





and Informatics Institute

USC Stevens Neuroimaging

USC Laboratory of Neuro Imaging

NIH National Institutes of Health

1. Data Ingest - ERMRest

Multiple repositories

Data Usage Agreements

Automated data clean up and

3. Data Analysis - Pipeline and

LONI Pipeline runs the analysis

on the amyloid PET and MRI

images and outputs amyloid

index values (SUVR: standard

region, for each subject using

• Genomics analysis using Globus

uptake value ratio) for each brain

computational resources at USC

Genomics on the Amazon cloud
Leverages elastic provisioner that optimizes performance and costs

Globus Genomics

Imaging datasets

٠

٠

ingest

Phenotypic, Genotypic and



USC Viterbi



rk•/88120/r83w

Abstract

Objective: Investigate if there are commonalities in patterns of *amyloid deposition* in individuals with Alzheimer's Disease or Parkinson's Disease that identify those individuals with or at risk for cognitive dysfunction

Approach: We created a *platform* for integration of multi-omic data, facilitate data discovery, cohort creation, enable rapid, scalable and reproducible analysis and finally publish results

Approach

We employed a systematic, reproducible approach that leveraged existing capabilities for understanding commonalities in amyloid deposition that included the following stages and challenges



The BDDS Platform

-		•														
nich na 154														¢	0 5	+
Santard	ad ly	ALVIN	e Fat													
130079	es 0.0	12102	100	2011	CHORNE											
30.401	ren i															
										2 m of						
				081											00	100
100854	**			1921	104-100 1038 W			12	10	10035			9219		-9924	
Indexed Drives by Cost	100	ng/ank	i bea	1021 410 %	A Data of organization	,		12 10 10 10 10 10	2.00 CTH 10 D	10(54			1929 w		1000	
Indicate Designed Designed	476 200 200 200	1004 00 1004 00	100	1921 410 %	4 20% 0527 1036 % 4 20% 01 organi 400 (4254)	100	20	14) 11 alta aprov 12 2941	CTN 20 22 9	10.595		•	00.100		Score Score No.	
Incodes Deversity Color Second Genomic	0790 2011 2012 2014	10/4010 1100 10 4.25 121	19	19211	4 2016-01-01 1638-16 2016-01-20[00 400] [442[41]	1	10	19.] 10.05-02704 10. 2041	029 So 30 9	10294		•	01100		Scale Scale N	
0000000 Diversity Color 6420:0 Genomic AC A(1)	0194 2000 2000 2000 2000 2000	425 (2) 102 (2)	100	1021 410 %	4 2014 0.07 4 10.00 W 10.00	100	10 04	193 10 shi aprov 10 2040 14	2.00 (75) (75)	NOR In too Int of Int of		1	1929-W		9994 Scart State N	
0100000 010000 040000 040000 040000 040000 040000 040000 040000	4794 2011 2010 2010 2010 2010 2010 2010 201	425 121 100/ 101 100/ 101	100		4 2014 10-07 16.36 W 4 2016 of project eff() 24.2043 7 0 7 0 7 0 7 0 7 0 7 0 7 0 7 0 7 0 7 0	10	210 04 10 40%	193 10 abs approve 10 2043 11 12 14 14 14 14 14 14 14 14 14 14 14 14 14	210 So So So NOSE	1036 m 100 m 101 m 10 m 10	1 500 Mill 500 Mill 6 7 10 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8	1	001308 001308 00010	\$ 385	10004 1000 11	
000000 00000 00000 00000 00000 00000 0000	0196 2010 2010 2010 2010 2010 2010 2010 201		100		4 2014 00.07 1016 10 2010 1010 00 2010 1010 1010 00 2010 1010 000 1010 000 1000 1000 1000 1000 1000 1000 1000 1000 1000 1000 1000 1000 1000 1000	100	20 00 64 10 476	143 10 10 10 10 10 10 10 10 10 10 10 10 10 1	2100 20 20 20 20 20 20 20 20 20 20 20 20 20 2	19294 mi Ton mi Ton mi Fil mi	1 000000 9 100 100 100 100 100 100 100 1	1	19294 04 108 04 10 019 10	1 200	10004 10000 10000 10000 10000 10000	
000000 00000 04000000	108 1794 2001 2002 2		100 (1) (1) (1) (1) (1)		4 2014 601 1021 102 102 1021 102	100 100 100 100	20 00 64 200 406 406 406	2011 10 101 20100 2011 2011 2011 2011 20	210 50 50 9 705 70566 70066	19894 miton	 1 1000 fram 1000	1	001300 001300 00110 01100 0100 0100 0100 0100 0100 0100 0100 0100 0100 0100 0100 0100 0100 0100 0100 0100 0100 0100 0100 0000	1.305	10004 5000 5000 5000 5000 5000 5000 5000	
000000 00000 0000 00000 00000 00000 0000	0194 0194 000 000 000 000 000 000 000 000 000 0		1 Des (1) (2) (2) (2) (4) (2) (4) (4)		4 2019 501 10.50 10 10.50 10 10.5		20 00 10 10 10 10 10 10 10 10 10 10 10 10	11. 10.10704 10.0000 10.00000 10.00000 10.00000 10.000000 10.00000000	200 20 200 20 2000 2000 2000 2000 2000	0.00 100 100 100 100 100 100 100 100 100	1 505 ftm 500000 9 100 100 100 100 100 100 100 100	1	001100 001100 00110 01000 0100 0100 01	130	10004 10004 10006 100000000	

2. Data Query and Exchange - BagIt

- ERMRest provides high-level API for auery
- Adopted the BagIt specification to create databags
- Enhanced BagIt to support big biomedical data





4. Data Publication and Integration – Globus Publication and ERMRest

- Results from analysis are integrated with ERMRest
- Results are published in Globus Publication service with appropriate metadata

🕘 globus		
	Please log In to submit a data	uset and view dataset
If you use this dataset please add this	citation to your publication:	
Glusman, Gustavo; Rodriguez, A http://hdl.handle.net/11466/T2N	lex, "PPMI Exome Variants, " 2015, ISSBB	Download
Title:	PPMI Exome Variants	
Authors:	Glusman, Gustavo Rodriguez, Alex	
Issue Date:	7-Oct-2015	
Publisher	Clobus	
UNI	http://bit.bandle.set/11466/T2N5588	
Appears in Collections:	8DDS Collection	
Tites in This litem:		
globuspublish#bdds-publication/bd	Idspub./published/publication534/	
show full nem record		

Analyzing role of Amyloid Burden



Discussion & Conclusions

- We created a powerful platform by integrating several existing services and capabilities
- Platform development resulted in adopting and extending BagIt format for biomedical big data
- Future work includes applying the platform for analyzing Alzheimer's data
- Reference implementation of the NIH Commons

Bibliography

- 1. The BagIt File Packaging Format (V0.97) available at :
 - http://tools.ietf.org/html/draft-kunzebagit-11
- 2. Minids: http://minid.bd2k.org