

## The Past, Present, and Future of the Brain Imaging Data Structure (BIDS)

Poldrack, Russell A.<sup>1</sup>, Markiewicz, Christopher J.<sup>1</sup>, Appelhoff, Stefan<sup>2</sup>, Ashar, Yoni K.<sup>3</sup>, Auer, Tibor<sup>4</sup>, Baillet, Sylvain<sup>5</sup>, Bansal, Shashank<sup>6</sup>, Beltrachini, Leandro<sup>7</sup>, Bénar, Christian G.<sup>8</sup>, Bertazzoli, Giacomo<sup>9,10,11,12,13</sup>, Bhogawar, Suyash<sup>14</sup>, Blair, Ross W.<sup>1</sup>, Bortoletto, Marta<sup>9</sup>, Boudreau, Mathieu<sup>15</sup>, Brooks, Teon L.<sup>1</sup>, Calhoun, Vince D.<sup>16</sup>, Castelli, Filippo Maria<sup>17,18</sup>, Clement, Patricia<sup>19,20</sup>, Cohen, Alexander L.<sup>21</sup>, Cohen-Adad, Julien<sup>15</sup>, D'Ambrosio, Sasha<sup>22,23</sup>, de Hollander, Gilles<sup>24</sup>, de la Iglesia-Vayá, María<sup>25</sup>, de la Vega, Alejandro<sup>26</sup>, Delorme, Arnaud<sup>27</sup>, Devinsky, Orrin<sup>28</sup>, Draschkow, Dejan<sup>29</sup>, Duff, Eugene Paul<sup>30</sup>, DuPre, Elizabeth<sup>1</sup>, Earl, Eric<sup>31</sup>, Esteban, Oscar<sup>32</sup>, Feingold, Franklin W.<sup>1</sup>, Flandin, Guillaume<sup>33</sup>, Galassi, Anthony<sup>31</sup>, Gallitto, Giuseppe<sup>34,35</sup>, Ganz, Melanie<sup>36,37</sup>, Gau, Rémi<sup>38</sup>, Gholam, James<sup>39</sup>, Ghosh, Satrajit S.<sup>40</sup>, Giacomel, Alessio<sup>41</sup>, Gillman, Ashley G.<sup>42</sup>, Gleeson, Pdraig<sup>43</sup>, Gramfort, Alexandre<sup>44</sup>, Guay, Samuel<sup>45</sup>, Guidali, Giacomo<sup>46</sup>, Halchenko, Yaroslav O.<sup>47</sup>, Handwerker, Daniel A.<sup>31</sup>, Hardcastle, Nell<sup>1</sup>, Herholz, Peer<sup>48</sup>, Hermes, Dora<sup>49</sup>, Honey, Christopher J.<sup>50</sup>, Innis, Robert B.<sup>31</sup>, Ioanas, Horea-Ioan<sup>47</sup>, Jahn, Andrew<sup>51</sup>, Karakuzu, Agah<sup>15</sup>, Keator, David B.<sup>52,53,54</sup>, Kiar, Gregory<sup>55</sup>, Kincses, Balint<sup>34,35</sup>, Laird, Angela R.<sup>56</sup>, Lau, Jonathan C.<sup>57</sup>, Lazari, Alberto<sup>58</sup>, Legarreta, Jon Haitz<sup>59</sup>, Li, Adam<sup>60</sup>, Li, Xiangrui<sup>61</sup>, Love, Bradley C.<sup>62</sup>, Lu, Hanzhang<sup>63</sup>, Maumet, Camille<sup>64</sup>, Mazzamuto, Giacomo<sup>65</sup>, Meisler, Steven L.<sup>66</sup>, Mikkelsen, Mark<sup>67</sup>, Mutsaerts, Henk<sup>68,69</sup>, Nichols, Thomas E.<sup>70</sup>, Nikolaidis, Aki<sup>71</sup>, Nilsonne, Gustav<sup>72,73</sup>, Niso, Guiomar<sup>74</sup>, Norgaard, Martin<sup>31,36</sup>, Okell, Thomas W.<sup>58</sup>, Oostenveld, Robert<sup>75,76</sup>, Ort, Eduard<sup>77</sup>, Park, Patrick J.<sup>78</sup>, Pawlik, Mateusz<sup>79</sup>, Pernet, Cyril R.<sup>37</sup>, Pestilli, Franco<sup>26</sup>, Petr, Jan<sup>80</sup>, Phillips, Christophe<sup>81</sup>, Poline, Jean-Baptiste<sup>82</sup>, Pollonini, Luca<sup>83,84</sup>, Raamana, Pradeep Reddy<sup>85</sup>, Ritter, Petra<sup>86,87,88,89,90</sup>, Rizzo, Gaia<sup>91,92</sup>, Robbins, Kay A.<sup>93</sup>, Rockhill, Alexander P.<sup>94</sup>, Rogers, Christine<sup>95</sup>, Rokem, Ariel<sup>96</sup>, Rorden, Chris<sup>97</sup>, Routier, Alexandre<sup>98</sup>, Saborit-Torres, Jose Manuel<sup>25</sup>, Salo, Taylor<sup>99</sup>, Schirner, Michael<sup>88,100,101,102,103</sup>, Smith, Robert E.<sup>104,105</sup>, Spisak, Tamas<sup>34,106</sup>, Sprenger, Julia<sup>107</sup>, Swann, Nicole C.<sup>108</sup>, Szinte, Martin<sup>107</sup>, Takerkart, Sylvain<sup>107</sup>, Thirion, Bertrand<sup>44</sup>, Thomas, Adam G.<sup>31</sup>, Torabian, Sajjad<sup>109</sup>, Varoquaux, Gael<sup>110</sup>, Voytek, Bradley<sup>111</sup>, Welzel, Julius<sup>112</sup>, Wilson, Martin<sup>113</sup>, Yarkoni, Tal<sup>114</sup>, Gorgolewski, Krzysztof J.<sup>1</sup>

1: Department of Psychology, Stanford University, Stanford, CA, USA

2: Max Planck Institute for Human Development, Berlin, Germany

3: University of Colorado Anschutz Medical Campus, Aurora, CO, USA

4: School of Psychology, University of Surrey, Guildford, UK

5: McConnell Brain Imaging Centre, Montréal Neurological Institute, McGill University, Montréal, Canada

6: Department of Bioengineering, University of California, San Diego, La Jolla, CA, USA

7: Cardiff University Brain Research Imaging Centre (CUBRIC), School of Physics and Astronomy, Cardiff University, Wales, UK

8: Institut de Neurosciences des Systèmes, Marseille, France

9: Neurophysiology Lab, IRCCS Istituto Centro San Giovanni di Dio Fatebenefratelli, Brescia, Italy

10: Center for Mind/Brain Sciences - CIMeC, University of Trento, Rovereto, TN, Italy

11: Brigham and Women's Hospital, Boston, MA, USA

12: Massachusetts General Hospital, Boston, MA, USA

13: Harvard Medical School, Boston, MA, USA

- 14: Rackspace Technology, San Antonio, TX, USA
- 15: NeuroPoly Lab, Polytechnique Montréal, Montréal, Quebec, Canada
- 16: Tri-institutional Center for Translational Research in Neuroimaging and Data Science (TReNDS), Georgia State, Georgia Tech, Emory, Atlanta, GA, USA
- 17: European Laboratory for Non-Linear Spectroscopy (LENS), University of Florence, Sesto Fiorentino, Italy
- 18: Bioretics srl, Cesena, Italy
- 19: Department of Medical Imaging, Ghent University Hospital, Ghent, Belgium
- 20: Department of Diagnostic Sciences, Ghent University, Ghent, Belgium
- 21: Department of Neurology, Boston Children's Hospital, Boston, MA, USA
- 22: Dipartimento di Scienze della Salute dell'Università degli Studi di Milano, Milan, Italy
- 23: Department of Clinical and Experimental Epilepsy, University College London, UK
- 24: Zurich Center for Neuroeconomics, Department of Economics, University of Zurich, Zurich, Switzerland
- 25: UMIB-FISABIO, Valencia, Spain
- 26: The University of Texas at Austin, Austin, TX, USA
- 27: SCCN, University of California, San Diego, La Jolla CA USA
- 28: Department of Neurology, NYU Langone Medical Center, New York, NY, USA
- 29: Department of Experimental Psychology, University of Oxford, Oxford, UK
- 30: UK Dementia Research Institute, Department of Brain Sciences, Imperial College London, London, UK
- 31: Intramural Research Program, National Institute of Mental Health, Bethesda, MD, USA
- 32: Department of Radiology, Lausanne University Hospital and University of Lausanne, Lausanne, Switzerland
- 33: Wellcome Centre for Human Neuroimaging, University College London, London, England, UK
- 34: Center for Translational Neuro- and Behavioral Sciences, University Medicine Essen, Essen, Germany
- 35: Department of Neurology, University Medicine Essen, Essen, Germany
- 36: Department of Computer Science, University of Copenhagen, Copenhagen, Denmark
- 37: Neurobiology Research Unit, Copenhagen University Hospital, Copenhagen, Denmark
- 38: Origamin Lab, The Neuro, McGill University, Montreal, Quebec, Canada
- 39: Cardiff University Brain Research Imaging Centre (CUBRIC), School of Psychology, Cardiff University, Wales, UK
- 40: Massachusetts Institute of Technology, Cambridge, MA, USA
- 41: Department of Neuroimaging, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, England, UK
- 42: The Australian e-Health Research Centre, Commonwealth Scientific and Industrial Research Organisation, Townsville, Queensland, Australia
- 43: Department of Neuroscience, Physiology and Pharmacology, University College London, London, England, UK
- 44: Inria, CEA, Université Paris-Saclay, Palaiseau, France
- 45: Université de Montréal, Montréal, QC, Canada
- 46: Department of Psychology & NeuroMI - Milan Centre for Neuroscience, University of

Milano-Bicocca, Milan, Italy

47: Center for Open Neuroscience, Department of Psychological and Brain Sciences, Dartmouth College, NH, USA

48: McConnell Brain Imaging Centre, Montréal Neurological Institute, McGill University, Montréal, Quebec, Canada

49: Department of Physiology and Biomedical Engineering, Mayo Clinic, Rochester, MN, USA

50: Department of Psychological & Brain Sciences, Johns Hopkins University, Baltimore, MD, USA

51: Functional MRI Laboratory, University of Michigan, Ann Arbor, MI, USA

52: Change Your Brain Change Your Life Foundation, Costa Mesa, CA, USA

53: Amen Clinics, Costa Mesa, CA, USA

54: Department of Psychiatry and Human Behavior, School of Medicine, University of California, Irvine, CA, USA

55: Center for Data Analytics, Innovation, and Rigor, Child Mind Institute, New York, NY USA

56: Department of Physics, Florida International University, Miami, FL, USA

57: Department of Clinical Neurological Sciences, Western University, London, Ontario, Canada

58: Wellcome Centre for Integrative Neuroimaging, FMRIB, Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, UK

59: Department of Radiology, Brigham and Women's Hospital, Mass General Brigham/Harvard Medical School, Boston, MA, USA

60: Columbia University, New York, NY, USA

61: Center for Cognitive and Behavioral Brain Imaging, The Ohio State University, Columbus, OH, USA

62: University College London, London, UK

63: Johns Hopkins University School of Medicine, Baltimore, MD, USA

64: Inria, Univ Rennes, CNRS, Inserm, IRISA UMR 6074, Empenn ERL U 1228, Rennes, France

65: National Research Council - National Institute of Optics (CNR-INO), Florence, Italy

66: Program in Speech and Hearing Bioscience and Technology, Harvard University, Cambridge, MA, USA

67: Department of Radiology, Weill Cornell Medicine, New York, NY, USA

68: Radiology and Nuclear Medicine, Vrije Universiteit Amsterdam, Amsterdam UMC location VUmc, Amsterdam, The Netherlands

69: Amsterdam Neuroscience, Brain Imaging, Amsterdam, The Netherlands

70: Big Data Institute, Li Ka Shing Centre for Health Information and Discovery, Nuffield Department of Population Health, University of Oxford, Oxford, UK

71: Center for the Developing Brain, Child Mind Institute, New York, NY, USA

72: Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden

73: Swedish National Data Service, Gothenburg University, Gothenburg, Sweden

74: Instituto Cajal, CSIC, Madrid, Spain

75: Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, Nijmegen, The Netherlands

76: NatMEG, Karolinska Institutet, Stockholm, Sweden

77: Heinrich Heine University, Department of Biological Psychology of Decision Making,

Düsseldorf, Germany

78: Western University, London, Ontario, Canada

79: Paris-Lodron-University of Salzburg, Department of Psychology, Centre for Cognitive Neuroscience, Salzburg, Austria

80: Helmholtz-Zentrum Dresden-Rossendorf, Institute of Radiopharmaceutical Cancer Research, Dresden, Germany

81: GIGA CRC in vivo imaging, Liege University, Liege, Belgium

82: Neuro Data Science ORIGAMI Laboratory, McConnell Brain Imaging Centre, Faculty of Medicine, McGill University, Montréal, Canada

83: Department of Engineering Technology, University of Houston, Houston, TX

84: Basque Center on Cognition, Brain and Language, Donostia-San Sebastián, Spain

85: University of Pittsburgh, Pittsburgh, PA, USA

86: Berlin Institute of Health at Charité, Universitätsmedizin Berlin, Charitéplatz 1, Berlin 10117, Germany

87: Department of Neurology with Experimental Neurology, Charité, Universitätsmedizin Berlin, Corporate member of Freie Universität Berlin and Humboldt Universität zu Berlin, Charitéplatz 1, Berlin 10117, Germany

88: Bernstein Focus State Dependencies of Learning and Bernstein Center for Computational Neuroscience, Berlin, Germany

89: Einstein Center for Neuroscience Berlin, Charitéplatz 1, Berlin 10117, Germany

90: Einstein Center Digital Future, Wilhelmstraße 67, Berlin 10117, Germany

91: Invicro, London, UK

92: Division of Brain Sciences, Imperial College London, London, UK

93: University of Texas at San Antonio, San Antonio, TX, USA

94: Department of Neurosurgery, Oregon Health & Science University, Portland, OR, USA

95: McGill Centre for Integrative Neuroscience (MCIN), Montréal Neurological Institute, McGill University, Montréal, QC, Canada

96: University of Washington, Department of Psychology and eScience Institute, Seattle, WA, USA

97: University of South Carolina, Department of Psychology, Columbia, SC, USA

98: Inria, Institut du Cerveau - Paris Brain Institute, Paris, France

99: Lifespan Informatics and Neuroimaging Center (PennLINC), Department of Psychiatry, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

100: Berlin Institute of Health (BIH) at Charité – Universitätsmedizin Berlin, Charitéplatz 1, 10117, Berlin, Germany

101: Department of Neurology with Experimental Neurology, Charité, Universitätsmedizin Berlin, Corporate member of Freie Universität Berlin and Humboldt Universität zu Berlin, Charitéplatz 1, 10117, Berlin, Germany

102: Einstein Center for Neuroscience Berlin, Charitéplatz 1, 10117, Berlin, Germany

103: Einstein Center Digital Future, Wilhelmstraße 67, 10117, Berlin, Germany

104: The Florey Institute of Neuroscience and Mental Health, Heidelberg, Victoria, Australia

105: The Florey Department of Neuroscience and Mental Health, The University of Melbourne, Parkville, Victoria, Australia

106: Institute for Diagnostic and Interventional Radiology and Neuroradiology, University

Medicine Essen, Essen, Germany

107: Institut de Neurosciences de la Timone (INT), UMR7289, CNRS, Aix-Marseille Université, France

108: University of Oregon, Department of Human Physiology, Eugene, OR, USA

109: University of California, Irvine, CA, USA

110: Inria, Université Paris Saclay, Saclay, France

111: Department of Cognitive Science, Halıcıoğlu Data Science Institute, and Neurosciences Graduate Program, University of California, San Diego, La Jolla, CA, USA

112: Kiel University, Kiel, Germany

113: University of Birmingham, Centre for Human Brain Health and School of Psychology, Birmingham, UK

114: Google X. Mountain View, CA USA

## **Abstract**

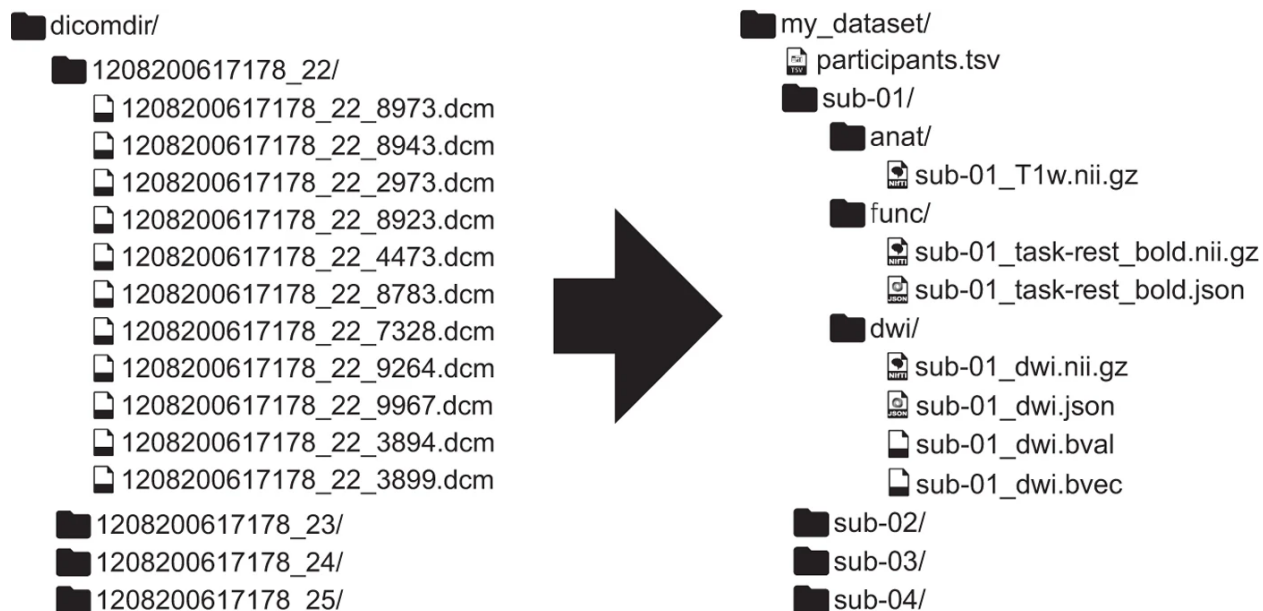
The Brain Imaging Data Structure (BIDS) is a community-driven standard for the organization of data and metadata from a growing range of neuroscience modalities. This paper is meant as a history of how the standard has developed and grown over time. We outline the principles behind the project, the mechanisms by which it has been extended, and some of the challenges being addressed as it evolves. We also discuss the lessons learned through the project, with the aim of enabling researchers in other domains to learn from the success of BIDS.

## **Main text**

The sharing of scientific research data is beneficial in numerous ways. Foremost, it maximizes the potential knowledge to be derived from the data, thus maximizing the benefit to the stakeholders who fund the research and the contributions of research participants. It also provides the means for researchers to attempt to reproduce the work of others in their field, which is an essential component of science. Further, it levels the scientific playing field by providing data to researchers from under-resourced environments or those without data acquisition capabilities, which they can use to develop novel analysis methods or test new scientific hypotheses. Given the rise of machine learning methods in science, another unsung benefit of data sharing is that it provides larger and more diverse training datasets, which can increase robustness and decrease overfitting and bias. For all of these reasons, the sharing of data has become increasingly common across science, as have demands from funding agencies and publishers that data be shared. Within the field of neuroimaging, data sharing efforts started around 2000 with the fMRI Data Center (Van Horn et al., 2013). Ten years later it began to flourish with the advent of the International Neuroimaging Data-sharing Initiative/Functional Connectomes Project (INDI/FCP) (Mennes et al., 2013). Subsequent projects have focused on prospective sharing of data, including the Human Connectome Project (Van Essen et al., 2013) and the Adolescent Brain Cognitive Development (ABCD) Study (Casey et al., 2018), which have had a major impact on the field by providing large quantities of neuroimaging data to researchers.

The sharing of data is a worthy goal, but only if the data are shared in a way that makes them FAIR (Findable, Accessible, Interoperable, and Reusable) (Wilkinson et al., 2016). One major contributor to FAIRness is the use of standard file formats, which allow researchers to reuse data across multiple software platforms. The field of neuroimaging research using MRI has benefitted from the longstanding convergence of the field on a standard file format for imaging data, the Neuroimaging Informatics Technology Initiative (NIfTI) format (Cox, R. W., Ashburner, J., Breman, H., & Fissell, K., 2004), which is used by nearly all major MRI analysis software packages. However, beyond file formats, the data and associated metadata must be organized in such a way that data recipients can quickly and accurately understand the contents of the data. The benefits of a clear, formal organization scheme for data include minimizing the burden of curation for researchers and for data sharing repositories, reducing the likelihood of errors due to misunderstanding or misinterpretation of the data, enabling the development of analysis tools that can automatically utilize the structure of the data to analyze them appropriately and with minimal user input, and affording the ability to automatically validate the data to determine whether they meet the standard (Gorgolewski et al., 2016).

In this paper we outline the history, current status, and future directions of the Brain Imaging Data Structure (BIDS), which has become a widely accepted community-driven data standard within the neuroscience community. The goal of this exposition is to provide a written history of the events that gave rise to BIDS and outline the events leading to its establishment, which we hope will be of interest to researchers in the neuroimaging field as well as to those in other fields working to establish successful new data standards.



**Figure 1.** A reproduction of Gorgolewski et al., 2016, Figure 1, showing an example mapping from DICOM to BIDS. BIDS is a community-driven standard for organizing, naming, and annotating neuroimaging data that places a heavy emphasis on human- and machine-readability. Since its initial publication, BIDS has expanded from structural, functional

and diffusion MRI to incorporate other MR methodologies such as arterial spin labeling and other recording technologies such as electrophysiology.

### The birth of BIDS

The birth of BIDS can be traced back ultimately to a social media post by Russ Poldrack on October 17, 2014<sup>1</sup>. The post referred to a talk that Chris Gorgolewski had given at the weekly Stanford cognitive/neuroscience seminar (“Frisem”). A reply to the post by Stuart Buck, then a program officer at the Laura and John Arnold Foundation, led to a discussion that ultimately resulted in a substantial grant from the Foundation to the Stanford group, with the aim of developing a new data sharing platform that would supercede the OpenfMRI data sharing platform (Poldrack et al., 2013) that the group had previously run. This new platform would ultimately become the OpenNeuro archive (Markiewicz et al., 2021) and the support from the Arnold Foundation would help start the work on BIDS.

One of the major challenges of the OpenfMRI project had been data curation. The project had developed an in-house data organization scheme (see Figure 1 in Poldrack et al., 2013), which reflected common practice in many labs but was built around a specific workflow for task fMRI analysis based on the FSL software package (Jenkinson et al., 2012). Standardization of file layout and study design metadata gained traction in other projects interested in formalizing the loading of data; e.g., PyMVPA (Hanke et al., 2009) version 2.6.1 (November 2014), included “Direct support for loading data and design from openfmri.org-style datasets”<sup>2</sup>. Since the scheme was not formally described, researchers wishing to deposit data could not easily transform their data to meet it. Instead, data depositors would send their data to the OpenfMRI team and a curator within the team would work with the depositor to transform the data to meet the informal standard. There was also no way to easily determine whether this transformation was correct, other than running it through the automated workflow and seeing whether the workflow ran successfully. This resulted in significant personnel costs and greatly limited the amount of data that could be ingested to the archive. This standard was also limited to a very specific type of MRI data and analysis workflow, and thus was not necessarily useful for a broad group of researchers.

When Chris Gorgolewski and Russ Poldrack began discussing the development of a new archive, they recognized that it was essential to substantially shift the burden of curation from the archive to the data owners, in order to make the archive financially sustainable in the long run. It became clear that this would require a detailed and general scheme for the organization of the intended data types, and that this scheme should support automated validation so that users can upload data and share them immediately without the need for manual curation. As it happened, the Stanford team was already participating in the Neuroimaging Data Sharing Task Force (or NIDASH), a collective effort under the umbrella of the International Neuroinformatics Coordinating Facility (INCF). Amongst other initiatives, this group was developing the Neuroimaging Data Model (NIDM) project, an informatics framework for the formal description of

---

<sup>1</sup> <https://twitter.com/russpoldrack/status/523263185764491264>

<sup>2</sup> <https://github.com/PyMVPA/PyMVPA/pull/240>

neuroimaging experiments and datasets (Maumet et al., 2016). As part of a series of meetings organized by INCF to further the progress of NIDM and other data sharing efforts, a meeting had been planned for January 2015 at Stanford in order to discuss the development of a NIDM model for fMRI experiments, based on the OpenfMRI use case. This was the meeting at which BIDS was first envisioned.

The Stanford meeting was held January 27th-30th, 2015, with significant support from the INCF<sup>3</sup>. The in-person attendees were Mathew Abrams, Michel Dumontier, Guillaume Flandin, Chris Gorgolewski, Karl Helmer, David Keator, Camille Maumet, Nolan Nichols, Russ Poldrack, Jean-Baptiste Poline, Ariel Rokem, and Vanessa Sochat; other invitees attending remotely included Satrajit Ghosh, Yaroslav Halchenko, Michael Hanke, David Kennedy, Angie Laird, Tom Nichols, and Jessica Turner. The meeting started with presentations on a number of ongoing relevant projects and their relation to the OpenfMRI use case. The first explicit mention of BIDS came on Day 2, when one of the subgroups was labeled as “Subgroup1: OBIDS (Open Brain Imaging Data Structure) format proposal (derived from the OpenfMRI format)”. A photo of a whiteboard drawing (Figure 2) shows the intended separation of standards, with a directory-based format (intended for most users) and a Resource Description Framework (RDF) based format (intended for computationally advanced users); the former is what would become BIDS.

Following the meeting in January 2015, there was a substantial effort to develop a set of examples that would be circulated with the initial draft of the specification, based primarily upon datasets from the OpenfMRI database. An additional meeting was held at the OHBM conference in Honolulu in June, 2015 and consequently at the INCF 2015 congress in Cairns, Australia, where both OBIDS and NIDM groups worked together as part of NIDASH to cross-fertilize both standards with basic metadata to describe MRI experiments and data. The BIDS specification draft and 22 example datasets were disseminated to the community for comments on September 21st 2015, along with an early version of the JavaScript-based bids-validator<sup>4</sup>, developed by Squishymedia. Squishymedia was a contractor also responsible for primary development of the OpenNeuro archive; it later ended operations in 2021, after which Nell Hardcastle (lead developer on the OpenNeuro project) joined the Stanford team. A NIDASH task force meeting was held in Chicago in October 2015 in advance of the Society for Neuroscience (SFN) Annual Meeting, and a BIDS leaflet<sup>5</sup> was distributed at the SFN meeting to promote the standard to a wider audience.

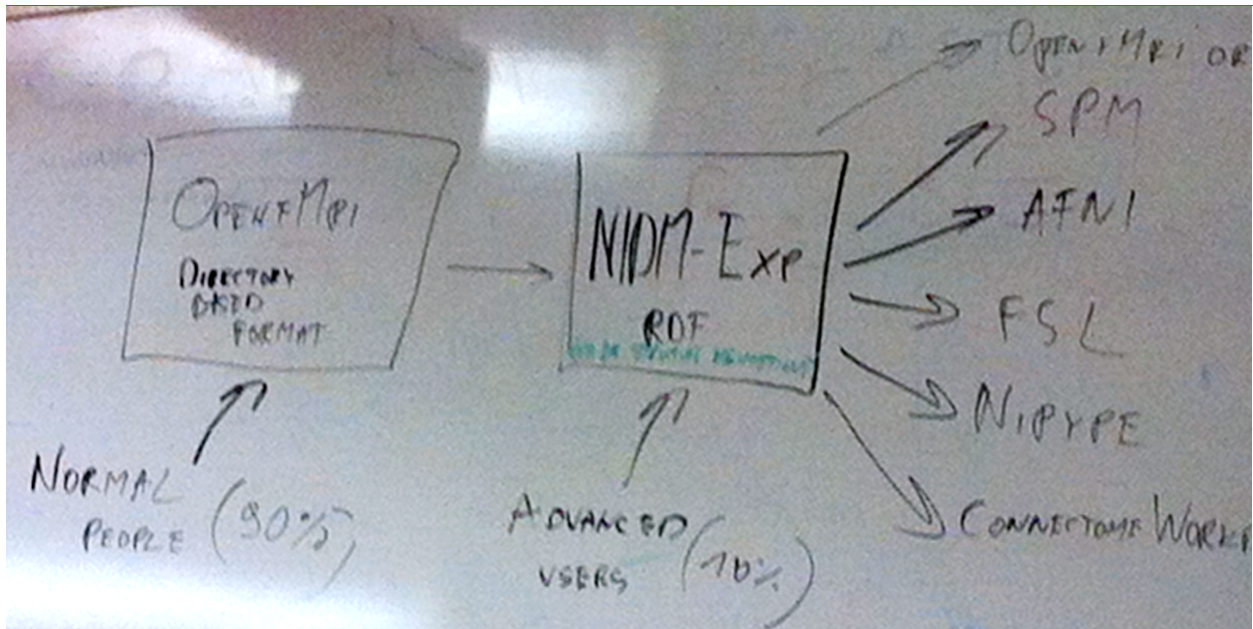
---

<sup>3</sup> The original agenda and notes from the meeting are available at <https://osf.io/kmavh/>

<sup>4</sup> <https://github.com/bids-standard/bids-validator/>

<sup>5</sup> A copy available from [https://web.archive.org/web/20230519223151/https://neuro.debian.net/\\_files/brochure\\_bids.pdf](https://web.archive.org/web/20230519223151/https://neuro.debian.net/_files/brochure_bids.pdf) and also listed on <https://centerforopenneuroscience.org/engage>.





**Figure 2:** A snapshot of the whiteboard at the initial BIDS meeting (January 27-30, 2015), outlining the intended separation of a directory-based format (which would become BIDS) and a formal RDF-based description (which would become NIDM-Experiment).

### The establishment of BIDS

The publication of a paper in *Scientific Data* in June 2016 (Gorgolewski et al., 2016) marked the initial official release of the standard. As BIDS developed, a set of principles became enshrined which have been central to subsequent decision-making in the project. The first principle is that *adoption is crucial*. This led to a focus on engaging the relevant research community, and on keeping the standard as close as possible to the practices already in use in the community, thus requiring engineering ingenuity to build a solution from simple existing components while avoiding technical complexity. For example, this led to the focus on human-readable text formats rather than more complicated formats such as RDF or XML (Extensible Markup Language).

The second principle is: *don't reinvent the wheel*. This led to the use of existing standards and file formats whenever possible, including NIFTI, JSON (JavaScript Object Notation), and tab-separated text files (TSV), along with other projects later incorporated, such as the Hierarchical Event Descriptor (HED) annotation system (Robbins et al., 2022). It is important to note the historical trajectory that led to the choice of these particular formats. For example, Digital Imaging and Communications in Medicine (DICOM) has long been the dominant imaging data format within the field of medical imaging. However, due to the overall complexity of DICOM and the early absence of explicit MRI support (with DICOM Working Group 16 on MRI only established in 1998), the academic and research neuroimaging field in the 1990s turned to

simpler formats, such as Analyze (Robb et al., 1989). Subsequently, in the early 2000s, this simpler format inspired the creation of the NIfTI format, which quickly became the standard for data exchange and storage in academic neuroimaging research. As advancements were made in MRI scanners, enhancing their capability to export acquired data in DICOM format, tools like `dcm2nii` and `dcm2niix` (Li et al., 2016) emerged to convert DICOM series into the simpler and widely supported NIfTI format. However, as NIfTI encapsulates only minimal metadata, compared to the comprehensive metadata represented within DICOM, these tools also needed to export the more complete version of the metadata into separate (“sidecar”) files, generally represented using JSON. Currently, over 70 metadata fields in BIDS sidecar files originate from or are based on metadata acquired from DICOM series. DICOM has become the primary format for data export from MRI and PET scanners, which are then converted to NIfTI for interoperability with commonly used software packages.

Finally, the group has focused on the work that will achieve the largest amount of impact, and on not letting the perfect be the enemy of the good; this principle is now referred to informally as the “80/20 rule”. This requires admitting that there will always be edge cases that can’t be well accommodated by the standard, but that this is fine as long as the standard works well for most people in most circumstances most of the time. Perhaps most important was the community outreach performed by Chris Gorgolewski as the leader of the project, whose dogged pursuit of feedback from a broad group of stakeholders led to a general feeling of inclusiveness and community-mindedness.

One particular idea for incentivizing the use of the standard was to develop software applications that could be easily applied to a BIDS dataset. This idea led to the concept of the *BIDS App* (Gorgolewski et al., 2017), which refers to a containerized software application of a specific analysis pipeline that is aware of the BIDS standard and can thus be applied directly to a BIDS dataset. This model has become highly successful, leading to widely used tools, including `fMRIPrep` (Esteban, Markiewicz, et al., 2019), `QSIprep` (Cieslak et al., 2021), `PyMVPA` (Torabian et al., 2023), and `MRIQC` (Esteban et al., 2017), and has almost certainly driven adoption of the standard by researchers who wish to use these software tools<sup>6</sup>. Those early efforts on BIDS Apps also catalyzed the development of interoperability layers for existing tools (e.g., the MNE-BIDS interoperability layer for MNE-Python; (Appelhoff et al., 2019) as well as generic libraries to facilitate interaction with BIDS datasets including `PyBIDS` (particularly `PyBIDS`; Yarkoni et al., 2019), `BIDS-Matlab` (Gau et al., 2022), and `nilearn` (RRID:SCR\_001362). In particular, interactions between BIDS and developers of existing tools were nurtured at a series of coding sprints organized by the Stanford Center for Reproducible Neuroscience in 2016 and 2017.

## BIDS Extensions

The initial BIDS standard was focused on specific types of MRI data, but even before its initial release it became clear that there would be a need for extensions to accommodate additional

---

<sup>6</sup> See <https://bids-apps.neuroimaging.io/apps/> for a complete list of currently available BIDS apps.

data types. Soon after the establishment of the standard, the BIDS team began to formalize the concept of extending BIDS through “BIDS Extension Proposals” (BEPs), inspired by the Python Enhancement Proposal (PEP) process. The first extension proposals to be intensively discussed and eventually merged into the specification were BEP008 (“Magnetoencephalography”) and BEP007 (“Hierarchical Event Descriptor” HED tags) in BIDS version 1.1, which were quickly followed by BEP006 (“Electroencephalography”) and BEP010 (“intracranial Electroencephalography”) in BIDS version 1.2. As the BEP numbers suggest, there were proposals even before these (e.g., BEPs 001 through 005); however, they progressed more slowly and were merged at later times or are still works in progress; see Table 1 for a list of merged BEPs and [https://bids.neuroimaging.io/get\\_involved.html](https://bids.neuroimaging.io/get_involved.html) for full documentation of all BEPs.

The process of developing a BEP was deliberately designed to be as inclusive and with as few technical hurdles for potential contributors as possible. Although the process has become more formal throughout the years and central documentation about the necessary steps and deliverables exists (<https://bids-extensions.readthedocs.io>), its core has stayed the same: Interested parties may suggest a BEP via the BIDS mailing list or central development repository on Github. After an initial check by a group that has come to be known as the “BIDS maintainers” (see below) of whether the proposed idea broadly fits the scope of BIDS, a BEP receives a number and an entry in the list of ongoing BEPs. The BEP itself is then developed using shared public documents that are advertised via the central BIDS website, and where any person may become a contributor by adding comments and suggestions. Each BEP is headed by a team of “leads” or “moderators,” who are responsible for tracking individual contributions, steering discussions, and accepting and rejecting suggestions to the document. The BEP leads are furthermore responsible for organizing virtual discussion rounds (e.g., video calls) with all contributors, and for explicitly reaching out to the broader community working with the relevant methods. Several community members of BIDS that are active to this day have initially found their way into the community by stumbling over a BEP document and offering their feedback and/or minor additions (i.e., simple typo corrections), further highlighting the inclusiveness of the community-driven process.

The first BIDS extension released for a new neuroimaging modality was for magnetoencephalography (MEG) (Niso et al., 2018). The design process was initiated around the creation of the Open MEG Archive (OMEGA; (Niso et al., 2016)) at the Montreal Neurological Institute, for which BIDS was identified as a principled framework to organize the multimodal data. Sylvain Baillet and Guiomar Niso thought that the development and adoption of a data organization standard for MEG would require the active participation of academic software developers in the field. An ad hoc working group first met around the idea of a possible MEG extension to BIDS at the International Conference on Biomagnetism held early October 2016 in Seoul, South Korea<sup>7</sup>. The first elements of specification for MEG-BIDS were shared as a preprint (Niso et al., 2018), which received comments and suggestions from the community. In July 2017, further input from the MEG community was collected through an online poll survey. The poll received 78 international entries, with the most salient results indicating strong interest

---

<sup>7</sup> [http://www.biomag2016.org/download/program/BIOMAG2016\\_Poster\\_Sessions.pdf](http://www.biomag2016.org/download/program/BIOMAG2016_Poster_Sessions.pdf), Poster Mo-P011

in a common standard for MEG data organization (~99%) and affirming a willingness to try a MEG-BIDS solution (97%). Although the initial thought was to develop a single BIDS extension encompassing all electrophysiological data modalities, the working group identified that each data type required specific adjustments that would be best addressed by their respective experts. The MEG extension to BIDS introduced substantial innovations that stemmed from the original MRI elements, considering the radically different nature of the instrument technology and time-resolved data types, paving the way for other electrophysiological modalities, such as EEG (Pernet et al., 2019) and iEEG (Holdgraf et al., 2019). With these three electrophysiology extensions (MEG, EEG, iEEG), BIDS became a common structure for a larger user group, including many different types of neuroimaging scientists beyond the original use case of MRI.

<b>BEP #</b>	<b>Title</b>	<b>Date merged</b>	<b>Publication</b>
BEP001	Quantitative MRI (qMRI)	2021-02-23 (v1.5.0)	(Karakuzu et al., 2022)
BEP003	Common Derivatives	2020-06-11 (v1.4.0)	
BEP005	Arterial Spin Labeling (ASL)	2021-02-23 (v1.5.0)	(Clement et al., 2022)
BEP006	Electroencephalography (EEG)	2019-03-04 (v1.2.0)	(Pernet et al., 2019)
BEP007	Hierarchical Event Descriptor (HED) Tags	2018-04-19 (v1.1.0)	(Robbins et al., 2022)
BEP008	Magnetoencephalography (MEG)	2018-04-19 (v1.1.0)	(Niso et al., 2018)
BEP009	Positron Emission Tomography (PET)	2021-04-22 (v1.6.0)	(Norgaard et al., 2022)
BEP010	intracranial Electroencephalography (iEEG)	2019-03-04 (v1.2.0)	(Holdgraf et al., 2019)
BEP018	Genetic information	2020-04-14 (v1.3.0)	(Moreau et al., 2020)
BEP030	Near Infrared Spectroscopy (NIRS)	2022-10-29 (v1.8.0)	

BEP031	Microscopy	2022-02-15 (v1.7.0)	(Bourget et al., 2022)
--------	------------	---------------------	------------------------

**Table 1.** Merged BIDS Extensions as of June 2023. A full link of all BEPs (including works in progress) is available at [https://bids.neuroimaging.io/get\\_involved.html](https://bids.neuroimaging.io/get_involved.html).

## BIDS Derivatives

The initial BIDS standard focused on “raw” data<sup>8</sup>. As a result of the excitement in the community for the BIDS standard, early growth of the project was horizontal, with a number of new data modalities being added over time. However, the need for a standard to organize processed data quickly became apparent, given that the end results of these studies are derived not directly from raw data, but rather through potentially numerous intermediate processing steps, possibly produced by one or more BIDS Apps. Discussions began in 2017 of a “layout file for derivatives”, and this led to the further development of the concept of a BIDS Derivatives Standard. In 2018, the US BRAIN Initiative funded a grant to develop BIDS standards for processed data, computational models, and statistical models<sup>9</sup>. This grant funded three meetings held in 2018 and 2019, in which members of the community convened to work on these new standards.

The description of processed data is much more challenging than the description of raw data, since there is an almost infinite set of possible processing operations that might be applied to a neuroimaging dataset and outputs that can be generated. Compounding this challenge are differences among practitioners of different modalities (for example, functional and diffusion imaging) for sharing processed results, particularly with regard to what is considered necessary metadata and provenance. At an August 2017 meeting at Stanford, it was agreed to split the BIDS Derivatives specification (BEP003) into a series of BEPs focused on particular modalities or use cases, to allow for independent development on largely independent components. An August 2018 meeting focused specifically on derivatives with a goal of advancing each sub-proposal, establishing common principles, and recombining into a single BEP, leading to a finalized document in December 2018 that was called “BIDS Derivatives Release Candidate 1” that was advertised for implementation. Following the OHBM 2019 meeting, in the absence of Chris Gorgolewski as a primary driver for integration of all derivatives at once (see next section), the sections of BIDS Derivatives were again broken into sub-proposals, with the common principles being the most important target for inclusion into BIDS. These Common Derivatives principles were finally released as part of BIDS 1.4.0 in June 2020.

<sup>8</sup> The concept of “raw” data remains controversial, with different researchers having very different interpretations of the term. The initial usage was meant to capture its most common usage in the MRI field, i.e. data as downloaded from the MRI scanner, usually in DICOM format; however, this usage conflicted with the term’s usage in other modalities. At a BIDS Derivatives meeting in Copenhagen, Denmark in 2023, an operational definition of the term “raw” was proposed, such that a “raw BIDS dataset” is one for which there is no source BIDS dataset specified.

<sup>9</sup> Original grant proposal and meeting summaries at <https://osf.io/c3dgx/>



The initial push for BIDS Derivatives also included two projects aimed at the description of computational artifacts related to neuroimaging data. The first was the BIDS Computational Models framework, which was meant to describe computational models used in neuroimaging data analysis (Poldrack et al., 2019). The Stanford group joined forces with Jonathan Cohen at Princeton University, who was developing a related project called PsyNeuLink<sup>10</sup>, and held a meeting at Princeton in April 2019. This was a wide-ranging meeting, which focused on establishing a set of use cases and example implementations that could be used to guide further development. Ultimately, this effort split in two different directions. One effort to develop a generic framework for the specification of a broad set of computational models, led by Jonathan Cohen, Pdraig Gleeson, and Tal Yarkoni, became the ModECI Model Description Framework (Gleeson et al., 2023); this is an example where a standard initially developed as part of BIDS became a freestanding separate project. A separate effort led by Petra Ritter aimed at establishing a more constrained framework for the specification of inputs to and outputs from computational network modeling software (such as the Virtual Brain; Sanz Leon et al., 2013; Schirner et al., 2022); this project is currently developing a BEP for review.

Another aim of the BIDS Derivatives project was to develop a framework to describe common statistical models that are applied to neuroimaging data. Statistical modeling of neuroimaging data follows broadly stereotyped patterns, with researchers making a relatively small number of significant decisions within established pipelines. However, the details of the pipelines' construction lead to a high degree of analytical flexibility (Botvinik-Nezer et al., 2020), which is difficult to capture using idiosyncratic user-generated code. The "BIDS Stats Model" specification was created on the observation that a majority of neuroimaging analysis methods can be represented by the General Linear Model (GLM) and common data flow operations. The effort kicked off with an initial document by Chris Gorgolewski, Tal Yarkoni, and Satra Ghosh in September 2016, which became BEP002. Subsequent efforts to develop the specification and the tooling were driven by Neuroscout (de la Vega et al., 2022), which was the first project to use the specification in production. An effort to substantially complete the specification was undertaken at a meeting at Stanford in October 2018, and was subsequently led by Tal Yarkoni, Alejandro de la Vega, and Chris Markiewicz until Yarkoni left academia in 2021. The specification was finalized at a Spring 2022 meeting at the University of Texas and published as a standalone website (<https://bids-standard.github.io/stats-models/>) with dual goals of creating user-facing documentation, validator, and technical specifications for implementations. Currently, the BIDS Stats Model specification is able to represent a majority of GLM models used in neuroimaging, and is implemented by PyBIDS/FitLins and BIDS-MATLAB/bidspm. Ongoing efforts are being made to develop tools to facilitate user-friendly model specification, as well as integration of the specification with the most popular neuroimaging analysis packages.

Other efforts have worked to add specifications for specific forms of derivative data across multiple modalities. One example is the BIDS Connectivity project, a BRAIN Initiative project led by Franco Pestilli of the University of Texas that aims to extend the BIDS standard to describe research objects commonly used in experiments related to brain connectivity. The

---

<sup>10</sup> <https://princetonuniversity.github.io/PsyNeuLink/>

BIDS connectivity extension(s) project comprises most major neuroimaging data modalities (sMRI, fMRI, DWI, PET, EEG, iEEG, and MEG) and covers data products pertaining to a broad range of analyses, including structural and functional connectivity matrices, seed-based connectivity maps, networks based on dimensionality reduction and tractograms and tractometry. The BIDS connectivity project is the first example of a large-scale coordination effort spanning multiple BEPs, with at least five BEPs encompassing different derivative types being advanced in a coordinated manner.

### Weathering an existential crisis

BIDS was envisioned from the beginning as a community project, but in practice its leadership fell heavily on the shoulders of its founder, Chris Gorgolewski. Chris's strong leadership role in the BIDS community ended with his departure from academia in March 2019, and the resulting leadership vacuum led to an existential crisis within the BIDS community, which had by then grown substantially. This discussion came to a head at a BIDS Town Hall meeting organized at the OHBM 2019 meeting in Rome<sup>11</sup>. Following an overview of the BIDS project, there was a spirited discussion about the future leadership of BIDS. This discussion continued on the BIDS Standard GitHub repository<sup>12</sup>; a particularly impactful argument was made by Kirstie Whitaker in favor of community governance and against the “tyranny of structurelessness” (Freeman, 1972) in which a small number of voices come to dominate the discussion.

In the wake of the OHBM meeting, Tom Nichols proposed the idea of a BIDS Steering Group, which would be elected based on a vote by the BIDS community. In parallel, a plan for community governance was developed by members of the community (see <https://bids.neuroimaging.io/governance>). The development of this process was led substantially by Franklin Feingold from the Stanford team, who served as the principal project manager for BIDS from 2018 through 2022. The first election was held in October 2019. The winning slate was chaired by Guiomar Niso including Melanie Ganz, Robert Oostenveld, Russ Poldrack, and Kirstie Whitaker, a team that represented the diversity of neuroimaging modalities already then in BIDS, including MRI, PET, MEG, and EEG. In 2021 an election was held to replace Kirstie Whitaker, which resulted in the election of Ariel Rokem. In 2022, an election was held to replace Melanie Ganz and Russ Poldrack, which resulted in the election of Yaroslav Halchenko and Cyril Pernet. The continued successful operation of the community in the absence of its founder demonstrates the strength of the community governance model.

The governance process also formalized the role of the BIDS maintainers. This was spurred by the move of the primary BIDS specification document from a collaboratively edited document to a website generated from a version-controlled GitHub repository, which increased the technical barriers to contribution. The initial maintainers' group consisted of Stefan Appelhoff, Franklin Feingold, Ross Blair, and Chris Markiewicz, who took responsibility for managing the repository and validator and facilitating contributions from users less familiar with working in GitHub. The work of the maintainers group includes social infrastructure, such as the BIDS website and

---

<sup>11</sup> Slides available at <https://osf.io/kmavh/>

<sup>12</sup> <https://github.com/bids-standard/bids-specification/pull/104>

social media, as well as running the steering group elections. The maintainers group remains a self-selected set of contributors that collectively make infrastructure decisions around the specification, validator, and example datasets, facilitate additions to the standard, and advise the steering group. Taylor Salo and Remi Gau joined the group in 2020, Anthony Galassi and Eric Earl in 2021, Christine Rogers and Nell Hardcastle in 2023.

One key strength of BIDS is the engineering acumen of the scientists involved, which is reflected in the way that the standard is now rendered for public consumption. Initially the published standard for the web site and a machine-readable JSON schema used by the JavaScript validator were kept in sync manually. The development of a generic, declarative schema to describe the standard (described further below) allowed the BIDS maintainers and other contributors to develop tools that generate some crucial specification text and tables directly from the very same schema as the validator and therefore avoid conflicts between the published standard and the validator. The unified schema also allowed downstream tools, such as HeuDiConv (RRID:SCR\_017427), to avoid hardcoding the standard, thus making them also more robust to changes to the BIDS standard.

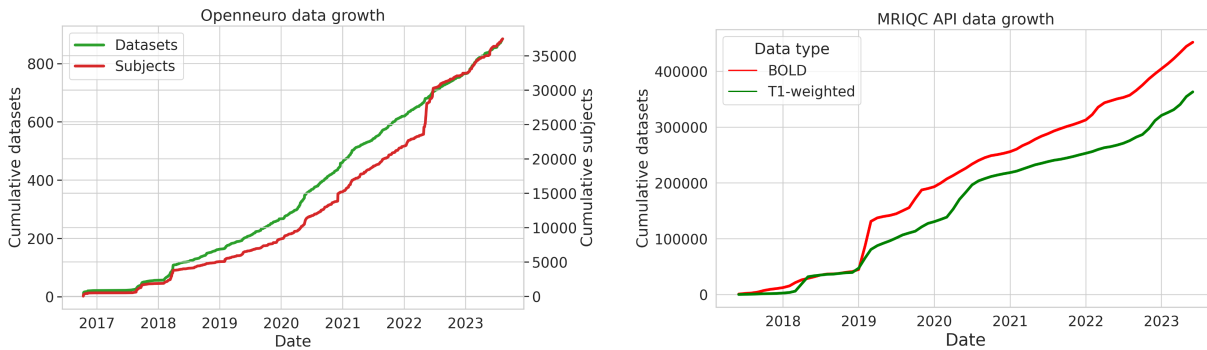
#### The present state of BIDS

At present, BIDS is a highly successful example of a community-driven standard for data organization; to our knowledge, it is one of the only grass-roots standards to have gained such broad acceptance within its field. This community support was recognized by the INCF when they endorsed BIDS as a best practice standard in 2018, and re-endorsed it in 2021. Not only has the adoption of BIDS continued to grow, but BIDS itself has grown to reflect developments in the field.

A massive amount of data is now shared in the BIDS format. The OpenNeuro archive (as of June 2023) shares data for more than 34,000 individuals from more than 850 BIDS datasets; Figure 3 shows the consistent increase in the size of this database over time. The ABCD-BIDS Community Collection (Feczko et al., 2021) shares a BIDS version of the ABCD dataset that includes longitudinal data from 11,877 children.

Another lens into BIDS usage comes from the statistics collected by the MRIQC BIDS app, which stores telemetry information about each MRI run (unless the user opts out) that is then shared via an open web API (Esteban, Blair, et al., 2019). Figure 3 shows the cumulative number of unique images (distinguished based on a checksum of each image) submitted to the MRIQC web API between 2018 and June 2023, which demonstrates a sustained and consistent growth in the number of datasets converted to BIDS. Given that these datasets represent only a subset of the total number of BIDS datasets, we would estimate that there are likely to be well over 100,000 datasets with millions of images that have been converted to BIDS at this point. These results highlight the fact that BIDS is being used by a significant number of researchers in the community.



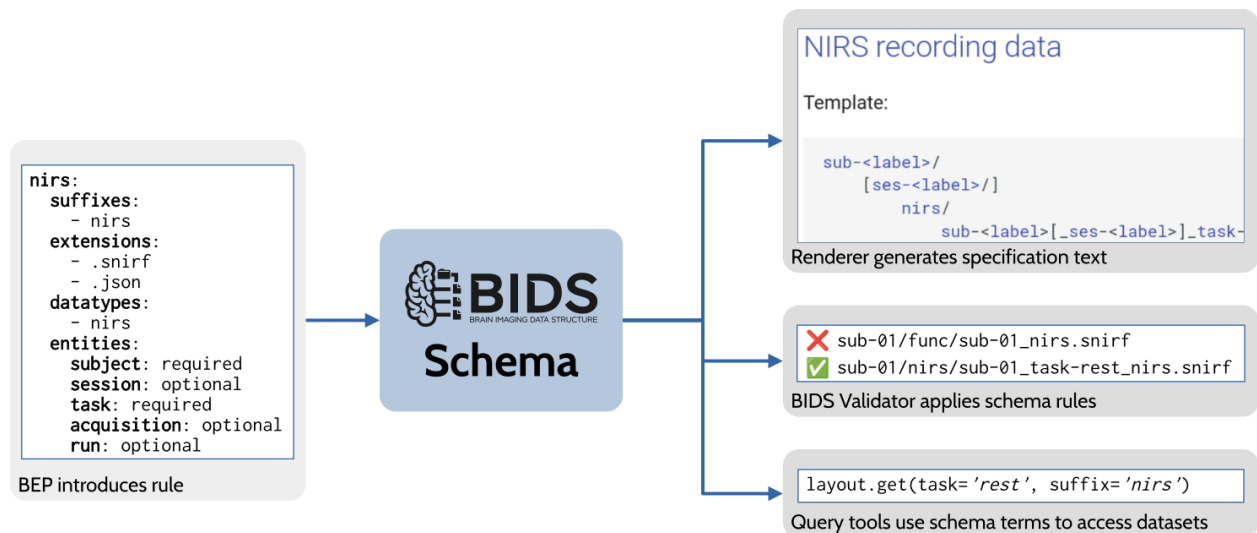


**Figure 3.** Growing usage of BIDS over time. Left: Growth of the OpenNeuro database since its inception in 2017, adapted from (Markiewicz et al., 2021). Right: Cumulative number of unique T1-weighted anatomical (t1w) and BOLD images from BIDS datasets submitted to the MRIQC web API (Esteban, Blair, et al., 2019) from 2018 to June 2023. Source data and code to generate figures available at <https://osf.io/x7fh8/>.

BIDS has also enabled a number of important data integrations that grow community adoption via the software ecosystem. For example, the cloud platforms [brainlife.io](https://brainlife.io) (Hayashi et al., 2023) and [nemar.org](https://nemar.org) (Delorme et al., 2022) have utilized BIDS to provide ready-to-use data services. The BIDS standard provides a common data format for ingestion and exchange; combined with DataLad (Halchenko et al., 2021) as a transfer mechanism, OpenNeuro datasets are made seamlessly available to [brainlife.io](https://brainlife.io) and NEMAR users. Further, the BIDS Apps specification (Gorgolewski et al., 2017) enabled BIDS-aware platforms such as [brainlife.io](https://brainlife.io) to make third-party BIDS Apps available to a large community of users. Other open-source platforms that integrate BIDS support, such as LORIS (Das et al., 2011) and CBRAIN (Sherif et al., 2014), also contribute open tools and workflows to onboard new user groups to BIDS (Rogers et al., 2022) as a basis for research collaboration. To take one final example, another BRAIN Initiative® data archive, DANDI (RRID:SCR\_017571), has effectively employed BIDS in harmony with standards from other subfields, such as Neurodata Without Borders (Teeters et al., 2015) and OME-Zarr (Moore et al., 2023), to facilitate integration across diverse domains of neuroscience data.

One major achievement of the BIDS community in recent years has been the development of a machine-readable schema to represent the standard. The early releases of BIDS consisted of a specification written in English and a validator written in JavaScript, which aspired to have a validation procedure for each English rule. As the standard grew to include more data types, the number of people with the expertise needed to adapt and maintain these parallel representations became vanishingly small. At the same time, each tool implementing BIDS recommendations resulted in yet another representation of the standard that required updating. These difficulties led to an effort to create a declarative (i.e., non-procedural) schema that is maintained as part of the specification document. This has four important benefits. First, it allows much easier inclusion of new elements to the standard (such as new BEPs). Second, it enables the consistent implementation of the validator across multiple languages. Third, it

makes it possible for the validator and any downstream tool using the schema to provide handling of the BIDS dataset specific to its BIDS version. Fourth, the inclusion of the schema with the specification encourages contributors who wish to propose a new rule to consider the difficulty of expressing that new rule within the schema. The schema-based validator has been implemented and is currently being tested, and is expected to replace the original validator in late 2023. It has already been adopted by the DANDI and OpenNeuro data archives, both of which contributed to the efforts to develop the schema and schema-based validator.



**Figure 4.** Overview of BIDS Schema usage. In this example, BEP 030 (NIRS) introduces a file naming rule to the schema as part of the BEP process. The schema rule is used to render a file naming template in the specification. The BIDS Validator uses the rule to identify valid NIRS data files while rejecting improperly named files. Finally, third-party tools, such as a query library, may ingest the updated schema to automatically gain access to new features of BIDS.

### The future of BIDS

As BIDS nears its tenth anniversary, its success has also led to increasing recognition of the limitations of the existing framework. This has in turn driven a growing discussion regarding the need for a new major version of BIDS (“BIDS 2.0”) that would introduce changes that are incompatible with the existing BIDS framework. A dedicated repository (<https://github.com/bids-standard/bids-2-devel/issues>) is collecting issues, a subset of which will be chosen for BIDS 2.0. It is likely that the discussion of a new version will continue over the next few years, given the significant difficulty that breaking changes would impose on tool developers.

As an example, one contentious issue that would require resolution in any successor specification is the “inheritance principle”. This allows metadata to be specified at a higher level in the data hierarchy, and be “inherited” by multiple data files to which it is applicable. On one hand, this introduces a degree of complexity to user comprehension and software interaction with the BIDS specification and BIDS datasets; on the other, it collapses redundant information

in a manner that is faithful to the hierarchical nature of the metadata at hand, which becomes increasingly pertinent for datasets of increasing complexity. Many such situations exist where identification of an issue or contention would not have been possible without hands-on experience within the BIDS ecosystem.

Another major challenge for the BIDS community is to decide where the standard should end. BIDS has already extended beyond what is classically considered as “neuroimaging data”, through the Microscopy extension (BEP031) and other extensions including genetic information (BEP018), eye tracking (BEP020), and movement data (BEP029). The success of BIDS has garnered interest from many researchers in developing related standards under the BIDS umbrella, but continued expansion of the scope of the standard also threatens to increase its complexity to a degree that it becomes difficult to change.

The question of where the standard ends is particularly relevant to continued standardization of BIDS Derivatives and their interaction with software tools.

1. *Hierarchical complexity*: The filesystem structure of the BIDS standard was designed to meet the requirements of BIDS Raw data, with data files arranged in an immutable hierarchy across datasets, then subject, (optionally) session, and finally imaging modality. While there may be tremendous flexibility in file naming within this structure, the hierarchy itself is inflexible. This did not pose an issue for BEP003 Common Derivatives, as each proposed derivative was a standalone piece of data derived from raw data from a single modality. However there are future prospects where this will no longer be the case, such as derivatives that are the result of explicitly multi-modal analysis, or data hierarchies more complex than that aforementioned (such as the result of a model fit that is spread across multiple data files), which cannot be faithfully represented in a BIDS structure with inheritance under the current standard.
2. *Derivatives as inputs*: From their inception, BIDS Apps were designed on the premise of taking as input a BIDS Raw dataset, and producing as output a final set of computed derivative data. The interest in BIDS Derivatives has been principally in facilitation of the sharing and unambiguous interpretation of such data. In parallel, a number of BIDS Apps have implemented the ability to identify and utilize the derivatives already computed by some other application rather than duplicating those requisite calculations internally. Combining these concepts presents an opportunity for the construction of large, complex, multi-modal processing pipelines incorporating disparate softwares. Just as many existing BIDS Apps are based on constructing and executing a directed acyclic graph of underlying individual commands from different neuroimaging software packages (Gorgolewski et al., 2017), larger analysis pipelines could be based on constructing and executing a directed acyclic graph of underlying BIDS Apps, with BIDS Derivatives serving as the means of translation between those Apps.
3. *Existing toolchains*: A large proportion of neuroimaging analyses are performed using one of a small set of existing software packages. Many of these have long-standing file

layout schemes and tools designed to operate on their own specific sets of derivatives. Conformance of such packages to a completely new data structure may therefore be too high an expectation. There are two key ways in which this problem should be considered. Firstly, any derivatives extension proposal should ideally have accompanying it a software tool for performing bidirectional conversions between the outputs of one or more major toolchains and the proposed specification; this would facilitate generation of conforming data by users and BIDS App developers alike, and subsequent manipulation / visualization of shared data using the originating toolchain. Secondly, an alternative approach would be to describe in the specification a way to encapsulate those results in the native toolchain format and merely annotate them with agreed common terms, an approach with precedents in BEPs 015 (Derivative mapping files) and 035 (Mega-analysis using non-compliant derivatives), and which may dovetail with BEP 028 (Provenance). This would permit *post hoc* annotation of derivatives that predate BIDS Derivatives and/or are generated by tools outside of the BIDS ecosystem; and while it may sacrifice ultimate filesystem convention conformity, it would not preclude the ongoing pursuit of such.

Another direction for future efforts is further integration with other related standards. Whereas the NIfTI file format has become standard within the MRI research community, DICOM has grown into the industry standard for a wide range of imaging modalities (such as physiology, etc.), while addressing the many shortcomings that had originally turned the neuroimaging research community towards simpler formats. As a result, many standardization efforts have been duplicated. As DICOM is the industry standard and more data will be arriving in DICOMs, coordination with developments in DICOM could help to ensure more rapid adoption of new imaging sequences and even modalities into the BIDS standard, which provides an umbrella organization at the study level. Continued work with related standards such as NWB and OME-Zarr will also be important to avoid unnecessary duplication of effort or conflicting recommendations for datasets and archives on the interface of neuroimaging and electrophysiology/microscopy.

The funding of continued BIDS development and maintenance remains a challenge. The BRAIN Initiative has funded many of the developments of BIDS, either directly (for BIDS development projects) or indirectly (through funding to data archives that have relied upon and contributed to BIDS). Much of the work to develop and maintain the standard is performed by the BIDS Maintainers; many of the maintainers are currently supported by related grants, which endangers the project given that grants usually have time windows of three to five years. The establishment of a foundation to support BIDS could be a useful future development.

#### Lessons Learned

Given the demonstrable success of BIDS, it is useful to ask: What lessons can be learned that might be useful for other standards projects? The first important point to acknowledge is that the success of any particular project derives in part from pure luck, so one should not overfit too heavily to the following. Nonetheless, we believe that there are several potentially important lessons to be learned.

## Main factors in the success of BIDS

**Absence of existing solutions.** Many standardization projects commence on the premise that there are multiple existing competing standards in that domain, each with their own strengths and weaknesses, and a new standard is sought that inherits more desirable attributes, takes into account lessons learned from those prior standards, and achieves more widespread adoption. The BIDS project was quite unique in that, within the neuroimaging community, there was effectively universal acceptance of ad hoc organization of neuroimaging data beyond the formatting of individual files. The benefits that were to be inherited through the prospect of field-wide harmonization were therefore implicitly attributed to BIDS itself.

**Clear use cases.** One of the main initial motivations to develop BIDS was to allow ingestion of datasets into data repositories at scale without the need for human curation. This not only provided a practical problem that defined the solution (as reflected by the project principles: “adoption is crucial”, “don’t reinvent the wheel”, and “80/20 rule”), but also enabled the team to fund the initial development of the standard via the OpenNeuro project. Subsequently, BIDS apps provided use cases for adoption by end-users, by enabling them to effortlessly process BIDS datasets using high-quality software tools.

**Solving a common end user problem.** Standardizing data organization was not only beneficial for data repositories and software platforms, but also for individual labs and projects. Even if data are never shared publicly, having a standard data organization scheme helps researchers within a lab group work together, and enables future reuse of the data within the lab. The development of BIDS absolved individual groups of the need to develop, document, and implement their own individual schemes, and provided tools (such as the validator and conversion tools) to help with this. Moving the project away from purely “open science” and “data sharing” framing by dropping “Open” from the name was crucial to reinforce the message that BIDS was not only for external sharing.

**Low technical barrier to entry.** The practices adopted by BIDS do not require sophisticated tools beyond those already widely in use in most laboratories, enabling rapid and widespread adoption. At the same time, community-based efforts arose to provide introductory materials to explain the core concepts of BIDS to a broad audience, such as the BIDS Starter Kit<sup>13</sup>. It is also worth acknowledging that community forums such as NeuroStars<sup>14</sup> have played a role in educating users as well as sharing best practices that fall outside the scope of the standard itself.

**Maturity and the size of the field.** When the BIDS efforts started, human neuroimaging was more than 30 years old. Standard patterns across experimental design and data types had already emerged, and common file formats were already widely adopted in some subfields (such as NIfTI within the MRI community). At the same time the overall size of the community (several thousand scientists, several hundred labs) allowed the team to gather detailed feedback from a significant portion of the community and required small overall alignment effort.

---

<sup>13</sup> <https://bids-standard.github.io/bids-starter-kit/>

<sup>14</sup> <https://neurostars.org/>

**Broad financial and institutional support.** BIDS has the benefit of early support by INCF (which supported a set of meetings of the neuroimaging data sharing (NIDASH) task force that included the first meeting at Stanford), and major funding from the Laura and John Arnold Foundation subsequently provided support for the early development of BIDS. Since then, BIDS development has been supported by grants from a number of institutions (NIMH, NSF, Novo Nordisk Foundation, French National Research Agency) to a number of different investigators<sup>15</sup>, which has broadened the base for support of the project and ensured that it didn't rely too heavily on any one particular grant or researcher.

**In person meetings with a diverse set of participants.** Key aspects of BIDS were drafted over many in-person meetings with participants traveling long and far to attend, and various BEPs have been started and developed spontaneously at conferences, workshops and hackathons. These in person meetings were key to brainstorming the different aspects of the standard and chart the work that happened later asynchronously.

**Open doors without “death by consensus”.** BIDS managed to finely balance having an open structure and listening to feedback from many members of the neuroimaging community, but at the same time avoiding trying to please everyone and creating a standard that was too flexible to be usable. An example of this was denying early calls to allow the MINC file format in addition to NIfTI; despite the technical superiority of MINC, this would have made supporting the standard by any software tool much harder and diminished its adoption. Achieving this outcome required deft leadership by Chris Gorgolewski, which relied heavily upon his trusted position in the community.

#### Stumbling points for the BIDS project

A number of challenges have arisen during the development of BIDS, which developers of other standards can also possibly learn from.

**Delayed adoption by large databanks.** With the exception of OpenNeuro, BIDS was not adopted by any large databases or consortia during its early development. Gaining the support of a large prospective data sharing project (such as UK Biobank or Human Connectome Project) at launch would have made it stronger and helped promote it. Instead, researchers have created “BIDS-ified” versions of these datasets, such as the ABCD-BIDS Community Collection (Feczko et al., 2021). A major challenge with a community-driven project like BIDS is the relatively slow pace of development (due to the need for community input and convergence), which can conflict with the desire of large projects to solve their problems quickly without constraints from outside. It would be useful in the future for funding agencies to require large data-generation projects to employ community standards in order to enhance the FAIRness of the resulting datasets.

---

<sup>15</sup> Full list of funding is available at <https://bids.neuroimaging.io/acknowledgments.html>

**Challenges of BIDS conversion.** Conversion of new datasets into the BIDS format has been, and often remains, a challenge for many researchers, which has limited even wider adoption of the standard. Converters have proliferated<sup>16</sup>, covering more use cases but also potentially increasing the burden on new users to choose which conversion tool to adopt. Earlier concerted efforts on developing more easily usable conversion tools could have increased early adoption.

**Lack of a machine-readable standard.** For many years, there was no common machine-readable instantiation of the standard. This led to misalignment between the standard, the validator, and the documentation, and made the implementation of changes in the standard difficult and time-consuming. This has been addressed by the schematization of the standard, but earlier development of a machine-readable standard may have improved the development workflow.

**Challenges of BEP management.** While BIDS Extension Proposals (BEPs) will remain the main driver of future BIDS development, they bring unique challenges for community management. Some BEPs may see development stagnate when faced with low developer availability or as project requirements are clarified. In the best case, this relative dormancy marks a clear growth point for the project, as developers realize that more work is necessary to clarify the BEP scope. In the worst case, BEPs may be abandoned by the proposing team; however, if this is telegraphed appropriately, other community members can choose to step in. Other more conceptual challenges include ensuring the same level of community consultation across BEPs. BEP leads who have not previously participated in BIDS development may not be familiar with its governance process. This is particularly concerning if BEPs are preemptively yoked to traditional incentives, with BEP leads committing to firm timelines for associated publications, grant deliverables, or graduate degree progress. Ensuring that BEPs maintain the same level of community consultation while sustaining the engagement of domain researchers motivated to develop BEPs is an ongoing challenge for BEP management.

## Conclusions

Few of us would have envisioned in 2015 that BIDS would be as successful as it has been, or that it would weather the departure of its initial founder so robustly. This success is a testament to the sustained efforts of the large number of individuals who have contributed in many different ways to the community and surrounding ecosystem of tools and data. We hope that BIDS can serve as a demonstration that communities of researchers can effectively develop standards that are essential to enable the effective and FAIR sharing of research data.

## Acknowledgments

Development of the BIDS Standard has been supported by the International Neuroinformatics Coordinating Facility, Laura and John Arnold Foundation, National Institutes of Health (R24MH114705, R24MH117179, R01MH126699, R24MH117295, P41EB019936, ZIAMH002977, R01MH109682, RF1MH126700), National Science Foundation (OAC-1760950,

---

<sup>16</sup> A current list of BIDS converters is available at <https://bids.neuroimaging.io/benefits#converters>

BCS-1734853, CRCNS-1429999, CRCNS-1912266), Novo Nordisk Fonden (NNF20OC0063277), French National Research Agency (ANR-19-DATA-0023, ANR 19-DATA-0021), Digital Europe TEF-Health (101100700), EU H2020 Virtual Brain Cloud (826421), Human Brain Project (SGA2 785907, SGA3 945539), European Research Council (Consolidator 683049), German Research Foundation (SFB 1436/425899996), SFB 1315/327654276, SFB 936/178316478, SFB-TRR 295/424778381), SPP Computational Connectomics (RI 2073/6-1, RI 2073/10-2, RI 2073/9-1), European Innovation Council PHRASE Horizon (101058240), Berlin Institute of Health & Foundation Charité, Johanna Quandt Excellence Initiative, ERAPerMed Pattern-Cog, and the Virtual Research Environment at the Charité Berlin – a node of EBRAINS Health Data Cloud.

### **Declaration of Competing Interests**

GR is an employee of Invicro. No other authors have competing interests to declare.

### **Use of generative AI language tools**

Generative AI tools were not used in the generation of the manuscript text, but were used in the formatting of the author affiliations and generation of code for the figures.

### **Data and Code Availability**

All data and code used to generate the figures in this paper are available at <https://osf.io/x7fh8/> under a permissive open source license.

### **Contributor Roles Taxonomy (CRediT) statement**

Authorship was offered to any individual who has contributed to the development of BIDS. Authors were asked to denote their contributions to the project according to the Contributor Roles Taxonomy (CRediT). Role abbreviations. C: Conceptualization, DC: Data curation for initial use and later re-use.), FA: Funding acquisition, M: Methodology, PA: Project administration, R: Resources, S: Software, SU: Supervision, V: Validation, W: Writing – original draft, RE: Writing – review & editing.

Russell A. Poldrack: C,FA,M,PA,R,SU,W. Christopher J. Markiewicz: C,M,PA,S,SU,V,W. Stefan Appelhoff: C,DC,M,PA,R,S,SU,V,W. Yoni K. Ashar: S,V,RE. Tibor Auer: C,DC,M,S,V,RE. Sylvain Baillet: C,DC,FA,PA,R,S,SU,RE. Shashank Bansal: S,V,RE. Leandro Beltrachini: M,RE. Christian G. Bénar: SU,RE. Giacomo Bertazzoli: C,M,RE. Suyash Bhogawar: C,DC,M,R,S,SU,V,RE. Ross W. Blair: C,PA,S,RE. Marta Bortoletto: C,RE. Mathieu Boudreau: C,M,SU,RE. Teon L. Brooks: M,S,RE. Vince D. Calhoun: FA,M,S,RE. Filippo Maria Castelli: DC,S,RE. Patricia Clement: M,PA,R,SU,RE. Alexander L. Cohen: S,V,RE. Julien Cohen-Adad: M,S,SU,RE. Sasha D'Ambrosio: DC,R,S,RE. Gilles De Hollander: C,DC,PA,R,SU,RE. Alejandro De La Vega: PA,S,SU,V,W. Arnaud Delorme: C,DC,M,R,S,W. Orrin Devinsky: FA,R,RE. Dejan Draschkow: PA,R,SU,RE. Eugene Paul Duff: C,M,RE. Elizabeth Dupre: PA,S,RE. Eric Earl:



C,DC,M,PA,S,RE. Oscar Esteban: C,DC,S,V,RE. Franklin W. Feingold: PA,RE. Guillaume Flandin: C,M,R,S,RE. Anthony Galassi: DC,S,SU,RE. Giuseppe Gallitto: C,M,RE. Melanie Ganz: C,FA,PA,R,SU,RE. James Gholam: C,M,RE. Satrajit S. Ghosh: C,DC,FA,R,S,SU,RE. Alessio Giacomel: C,DC,S,RE. Ashley G. Gillman: S,RE. Pdraig Gleeson: V,RE. Alexandre Gramfort: FA,R,S,V,RE. Samuel Guay: R,S,RE. Giacomo Guidali: M,RE. Yaroslav O. Halchenko: C,FA,R,S,RE. Daniel A. Handwerker: M,RE. Nell Hardcastle: DC,S,W. Peer Herholz: DC,S,RE. Dora Hermes: DC,M,R,RE. Christopher J. Honey: R,RE. Robert B. Innis: DC,FA,PA,R,RE. Horea-loan Ioanas: R,S,V,RE. Andrew Jahn: V,RE. Agah Karakuzu: C,M,PA,R,RE. David B. Keator: S,RE. Gregory Kiar: C,M,S,RE. Balint Kincses: C,RE. Angela R. Laird: DC,R,W. Jonathan C. Lau: C,V,RE. Alberto Lazari: DC,RE. Jon Haitz Legarreta: S,RE. Adam Li: DC,S,V,RE. Xiangrui Li: S,RE. Bradley C. Love: M,RE. Hanzhang Lu: M,R,S,RE. Camille Maumet: C,FA,M,SU,RE. Giacomo Mazzamuto: C,DC,S,RE. Steven L. Meisler: S,RE. Mark Mikkelsen: C,M,RE. Henk Mutsaerts: C,M,PA,SU,RE. Thomas E. Nichols: C,FA,M,R,SU,RE. Aki Nikolaidis: R,RE. Gustav Nilsonne: PA,RE. Guiomar Niso: C,DC,M,PA,R,S,SU,W. Martin Norgaard: C,DC,M,PA,R,S,SU,RE. Thomas W. O'kell: M,R,S,RE. Robert Oostenveld: C,DC,M,PA,R,S,SU,RE. Eduard Ort: S,RE. Patrick J. Park: C,R,S,RE. Mateusz Pawlik: S,RE. Cyril R. Pernet: C,DC,M,PA,S,SU,W. Franco Pestilli: C,FA,M,PA,R,SU,RE. Jan Petr: C,DC,S,SU,V,RE. Christophe Phillips: C,PA,SU,RE. Jean-Baptiste Poline: C,M,R,SU,W. Luca Pollonini: PA,SU,RE. Pradeep Reddy Raamana: C,DC,S,RE. Petra Ritter: C,DC,FA,M,SU,V,RE. Gaia Rizzo: C,M,RE. Kay A. Robbins: C,DC,M,R,S,V,RE. Alexander P. Rockhill: R,S,V,RE. Christine Rogers: DC,S,RE. Ariel Rokem: C,DC,FA,PA,R,S,SU,RE. Chris Rorden: DC,M,S,V,RE. Alexandre Routier: C,RE. Jose Manuel Saborit-Torres: C,DC,FA,S,RE. Taylor Salo: R,S,W. Michael Schirner: C,DC,M,RE. Robert E. Smith: S,W. Tamas Spisak: C,M,RE. Julia Sprenger: C,S,RE. Nicole C. Swann: FA,SU,RE. Martin Szinte: R,S,RE. Sylvain Takerkart: C,FA,S,SU,RE. Bertrand Thirion: C,S,RE. Adam G. Thomas: FA,PA,R,S,SU,RE. Sajjad Torabian: C,M,S,V,RE. Gael Varoquaux: C,S,RE. Maria De La Iglesia Vaya: C,DC,FA,R,S,RE. Bradley Voytek: S,RE. Julius Welzel: C,R,RE. Martin Wilson: DC,S,SU,RE. Krzysztof J. Gorgolewski: C,DC,FA,M,PA,R,S,SU,V,W.

## References

Appelhoff, S., Sanderson, M., Brooks, T. L., van Vliet, M., Quentin, R., Holdgraf, C., Chaumon, M., Mikulan, E., Tavabi, K., Höchenberger, R., Welke, D., Brunner, C., Rockhill, A. P., Larson, E., Gramfort, A., & Jas, M. (2019). MNE-BIDS: Organizing electrophysiological data into the BIDS format and facilitating their analysis. *Journal of Open Source Software*, 4(44). doi: 10.21105/joss.01896

Botvinik-Nezer, R., Holzmeister, F., Camerer, C. F., Dreber, A., Huber, J., Johannesson, M.,

- Kirchler, M., Iwanir, R., Mumford, J. A., Adcock, R. A., Avesani, P., Baczkowski, B. M., Bajracharya, A., Bakst, L., Ball, S., Barilari, M., Bault, N., Beaton, D., Beitner, J., ... Schonberg, T. (2020). Variability in the analysis of a single neuroimaging dataset by many teams. *Nature*, *582*(7810), 84–88.
- Bourget, M.-H., Kametsky, L., Ghosh, S. S., Mazzamuto, G., Lazari, A., Markiewicz, C. J., Oostenveld, R., Niso, G., Halchenko, Y. O., Lipp, I., Takerkart, S., Toussaint, P.-J., Khan, A. R., Nilsonne, G., Castelli, F. M., BIDS Maintainers, & Cohen-Adad, J. (2022). Microscopy-BIDS: An Extension to the Brain Imaging Data Structure for Microscopy Data. *Frontiers in Neuroscience*, *16*, 871228.
- Casey, B. J., Cannonier, T., Conley, M. I., Cohen, A. O., Barch, D. M., Heitzeg, M. M., Soules, M. E., Teslovich, T., Dellarco, D. V., Garavan, H., Orr, C. A., Wager, T. D., Banich, M. T., Speer, N. K., Sutherland, M. T., Riedel, M. C., Dick, A. S., Bjork, J. M., Thomas, K. M., ... ABCD Imaging Acquisition Workgroup. (2018). The Adolescent Brain Cognitive Development (ABCD) study: Imaging acquisition across 21 sites. *Developmental Cognitive Neuroscience*, *32*, 43–54.
- Cieslak, M., Cook, P. A., He, X., Yeh, F.-C., Dhollander, T., Adebimpe, A., Aguirre, G. K., Bassett, D. S., Betzel, R. F., Bourque, J., Cabral, L. M., Davatzikos, C., Detre, J. A., Earl, E., Elliott, M. A., Fadnavis, S., Fair, D. A., Foran, W., Fotiadis, P., ... Satterthwaite, T. D. (2021). QSIprep: an integrative platform for preprocessing and reconstructing diffusion MRI data. *Nature Methods*, *18*(7), 775–778.
- Clement, P., Castellaro, M., Okell, T. W., Thomas, D. L., Vandemaele, P., Elgayar, S., Oliver-Taylor, A., Kirk, T., Woods, J. G., Vos, S. B., Kuijjer, J. P. A., Achten, E., van Osch, M. J. P., BIDS maintainers, Detre, J. A., Lu, H., Alsop, D. C., Chappell, M. A., Hernandez-Garcia, L., ... Mutsaerts, H. J. M. M. (2022). ASL-BIDS, the brain imaging data structure extension for arterial spin labeling. *Scientific Data*, *9*(1), 543.
- Cox, R. W., Ashburner, J., Breman, H., & Fissell, K. (2004). *A (sort of) new image data format*

*standard: NIFTI-1*. 10th Annual Meeting of the Organization for Human Brain Mapping. Retrieved from [https://nifti.nih.gov/nifti-1/documentation/hbm\\_nifti\\_2004.pdf](https://nifti.nih.gov/nifti-1/documentation/hbm_nifti_2004.pdf)

Das, S., Zijdenbos, A. P., Harlap, J., Vins, D., & Evans, A. C. (2011). LORIS: a web-based data management system for multi-center studies. *Frontiers in Neuroinformatics*, 5, 37.

de la Vega, A., Rocca, R., Blair, R. W., Markiewicz, C. J., Mentch, J., Kent, J. D., Herholz, P., Ghosh, S. S., Poldrack, R. A., & Yarkoni, T. (2022). Neuroscout, a unified platform for generalizable and reproducible fMRI research. *eLife*, 11. doi: 10.7554/eLife.79277

Delorme, A., Truong, D., Youn, C., Sivagnanam, S., Stirm, C., Yoshimoto, K., Poldrack, R. A., Majumdar, A., & Makeig, S. (2022). NEMAR: an open access data, tools and compute resource operating on neuroelectromagnetic data. *Database: The Journal of Biological Databases and Curation*, 2022. doi: 10.1093/database/baac096

Esteban, O., Birman, D., Schaer, M., Koyejo, O. O., Poldrack, R. A., & Gorgolewski, K. J. (2017). MRIQC: Advancing the automatic prediction of image quality in MRI from unseen sites. *PloS One*, 12(9), e0184661.

Esteban, O., Blair, R. W., Nielson, D. M., Varada, J. C., Marrett, S., Thomas, A. G., Poldrack, R. A., & Gorgolewski, K. J. (2019). Crowdsourced MRI quality metrics and expert quality annotations for training of humans and machines. *Scientific Data*, 6(1), 30.

Esteban, O., Markiewicz, C. J., Blair, R. W., Moodie, C. A., Isik, A. I., Erramuzpe, A., Kent, J. D., Goncalves, M., DuPre, E., Snyder, M., Oya, H., Ghosh, S. S., Wright, J., Durnez, J., Poldrack, R. A., & Gorgolewski, K. J. (2019). fMRIPrep: a robust preprocessing pipeline for functional MRI. *Nature Methods*, 16(1), 111–116.

Feczko, E., Conan, G., Marek, S., Tervo-Clemmens, B., Cordova, M., Doyle, O., Earl, E., Perrone, A., Sturgeon, D., Klein, R., Harman, G., Kilamovich, D., Hermsillo, R., Miranda-Dominguez, O., Adebimpe, A., Bertolero, M., Cieslak, M., Covitz, S., Hendrickson, T., ... Fair, D. A. (2021). Adolescent Brain Cognitive Development (ABCD) Community MRI Collection and Utilities. In bioRxiv (p. 2021.07.09.451638). doi: 10.1101/2021.07.09.451638

- Freeman, J. (1972). THE TYRANNY OF STRUCTURELESSNESS. *Berkeley Journal of Sociology*, 17, 151–164.
- Gau, R., Flandin, G., Janke, A., tanguyduval, Oostenveld, R., Madan, C., Niso Galán, G., Szczepanik, M., Mutsaerts, H., Beliy, N., Norgaard, M., Pernet, C., & Chrisophe, P. (2022). *bids-matlab*. Zenodo. doi: 10.5281/ZENODO.5910584
- Gleeson, P., Crook, S., Turner, D., Mantel, K., Raunak, M., Willke, T., & Cohen, J. D. (2023). Integrating model development across computational neuroscience, cognitive science, and machine learning. *Neuron*, 111(10), 1526–1530.
- Gorgolewski, K. J., Alfaro-Almagro, F., Auer, T., Bellec, P., Capotă, M., Chakravarty, M. M., Churchill, N. W., Cohen, A. L., Craddock, R. C., Devenyi, G. A., Eklund, A., Esteban, O., Flandin, G., Ghosh, S. S., Guntupalli, J. S., Jenkinson, M., Keshavan, A., Kiar, G., Liem, F., ... Poldrack, R. A. (2017). BIDS apps: Improving ease of use, accessibility, and reproducibility of neuroimaging data analysis methods. *PLoS Computational Biology*, 13(3), e1005209.
- Gorgolewski, K. J., Auer, T., Calhoun, V. D., Craddock, R. C., Das, S., Duff, E. P., Flandin, G., Ghosh, S. S., Glatard, T., Halchenko, Y. O., Handwerker, D. A., Hanke, M., Keator, D., Li, X., Michael, Z., Maumet, C., Nichols, B. N., Nichols, T. E., Pellman, J., ... Poldrack, R. A. (2016). The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. *Scientific Data*, 3, 160044.
- Halchenko, Y., Meyer, K., Poldrack, B., & Solanky, D. (2021). DataLad: distributed system for joint management of code, data, and their relationship. *Journal of Open Source*. Retrieved from <https://par.nsf.gov/servlets/purl/10303143>
- Hanke, M., Halchenko, Y. O., Sederberg, P. B., Hanson, S. J., Haxby, J. V., & Pollmann, S. (2009). PyMVPA: A python toolbox for multivariate pattern analysis of fMRI data. *Neuroinformatics*, 7(1), 37–53.
- Hayashi, S., Caron, B., Heinsfeld, A. S., Vinci-Booher, S., McPherson, B. C., Bullock, D. N.,

- Berto, G., Niso, J. G., Hanekamp, S., Levitas, D., Kitchell, L., Leong, J., Silva, F. N., Koudoro, S., Willis, H., Jolly, J., Pisner, D., Zuidema, T., Kurzwaski, J., ... Pestilli, F. (2023). brainlife.io: A decentralized and open source cloud platform to support neuroscience research. *ArXiv*. doi: 10.1186/s13643-020-01332-7
- Holdgraf, C., Appelhoff, S., Bickel, S., Bouchard, K., D'Ambrosio, S., David, O., Devinsky, O., Dichter, B., Flinker, A., Foster, B. L., Gorgolewski, K. J., Groen, I., Groppe, D., Gunduz, A., Hamilton, L., Honey, C. J., Jas, M., Knight, R., Lachaux, J.-P., ... Hermes, D. (2019). iEEG-BIDS, extending the Brain Imaging Data Structure specification to human intracranial electrophysiology. *Scientific Data*, 6(1), 102.
- Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W., & Smith, S. M. (2012). FSL. *NeuroImage*, 62(2), 782–790.
- Karakuzu, A., Appelhoff, S., Auer, T., Boudreau, M., Feingold, F., Khan, A. R., Lazari, A., Markiewicz, C., Mulder, M., Phillips, C., Salo, T., Stikov, N., Whitaker, K., & de Hollander, G. (2022). qMRI-BIDS: An extension to the brain imaging data structure for quantitative magnetic resonance imaging data. *Scientific Data*, 9(1), 517.
- Li, X., Morgan, P. S., Ashburner, J., Smith, J., & Rorden, C. (2016). The first step for neuroimaging data analysis: DICOM to NIfTI conversion. *Journal of Neuroscience Methods*, 264, 47–56.
- Markiewicz, C. J., Gorgolewski, K. J., Feingold, F., Blair, R., Halchenko, Y. O., Miller, E., Hardcastle, N., Wexler, J., Esteban, O., Goncavles, M., Jwa, A., & Poldrack, R. (2021). The OpenNeuro resource for sharing of neuroscience data. *eLife*, 10. doi: 10.7554/eLife.71774
- Maumet, C., Auer, T., Bowring, A., Chen, G., Das, S., Flandin, G., Ghosh, S., Glatard, T., Gorgolewski, K. J., Helmer, K. G., Jenkinson, M., Keator, D. B., Nichols, B. N., Poline, J.-B., Reynolds, R., Sochat, V., Turner, J., & Nichols, T. E. (2016). Sharing brain mapping statistical results with the neuroimaging data model. *Scientific Data*, 3, 160102.
- Mennes, M., Biswal, B. B., Castellanos, F. X., & Milham, M. P. (2013). Making data sharing

work: the FCP/INDI experience. *NeuroImage*, 82, 683–691.

Moore, J., Basurto-Lozada, D., Besson, S., Bogovic, J., Bragantini, J., Brown, E. M., Burel, J.-M., Casas Moreno, X., de Medeiros, G., Diel, E. E., Gault, D., Ghosh, S. S., Gold, I., Halchenko, Y. O., Hartley, M., Horsfall, D., Keller, M. S., Kittisopikul, M., Kovacs, G., ... Swedlow, J. R. (2023). OME-Zarr: a cloud-optimized bioimaging file format with international community support. *bioRxiv : The Preprint Server for Biology*. doi: 10.1101/2023.02.17.528834

Moreau, C. A., Jean-Louis, M., Blair, R., Markiewicz, C. J., Turner, J. A., Calhoun, V. D., Nichols, T. E., & Pernet, C. R. (2020). The genetics-BIDS extension: Easing the search for genetic data associated with human brain imaging. *GigaScience*, 9(10). doi: 10.1093/gigascience/giaa104

Niso, G., Gorgolewski, K. J., Bock, E., Brooks, T. L., Flandin, G., Gramfort, A., Henson, R. N., Jas, M., Litvak, V., T Moreau, J., Oostenveld, R., Schoffelen, J.-M., Tadel, F., Wexler, J., & Baillet, S. (2018). MEG-BIDS, the brain imaging data structure extended to magnetoencephalography. *Scientific Data*, 5, 180110.

Niso, G., Rogers, C., Moreau, J. T., Chen, L.-Y., Madjar, C., Das, S., Bock, E., Tadel, F., Evans, A. C., Jolicoeur, P., & Baillet, S. (2016). OMEGA: The Open MEG Archive. *NeuroImage*, 124(Pt B), 1182–1187.

Norgaard, M., Matheson, G. J., Hansen, H. D., Thomas, A., Searle, G., Rizzo, G., Veronese, M., Giacomel, A., Yaqub, M., Tonietto, M., Funck, T., Gillman, A., Boniface, H., Routier, A., Dalenberg, J. R., Betthausen, T., Feingold, F., Markiewicz, C. J., Gorgolewski, K. J., ... Ganz, M. (2022). PET-BIDS, an extension to the brain imaging data structure for positron emission tomography. *Scientific Data*, 9(1), 65.

Pernet, C. R., Appelhoff, S., Gorgolewski, K. J., Flandin, G., Phillips, C., Delorme, A., & Oostenveld, R. (2019). EEG-BIDS, an extension to the brain imaging data structure for electroencephalography. *Scientific Data*, 6(1), 103.

- Poldrack, R. A., Barch, D. M., Mitchell, J. P., Wager, T. D., Wagner, A. D., Devlin, J. T., Cumba, C., Koyejo, O., & Milham, M. P. (2013). Toward open sharing of task-based fMRI data: the OpenfMRI project. *Frontiers in Neuroinformatics*, 7, 12.
- Poldrack, R. A., Feingold, F., Frank, M. J., Gleeson, P., de Hollander, G., Huys, Q. J., Love, B. C., Markiewicz, C. J., Moran, R., Ritter, P., Rogers, T. T., Turner, B. M., Yarkoni, T., Zhan, M., & Cohen, J. D. (2019). The importance of standards for sharing of computational models and data. *Computational Brain & Behavior*, 2(3-4), 229–232.
- Robbins, K., Truong, D., Jones, A., Callanan, I., & Makeig, S. (2022). Building FAIR Functionality: Annotating Events in Time Series Data Using Hierarchical Event Descriptors (HED). *Neuroinformatics*, 20(2), 463–481.
- Robb, R. A., Hanson, D. P., Karwoski, R. A., Larson, A. G., Workman, E. L., & Stacy, M. C. (1989). Analyze: a comprehensive, operator-interactive software package for multidimensional medical image display and analysis. *Computerized Medical Imaging and Graphics: The Official Journal of the Computerized Medical Imaging Society*, 13(6), 433–454.
- Rogers, C., Wickenheiser, A., Fesselier, L., Callegaro, J., Madjar, C., Das, S., & Evans, A. C. (2022). *EEG2BIDS Wizard: a flexible open-source cross-platform tool for anonymization and BIDS standardization*. 14th Annual Meeting of the Organization for Human Brain Mapping.
- Sanz Leon, P., Knock, S. A., Woodman, M. M., Domide, L., Mersmann, J., McIntosh, A. R., & Jirsa, V. (2013). The Virtual Brain: a simulator of primate brain network dynamics. *Frontiers in Neuroinformatics*, 7, 10.
- Schirner, M., Domide, L., Perdakis, D., Triebkorn, P., Stefanovski, L., Pai, R., Prodan, P., Valean, B., Palmer, J., Langford, C., Blickensdörfer, A., van der Vlag, M., Diaz-Pier, S., Peyser, A., Klijjn, W., Pleiter, D., Nahm, A., Schmid, O., Woodman, M., ... Ritter, P. (2022). Brain simulation as a cloud service: The Virtual Brain on EBRAINS. *NeuroImage*, 251, 118973.

- Sherif, T., Rioux, P., Rousseau, M.-E., Kassis, N., Beck, N., Adalat, R., Das, S., Glatard, T., & Evans, A. C. (2014). CBRAIN: a web-based, distributed computing platform for collaborative neuroimaging research. *Frontiers in Neuroinformatics*, *8*, 54.
- Teeters, J. L., Godfrey, K., Young, R., Dang, C., Friedsam, C., Wark, B., Asari, H., Peron, S., Li, N., Peyrache, A., Denisov, G., Siegle, J. H., Olsen, S. R., Martin, C., Chun, M., Tripathy, S., Blanche, T. J., Harris, K., Buzsáki, G., ... Sommer, F. T. (2015). Neurodata Without Borders: Creating a Common Data Format for Neurophysiology. *Neuron*, *88*(4), 629–634.
- Torabian, S., Vélez, N., Sochat, V., Halchenko, Y. O., & Grossman, E. D. (2023). The PyMVPA BIDS-App: a robust multivariate pattern analysis pipeline for fMRI data. *Frontiers in Neuroscience*, *17*. doi: 10.3389/fnins.2023.1233416
- Van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E. J., Yacoub, E., Ugurbil, K., & WU-Minn HCP Consortium. (2013). The WU-Minn Human Connectome Project: an overview. *NeuroImage*, *80*, 62–79.
- Van Horn, J. D., & Gazzaniga, M. S. (2013). Why share data? Lessons learned from the fMRIDC. *NeuroImage*, *82*, 677–682.
- Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., ... Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, *3*, 160018.
- Yarkoni, T., Markiewicz, C. J., de la Vega, A., Gorgolewski, K. J., Salo, T., Halchenko, Y. O., McNamara, Q., DeStasio, K., Poline, J.-B., Petrov, D., Hayot-Sasson, V., Nielson, D. M., Carlin, J., Kiar, G., Whitaker, K., DuPre, E., Wagner, A., Tirrell, L. S., Jas, M., ... Blair, R. (2019). PyBIDS: Python tools for BIDS datasets. *Journal of Open Source Software*, *4*(40). doi: 10.21105/joss.01294



