How does the national centre for Excellence in Research on Parkinson's disease benefit from forward-thinking IT infrastructure and data capture?

Dr. Reinhard Schneider

Luxembourg Centre for Systems Biomedicine (LCSB)

LCSB in a nutshell

Interdisciplinary research centre of the University of Luxembourg

Biomedical and Systems Biology Research

Aim: *Personalized medicine*

part of the Biohealth Initiative of Luxembourg

Founded in Sept. 2009

Scientific strategy of LCSB

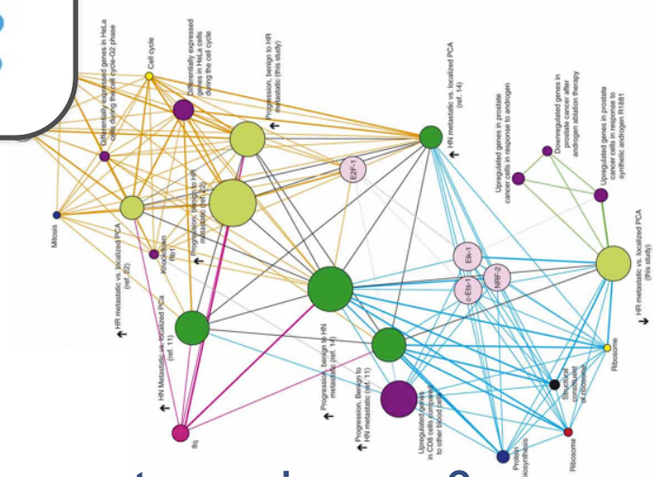
Clinical translational research

Experimental Biology

LCSB

Discoveries

Technology platforms



Computer science & Bioinformatics

Bioinformatics Core facilities (30+ FTE's)

Data integration and management
Organize, store and categorize large amount of data (PetaByte scale). Providing access and management to large compute farms.

Automatics pipelines for larges scale data-analysis
Setup of automatic procedures to filter and extract the most relevant information out of large heterogeneous datasets

Network (re-)construction
Extract known and predicted networks (protein, protein-protein, protein-chemical,...) from databases and by applying text-mining technologies

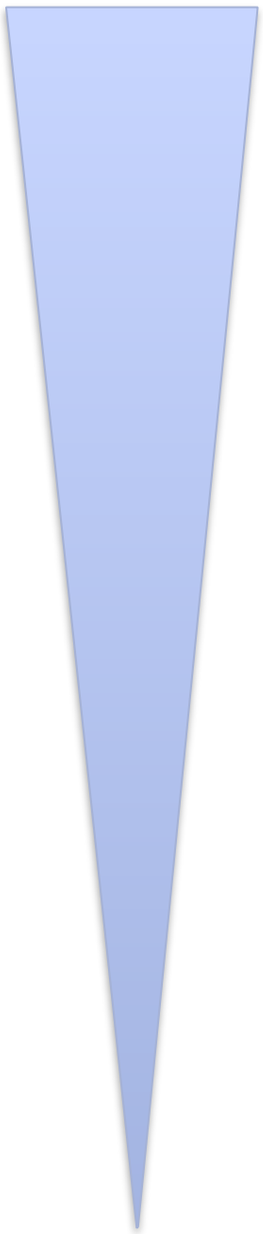
Large scale visualization tools for heterogenous data
Development of 2D and 3D visualization tool for data exploration and hypothesis generation

Text-mining
Crunching large scale full-text corpora of hundreds of thousand of articles to extract knowledge map and relationships between diseases, genes, proteins etc.

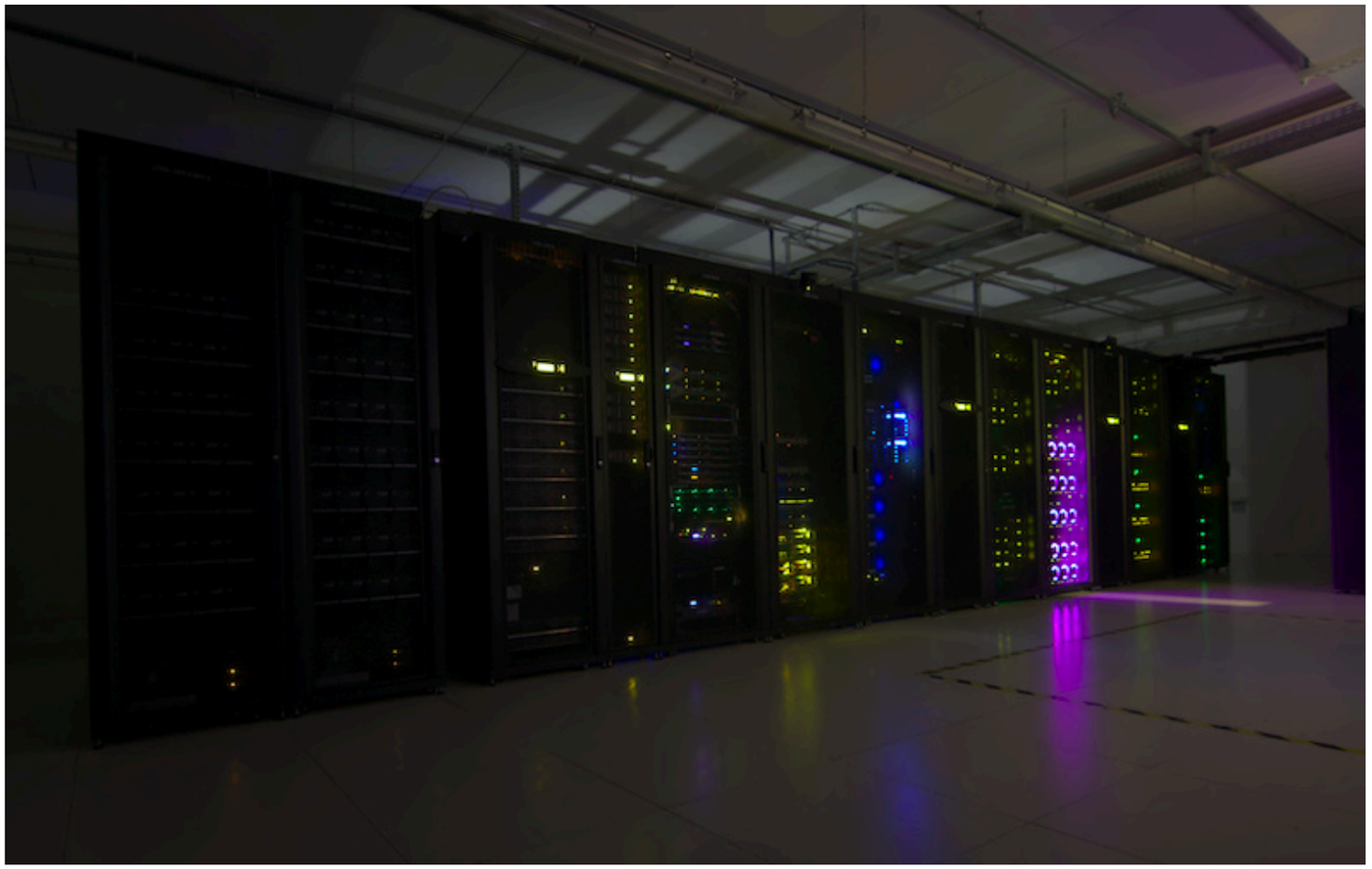
Data Analysis partner in several projects
FP7: eTRIKS, EpiPGX; coGIE; betaJUDO

Development of dedicated problem oriented research tools

Challenges in **translational medicine**



- Data quality
- Data standards
- Data openness, sharing, access, legal issues
- Data protection, data privacy
- Ontologies, matching vocabulary
- Amount of data
- Analysis of data



Cluster with >5000 cores
Several large memory machine 1-4 TB RAM
~5 PetaByte storage

Genome Pipeline

Ultra-Rapid Genome and Exome Data Analysis

22min
FASTQ → VCF
WHG @ 30x Coverage

99.37%
Precision
SNP + INDEL Combined

98.31%
Sensitivity
SNP + INDEL Combined

UNIVERSITÉ DU LUXEMBOURG

LCSB

"healthy" "diseased"

Cell types and organs

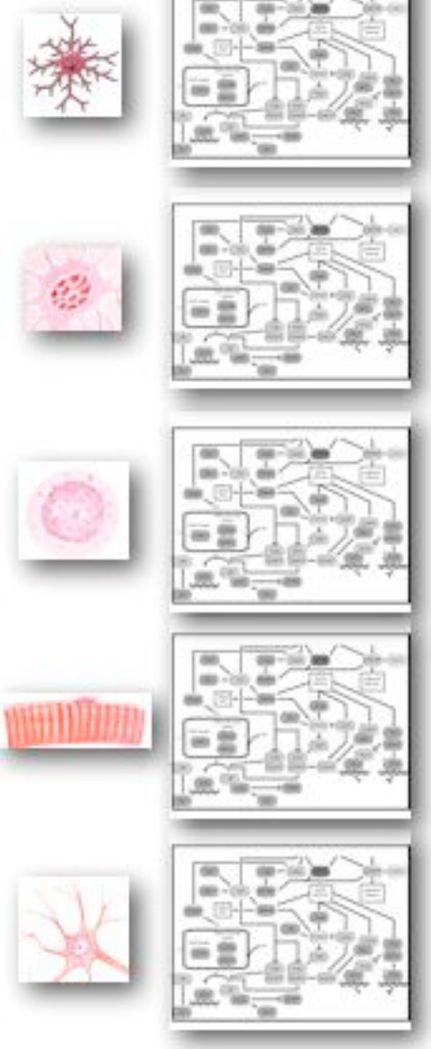
"normal" cellular biochemical state

"perturbed" cellular biochemical state

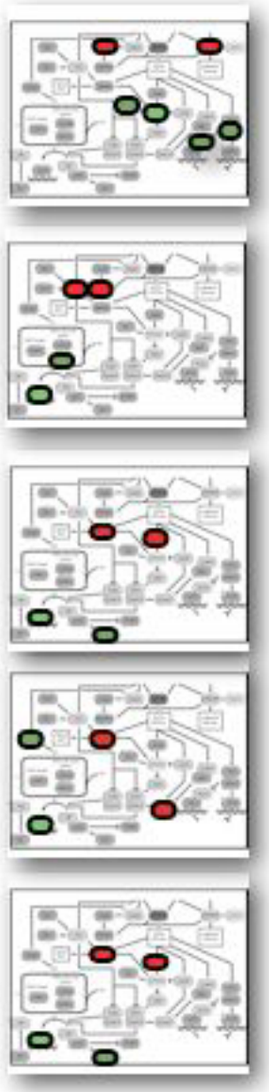
Time

higher concepts of molecular dysfunction, disease phenotypes and symptoms

Diseases and medical treatments



Perturbation of the system by genomics, environmental or life style factors



minutes, days, weeks or many years

Apotosis

Anxiety

bradykinesia

Epilepsy

blurry vision

weight loss

polyuria

Diabetes

Fatigue

dysphagia

atrophy

ALS

fasciculations

protein degradation

Dementia

Alzheimer

pneumonia

Huntington

mitochondrial dysfunction

orthostatic hypotension

Tremor

Parkinson

rigid muscle

REM behavior disorder

Biological Ontologies

missing links

Medical Ontologies

Data Tsunami

Bioinformatician sorting out the ”good stuff”

Curation

Making biological knowledge machine-readable and machine-interpretable

LCSB

UNIVERSITÉ DU LUXEMBOURG

Pets

~6500 pets reported
600 distinct pets reported
When filtered by:
“dog”, 112, misspellings
and dog-like entries

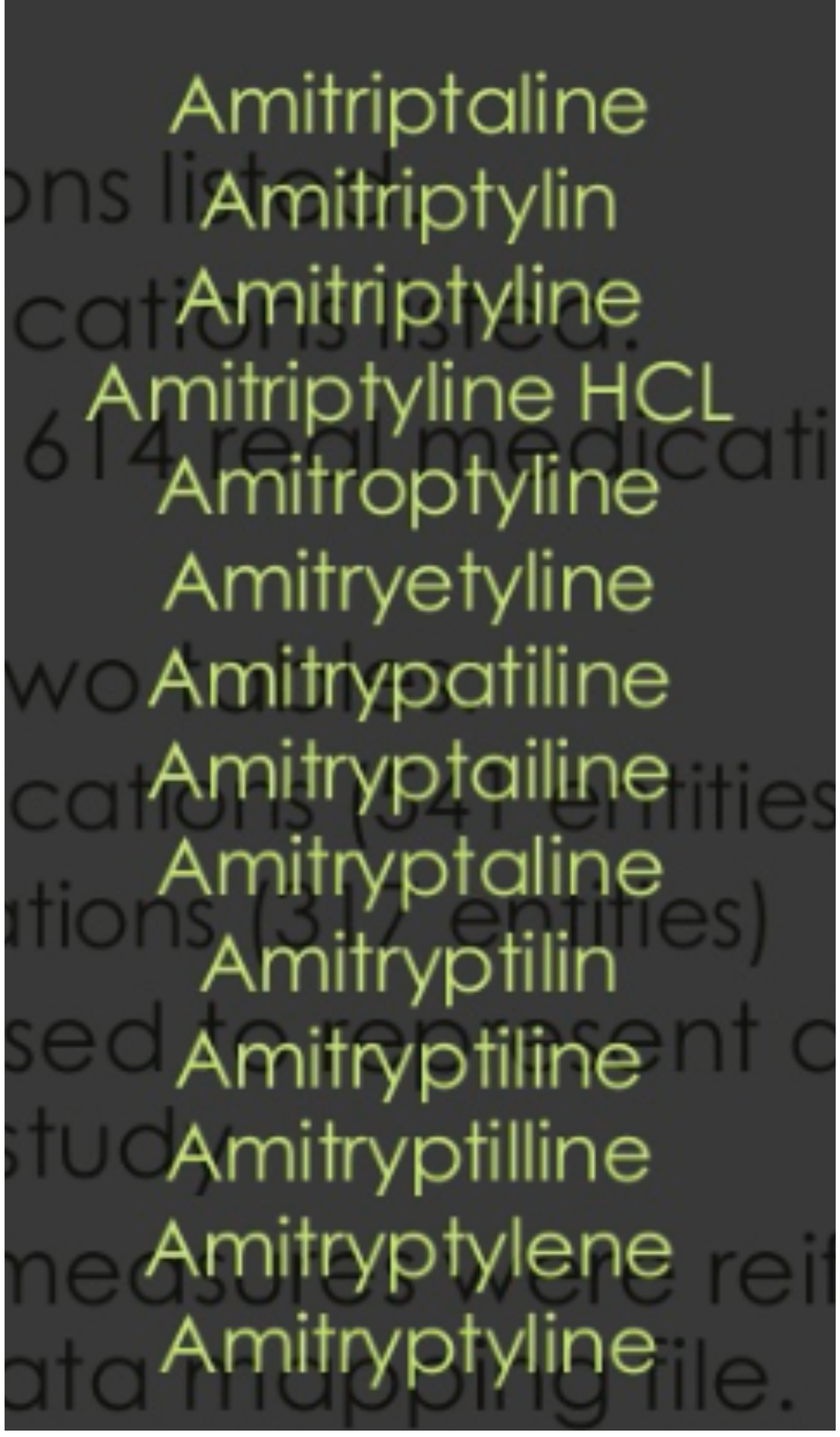
All pets mapped to a 31 term vocabulary

Alligator	Crayfish	Possum
Amphibian	Dog	Rabbit
Assorted	Farm Animal	Raccoon
Badger	Ferret	Reptile
Bird	Fish	Rodent
Butterflies	Harbor Seal	Skunk
Cat	Hedgehog	Snake
Chameleon	Hermit Crab	Spider
Chimpanzee	Lizard	Sting Ray
Chinchilla	Monkey	Turtle
		Worm

Accelerated Cure Project Multiple Sclerosis, S. Wicks; http://www.acceleratedcure.org

Medication

misspellings



Accelerated Cure Project Multiple Sclerosis, S. Wicks; http://www.acceleratedcure.org

Standard Blood Test in University Hospitals

Measurment	Lab 1	unit	Lab 2	unit	Lab 3	unit	STATUS	comment
Adiponectin	☹		☹	µg/mL	☺	mg/L	?	
ALAT (alaninaminotransferase)	☺	ukat/L	☹		☺	ukat/L	OK	
Albumin	☺	g/L	☺	%	☺	g/L	OK	is it really %?
Apolipoprotein A1	☺	g/L	☺	mg/dL	☺	g/L	OK	
Apolipoprotein B	☺	g/L	☺	mg/dL	☺	g/L	OK	
Apolipoprotein B/ApoA1 ratio	☺	-	☺	-	☺	-	OK	was calculated for S and L
ASAT (aspartateaminotransferase)	☺	ukat/L	☹		☺	ukat/L	OK	
Cholesterol total	☺	mmol/L	☺	mg/dL	☺	mmol/L	OK	
C-Peptide	☺	nmol/L	☺	**pmol/L?**	☺	nmol/L	OK	or pmol/mL?
C-Reactive Protein hs-CRP	☺	mg/L	☺	mg/dL	☺	mg/L	OK	
Creatinine	☹		☺	mg/dL	☺	umol/L	OK	
Cystatin C	☺	mg/L	☺	mg/L	☺	mg/L	OK	
Estradiole E2	☺	pmol/L	☺	pg/mL	☺	pmol/L	OK	
Follicle Stimulating Hormone FSH	☺	IE/L	☺	mU/mL	☺	U/L	OK	
FT3 (free trijodtyronin)	☹		☺	pmol/L	☺	pmol/L	OK	
FT4 (free tyroxin)	☹		☺	ng/dL	☺	ng/dL	OK	
GFR (Glomerular filtration rate)	☺	mL/mi/1,73	☹		☹		?	
Glucose	☺	mmol/L	☺	mg/dL	☺	mmol/L	OK	
Glutamyltransferase GGT	☺	ukat/L	☺	U/L	☺	ukat/L	OK	
GOT glutamic oxaloacetate transaminase			☺	U/L			?	
GPT glutamic pyrovate transaminase			☺	U/L			?	
HbA1c			☹	%			?	exclude?
HDL-cholesterol	☺	mmol/L	☺	mg/dL	☺	mmol/L	OK	recalc
Hemolysis	☺	-						
Ikterus (bilirubin)	☺	-						
Insulin	☺	mE/L	☺	µU/mL	☺	pmol/L	?	
Interleukin 1-beta IL-1b			☹					exclude?
Interleukin 6 IL-6			☺	?				
LDL-cholesterol	☹		☺	mg/dL	☺	mmol/L	?	
Leptin LEPT	☹		☺	ng/mL	☺	ng/mL	OK	
LH Lutropin	☺	IE/L	☺	mU/mL	☺	U/L	OK	
Lipoprotein A LPA	☹		☺	mg/dL	☺	g/L	OK	
Proinsulin	☺	pmol/L	☺	pmol/L	☺	pmol/L	OK	Lab 1 high
Prolactin PROL	☹		☺	µU/mL	☺	mU/L	OK	wrong units?
SHBG (steroid hormone binding globulin)	☺	nmol/L	☺	nmol/L	☺	nmol/L	OK	
Testosterone	☺	ng/mL	☺	ng/mL	☺	nmol/L	OK	recalc
TNF-alpha (tumour necrosis factor)			☺	?				
Triglyceride	☺	mmol/L	☺	mg/dL	☺	mmol/L	OK	recalc
TSH (tyroid stimulating hormone)	☺	mIE/L	☺	mU/L	☺	mU/L	OK	
UREA (Harnstoff HST)	☹		☺	mg/dL	☺	mmol/L	OK	recalc
Uric acid (Urat Harnsaure HRS)	☺	umol/L	☺	mg/dL	☺	umol/L	OK	recalc
UU - Pt-GFR(CystC-beräkn) (see also GFR)	☺	mL/mi/						
UU - GLP-1 (glucagon like peptide -1 in P800 tubes)	☹	?					when?	
UU -FFA Free fatty acids	☺	mg/dL						
UU -steroid hormone pattern (12 different)	☺	pg-ng/mL						

☺	= values reported
☹	= values NOT reported
?	= values reported but without unit

STATUS = OK (good correlation between the reporting laboratories, at least 2/3), OK (probably fine but something unclear), ? (not ready),

UNIVERSITÉ DU LUXEMBOURG

SB

Data curation

Information Technology for Translational Medicine

Data cleansing

- Errors/inconsistencies/ambiguities
  - Replacing semantically identical concepts by unique concept
    - "NA", "n/a" and "not applicable" -> *"Not Applicable"*
  - Consistent data type for given variable
    - Numeric variable: "2.3"; ">2.9"; "NA"; 2.1
    - Date-type variable: "23-12-2015"; "12- 23-2015"
  - Eliminating typos
    - Dexamethasone, Dextramethason, DEXAMETHASONE -> *Dexamethasone*
  - Detecting out of range numeric values
    - Diastolic blood pressure: "640 mmHg"
- Converting to standard units
  - Blood glucose level: "110 mg/dl" -> "6.1 mmol/L"

PARKINSON

National Centre of Excellence in Research

8-year research program

NCER-PD: early-stage diagnosis of Parkinson's disease (PD) and improvement in the stratification of PD will focus on the identification and validation of clinical and molecular traits (biomarker signatures) of PD patients to improve diagnosis and characterization of their condition.

Partners:
Luxembourg Centre for Systems Biomedicine (LCSB) (coordinator)
The Luxembourg Institute of Health (LIH),
The Integrated BioBank of Luxembourg (IBBL)
Centre Hospitalier de Luxembourg (CHL).

uni.lu UNIVERSITÉ DU LUXEMBOURG

LCSB

NCER-PD Web Portal

- Web-server was setup with Joomla
- Available at www.ncer-pd.lu

HELP-PD Cohort

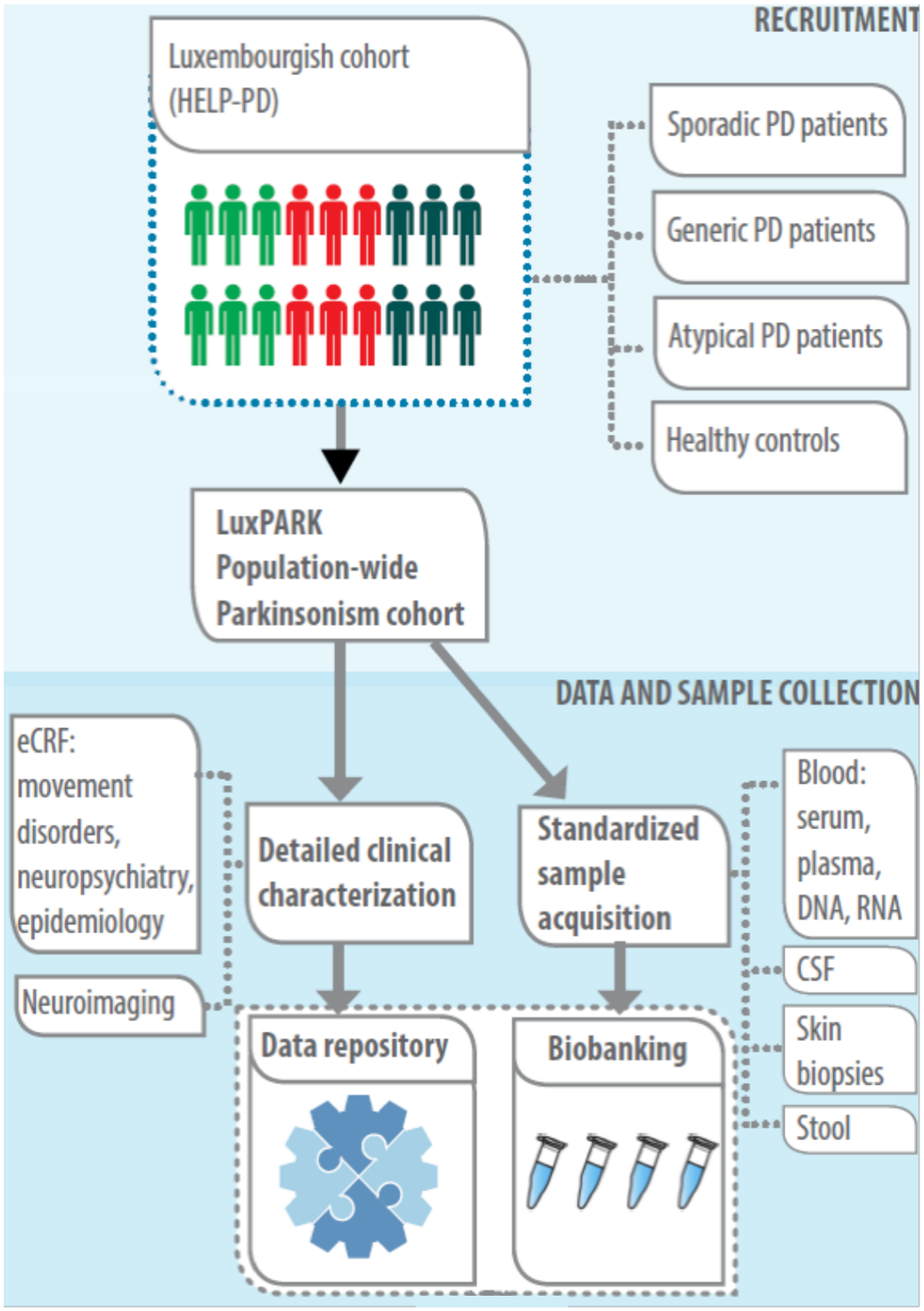
Prospective recruitment over 4 years

Longitudinal follow-up

Clinical data

NON-MOTOR

- Motor
- Cognition
- Neurosensory
- Psychiatric
- Sleep
- Autonomic
- Quality of Life
- Environment



800 patients with typical and atypical PS

800 Control subjects

Biosampling

Blood: DNA, RNA

Urine

Saliva

Nasal Washes

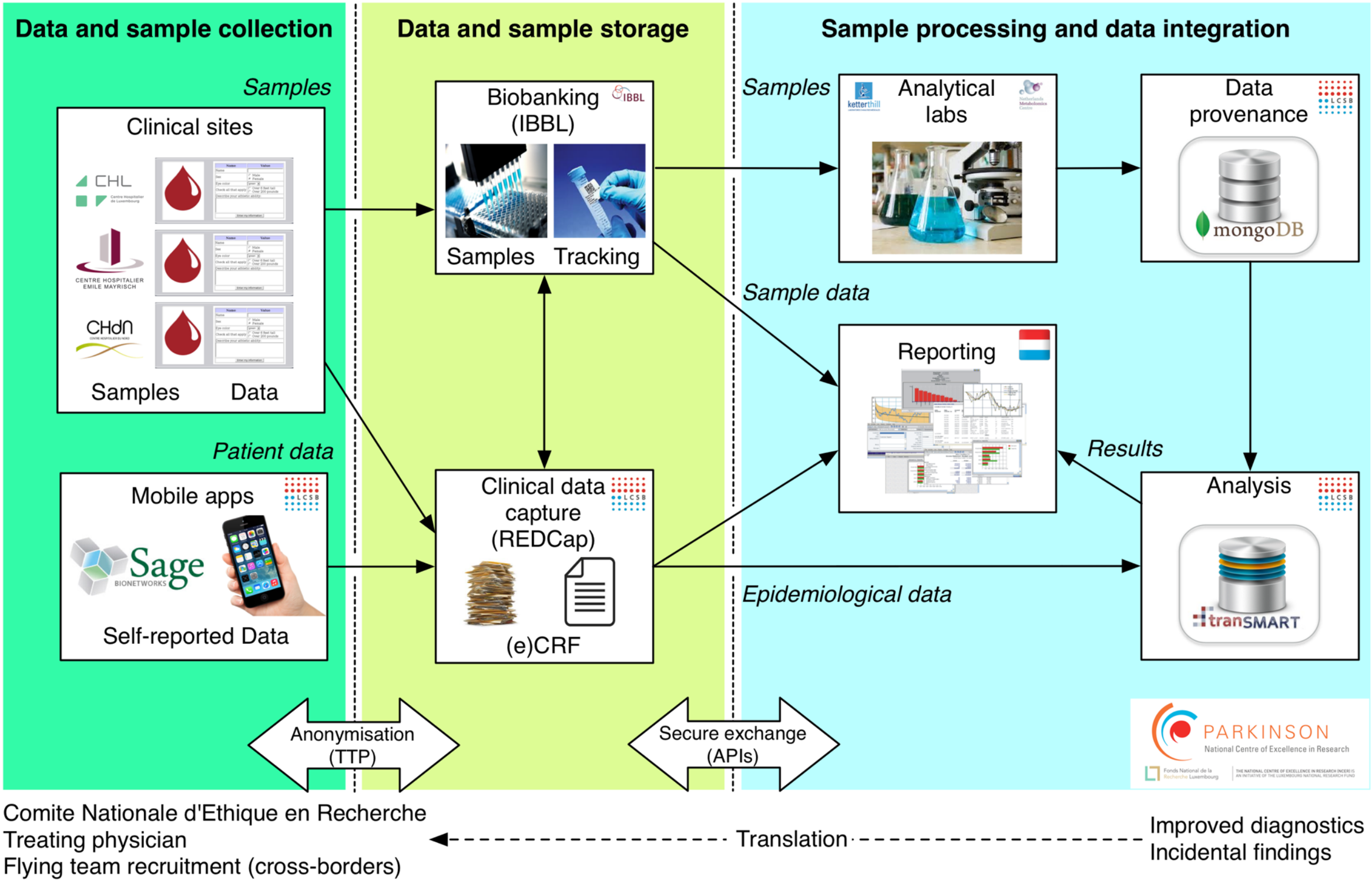
Stool

CSF

Skinbiopsy

Colonbiopsy

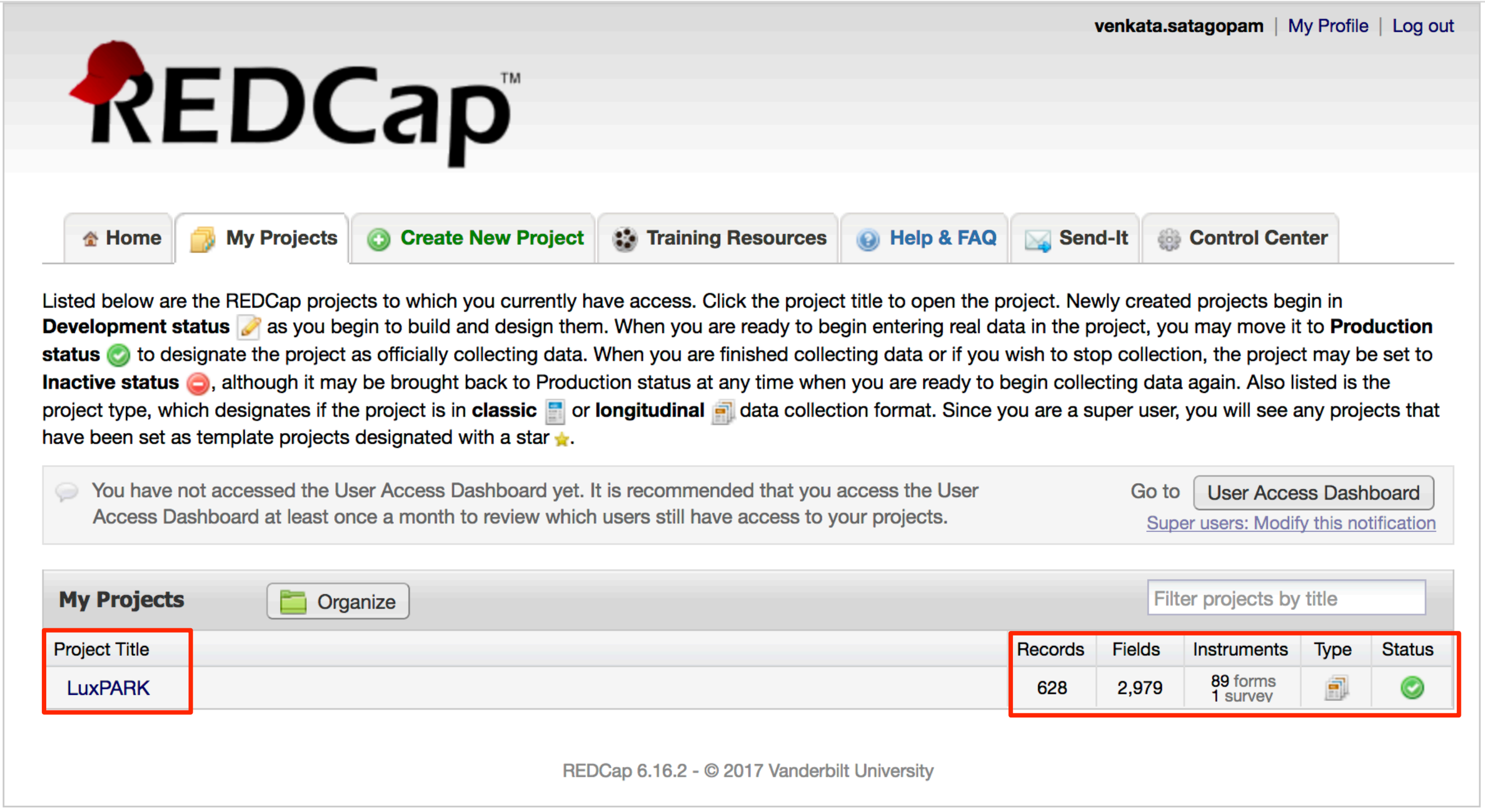
Example: NCER-PD



REDCap Server

- An Electronic Data Capture (EDC) system
- LxPARK eCRF (electronic Case Report Form) has been setup in French, German and English languages
- Available at https://pd-redcap.uni.lu

LuxPARK eCRF in REDCap



MDS-UPDRS - Part III: Motor Examination

3a Is the patient on medication for treating the symptoms of Parkinson's Disease? ☐ Yes ☐ No

3b If the patient is receiving medication for treating the symptoms of Parkinson's Disease, mark the patient's clinical state using the following definitions:

☐ ON: On is the typical functional state when patients are receiving medication and have a good response.
☐ OFF: Off is the typical functional state when patien ts have a poor response in spite of taking medications.

3c Is the patient on Levodopa? ☐ Yes ☐ No

3.C1 If yes, minutes since the last levodopa dose: _____

3.1 Speech

☐ 0: Normal: No speech problems.
☐ 1: Slight: Loss of modulation, diction or volume, but still all words easy to understand.
☐ 2: Mild: Loss of modulation, diction, or volume, with a few words unclear, but the overall sentences easy to follow.
☐ 3: Moderate: Speech is difficult to understand to

MDT-PD (Munich Dysphagia Test - Parkinson's Disease)

This is a test for subjects with a PSP diagnosis only!

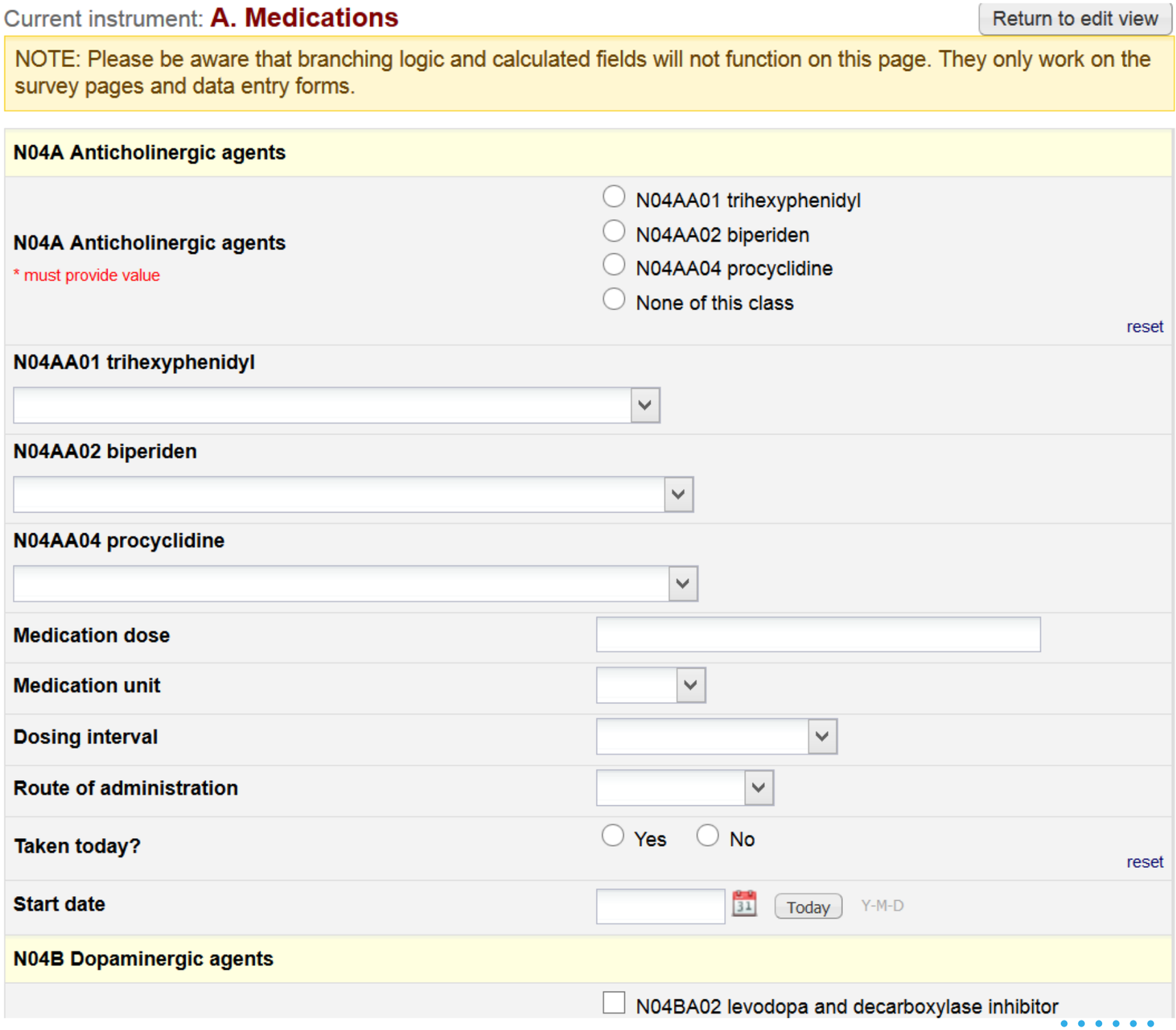
I DIFFICULTY SWALLOWING food and liquids

1. I have difficulties with the chewing and swallowing of solid/ fibrous/ crumbly food. (e.g.apples, meat, cracker/ chips)

☐ (almost) never
☐ occasional/monthly (once/multiple)
☐ frequently/weekly (once/multiple)
☐ very often/daily (once multiple)

Ontologies

- Example: Medications
- WHO ATC /DDD Index is being used:
  - ATC code
  - ATC level name (INN/ generic name)



UNIVERSITÉ DU LUXEMBOURG

Data Quality

Data Quality

This module will allow you to execute data quality rules upon your project data to check for discrepancies in your data. Listed below are some pre-defined data rules that you may utilize and run. You may also create your own rules or edit, delete, or reorder the rules you have already created. To find discrepancies for a given rule, simply click the Execute button next to it, or click the Execute All Rules button to fire all the rules at once. It will provide you with a total number of discrepancies found for each rule and will allow you to view the details of those discrepancies by clicking the View link next to each. Read more detailed instructions.

Data Quality Rules — **Execute rules:** All | All except A&B | All custom | Clear

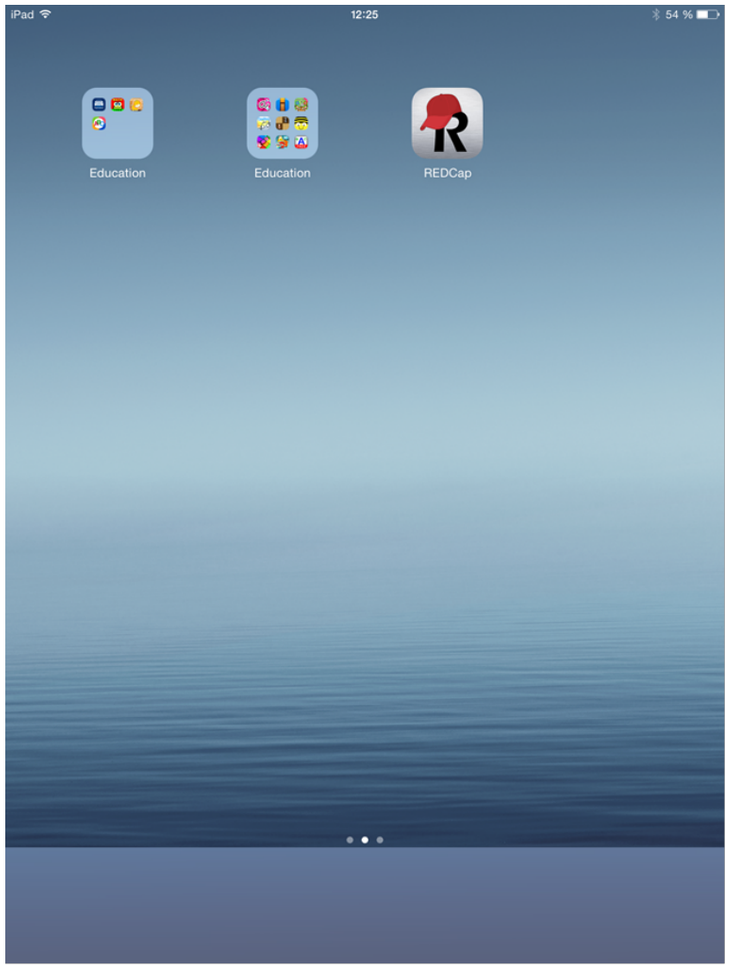
	Rule #	Rule Name	Rule Logic (Show discrepancy only if...)	Real-time execution ?	Total Discrepancies	Delete rule?
	A	Missing values*	-		Execute	
	B	Missing values* (required fields only)	-		Execute	
	C	Field validation errors (incorrect data type)	-		Execute	
	D	Field validation errors (out of range)	-		Execute	
	E	Outliers for numerical fields (numbers, integers, sliders, calc fields)	-		Execute	
	F	Hidden fields that contain values**	-		Execute	
	G	Multiple choice fields with invalid values	-		Execute	
	Add	Enter descriptive name for new rule (e.g., Participants below age 18)	Enter logic for new rule (e.g., [age] < 18) How do I use special functions?	Execute in real time on data entry forms ?		

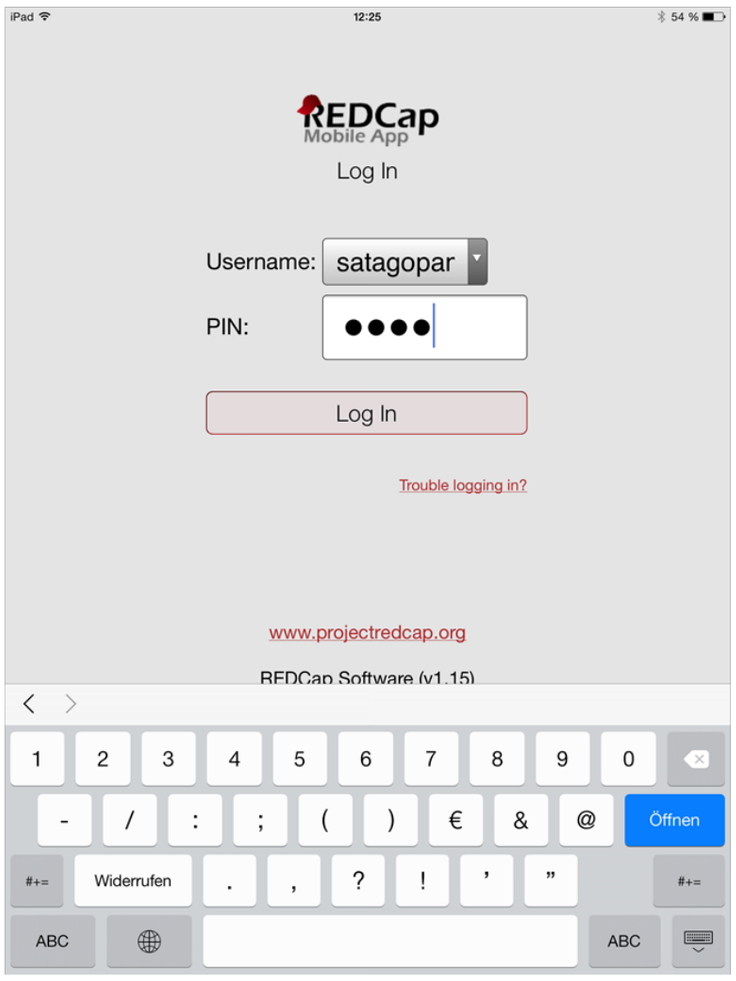
UNIVERSITÉ DU LUXEMBOURG

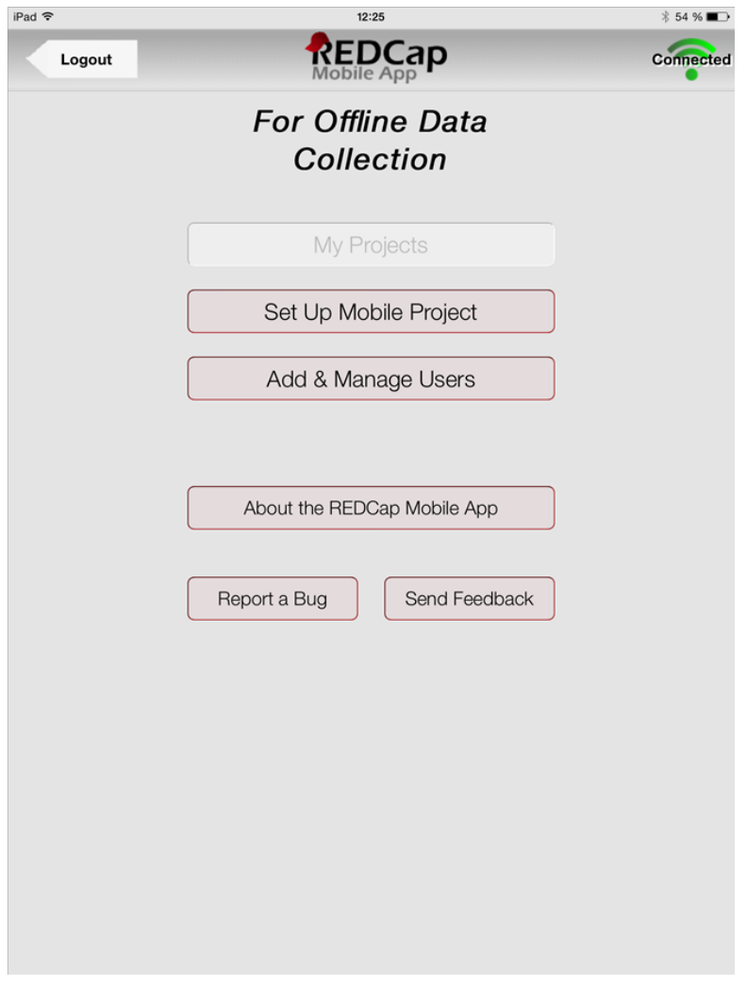
LCSB

REDCap Mobile App

- It facilitates collection of data offline and sync with online REDCap once connected to internet



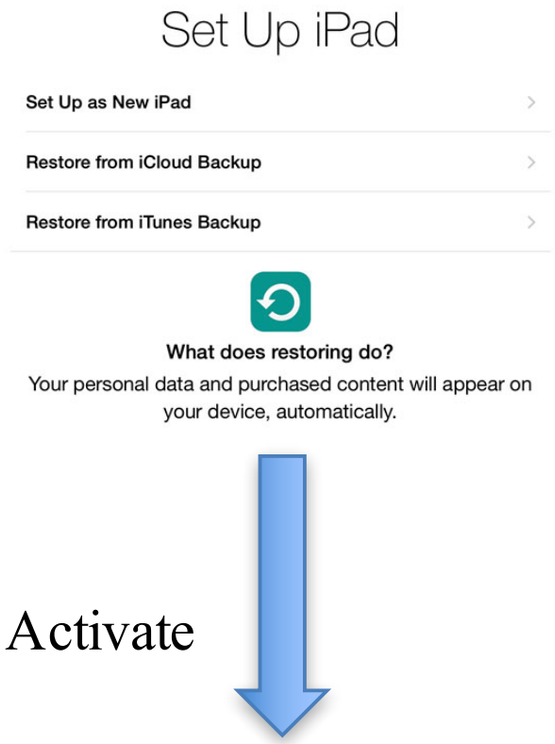




Enrollment of iPads

Unpack

Setup



Activate

Secured iPad by LCSB

MDM

MobileIron

Apple DEP

DEP

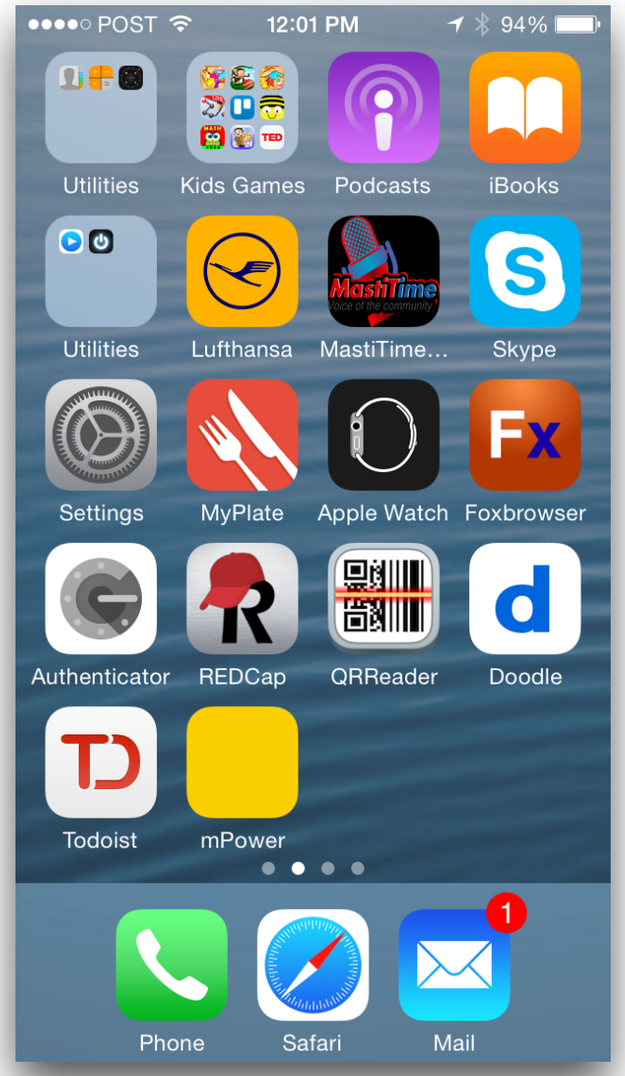
Device Enrolment Program

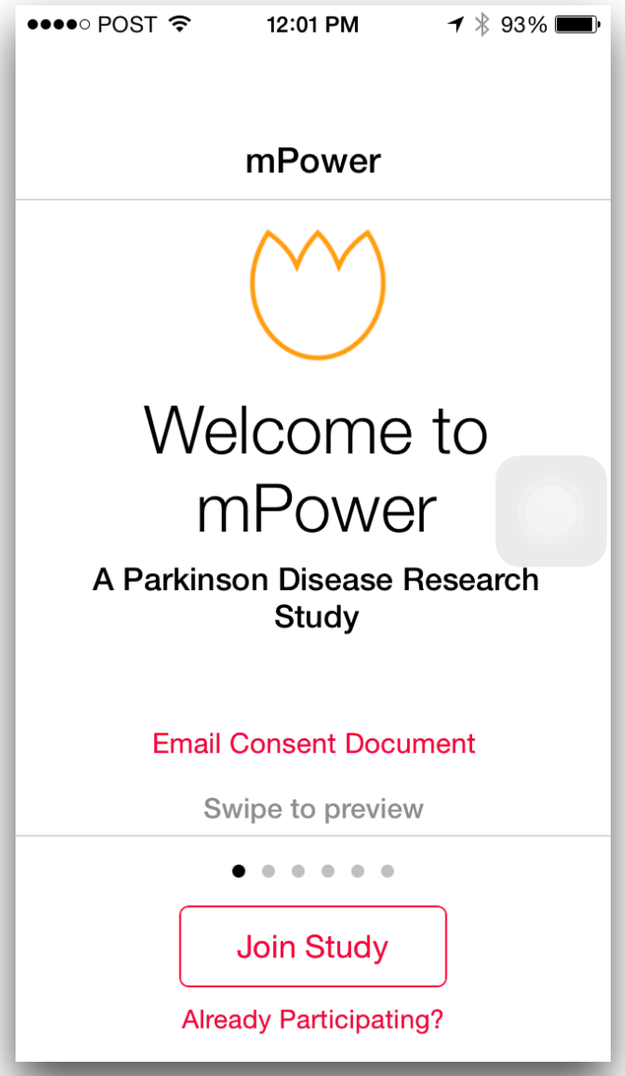
UNIVERSITÉ DU LUXEMBOURG

LCSB

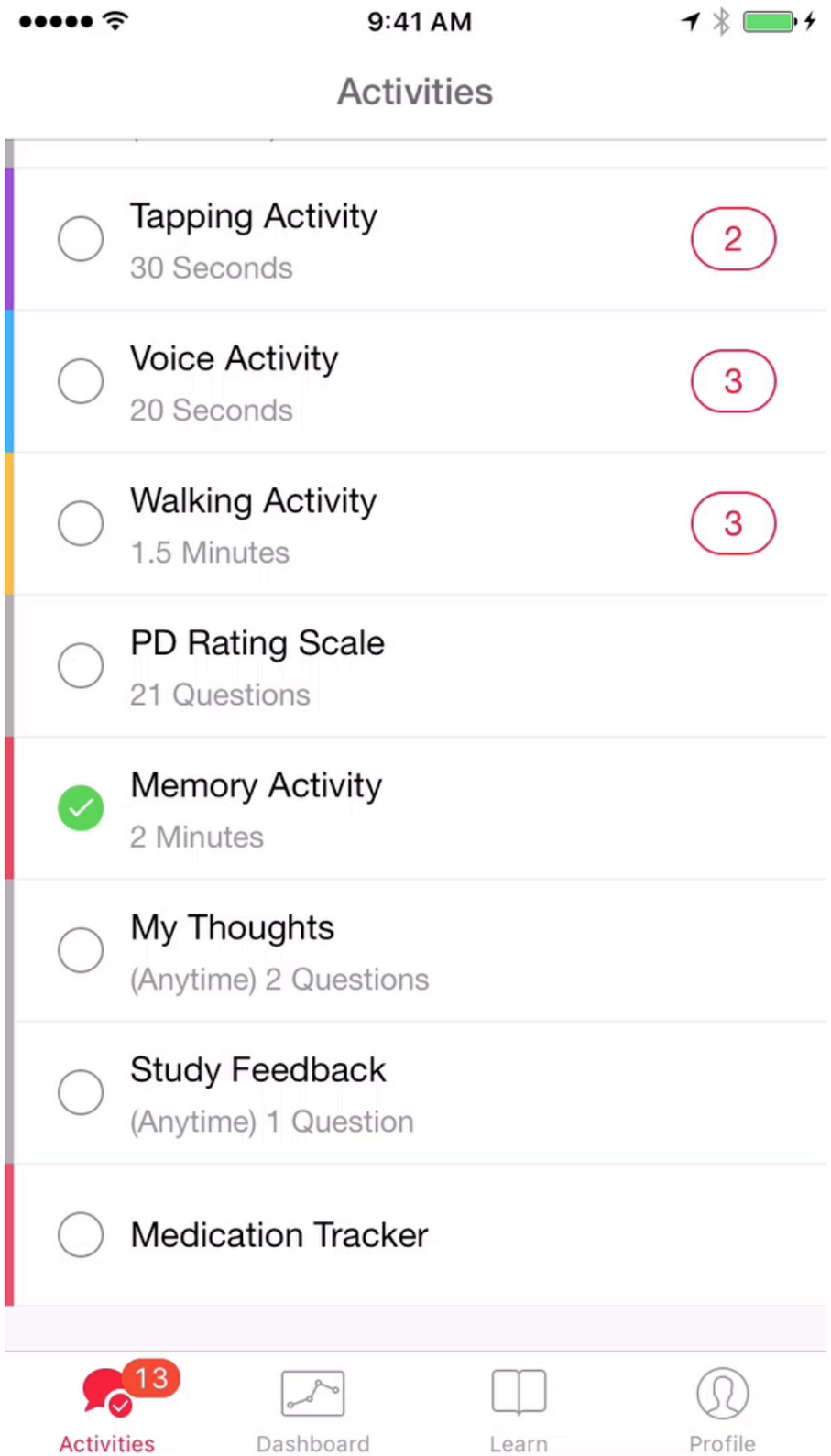
Wearable Medical Devices

- Collaboration with SAGE Bionetworks to use mPower (Mobile Parkinson Observatory for Worldwide Evidenced-based Research) mobile app





HELP-PD Ongoing Projects: mPower App



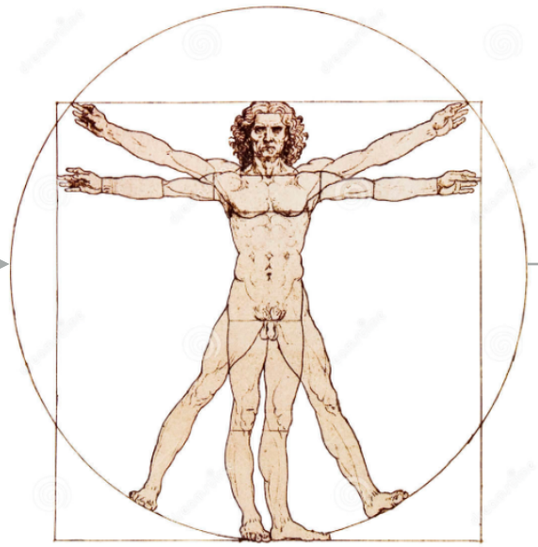
mPower

Sage BIONETWORKS

Clinical Team

LUXEMBOURG INSTITUTE OF HEALTH
RESEARCH DEDICATED TO LIFE

CHL Centre Hospitalier de Luxembourg



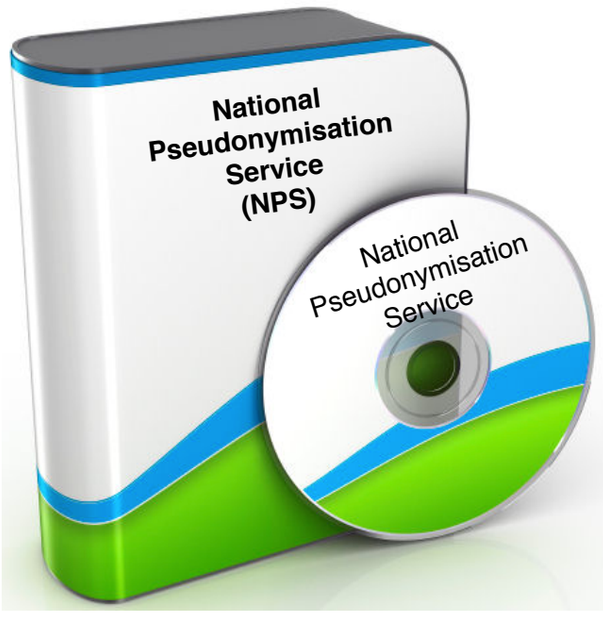
Subject

mPower

SAGE Private Cloud

Pseudonym

Request for Pseudonym



Anonymisation Service

Request for Pseudonym

Pseudonym

NCER-PD Repository

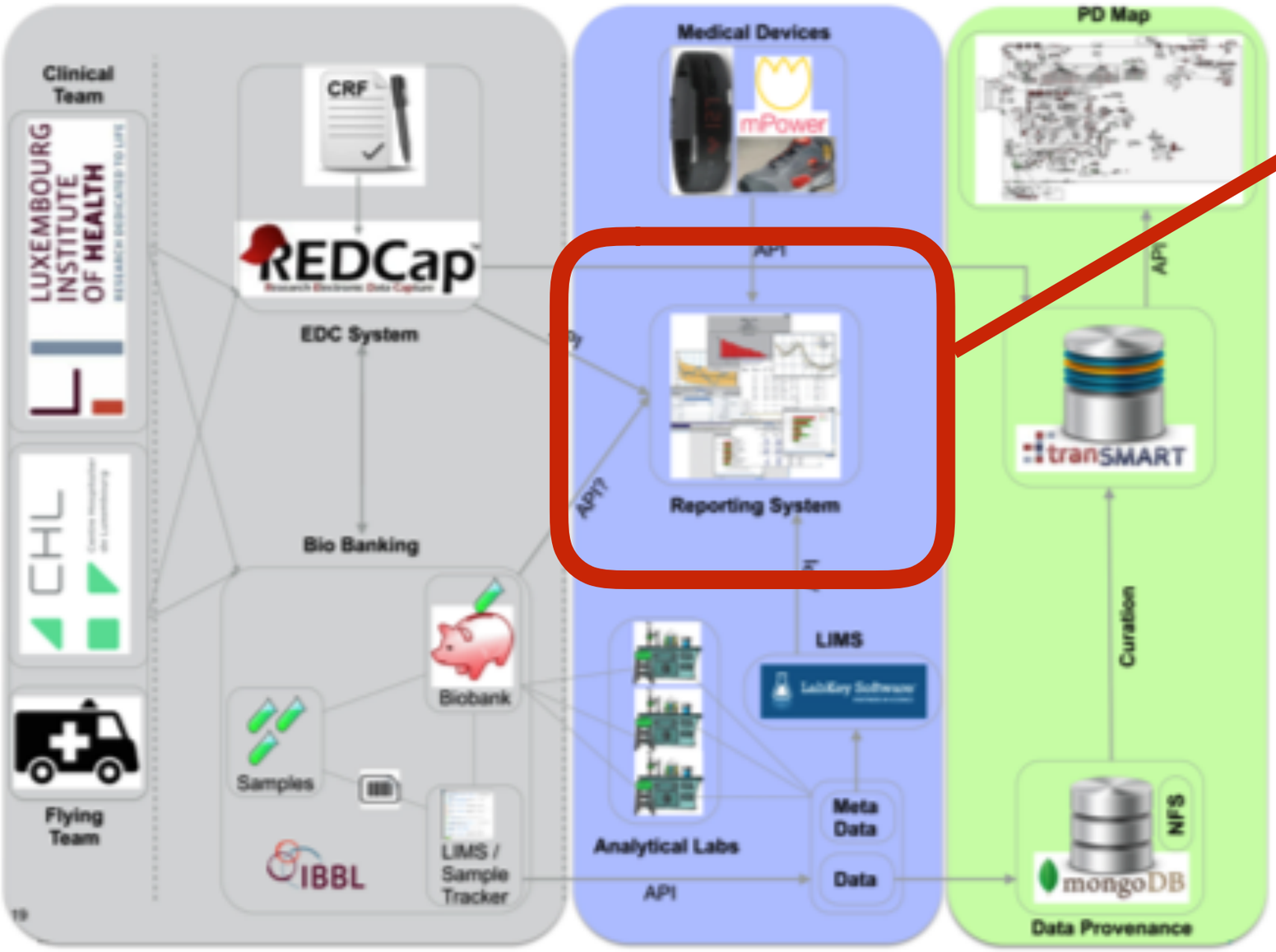
Flying Team

eGaIT –
embedded Gait analysis using IT

EMERGING FIELDS INITIATIVE

Bayerische Forschungsstiftung

NCER-PD Reporting System



- Central piece where all data meets
- Goals
  - Integrate
  - Store
  - Report
  - Analyze

Ada | Studies | Extras | Admin | Links | Contact

Views | Analytics | Dictionary | Category Tree | Setting

- Lux Park
  - Clinical
  - IBBL
    - Biosample
    - Biosample Test
  - mPower
- DeNoPa
- mPower Portal
- PPMI
- ML

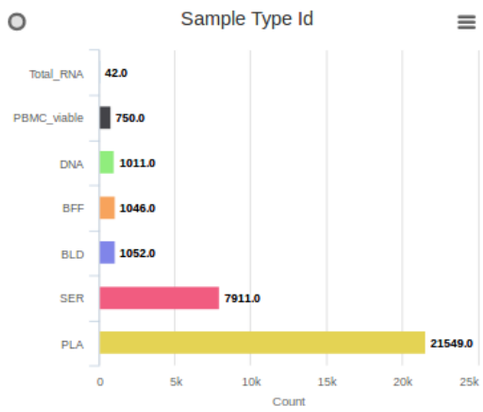
44,281 Biosample Items found

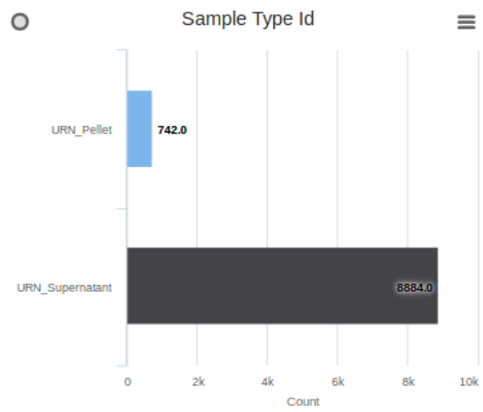
Export

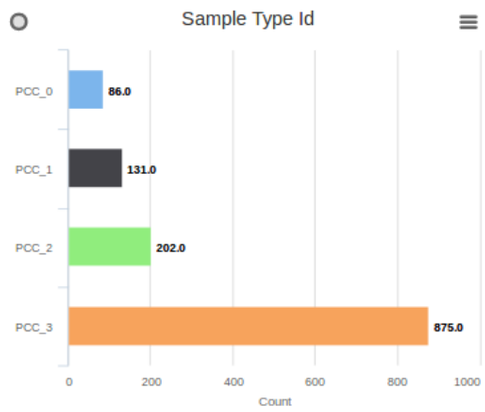
Count: 33361
Filter: Blood

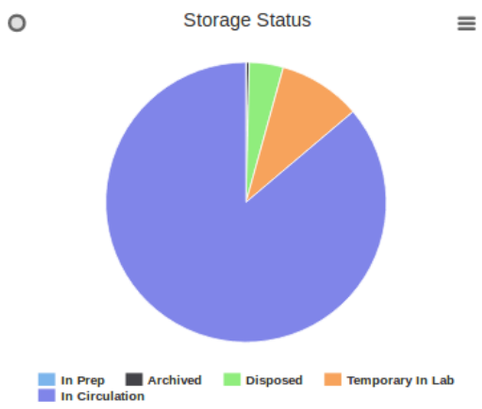
Count: 9626
Filter: Urine

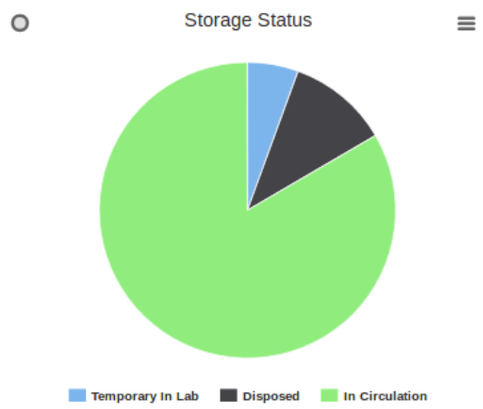
Count: 1294
Filter: Skin

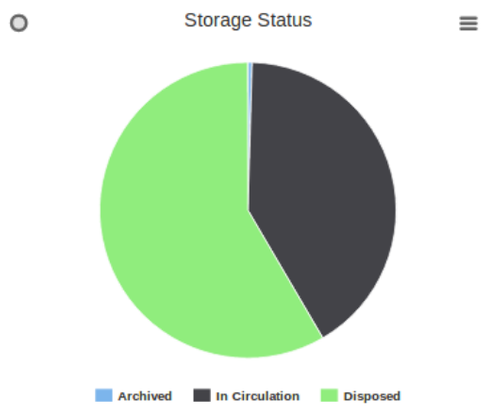


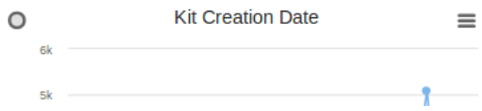


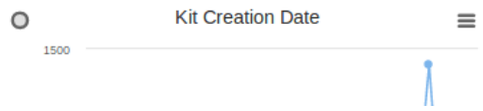


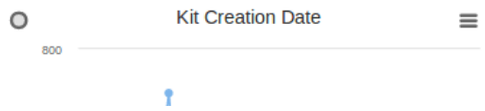












Ada: Dictionary update for 'lux_park.mpower_voice_activity' successfully finished.

Ada: Dictionary update for 'lux_park.mpower_memory_activity' successfully finished.

Ada: Import of data set 'Memory Activity' successfully finished.

Ada: Dictionary update for 'lux_park.mpower_my_thoughts' successfully finished.

Ada: Import of data set 'My Thoughts' successfully finished.

Ada: Dictionary update for 'lux_park.mpower_tapping_activity2' successfully finished.

Ada: Dictionary update for 'lux_park.mpower_walking_activity2' successfully finished.

Ada: Import of data set 'Walking Activity2' successfully finished.

Ada: Dictionary update for 'lux_park.mpower_memory_activity2' successfully finished.

Ada: Import of data set 'Memory Activity2' successfully finished.

Ada: Dictionary update for 'lux_park.mpower_voice_activity2' successfully finished.

Ada: Import of data set 'Voice Activity2' successfully finished.

Ada: Dictionary update for 'lux_park.mpower_tremor_activity2' successfully finished.

Ada: Import of data set 'Tremor Activity2' successfully finished.

Say what you want

Translational Medicine IT @ LCSB in Nutshell

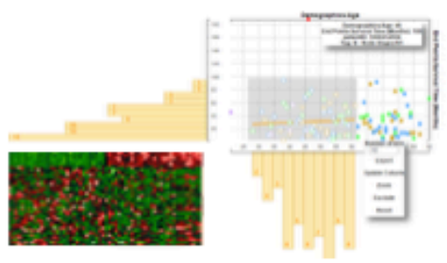
Data Capture

- Clinical Data → REDCap EDC System
- OMICS Data → Data Provenance
- Imaging Data → XNAT Imaging IT
- Meta Data → Data Catalog
- Mobile Data → mPower Synapse Mobile EDC

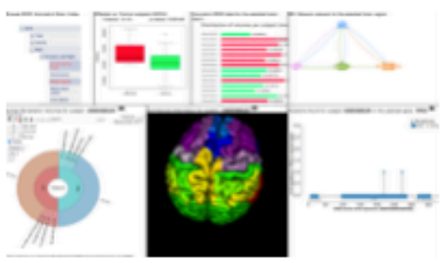
Data Integration

- Ada — Study Reporting and Exploration
- tranSMART — Cohort and Hypothesis Generation
- Variant-DB — Genetic Variants and Annotation

Visual-Analytics



Interactive Analysis: SmartR



Integrated Visualization: BrainWiz

Intuitive Pipeline: Galaxy

ETL

API

Successful Stories and Use cases

PARKINSON
National Centre of Excellence in Research

GEoPD
Genetic Epidemiology of Parkinson's Disease

SysMedPD

eTRIKS
European Translational Information and Knowledge Management Services

efpia

imi

elixir LUXEMBOURG

AETIONOMY

EpiPGX

BioCog

UNIVERSITÉ DU LUXEMBOURG

betaJuDO

INTERNATIONAL LEAGUE AGAINST EPILEPSY ILAE ESTABLISHED 1909

LCSB