purpose of public sharing – most visibly in the Human Connectome Project (HCP) [3] - but also in the Pediatric Imaging, Neurocognition and Genetics (PING) study [4] There are a number of funded efforts to collect old data and re-release as public databases, notably the INDI [5] efforts (which include the popular ABIDE [6] and functional connectomes 1000 datasets [7]). The BRAIN initiative [8] aims to collect data that will be a challenge to store, let alone analyze. There are even online journals focused on publishing datasets (e.g. Nature Scientific Data), or with options to release data (e.g. F1000 "Data Notes").

NiData is a Python package that provides a single interface accessing data from a variety of open data sources. The software framework makes it easy to add new data sources, simple to define and to provide access to multiple datasets from a single data source. Software dependencies are managed on a per-dataset basis, allowing downloads and examples to use any public packages without requiring installation of packages required by unused datasets. The interface also allows selective download of data (by subject or type) and caches files locally, allowing easy management of big datasets.

Results

We focused on exposing new methods for downloading data from the HCP, supporting access via Amazon S3 and HTTP/XNAT [9]. We were able to provide a downloader that accepts login credentials and downloads files locally. We created an example that interacts with DIPY [10] to produce diffusion imaging results on a single subject from the HCP. We also worked at collecting common data sources, as well as individual datasets stored at each data source, into NiData's "data sources" wiki page. We incorporated downloads, documentation, and examples from the nilearn package and began discussion of making a more extensible object model.

Since the hackathon, we have created such an object model and migrated all code to use it, and a Sphinx- based website is under development. The current object model makes it easier to write general-purpose fetchers (e.g. HTTP, XNAT, Amazon S3) that can be extended to access specific databases (e.g. COINS [11], LORIS [12], ADNI [13]).

Conclusions

Projects like NiData improve curated data access and increase the effectivity of big data projects with open source data.

Availability of supporting data

More information about this project can be found at: http://github.com/nidata/nidata

Competing interests

None.

Author's contributions

BC and AR wrote the software and the report.

Acknowledgements

Report from 2015 OHBM Hackathon (HI). The authors would like to thank the organizers and attendees of the 2015 OHBM Hackathon.

References

1. Poldrack RA, Barch DM, Mitchell JP, Wager TD, Wagner AD, Devlin JT, Cumba C, Koyejo O, Milham MP. Toward open sharing of task-based fMRI data: the OpenfMRI project. Front Neuroinform. 2013; 7.
2. Buccigrossi R, Ellisman M, Grethe J, Haselgrove C, Kennedy DN, Martone M, Preuss N, Reynolds K, Sullivan M, Turner J, Wagner K. The Neuroimaging Informatics Tools and Resources Clearinghouse (NITRC). AMIA Annu Symp Proc. 2008.
3. Van Essen DC, Smith SM, Barch DM, Behrens TE, Yacoub E, Ugurbil K, Van Essen D, Barch D, Corbetta M, Goate A, Heath A, Larson-Prior L, Marcus D, Petersen S, Prior F, Province M, Raichle M, Schlaggar B, Shimony J, Snyder A, Adeyemo B, Archie K, Babajani-Feremi A, Bloom N, Bryant JE, Burgess G, Cler E, Coalson T, Curtiss S, Danker S, Denness R, Dierker D, Elam J, Evans T, Feldt C, Fenlon K, Footer O, Glasser M, Gordon E, Gu P, Guilday C, Harms M, Hartley T, Harwell J, Hileman M, Hodge M, Hood L, Horton W, House M, Laumann T, Lugo M, Marion S, Miezin F, Nolan D, Nolan T, Power J, Ramaratnam M, Reid E, Schindler J, Schmitz D, Schweiss C, Serati J, Taylor B, Tobias M, Wilson T, Ugurbil K, Garwood M, Harel N, Lenglet C, Yacoub E, Adriany G, Auerbach E, Moeller S, Strupp J, Smith S, Behrens T, Jenkinson M, Johansen-Berg H, Miller K, Woolrich M, Andersson J, Duff E, Hern,ez M, Jbabdi S, Robinson E, Salimi-Khorshidi R, Sotiropoulos S, Romani GL, Della Penna S, Pizzella V, de Pasquale F, Di Pompeo F, Marzetti L, Perruci G, Bucholz R, Roskos T, Kiser T, Luo QJ, Stout J, Oostenveld R, Beckmann C, Schoffelen JM, Fries P, Michalareas G, Sapiro G, Sporns O, Nichols T, Farber G, Bjork J, Blumensath T, Chang A, Chen L, Feinberg D, Kull L, Wig G, Xu JG, Basser P, Bullmore E, Evans A, Gazzaniga M, Glahn D, Hawrylycz M, Hennig J, Parker G, Poldrack R, Salmelin R. The WU-Minn Human Connectome Project: an overview. Neuroimage. 2013; 80: 62–79.
4. Jernigan TL, Brown TT, Hagler DJ, Akshoomoff N, Bartsch H, Newman E, Thompson WK, Bloss CS, Murray SS, Schork N, Kennedy DN, Kuperman JM, McCabe C, Chung Y, Libiger O, Maddox M, Casey BJ, Chang L, Ernst TM, Frazier JA, Gruen JR, Sowell ER, Kenet T, Kaufmann WE, Mostofsky S, Amaral DG, Dale AM. The Pediatric Imaging, Neurocognition, and Genetics (PING) Data Repository. Neuroimage. 2016; 124: 1149–1154.
5. Mennes M, Biswal BB, Castellanos FX, Milham MP. Making data sharing work: the FCP/INDI experience. Neuroimage. 2013; 82: 683–691.
6. Di Martino A, Yan CG, Li Q, Denio E, Castellanos FX, Alaerts K, Anderson JS, Assaf M, Bookheimer SY, Dapretto M, Deen B, Delmonte S, Dinstein I, Ertl-Wagner B, Fair DA, Gallagher L, Kennedy DP, Keown CL, Keysers C, Lainhart JE, Lord C, Luna B, Menon V, Minshew NJ, Monk CS, Mueller S, Muller RA, Nebel MB, Nigg JT, O'Hearn K, Pelphrey KA, Peltier SJ, Rudie JD, Sunaert S, Thioux M, Tyszka JM, Uddin LQ, Verhoeven JS, Wenderoth N, Wiggins JL, Mostofsky SH, Milham MP. The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. Mol Psychiatry. 2014; 19: 659–667.
7. Biswal BB, Mennes M, Zuo XN, Gohel S, Kelly C, Smith SM, Beckmann CF, Adelstein JS, Buckner RL, Colcombe S, Dogonowski AM, Ernst M, Fair D, Hampson M, Hoptman MJ, Hyde JS, Kiviniemi VJ, Kotter R, Li SJ, Lin CP, Lowe MJ, Mackay C, Madden DJ, Madsen KH, Margulies DS, Mayberg HS, McMahon K, Monk CS, Mostofsky SH, Nagel BJ, Pekar JJ, Peltier SJ, Petersen SE, Riedl V, Rombouts SA, Rypma B, Schlaggar BL, Schmidt S, Seidler RD, Siegle GJ, Sorg C, Teng GJ, Veijola J, Villringer A, Walter M, Wang L, Weng XC, Whitfield-Gabrieli S, Williamson P, Windischberger C, Zang YF, Zhang HY, Castellanos FX, Milham MP. Toward discovery science of human brain function. Proc Natl Acad Sci USA. 2010; 107: 4734–4739.
8. Insel TR, L,is SC, Collins FS. Research priorities. The NIH BRAIN Initiative. Science. 2013; 340: 687–688.
9. Marcus DS, Olsen TR, Ramaratnam M, Buckner RL. The Extensible Neuroimaging Archive Toolkit: an informatics platform for managing, exploring, and sharing neuroimaging data. Neuroinformatics. 2007; 5.
10. Garyfallidis E, Brett M, Amirbekian B, Rokem A, van der Walt S, Descoteaux M, Nimmo-Smith I. Dipy, a library for the analysis of diffusion MRI data. Front Neuroinform. 2014; 8.
11. Scott A, Courtney W, Wood D, de la Garza R, Lane S, King M, Wang R, Roberts J, Turner JA, Calhoun VD. COINS: An Innovative Informatics and Neuroimaging Tool Suite Built for Large Heterogeneous Datasets. Front Neuroinform. 2011; 5.
12. Das S, Zijdenbos AP, Harlap J, Vins D, Evans AC. LORIS: a web-based data management system for multi-center studies. Front Neuroinform. 2011; 5.
13. Jack CR, Bernstein MA, Fox NC, Thompson P, Alex,er G, Harvey D, Borowski B, Britson PJ, L Whitwell J, Ward C, Dale AM, Felmlee JP, Gunter JL, Hill DL, Killiany R, Schuff N, Fox-Bosetti S, Lin C, Studholme C, DeCarli CS, Krueger G, Ward HA, Metzger GJ, Scott KT, Mallozzi R, Blezek D, Levy J, Debbins JP, Fleisher AS, Albert M, Green R, Bartzokis G, Glover G, Mugler J, Weiner MW. The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI methods. J Magn Reson Imaging. 2008; 27: 685–691.

A3

Integrating the Brain Imaging Data Structure (BIDS) standard into C-PAC

Daniel Clark[1], Krzysztof J. Gorgolewski[2], R. Cameron Craddock[1,3]
[1]Center for the Developing Brain, Child Mind Institute, New York, NY, USA; [2]Poldrack Lab, Department of Psychology, Stanford University, Stanford, CA, USA; [3]Computational Neuroimaging Lab, Center for Biomedical Imaging and Neuromodulation, Nathan Kline Institute for Psychiatric Research, Orangeburg, NY, USA
Correspondence: Daniel Clark (daniel.clark@childmind.org) – Center for the Developing Brain, Child Mind Institute, New York, NY, USA
GigaScience 2016, **5(Suppl 1):**A3

Introduction
Data acquired during neuroimaging experiments can be organized in many ways. This stems from differences in scanner software, various DICOM and NIFTI tools, and custom data organizing scripts within different laboratories. The Brain Imaging Data Structure (BIDS) specification [1] provides a simple, straightforward solution to this problem by introducing an intuitive standard for neuroimaging data organization. The widespread adoption of BIDS can be facilitated through incorporating this standard into software projects used for neuroimaging analysis. These software packages will in turn benefit from the homogenous data structure and ease of specifying data acquisition parameters afforded by BIDS. The goal of this Brainhack project was to integrate BIDS into the Configurable Pipeline for the Analysis of Connectomes (C-PAC) [2] a Python Package? built on Nipype [3] for the high-throughput analysis of resting state fMRI data.

Approach
Processing data with C-PAC begins with specifying the paths of the anatomical and functional files to be processed, along with their corresponding acquisition parameters. This is accomplished in a semi-automatic procedure in which the user specifies templates that describe the file organization and then a script walks this structure to find the data. The resulting subject list can then be partnered with a pipeline configuration and submitted to C-PAC for processing. We extended this functionality to natively understand BIDS, so that data that conforms to this standard can be configured to run through C-PAC with minimal user intervention.

C-PAC with BIDS

A BIDS flag was added to the subject list builder along with a text box for the user to specify the base directory of the data file structure. The BIDS file hierarchy is then traversed to build anatomical and functional file pattern templates. These templates are returned to the main subject list builder function, which runs the same way as if using user specified file path templates. This approach minimized modifications to the data-gathering algorithm while providing for a robust way to ensure all data is found and returned properly. Additional scanning parameters that are required to complete the processing (repetition time, slice timing information, etc.) are read from BIDS specified JSON files that are stored alongside the imaging data.

The new implementation takes advantage of one of many BIDS utilities openly available: the BIDS meta-data tool [4] [https://github.com/INCF/bidsutils] This tool provides the subject, session, and run-level indicators to the builder without needing the user to manually enter any keywords; it takes advantage of the fixed organization scheme and the presence of JSON files to deliver all of this information reliably and efficiently. The tool is written in Python, which provided for easy integration into the C-PAC source code. It works for BIDS datasets stored locally as well as those available through remotely through Amazon S3.

Results
The updated C-PAC GUI reflects the "BIDS" and "Custom" options - as seen in Fig. 2 - with the "Custom" option allowing users to specify their data structure as in previous versions of C-PAC. In the future this option would be more elegantly displayed via a radio button with the input fields dynamically changing to reflect the type of input desired.

The code changes were fairly straightforward and were cleanly inserted into the current builder module [https://github.com/FCP-INDI/C-PAC/blob/test_dev/CPAC/utils/build_sublist.py] The implementation developed during Brainhack is feature full, but will require more testing in the future.

Conclusions
Incorporating the BIDS subject list builder into C-PAC is a great step forward in bringing the standard to a broader audience. Throughout the integration process, other technologies were discovered that could further enable input data gathering across a wide range of file system types, including FTP, SFTP, Zip, S3, and an array of virtual file-systems. With further development, the overhead of preprocessing one's own neuroimaging data for discovery science can be minimized so scientists can focus on the results.

Availability of supporting data
More information about this project can be found at: https://bids.neuroimaging.io.

Competing interests
None.

Author's contributions
RCC and KJG provided supervision and reference, DJC and KJG wrote the software, DJC and KJG performed tests, and DJC wrote the report.

Acknowledgements
Report from 2015 Brainhack Americas (MX). The authors would like to thank the organizers and attendees of Brainhack MX and the developers of C-PAC. This project was funded in part by an Educational Research Grant from Amazon Web Services.

References

1. Gorgolewski KJ, Poline JB, Keator DB, Nichols BN, Auer T, Craddock RC, Fl,in G, Ghosh SS, Sochat VV, Rokem A, Halchenko YO, Hanke M, Haselgrove C, Helmer K, Maumet C, Nichols TE, Turner JA, Das S, Kennedy DN, Poldrack RA. Brain Imaging Data Structure - a new standard for describing and organizing human neuroimaging data. Frontiers in Neuroscience.
2. Craddock RC, Sikka S, Cheung B, Khanuja R, Ghosh SS, Yan CG, Li Q, Lurie D, Vogelstein J, Burns R, Colcombe S, Mennes M, Kelly C, Di Martino A, Castellanos FX, Milham MP. Towards Automated Analysis of Connectomes: The Configurable Pipeline for the Analysis of Connectomes (C-PAC). Frontiers in Neuroinformatics. 2013.
3. Gorgolewski K, Burns CD, Madison C, Clark D, Halchenko Y), Waskom ML, Ghosh SS. Nipype: a flexible, lightweight and extensible neuroimaging data processing framework in python. Front Neuroinform. 2011; 5.
4. Gorgolewski KJ. bidsutils. GitHub; 2015. https://github.com/INCF/bidsutils.

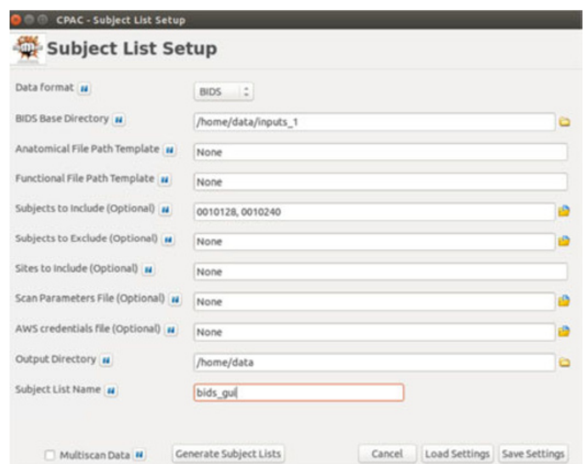

Fig. 2 (abstract A3). C-PAC subject list builder using BIDS directory

A4
Optimized implementations of voxel-wise degree centrality and local functional connectivity density mapping in AFNI

R. Cameron Craddock[1,2], Daniel J. Clark[2]

[1]Computational Neuroimaging Lab, Center for Biomedical Imaging and Neuromodulation, Nathan Kline Institute for Psychiatric Research, Orangeburg, NY, USA; [2]Center for the Developing Brain, Child Mind Institute, New York, NY, USA

Correspondence: R. Cameron Craddock (ccraddock@nki.rfmh.org) – Computational Neuroimaging Lab, Center for Biomedical Imaging and Neuromodulation, Nathan Kline Institute for Psychiatric Research, Orangeburg, NY, USA

GigaScience 2016, **5(Suppl 1):**A4

Introduction
Degree centrality (DC) [1] and local functional connectivity density (lFCD) [2] are statistics calculated from brain connectivity graphs that measure how important a brain region is to the graph. DC (a.k.a.