação desenvolvida é implementar de forma eficiente e automatizada a organização, armazenamento e partilha de estudos multimodais de Ressonância Magnética cerebral. Como tal, não deverá ser incluída na interface informação que não seja necessária à implementação deste sistema.

O número de parâmetros a definir pelo utilizador de ser também minimizado, facilitando a execução de todas as etapas. No desenvolvimento da interface do BrainArchive, optou-se por uma divisão lógica da mesma, consoante o processo em causa e utilidade pretendida para cada elemento.

Na Figura 4.15 encontra-se a primeira janela que qualquer utilizador irá visualizar quando pretender aceder ao sistema. Neste painel o utilizador irá fazer a Autenticação no sistema. Caso a autenticação falhe será apresentada uma mensagem assinalando o sucedido.

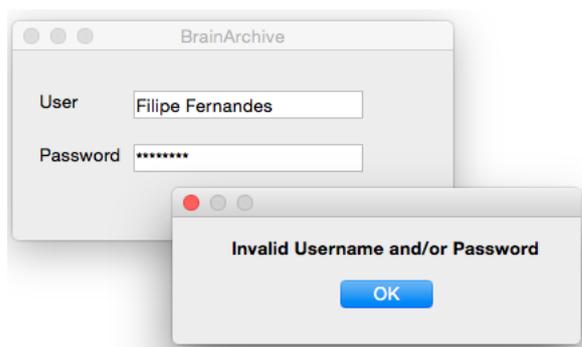


Figura 4.15: Esboço da interface de autenticação e mensagem de falha da mesma na aplicação BrainArchive.

Após o sucesso de autenticação, será apresentado ao utilizador uma janela

como a apresentado na Figura 4.16. A aplicação BrainArchive possui dois separadores principais: *Repository*, onde é apresentada informação relativa aos ficheiros contidos na pasta *MyBrain* e permite exportação dos mesmos para o repositório; *Query*, onde é permitido ao utilizador pesquisar ficheiros presentes no repositório e importar os mesmos para a pasta *MyBrain* local. Por defeito após autenticação a aplicação encontra-se no separador *Repository*. Neste é exibido ao utilizador todos os estudos que compõem a sua pasta local *MyBrain* além de todas as opções que permitem a exportação de ficheiros, análises, sujeitos ou mesmo estudos completos. No lado esquerdo da janela (Figura 4.16.A) encontra-se a estrutura de pastas contendo os estudos locais presentes na pasta *MyBrain*. Idealmente estes encontram-se na organizados por *Estudo/Sujeito/Análise/Ficheiro* como apresentado na Seção 4.4. No canto superior direito (Figura 4.16.B) foram colocados os elementos que permitem ao utilizador realizar a operação de exportação.

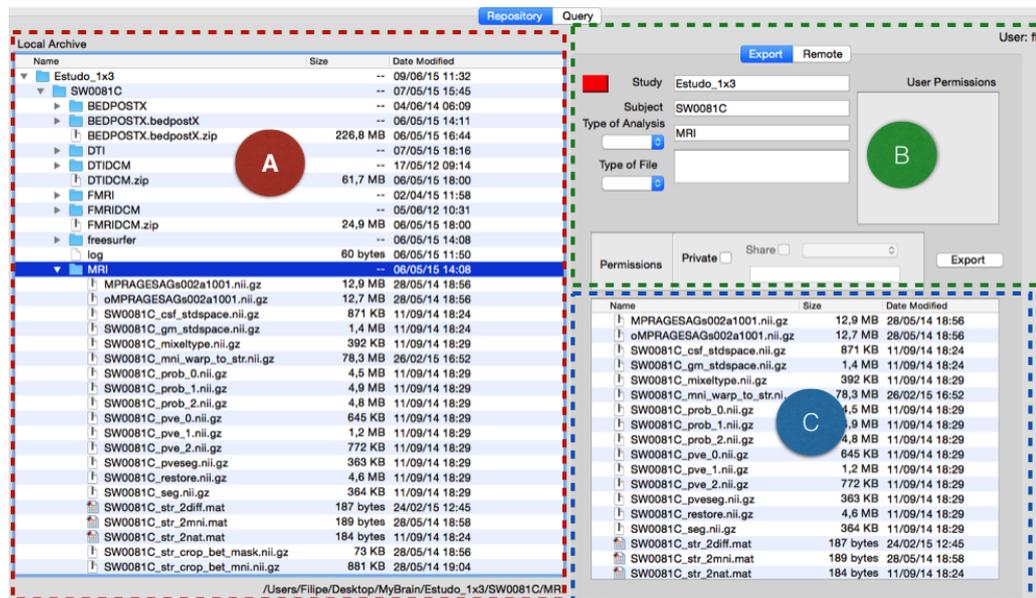


Figura 4.16: Divisão lógica da interface inicial da aplicação BrainArchive. (A) representação dos estudos presentes localmente para um utilizador. (B) seção de exportação de ficheiros/estudos. (C) estruturação de subpastas descendentes de localização selecionada em (A).

Como evidenciado na Figura 4.16.B, a seleção da pasta *MRI* do estudo

*Estudo\_1x3* e sujeito *SW0085C* resulta no preenchimento automático dos campos *Study*, *Subject* e *Type of Analysis*. Desta forma é possível exportar os ficheiros de interesse presentes nesta pasta sem ser necessária intervenção do utilizador.

O quadrado presente no canto superior esquerdo desta divisão permite ao utilizador descobrir o estado de exportação de determinado ficheiro ou análise. Neste sentido, na seleção de um ficheiro, caso este já se encontre presente no repositório, o quadrado apresentará cor verde, ou vermelha caso ainda não tenha sido exportado. Por outro lado, caso a seleção recaia numa pasta como e.g. FMRI, um novo estado pode ser apresentado (amarelo), evidenciando que alguns ficheiros presentes já se encontram exportados mas não a sua totalidade. É também possível alterar a permissão de acesso aos ficheiros exportados (por defeito com permissão pública, mas alterável para privado ou alargado a utilizadores específicos).

Por último, no canto inferior direito ( Figura 4.16.C ) é apresentado ao utilizador a estruturação de subpastas descendentes da localização selecionada (caso seja selecionado um ficheiro nada será apresentado).

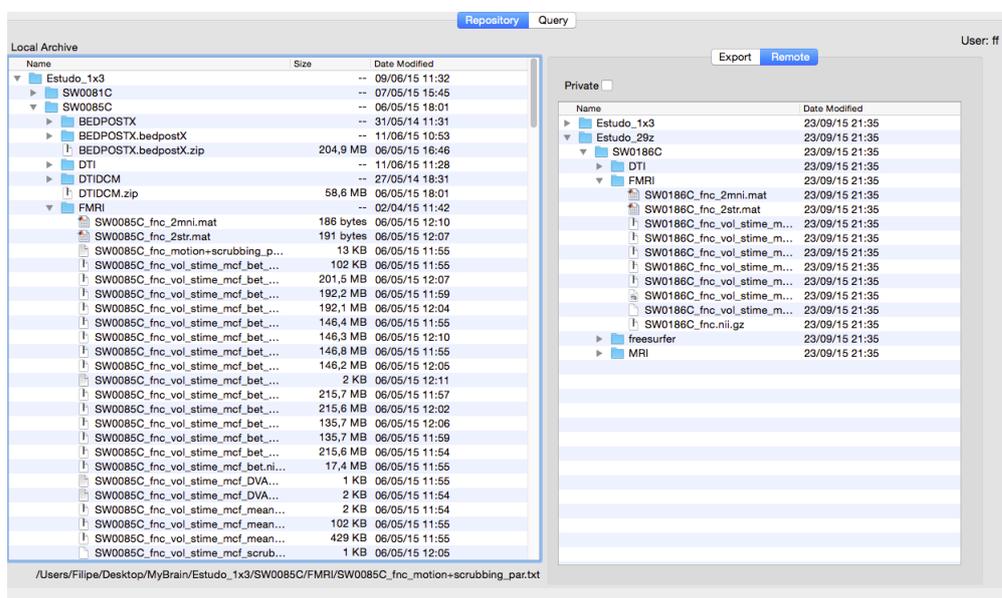


Figura 4.17: Interface da aplicação BrainArchive, para visualização de ficheiros presentes no repositório a que um utilizador tem acesso.

Na janela apresentada na Figura 4.16, os painéis B e C correspondem a um separador denominado *Export*. Ao mesmo nível deste existe um separador denominado *Remote* (Figura 4.17). Caso este separador seja seleccionado é exibido ao utilizador os estudos aos quais tem acesso que já se encontram remotamente partilhados no repositório. O utilizador pode ainda alternar entre visualizar os estudos por ele exportados ou dos quais tem acesso através de partilha por outros utilizadores.

Tendo novamente por base a janela apresentada na Figura 4.16, é possível ao utilizador seleccionar o separador *Query*. Nesta seção da interface evidenciada na Figura 4.18 é apresentado ao utilizador um conjunto de soluções para a importação de ficheiros, análises, sujeitos, estudos ou tudo ao qual o mesmo possui acesso, além de permitir a alteração de permissões de ficheiros por ele exportados.

The screenshot shows the 'Query' tab of the BrainArchive interface. On the left, there is a search filter section (A) with fields for User, Study, Subject, Analysis, and Type of file. Below this is a table with columns for User, Study, and Subject, containing entries like 'BrainMongoTestUse', 'Estudo\_29z', and 'SW0186C'. At the bottom left, there is a 'Change File Permissions' section (B) with options for 'To Public' and 'Add User', and buttons for 'Change All' and 'Change Selected'. The main area (C) is a table listing files from a 'Remote Archive' with columns for ID, Study, Subject, Analysis, User, FileType, File, and UploadDate. The table contains 20 rows of data, including entries like 'Estudo\_1x3', 'Estudo\_29z', 'SW0081C', 'SW0186C', 'DTI', 'ff', 'Mean diffusivity map', 'Diffusion weights', 'Diffusion gradients', 'Raw data', 'Raw b0 image', 'Brain Mask', 'Fractional anisotropy map', 'First eigenvalue map', 'Second eigenvalue map', 'Third eigenvalue map', 'Mean diffusivity map', 'Radial diffusivity map', 'B0 image', 'Diffusion tensors', 'First eigenvector', 'Second eigenvector', 'Third eigenvector', 'Eddy current correction transformations', 'Preprocessed data', and 'Rotated diffusion gradients'. At the bottom right, there are buttons for 'Import Selected File(s)' and 'Import All File(s)'.

Figura 4.18: Interface de interrogação e importação da aplicação BrainArchive. (A) seção de filtragem de ficheiros. (B) alteração de permissões de acesso a ficheiros. (C) seção de visualização e importação de ficheiros presentes no repositório.

Na Figura 4.18 é apresentada a divisão da interface por funcionalidade. Na seção superior esquerda (Figura 4.18.A) encontram-se todas as opções que permitem ao utilizador filtrar a pesquisa de ficheiros por *Utilizador, Estudo, Sujeito/Caso, Análise* e *Tipo de ficheiro*, sendo apenas exibidos os ficheiros aos quais o mesmo tem permissão de acesso.

A correspondente seleção destes parâmetros actualiza automaticamente a lista de ficheiros apresentados na secção representada na Figura 4.18.C. Por sua vez esta lista contém todos os ficheiros filtrados, revelando o estudo, sujeito, utilizador e tipo de ficheiro a que pertence, bem como a data de exportação do mesmo para a base de dados. Tendo em conta a necessidade de importação de ficheiros foram adicionados botões (*Import Selected File(s)* e *Import All Files*) colocados no quadrante inferior direito desta secção (4.18.C). Estes botões permitem respectivamente: importar o(s) ficheiro(s) seleccionados na lista apresentada ; importar todos ficheiros exibidos na mesma lista.

O processo de importação decorre de forma transparente para o utilizador, com alocação estruturada na pasta *MyBrain* segundo a lógica estabelecida na Seção 4.4. Simultaneamente a divisão correspondente à Figura 4.16.A é actualizada revelando os ficheiros importados.

Por último, caso a opção *private* (presente na divisão 4.18.A) seja seleccionada, apenas os ficheiros por ele exportados serão apresentados e a área correspondente à Figura 4.18.B torna-se configurável para o utilizador. Nesta secção o utilizador poderá alterar a permissão de acesso a ficheiros. É assim possível seleccionar determinado(s) ficheiro(s) bem como todos os ficheiros apresentados, alterando a permissão de acesso aos mesmos ( público, privado ou disponível a utilizadores específicos).

Por forma a facilitar o processo de integração entre a aplicação e a pasta MyBrain foi desenvolvido um método exibido no Menu *Import Local Study*. Esta opção permite que sejam importados estudos locais (presentes no computador de forma possivelmente desestruturada e separada) para a pasta *MyBrain* e conseqüente utilização na aplicação BrainArchive. Inicialmente é requerido ao utilizador que forneça o nome do estudo em causa bem como o número de sujeitos a importar (Figura 4.19.A).

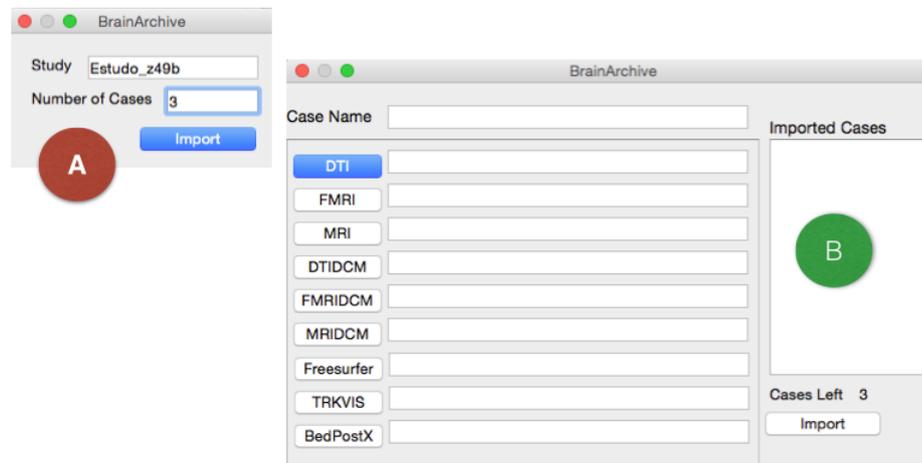


Figura 4.19: Janelas da interface BrainArchive para importação local assistida. (A) definição do estudo e número de casos. (B) seleção de sujeito/caso e diretorias a importar.

Após o preenchimento destes campos uma nova janela será apresentada (Figura 4.19.B). Neste momento o utilizador pode fazer uso de 9 botões desenvolvidos respectivamente para a importação da diretoria em causa, selecionando para cada caso o nome atribuído ao mesmo. No final de importação a pasta *MyBrain* possuirá todos os ficheiros importados de forma organizada segundo a estrutura apresentada na [Seção 4.4](#).

Foi também adicionado ao Menu a opção *Logout*. Tal como o nome sugere, esta opção permite ao utilizador terminar a sessão, voltando a ser apresentada a janela de autenticação (Figura 4.15).

Os elementos visuais e a sua disposição na interface BrainArchive serão analisados com maior pormenor à medida que forem explicitadas as principais funcionalidades da aplicação. Foi ainda desenhado um ícone para identificação do BrainArchive, que se apresenta na Figura 4.20 e cuja aparência remete para interligação da conectividade cerebral numa base de dados.



Figura 4.20: Ícone da aplicação BrainArchive

## 4.7 Principais funcionalidades

A utilidade de uma aplicação e a forma como permite ao utilizador atingir os seus objectivos encontra-se directamente relacionada com as funcionalidades que a aplicação implementa. No caso do BrainArchive, as funcionalidades implementadas refletem em grande parte a estrutura definida na [Seção 4.4](#).

Um estudo neuroimagiológico de Ressonância Magnética é tipicamente composto por vários casos/sujeitos. Para cada caso são necessárias aquisições de diferentes modalidades e realizados diferentes passos de pré-processamento e posterior análise. Daqui resulta não só um grande volume de dados mas também um elevado número de ficheiros. Como tal, é conveniente a organização de uma hierarquia de estudo sob uma diretoria de trabalho.

Neste sentido, pretende-se que cada utilizador organize os seus estudos (como apresentado na [Seção 4.4](#)) numa diretoria designada *MyBrain*. A aplicação desenvolvida não só tem em conta esta organização como tira partido dela. De seguida pretende-se apresentar as funcionalidades essenciais desenvolvidas na aplicação BrainArchive:

- **Autenticação** - para aceder à aplicação BrainArchive, cada utilizador deverá autenticar-se no sistema fornecendo para tal o seu nome de utilizador e respectiva palavra passe. Tal como referido a segurança de comunicação entre a aplicação BrainArchive e o repositório é assegurada por [SSL](#);
- **Incorporação da diretoria *MyBrain* e exportação de ficheiros** - após o sucesso na autenticação é apresentado ao utilizador a janela

inicial da aplicação BrainArchive. Nesta serão exibidos os ficheiros presentes na sua diretoria local *MyBrain*, sendo possível a exportação dos mesmos para o repositório. Neste sentido, e caso a organização de estudos siga a estrutura detalhada na [Seção 4.4](#), é possível exportar de forma automática análises, sujeitos e mesmo estudos completos (possivelmente constituídos por inúmeros sujeitos e análises). Desta forma, o processo de exportação de grandes volumes de informação é simplificado, podendo ser realizado num único passo.

Aquando do processo de exportação de ficheiros, o utilizador pode definir a permissão de acesso aos mesmos, podendo estes ser de domínio público ou privado (disponível apenas para si próprio ou alargado a utilizadores específicos);

- **Importação assistida de ficheiros para a diretoria *MyBrain*** - a aplicação BrainArchive disponibiliza um modo de importação de estudos locais para a diretoria *MyBrain*, segundo a hierarquia *Estudo/Sujeito/Análise/Ficheiros* descrita na [Seção 4.4](#). Desta forma é possível definir o identificador do estudo e respectivos sujeitos/casos, selecionando para cada as diferentes análises a importar.

Desta forma, os respectivos ficheiros serão copiados para a diretoria *MyBrain* com a estrutura correcta e reconhecida pela aplicação BrainArchive;

- **Extração de ficheiros presentes no repositório** - toda a informação presente no repositório pode em cada instante ser importada para a diretoria local na máquina no qual está a ser utilizada a aplicação. Neste sentido é apresentada uma interface simples no qual cada ficheiro pode ser filtrado por utilizador, estudo, sujeito, análise ou tipo de ficheiro, importando assim apenas os dados necessários de forma intuitiva e automática. Estes ficheiros são então colocados na diretoria *MyBrain*, mantendo a estrutura com o qual foram anteriormente exportados para o repositório.

Além da importação de ficheiros, cada utilizador pode alterar a per-

missão de acesso a ficheiros por ele anteriormente exportados para o repositório;

Maior detalhe relativo às funcionalidades implementadas na aplicação BrainArchive é apresentado no [Apêndice C](#).

# Capítulo 5

## Resultados e Conclusões

O presente capítulo enquadra e caracteriza de forma sinóptica o trabalho efectuado e apresenta resultados e conclusões do presente trabalho. Para além disso, são apresentadas propostas de trabalho futuro.

### 5.1 Sinopse

No presente trabalho, e seguindo as linhas em epígrafe, é possível resumir os resultados obtidos, fazendo um pequeno sumário dos objectivos alcançados em cada capítulo:

- **Capítulo 1**

Foi apresentado o enquadramento deste trabalho com a introdução ao conceito de Informática Médica, o seu domínio e área de investigação. De seguida, foi apresentado o contexto da disciplina de Neuroimagem, com especial destaque à técnica de Ressonância Magnética. A motivação para a escolha da técnica de [MRI](#) e diversas modalidades de análise é também apresentada neste capítulo.

De seguida foi apresentado o problema que este trabalho se propunha solucionar, tendo em conta diversos problemas detectados no armazenamento e partilha de estudos neuroimagingológicos.

Por último, neste capítulo foram apresentados os objectivos do trabalho, a metodologia de investigação utilizada e a estrutura do documento.

- **Capítulo 2**

No segundo capítulo foi abordada a temática imagiológica de Ressonância Magnética, com destaque para as modalidades de aquisição estrutural, funcional e de difusão (cuja informação mais detalhada se encontra presente no [Apêndice A](#)). Foi então descrito um fluxo de processamento para estudos multimodais de [MRI](#), constituído por diversas etapas de pré-processamento e técnicas de análise posteriores. Para cada momento/passo do fluxo de processamento foi referida a informação considerada essencial na execução do mesmo, sob a forma de ficheiros.

Por último foi apresentado o volume de dados gerados para estudos multimodais típicos e soluções de mercado existentes para armazenamento e partilha de informação neuroimagiológica.

- **Capítulo 3**

O terceiro capítulo explora conceitos tecnológicos e soluções, encontrando-se dividido em duas partes. Na primeira são apresentadas abordagens de arquiteturas em rede, comparando para tal o modelo cliente-servidor e [P2P](#). São referidas vantagens e desvantagens associadas a ambos os modelos.

Numa segunda parte é realizada uma análise detalhada de diversas soluções de bases de dados, com especial destaque para a comparação da vertente de base de dados relacionais e [NoSQL](#)

- **Capítulo 4**

No capítulo 4 é apresentada e descrita a aplicação informática desenvolvida, [BrainArchive](#). Inicialmente é elaborada uma introdução relativa à contextualização do sistema proposto. De seguida é apresentada uma

análise de requisitos do sistema a cumprir, bem como a base tecnológica utilizada no desenvolvimento da aplicação. Desta forma é apresentada a linguagem de programação Python bem como as respectivas ferramentas utilizadas no desenvolvimento da aplicação informática e correspondente interface gráfica. Além disso é também explorada a base de dados MongoDB, evidenciando a motivação da sua escolha na construção de um repositório de ficheiros neuroimagiológicos.

É ainda abordada a hierarquia e nomenclatura de estudos para a organização do fluxo de processamento presente no Capítulo 2. Desta nomenclatura resulta um dicionário de ficheiros representado na íntegra no [Apêndice B](#). Além disso é apresentada a arquitetura da aplicação e implementação do Repositório na base de dados MongoDB.

Por último são analisados os componentes e principais funcionalidades da aplicação (detalhados em maior profundidade no [Apêndice C](#)), demonstrando de forma intuitiva o processo de utilização da mesma.

## 5.2 Resultados

A Ressonância Magnética é actualmente o método mais utilizado na exploração e avaliação de patologias, estrutura e função cerebral. Estudos neuroimagiológicos de [MRI](#) envolvem geralmente a utilização de muitos sujeitos e realização de múltiplas aquisições. Tal pressuposto aliado à execução de diversas técnicas de análise, etapas de pré-processamento e processos computacionalmente exigentes, implica cada vez mais a existência de soluções para manutenção e organização de dados.

Além disso, a crescente necessidade de partilha de aquisições, porções de análise ou mesmo estudos completos, acresce a relevância de desenvolvimento de sistemas que permitam de forma simplificada a distribuição de informação neuroimagiológica.

O trabalho realizado no âmbito da presente dissertação teve como resultado o desenvolvimento de uma aplicação e repositório, para um sistema completo de estruturação, armazenamento, e distribuição automática de da-

dos multimodais de Ressonância Magnética cerebral.

De forma a ser possível orientar o desenvolvimento da aplicação e implementar o repositório de estudos, foi necessário estudar o fluxo de processamento proposto na [Seção 2.6](#). Esta análise permitiu detalhar os ficheiros que se consideram necessários, para uma completa reconstrução do fluxo de processamento considerado. Uma vez seleccionados estes ficheiros foi possível definir uma hierarquia de estudos, no qual cada documento possui uma nomenclatura própria, caracterizada por uma descrição concisa.

De seguida, foi possível desenvolver uma arquitetura de armazenamento e partilha de informação, através de um serviço de computação em nuvem, cujos componentes funcionais se encontram distribuídos por diferentes entidades. Neste sentido foi implementada um sistema de base de dados MongoDB, no qual cada ficheiro armazenado possui um conjunto de metadados associados, que permitem definir de forma rigorosa esse mesmo documento. Na [figura 5.1](#) é possível verificar o processo de exportação/importação de ficheiros implementado na arquitetura desenvolvida.

No entanto a arquitetura desenvolvida permite estender este processo para a exportação/importação automática de estudos, tirando partido da hierarquia e nomenclatura definida. Desta forma, é possível num único passo exportar e.g. um estudo completo. Este mesmo estudo pode por sua vez ser constituído por uma grande variedade de sujeitos constituídos por múltiplos ficheiros. Face a esta característica todo o processo de partilha de ficheiros é facilitado, uma vez que não é necessário exportar/importar cada ficheiro individualmente. Além disso, cada ficheiro exportado automaticamente, apresenta um conjunto de informação que o caracteriza, sem necessidade de definição manual por parte do utilizador. Esta mesma informação permite filtrar os ficheiros presentes no repositório, através de parâmetros que os definem, obtendo deste os dados que se pretende. Este processo encontra-se esquematizado na [figura 5.2](#). Nesta [figura \(5.2\)](#) é possível verificar a relação entre o processo de exportação e importação de estudos, tendo em conta a informação presente em cada ficheiro.

Além das funcionalidades apresentadas, foram ainda exploradas as questões de segurança e permissão de acesso a ficheiros. Neste sentido pretendeu-

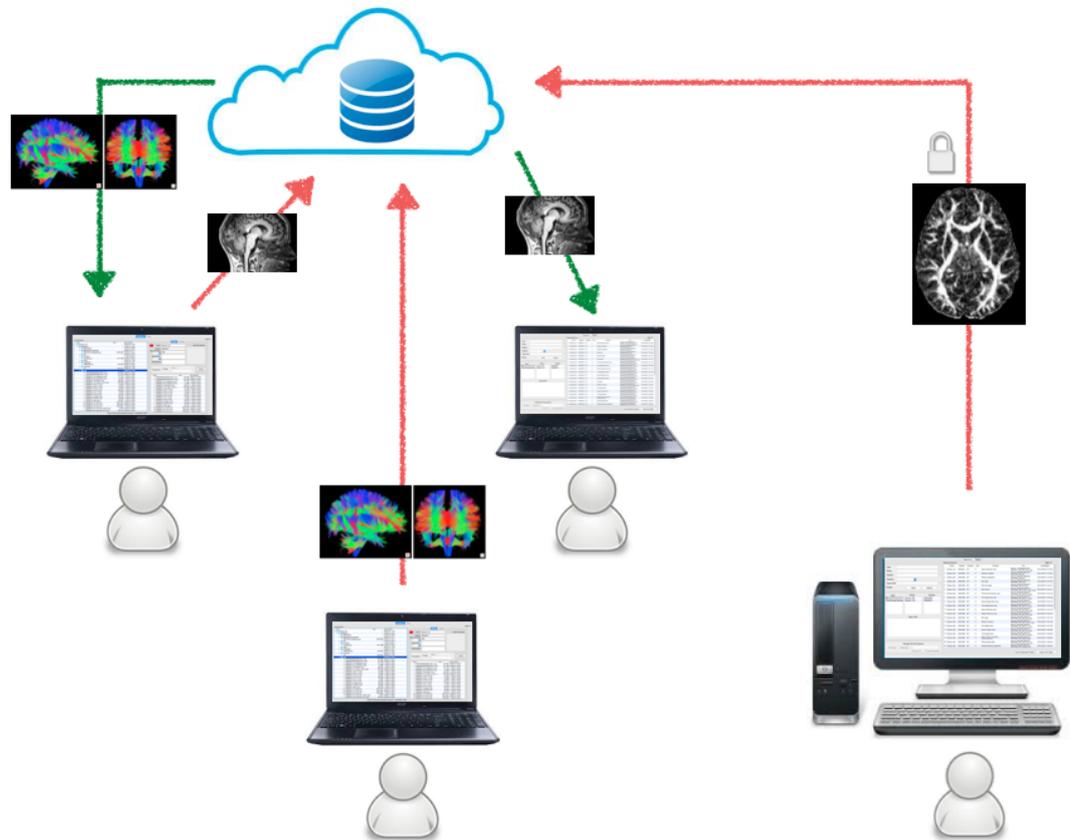


Figura 5.1: Processo de exportação/importação de ficheiros entre utilizadores da aplicação BrainArchive. Linhas a vermelho representam processos de exportação e a verde de importação.

se assegurar a segurança de acesso aos dados utilizando [SSL](#) na ligação entre a aplicação BrainArchive e a base de dados MongoDB. Por outro lado cada utilizador pode em cada momento alterar as permissões de acesso aos dados por si exportados para o repositório. A aplicação desenvolvida apresenta uma interface de aspecto compacto e simples, focada nas funcionalidades que se propõe resolver.

Analisando as soluções apresentadas na [Seção 2.8](#) é possível denotar funcionalidades semelhantes entre as mesmas e a aplicação BrainArchive. Tal como a aplicação BrainArchive, a arquitetura ANIMA explorada na [Subse-](#)

ção 2.8.1, permite aos seus utilizadores pesquisar por dados submetidos, uma vez que a cada elemento se encontra associado um conjunto de metadados descrevendo o seu conteúdo. No mesmo sentido, as arquiteturas GAAIN e XNAT descritas respectivamente na Subsecção 2.8.3 e 2.8.4, permitem a captura de dados de diversas fontes e posterior centralização de informação.

Apesar da existência de competências genéricas entre estas aplicações, a solução BrainArchive foi desenvolvida com a pretensão de facilitar ao utilizador todo o processo organizativo e de partilha de estudos. Desta forma a incorporação de uma estrutura e nomenclatura de estudos definida, permite a investigadores organizar e partilhar informação de forma automática, sem necessidade de definição manual e individual de cada ficheiro, reduzindo significativamente o tempo despendido neste processo.

Em suma, a solução desenvolvida pretende ser um método inovador na organização e distribuição de fluxos de processamento multimodais já implementados. Desta forma, não são introduzidas alterações nos passos de investigação até então realizados, mas sim na forma de manutenção e partilha de informação. Esta alternativa permite a conjugação entre ciclos de processamento e a estruturação definida de ficheiros resultantes, tornando o reaproveitamento, partilha e colaboração entre profissionais uma tarefa simplificada. Neste sentido é reduzido o esforço e tempo suplementar necessário para manter a coerência da informação gerada, ao mesmo tempo que se diminui a probabilidade de se cometerem erros inadvertidos na combinação de estudos e perda de informação. A replicação de ficheiros locais no repositório permite a investigadores armazenar de forma segura informação estruturada, assim como partilhar e extrair dados, através de uma interface simples e intuitiva. Os ficheiros presentes no repositório são por sua vez pesquisáveis através da seleção de parâmetros que os caracterizam como e.g. o estudo, sujeito, análise e mesmo tipo de ficheiro. Após seleção da informação pretendida, a mesma pode ser extraída do repositório mantendo a estrutura e nomenclatura com o qual foi anteriormente exportada para o repositório.



Figura 5.2: Processo de exportação e importação de estudos através da aplicação BrainArchive, baseado em hierarquia e nomenclatura definida.

## 5.3 Conclusões

De uma forma geral, foi possível encontrar soluções viáveis para a resolução dos problemas e desafios colocados no início deste trabalho, tendo sido dada preferência a tecnologias modernas e de uso livre. O sistema desenvolvido exigiu o estudo e desenvolvimento de um conjunto de áreas de conhecimento abrangente.

Tendo em conta o que aqui foi apresentado, é possível afirmar que o protótipo desenvolvido, englobando a aplicação BrainArchive e repositório MongoDB, constituem um sistema inovador e de características vantajosas a nível funcional, de integração e disponibilidade.

Terminado o trabalho da presente dissertação, e após um período reflexivo, foi possível concluir que o mesmo constitui um importante passo na disseminação de informação de imagem cerebral. Durante todo o processo de desenvolvimento, novas ideias e possíveis melhoramentos foram surgindo, cuja implementação seria positiva. De seguida são expostas as principais ideias identificadas.

Da execução de estudos de **MRI** (principalmente multimodais), resulta uma grande quantidade de informação. Desta forma, caso a utilização deste sistema seja alargado a muitas entidades será necessário desenvolver uma solução para armazenamento massivo de dados. Como referido anteriormente, as bases de dados MongoDB possuem um mecanismo designado por *sharding*, no qual a base de dados é particionada horizontalmente, por várias instâncias separadas. Por conseguinte, é possível "dividir" a informação presente na base de dados em múltiplos servidores, tirando uso da capacidade de armazenamento de cada um.

Em termos de funcionalidade seria também interessante alargar a solução desenvolvida ao nível da prestação de serviços. Desta forma à aplicação poderão ser introduzidos métodos para controlo do estado de execução de determinado estudo, tirando proveito de capacidades gráficas e elaboração de relatórios.

Com a aplicação desenvolvida pretende-se alertar para os claros benefícios da utilização de soluções informáticas integradas e automatizadas no armazenamento e partilha de dados das modalidades neuroimagiológicas de **MRI**. Este tipo de soluções constituem ferramentas extremamente úteis, facilitando a interação entre investigadores e instituições de saúde. Este aspecto conduz à disseminação de informação, potenciando o surgimento de importantes desenvolvimentos em investigação científica e na prática clínica, cujos benefícios se repercutem em melhores cuidados de saúde. O sistema projectado está orientado a estudos multimodais de **MRI**. No entanto, pode ser facilmente estendido a outras áreas de conhecimento, que envolvam a necessidade de armazenamento e partilha de informação padronizada.

## Referências

- [1] Reinhold Haux, Medical informatics: Past, present, future, *International Journal of Medical Informatics*, Volume 79, Issue 9, September 2010, Pages 599-610.
- [2] Shortliffe, E.H., Cimino, J.J., *Biomedical Informatics Computer Applications in Health Care and Biomedicine*. New York: Springer (2006).
- [3] Hoyt, R., Sutton, M., Yoshihashi, A.: *Medical Informatics: Practical Guide for the Healthcare Professional*. Lulu.com (2008).
- [4] Jamal, A., McKenzie, K., Clark, M.: The impact of health information technology on the quality of medical and health care: a systematic review. *Heal. Inf. Manag. J.* 38, (2009).
- [5] Greenes, R. A., Shortliffe, E. H. Medical Informatics An Emerging Academic Discipline and Institutional Priority. *Journal of the American Medical Association* , 1114-1120. (1990).
- [6] Degoulet, P., Fieschi, M. Critical dimensions in medical informatics. *International Journal of Medical Informatics* , 21-26.(1997).
- [7] Vanderby S., Pena Sanchez J., Kalra N., Babyn P., Finding the Truth in Medical Imaging: Painting the Picture of Appropriateness for Magnetic Resonance Imaging in Canada, *Canadian Association of Radiologists Journal*, 12 August 2015.
- [8] Elliott, A. Medical Imaging. *Nuclear Instruments and Methods in Physics Research* , 1- 13 (2005).

- [9] Richter, R. W., Richter, B. Z. Alzheimer's disease: a physician's guide to practical management. Humana Press (2004).
- [10] Horwitz, B. The elusive concept of brain connectivity. *NeuroImage* , 466-470 (2003).
- [11] Chung M, Dahabreh IJ, Hadar N, et al. Emerging MRI Technologies for Imaging Musculoskeletal Disorders Under Loading Stress [Internet]. Rockville (MD): Agency for Healthcare Research and Quality (US); 2011 Nov.
- [12] Tetsuya Takahashi, Complexity of spontaneous brain activity in mental disorders, *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, Volume 45, 1 August 2013, Pages 258-266.
- [13] Klauschen, F., Goldman, A., Barra, V., Lindenberg, A. M., Lunder-vold, A. Evaluation of Automated Brain MR Image Segmentation and Volumetry Methods. *Human Brain Mapping* , 1310-1327. (2009).
- [14] Ashburner J, Friston KJ. Why voxel-based morphometry should be used. *Neuroimage*. (2001).
- [15] Certaines, J. D., Bovee, W., Podo, F. *Magnetic Resonance Spectroscopy in Biology and Medicine*. Oxford: Pergamon Press. (1992).
- [16] Wedad S. Salem, Ahmed F. Seddik, Hesham F. Ali. A Review on Brain MRI Image Segmentation. Computers and Systems Department, Electronics Research Institute, Cairo, Egypt. Biomedical Engineering Department, Helwan University, Egypt. (2013)
- [17] Steen, R. G., Mull, C., McClure, R., Hamer, R. M., Lieberman, J. A. Brain volume in first-episode schizophrenia. *The British Journal of Psychiatry* , 510-518. (2006).
- [18] Qian Shen, Weizhao Zhao, David A. Loewenstein, Elizabeth Potter, Maria T. Greig, Ashok Raj, Warren Barker, Huntington Potter, Ranjan Duara, Comparing new templates and atlas-based segmentations

in the volumetric analysis of brain magnetic resonance images for diagnosing Alzheimer disease, *Alzheimer's Dementia*, Volume 8, Issue 5, September 2012, Pages 399-406.

- [19] Sylvie Goldman, Liam M. Brien, Pauline A. Filipek, Isabelle Rapin, Martha R. Herbert, Motor stereotypies and volumetric brain alterations in children with Autistic Disorder, *Research in Autism Spectrum Disorders*, Volume 7, Issue 1, January 2013, Pages 82-92.
- [20] Jones, D. Studying connections in the living human brain with diffusion MRI . *Cortex* , 44, 936-952. (2008)
- [21] Eriksson SH, Rugg-Gunn EJ, Symms MR, et al. Diffusion tensor imaging in patients with epilepsy and malformations of cortical development. *Brain* 2001.
- [22] Guo AC, Jewells VL, Provenzale JM. Analysis of normal-appearing white matter in multiple sclerosis: comparison of diffusion tensor MR imaging and magnetization transfer imaging. *AJNR Am J Neuroradiology* (2001).
- [23] Foong J, Maier M, Clark C, et al. Neuropathological abnormalities of the corpus callosum in schizophrenia: a diffusion tensor imaging study. *J Neurol Neurosurg Psychiatry* 2000.
- [24] Filippi CG, Edgar MA, Ulug AM, et al. Appearance of meningiomas on diffusion-weighted images: correlating diffusion constants with histopathologic findings. *AJNR Am J Neuroradiol* 2001.
- [25] Fingelkurts, A. A., Fingelkurts, A. A., Kahkonen, S. Functional connectivity in the brain-is it an elusive concept? *Neuroscience and Biobehavioral Reviews* (28), 827-836. (2005).
- [26] Newton, H. B., Jolesz, F. A. *Handbook of neuro-oncology neuroimaging*. Academic Press. (2007).

- [27] Lowe, M., Dzemidzic, M., Lurito, J., Mathews, V., Phillips, M. Correlations in Low-Frequency BOLD Fluctuations Reflect Cortico-Cortical Connections. *NeuroImage*, 12(5), 582-587. (2000)
- [28] Faro, S. H., Mohamed, F. B., Law, M., Ulmer, J. T. *Functional Neuroradiology: Principles and Clinical Applications* (1st ed.). Springer. (2011).
- [29] Huettel, S. A., Song, A. W., McCarthy, G. *Functional magnetic resonance imaging*. Sinauer Associates. (2009).
- [30] Sabsevitz, D., Swanson, S., Hammeke, T. A., Spanaki, M. V., Possing, E. T., Morris, G. L., Mueller, W. M., et al. Use of preoperative functional neuroimaging to predict language deficits from epilepsy surgery. *Neurology*, 60(11), 1788 -1792. (2003).
- [31] Castelo-Branco, M., Formisano, E., Backes, W., Zanella, F., Neuenschwander, S., Singer, W., Goebel, R. Activity patterns in human motion-sensitive areas depend on the interpretation of global motion. *Proceedings of the National Academy of Sciences of the United States of America*, 99(21), 13914-13919. (2002).
- [32] Rogers, B. P., Morgan, V. L., Newton, A. T., Gore, J. C. Assessing functional connectivity in the human brain by fMRI. *Magnetic Resonance Imaging* (25), 1347-13. (2007).
- [33] Sporns, O. *Networks of the Brain*. MIT Press. (2010).
- [34] Friston, K. J., Frith, C. D., Liddle, P. F., Frackowiak, R. S. Functional connectivity: the principal-component analysis of large (PET) data sets. *Journal of Cerebral Blood Flow and Metabolism: Official Journal of the International Society of Cerebral Blood Flow and Metabolism*, 13(1), 5-14. (1993).
- [35] Sporns, O., Tononi, G., Kotter, R. The human connectome: A structural description of the human brain. *PLoS Computational Biology*, 1(4), e42. (2005).

- [36] He, Y., Evans, A. Graph theoretical modeling of brain connectivity. *Current Opinion in Neurology* , 1-10. (2010).
- [37] Zalesky, A., Cocchi, L., Fornito, A., Murray, M. M., Bullmore, E. Connectivity differences in brain networks. *NeuroImage* , 1055-1062. (2012).
- [38] Takashi Itahashi, Takashi Yamada, Motoaki Nakamura, Hiromi Watanabe, Bun Yamagata, Daiki Jimbo, Seiji Shioda, Miho Kuroda, Kazuo Toriizuka, Nobumasa Kato, Ryuichiro Hashimoto, Linked alterations in gray and white matter morphology in adults with high-functioning autism spectrum disorder: A multimodal brain imaging study, *NeuroImage: Clinical*, Volume 7, 2015, Pages 155-169, ISSN 2213-1582.
- [39] Andrew T. Reid, Danilo Bzdok, Sarah Genon, Robert Langner, Veronika I. Muller, Claudia R. Eickhoff, Felix Hoffstaedter, Edna-Clarisse Cieslik, Peter T. Fox, Angela R. Laird, Katrin Amunts, Svenja Caspers, Simon B. Eickhoff, ANIMA: A data-sharing initiative for neuroimaging meta-analyses, *NeuroImage*, Available online 29 July 2015, ISSN 1053-8119.
- [40] Gregory A. Book, Michael C. Stevens, Michal Assaf, David C. Glahn, Godfrey D. Pearlson, Neuroimaging data sharing on the neuroinformatics database platform, *NeuroImage*, Available online 16 April 2015, ISSN 1053-8119.
- [41] Michael R. Hodge, William Horton, Timothy Brown, Rick Herrick, Timothy Olsen, Michael E. Hileman, Michael McKay, Kevin A. Archie, Eileen Cler, Michael P. Harms, Gregory C. Burgess, Matthew F. Glasser, Jennifer S. Elam, Sandra W. Curtiss, Deanna M. Barch, Robert Oostenveld, Linda J. Larson-Prior, Kamil Ugurbil, David C. Van Essen, Daniel S. Marcus, ConnectomeDB - A Sharing human brain connectivity data, *NeuroImage*, Available online 29 April 2015, ISSN 1053-8119.

- [42] Rex, D. E., Ma, J. Q., Toga, A. W. The LONI Pipeline Processing Environment. *NeuroImage*, 19(3), 1033-1048. (2003).
- [43] Marques P, Soares JM, Alves V and Sousa N. BrainCAT a tool for automated and combined functional magnetic resonance imaging and diffusion tensor imaging brain connectivity analysis. *Front. Hum. Neurosci.* 7:794. (2013)
- [44] Avison, D., Lau, F., Myers, M., Nielsen, P.A.: *Action research*, (1999).
- [45] Somekh, B.: *Action research: A methodology for change and development*. (2006).
- [46] McNiff, J.: *Action research for professional development Concise advice for new action researchers*. Routledge (2002).
- [47] Kuperman, V. *Magnetic Resonance Imaging - Physical Principles and Applications*. Maryland: Academic Press. (2000).
- [48] Cole, D. M., Smith, S. M., Beckmann, C. F. Advances and pitfalls in the analysis and interpretation of resting-state FMRI data. *Frontiers in Systems Neuroscience*, 4, 8. (2010)
- [49] Greicius, M. D., Supekar, K., Menon, V., Dougherty, R. F. Resting-State Functional Connectivity Reflects Structural Connectivity in the Default Mode Network. *Cerebral Cortex*, 19(1), 72 -78. (2009).
- [50] Fischl, B. FreeSurfer. *NeuroImage* , 774-781. (2012).
- [51] M., J., F., B. C., Behrens, T. E., W., W. M., Smith, S. M. FSL. *NeuroImage* , 782-790. (2012).
- [52] Matthews, P. M., Jezzard, P. *Functional magnetic resonance imaging. Neuroscience for Neurologists* , 6-12. (2004).
- [53] Ramsey, N., Hoogduin, H., Jansma, J. *Functional MRI experiments: acquisition, analysis and interpretation of data. European Neuropsychopharmacology* , 517-526. (2002).

- [54] Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcoix, N., et al. Automated Anatomical Labeling of Activations in SPM Using a Macroscopic Anatomical Parcellation of the MNI MRI Single-Subject Brain. *NeuroImage* , 273-289. (2002).
- [55] Mccauslandcenter.sc.edu, (2015). dcm2nii DICOM to NIfTI conversion. [online] Available at: <http://www.mccauslandcenter.sc.edu/mricro/mricron/dcm2nii.html> [Accessed 13 Aug. 2015].
- [56] The R-fMRI Network, Data Processing Assistant for Resting-State fMRI (DPARSF) V4.0. [online] Available at: <http://rfmri.org/DPARSF> [Accessed 18 Sep. 2015].
- [57] Cox, R. W. AFNI: What a long strange trip it's been. *NeuroImage* , 743-747. (2012).
- [58] Talairach, J., Tournoux, P. Co-planar stereotaxic atlas of the human brain: 3-dimensional proportional system : an approach to cerebral imaging. Thieme. (1988).
- [59] Fmrib.ox.ac.uk, (2015). MELODIC - Probabilistic Independent Component Analysis for FMRI. [online] Available at: <http://www.fmrib.ox.ac.uk/analysis/techrep/tr04ss2/tr04ss2/node7.html> [Accessed 5 Aug. 2015].
- [60] Le Bihan, D., Mangin, J.-F., Poupon, C., Clark, C. A., Pappata, S. P., Molko, N. M., et al. Diffusion Tensor Imaging: Concepts and Applications. *Journal of Magnetic Resonance Imaging*. (2001).
- [61] Sundgren, P. C., Dong, Q., Gomez-Hassan, D., Mukherji, S. K., Maly, P., Welsh, R. Diffusion tensor imaging of the brain: review of clinical applications. *Neuroradiology*, 46(5), 339-350. (2004).
- [62] Med.inria.fr, (2015). medInria - multi-platform medical image processing and visualization software. [online] Available at: <https://med.inria.fr/> [Accessed 9 Aug. 2015].

- [63] Trackvis.org, (2015). TrackVis. [online] Available at: <http://trackvis.org/> [Accessed 2 Sep. 2015].
- [64] Gaain.org, (2014). GAAIN Architecture. [online] Available at: <http://gaain.org/platform/architecture/> [Accessed 8 Oct. 2015].
- [65] Marcus, D., Olsen, T., Ramaratnam, M. and Buckner, R. The extensible neuroimaging archive toolkit. *Neuroinform*, 5(1), pp.11-33. (2007).
- [66] J. Blazewicz, P. Dell'Olmo, M. Drozdowski Scheduling of client-server applications *International Journal in Operations Research*, 6 (1999).
- [67] Adler, R.M., "Distributed coordination models for client/server computing," *Computer*, vol.28, no.4, pp.14,22, Apr 1995
- [68] C. Smith, P. Duchessi The initiation and adoption of client/server technology in organizations *Information and Management*, 35 (1998).
- [69] D.S. Touceda, J.M.S. Camara, M. Soriano Decentralized certification scheme for secure admission in on-the-fly peer-to-peer systems *Peer-to-Peer Netw. Appl.*, 5 (2) (2012).
- [70] Jin Li, "On peer-to-peer (P2P) content delivery", Springer Science Business Media, LLC 2008.
- [71] Robin Jan Maly, "Comparison of Centralized (Client-Server) and Decentralized (Peer-to-Peer) Networking", ETH Zurich, Switzerland, March 2003.
- [72] Codd, E. A relational model of data for large shared data banks. *Communications of the ACM*, 13(6), pp.377-387. (1970).
- [73] Maier, David. *The theory of relational databases*. Vol. 11. Rockville: Computer science press, 1983.
- [74] Leavitt, N. Will NoSQL databases live up to their promise?. *Computer*, 43(2), 12-14. (2010).

- [75] The Power of Small Data.(2014) [online] Available at: <http://pt.slideshare.net/R2integrated/r2i-big-datasmalldata> [Accessed 8 Oct. 2015].
- [76] Moniruzzaman, Akhter Hossain. NoSQL Database: New Era of Databases for Big data Analytics - Classification, Characteristics and Comparison. International Journal of Database Theory and Application Vol. 6, No. 4. (2013).
- [77] R. Cattell. "Scalable SQL and NoSQL Data Stores,"ACM SIGMODRecord, vol. 39. 2010)
- [78] R.L .Villars, C.W. Olofson, M. Eastwood, Big data: what it is and why you should care, White Paper, IDC, 2011, MA, USA.
- [79] Hecht, R., Jablonski, S. NoSQL evaluation: A use case oriented survey. In Cloud and Service Computing (CSC), 2011 International Conference on (pp. 336-341). IEEE. (2011, December).
- [80] Konstantinou, I., Angelou, E., Boumpouka, C., Tsoumakos, D., Koziris, N. On the elasticity of nosql databases over cloud management platforms. In Proceedings of the 20th ACM international conference on Information and knowledge management (pp. 2385-2388). ACM. (2011, October)
- [81] J. Gray and A. Reuter. Transaction Processing: Concepts and Techniques. Morgan Kaufmann (1993).
- [82] Eric Brewer, CAP twelve years later: How the "rules"have changed, IEEE Explore, Volume 45, Issue 2 (2012), pg. 23-29.
- [83] Maxivak.com. RDBMSs vs. NoSQL Databases: Overview : Max Ivak Personal Site. [online] Available at: <http://maxivak.com/rdbms-vs-nosql-databases/> [Accessed 8 Aug. 2015].
- [84] Charles, R. The question of database transaction processing: an ACID, BASE, NoSQL primer. Dataversity Educationn, LCC (2013).

- [85] Ippolito, Bob: Drop ACID and think about Data. March 2009. Talk at Pycon on 2009-03-28. <http://blip.tv/file/1949416/>.
- [86] Chang, Fay, et al. "Bigtable: A distributed storage system for structured data." *ACM Transactions on Computer Systems (TOCS)* 26.2 (2008).
- [87] Stonebraker, M. SQL databases v. NoSQL databases. *Communications of the ACM*, 53(4), p.10. (2010).
- [88] G. DeCandia, et al., "Dynamo: amazon's highly available key-value store," in *SOSP '07 Proceedings of twenty-first ACM SIGOPS*, New York, USA, pp. 205-220. (2007).
- [89] Padhy, R. P., Patra, M. R., Satapathy, S. C. RDBMS to NoSQL: Reviewing Some Next-Generation Non-Relational Database. *International Journal of Advanced Engineering Science and Technologies*, 11(1), 15-30. (2011).
- [90] Han, J., Haihong, E., Le, G., Du, J. Survey on nosql database. In *Pervasive Computing and Applications (ICPCA), 2011 6th International Conference on* (pp. 363-366). IEEE. (2011, October).
- [91] [Mongodb.com](https://www.mongodb.com), (2015). MongoDB Architecture | MongoDB. [online] Available at: <https://www.mongodb.com/mongodb-architecture> [Accessed 8 Oct. 2015].
- [92] [Mongodb.com](https://www.mongodb.com), (2015). Building MongoDB Applications with Binary Files Using GridFS: Part 2 | MongoDB. [online] Available at: <https://www.mongodb.com/blog/post/building-mongodb-applications-binary-files-using-gridfs-part-2> [Accessed 8 Oct. 2015].
- [93] T. Dierks, E. Rescorla. "The Transport Layer Security (TLS) Protocol, Version 1.2". (August 2008).
- [94] Paulsen JS, Zimbelman JL, Hinton SC, Langbehn DR, Leveroni CL, Benjamin ML, Reynolds NC, Rao SM. fMRI biomarker of early neuro-

- nal dysfunction in presymptomatic Huntington Disease. *AJNR Am J Neuroradiol.* 2004.
- [95] Healthcare.siemens.com, (2015). MAGNETOM Avanto Upgrade - Siemens Healthcare Global. [online] Available at: <https://www.healthcare.siemens.com/magnetic-resonance-imaging/options-and-upgrades/upgrades/magnetom-avanto-upgrade> [Accessed 5 Oct. 2015].
- [96] Weishaupt, D., Koachli, V. D., Marincek, B. How does MRI work?: An introduction to the physics and function of magnetic resonance imaging. Berlin: Springer. (2006).
- [97] Pham, D. L., Xu, C., Prince, J. L. Current Methods in Medical Imaging Segmentation. *Annu. Rev. Biomed. Eng* , 315-337. (2000).
- [98] Andrews T, Halpern S, Purves D. Correlated size variations in human visual cortex, lateral geniculate nucleus, and optic tract. *J Neuroscience* (1997).
- [99] Caviness, V. S., Lange, N. T., Makris, N., R., H. M., Kennedy, D. N. MRI-based brain volumetrics: emergence of a developmental brain science. *Brain Development.* (1999).
- [100] Armstrong, E. Relative brain size and metabolism in mammals. *Science.* (1983).
- [101] Fjell, A., Walhovd, K., Reinvang, I., Lundervold, A., A.M., D., Quinn, B., et al. Age does not increase rate of forgetting over weeks—neuroanatomical volumes and visual memory across the adult life-span. *J Int Neuropsychol Soc* , 2-15.(2005).
- [102] Wright IC, McGuire PK, Poline JB, Traverso JM, Murray RM, Frith CD, Frackowiak RS, Friston KJ. A voxel-based method for the statistical analysis of gray and white matter density applied to schizophrenia. *Neuroimage.* (1995)

- [103] Good CD, Johnsrude IS, Ashburner J, Henson RN, Friston KJ, Frackowiak RS. A voxel-based morphometric study of ageing in 465 normal adult human brains. *Neuroimage*. (2001).
- [104] Ashburner, J., Friston, K. J. Voxel-based morphometry—the methods. *NeuroImage*, 11(6 Pt 1), 805-821. (2000).
- [105] Salmond CH, Ashburner J, Vargha-Khadem F, Connelly A, Gadian DG, Friston KJ. Distributional assumptions in voxel-based morphometry. *Neuroimage*. (2002).
- [106] Whitwell JL, Jack CR Jr. Comparisons between Alzheimer disease, frontotemporal lobar degeneration, and normal aging with brain mapping. *Top Magn Reson Imaging*. (2005).
- [107] Whitwell JL, Josephs KA. Voxel-based morphometry and its application to movement disorders. *Parkinsonism Relat Disord* 3(Suppl 13). (2007)
- [108] Keller SS, Roberts N. Voxel-based morphometry of temporal lobe epilepsy: an introduction and review of the literature. *Epilepsia*. (2008)
- [109] Prinster A, Quarantelli M, Orefice G, Lanzillo R, Brunetti A, Mollica C, Salvatore E, Morra VB, Coppola G, Vacca G, Alfano B, Salvatore M. Grey matter loss in relapsing-remitting multiple sclerosis: a voxel-based morphometry study. *Neuroimage*. (2006).
- [110] Williams LM. Voxel-based morphometry in schizophrenia: implications for neurodevelopmental connectivity models, cognition and affect. *Expert Rev Neurother*. (2008).
- [111] Querbes O, Aubry F, Pariente J, Lotterie JA, Duret V, Puel M, Berry I, Fort JC, Celsis P. Early diagnosis of Alzheimer's disease using cortical thickness: impact of cognitive reserve., *Alzheimer's Disease Neuroimaging Initiative Brain*. 2009 Aug; 132(Pt 8):2036-47.

- [112] Salat DH, Buckner RL, Snyder AZ, Greve DN, Desikan RS, Busa E, Morris JC, Dale AM, Fischl B. Thinning of the cerebral cortex in aging. *Cereb Cortex*. 2004 Jul; 14(7):721-30.
- [113] Hutton, C., Draganski, B., Ashburner, J., Weiskopf, N. A comparison between voxel- based cortical thickness and voxel-based morphometry in normal aging . *NeuroImage* , 371- 380. (2009).
- [114] Kim JS, Singh V, Lee JK, Lerch J, Ad-Dab'bagh Y, MacDonald D, Lee JM, Kim SI, Evans. Automated 3-D extraction and evaluation of the inner and outer cortical surfaces using a Laplacian map and partial volume effect classification. *ACNeuroimage*. 2005 Aug 1; 27(1):210-21.
- [115] Lee J, Ehlers C, Crews F, et al. Automatic cortical thickness analysis on rodent brain. *Proceedings - Society of Photo-Optical Instrumentation Engineers*. 2011.
- [116] Buxton R, Frank L. A model for the coupling between cerebral blood flow and oxygen metabolism during neural stimulation. *J Cereb Blood Flow Metab*. 1997; 17:64.
- [117] Amaro Jr, E., Barker, G. J. Study design in fMRI: Basic Principles. *Brain and Cognition* , 220-232. (2006).
- [118] Bullmore, E., Sporns, O., Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat. Rev. Neurosci*. 10 (3). 2009.
- [119] Greicius, M., Resting-state functional connectivity in neuro- psychiatric disorders. *Curr. Opin. Neurol*. 21 (4). 2008.
- [120] Damoiseaux, J., Rombouts, S., Barkhof, F., Scheltens, P., Stam, C., Smith, S., et al. Consistent resting-state networks across healthy subjects . *Proc Natl Acad Sci USA* , 13848- 13853. (2006).
- [121] Liao, W., Chen, H., Feng, Y., Mantini, D., Gentili, C., Pan, Z., et al. Selective aberrant functional connectivity of resting state networks in social anxiety disorder . *NeuroImage* , 1549- 1558. (2010).

- [122] Worsley, K. J., Liao, C. H., Aston, J., Petre, V., Duncan, G. H., Morales, F., Evans, A. C. A general statistical analysis for fMRI data. *NeuroImage*, 15(1), 1-15. (2002).
- [123] Lindquist, M. A. The Statistical Analysis of fMRI Data. *Statistical Science*, 23(4), 439-464. (2008).
- [124] Josephs, O., Turner, R., et al., Event-related fMRI. *Hum. Brain Mapp.* 5 (4). 1997.
- [125] Biswal, B., Yetkin, F. Z., Haughton, V. M., Hyde, J. S. Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magnetic Resonance in Medicine: Official Journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine*, 34(4), 537-541. (1995).
- [126] Shehzad, Z., Kelly, A. M. C., Reiss, P. T., Gee, D. G., Gotimer, K., Uddin, L. Q., Lee, S. H., et al. The resting brain: unconstrained yet reliable. *Cerebral Cortex (New York, N.Y.: 1991)*, 19(10), 2209-2229. (2009).
- [127] Calhoun, V.D., Adali, T., Pearlson, G.D., Pekar, J.J., A method for making group inferences from functional MRI data using independent component analysis. *Hum. Brain Mapp.* 14 (3). 2001.
- [128] De Luca, M., Beckmann, C.F., De Stefano, N., Matthews, P.M., Smith, S.M., fMRI resting state networks define distinct modes of long-distance interactions in the human brain. *Neuro- image* 29 (4). 2006.
- [129] Bell, A. J., Sejnowski, T. J. An information-maximization approach to blind separation and blind deconvolution. *Neural Computation*, 7, 1129–1159. (1995).
- [130] Hyvarinen, A., Oja, E. A Fast Fixed-Point Algorithm for Independent Component Analysis. *Neural Computation*, 9, 1483–1492. (1997).

- [131] Fox, M. D., Raichle, M. E. Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nat Rev Neurosci*, 8(9), 700-711. (2007).
- [132] van den Heuvel, M.P., Hulshoff Pol, H.E., Exploring the brain network: A review on resting-state fMRI functional connectivity, *Eur. Neuropsychopharmacol.* (2010).
- [133] Achard, S., Bullmore, E., Efficiency and cost of economical brain functional networks. *PLoS Comput. Biol.* 3 (2), e17. 2007.
- [134] Achard, S., Salvador, R., Whitcher, B., Suckling, J., Bullmore, E., A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs. *J. Neurosci.* 26 (1). 2006.
- [135] Liu, Y., Liang, M., Zhou, Y., He, Y., Hao, Y., Song, M., Yu, C., Liu, H., Liu, Z., Jiang, T., Disrupted small-world networks in schizophrenia. *Brain* 131 (4), 945. 2008.
- [136] Eguiluz, V.M., Chialvo, D.R., Cecchi, G.A., Baliki, M., Apkarian, A.V., Scale-free brain functional networks. *Phys. Rev. Lett.* 94 (1), 018102. 2005.
- [137] Buckner, R.L., Sepulcre, J., Talukdar, T., Krienen, F.M., Liu, H., Hedden, T., Andrews-Hanna, J.R., Sperling, R.A., Johnson, K.A., Cortical hubs revealed by intrinsic functional connectivity: mapping, assessment of stability, and relation to Alzheimer's disease. *J. Neurosci.* 29 (6). 2009.
- [138] Moseley, M. E., Kucharczyk, J., Asgari, H. S., Norman, D. Anisotropy in diffusion-weighted MRI. *Magnetic Resonance in Medicine: Official Journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine*, 19(2), 321-326. (1991)

- [139] Makris, N., Papadimitriou, G. M., Worth, A. J., Jenkins, B. G., L., G., Sorensen, A. G., et al. Diffusion Tensor Imaging. *Neuropsychopharmacology: The Fifth Generation of Progress* , 357-372. (2002).
- [140] Pierpaoli, C., Jezzard, P., Basser, P. J., Barnett, A., Di Chiro, G. Diffusion tensor MR imaging of the human brain. *Radiology*, 201(3), 637-648. (1996).
- [141] Mukherjee, P., Chung, S. W., Berman, J. I., Hess, C. P., Henry, R. G. Diffusion tensor MR imaging and fiber tractography: technical considerations. (2008).
- [142] Tournier, J., Mori, S., Leemans, A. Diffusion tensor imaging and beyond. *Magnetic Resonance in Medicine*, 65(6), 1532-1556. (2011).
- [143] Conturo, T. E., Lori, N. F., Cull, T. S., Akbudak, E., Snyder, A. Z., Shimony, J. S., McKinstry, R. C., et al. Tracking neuronal fiber pathways in the living human brain. *Proceedings of the National Academy of Sciences*, 96(18), 10422 -10427. (1999).
- [144] Correia, S., Lee, S. Y., Voorn, T., Tate, D. F., Paul, R. H., Zhang, S., Salloway, S. P., et al. Quantitative tractography metrics of white matter integrity in diffusion-tensor MRI. *NeuroImage*, 42(2), 568-581. (2008).
- [145] Smith, S. M., Jenkinson, M., Johansen-Berg, H., Rueckert, D., Nichols, T. E., Mackay, C. E., Watkins, K. E., et al. Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. *NeuroImage*, 31(4), 1487-1505. (2006).

# Apêndice A

## Ressonância Magnética

### A.1 Princípios de funcionamento

Entre os órgãos do corpo humano, o cérebro destaca-se como um dos mais fascinantes e, simultaneamente, intrigante. Essencial à vida e ao controlo do funcionamento de todo o corpo, é também um dos mais difíceis de compreender e estudar na sua função e estrutura. Tal se deve não só à dificuldade de acesso como também à complexa interação entre todos os seus elementos na realização de tarefas sensoriais, motoras e cognitivas. A neuroimagem é a disciplina responsável pela investigação da estrutura e respectivas funções do cérebro, utilizando técnicas como [Positron Emission Tomography](#), [Computed Tomography](#), [Electroencephalography](#) e [Magnetic Resonance Imaging](#) [10].

Nas últimas décadas, a técnica de [MRI](#) tem-se tornado uma das mais relevantes modalidades de neuroimagem, contribuindo para tal vantagens inerentes à própria técnica. Além de ser uma técnica não invasiva, não apresenta efeitos adversos conhecidos comparativamente a [CT](#), que utiliza radiação ionizante [47]. As principais vantagens de [MRI](#) prendem-se com o facto de proporcionar um bom contraste em tecidos moles e apresentar excelente resolução espacial. Desta forma é capaz de caracterizar e discriminar tecidos através das suas propriedades físicas e bioquímicas [15].

Imagens de ressonância magnética são obtidas através da colocação de um paciente ou área de interesse num campo magnético forte, estático e uniforme

gerado por um electroímã robusto. Durante um processo designado por *excitação* o sistema de MRI emite energia sob a forma de ondas rádio na frequência de ressonância dos prótons. Estas ondas rádio "empurram" os prótons para fora do seu alinhamento, absorvendo desta energia. Quando o sinal rádio é desligado, os prótons retornam ao seu alinhamento original, libertando no processo a energia anteriormente absorvida. Esta energia pode ser quantificada através de receptores electromagnéticos colocados à volta do sujeito. Um computador recebe os sinais de MRI como dados matemáticos compilando-os em imagens [94]. Na Figura A.1 é apresentado um sistema de MRI



Figura A.1: Siemens Magnetom Avanto (retirado de [95])

De um ponto de vista simplista, MRI é uma técnica espectroscópica que resulta na absorção selectiva e futura libertação de energia pelo núcleo de átomos, que quando sujeitos a um campo magnético intenso, apresentam diferente intensidade resultante da sua própria constituição. As imagens obtidas reflectem a constituição de “fatias” do objecto em estudo, com uma

determinada espessura. Essas porções podem, por sua vez, ser divididas em pequenas unidades volumétricas cuja constituição será refletida na intensidade do sinal emitido. Conseqüentemente, cada pixel reproduzido é determinado pela intensidade do sinal captado [96].

A abrangência de diferentes modalidades de MRI permite a realização de diferentes estudos com propósito de exploração de distintas características cerebrais. Estas modalidades dividem-se em dois grandes grupos: modalidades de estudo estrutural e funcional.

## A.2 Ressonância Magnética estrutural

*structural Magnetic Resonance Imaging* é uma poderosa ferramenta utilizada na pesquisa cerebral e neurologia clínica. Esta fornece um método não invasivo para obtenção de imagens com alta resolução da estrutura anatômica do cérebro bem como de patologias neurológicas. Muitas sequências de sMRI são volumétricas, o que significa que as medições podem ser realizadas em estruturas cerebrais específicas para cálculo de volumes de tecido. A capacidade de segmentar e classificar diferentes regiões a partir de sMRI tem grande importância em estudos de desenvolvimento cerebral, degeneração neuronal e na avaliação de doenças neurológicas e psicológicas [13].

A segmentação de imagens cerebrais é um dos passos mais importantes e desafiantes em estudos neurológicos, uma vez que estas contêm geralmente ruído, heterogeneidades e por vezes desfasamentos. O objetivo da segmentação é simplificar ou alterar a representação de uma imagem em algo mais significativo e fácil de analisar. Por sua vez, a seleção ótima de características (*features*), tecidos, elementos cerebrais e não cerebrais são considerados os principais obstáculos na segmentação de imagens cerebrais. A segmentação da estrutura cerebral de MRI possui uma importância primordial uma vez que esta técnica se distingue por si só de outras modalidades (excelente contraste de tecidos moles, alta resolução espacial e não invasiva), e pode ser aplicada em análises de volumetria de tecidos cerebrais, em doenças como esclerose múltipla, esquizofrenia, epilepsia, parkinson, alzheimer ou atrofia cerebral [16]. Independentemente do método utilizado, a segmentação tridi-

mensional de **MRI** cerebral apresenta-se, tal como referido, como um grande desafio. A quantidade de estruturas com diferente forma e aparência, em conjunto com a necessidade de alta precisão e qualidade dos resultados justificam a diversidade de métodos existentes. A existência de factores como ruído próprio da aquisição, artefactos e a presença de outros tecidos como o crânio dificultam a segmentação de imagens de **MRI**. Assim sendo, antes de aplicar qualquer método de segmentação é necessário aplicar alguns passos de pré-processamento. Estes não são exclusivos a análises de **sMRI**, sendo aplicados a outras modalidades de **MRI**.

Os métodos de segmentação podem ser divididos em 2 grandes grupos: métodos manuais/semi-automáticos e automáticos. O método de segmentação mais tradicional envolve a delimitação manual por parte de um especialista, que analisa cada imagem individualmente e realiza a segmentação. Compreensivamente, este método é muito ineficaz temporalmente, sendo usado apenas como uma referência ou *Gold-standard*. Métodos semi-automáticos necessitam da interação do utilizador no início ou durante o processo. Estes não são necessariamente utilizados de forma isolada podendo ser combinados para atingir melhores resultados. Métodos de segmentação automática, pretendem segmentar as imagens sem necessidade de interação do utilizador para definir regiões de interesse ou pré segmentações. Estes métodos são normalmente mais complexos e requerem a integração de várias ferramentas [97]. Na Figura A.2 são apresentadas imagens de segmentação cerebral segundo os atlas Destrieux e Desikan.

### A.2.1 Volumetria

A técnica de volumetria cerebral baseada em **MRI** é uma metodologia estabelecida de grande versatilidade e confiabilidade com um leque de aplicação alargado a nível clínico e em neurociências. Esta refere-se à organização e análise de relações quantitativas e qualitativas entre volume e forma dos componentes estruturais do cérebro. O fundamento teórico subjacente à volumetria cerebral assenta no conceito de que o volume é uma propriedade do cérebro, cujos componentes são regulados por processos de evolução e

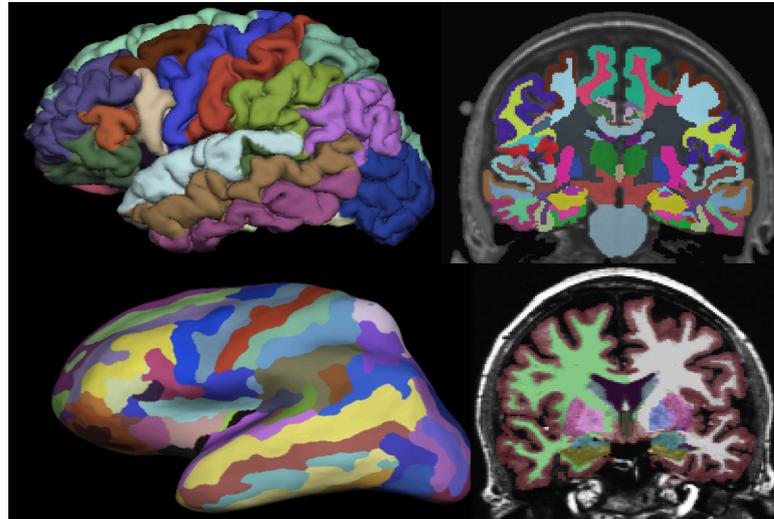


Figura A.2: *Output* do fluxo de processamento Freesurfer: imagens de segmentação cerebral segundo os atlas Destrieux e Desiken.

desenvolvimento. O volume de uma estrutura cerebral refere diretamente o tamanho, forma, padrão de arranjo e densidade dos seus diversos componentes celulares [98]. Neste sentido, o volume característico das diferentes estruturas está diretamente associado à sua capacidade de processar informação, refletindo o resultado do processo de evolução e devendo ser uniforme para indivíduos da mesma espécie. Assim, as características volumétricas serão uma manifestação direta dos processos histológicos e genéticos [99, 100].

Estudos em volumetria podem ser realizados em diferentes escalas, desde a análise a diferentes estruturas obtidas por esquemas de segmentação a análise da totalidade do córtex. Estes são essencialmente relevantes na exploração de doenças do foro neurológico como a esquizofrenia [17], alzheimer [18] e autismo [19], assim como em estudos relacionados com o envelhecimento do cérebro [101].

### A.2.2 Voxel-based Morphometry

VBM é uma técnica automática com grande popularidade desde a sua introdução [102], impulsionada pela relativa facilidade de utilização e obtenção de resultados. Nesta é utilizada estatística por forma a identificar diferenças

anatômicas cerebrais entre grupos de indivíduos, o que por sua vez pode ser utilizado para inferir a presença de uma atrofia ou, menos comumente, a expansão de tecidos em sujeitos com determinada doença. A modalidade de **VBM** permite a identificação de diferenças locais na densidade de substância cinzenta e substância branca, através de testes estatísticos que comparam todos os vóxeis de diferentes grupos [14]. Por forma a comparar diferentes grupos, o passo inicial na **VBM** consiste na segmentação do cérebro em substância cinzenta, branca e líquido cefalorraquidiano, nos quais são testadas as diferentes hipóteses. Posteriormente, é realizada normalização espacial [103], transformando todas as imagens para um template comum. Para que estas correspondam o mais fidedignamente ao template selecionado, as imagens são distorcidas. Caso a intensidade de cada vóxel seja mantida, esta refletirá a intensidade de substância cinzenta. Por outro lado, caso a intensidade de cada vóxel for modulado consoante o grau de distorção a que foi sujeito, as intensidades corresponderão a alterações volumétricas. Por último, as imagens são suavizadas [104] utilizando a média dos vóxeis vizinhos segundo um *kernel* de suavização. Este passo torna os dados adequados ao modelo Gaussiano, processando-os de forma normalmente distribuída, aumentando a validade de testes paramétricos e reduzindo a variabilidade inter-individual [105].

A análise estatística é tipicamente realizada aplicando **GLM** e a teoria Gaussiana. Esta técnica vem sendo aplicada a diferentes doenças, incluindo doenças neurodegenerativas [106], doenças motoras [107], epilepsia [108], esclerose múltipla [109], e esquizofrenia [110], contribuindo para a compreensão das alterações cerebrais ocorrentes nestas doenças e como este se relaciona com características clínicas específicas.

### A.2.3 Cortical thickness analysis

A espessura cortical (*Cortical thickness*) é uma característica morfológica de relevo em diversos estudos cerebrais, sendo uma das aplicações mais interessantes da análise estrutural de imagens de **MRI**. Esta medida é útil na avaliação do desenvolvimento de doenças como Alzheimer (conhecidas por causar atrofia em regiões corticais), na avaliação de tratamentos e no estudo

do desenvolvimento e envelhecimento cerebral [111,112].

Os métodos de medição de espessura cortical baseados em **MRI** podem ser classificados em dois grupos: *voxel-based methods* e *surface-based methods*. A **Voxel Based Cortical Thickness (VBCT)** análoga à **VBM**, estabelece fronteiras entre a substância cinzenta e branca usando informação dos vóxeis. A espessura é depois calculada dependendo da distância entre as diferentes fronteiras. O resultado será um mapa onde a cada vóxel na substância cinzenta é atribuído um valor de espessura. Tal como na **VBM** é possível utilizar normalização espacial para permitir comparar diferentes estudos [113]. A segunda metodologia baseia-se em técnicas de superfície. Esta utiliza informação da imagem e geometria de superfícies para criar modelos geométricos que correspondam à substância branca e cinzenta. A espessura do córtex é definida avaliando a distância entre pontos ou vértices de superfície [114]. De forma a colmatar problemas evidenciados por cada uma destas abordagens, surgiram métodos que resultam da combinação destas técnicas [115].

## A.3 Ressonância Magnética funcional

### A.3.1 Princípios básicos

O nosso cérebro é uma rede constituída por regiões espacialmente distribuídas, mas funcionalmente ligadas que partilham continuamente informação entre elas. A modalidade de **fMRI** é actualmente uma das técnicas mais utilizadas para o mapeamento do funcionamento do cérebro humano na área das neurociências. Os seus métodos podem medir a perfusão, volume ou oxigenação (neste, a técnica **BOLD** é a mais comum) do sangue no cérebro. São métodos conhecidos por possuírem excelente resolução espacial mas fraca resolução temporal. A **fMRI** tem sido utilizada para investigar e testar hipóteses da topografia funcional do cérebro e investigar mecanismos neurofisiológicos [52, 53], através da medição da resposta hemodinâmica cerebral para obter um mapeamento da atividade/função cerebral.

Do aumento de actividade cerebral resultam duas consequências primárias, detectadas com recurso a **MRI**: aumento do fluxo sanguíneo cerebral

e alterações na concentração de oxigenação sanguíneo. O método **BOLD** é sensível a variações na oxigenação do sangue e baseia-se nas diferentes propriedades magnéticas da oxi e desoxihemoglobina. Este segundo possui menor susceptibilidade magnética, criando perda de coerência na precessão dos prótons e, conseqüentemente, uma perda do sinal captado. Por outro lado, um aumento do sangue oxigenado traduz-se num acréscimo de intensidade do sinal medido. Durante a ativação sináptica ocorre um aumento na extração de oxigênio no sangue que provoca um aumento drástico localizado do fornecimento de sangue oxigenado, que por sua vez se irá traduzir num aumento de intensidade do sinal de **fMRI** [116].

Os estudos **fMRI** dividem-se em duas grandes áreas: estudos de repouso (*rsfMRI*), onde se caracterizam as dinâmicas que existem por defeito no cérebro, e estudos com estímulos (*task-related fMRI (trfMRI)*), onde se avalia a resposta do cérebro a estímulos específicos.

### A.3.2 Task-related fMRI

Estudos **trfMRI**, procuram induzir diferentes estados neuronais através de incentivos visuais, auditivos, ou de natureza reactiva, como responder a estímulos ou memorizar palavras. Terminado o procedimento de aquisição, as imagens são tratadas e processadas com o objetivo de identificar áreas cerebrais ativadas em consequência dos estímulos apresentados. Pretende-se assim obter mapas de activação por comparação de sinais registados durante os diferentes estados.

A construção, organização temporal e previsões de comportamento das tarefas executadas pelo sujeito designa-se de paradigma [117], classificados num de dois tipos: paradigma de blocos e paradigma de eventos. Num paradigma de blocos, o sujeito executa duas ou mais funções diferentes, de forma individual e repetitiva em blocos de tempo consideravelmente longos (20 a 30 segundos). Os níveis relativos de activação durante cada estímulo são então comparados entre si para obter indicações sobre a activação diferencial de uma região cortical. Estas repeticoes são intervaladas com períodos de repouso que estabelecem um nível de base da actividade neuronal.

Num paradigma de eventos, são apresentados estímulos de curta duração (1 a 5 segundos) de uma ou mais categorias com intervalos variáveis entre eles, a fim de medir a evolução no tempo da ativação de uma determinada região cortical. Assim, a vantagem de um paradigma de blocos é a localização das áreas funcionais de resposta, enquanto a principal vantagem de um paradigma de eventos é fornecer indicações sobre a natureza da resposta que se seguiu. O principal factor que possibilita a realização deste tipo de paradigmas em [MRI](#) e a rapidez de aquisição de imagens, impossível em técnicas como [PET](#) [117].

Terminado o procedimento de aquisição, as imagens são tratadas e processadas com o objetivo de identificar áreas cerebrais ativadas em consequência dos estímulos apresentados. Pretende-se assim obter mapas de activação por comparação de sinais registados durante os diferentes estados.

### A.3.3 Resting-state fMRI

As comunicações funcionais entre regiões cerebrais são prováveis de desempenhar um papel fundamental no complexo processo cognitivo, prosperando na contínua integração de informação ao longo das diferentes seções do cérebro. Curiosamente, cerca de 15 anos após a invenção de [fMRI](#), estudos começaram a examinar a possibilidade de avaliar a conectividade funcional entre regiões do cérebro como a co-activação de séries temporais, espontâneas e funcionais de [MRI](#), registadas durante o repouso.

A Conectividade funcional é definida como a dependência temporal de padrões neuronais activos de regiões cerebrais anatomicamente separadas [34]. . Examinar o cérebro humano como uma rede integrada de regiões neuronais funcionalmente interactivas pode fornecer conhecimento acerca da comunicação neuronal em grande escala. Esta abordagem providencia uma plataforma de correlação de como a conectividade funcional e integração de informação se relacionam com o comportamento humano, e como esta organização pode ser alterada como resultado de doenças neurodegenerativas [118, 119].

Na realização de aquisições [rsfMRI](#), voluntários são instruídos a relaxar

e não pensar em algo particular, enquanto o nível de actividade espontânea cerebral (flutuações de sinal) é registado. A associação de componentes de baixa frequência (entre 0.01 e 0.1 Hz) a conjuntos de estruturas da substância cinzenta com funções conhecidas, como as áreas motoras, visuais, linguagem e processamento auditivo, tem apoiado a teoria que estas flutuações estão associadas à actividade neuronal base do cérebro [120].

A actividade intrínseca em repouso é organizada espacialmente em padrões coerentes específicos denominados de Redes Neurais de Repouso (*Resting State Networks (RSN)*) [121]. As *RSN* refletem sistemas centrais de percepção e processamento cognitivo. Estas ilustram de forma consistente, por exemplo, padrões de conectividade funcional em núcleos do tálamo e cerebelo permitindo a investigação de ligações cortico-cerebelares e subcorticais em maior detalhe do que medidas de conectividade estrutural. Das *RSN* conhecidas, a *Default-mode network (DMN)* é a que tem recebido maior atenção da comunidade científica. Esta rede tem a particularidade de possuir um maior sinal *BOLD* em repouso do que durante a realização de tarefas, levando a hipótese de que esta se encontra activa durante descanso e suspensa quando tarefas específicas estão a ser realizadas [120]. Na Figura A.3 é possível observar *RSN* obtidas através de *ICA*.

### A.3.4 General Linear Model

O *GLM* é o método de análise mais utilizado em *fMRI* [122]. Nesta metodologia, as séries temporais são modeladas como uma combinação linear de diferentes componentes de sinal e é testada a hipótese da actividade numa região cerebral se encontrar sistematicamente relacionada com alguma destas funções [123].

Durante uma aquisição de *trFMRI*, um evento neuronal produz uma resposta hemodinâmica que dura mais de vinte segundos. Usando *GLM* para análise estatística, a série temporal de cada vóxel é adaptada a um modelo que se relaciona com as respostas neuronais esperadas [124]. Na sua versão mais simples, o modelo *GLM* assume que a função de estímulo e a resposta hemodinâmica (*Hemodynamic Response Function (HRF)*) são conheci-

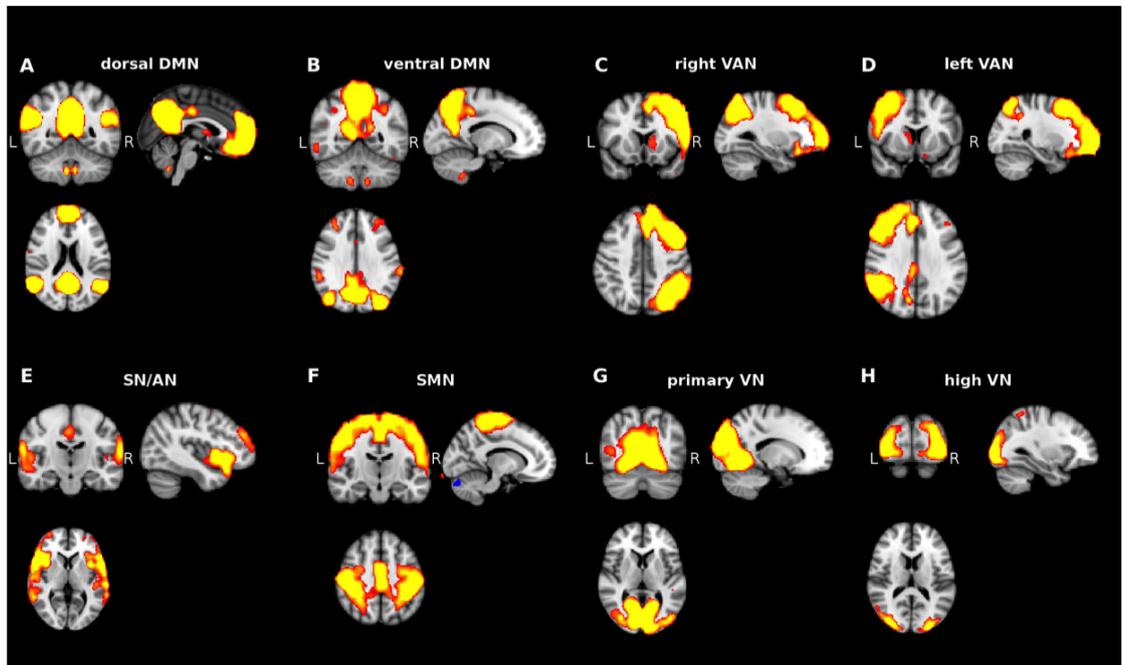


Figura A.3: Resting state networks obtidas através de análise de componentes independentes.

das. Pressupõe-se que o estímulo é equivalente ao paradigma experimental, enquanto que a *HRF* é modelada utilizando a *HRF* canónica. Da convolução do paradigma experimental com a *HRF* canónica resulta um conjunto de respostas esperadas que são introduzidas numa matriz de *design*. Este procedimento encontra-se esquematicamente representado na Figura A.4

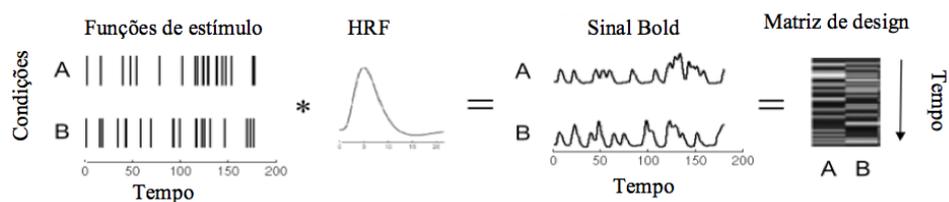


Figura A.4: Convulsão do sinal *HRF* com a matriz de respostas ao paradigma experimental que dá origem ao sinal funcional esperado. Adaptado de [123]

Apesar das potencialidades do , este método é, no entanto, muito rígido.

As conclusões são retiradas com base no modelo apresentado sem garantias que esse modelo seja o mais adequado. Além disso, esta análise assume que a matriz de design é a mesma para todo o cérebro sabendo-se, atualmente, que a **HRF** difere consoante a região cerebral.

### A.3.5 Seed Correlation Analysis

As análises de correlação **Seed Correlation Analysis (SCA)** são outro tipo de abordagem dependente de um modelo, baseada na seleção a priori de uma região (*seed*), na qual séries temporais são extraídas [125]. Assim, a metodologia mais simples para estabelecimento de conexões funcionais passa por correlacionar as séries temporais de repouso de uma determinada região cerebral de interesse com as séries temporais de cada vóxel cerebral.

A região de interesse (*seed*) pode ser definida a priori ou pode ser selecionada através de um mapa de activação *task-dependent* tradicional, obtido num ensaio **fMRI** separado. Posteriormente, as séries temporais *resting-state* da região selecionada pode ser correlacionada com as séries temporais de todos os outros voxéis, resultando num mapa de conectividade funcional (*fc-map*). O **functional connectivity Map (fcMap)** resultante providencia informação sobre as regiões funcionalmente relacionadas com a região selecionada, e qual a sua extensão.

A simplicidade deste tipo de análise aliado com a obtenção de resultados de forma direta, formam a principal vantagem de métodos *seed-dependent*. Os **functional connectivity Maps (fcMaps)** providenciam uma visão clara das regiões funcionalmente conectadas com a região de interesse, tornando-o um modo elegante de exploração de conectividade funcional no cérebro humano. Recentemente comprovou-se que este tipo de análise permite identificar as relações de conectividade das redes neuronais de repouso com elevado grau de confiança [126]. A desvantagem mais significativa prende-se com o facto da presença de ruído na região da *seed* poder influenciar negativamente a série temporal desta e, conseqüentemente, comprometer os resultados. Além disso, a localização da *seed*, o volume ou mesmo diferenças entre sujeitos podem introduzir graus de variabilidade elevados. Por outro lado, a informação de

um é limitada às conexões funcionais da região selecionada, tornando difícil a exploração de padrões de conectividade funcional a uma escala cerebral global.

### A.3.6 Independent Component Analysis

Por forma a examinar padrões de conectividade cerebral numa escala global, foram introduzidos métodos independentes de modelo, não existindo necessidade de definição de uma região (*seed*) a priori. Os métodos de ICA são porventura os mais consistentes e utilizados na identificação de RSN [127, 128]. No contexto das análises de conectividade funcional, estes métodos são utilizados com o objetivo de procurar uma mistura de fontes que permitam explicar os padrões das redes neuronais de repouso.

O pressuposto base desta análise é o de que as fontes de sinal são estatisticamente independentes com distribuições não Gaussianas. A determinação das componentes decorre num processo iterativo em que se procura maximizar o grau de independência entre as componentes, podendo ser utilizadas várias métricas diferentes para o efeito. Vários algoritmos podem ser empregues para o cálculo das componentes, contudo os mais utilizados em fMRI são o Infomax [129] e o FastICA [130]

As análises ICA podem classificar-se em espaciais e temporal, consoante se considera que as fontes de sinal são independentes no espaço ou no tempo. Estes são métodos poderosos, uma vez que podem ser aplicados a dados da totalidade do cérebro. Além disso, os sinais temporais das componentes relevantes podem ser facilmente selecionados para análises de grupo entre controlos e pacientes. A principal desvantagem destes métodos é a maior dificuldade de compreensão das componentes independentes geradas, comparativamente aos fcMaps gerados com regiões *seed*, por se tratarem de representações mais complexas dos dados [131]

### A.3.7 Análise de grafos

Novos avanços em técnicas de análise de rsfMRI possibilitaram a exploração da estrutura global da rede cerebral, com grande detalhe espacial,

utilizando métodos analíticos de grafos. Teoria de grafos providencia um ponto de vista teórico no qual a topologia de redes complexas pode ser examinada, e pode revelar informação importante acerca da organização local e global de redes cerebrais funcionais. Recorrendo a teoria de grafos, redes cerebrais funcionais podem ser definidas como um grafo  $G = (V, E)$ , com  $V$  a ser a coleção de nós refletindo as regiões cerebrais, e  $E$  as conexões cerebrais entre essas regiões [132]. Neste tipo de abordagem, os nós da rede cerebral podem ser representados como regiões sub-corticais, podendo ser um pequeno número de regiões cerebrais de larga escala num modelo cortical pré-definido, voxéis *fMRI*, ou uma abordagem híbrida.

Por outro lado, o nível de conectividade funcional entre 2 regiões é obtido como o nível de correlação entre as séries temporais das duas regiões cerebrais. O cálculo do nível de conectividade funcional entre todos os pares possíveis e a determinação da sua existência, resulta numa representação gráfica da rede funcional e permite a exploração da sua organização recorrendo a teoria de grafos [133].

Estudos pioneiros aplicaram técnicas de análise de grafos avançada a dados de *rsfMRI* [134–136], revelando novas descobertas acerca da organização geral de redes cerebrais funcionais. Estes estudos, revelaram que a rede cerebral é organizada segundo uma distribuição eficiente *small-world*. Estas redes são conhecidas por um alto nível de conectividade local, com distâncias curtas entre nós. Análise de grafos em *rsfMRI* revelaram um numero elevado de *hub-regions* altamente conectadas, sendo sugerido que estas regiões especializadas podem ser afectados pela doença de Alzheimer, resultando num decréscimo de eficiência funcional nestes pacientes [137].

De um ponto de vista geral, estudos de *rsfMRI* recorrendo a análise de grafos sugerem uma organização eficiente na comunicação funcional cerebral, indicando que o cérebro humano não é uma rede aleatória, mas sim uma rede optimizada para um grande nível de eficiência local e global.

Na Figura A.5 é apresentada a vista sagital de um conectoma funcional.

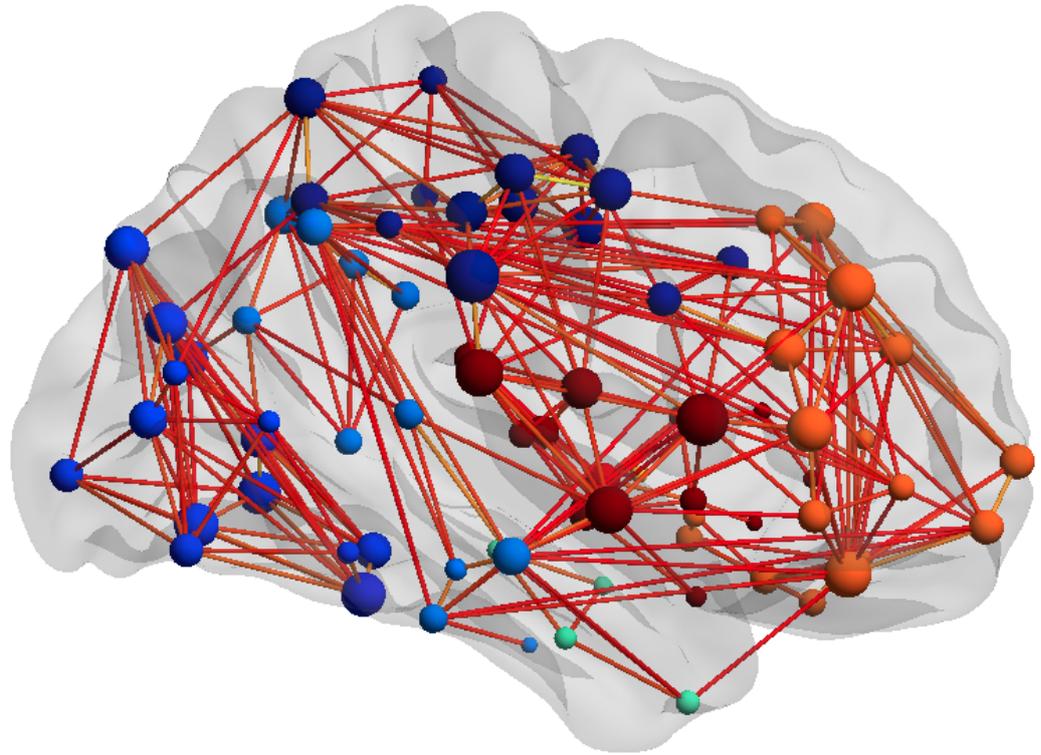


Figura A.5: Vista sagital de um conectoma funcional. Cores dos nós identificam o módulo a que pertencem e cor das ligações reflete a intensidade da ligação.

## A.4 Ressonância Magnética de difusão

A ressonância magnética de difusão teve o seu início nos anos 80 com o intuito de permitir estudos funcionais do cérebro, assentando no pressuposto de que poderia ser utilizada para avaliar o fluxo sanguíneo cerebral. É no entanto um padrão de avaliação de alterações na substância branca cerebral, acompanhamento de vítimas de acidente vascular cerebral e criação de mapas de ligação cerebrais [60].

A imagem por Tensor de Difusão (DTI) baseia-se no princípio de difusão das moléculas de água, um processo físico que ocorre naturalmente em condições de equilíbrio térmico, caracterizado por movimentos moleculares aleatórios. Num meio livre de barreiras aos movimentos moleculares, o processo

de difusão das partículas de água diz-se isotrópico. Nestas condições, o deslocamento probabilístico das moléculas de água ao longo de um determinado período de tempo é igual para todas as direções. Em estruturas biológicas como os ventrículos cerebrais e a substância cinzenta em geral existe pouca restrição ao movimento das moléculas de água e, naturalmente, a difusão da água nestas estruturas é altamente isotrópica (uniforme em todas as direções). Por outro lado, na substância branca, devido ao arranjo específico das fibras axonais, a difusão é altamente anisotrópica [138]. É o movimento preferencial das moléculas de água ao longo das estruturas fibrosas que torna a **MRI** de difusão uma modalidade imagiológica particularmente útil no estudo destas estruturas.

O estudo da difusão de moléculas de água *in vivo* tornou o **DTI** uma ferramenta singular no estudo de várias patologias incluindo esclerose múltipla, dislexia, derrames cerebrais e esquizofrenia [61].

#### A.4.1 Aquisição e Tensor de difusão

Na presença de um campo magnético com gradiente, a movimentação das moléculas de água origina o desfasamento do sinal de **MRI**, produzindo uma atenuação  $A$  caracterizada por  $A = \exp(-bD)$ , que é proporcional à magnitude de difusão da água na direção do gradiente aplicado [139]. A variável  $b$  caracteriza o gradiente de pulsos aplicado e  $D$  é um tensor simétrico 3x3 que descreve a difusão nas várias direções.

Na construção da imagem de são necessárias, no mínimo, imagens do efeito de difusão ao longo de seis direções diferentes, em relação a um sistema de eixos (orientações dos eixos  $x, y$  e  $z$ , tal como  $xy, xz$  e  $yz$ ). Em conjunto com estas é utilizada uma imagem base sem qualquer informação de difusão ( $b=0$ ). Das imagens ponderadas são calculados os diferentes componentes do tensor de difusão  $D$ . Com o tensor de difusão podem ser calculadas a anisotropia fraccional, a difusividade média e a direção de difusão para cada vóxel, obtida pela diagonalização do tensor de difusão e representadas pelos vetores ( $e_1, e_2$  e  $e_3$ ) e valores próprios (1, 2 e 3), que irão corresponder às direções e magnitudes principais de difusão. [60].

Para a representação visual da informação do tensor para cada vóxel são utilizadas elipsóides como evidenciado na Figura A.6, que representam a distância percorrida por difusão em cada uma das direções, sendo o seu eixo principal, o eixo onde ocorre maior difusão.

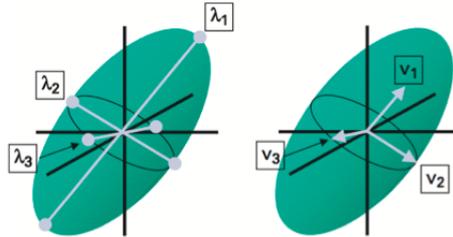


Figura A.6: Elipsóides que resultam da combinação dos vetores (a) e valores próprios (b) obtidos na análise de tensores de difusão. Representação gráfica da anisotropia fracional.

#### A.4.2 Visualização de dados de DTI

A elevada dimensionalidade dos dados e as relações complexas no domínio do tensor de difusão são os maiores desafios que se colocam à utilização da modalidade de DTI. Consequentemente várias foram as alternativas propostas ao longo dos últimos anos para facilitar a visualização de informação contida no tensor, variando em complexidade e no contexto onde são usadas.

Na área imagiológica, os profissionais de saúde, em especial, os radiologistas têm muita prática na visualização de índice escalares em imagens bidimensionais. Trata-se de uma forma de visualização clara, simples e de fácil interpretação embora a quantidade de informação que transmite seja limitada. Consequentemente, a visualização de índices escalares tem sido muito utilizada, em especial como método complementar à visualização multidimensional dos dados. Após a determinação dos vectores e valores próprios que permitem descrever os tensores, um vasto leque de índices escalares podem ser calculados.

Uma das métricas mais simples e conhecida é a difusividade média ( $MD$ ), e é uma quantidade independente de direções e obtida utilizando os três

valores próprios do tensor de difusão para calcular um valor médio. Outras duas medidas utilizadas são a anisotropia relativa (*Relative Anisotropy (RA)*) e a anisotropia fraccional (*FA*). *FA* mede a fração do tensor que pode ser atribuído a difusão anisotrópica e assume valores entre 0 (completamente isotrópica) e 1 (infinitamente anisotrópica). *RA* representa o rácio entre as partes anisotrópicas e isotrópicas, sendo um desvio padrão normalizado [20,140].

Apesar de ser facilmente interpretável, um mapa escalar permite apenas representar uma característica do tensor de difusão. Contudo é do interesse geral que toda a informação do tensor de difusão possa ser visualizada numa única representação. Para tal foi introduzido o método de visualização por glifos [141]. Os glifos são objetos gráficos parametrizados que descrevem completamente um tensor de difusão através e.g. do seu tamanho, forma, cor, textura, localização. Uma vez que os glifos possuem normalmente 6 ou mais graus de liberdade, o tensor de difusão pode ser descrito através deste método [140].

### A.4.3 Tratografia

A constante evolução dos meios computacionais disponíveis veio facilitar a visualização tridimensional dos dados de *DTI*. Neste âmbito, a tratografia (*fiber tracking* ou *tractography*) é a técnica mais utilizada. A tratografia consiste em estimar trajetórias tridimensionais para estruturas anisotrópicas, como as fibras da substância branca, a partir da orientação dos tensores de cada vóxel [60]. Esta técnica é constituída por três fases distintas [142]: **seeding**, consiste na definição dos pontos iniciais (*seeds*) a partir do qual são geradas as fibras; **propagação**, geração gradual de fibras através de um algoritmo específico; *terminação*, propagação terminada quando um critério de paragem é alcançado, evitando assim que sejam geradas representações das fibras onde a robustez do campo vetorial não é garantido.

Existem dois tipos de metodologias possíveis para calcular a orientação dos tensores: determinísticas e probabilísticas. Métodos determinísticos, iniciam o processo com a escolha de um conjunto de *seedpoints* e regiões de

interesse (*roi*), que indicam os locais onde o processo se irá iniciar e o espaço a que estão limitadas. A partir destes pontos é feita uma reconstrução vóxel a vóxel utilizando informação do tensor e adaptando modelos matemáticos para estabelecer um caminho. Apesar dos resultados clínicos úteis que apresenta, estes métodos não permitem criar vários ramos a partir de um vóxel e não apresentam medida de confiança do resultado obtido [143].

Os métodos probabilísticos constroem, a partir de um *seedpoint*, uma grande quantidade de caminhos possíveis. Embora existam diversas abordagens, todas elas produzem um mapa que quantifica o número de caminhos que passam em cada vóxel, gerados a partir de um *seedpoint*. Neste sentido, representam a probabilidade de ligação entre diferentes vóxeis, apresentando sempre um grau de confiança para cada caminho obtido [20] Na Figura A.7 é possível verificar o resultado de aplicação de algoritmo de tratografia probabilística do software Trackvis.

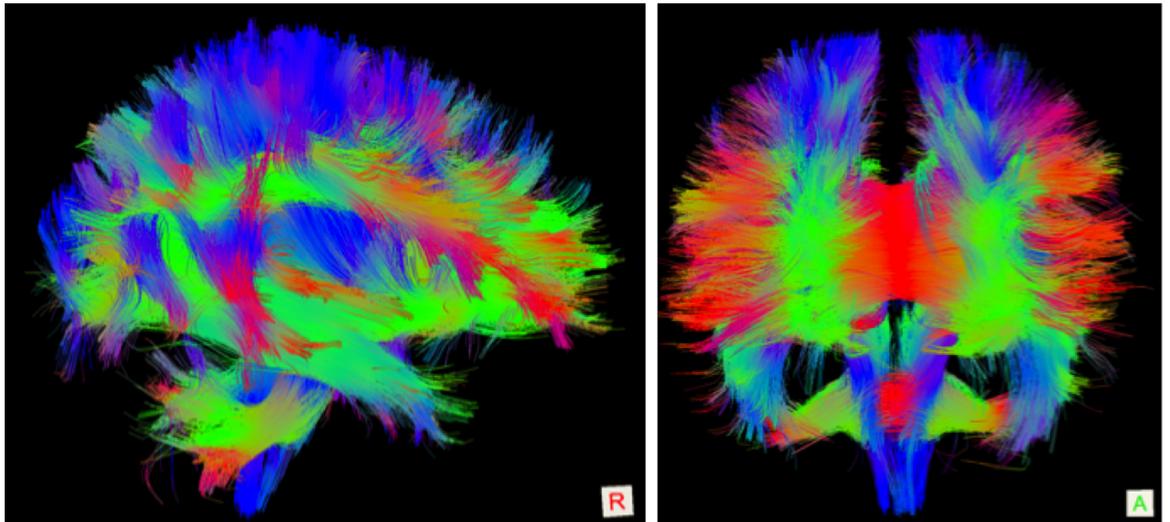


Figura A.7: Vistas sagital (a) e coronal (b) do resultado da aplicação de algoritmo de tratografia probabilística do Trackvis.

#### A.4.4 Análise quantitativa

Não obstante da relevância de técnicas de visualização de dados de DTI intuitivas como o mapas de índices escalares e a tratografia, na maior parte dos casos, os resultados não são óbvios pela simples observação visual. Neste sentido técnicas para quantificar analiticamente esses resultados foram introduzidas, tais como a abordagem por ROIs, VBM, tratografia quantitativa e TBSS.

A abordagem por ROIs é popular quando o objectivo de um estudo envolve a análise de uma região anatômica conhecida, ou seja, quando se pretende testar uma hipótese, definida a priori, envolvendo uma região bem delimitada. Assim, após delimitação de uma região cerebral específica incidirá a análise quantitativa. Posteriormente vários tipos de métricas (MD e FA, por exemplo) podem ser extraídas como forma de caracterizar o processo de difusão em estudo [142]. Esta técnica é, no entanto, dispendiosa temporalmente e requer grande experiência e conhecimento anatômico cerebral por parte do técnico ou investigador.

A abordagem VBM segue o princípio oposto à abordagem por ROIs, permitindo o estudo exploratório do volume cerebral como um todo através de uma análise baseada nos vóxeis, sem ser necessária uma hipótese inicial. [104]. Não sendo necessárias hipóteses definidas à priori, tipicamente este tipo de análise envolve as seguintes etapas:

- normalização das imagens para um espaço comum de referência, para que análise não seja comprometida por variabilidade anatômica entre sujeitos;
- suavização espacial para que a informação de elevada frequência espacial não interfira na análise;
- segmentação (opcional mas geralmente utilizada) para produzir mapas independentes para a substância branca, substância cinzenta e fluido cerebroespinal;
- análise estatística dos dados vóxel-a-vóxel.

Tratando-se de uma metodologia automatizada, os erros devido à ação do técnico/investigador são reduzidos. De salientar contudo que, no caso do processo de normalização não ser efetuado com precisão, os resultados podem ser comprometidos pela variabilidade entre sujeitos.

Nos métodos de tratografia quantitativa pretende-se a partir dos dados dos tensores, traçar linhas a partir dos dados dos tensores com o intuito de representar os feixes de substância branca subjacentes. Das abordagens deste tipo, uma das mais comuns é a análise por feixes de interesse (*Tracts of Interest (TOI)*). Posteriormente à seleção do TOI, as métricas escalares podem então ser extraídas dos vóxeis que compõe esse TOI, permitindo descrever o processo de difusão nesse feixe [144]. O aspecto mais positivo desta técnica é o facto de permitir a seleção de feixes individuais que, através da utilização de escalares, seriam de difícil localização devido à sua geometria complexa. Por outro lado, tal como na seleção de ROIs, a seleção de TOI requer alguém experiente para a isolar corretamente o feixe de interesse.

A modalidade de TBSS é um tipo de análise vóxel-a-vóxel que permite estudar feixes de grandes dimensões do cérebro humano em estudos de grupo [145]. O pressuposto principal para a utilização desta técnica é o de que o valor máximo de FA surge no centro dos feixes de matéria branca e um modelo de *skeleton* (estrutura que serve de referência/suporte) tridimensional pode ser construído a partir destes valores de FA. A etapa inicial da análise consiste num alinhamento não linear dos mapas de FA de cada indivíduo. É então criado um mapa de FA com a média dos outros valores de FA que é depois utilizado para gerar o *skeleton* referido. Posteriormente, os mapas cerebrais individuais são alinhados com o *skeleton* através de um alinhamento local dos valores máximos de FA dos feixes de cada sujeito com o *skeleton*. O mesmo processo se aplica-se a outros mapas escalares. O mapeamento de máximos locais com skeletons torna a análise TBSS única, permitindo análises de grupo vóxel-a-vóxel e não sendo necessárias mais etapas de normalização espacial.



# Apêndice B

## Nomenclatura e descrição de arquivos

De seguida pretende-se apresentar o dicionário de nomenclatura de arquivos, criado através do fluxo de processamento multimodal apresentado na [Seção 2.6](#). Nas Figuras [2.3](#), [2.4](#) e [2.5](#) presentes na [Seção 2.6](#), é respectivamente detalhado o fluxo de processamento para aquisições funcionais, de difusão e estruturais. Para cada uma destas são apresentados os arquivos essenciais, resultantes dos mais diversos passos. Nesse sentido, cada passo foi enumerado e corresponderá nas Figuras de seguida apresentadas, aos momentos de produção dos arquivos respectivos.

Assim para cada arquivo será apresentada a nomenclatura a utilizar, a designação do mesmo e que será utilizada como metadado na aplicação BrainArchive, o software utilizado na sua produção e a *tag* correspondente ao momento de produção do arquivo no fluxo de processamento apresentado na [Seção 2.6](#). De referir ainda que os arquivos resultantes da conversão de imagens [DICOM](#) para NIfTI se encontram representados pela *tag* (-). As imagens [DICOM](#) para cada aquisição devem ser guardadas como referido, numa pasta *FMRIDCM*, *DTIDCM* ou *MRIDCM*, respectivamente para aquisições funcionais, de difusão ou estruturais. Desta forma na [Tabela B.1](#) é apresentado o dicionário de arquivos para análise *FMRI*, na [Tabela B.2](#) o dicionário para análise DTI, enquanto que na [Tabela B.3](#) é evidenciado o dicionário criado

152 APÊNDICE B. NOMENCLATURA E DESCRIÇÃO DE FICHEIROS

para análise . Por último na Tabela B.4 é apresentado o dicionário resultante da aplicação de TrackVis, Freesurfer e gbedpostx respectivamente.

Tabela B.1: Dicionário de ficheiros para a análise *FMRI*

Ficheiro	Designação	Software	Tag
..._fnc_2mni.mat	Transformation matrix to MNI	FSL flirt	5
..._fnc_2str.mat	Transformation matrix to structural space	FSL flirt	5
..._fnc.nii.gz	Raw data	dcm2nii	-
..._fnc_vol_stime_mcf.par	Motion parameters	FSL mcflirt	3
..._fnc_vol_stime_mcf_rot.png	Rotational movement	FSL mcflirt	3
..._fnc_vol_stime_mcf_trans.png	Translational movement	FSL mcflirt	3
..._fnc_vol_stime_mcf_dvars.txt	DVARs estimation	FSL fsl_motion_outlie	3
..._fnc_vol_stime_mcf_outliers.txt	Motion outliers	FSL fsl_motion_outlie	3
..._fnc_vol_stime_mcf.nii.gz	Movement correced image	FSL mcflirt	3
..._fnc_vol_stime_mcf_mean_bet_mask.nii.gz	Brain mask	FSL fslmaths	4
..._fnc_vol_stime_mcf_bet_mni.nii.gz	Normalized data	FSL flirt/fnirt	5
..._fnc_vol_stime_mcf_bet_mni_smooth_filter.nii.gz	Preprocessed data without denoising	FSL fslmaths	7,8
..._fnc_vol_stime_mcf_bet_denoise_filtered.nii.gz	Preprocessed data without smoothing	FSL fslmaths	9
..._fnc_vol_stime_mcf_bet_mni_filtered.nii.gz	Preprocessed data without smoothing and denois	FSL fslmaths	9
..._fnc_vol_stime_mcf_bet_mni_denoised_filter.nii.gz	Preprocessed data without smoothing	FSL fslmaths	9
..._fnc_vol_stime_mcf_bet_mni_denoised_smooth_filter.nii	Preprocessed data	FSL fslmaths	7,8
..._fnc_vol_stime_mcf_bet_mean.nii.gz	Mean image	FSL fslmaths	5
..._fnc_meants.txt	AAL mean time-series	fsl fslmeants	10

Tabela B.2: Dicionário de ficheiros para a análise *DTI*

Ficheiro	Designação	Software	Tag
..._diff_2mni.mat	Transformation matrix to MNI	FSL flirt	13
..._diff_2str.mat	Transformation matrix to structural space	FSL flirt	13
..._diff_b0_bet_mask.nii.gz	Brain mask	FSL bet	12
..._diff_b0.nii.gz	Raw b0 image	dcm2nii	-
..._diff_dtifit_FA.nii.gz	Fractional anisotropy map	FSL fdt	16
..._diff_dtifit_L1.nii.gz	First eigenvalue map	FSL fdt	15
..._diff_dtifit_L2.nii.gz	Second eigenvalue map	FSL fdt	15
..._diff_dtifit_L3.nii.gz	Third eigenvalue	FSL fdt	15
..._diff_dtifit_MD.nii.gz	Mean diffusivity map	FSL fdt	16
..._diff_dtifit_RD.nii.gz	Radial diffusivity map	FSL fdt	16
..._diff_dtifit_S0.nii.gz	B0 image	FSL fdt	15
..._diff_dtifit_tensor.nii.gz	Diffusion tensors	FSL fdt	14
..._diff_dtifit_V1.nii.gz	First eigenvector	FSL fdt	15
..._diff_dtifit_V2.nii.gz	Second eigenvector	FSL fdt	15
..._diff_dtifit_V3.nii.gz	Third eigenvector	FSL fdt	15
..._diff_eddy_bet.nii.gz	Preprocessed data	FSL eddy_correct	13
..._diff_rotated.bvec	Rotated diffusion gradients	fsl rotate_bvecs	13
..._diff.bval	Diffusion weights	dcm2nii	-
..._diff.bvec	Diffusion gradients	dcm2nii	-
..._diff.nii.gz	Raw data	dcm2nii	-
..._diff_eddy.ecclog	Eddy current correction transformations	FSL eddy_correct	11
..._diff_trackvis.bvec	Trackvis ready diffusion gradients		17

Tabela B.3: Dicionário de ficheiros para a análise *MRI*

Ficheiro	Designação	Software	Tag
..._str_crop.nii.gz	Raw data	dcm2nii	
..._str_warp.nii.gz	Non-linear transformation to MNI	FSL fnirt	23
..._str_2mni.mat	Transformation matrix to MNI	FSL flirt	23
..._str_2diff.mat	Transformation matrix to diffusion space	FSL flirt	23
..._str_2fnc.mat	Transformation matrix to functional space	FSL flirt	23
..._str_crop_bet_mask.nii.gz	Brain mask	FSL fslmaths	24
..._str_crop_bet_mni.nii.gz	Normalized data	FSL fnirt/flirt	23
..._str_pve_0.nii.gz	CSF partial volume estimate	FSL FAST	25
..._str_pve_1.nii.gz	GM partial volume estimate	FSL FAST	25
..._str_pve_2.nii.gz	WM partial volume estimate	FSL FAST	25
..._str_prob_0.nii.gz	CSF tissue probability map	FSL FAST	25
..._str_prob_1.nii.gz	GM tissue probability map	FSL FAST	25
..._str_prob_2.nii.gz	WM tissue probability map	FSL FAST	25

Tabela B.4: Dicionário de ficheiros de utilização do software TrackVis, Freesurfer e BEDPOSTX, respectivamente

Ficheiro	Designação	Software	Tag
..._diff_dtk_track.trk	Deterministic tractography	trackvis	18
..._diff_dtk_fa_color.trk	Colored fractional anisotropy map	trackvis	18
aparc.a2009s+aseg.mgz	Parcellated cortical ribbon Destrieux	Freesurfer	20
aparc+aseg.mgz	Parcellated cortical ribbon Desikan-Killiany	Freesurfer	21
T1.mgz	T1 volume	Freesurfer	22
All files	Bayesian Estimation of diffusion parameters	FSL bedpostx	19

# Apêndice C

## BrainArchive - Descrição Funcional

Neste capítulo pretende-se descrever pormenorizadamente as funcionalidades desenvolvidas na aplicação BrainArchive (apresentadas em menor detalhe na [Seção 4.7](#)), destacando as soluções encontradas para a sua implementação.

### C.1 Autenticação e exportação de um ficheiro

Para que um utilizador, neste caso *User1*, inicie funções na aplicação BrainArchive é necessário a sua autenticação como evidenciado na [Figura C.1](#). Pretende-se assim demonstrar as capacidades da aplicação através da recorrência a dois utilizadores (*User1* e *User2*).

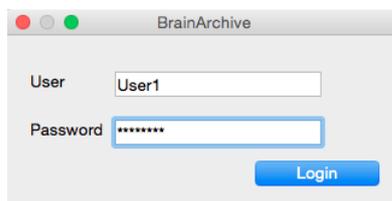


Figura C.1: Autenticação de um utilizador na aplicação BrainArchive

Após o sucesso na autenticação é possível ao utilizador tirar proveito das capacidades da aplicação. Como referido anteriormente, o BrainArchive faz

uso de uma diretoria de trabalho, designada (*MyBrain*) para organização de estudos segundo uma estrutura definida. Na Figura C.2 é possível verificar que a diretoria *MyBrain* (apresentada na Figura C.2.A) é representada fidedignamente na aplicação BrainArchive (Figura C.2.B).

De seguida pretende-se demonstrar a situação no qual um utilizador pretende exportar um ficheiro para o repositório. Para tal, é apenas necessário seleccionar o ficheiro em causa e caso este siga os parâmetros definidos na Seção 4.4 será automaticamente identificado conforme aquilo que representa.

Desta forma, e.g. caso se pretenda armazenar/partilhar a *Brain Mask* obtida em aquisições funcionais para o sujeito *SW0085C* do estudo *Estudo\_29z*, bastará escolher o ficheiro *SW0085\_fnc\_vol\_stime\_mcf\_mean\_bet\_mask.nii.gz* e posteriormente seleccionar o botão *Export*, como evidenciado na Figura C.3.

Como observável na Figura C.3 o quadrado referente ao estado de exportação do ficheiro está vermelho, simbolizando que este ainda não se encontra presente no repositório. Após a sucesso na exportação do ficheiro, este altera a sua cor para verde (Figura C.4). Além disso, o ficheiro foi exportado sem alteração nas opções de permissão, pelo que ficará disponível para todos os utilizadores com acesso ao repositório. Mais detalhes sobre permissões de acesso a ficheiros são apresentados em seções seguintes.

A exportação do ficheiro pode ser verificada na Figura C.4 no qual o quadrado relativo ao estado de exportação se alterou para verde, assim como na Figura C.5, onde se observa a existência do ficheiro de *Brain Mask* anteriormente exportado no repositório.

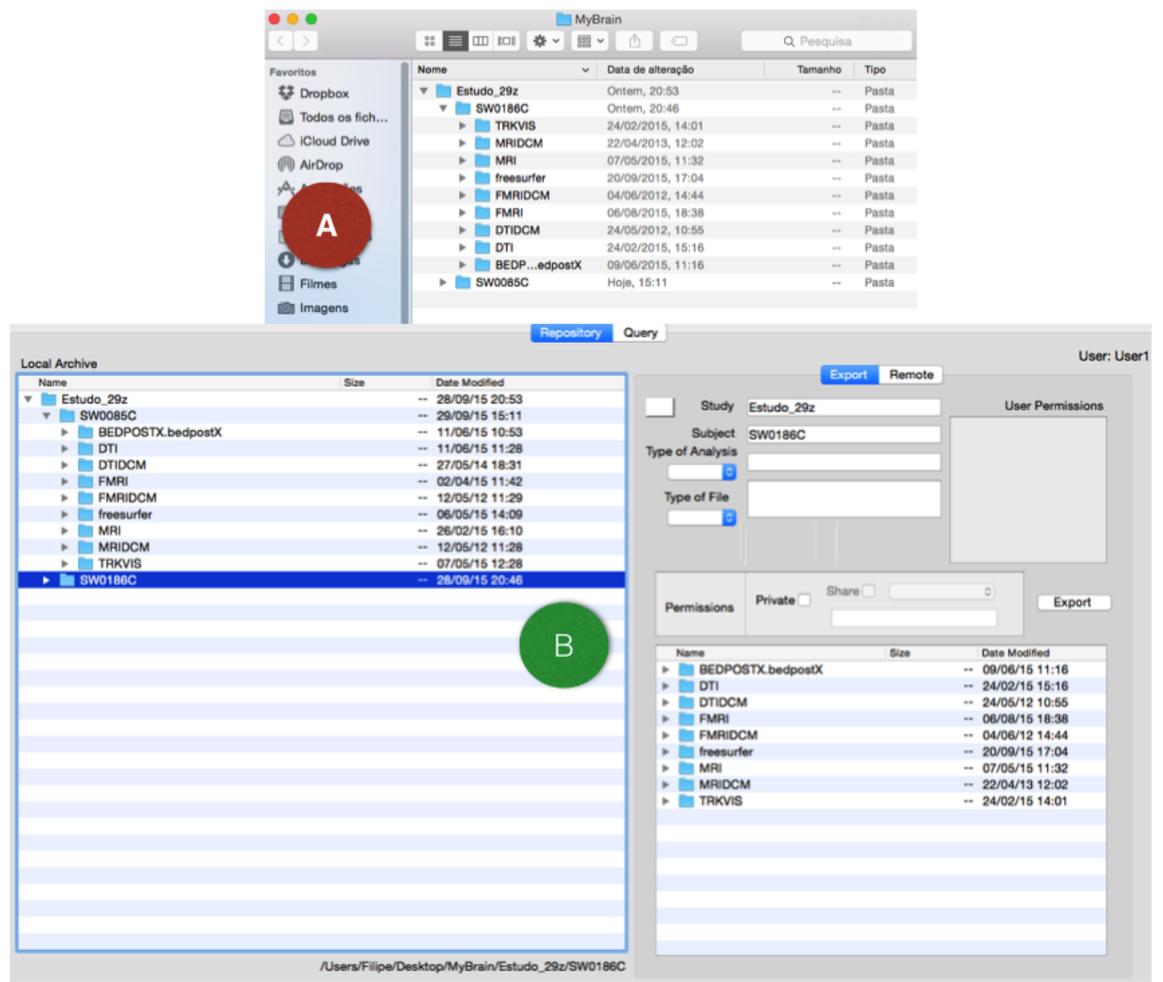


Figura C.2: Confirmação de paralelismo entre a diretoria *MyBrain* e a aplicação BrainArchive. (A) Exemplo de estruturação na diretoria *MyBrain*. (B) Interface inicial da aplicação BrainArchive

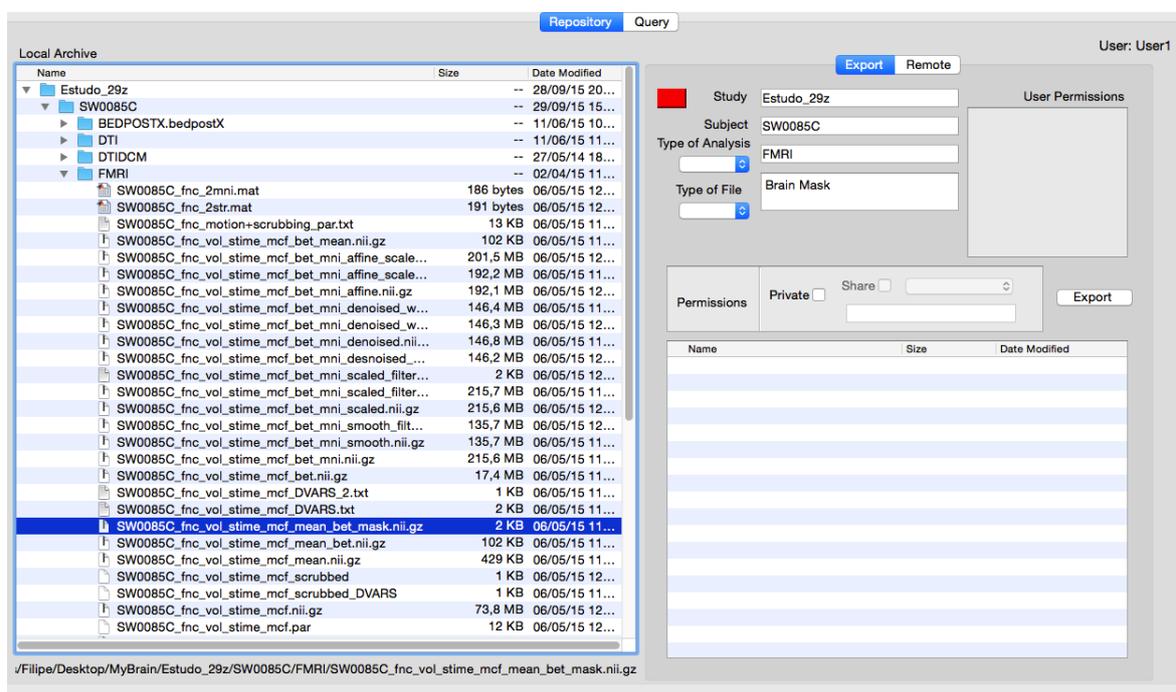


Figura C.3: Exportação de um ficheiro representado na interface da aplicação BrainArchive.

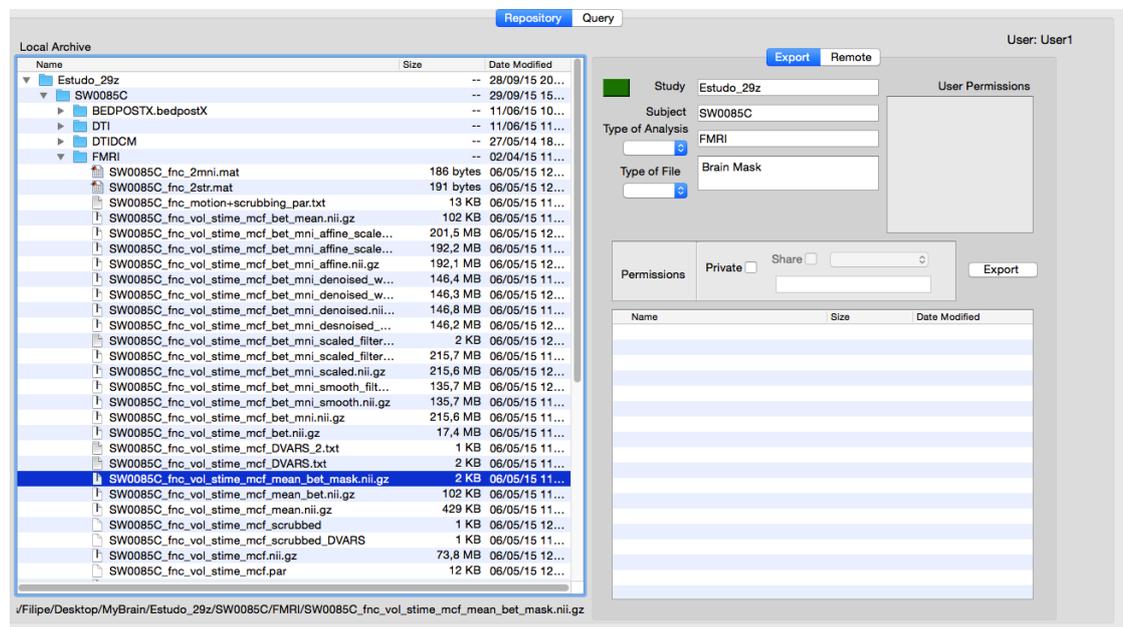


Figura C.4: Alteração no estado de exportação (quadrado de estado a verde) representado na interface BrainArchive

The screenshot displays the BrainArchive application interface. At the top, there are tabs for 'Repository' and 'Query'. The main area is titled 'Remote Archive' and shows a table of search results. The table has columns for Study, Subject, Analysis, User, FileType, File, and UploadDate. A single result is shown for the file 'Estudo\_29z/SW0085C/FMRI/SW0085C\_fnc\_vol\_stime\_mcf\_mean\_bet\_mask.nii.gz'.

Study	Subject	Analysis	User	FileType	File	UploadDate	
1	Estudo_29z	SW0085C	FMRI	User1	Brain Mask	Estudo_29z/SW0085C/FMRI/SW0085C_fnc_vol_stime_mcf_mean_bet_mask.nii.gz	2015-09-29 16:08:26

Below the table, there are several sections for file management:

- Change File Permissions:** Includes checkboxes for 'To Public' and 'Add User', and buttons for 'Change All' and 'Change Selected'.
- Import Options:** Buttons for 'Import Selected File(s)' and 'Import All File(s)'.

Search filters on the left include fields for User, Study, Subject, Analysis, and Type of file, along with 'Clear' and 'Search' buttons.

Figura C.5: Interface de interrogação da aplicação BrainArchive ( Confirmação de presença de ficheiro exportado para o repositório)

## C.2 Importação local assistida

Nesta secção pretende-se demonstrar a utilização do módulo de importação *Import Local Study*. Através deste módulo é possível ao utilizador organizar estudos de acordo com a hierarquia definida na [Seção 4.4](#).

Para tal, o utilizador *User2* será usado na demonstração deste processo. Como evidenciado na [Figura C.6](#) a diretoria *MyBrain* e correspondente representação na aplicação BrainArchive encontra-se vazia.

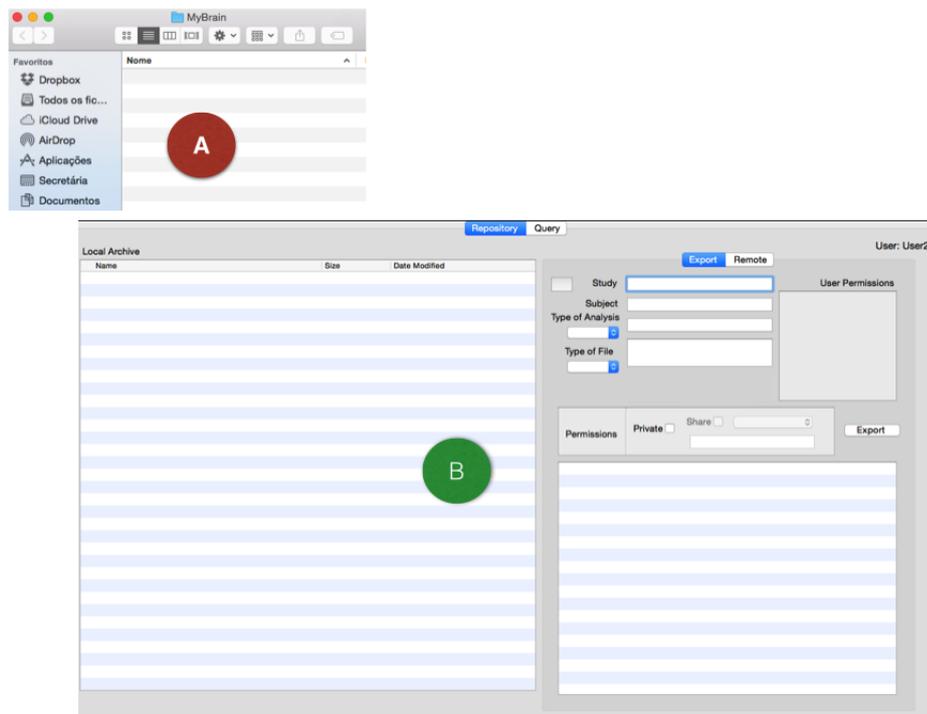


Figura C.6: Paralelismo evidenciado entre a diretoria *MyBrain* e interface da aplicação BrainArchive (ambas vazias). (A) diretoria *MyBrain*. (B) interface da aplicação BrainArchive.

Por forma a simplificar o processo de integração de estudos na diretoria *MyBrain* foi utilizado o módulo *Import Local Study* como evidenciado na [Figura C.7](#).

Na [Figura C.7.A](#) é visível a definição de dois atributos: *Study* e *Number of cases*, respectivamente *Estudo\_1x3* e 2. Após o preenchimento destes

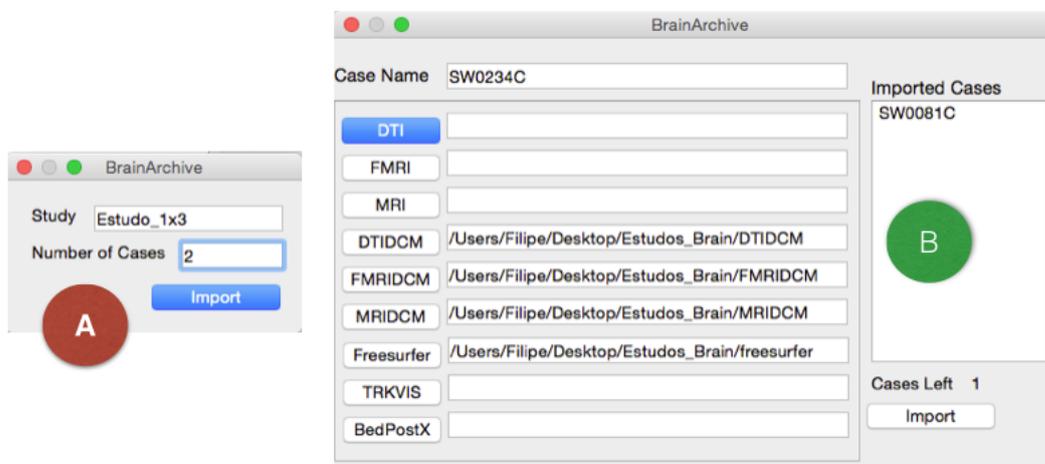


Figura C.7: Janelas da aplicação BrainArchive para importação local assistida de estudos. (A) Definição do estudo e número de sujeitos/casos. (B) Definição de cada sujeito/caso e seleção das opções a importar.

campos é necessário ao utilizador atribuir o nome de cada caso, bem como utilizar os botões respectivos para importação de cada opção de "análise".

Assim, é possível observar na Figura C.7.B que o sujeito/caso *SW0081C* se encontra importado. Para este foi selecionada a localização da diretoria contendo os ficheiros respectivos a *FMRI* e *DTI*. Da mesma forma é visível a importação do sujeito/caso *SW0234C* onde serão importadas as opções *DTIDCM*, *FMRIDCM*, *MRIDCM* e *Freesurfer*.

Após a importação do estudo *Estudo\_1x3*, este ficará disponível na diretoria MyBrain e consequentemente na aplicação *BrainArchive*. O sucesso de importação pode ser confirmado na Figura C.8, onde é visível a correcta estruturação do estudo *Estudo\_1x3* e respectivos sujeitos/casos *SW0081C* e *SW0234C*.

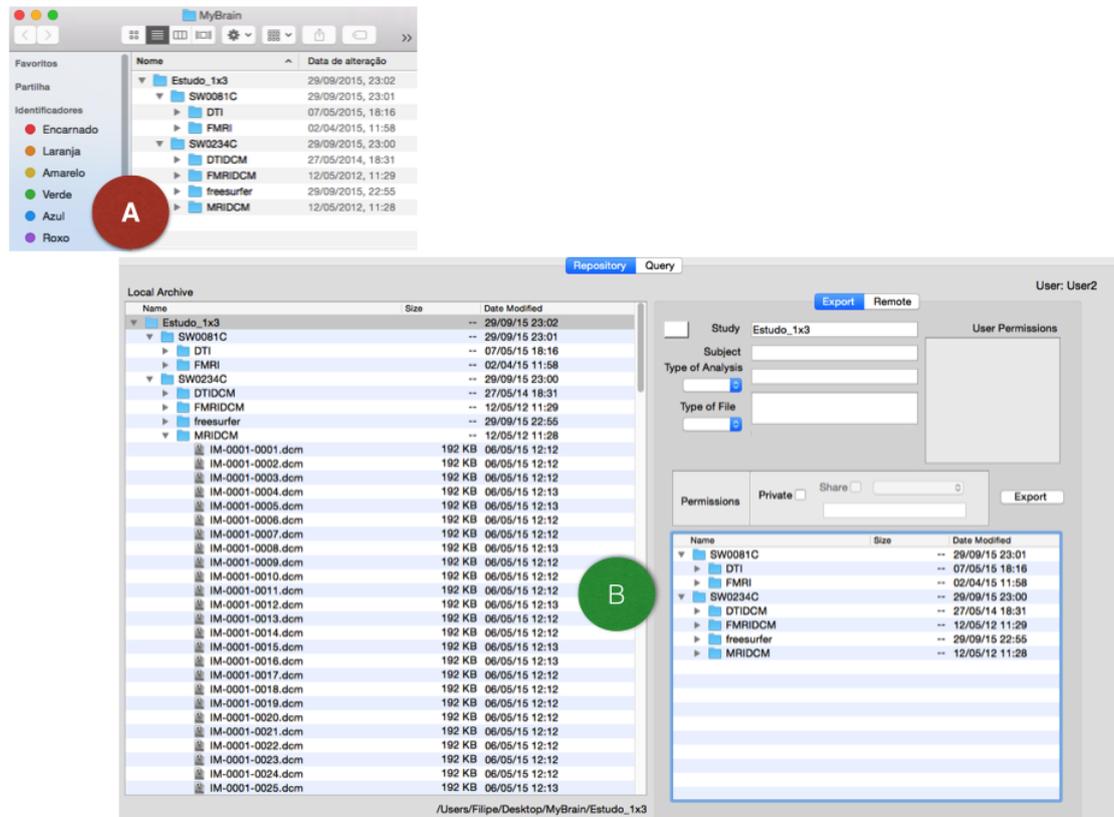


Figura C.8: Confirmação de importação local de Estudo na interface inicial da aplicação BrainArchive.

### C.3 Exportação automática de estudo

Nesta seção pretende-se demonstrar a capacidade de exportação de automática de um estudo. Para tal será novamente utilizado o utilizador *User2*. Assim, partindo do estado apresentado na Figura C.8 é permitido ao utilizador *User2* exportar para o repositório o estudo *Estudo\_1x3* de forma automática.

Este estudo apresenta dois sujeitos/casos *SW0081C* e *SW0234C*. Para o primeiro caso estão presentes as "análises" *FMRI* e *DTI*, enquanto que para o segundo caso encontram-se disponíveis para exportação as opções *DTIDCM*,

*FMRIDCM*, *MRIDCM* e *Freesurfer*.

Desta forma é possível ao utilizador seleccionar este estudo e exportar de forma automática toda a informação útil nele presente. Desta forma cada ficheiro exportado apresentará a descrição associada permitindo o seu futuro acesso por Estudo, Sujeito, Análise, Utilizador ou tipo de ficheiro.

É novamente permitido ao utilizador definir a permissão de acesso ao estudo a exportar. Neste sentido será seleccionada a opção *private*, pelo que até alteração posterior por parte do utilizador *User2*, o acesso a todos os ficheiros exportados será restrito ao mesmo.

The screenshot shows the BrainArchive interface with the following details:

- Search Filters:**
  - User: User2
  - Study: Estudo\_1x3
  - Subject: SW0081C, SW0234C
  - Analysis: DTI, FMRIDCM, freesurfer
  - Type of file: DTI, FMRIDCM, DTIDCM, FMRIDCM, DTI Dicom Zip, FMRIDCM, Parcellated cortical ribbon Desikan-Killiany, Parcellated cortical ribbon Destrieux, data mgz format, T1 volumes, MRI Dicom Zip, Brain Mask
  - Private:
- File List Table:**

Study	Subject	Analysis	User	FileType	File	UploadDate	
18	Estudo_1x3	SW0081C	DTI	User2	Third eigenvector	Estudo_1x3/SW0081C/DTI/SW0081C_diff_sdtit_V3.nii.gz	2015-10-06 22:04:03
19	Estudo_1x3	SW0081C	DTI	User2	Eddy current correction transformations	Estudo_1x3/SW0081C/DTI/SW0081C_diff_eddy_eccllog	2015-10-06 22:04:04
20	Estudo_1x3	SW0081C	DTI	User2	Preprocessed data	Estudo_1x3/SW0081C/DTI/SW0081C_diff_eddy_bet.nii.gz	2015-10-06 22:04:30
21	Estudo_1x3	SW0081C	DTI	User2	Rotated diffusion gradients	Estudo_1x3/SW0081C/DTI/SW0081C_diff_rotated_bvec	2015-10-06 22:04:31
22	Estudo_1x3	SW0081C	FMRI	User2	Raw data	Estudo_1x3/SW0081C/FMRI/SW0081C_fnc.nii.gz	2015-10-06 22:05:02
23	Estudo_1x3	SW0081C	FMRI	User2	Transformation matrix to MNI	Estudo_1x3/SW0081C/FMRI/SW0081C_fnc_2mmi.mat	2015-10-06 22:05:02
24	Estudo_1x3	SW0081C	FMRI	User2	Transformation matrix to structural space	Estudo_1x3/SW0081C/FMRI/SW0081C_fnc_2str.mat	2015-10-06 22:05:02
25	Estudo_1x3	SW0081C	FMRI	User2	Motion parameters	Estudo_1x3/SW0081C/FMRI/SW0081C_fnc_vol_stime_mcf.par	2015-10-06 22:05:03
26	Estudo_1x3	SW0081C	FMRI	User2	Mean image	Estudo_1x3/SW0081C/FMRI/SW0081C_fnc_vol_stime_mcf_be...	2015-10-06 22:05:03
27	Estudo_1x3	SW0081C	FMRI	User2	Normalized data	Estudo_1x3/SW0081C/FMRI/SW0081C_fnc_vol_stime_mcf_be...	2015-10-06 22:10:48
28	Estudo_1x3	SW0081C	FMRI	User2	Preprocessed data without denoising	Estudo_1x3/SW0081C/FMRI/SW0081C_fnc_vol_stime_mcf_be...	2015-10-06 22:14:49
29	Estudo_1x3	SW0081C	FMRI	User2	Brain Mask	Estudo_1x3/SW0081C/FMRI/SW0081C_fnc_vol_stime_mcf_m...	2015-10-06 22:14:49
30	Estudo_1x3	SW0234C	DTIDCM	User2	DTI Dicom Zip	Estudo_1x3/SW0234C/DTIDCM.zip	2015-10-06 22:16:51
31	Estudo_1x3	SW0234C	FMRIDCM	User2	FMRIDCM Zip	Estudo_1x3/SW0234C/FMRIDCM.zip	2015-10-06 22:17:48
32	Estudo_1x3	SW0234C	freesurfer	User2	Parcellated cortical ribbon Desikan-Killiany	Estudo_1x3/SW0234C/freesurfer/SW0234C/mri/aparc+aseq.mgz	2015-10-06 22:18:04
33	Estudo_1x3	SW0234C	freesurfer	User2	Parcellated cortical ribbon Destrieux	Estudo_1x3/SW0234C/freesurfer/SW0234C/mri/aparc.a2009+ase...	2015-10-06 22:18:05
34	Estudo_1x3	SW0234C	freesurfer	User2	data mgz format	Estudo_1x3/SW0234C/freesurfer/SW0234C/mri/wg.mgz	2015-10-06 22:18:17
35	Estudo_1x3	SW0234C	freesurfer	User2	T1 volumes	Estudo_1x3/SW0234C/freesurfer/SW0234C/mri/T1.mgz	2015-10-06 22:18:29
36	Estudo_1x3	SW0234C	MRIDCM	User2	MRI Dicom Zip	Estudo_1x3/SW0234C/MRIDCM.zip	2015-10-06 22:19:00
37	Estudo_29z	SW0085C	FMRI	User1	Brain Mask	Estudo_29z/SW0085C/FMRI/SW0085C_fnc_vol_stime_mcf_m...	2015-09-29 16:06:26
- Permissions:** Change File Permissions: To Public  Add User  Change All Change Selected
- Actions:** Import Selected File(s) Import All File(s)

Figura C.9: Confirmação de exportação de Estudo na interface de interrogação da aplicação BrainArchive.

Após o sucesso de exportação é possível observar a presença dos respectivos ficheiros no repositório na Figura C.9. Nesta foram acrescentadas linhas a vermelho por forma a melhor ilustrar informação relevante. Assim é possível observar na seção esquerda da interface a presença do utilizador *User2*, estudo *Estudo\_1x3* e sujeitos *SW0081C* e *SW0234C*.

Da mesma forma, na seção direita da interface, correspondente ao re-

#### C.4. ALTERAÇÃO DE PERMISSÕES EM FICHEIROS PRESENTES NO REPOSITÓRIO

positório, é visível a existência remota dos diversos ficheiros exportados. Como referido anteriormente para o sujeito *SW0081C* foi exportada a "análise" *FMRI* e *DTI*. Tal é observável na Figura C.9 onde se encontra presente e.g. o ficheiro correspondente aos gradientes de rotação de difusão (*Rotated diffusion gradients - DTI*) e a matriz de transformação para o espaço MNI (*Transformation matrix to structural space - FMRI*).

Da mesma forma, para o sujeito *SW0234C* foram exportadas as opções *DTIDCM*, *FMRIDCM*, *MRIDCM* e *Freesurfer*. Novamente na Figura C.9 é possível comprovar o sucesso de exportação das mesmas. Para cada diretoria contendo as imagens *DICOM* funcionais, estruturais e de difusão é automaticamente criado um ficheiro comprimido no formato *Zip*, sendo depois exportado. O resultado de exportação *Freesurfer* é também comprovado com presença e.g. do ficheiro correspondente a *Parcellated cortical ribbon Destrieux*.

Além dos ficheiros resultantes da exportação do estudo *Estudo\_1x3* por parte do utilizador *User2*, é também possível aceder ao ficheiro de (*Brain Mask - FMRI*) exportado anteriormente pelo utilizador *User1* e cuja permissão de acesso foi mantida pública.

## C.4 Alteração de permissões em ficheiros presentes no repositório

Na seção anterior foi apresentada a exportação automática de um estudo por parte do utilizador *User2*. Este foi realizado de modo "privado", por forma a que todos os ficheiros exportados apenas possam ser acessados pelo próprio até alteração posterior.

Desta forma, e tomando como exemplo o utilizador *User1*, este apenas verá o ficheiro de *Brain Mask - FMRI* por ele anteriormente exportado, tal como evidenciado na Figura C.10.

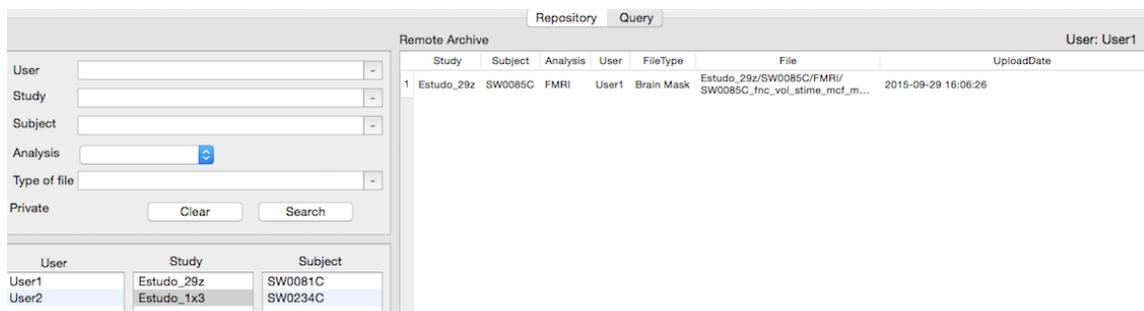


Figura C.10: Confirmação de inacessibilidade de ficheiros exportados por um utilizador por parte de outro, na interface de interrogação da aplicação BrainArchive.

Consideremos a situação no qual o utilizador *User2* decide partilhar com o utilizador *User1* alguns dos ficheiros exportados na seção anterior. De seguida pretende-se demonstrar o processo de alteração de permissão de ficheiros presentes no repositório, tendo como exemplo novamente *User1* e *User2*. Por forma a alterar a permissão de ficheiros, o utilizador *User2* deverá aceder à aplicação BrainArchive na aba *Query*, seleccionando a opção *private* como demonstrado na Figura C.11.

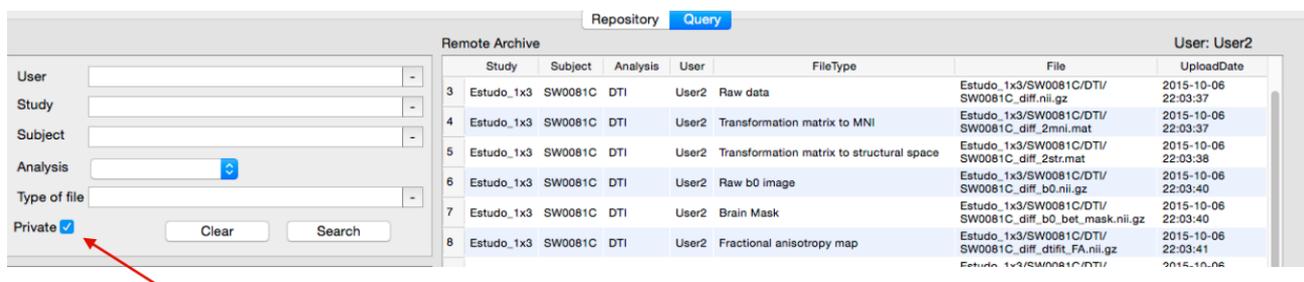


Figura C.11: Visualização exclusiva de ficheiros exportados por um utilizador, na interface de interrogação da aplicação BrainArchive.

Após a seleção desta opção a seção representada na Figura C.12 fica ativa e assim disponível para interação. Nesta é possível alterar as permissões para um nível público ou possibilitar o seu acesso a utilizadores específicos.

Neste caso pretende-se partilhar informação exclusivamente com a entidade *User1*, pelo que deverá ser seleccionada a opção *Add User* e introduzido

#### C.4. ALTERAÇÃO DE PERMISSÕES EM FICHEIROS PRESENTES NO REPOSITÓRIO

o nome do mesmo, como exemplificado na Figura C.12.

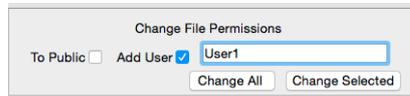


Figura C.12: Esboço de extensão de permissões a um utilizador por parte de outro, na Interface de interrogação da aplicação BrainArchive.

Selecionado o utilizador(es) ao qual/quais se pretende partilhar ficheiros é então necessário seleccionar qual a informação a partilhar. Neste processo é permitido ao utilizador partilhar desde um único ficheiro a todos os que a ele lhe dizem respeito, ou seja, que tenham sido exportados pelo próprio.

Neste exemplo pretende-se demonstrar a partilha de todos os ficheiros *DTI* pertencentes ao sujeito *SW0081C* do estudo *Estudo\_1x3*.

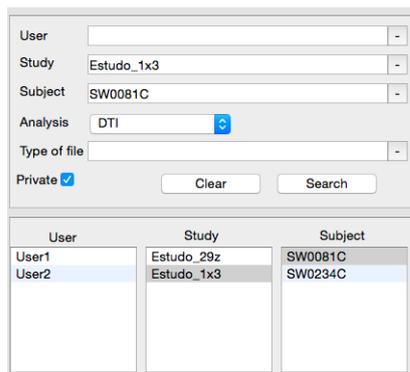


Figura C.13: Filtragem de ficheiros segundo parâmetros desejados, na interface de interrogação da aplicação BrainArchive.

Assim, tal como evidenciado na Figura C.13 são seleccionados os atributos que permitem exibir exclusivamente os ficheiros *DTI* pertencentes ao sujeito *SW0081C* do estudo *Estudo\_1x3*, exportados por *User2*.

Filtrada que está a informação, e uma vez que se pretende partilhar todos os ficheiros que cumprem com os atributos seleccionados, é apenas necessário ao utilizador optar pela opção *Change All*. Caso se pretenda partilhar apenas alguns ficheiros, o utilizador deverá escolher os mesmos e optar pela opção *Change Selected*.

Após o sucesso na alteração de permissão para estes ficheiros, a entidade *User1* passará a ter acesso aos mesmos, podendo extraí-los e usá-los da forma que pretenda. A confirmação desta modificação é evidenciada na Figura C.14, na qual é possível a *User1* aceder aos ficheiros de *DTI*, pertencentes ao sujeito *SW0081C* do estudo *Estudo\_1x3* exportados por *User2*.

Repository		Query		User: User1			
Study	Subject	Analysis	User	FileType	File	UploadDate	
3	Estudo_1x3	SW0081C	DTI	User2	Raw data	Estudo_1x3/SW0081C/DTI/SW0081C_diff.nii.gz	2015-10-06 22:03:37
4	Estudo_1x3	SW0081C	DTI	User2	Transformation matrix to MNI	Estudo_1x3/SW0081C/DTI/SW0081C_diff_2mni.mat	2015-10-06 22:03:37
5	Estudo_1x3	SW0081C	DTI	User2	Transformation matrix to structural space	Estudo_1x3/SW0081C/DTI/SW0081C_diff_2str.mat	2015-10-06 22:03:38
6	Estudo_1x3	SW0081C	DTI	User2	Raw b0 image	Estudo_1x3/SW0081C/DTI/SW0081C_diff_b0.nii.gz	2015-10-06 22:03:40
7	Estudo_1x3	SW0081C	DTI	User2	Brain Mask	Estudo_1x3/SW0081C/DTI/SW0081C_diff_b0_bet_mask.nii.gz	2015-10-06 22:03:40
8	Estudo_1x3	SW0081C	DTI	User2	Fractional anisotropy map	Estudo_1x3/SW0081C/DTI/SW0081C_diff_dtifit_FA.nii.gz	2015-10-06 22:03:41
9	Estudo_1x3	SW0081C	DTI	User2	First eigenvalue map	Estudo_1x3/SW0081C/DTI/SW0081C_diff_dtifit_L1.nii.gz	2015-10-06 22:03:42
10	Estudo_1x3	SW0081C	DTI	User2	Second eigenvalue map	Estudo_1x3/SW0081C/DTI/SW0081C_diff_dtifit_L2.nii.gz	2015-10-06 22:03:43
11	Estudo_1x3	SW0081C	DTI	User2	Third eigenvalue map	Estudo_1x3/SW0081C/DTI/SW0081C_diff_dtifit_L3.nii.gz	2015-10-06 22:03:44
12	Estudo_1x3	SW0081C	DTI	User2	Mean diffusivity map	Estudo_1x3/SW0081C/DTI/SW0081C_diff_dtifit_MD.nii.gz	2015-10-06 22:03:46
13	Estudo_1x3	SW0081C	DTI	User2	Radial diffusivity map	Estudo_1x3/SW0081C/DTI/SW0081C_diff_dtifit_RD.nii.gz	2015-10-06 22:03:47
14	Estudo_1x3	SW0081C	DTI	User2	B0 image	Estudo_1x3/SW0081C/DTI/SW0081C_diff_dtifit_S0.nii.gz	2015-10-06 22:03:48
15	Estudo_1x3	SW0081C	DTI	User2	Diffusion tensors	Estudo_1x3/SW0081C/DTI/SW0081C_diff_dtifit_tensor.nii.gz	2015-10-06 22:03:54
16	Estudo_1x3	SW0081C	DTI	User2	First eigenvector	Estudo_1x3/SW0081C/DTI/SW0081C_diff_dtifit_V1.nii.gz	2015-10-06 22:03:57
17	Estudo_1x3	SW0081C	DTI	User2	Second eigenvector	Estudo_1x3/SW0081C/DTI/SW0081C_diff_dtifit_V2.nii.gz	2015-10-06 22:04:00
18	Estudo_1x3	SW0081C	DTI	User2	Third eigenvector	Estudo_1x3/SW0081C/DTI/SW0081C_diff_dtifit_V3.nii.gz	2015-10-06 22:04:03
19	Estudo_1x3	SW0081C	DTI	User2	Eddy current correction transformations	Estudo_1x3/SW0081C/DTI/SW0081C_diff_eddy_ecclog	2015-10-06 22:04:04
20	Estudo_1x3	SW0081C	DTI	User2	Preprocessed data	Estudo_1x3/SW0081C/DTI/SW0081C_diff_eddy_bet.nii.gz	2015-10-06 22:04:30
21	Estudo_1x3	SW0081C	DTI	User2	Rotated diffusion gradients	Estudo_1x3/SW0081C/DTI/SW0081C_diff_rotated.bvec	2015-10-06 22:04:31
22	Estudo_29z	SW0085C	FMRI	User1	Brain Mask	Estudo_29z/SW0085C/FMRI/SW0085C_fnc_vol_stime_mcf_m...	2015-09-29 16:06:26

Figura C.14: Confirmação de partilha de ficheiros exportados por um utilizador para outro, na interface de interrogação da aplicação BrainArchive.

## C.5 Extração de ficheiros presentes no repositório

Uma dos principais pressupostos da aplicação BrainArchive é permitir a partilha, extração e uso de ficheiros/estudos adicionados ao repositório por diferentes entidades, mantendo em todo o processo a estruturação transparente dos ficheiros.

Na seção anterior foi exemplificada a alteração na permissão de acesso a ficheiros presentes no repositório. De seguida pretende-se ilustrar o processo de extração de ficheiros por parte da entidade *User1*.

Este passo é novamente simples e intuitivo sendo apenas necessário ao utilizador filtrar e selecionar os ficheiros que pretende. Caso pretenda importar todos os ficheiros não é necessária seleção de ficheiros, bastando apenas filtrar de acordo com o pretendido e optar pela opção *Import All File(s)*.

The screenshot shows the BrainArchive interface with the 'Query' tab selected. The interface is divided into several sections:

- Search Filters:** Includes fields for User, Study, Subject, Analysis, and Type of file. There are also 'Clear' and 'Search' buttons.
- User Selection:** A table with columns 'User', 'Study', and 'Subject'. It lists 'User1' and 'User2' with their respective study and subject IDs.
- File List Table:** A table with columns: Study, Subject, Analysis, User, FileType, File, and UploadDate. It contains 22 rows of data, including files like 'Raw data', 'Transformation matrix to MNI', 'Brain Mask', and 'Fractional anisotropy map'.
- Permissions:** A section for 'Change File Permissions' with 'To Public' and 'Add User' options, and 'Change All' and 'Change Selected' buttons.
- Import Buttons:** 'Import Selected File(s)' and 'Import All File(s)' buttons at the bottom right.

Figura C.15: Extração de ficheiros por parte de um utilizador, na interface de interrogação da aplicação BrainArchive

No exemplo demonstrado na Figura C.15, o utilizador *User1* pretende extrair alguns dos ficheiros *DTI* partilhados pela entidade *User2*. Desta forma são selecionados seis ficheiros pretendidos: *Raw data*, *Fractional anisotropy map*, *Mean diffusivity map*, *Diffusion tensors*, *Eddy current correction transformations* e *Preprocessed data*.

Após a seleção dos mesmos basta optar pela opção *Import Selected File(s)*, para que a extração dos ficheiros em causa ocorra.

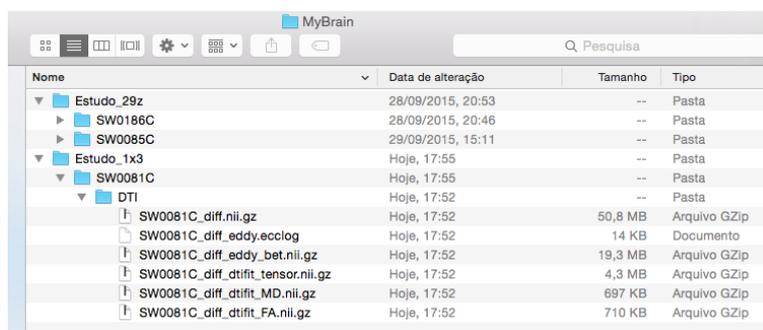


Figura C.16: Confirmação de extração de ficheiros para a diretoria *MyBrain*

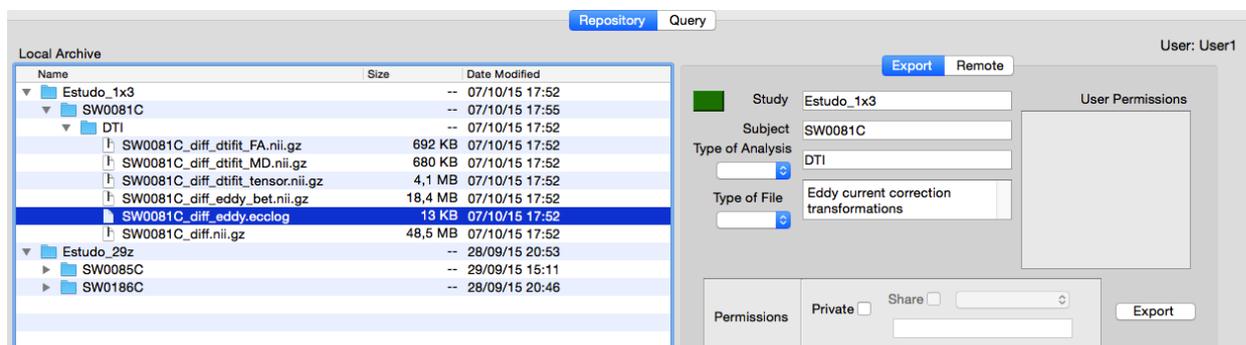


Figura C.17: Resultado de extração de ficheiros evidenciado na interface inicial da aplicação BrainArchive

As Figuras C.16 e C.17 permitem observar o sucesso da extração dos seis ficheiros pretendidos. Este processo transparente mantém a estruturação dos mesmos sendo visível a manutenção de hierarquia *Estudo/Sujeito/Análise/-Ficheiro*. Concluído o processo de extração, os ficheiros podem ser utilizados pelo utilizador de forma pretendida pelo próprio

## C.6 Outras considerações de utilização

Nesta seção pretende-se esclarecer dois aspectos :exportação de ficheiros que não cumpram com a estruturação hierárquica definida; exportação de ficheiros que se encontrem já presentes no repositório.

### C.6.1 Exportação de ficheiros que não cumpram com a estruturação hierárquica definida

Referente ao primeiro caso, um utilizador pode colocar na sua diretoria *MyBrain* ficheiros cuja nomenclatura não se encontra definida na arquitetura *BrainArchive*. Desta forma é possível a um utilizador usar ficheiros que não cumpram com os requisitos previstos, assegurando assim e.g. a incorporação de ficheiros obtidos previamente à utilização de *BrainArchive*.

O processo de exportação destes ficheiros terá de ser compreensivelmente manual, no entanto o mesmo será exportado de forma estruturada, assegurando assim a manutenção da coerência do repositório.

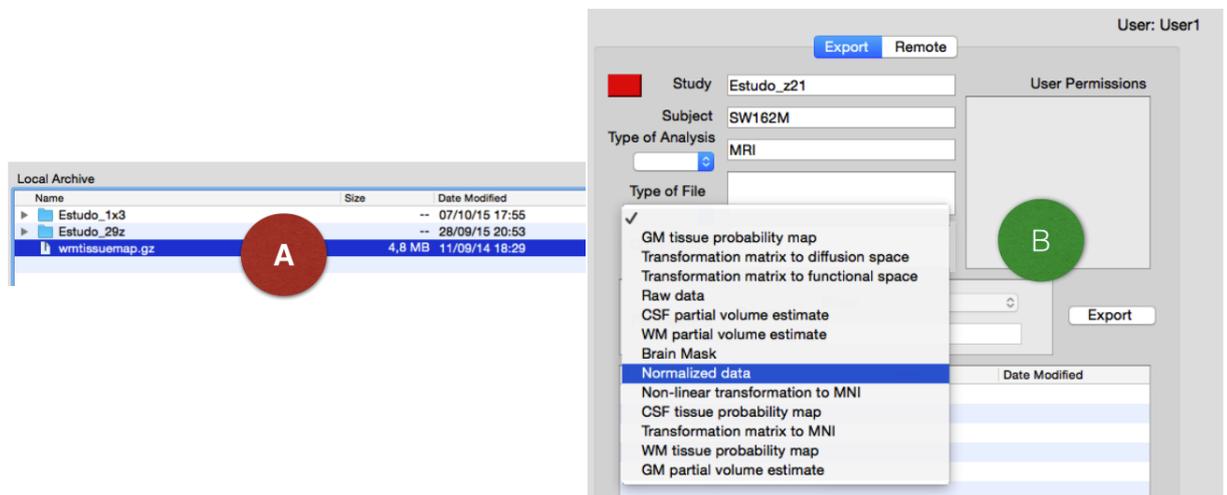


Figura C.18: Exportação de ficheiro sem estruturação definida. (A) Visualização do ficheiro adicionado à diretoria *MyBrain*. (B) Processo de exportação manual do ficheiro na interface inicial da aplicação *BrainArchive*

Na Figura C.18.A é visível a incorporação de um ficheiro designado *wmtis-suemap*. Este não se encontra numa hierarquia definida e a sua nomenclatura não é reconhecida. No entanto este pode ser exportado caso o utilizador preencha manualmente os campos necessários como ilustrado na Figura C.18.B. Nesta o utilizador *User1* define o ficheiro como pertencente ao estudo *Estudo\_z21*, sujeito *SW162M*, análise *MRI* e contendo informação referente a dados *MRI* normalizados.

Após a exportação do ficheiro, este será automaticamente movido para uma nova localização com a hierarquia definida pelo utilizador (*Estudo\_z21*, sujeito *SW162M*, análise *MRI*) e possuindo a nomenclatura própria do ficheiro a que corresponde. Tal é observável na Figura C.19.A.

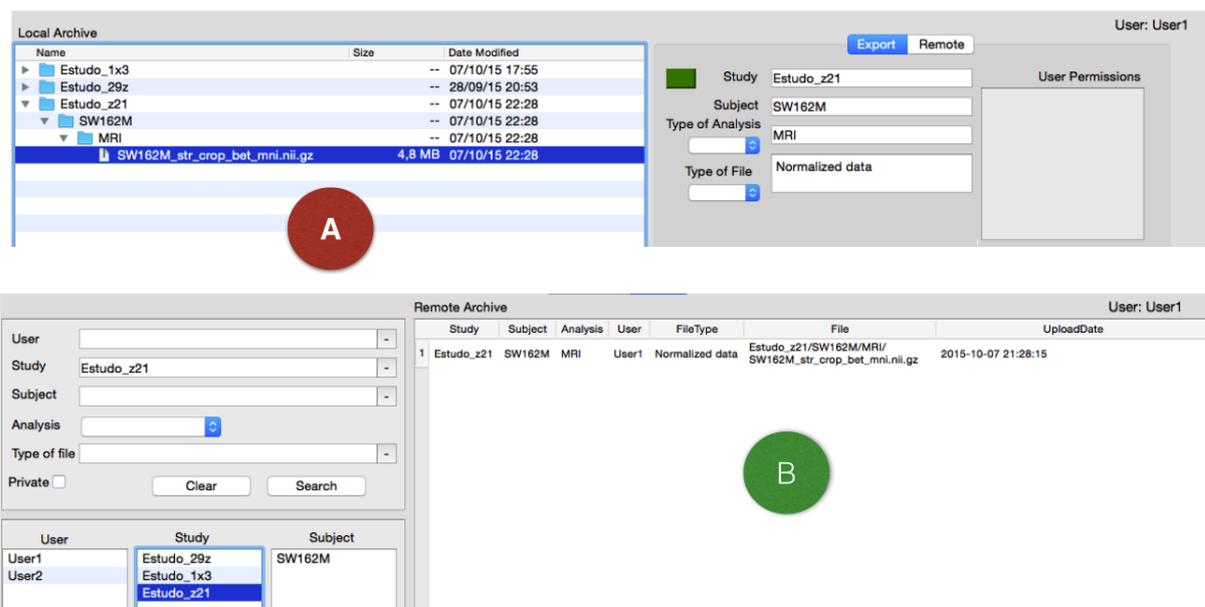


Figura C.19: Confirmação do sucesso de exportação de ficheiro desestruturado. (A) realocação do ficheiro segundo a estrutura definida pelo utilizador na interface inicial da aplicação BrainArchive. (B) presença do ficheiro no repositório evidenciado na interface de interrogação da aplicação BrainArchive.

Da mesma forma, o sucesso de exportação pode ser confirmado na Figura C.19.B, na qual é visível a presença do ficheiro anteriormente exportado.

### C.6.2 Exportação de ficheiros que se encontrem já presentes no repositório

Pretende-se de seguida esclarecer a situação na qual por engano ou porque foi obtida uma nova versão de um ficheiro, é submetida a exportação de um documento já presente no repositório.

Por forma a diferenciar as duas situações acima referenciadas optou-se por utilizar o algoritmo de *hash* unidirecional [Message-Digest algorithm 5 \(MD5\)](#). Este algoritmo funciona como impressão digital de um ficheiro, pelo que dois arquivos exatamente iguais apresentam o mesmo MD5.

Desta forma é possível comparar a versão entre o ficheiro presente no repositório e o ficheiro a exportar. Caso os dois documentos apresentem o mesmo MD5, então estes são exatamente iguais pelo que o processo de exportação é cancelado de forma transparente.

Por outro lado se os MD5 forem distintos o novo documento será exportado para o repositório.