Grant agreement no. 283562

N4U

neuGRID for you: expansion of neuGRID services and outreach to new user communities

Combination of Collaborative Project and Coordination and Support Action

Objective INFRA-2011-1.2.1 – e-Science environments

Start date: July 1st 2011 **- Duration**: 42 months

Deliverable data

Deliverable reference number and title: 7.1 NS-NDD Data Portfolio Report
Due date: 30 June 2012
Actual submission date: 26 July 2012
Organisation name of lead contractor for this deliverable: CEA
Dissemination level: Public

Authors

JF Mangin (CEA), A Redolfi (FbF), GB Frisoni (FbF)

Approval

Workpackage Leader: JF Mangin (CEA)
Project Coordinator: GB Frisoni (FbF)
PMT members: Giovanni B Frisoni (P1 FBF), David Manset (P2 maatG), Richard McClatchey (P3 UWE)

History Record

	Date	Notes
Version 1	June 29, 2012	First release by JFM
	June 29, 2012	Comments by GBF
Version 2	July 12, 2012	Second release by JFM
Version 3	July 14, 2012	Edits by GBF
	July 16, 2012	Comments by CFc
Version 4	July 20, 2012	Edits by AR
Version 5	July 20, 2012	Final Edits by GBF, final version

Table of Contents

1. Executive Summary

The consortium maintains three lists of datasets of interest, one for each neuroscientific application, including priorities for integration. Priorities have been updated during all the quarterly meetings. Two typical datasets are currently used to set up the N4U procedure for dataset integration: one monomodal (only one imaging modality) with cross-sectional and longitudinal data (OASIS), one multimodal (ADHD). The procedure of data uploading and raw quality control has been defined to rely on CATI's engineers and softwares but also on the definition of N4U's data format and database schema organisation. In addition, the development of a specific quality control environment (QCE) has been started to ensure high quality results concerning the outputs of the main N4U pipelines, not always 100% reliable, in analyzing large amount of data.

2. Introduction

One of the services provided by N4U to the users is access to large public databases already imported in the grid facility. The list of interesting databases initially included in Appendix of the DoW has been constantly upgraded and is regularly enriched with new opportunities (see Annex I).

The table reports 22 datasets categorized according to interest to target communities:

- Ns-NND: Neurodegenerative disease community
- Ns-WMD: White Matter community
- Ns-PSY: Psychiatric community

For each dataset specific parameters have been investigated:

1) The data web source
2) Its possible scientific impact
2) The relative importance and priority of the dataset in the N4U platform
3) The clinical and epidemiological characteristics of the dataset plus additional information on group size and acquisition type
4) The accessibility of the dataset (Public / Private)
5) A "Usage Example".

Specific priorities have been defined in terms of database importation. Here importation means that the public databases are preformatted for seamless application of any item of N4U's analysis portfolio.

A preliminary work before adding a new database to N4U's offer is an initial sanity check and a quality control aiming at predicting the behavior of the analysis portfolio. This quality control aims at filtering out low quality images bound to trigger trouble for the user. Whenever the amount of rejected images is deemed too high, the whole database importation will be canceled.

Among the others, two datasets of primary importance for the definition of the upload pipeline were chosen:

1. **OASIS:** dataset with only T13D volumetric scan. It's a very comprehensive study characterized by both cross-sectional and longitudinal arm. The population covers a very wide range of aging starting from Young, Middle Aged, not-demented and Demented Older Adults.
2. **ADHD200:** dataset characterized by both functional imaging datasets (fMRI) and anatomical imaging across many independent sites. Preliminary QC assessments (questionable versus usable) based upon a visual inspection by an expert of every Time-Series scans are already included. These *a-priori* judgments can be used to make comparison with the method that will be developed in N4U and extended to other datasets.

It is clear that for their intrinsic characteristics these two datasets are very well-suited to the

fine-tuning definition process at this stage.

The collaboration between N4U and the French platform CATI (led by CEA) aims at building upon CATI's facilities to provide N4U with an efficient mechanism for database importation. CATI's workflow includes several levels of Quality Control (raw data and analysis pipelines) performed by clinical research assistants and technicians endowed with dedicated softwares developed by CATI's engineer. Since the CATI platform was created just recently (January 2011), the first releases of these softwares occurred just in 2012. Then CATI's engineers began working on deriving some of these softwares for the specific needs of N4U. The aforementioned prioritized datasets (OASIS and ADHD200) are used for this tuning. The first version of N4U's importation procedure should be ready at the beginning of 2013.

3. Methodological Approach

3.1 Procedures for dataset upload in the production phase

During quarterly meetings, the lists of selected datasets are revised by the consortium in order to update the integration priorities. These priorities are defined according to several criteria: scientific interest, novelty of the dataset, quality of the acquisition, amount of work to clean up the dataset according to N4U needs, and specific IP issues. Moreover, CEA has screened the datasets to detect and report on strengths and weaknesses. What follows is a description of the procedures that have been defined for the first datasets integration.

CEA technician in charge of evaluation or integration will:

1. Download the dataset images and provide a sanity check report;
2. Download the clinical information and provide a report on the content;
3. Check whether the data are really anonymised;
4. Perform a low-level quality control using QualiCati platform (see figure 1 below);
5. Perform a high-level quality control using CATI's pipelines;
6. Deliver a QC report of the scan to be added to the clinical data. In this report trouble-shootings will be reported (e.g.: image degradation due to: Movement artifact, Wrap around, Low SNR, Signal loss, Metal artifact, Image degradation, Ghosting effects, etc..);
7. Prepare the dataset according to N4U's standard in order for MaatG and UWE to push the dataset on the grid in the Data Atlas schema.

This procedure is currently tuned by CATI's engineers with OASIS and ADHD datasets, which are the two first datasets in the process of being imported in N4U.

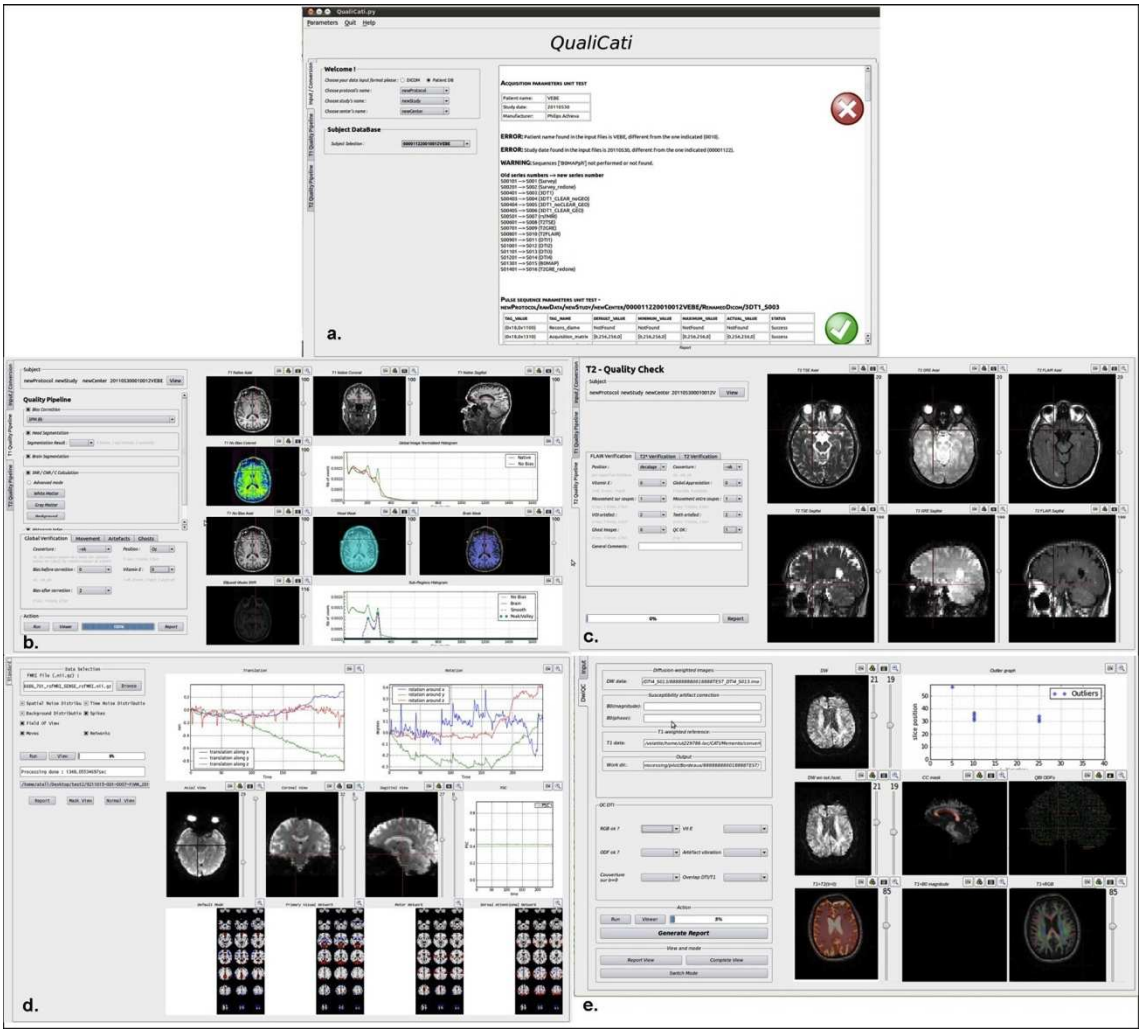

Figure 1: A snapshot of the QC software of CATI to be used during N4U dataset importation protocol

3.2 Datasets being uploaded into N4U

As reported above for NS-NDD, the OASIS dataset is in the process of being integrated first (http://www.oasis-brains.org/). This is a very popular dataset freely available to the scientific community. Thanks to its monomodal specificity, it is used as a first simple benchmark to validate N4U's procedures to integrate new datasets.

OASIS includes:

1) A Cross-sectional MRI Data in Young, Middle Aged, Nondemented and Demented Older Adults (416 subjects);
2) A Longitudinal MRI Data in Nondemented and Demented Older Adults (150 subjects).

In addition to the raw data QC level provided by QualiCATI, OASIS and ADHD T1-weighted scans have just been finished to be analyzed using two *ad-hoc* pipelines providing grey matter / white matter classification (SPM and BrainVISA/Morphologist), in order to use the agreement between the results as an index of the scan quality. At the same time, a visual quality control was performed by one clinical research assistant of CATI. It was shown that the low quality images detected by the human evaluator were predicted by the low agreement between the two pipelines. This approach will provide an efficient way to perform QC of massive databases using limited human resources: a first automatic QC will select the most suspicious data to be controlled by the clinical research assistant.

A second freely available multimodal dataset used to validate the process further was the ADHD200 study with 776 anatomical + resting state data, including 281 Attention Deficit Disorders. Having been verified the consistency of this data assurance pipeline, both datasets are about to be loaded in the neuGRID4you platform.

3.3 Datasets being prioritized for future upload into N4U

The procedures set up during this validation work will be used to integrate other datasets. In order to improve the attractiveness of N4U, the consortium is negotiating the rights to perform the first public release of large scale datasets, which would increase the visibility of the platform. Current targets are:

1) **The LADIS dataset**[1], a very well known success story in the neuroscientific community. The LADIS project is led by Leonardo Pantoni and Domenico Inzitari at the Universuty of Florence, Italy, and has Frederik Barkhof, from N4U partner VUMC, as one of the key steering committee members. The LADIS project (Leukoaraiosis And DISability in the elderly) aimed at assessing the role of age-related white matter changes (ARWMC) as an independent predictor of the transition to disability in initially nondisabled elderly 65 to 84 years old. Subjects who were not impaired or impaired on only 1 item of the Instrumental Activity of Daily Living (IADL) scale, presenting with different grades of ARWMC severity, were enrolled. Eleven European centers were involved. Patients were assessed at baseline using an extensive set of clinical and functional tests including global functioning, cognitive, motor, psychiatric and quality of life measures. MRI studies were performed at baseline and have been repeated at the end of the follow-up period to evaluate changes of ARWMC and other lesions.

   Six-hundred and thirty-nine subjects were enrolled in hospital-based settings and were followed up for up to 3 years. The large and comprehensive set of measures in LADIS has enabled a comprehensive description of their functional and clinical features to be examined in relation to different morphological patterns and severity of ARWMC. The longitudinal design has provided insight into the possible role of ARWMC and their progression as an independent contributor to disability in the elderly, eventually helping to develop preventive strategies to reduce the burden of disability in late life. The study

[1] Pantoni L, Basile AM, Pracucci G, Asplund K, Bogousslavsky J, Chabriat H, Erkinjuntti T, Fazekas F, Ferro JM, Hennerici M, O'brien J, Scheltens P, Visser MC, Wahlund LO, Waldemar G, Wallin A, Inzitari D. Impact of age-related cerebral white matter changes on the transition to disability -- the LADIS study: rationale, design and methodology. Neuroepidemiology. 2005;24(1-2):51-62.

have also helped standardize, on an international basis, tools and criteria to identify early stages of disability.

The project has produces more than 120 scientific papers in the international peer-reviewed literature, and counting. The neuroscientific community feels that despite this impressive amount of results the LADIS dataset still lends itself to investigation, particularly with the most advanced image analysis tools.

2) **The EDSD dataset** (European DTI study in Dementia (EDSD). EDST[2] is a newly established framework of nine European centers: Amsterdam (Netherlands), Brescia (Italy), Dublin (Ireland), Frankfurt (Germany), Freiburg (Germany), Milano (Italy), Mainz (Germany), Munich (Germany), and Rostock (Germany), with one center including data from two different MRI scanners. At present, the data bank includes 335 Diffusion Tensor Imaging and 335 structural MRI scans from patients with AD and healthy elderly subjects. EDST is led by Stefan Teipel in Rostok, Germany. The dataset is just starting producing its very first papers, but it is regarded as one of the very first examples of its kind internationally and the first in Europe, and therefore with high potential.

The rapid evolution of the mentalities related to data sharing should multiply the opportunities during the next years. We are especially attentive to the International Neuroimaging Data-sharing Initiative, which is performing a very active scanning of public datasets (http://fcon_1000.projects.nitrc.org/indi/summerofsharing2012.html).

3.5 Quality Control Service (QCE) Deployment

A pipeline-based QC is also developed through collaboration between CEA and FBF through Alberto Redolfi, doctoral student at CEA. This line of work is developed along the ambitious purpose called: “N4U Quality Control Environment (QCE)”, which will be used to perform a systematic evaluation of the outputs of the most important pipelines integrated in the N4U portfolio.

To implement the QCE, which will also be proposed as a service at the end of the project, we have decided to use the Python Language exploiting in particular the Scikitslearn (machine learning functions in Python), Numpy and Scipy packages (fig. 2).

[2] Teipel SJ, Reuter S, Stieltjes B, Acosta-Cabronero J, Ernemann U, Fellgiebel A, Filippi M, Frisoni G, Hentschel F, Jessen F, Klöppel S, Meindl T, Pouwels PJ, Hauenstein KH, Hampel H. Multicenter stability of diffusion tensor imaging measures: a European clinical and physical phantom study. Psychiatry Res. 2011 Dec 30;194(3):363-71. Epub 2011 Nov 9.

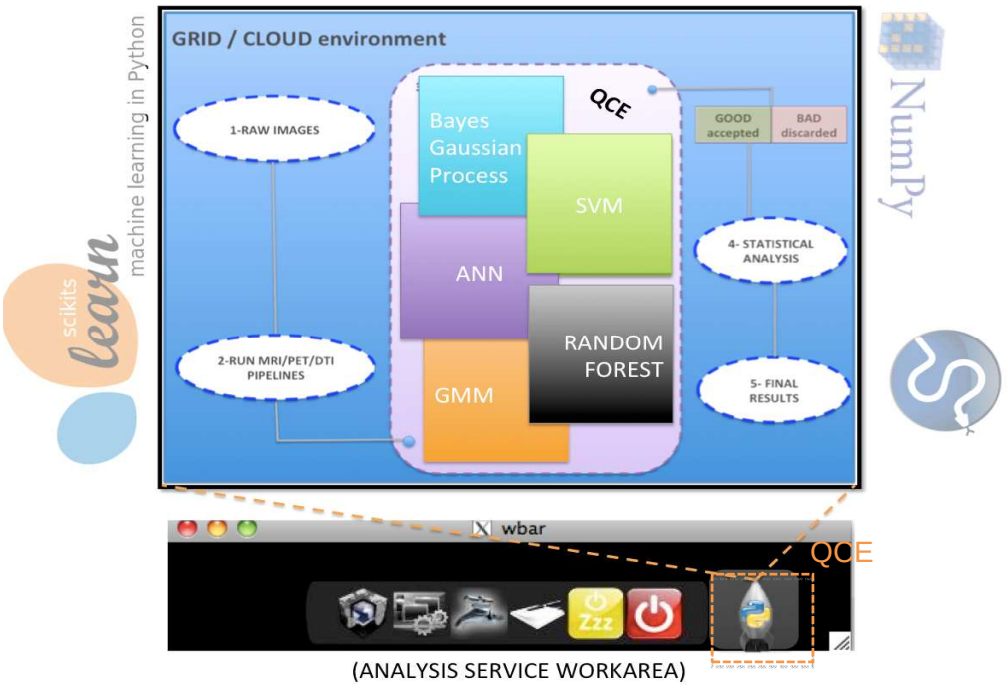

Figure 2: QCE schema structure. The python based environment should be integrated in the N4U analysis service area allowing the evaluation of the correctness of the different pipelines' outputs offered by N4U to the neuroscientists.

The QCE has been designed to be based on specific machine learning approaches (e.g.: Support Vector Machine (SVM), Random Forest (RF), Gaussian Mixture Models (GMM) and Artificial Neural Network (ANN)) to quality assess the pipeline outputs.

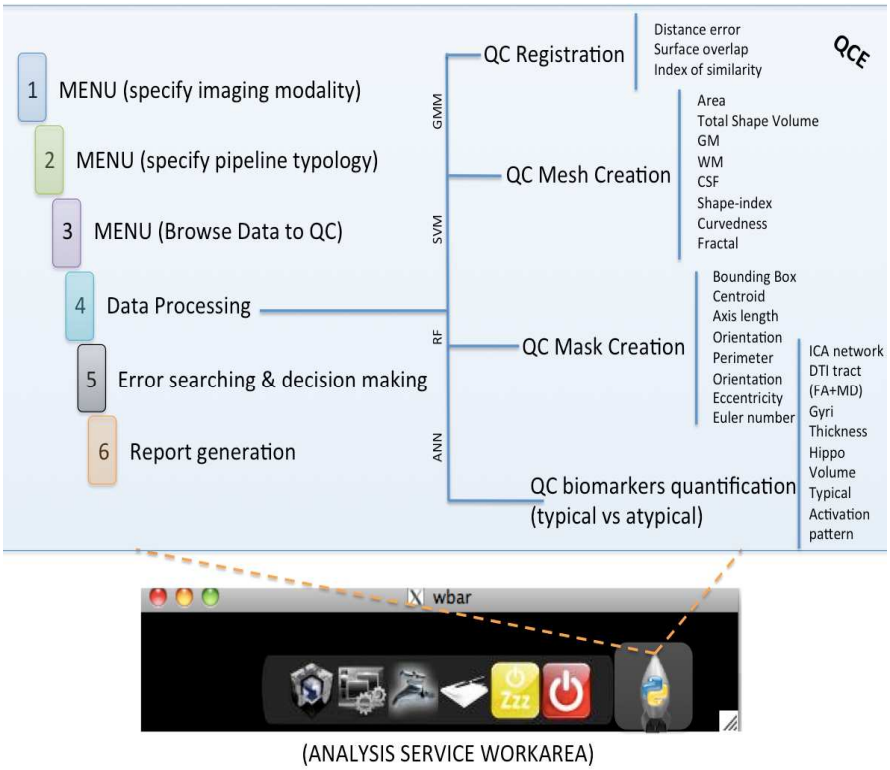

Figure 3: main steps of the QCE that users should be able to play with through specific interfaces.

The QCE has been thought to be of great granularity in the sense that it will be able to measure specific output properties and features (e.g.: Index of similarity, Surface Overlap, Total Shape Volume, Shape index, Bounding Box, Centroid, Axis length, Orientation, Euler Number, eccentricity, etc..) and to assess where the faults happen in the different steps of the pipelines (e.g.: Registration, mask/mesh creation, typical vs atypical biomarkers measurements). This tool has been designed (Figure 3) to enable end users to enhance the significance of the results obtained from the main N4U pipelines (e.g.: CIVET, ACM-AdaBoost, VBM, DTI-TK and Melodic) and to strength subsequent statistical analysis on the high-throughput results obtained by the neuGRID platform.

4. Conclusions

It should be noted that the current organization relies on a close collaboration between N4U and CATI. All the main datasets have been prioritized in order to clearly regulate the phase of uploading. The quality assurance of the dataset scans was defined and tested successfully. The OASIS and ADHD200 datasets are being uploaded at N4U after a close collaboration with MaatG and UWE partners. Last but not least, the development of the quality control environment (QCE) of the pipeline outputs relying on the machine learning approach is certainly an added value for the entire N4U infrastructure.

ANNEX I

Table 1. Image datasets relevant to the NS-NDD, NS-WMD, and NS-PSY communities. Datasets are categorized by community of interest and loading priority.

DataSet	*Community*	*Impact*	*Clinical epidemiological characteristics*	*Priority*	*Accessibility*	*Usage Example*
1000 Functional Connectomes Project (http://fcon_1000.projects.nitrc.org**)**	NS-NDD	This dataset will equip researchers with a means of exploring and refining rest-fMRI approaches.	Unrestricted public release of 1200 'resting state' functional MRI 4D-Images independently collected at 33 sites. All images form the broader imaging community complete access to a large-scale functional imaging dataset. Age, sex and imaging centre information are provided for each of the images. In accordance with HIPAA guidelines, all imagess are anonymous, with no protected health information included.	High	Public	Development of common paradigms for interrogating the myriad functional systems in the brain with *a priori* hypotheses.
ADHD200 (http://fcon_1000.projects.nitrc.org/indi/adhd200/**)**	**NS-NDD**	Attention Deficit Hyperactivity Disorder affects at least 5-10% of school-age, with annual direct costs exceeding $36 billion/year.	The ADHD-200 Sample is a grassroots initiative, dedicated to accelerating the scientific community's understanding of the neural basis of ADHD through the implementation of open data-sharing and discovery-based science. Towards this goal, we are pleased to announce the unrestricted public release of 776 resting-state fMRI and anatomical datasets aggregated across 8 independent imaging sites, 491 of which were obtained from typically developing individuals and 285 in children and adolescents with ADHD (ages: 7-21 years old). Accompanying phenotypic information includes: diagnostic status, dimensional ADHD symptom measures, age, sex, intelligence quotient (IQ) and lifetime medication status.	**High**	**Public**	Development of a model of the pathophysiology that help the clinical community to inform patients and guiding clinicians in the decision-making regarding treatment.
ADNI-GO (https://ida.loni.ucla.edu/login.jsp?project=ADNI)	NS-NDD	This dataset will better explore earlier stages of MCI; to study novel imaging markers	Expands ADNI01 (already ported in neuGRID) by 200 additional patients with early MCI; all patients undergo structural MRI scans at 3 T at four time points, amyloid imaging with a fluorinated ligand, resting0state fMRI in Philips scanners, diffusion MRI in GE scanners, and CSF studies	High	Public	Integrated combination of clinical–cognitive, CSF–plasma biomarker, MRI, amyloid–FDG0PET, and genetic measures for early diagnosis and Alzheimer's disease and dementia tracking
ADNI-2 (https://ida.loni.ucla.edu/login.jsp?project=ADNI)	NS-NDD	This dataset will extend the observation window of the MCI stage to earlier and later stages	ADNI02 will study and follow 500 additional individuals	High	Public	
OASIS (http://www.oasis-brains.org/**)**	NS-NDD	Aims to facilitate future discoveries in basic and clinical neuroscience	OASIS (Open Access Series of Imaging Studies) consists of (I) a cross-sectional collection of 416 subjects aged 18 to 96. 100 of the included subjects over the age of 60 have been clinically diagnosed with very mild to moderate Alzheimer's disease (AD). (II) A longitudinal collection of 150 subjects aged 60 to 96. Each subject was scanned on two or more visits, separated by at least one year for a total of 373 imaging sessions. For each subject, 3 or 4 individual T1-weighted MRI scans obtained in single scan sessions are included. The subjects are all right-handed and include both men and women. 72 subjects were nondemented, 64 were demented, including 51 individuals with mild to moderate Alzheimer's disease.	High	Public	Automated calculation of biomarkers to demonstrate the use of these data for assessing differences associated with normal aging versus Alzheimer's disease.

European ADNI – PHARMACOG (http://cordis.europa.eu/wire/index.cfm?fuseaction=article.Detail&rcn=19336**)**	NS-NDD	Identification of biomarkers sensitive to disease progression	The PharmaCOG dataset is composed by 150 subjects between the ages of 55 and 90 affected by amnestic Mild Cognitive Impairment (disease. All the subjects are divided in two groups based on their $A\beta_{1-42}$ levels: (I) low $A\beta_{1-42}$ (positive aMCI patients) and (II) High $A\beta_{1-42}$ (Negative aMCI patients). The threshold of $A\beta_{1-42}$ used to divide the subject will be 500 (ng/L). For all the subjects recruited in PharmaCOG the following information are collected: - Magnetic resonance imaging data such as T13D, T2*, DTI, rs-fMRI, amyloid PET images; - EEG / ERP records; - BLOOD related data; - CSF related data; - Bio-statistical data derived from the analytic processing steps.	High	Proprietary	Identification of biomarkers with sufficient sensitivity to track AD progression from its earliest stage.
LADIS database (www.ladis.eu**)**	NS-WMD	Assessment of disability in older people with age related changes in white matter	The Ladis dataset is composed by 639 non-disabled older patients (mean age 74.1) in whom brain magnetic resonance imaging showed mild, moderate, or severe age related changes in white matter	High	Proprietary	Understanding of the transition from no disability to disability.
EDSD database	NS-WMD	Assessment of disability in older people	The EDSD data bank includes 335 Diffusion Tensor Imaging (DTI) and 335 structural MRI scans (MPRAGE) from patients with AD and healthy elderly subjects.	High	Proprietary	Understanding of the transition from no disability to disability.
Multisite Imaging Research In the Analysis of Depression (MIRIAD) (http://nirlarc.duhs.duke.edu/nirle/**)**	NS-PSY	Multiple institution effort for automated analyses of PD and T2-weighted MR from a longitudinal study of depression in late life	A multiple institution study of structural MRI, including raw PD and T2 MRIs and derived measures of white matter changes, basal ganglia and other regions. Demographic and extensive clinical assessment data is available for each case.	High	Public	MIRIAD will be used to track regions of interest and cerebral tissues involved in the depression in elderly patients.
fBIRN Phase II (http://fbirnbdr.nbirn.net:8080/BDR/**)**	NS-PSY	A multi-centric effort for the characterization of schizophrenia diasease	The FBIRN multi-site dataset of subjects with schizophrenia and controls includes functional MRI images, behavioral data, demographic, and clinical assessments on 253 subjects from around the US. Subjects were recruited locally, evaluated with a standardized clinical assessment battery, and scanned twice at each of the 9 participating FBIRN sites. A standard FBIRN protocol was used for image acquisition with controlled variation at different sites and included a sensorimotor task (SM), an Auditory Oddball task (AO), a Serial Item Recognition Paradigm (SIRP), and a breathhold task (BH).	High	Public	Collection, storage, sharing and management of large number of images aimed to characterize the Schizophrenia pathology.
Efficient Longitudinal Upload of Depression in the Elderly (ELUDE) (http://nirlarc.duhs.duke.edu/nirle/**)**	NS-PSY	Longitudinal study of late-life depression	There are 281 depressed subjects and 154 controls included. An MR scan of each subject was obtained every 2 years for up to 8 years (total of 1093 scans). Clinical assessments occurred more frequently and consists of a battery of psychiatric tests including several depression-specific tests.	High	Public	ELUDE will be analyzed to better characterized the depression in elderly patients.
Multisite Imaging Research In the Analysis of Depression (MIRIAD) (http://nirlarc.duhs.duke.edu/nirle/**)**	NS-PSY	Multiple institution effort for automated analyses of PD and T2-weighted MR from a longitudinal study of depression in late life	A multiple institution study of structural MRI, including raw PD and T2 MRIs and derived measures of white matter changes, basal ganglia and other regions. Demographic and extensive clinical assessment data is available for each case.	High	Public	MIRIAD will be used to track regions of interest and cerebral tissues involved in the depression in elderly patients.

Human Imaging database (HID) (http://nbirn.net/tools/human_imaging_schema/index.shtm**)**	NS-NDD	Extensible database management system developed to handle the increasingly large datasets collected as part of the FBIRN	HID contains several datasets including (I) BrainScape Resting State fMRI Dataset 1: This dataset is characterized by seventeen healthy subjects with four resting state fixation scans plus one T1 scan and one T2 scan. The data were collected as part of a study on the behavioral effects of spontaneous BOLD fluctuations. (II) BrainScape Resting State fMRI Dataset 2: This dataset includes 10 healthy subjects scanned 3 times with 3 conditions: eyes open, eyes closed, and fixating. Imaging data are the primary data type available with a minimal set of descriptive metadata and, in some cases, associated clinical, genetic, or other biomedical data.	Medium	Public	Collection, storage, querying, sharing, management, tracking and analysis of large image datasets aimed to characterize brain function
Italian Brain Normative Archive of Normal Aging (IBNA)	NS-NDD	A robust reference for studies of brain morphometry in clinical populations	Database of normal aging subjects with 900 cases. This is a national archive of MR images in digital format with clinical and genetic data.	Medium	Proprietary	Correlation of variables to understand the phases of a normal aging
MIDAS (CasiLab) (http://www.insight-journal.org/midas/community/view/21**)**	NS-NDD	Database of 100 healthy subjects divided by age and sex.	100 patients divide into five group age (decades) and by sex, handedness and race. It consists of T1 and T2 acquired at 1x1x1 mm^3 at 3T, Angiography (MRA) acquired at 0.5 x 0.5 x 0.8 mm^3, and Diffusion Tensor Imaging (DTI) using 6 directions and a voxel size of 2x2x2 mm^3	Medium	Public	This database may prove useful in assessing the effects of healthy aging
MAGNIMS (http://www.magnims.eu/**)**	NS-WMD	Aims to facilitate discoveries in the MS inflammatory mechanism	MAGNIMS (magnetic resonance in multiple sclerosis) is an European multicenter study group on the relationship between inflammatory parameters and atrophy/neurodegeneration in 100 multiple sclerosis patients assessed at multiple time points in UK, NL, SP, IT, SW, AU, DK. These patients are divided into groups well defined on the basis of clinical and neuroradiological parameters.	Medium	Proprietary	Development of MS biomarkers using very innovative methodologies.
MSDN (https://www.imedweb.it/**)**	NS-WMD	The primary impact is the definition of guidelines for the treatment and the improvement of patients' life with Multiple Sclerosis	Multicenter database of Multiple Sclerosis patients from 26 Italian national centers. The data are anonymous. MSDN is characterized by a sample of 10078 patients. These subjects are very similar to other national and international studies described and this confirm that this MS cohort is highly representative of the disease. 6282 patients had at least two or more visits with on EDSS assessments and a follow-up of at least one year. 3500 (55%) of them are treated with immunomodulators / immunosuppressants. Approximately 39% of patients are treated with IFNβ therapy for a period longer than 2 years and 29% for a period longer than 4 years.	Medium	Proprietary	Collection, storage, querying, sharing, management and analysis of large images aimed to characterize the MS pathology.
Neuro-imaging Research Group (Utrecht database)	NS-PSY	Study the patho-physiological plasticity of the brain	Cohorts of healthy subjects, patients with psychiatric disorders, and their family members, including monozygotic and dizygotic twins and singleton siblings followed longitudinally at a 2-5 year interval. To date, the continuously growing database consists of over 3'000 MRI brain scans.	Medium	Proprietary	Structural MRI investigation on schizophrenia.
Sylvia Lawry Centre Dataset for MS Research (daumer@slcmsr.org**)**	NS-PSY	Collection of 45 databases supplied by pharmaceutical companies, universities, clinicians and researchers.	Database of MS patient information (over 20000). It comprises virtually all the MS patients who participated in the placebo arms of major clinical trials in the last 20 years.	Medium	Proprietary	Better refinement of MRI markers of disease activity and progression.
Bipolar Disorder Neuro-imaging	NS-PSY	Meta-analysis and database of MRI	The database contains information of 141 studies which have investigated brain structure (using MRI and CT scans) in patients with bipolar disorder compared to a control group.	Low	Public	To investigate structural brain changes in bipolar disorder and

Database (BiND) (http://sites.google.com/site/bipolardatabase/)		studies	Ninety-eight studies and 47 brain structures are included in the meta-analysis.			assess the effect of medication use and demogr.and clinical variables.
Bipolar Disorder Phenome Database (http://bioinformoodics.jhmi.edu/gm/phenome/)	NS-PSY	Johns Hopkins and NIMH researchers have jointly created this database, which posts the clinical phenotypes of over 5,000 people recruited for bipolar disorder genetics studies.	It consists of 5,721 subjects (3,186 affected) in 1,177 families, 197 variables, and 1,127,037 data points. There are various possibilities for new research with this database, but it should be remembered that this is only a phenomenological database.	Low	Public	Researchers can explore connections between clinical variables and genetics with adequate numbers of subjects to detect even moderate genetic effects
Brainweb-Simulated Brain Database (http://mouldy.bic.mni.mcgill.ca/brainweb/)	NS-NDD	Validation of neuroimaging analyses on *in vivo* acquired data	SBD contains simulated brain MRI data based on two anatomical models: normal and MS. For both of these, full 3-dimensional data volumes have been simulated using three sequences (T1-, T2-, and proton-density- (PD-) weighted) and a variety of slice thicknesses, noise levels, and levels of intensity non-uniformity. These data are available for viewing in three orthogonal views (transversal, sagittal, and coronal), and for downloading.	Low	Public	Validation of new algorithms for neurodegenerative diseases
Major Depressive Disorder Neuroimaging Database (MaND) (http://sites.google.com/site/depressiondatabase/**)**	NS-PSY	Meta-analyses and database of MRI studies	The database contains information of 225 studies which have investigated brain structure (using MRI and CT scans) in patients with major depressive disorder compared to a control group. 143 studies and 63 brain structures are included in the meta-analysis.	Low	Public	To investigate Major depressive disorder with MRI and CT scans