TemplateFlow: FAIR-sharing of multi-scale, multi-species brain models

Rastko Ciric[1,2*], William H. Thompson[3], Romy Lorenz[1,4,5], Mathias Goncalves[1], Eilidh MacNicol[6], Christopher J. Markiewicz[1], Yaroslav O. Halchenko[7], Satrajit S. Ghosh[8,9], Krzysztof J. Gorgolewski[1], Russell A. Poldrack[1], Oscar Esteban[10*]

***For correspondence:**
rastko@stanford.edu (RC);
phd@oscaresteban.es (OE)

[1]Department of Psychology, Stanford University, CA, USA; [2]Department of Bioengineering, Stanford University, CA, USA; [3]Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden; [4]MRC CBU, University of Cambridge, Cambridge, UK; [5]Department of Neurophysics, MPI, Leipzig, Germany; [6]Department of Neuroimaging, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK; [7]Department of Psychological and Brain Sciences, Dartmouth College, Hanover, NH, USA; [8]McGovern Institute for Brain Research, Massachusetts Institute of Technology: MIT, Cambridge, MA, USA; [9]Department of Otolaryngology, Harvard Medical School, Boston, MA, USA; [10]Department of Radiology, University Hospital of Lausanne and University of Lausanne, Lausanne, Switzerland

Abstract Reference anatomies of the brain and corresponding atlases play a central role in experimental neuroimaging workflows and are the foundation for reporting standardized results. The choice of such references —i.e., templates— and atlases is one relevant source of methodological variability across studies, which has recently been brought to attention as an important challenge to reproducibility in neuroscience. *TemplateFlow* is a publicly available framework for human and nonhuman brain models. The framework combines an open database with software for access, management, and vetting, allowing scientists to distribute their resources under FAIR —findable, accessible, interoperable, reusable— principles. *TemplateFlow* supports a multifaceted insight into brains across species, and enables multiverse analyses testing whether results generalize across standard references, scales, and in the long term, species, thereby contributing to increasing the reliability of neuroimaging results.

Brains are morphologically variable, exhibiting diversity in such features as overall size (***Lüders et al., 2002***), sulcal curvature (***Tosun et al., 2015***), and functional topology (***Tavor et al., 2016***; ***Mars et al., 2018***). Morphological variability manifests not only in differences between brains but also in the way that a brain changes across its lifespan, as it is remodelled by development, aging, and degenerative processes (***Courchesne et al., 2000***; ***Good et al., 2001***; ***Sowell et al., 2003***). These morphological differences often correspond with the effects of interest in neuroimaging studies and hinder direct spatial comparisons between brain maps (***Brett et al., 2002***). The substantial variability within and between individual brains necessitates a means of formalizing population-level knowledge about brain anatomy and function. Neuroscientists have answered this need by creating brain atlases as references for understanding and contextualizing morphological variability. Atlases map landmarks, features, and other knowledge about the brain as annotations that are consistent across individual brains.

The development of atlases in neuroscience has accelerated knowledge discovery and dissemination. Early endeavors, epitomized by the groundbreaking work of ***Brodmann*** (***2006***, originally

published in German in 1909) and complemented by ***Von Economo and Koskinas*** (***2008***, originally published in German in 1925), leveraged careful scrutiny of microanatomy and cytoarchitectonic properties in small numbers of brains. Concurrent macroanatomical approaches, by contrast, identified common features in nuclear boundaries and cortical gyrification. Modern atlases advanced on these approaches by incorporating stereotaxy, defining a basis set of coordinate axes over the brain and anchoring neural landmarks to coordinates. Talairach's assiduous postmortem examination of a single brain produced a stereotaxic atlas that saw wide use (***Talairach et al., 1957***). Stereotaxy was a fundamental feature to unfold surgical neuronavigation systems. ***Schurr and Merrington*** (***1978***) developed the first sterotaxic apparatus to surgically induce targeted brain lesions on cats. This early antecedent of neuronavigation informed early sectional atlases of the cat and macaque brains. Since then, neuroscientists have directed great efforts to improve existing (***Talairach and Tournoux, 1988***) and generate new atlases of the neurotypical adult human (***Landman et al., 2012***) and nonhuman (***Paxinos and Watson, 1997***; ***Martin and Bowden, 2000***) brain; as well as developing, aging, and neurologically atypical brains. For instance, new atlases and representative stereotaxic maps can be created for diseased (***Dickie et al., 2015***), infant (***Matsuzawa et al., 2001***; ***Fonov et al., 2011***; ***Shi et al., 2011***), and elderly (***Buckner et al., 2004***) human populations or to capture the rapid postnatal development of nonhuman species (***Calabrese et al., 2013***; ***Szulc et al., 2015***). Advancing beyond the volumetric constraints of stereotaxy, researchers of primate neocortex have also devised standard spaces based on geometric reconstructions of the cortical surface. This surface-based approach has the advantage of respecting the intrinsic topology of cortical folds, a development that has led to further improvements in spatial localization (***Coalson et al., 2018***). On account of its relatively high spatial resolution, its capacity to image the entire brain, and its non-invasive acquisition protocols, magnetic resonance imaging (MRI) has revolutionized neuroscience in general and the atlasing endeavor (***Evans et al., 2012***) in particular. In combination with software instruments' progress to map homologous features between subjects supported by regular grids (***Avants et al., 2008***) or reconstructed anatomical surfaces (***Robinson et al., 2014***), MRI has enabled researchers to create population-average maps of a particular image modality and/or particular sample with relative ease. These maps, called "*templates*", are typically created by averaging features across individuals that are representative of the population of interest to a study (***Dickie et al., 2017***). As a result, atlasing endeavours have been made contingent on templates, and have largely shifted away from the search for a single universal neuroanatomical pattern, instead making use of increasingly large samples with the aim of representing a population average of the distribution of morphological patterns.

Such resources as atlases and templates, which provide standardized prior knowledge, have become an indispensable component of modern neuroimaging data workflows for two cardinal reasons. First, group inference in neuroimaging studies requires that individuals' features are aligned into a common spatial frame of reference where their location can be called standard (***Brett et al., 2002***). Second, templates engender a stereotaxic coordinate system in which atlases can be delineated or projected. Associating atlases with template coordinates also facilitates the mapping of prior population-level knowledge about the brain into images of individual subjects' brains (for instance, to sample and average the functional MRI signal indexed by the regions defined in an atlas; ***Yeo et al.*** (***2011***)).

Because they are integral to analytic workflows, templates and atlases are frequently distributed as part of neuroimaging software libraries. For the most part, the developers of these libraries have substantial commitments apart from template aggregation and curation; thus, most libraries are practically limited in the subset of templates and atlases that they include. As an unfortunate consequence of this distribution model, access to and reuse of templates and atlases has become tightly coupled to a user's choice of software library. As an alternative to this software-bound distribution model, some laboratories and institutions maintain repositories where templates and atlases can be downloaded. The Montreal Neurological Institute (MNI) has spearheaded this mode of distribution and offers a large portfolio of human and non-human templates and atlases (***Evans***

et al., 1993; ***Mazziotta et al., 1995***; ***Holmes et al., 1998***; ***Collins et al., 1999***; ***Mazziotta et al., 2001***; ***Fonov et al., 2011***) accessible via a web site (***MNI, n.d.***). These templates and atlases have evolved iteratively (***Evans et al., 2012***), preserving spatial alignment to the "MNI Average Brain (305 MRI) Stereotaxic Registration Model" ("MNI305"; ***Evans et al., 1993***). As a consequence, it is common to find references to them in the literature under the umbrella term of "*MNI space*". Indeed, the default templates distributed with the popular FSL (***Jenkinson et al., 2012***) and SPM (***Friston et al., 2006***) software packages are generally referred to as being "*defined in MNI space*," even though the specific templates differ.

The limitations of the software-bound distribution model underscore three separate problems that arise in common practice. First, software-default templates are not generalizable to many use cases. When the population targeted by a study substantially deviates from neurotypical human adults (e.g., infants, elderly, or nonhuman animals), using an inadequate reference such as the default *MNI space* offered by a software library can introduce so-called "template effects" that bias morphometric analyses and produce incorrect results (***Yoon et al., 2009***). There is not yet any standard distance function that can objectively determine whether a template choice is phenotypically proximal to the study's sample, and thus whether template effects will be relevant. For example, since most MNI templates are created with a sample of adults of European ancestry, a study involving East Asian adults might require a non-default template. Because of the relative scarcity of nonhuman imaging resources, exposure to template effects is even more pressing in the nonhuman context: e.g., is it appropriate to use a mouse template for the spatial standardization of rat images? Not only are nonhuman templates and atlases scarce, accommodation of such resources in popular software tools is generally limited. For instance AFNI (***Cox and Hyde, 1997***) includes a rat template that can be applied in some contexts, while SPM provides functionality only through third-party add-ons (e.g., ***Sawiak et al., 2009***). Second, deviating from software defaults places a knowledge burden on the user. Once the researcher has selected a reference standard space that is suitable for their study population, if their choice is not included by default with the software they plan to use, they must then locate and download the reference template or atlas and integrate it within their analytic pipeline. This kind of excursion from defaults is far from frictionless and will often require expertise in template spaces and pipeline informatics. The required expertise is greater still when a researcher is working with an under-represented population for which no suitable template currently exists. In this situation, researchers often develop and make available new templates and atlases based on their own data samples, afterward distributing the new data assets using institutional websites or data storage systems such as FigShare (RRID:SCR_004328) or Dryad (RRID:SCR_005910). The lack of a centralized index for such templates propagates a share of the knowledge burden to researchers who stand to benefit from reusing them. Users must instead be aware not only of the prior existence of a template, but also where to locate it and the methods required to access it. Finally, as illustrated by the case of *MNI space*, it is not always clear what template a study is using. Since the templates most often used in the literature are software library defaults, reporting of spatial standardization is generally implicit (e.g., ***Carp, 2012b***, for functional MRI studies). In addition, template and atlas curators do not generally mint universally unique identifiers (such as the Research Resource Identifier, RRID; ***Bandrowski and Martone, 2016***) to precisely report spatial standardization and analysis. Therefore, deviating from software defaults has some potential to endanger reproducibility of studies due to template/atlas accessibility (and continuity thereof through time) and the risk of misreporting. Additional concerns regarding the reproducibility of spatial standardization in research include unlicensed distribution and provenance tracking. Errors in template and atlas resources are not common, but have been reported (e.g., ***Rohlfing, 2013***; ***Halchenko, 2013***). Using version control for templates and atlases has traditionally been considered too onerous and requires an expertise that may exceed the resources of research teams.

Overall, current practices in management and stewardship of group-standardized data (templates, atlases, and associated resources) do not follow "Findability, Accessibility, Interoperability,

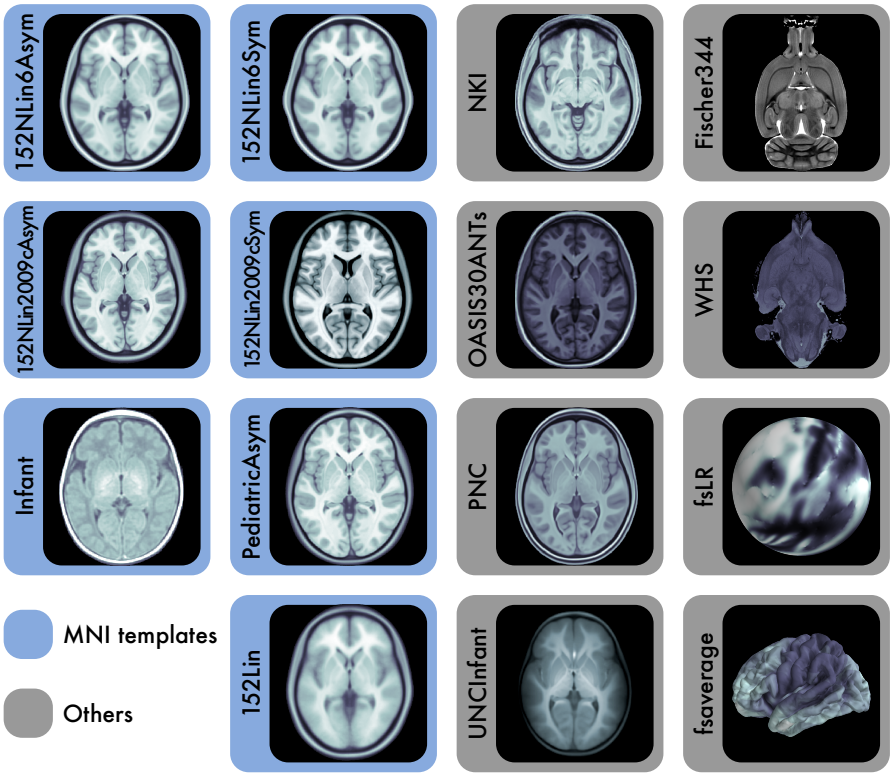


Figure 1. Representative views of 15 templates currently available in the *TemplateFlow Archive*. The 7 templates highlighted in blue are constituents of the Montreal Neurological Institute (MNI) portfolio. `WHS` (Waxholm space) and `Fischer344` correspond to rat templates. `fsaverage` and `fsLR` are surface templates; the remaining templates are volumetric. Each template is distributed with atlas labels, segmentations, and metadata files. The 15 templates displayed here are only a small fraction of those created as stereotaxic references for the neuroimaging community.

and Reusability (FAIR) Guiding Principles" (***Wilkinson et al., 2016***), making it difficult for researchers to locate and use these data assets and thereby reducing their long-term value. To address the need for a centralized resource for the archiving and redistribution of templates and atlases that allows programmatic access to human and nonhuman resources, we have developed *TemplateFlow*. This resource implements the FAIR Guiding Principles, effectively decoupling standardized spatial data from software libraries while maximizing the flexibility of emerging processing and analysis workflows (e.g. ***Esteban et al., 2017***, ***2019***). *TemplateFlow* comprises a cloud-based repository of human and nonhuman imaging templates —the "*TemplateFlow Archive*", Figure 1— paired with a Python-based library —the "*TemplateFlow Client*"— for programmatically accessing template resources. The resource is complemented with a "*TemplateFlow Manager*" tool to upload new or update existing resources. When adding a new template, the *Manager* initiates a peer-reviewed contribution pipeline where experts are invited to curate and vet new proposals. These software components, as well as all template resources, are version-controlled. Therefore, not only does *TemplateFlow* enable access to templates "off-the-shelf" by humans and machines, it also permits researchers to share their resources with the community. To implement several of the FAIR Principles, the *TemplateFlow Archive* features a tree-directory structure, metadata files, and data files following an organization inspired by the Brain Imaging Data Structure (BIDS; ***Gorgolewski et al., 2016***). BIDS is a widespread standard that balances the needs for human- and machine-readability. The online documentation hub and the resource browser located at TemplateFlow.org provide further details for users.

Table 1. Digital templates included in *TemplateFlow*. *TemplateFlow* is designed to maximise the discoverability and accessibility of new templates, minimise redundancies in template creation, and promote standardisation of processing workflows. To enhance visibility of existing templates, *TemplateFlow* includes a web-based browser indexing all files in the *TemplateFlow Archive* (templateflow.org/browse/).

Template ID	Description
`MNI152Lin`	Neurotypical adult human template created as the average from a linear mapping of 152 subjects from the MNI cohort of the ICBM registered to the earlier MNI305 template (***Mazziotta et al., 1995***, ***2001***).
`MNI152NLin6Asym`	FSL's version of the MNI152 neurotypical adult human template created using iterative nonlinear registration and averaging (***Evans et al., 2012***).
`MNI152NLin6Sym`	Symmetric version of `MNI152NLin6Asym` (***Evans et al., 2012***).
`MNI152NLin2009cAsym`	Update of the MNI152 neurotypical adult template with nonlinear registration. The mapping and averaging proceeded over 40 iterations beginning from the earlier MNI152 template (***Fonov et al., 2011***; ***Collins et al., 1999***).
`MNI152NLin2009cSym`	Symmetric version of `MNI152NLin2009cAsym` (***Fonov et al., 2011***; ***Collins et al., 1999***).
`MNIInfant`	Series of human infant templates created from 11 cohorts of infants and young children. Each cohort spans a different age range between 0 and 60 months (***Fonov et al., 2011***).
`MNIPediatricAsym`	Series of human pediatric templates created from 6 partially overlapping cohorts of children and young adults. Each cohort spans a different age range between 4.5 and 18.5 years (***Fonov et al., 2011***; ***Collins et al., 1999***).
`NKI`	Template created for the NKI-Rockland sample using ANTs diffeomorphic registration and averaging (***Avants et al., 2011***; ***Nooner et al., 2012***).
`OASIS30ANTs`	Template created using ANTs diffeomorphic registration and averaging for the Open Access Series of Imaging Studies (OASIS) (***Avants et al., 2011***; ***Marcus et al., 2007***).
`PNC`	Pediatric and young adult template created using ANTs diffeomorphic registration and averaging for the Philadelphia Neurodevelopmental Cohort (***Satterthwaite et al., 2016***).
`UNCInfant`	Series of human infant templates created from a 95-subject longitudinal sample comprising three scans: as neonates, as one-year-olds, and as two-year-olds (***Shi et al., 2011***).
`WHS`	Waxholm space template created as an atlas of the Sprague-Dawley rat brain (***Kjonigsen et al., 2015***; ***Osen et al., 2019***; ***Papp et al., 2014***, ***2015***).
`Fischer344`	Rat template created as the average of 41 four-month-old animals from the Fischer 344 strain (***Goerzen et al., 2020***).
`fsLR`	Surface-based Freesurfer template created for the Human Connectome Project (HCP) (***Van Essen et al., 2012***).
`fsaverage`	Surface-based average Freesurfer template (***Fischl et al., 1999***).

Results

Management and stewardship following the FAIR Guiding Principles

The specific and measurable FAIR principles of ***Wilkinson et al.*** (***2016***) are reproduced in Supplementary Box S1. Principles are indexed by their category and numbering, e.g., principle I3 – "*(meta)data include qualified references to other (meta)data*" belongs in the "to be Interoperable" category. We describe how *TemplateFlow* implements each of these specific principles in the following. Every template and all associated metadata, atlases, etc. are assigned a unique and persistent identifier (principle F1). BIDS prescribes a file naming scheme comprising a series of key-value pairs (called "entities") that are ordered hierarchically. Following BIDS' patterns, the template identifier is an alphanumeric label that is unique across the *Archive*, signified with the key `tpl-` (e.g., `tpl-MNI152Lin`). *TemplateFlow* therefore adapts BIDS to the specific domain of templates and atlases, affording the tool with a robust implementation of the principles I1-3 (interoperability). Table 1 enumerates several templates currently distributed with the *Archive*, and their corresponding unique identifiers. The unique identifier resolves the issue of inaccurate reporting, as it unambiguously designates one specific template. In addition, because the repository is versioned, researchers can easily retrieve and report the exact version of the template or atlas that was used in their study. Suppl. Table S2 summarizes the available entities and shows a segment of the file organization of the *Archive*. For each template, the *TemplateFlow* database includes reference volumetric template images (e.g., one T1-weighted and one T2-weighted average map; all must be in register), a set of atlas labels and voxelwise annotations defined with reference to the template image, and additional files containing the template and atlas metadata. Correspondingly, *TemplateFlow* allows surface-based resources such as average features, geometry files, annotations, or metadata.

Template resources are described with rich metadata (principles F2 and R1), ensuring that the data usage license is clear and accessible (R1.1), data and metadata are associated with detailed provenance (R1.2), and data and metadata follow a domain-relevant structure transferred from the neuroimaging community standards of BIDS (R1.3). Figure 2 summarizes the data types and metadata that can be stored in the *Archive*. Figure 3 provides an overview of the *Archive*'s metadata specification, showing that metadata clearly and explicitly include the identifier of the data they describe (F3). Data and metadata are retrievable using several open, free, standard communications protocols without need for authentication (A1) by using *DataLad* (***Halchenko et al., 2021***). Cloud storage for the *Archive* is supported by the Open Science Framework (osf.io) and Amazon's Simple Storage Service (S3). Version control, replication, and synchronisation of template resources across filesystems is managed with *DataLad*. Leveraging *DataLad*, metadata are stored on GitHub, ensuring accessibility to metadata even when corresponding data are no longer available (A2). *DataLad* is based on *Git* and *Git-Annex*, which index all data and metadata (F4). Although *DataLad* also provides searching tools that are applicable to *TemplateFlow*'s resources, the *TemplateFlow* framework provides a client tool that facilitates searching and querying.

An indexed resource, searchable with a *TemplateFlow "Client"*

TemplateFlow's Python client provides human users and software tools with reliable and programmatic access to the archive. The client can be integrated seamlessly into image processing workflows to handle requests for template resources on the fly. It features an intuitive application programming interface (API) that can query the *TemplateFlow Archive* for specific files (Figure 5). The BIDS-inspired organization enables easy integration of tools and infrastructure designed for BIDS (e.g., the Python client uses *PyBIDS* (***Yarkoni et al., 2019***) to implement queries like those listed in Suppl. Table S2). To query *TemplateFlow*, a user can submit a list of arguments corresponding to the BIDS-like key-value pairs in each entity's file name (e.g., `atlas="Schaefer2018"` to return files containing voxelwise annotations by ***Schaefer et al.*** (***2018***)).

To integrate template resources into neuroimaging workflows, traditional approaches required deploying an oftentimes voluminous tree of prepackaged data to the filesystem. By contrast, the *TemplateFlow* client implements lazy loading, which permits the base installation to be extremely

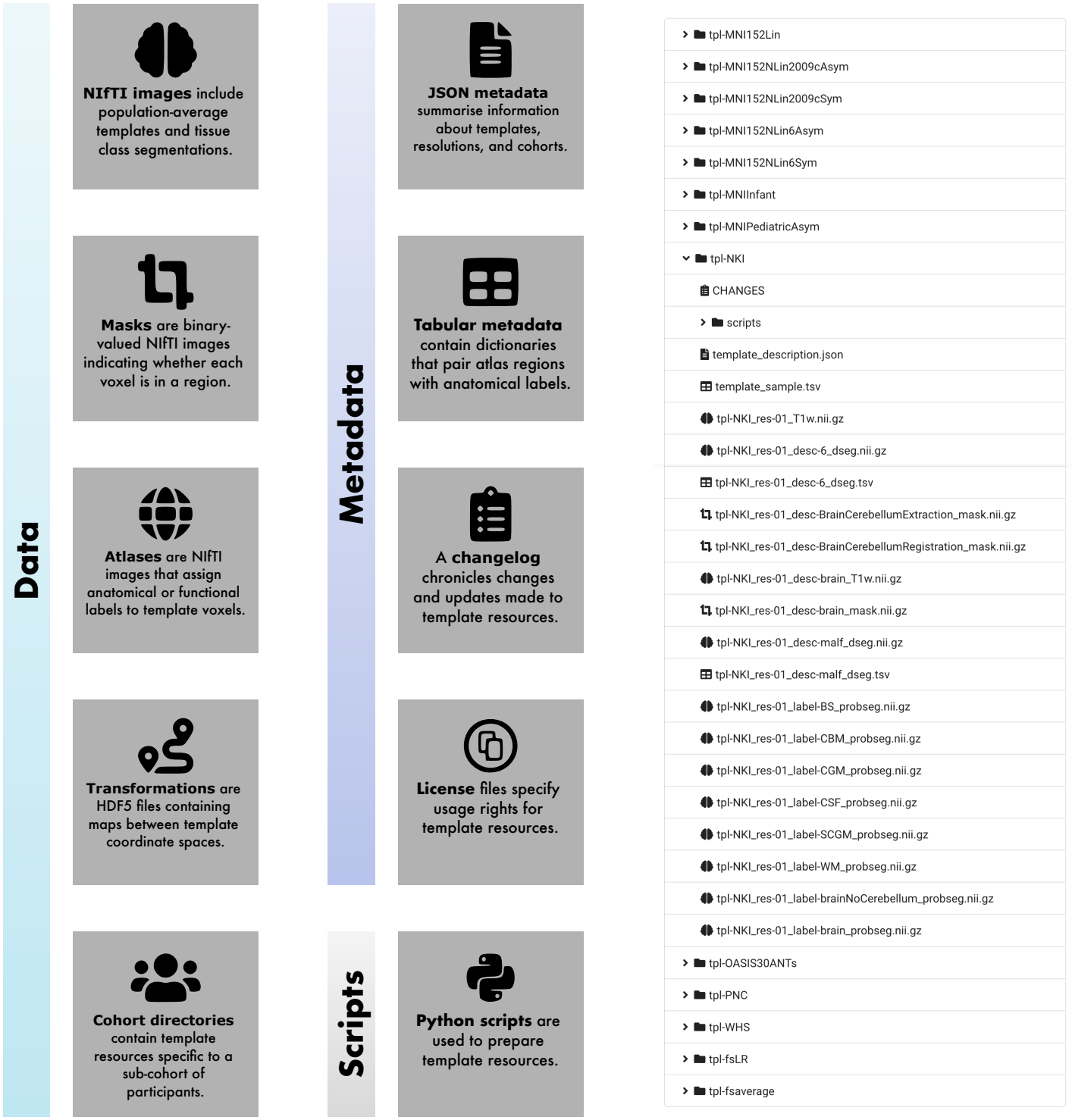


Figure 2. The *TemplateFlow Archive* contains template resources. Left, common file formats included in the *TemplateFlow Archive*. Right, view of the *TemplateFlow Archive*'s browser, accessible at TemplateFlow.org, with a single template resource directory expanded. Template data are archived using a BIDS-like directory structure, with top-level directories for each template. Each directory contains image files, annotations, and metadata for that template. Following BIDS specifications, volumetric data are stored in NIfTI-1 format. Further surface-based data types are supported with GIFTI (surfaces) and CIFTI-1 (mixed volumetric-and-surface data).

```
{
  "Authors": [
    "Fonov V",
    "Evans AC",
    "Botteron K",
    "Almli CR",
    "McKinstry RC",
    "Collins DL"
  ],
  "Curators": [
    "Esteban O"
  ],
  "Identifier": "MNIPediatricAsym",
  "License": "MIT-derived. See LICENSE file",
  "Name": "MNI's unbiased standard MRI template for
           pediatric data from the 4.5 to 18.5y age
           range",
  "RRID": "SCR_008796",
  "ReferencesAndLinks": [
    "https://doi.org/10.1016/j.neuroimage.2010.07.033",
    "https://doi.org/10.1016/S1053-8119(09)70884-5",
    "http://nist.mni.mcgill.ca/?p=974",
    "https://doi.org/10.1007/3-540-48714-X_16"
  ],
  "TemplateFlowVersion": "1.0.0",
  "cohort": {
    "1": {
      "age": [
        4.5,
        18.5
      ],
      "name": "whole age range",
      "units": "yr"
    },
    "2": {
      "age": [
        4.5,
        8.5
      ],
      "name": "prepuberty",
      "units": "yr"
    },
  . . .
  },
  "res": {
    "1": {
      "origin": [
        -98.0,
        -134.0,
        -72.0
      ],
      "shape": [
        197,
        233,
        189
      ],
      "zooms": [
        1.0,
        1.0,
        1.0
      ]
    },
  . . .
  }
}
```

General metadata

Field	Type	Description
Authors	Array (String)	Names of authors who created the template.
Curators	Array	Names of TemplateFlow curators who contributed or manage the dataset.
Identifier	String	Unique human-readable template identifier within TemplateFlow.
License	String	License under which template resources are available.
Name	String	Full descriptive name of the template.
RRID	String	Research Resource Identifier for the TemplateFlow dataset.
ReferencesAndLinks	Array	Publications to reference when using the template, and salient links for template information.
TemplateFlowVersion	String	Version of TemplateFlow under which the dataset was distributed.

Cohort metadata

Field	Type	Description
cohort	Object	Top-level field containing all cohort metadata objects.
1, 2, ...	Object	Cohort identifiers. Each has a subdirectory in the template data directory, and each has a metadata object nested in the `cohort` field.
age	Array (Number)	2-tuple array indicating the lower and upper bounds for participant age in the cohort, if the cohorts are stratified by age.
name	String	Full descriptive name of the cohort..
units	String	Units for cohort age bounds.

Resolution metadata

Field	Type	Description
res	Object	Top-level field containing all resolution metadata objects.
1, 2, ...	Object	Resolution identifiers. Each has a metadata object nested in the `res` field. The metadata for each resolution apply to all images whose name includes `res-<identifier>`. The identifier itself does not necessarily correspond to the voxel size.
origin	Array	(x, y, z) spatial location of the voxel origin relative to the physical origin in mm.
shape	Array	(x, y, z) shape of the image in voxels.
zooms	Array	(x, y, z) size of each voxel in mm.

Figure 3. Overview of the metadata specification of the *TemplateFlow Archive*. *TemplateFlow*'s metadata are formatted as JavaScript Object Notation (JSON) files located within each template set. An example `template_description.json` metadata file is displayed at the left for `MNIPediatricAsym`. In addition to general template metadata, datasets can contain cohort-level and resolution-level metadata, which are nested within the main metadata dictionary and apply only to subsets of images in the dataset.

lightweight. Instead of distributing neuroimaging data with the installation, *TemplateFlow* allows the user to dynamically pull from the cloud-based storage only those resources they need, as they need them. After a resource has been requested once, it remains cached in the filesystem for future utilization.

We demonstrate benefits of centralizing templates in general, and the validity of the *TemplateFlow* framework in particular, via its integration into *fMRIPrep* (***Esteban et al., 2019***), a functional MRI preprocessing tool. This integration provides *fMRIPrep* users with flexibility to spatially normalize their data to any template available in the *Archive* (see Box 1). This integration has also enabled the development of *fMRIPrep* adaptations, for instance to pediatric populations or rodent imaging (***MacNicol et al., 2021***), using suitable templates from the archive. The uniform interface provided by the BIDS-like directory organisation and metadata enables straightforward integration of new templates into workflows equipped to use *TemplateFlow* templates. Further examples of tools leveraging *TemplateFlow* include *MRIQC* (***Esteban et al., 2017***) for quality control of MRI; *PyNets* (***Pisner and Hammonds, 2020***), a package for ensemble learning of functional and structural connectomes; *ASLPrep* (***Adebimpe et al., 2021***), an ASL pre-processing pipeline that makes use of *TemplateFlow* through *sMRIPrep* —the spin-off structural pipeline from *fMRIPrep*; and *NetPlotBrain* (***Thompson and Fanton, 2021***), which uses *TemplateFlow* to display spatially standardized brain network data.

A framework for researchers who generate and share spatially-standardized data

A centralized repository for neuroimaging templates should also address the needs of template creators, enabling peer-reviewed integration of new templates with minimal informatic overhead. Inspired by the Conda-forge community repository and the Journal of Open Source Software, the GitHub-based "templateflow" organisation is a site for dialogue between members of the neuroimaging community and *TemplateFlow Archive* curators. GitHub issues offer any community member the ability to share their needs with developers and *Archive* curators, for instance by identifying templates or workflow features for potential inclusion in the project. "Pull requests" provide a means for members of the community to directly contribute code or template resources to the *TemplateFlow Archive*.

This peer-reviewed contribution process is facilitated through the Python-based *TemplateFlow Manager*. The *TemplateFlow Manager* automates the work of synchronizing data from a local directory to cloud storage in OSF. Furthermore, it creates a GitHub repository containing git-annex pointers that enable *DataLad* to download template data from cloud storage to any machine with a copy of the repository. Finally, it opens a new pull request to propose adding the newly contributed template repository into the main *TemplateFlow Archive* (Figure 6). Synchronization of spatial data assets to the *TemplateFlow Archive* affords data producers an immediate way to distribute their data according to FAIR principles and thereby increase its reach.

Unambiguous and precise reporting of spatially standardized processing and analysis.

To explore the coupling between software libraries and standard spaces, we conducted a topic modeling analysis (***Blei et al., 2003***) of *MNI* space in the neuroimaging literature. We identified a coupling between the reporting of spatial standardization and software libraries across 6,048 articles containing the term *MNI* and published in two leading domain-specific journals (*NeuroImage* and *NeuroImage: Clinical*). To demonstrate the heterogeneity in the reported standardization to *MNI space*, we sorted topics according to their dominance in articles (i.e., the topic with the highest model score in MNI-related sentences; Figure 4). Out of 15 topics we modeled, two of the most dominant topics contained software tool names as well as the names of related scientists. As shown in Figure 4, around 500 articles (each term) contained either "SPM" (9% of the documents) or "FSL" (8%). Interestingly, the two words do not ever appear together, suggesting that researchers stick with one or another in their analyses. Additional topics that seemingly relate to the provenance of templates and atlases —beyond the ubiquitous "Montreal", "Neurological" and

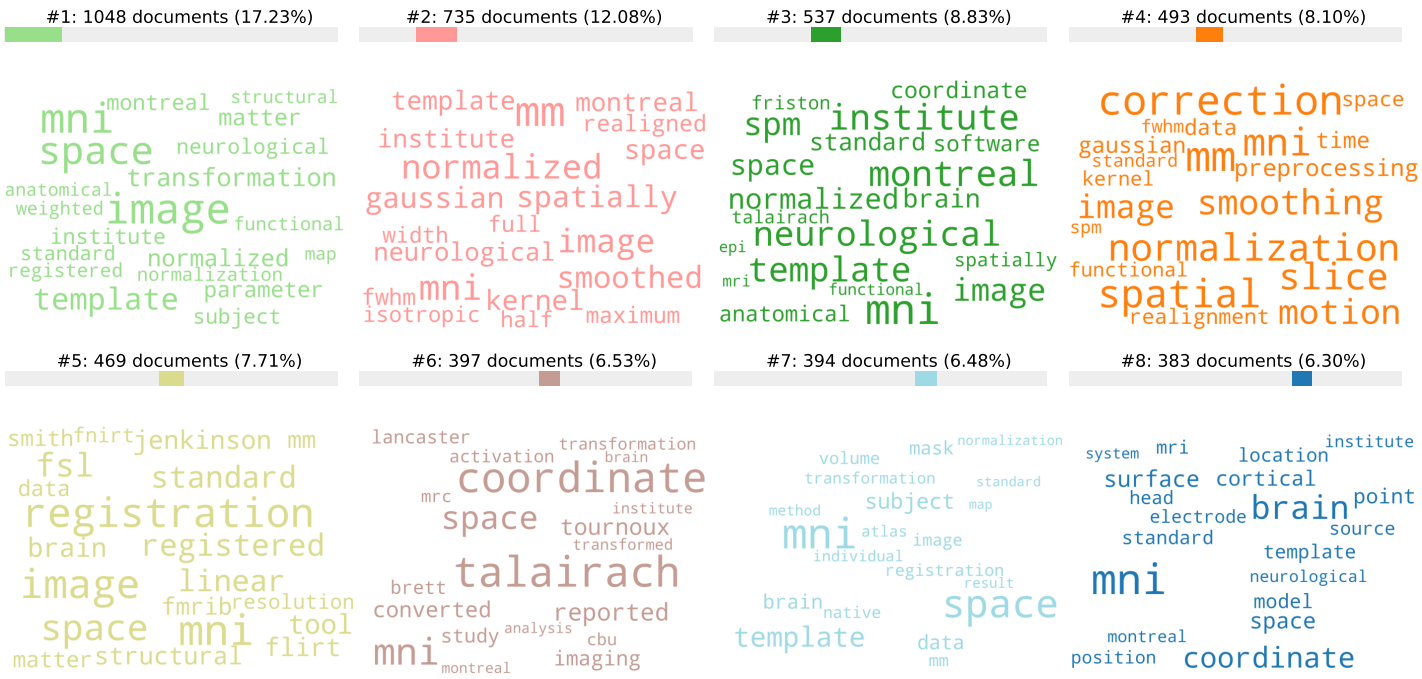


Figure 4. The FSL and SPM software tools associate with dominant topics of sentences including the term "MNI" across the literature. We performed topic modeling with latent Dirichlet allocation (LDA; ***Blei et al., 2003***) on text sentences extracted from 6,048 articles that contained the word "MNI". For each topic identified, the 20 words with the highest loadings on that topic are displayed in a word cloud with larger font size indicating higher loading of the word on the corresponding topic. Word clouds are sorted by descending topic dominance. Ranking and relative dominance are shown above each topic's cloud. Only the top 8 dominant topics are shown here (full model is reported in Suppl. Figure S1). Two top-dominant topics —#3 and #5— are associated with SPM and FSL respectively.

"Institute" for MNI— are those that ranked #4, #13, #14, which include "SPM" and other terms such as "McGill", "Wellcome", "UCL", or the "parametric" in SPM (see Suppl. Figure S1). The remainder of topics appears to relate to miscellaneous aspects of spatial standardization, such as "anatomical", "smoothness", "map/mapping", "standard", "normalization", "(re)align/alignment", etc., with no apparent relationship to the actual origin of the resource. These interpretations suggest that "*MNI space*" can refer to any of a family of templates and is not a unique identifier. As a matter of fact, studies carried out with SPM96 (***Friston et al., 2006***) and earlier versions report their results in *MNI space* with reference to the single-subject Colin 27 average template (***Holmes et al., 1998***). However, beginning with SPM99, SPM updated its definition of *MNI space* to the template that MNI released in 2001: an average of 152 subjects from the ICBM database, aligned by means of linear registration. In SPM12 (the latest release at the time of writing), the meaning of *MNI template* varies by submodule: different modules alternately use the Linear MNI152 template (***Mazziotta et al., 1995***) and a new, nonlinear revision from 2009 (***Fonov et al., 2011***). By contrast, the *MNI template* bundled with the FSL toolbox was developed by Dr. A. Janke in collaboration with MNI researchers (***Evans et al., 2012***). Although it was generated under the guidance of and using the techniques of the 2006 release of nonlinear MNI templates, this template is not in fact part of the official portfolio distributed by MNI. Nonetheless, our results suggest that the MNI templates bundled with SPM and FSL have historically gained broader currency as a result of the widespread use of these software libraries.

Discussion

The use of templates and atlases is ubiquitous in neuroimaging, and the emerging challenges regarding template use accordingly merit immediate attention. In an early perspective, Van Essen identified a set of desiderata for brain templates (***Van Essen, 2002***). Above and beyond anatomical

fidelity, he called for connecting templates in an aggregation of databases with "powerful and flexible options for searching, selecting, and visualizing data". Finally, he stressed the importance of resource accessibility. *TemplateFlow* provides a framework that satisfies all of the aforementioned desiderata while following the FAIR principles (***Wilkinson et al., 2016***). We elaborate that most of the issues concerning the reliability of neuroimaging research relating to standardized spatial references stem from the lack of a centralized repository designed to meet FAIR principles. We show how we effectively implement such principles with the adoption of a BIDS-like structure for the data and metadata in the resource, and with *DataLad* to support the core of the data management system. We complete the implementation with an easy-to-use client tool.

When researchers develop a new brain template or atlas for public dissemination, there exists no standard channel or format for distributing their work. With no central repository or uniform organizational scheme, template creators are often tasked with the responsibilities of maintaining template resources and managing access on an ad hoc basis and sustaining them over time with limited to no support. While the quality of peer reviewed template resources is assessed once prior to publication, reviewers often focus on perceived academic merit to the exclusion of FAIR principles. This can lead to poor resource adoption and low community value even for high-quality resources. Informal vetting of resources is prone to more clerical errors such as missing or corrupted files and/or metadata, or unlicensed distribution, which nonetheless may make the resource unusable. Conversely, users are confronted with a surfeit of available templates and atlases, many with unclear provenance, absent licensing terms, and the attendant challenges of accessing them and integrating them into workflows.

Without a uniform distribution format, integrating a template into software requires a custom solution for every new template, increasing the burden on developers. Consistently with the previous investigation by ***Carp*** (***2012b***) in the domain of functional MRI, our text mining exploration illustrates a strong coupling between software library and the templates and atlases of choice. Indeed, ***Carp*** (***2012b***) analysed 241 functional MRI studies of which 90.9% reported normalizing brain images to a common template. Of those, 79.0% indicated the target space used for spatial normalization. Few studies reported critical parameters, and only 50 specified the template image: 26.0% used "the MNI152 template", and 26.0% the "SPM library's echo-planar imaging template". Unfortunately, template selection is seen as a default parameter of the software library, which lends itself to assuming that the target normalization space is implicitly reported by identifying the software tools of choice. ***Bowring et al.*** (***2019***) wrote comparable pipelines in three software suites (AFNI, FSL and SPM) in order to identify challenges to reproducing published studies with openly shared raw data. When discussing differences among pipelines, they noted that, "while all packages are purportedly using the same MNI atlas space, an appreciable amount of activation detected by AFNI and FSL fell outside of SPM's analysis mask".

This coupling seems to also limit the utilization of templates other than those defaulted by the software. Custom templates (i.e., those not included as default option for the software tool) range from population-specific templates to ad hoc templates created by averaging images of the study at hand. In some settings, the use of default templates risks introducing "template effects" that confound the interpretation of results (such as those introduced in pediatric imaging studies by using an adult template, ***Yoon et al., 2009***). As the target population moves away from that with which a default template was created, "template effects" become more concerning and custom templates more necessary. The problem is exacerbated in the case of nonhuman imaging, as the scarcity (or absence) of specific templates available within software packages hinders already challenging translational endeavors. Further, the consistency across templates and atlases is reportedly low (***Bohland et al., 2009***), and although there has not been any programmatic comparison to understand the extent to which this inconsistency alters the spatial interpretation of results, it is reasonable that templates and atlases introduce a decision point and therefore are sources of some analytical variability.

One ostensible caveat regarding centralized and FAIR-principled knowledge repositories such

as *TemplateFlow* is that, by increasing the findability and reuse potential of data resources, they also open the door to increased methodological flexibility. ***Carp*** empirically investigated the consequences of such methodological flexibility in neuroimaging, demonstrating that decision points in workflows can lead to substantial variability in analysis outcomes. In a contemporaneous paper, ***Carp*** (***2012b***) contextualized these findings vis-à-vis the inflated risk of false positives, underscoring that analytical variability degrades the reproducibility of studies only *in combination with* (intended or unintended) *selective reporting* of methods and results. Selective reporting, in this particular application, would mean that a researcher explores the results with reference to several templates or atlases and reports only those that confirm the research's hypotheses. *TemplateFlow*'s standardization preempts the problem of *unintended* selective reporting: the provenance of all resources is tracked, all resources are accessible, and comprehensive metadata are generated.ß More recently, ***Botvinik-Nezer et al.*** (***2020***) advocated for another solution to the problem of analytical variability: "multiverse" analyses, wherein many combinations of methodological choices are all thoroughly reported and cross-compared when presenting results. Applied to the particular choice of template and atlas combinations, it would thus be desirable to report neuroimaging results with reference to several standard spaces and determine whether the interpretations hold across those references and atlases. *TemplateFlow*'s interoperability empowers users to incorporate this type of analysis into their research by easily making template or atlas substitutions for cross-comparison. For instance, Box 1 shows how *TemplateFlow* works with *fMRIPrep* to automate preprocessing of outputs in multiple standard spaces. This facilitates assessment of the robustness of a result with respect to the template or atlas of choice in accordance with the multiverse approach.

Limitations

TemplateFlow affords researchers substantial analytical flexibility in the choice of standard spaces of reference. Such flexibility helps researchers minimize "template effects" —by easily inserting the most adequate template— but also opens opportunities for incomplete reporting of experiments. Using *DataLad* or the *TemplateFlow Client*, researchers have at their disposal the necessary tooling for precise reporting: unique identifiers, provenance tracking, version tags, and comprehensive metadata. Therefore, the effectiveness of *TemplateFlow* to mitigate selective reporting is bounded by the user's discretion. Similarly, the resource is limited at the time of writing to MRI templates, but it readily supports such other modalities as nuclear imaging maps (e.g., PET/SPECT), and would support others with minimal adaptations (e.g., protein expression maps, or 3D reconstructions from histology).

As a research resource, the scope of this manuscript is limited to describing the framework and infrastructure of *TemplateFlow*, highlighting how neuroscientists can leverage this new data archive and the tooling around it. Therefore, some fundamental issues related to this work must be left for future investigation: (i) the overarching problems of cross-template and cross-atlas consistency; (ii) comparative evaluation of methodological alternatives for producing new templates, atlases and related data; (iii) providing neuroimagers with more objective means to determine the most appropriate template and atlas choices that apply to their research, as well as better understanding "template effects"; (iv) the adequacy of original (MRI, nuclear imaging, etc.) and derived (regularly gridded images, surfaces, etc.) modalities for a specific research application; or (v) the study of the validity and reliability of inter-template registration, as well as the evaluation of such a component of the *TemplateFlow* framework.

In more practical terms, *TemplateFlow* is limited to the redistribution of resources under permissive licenses, without any access restrictions.

Conclusion

We introduce an open framework for the archiving, maintenance and sharing of neuroimaging templates and atlases called *TemplateFlow* that is implemented under FAIR data sharing principles. We describe the current need for this resource in the domain of neuroimaging, and further discuss the implications of the increased analytical flexibility this tool affords. These two facets of repro-

ducibility —availability (under FAIR guiding principles) of prior knowledge required by the research workflow, and the analytical flexibility such availability affords— are ubiquitous concerns across disciplines. *TemplateFlow*'s approach to addressing both establishes a pattern broadly transferable beyond neuroimaging. We envision *TemplateFlow* as a core research tool undergirding multiverse analyses —assessing whether neuroimaging results are robust across population-wide spatial references— as well as a stepping stone towards the quest of mapping anatomy and function across species.

Acknowledgments

The development of this resource was supported by the Laura and John Arnold Foundation (RAP and KJG), the NIBIB (R01EB020740, SSG; 1P41EB019936-01A1 SSG, YOH), NIMH (RF1MH121867 RAP, OE; R24MH114705 and R24MH117179, RAP; 1RF1MH121885 SSG), NINDS (U01NS103780, RAP), and NSF (CRCNS 1912266, YOH). RL is funded by the Wellcome Trust (209139/Z/17/Z). EM was supported by the UK Medical Research Council (MR/N013700/1) and King's College London. OE acknowledges financial support from the SNSF Ambizione project "Uncovering the interplay of structure, function, and dynamics of brain connectivity using MRI" (grant number PZ00P2_185872).

Ethical compliance

We complied with all relevant ethical regulations. This resource reused publicly available data derived from studies acquired at many different institutions. Protocols for all of the original studies were approved by the corresponding ethical boards.

Code & data availability statement

All the software components discussed in this paper are available under the Apache 2.0 license, accessible as repositories of https://github.com/templateflow. All templates and associated data are available under corresponding open licenses and accessible as described in the manuscript.

Author contributions

Conceptualization: RC, CJM, KJG, RAP, OE Data curation: RC, EM, OE Topics analysis: RC, RL, OE Funding acquisition: KJG, RAP, OE Methodology: RC, OE Project administration: RAP, OE Resources: YOH, SSG, KJG, RAP, OE Software & Documentation: RC, WHT, MG, EM, CJM, YOH, OE Supervision: RAP, OE Validation: RC, WHT, EM, OE Visualization: RC, RL. Writing – original draft: RC & OE Writing – review & editing: RC, RL, WHT, MG, EM, CJM, YOH, SSG, KJG, RAP, OE.

References

Adebimpe A, Bertolero M, Dolui S, Cieslak M, Murtha K, Baller EB, Boeve B, Boxer A, Butler ER, Cook P, Colcombe S, Covitz S, Davatzikos C, Davila DG, Elliott MA, Flounders MW, Franco AR, Gur RE, Gur RC, Jaber B, et al. ASLPrep: A Generalizable Platform for Processing of Arterial Spin Labeled MRI and Quantification of Regional Brain Perfusion. bioRxiv. 2021 May; p. 2021.05.20.444998. doi:10.1101/2021.05.20.444998.

Avants BB, Epstein CL, Grossman M, Gee JC. Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. Med Image Anal. 2008; 12(1):26–41. doi:10.1016/j.media.2007.06.004.

Avants BB, Tustison NJ, Song G, Cook PA, Klein A, Gee JC. A reproducible evaluation of ANTs similarity metric performance in brain image registration. NeuroImage. 2011 Feb; 54(3):2033–44. doi:10.1016/j.neuroimage.2010.09.025.

Bandrowski AE, Martone ME. RRIDs: A Simple Step toward Improving Reproducibility through Rigor and Transparency of Experimental Methods. Neuron. 2016 May; 90(3):434–436. doi:10.1016/j.neuron.2016.04.030.

Blei DM, Ng AY, Jordan MI. Latent Dirichlet Allocation. J Mach Learn Res. 2003; 3(Jan):993–1022. https://jmlr.org/papers/v3/blei03a.html.

Bohland JW, Bokil H, Allen CB, Mitra PP. The Brain Atlas Concordance Problem: Quantitative Comparison of Anatomical Parcellations. PLoS One. 2009 Sep; 4(9):e7200. doi:10.1371/journal.pone.0007200.

Botvinik-Nezer R, Holzmeister F, Camerer CF, Dreber A, Huber J, Johannesson M, Kirchler M, Iwanir R, Mumford JA, Adcock RA, Avesani P, Baczkowski BM, Bajracharya A, Bakst L, Ball S, Barilari M, Bault N, Beaton D, Beitner J, Benoit RG, et al. Variability in the analysis of a single neuroimaging dataset by many teams. Nature. 2020 Jun; 582(7810):84–88. doi:10.1038/s41586-020-2314-9.

Bowring A, Maumet C, Nichols TE. Exploring the impact of analysis software on task fMRI results. Human Brain Mapping. 2019; 40(11):3362–3384. doi:https://doi.org/10.1002/hbm.24603.

Brett M, Johnsrude IS, Owen AM. The problem of functional localization in the human brain. Nat Rev Neurosci. 2002; 3:243–249. doi:10.1038/nrn756.

Brodmann K. Brodmann's: Localisation in the Cerebral Cortex. Springer US; 2006. doi:10.1007/b138298.

Buckner RL, Head D, Parker J, Fotenos AF, Marcus D, Morris JC, Snyder AZ. A unified approach for morphometric and functional data analysis in young, old, and demented adults using automated atlas-based head size normalization: reliability and validation against manual measurement of total intracranial volume. NeuroImage. 2004 Oct; 23(2):724–738. doi:10.1016/j.neuroimage.2004.06.018.

Calabrese E, Badea A, Watson C, Johnson GA. A quantitative magnetic resonance histology atlas of postnatal rat brain development with regional estimates of growth and variability. NeuroImage. 2013 May; 71:196–206. doi:10.1016/j.neuroimage.2013.01.017.

Carp J. On the Plurality of (Methodological) Worlds: Estimating the Analytic Flexibility of fMRI Experiments. Front Neurosci. 2012; 6. doi:10.3389/fnins.2012.00149.

Carp J. The secret lives of experiments: Methods reporting in the fMRI literature. NeuroImage. 2012 Oct; 63(1):289–300. doi:10.1016/j.neuroimage.2012.07.004.

Coalson TS, Van Essen DC, Glasser MF. The impact of traditional neuroimaging methods on the spatial localization of cortical areas. Proceedings of the National Academy of Sciences. 2018; 115(27):E6356–E6365. doi:10.1073/pnas.1801582115.

Collins DL, Zijdenbos AP, Baaré WFC, Evans AC. ANIMAL+INSECT: Improved Cortical Structure Segmentation. In: Kuba A, Šáamal M, Todd-Pokropek A, editors. *Information Processing in Medical Imaging* Lecture Notes in Computer Science, Berlin, Heidelberg: Springer; 1999. p. 210–223. doi:10.1007/3-540-48714-X_16.

Courchesne E, Chisum HJ, Townsend J, Cowles A, Covington J, Egaas B, Harwood M, Hinds S, Press GA. Normal Brain Development and Aging: Quantitative Analysis at in Vivo MR Imaging in Healthy Volunteers. Radiology. 2000 Sep; 216(3):672–682. doi:10.1148/radiology.216.3.r00au37672.

Cox RW, Hyde JS. Software tools for analysis and visualization of fMRI data. NMR Biomed. 1997; 10(4-5):171–178. doi:10.1002/(SICI)1099-1492(199706/08)10:4/5<171::AID-NBM453>3.0.CO;2-L.

Dickie DA, Job DE, Gonzalez DR, Shenkin SD, Wardlaw JM. Use of Brain MRI Atlases to Determine Boundaries of Age-Related Pathology: The Importance of Statistical Method. PLoS One. 2015 May; 10(5):e0127939. doi:10.1371/journal.pone.0127939.

Dickie DA, Shenkin SD, Anblagan D, Lee J, Blesa Cabez M, Rodriguez D, Boardman JP, Waldman A, Job DE, Wardlaw JM. Whole Brain Magnetic Resonance Image Atlases: A Systematic Review of Existing Atlases and Caveats for Use in Population Imaging. Front Neuroinform. 2017; 11. doi:10.3389/fninf.2017.00001.

Esteban O, Birman D, Schaer M, Koyejo OO, Poldrack RA, Gorgolewski KJ. MRIQC: Advancing the Automatic Prediction of Image Quality in MRI from Unseen Sites. PLoS One. 2017 Aug; 12(9):e0184661. doi:10.1371/journal.pone.0184661.

Esteban O, Markiewicz CJ, Blair RW, Moodie CA, Isik AI, Erramuzpe A, Kent JD, Goncalves M, DuPre E, Snyder M, Oya H, Ghosh SS, Wright J, Durnez J, Poldrack RA, Gorgolewski KJ. fMRIPrep: a robust preprocessing pipeline for functional MRI. Nat Meth. 2019 Jan; 16(1):111–116. doi:10.1038/s41592-018-0235-4.

Evans AC, Collins DL, Mills SR, Brown ED, Kelly RL, Peters TM. 3D statistical neuroanatomical models from 305 MRI volumes. In: *IEEE Conference Record Nuclear Science Symposium and Medical Imaging Conference*, vol. 3 San Francisco, CA, USA; 1993. p. 1813–1817. doi:10.1109/NSSMIC.1993.373602.

Evans AC, Janke AL, Collins DL, Baillet S. Brain templates and atlases. NeuroImage. 2012 Aug; 62(2):911–922. doi:10.1016/j.neuroimage.2012.01.024.

Fischl B, Sereno MI, Dale AM. Cortical surface-based analysis II: Inflation, flattening, and a surface-based coordinate system. NeuroImage. 1999; 9(2):195–207.

Fonov V, Evans AC, Botteron K, Almli CR, McKinstry RC, Collins DL. Unbiased average age-appropriate atlases for pediatric studies. NeuroImage. 2011 Jan; 54(1):313–327. doi:10.1016/j.neuroimage.2010.07.033.

Friston KJ, Ashburner J, Kiebel SJ, Nichols TE, Penny WD. Statistical parametric mapping : the analysis of functional brain images. London: Academic Press; 2006.

Goerzen D, Fowler C, Devenyi GA, Germann J, Madularu D, Chakravarty MM, Near J. An MRI-Derived Neuroanatomical Atlas of the Fischer 344 Rat Brain. Sci Rep. 2020 Apr; 10(1):6952. doi:10.1038/s41598-020-63965-x.

Good CD, Johnsrude IS, Ashburner J, Henson RN, Friston KJ, Frackowiak RS. A voxel-based morphometric study of ageing in 465 normal adult human brains. NeuroImage. 2001; 14(1):21–36. doi:10.1006/nimg.2001.0786.

Gorgolewski KJ, Auer T, Calhoun VD, Craddock RC, Das S, Duff EP, Flandin G, Ghosh SS, Glatard T, Halchenko YO, Handwerker DA, Hanke M, Keator D, Li X, Michael Z, Maumet C, Nichols BN, Nichols TE, Pellman J, Poline JB, et al. The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. Sci Data. 2016 Jun; 3:160044. doi:10.1038/sdata.2016.44.

Halchenko YO, FSL user mailing list: Incorrect probabilities in Harvard-Oxford-sub Left hemisphere; 2013. [Online] Available: https://www.jiscmail.ac.uk/cgi-bin/webadmin?A2=FSL;2bb44bee.1301 (Accessed: 2021-07-05).

Halchenko YO, Meyer K, Poldrack B, Solanky DS, Wagner AS, Gors J, MacFarlane D, Pustina D, Sochat V, Ghosh SS, Mönch C, Markiewicz CJ, Waite L, Shlyakhter I, Vega Adl, Hayashi S, Häusler CO, Poline JB, Kadelka T, Skytén K, et al. DataLad: distributed system for joint management of code, data, and their relationship. Journal of Open Source Software. 2021 Jul; 6(63):3262. doi:10.21105/joss.03262.

Holmes CJ, Hoge R, Collins L, Woods R, Toga AW, Evans AC. Enhancement of MR Images Using Registration for Signal Averaging. J Comput Assist Tomogr. 1998 Mar; 22(2):324–333. https://insights.ovid.com/crossref?an=00004728-199803000-00032.

Jenkinson M, Beckmann CF, Behrens TEJ, Woolrich MW, Smith SM. FSL. NeuroImage. 2012 Aug; 62(2):782–790. doi:10.1016/j.neuroimage.2011.09.015.

Kjonigsen LJ, Lillehaug S, Bjaalie JG, Witter MP, Leergaard TB. Waxholm Space atlas of the rat brain hippocampal region: Three-dimensional delineations based on magnetic resonance and diffusion tensor imaging. NeuroImage. 2015 Mar; 108:441–449. doi:10.1016/j.neuroimage.2014.12.080.

Landman BA, Ribbens A, Lucas B, Davatzikos C, Avants B, Ledig C, Ma D, Rueckert D, Vandermeulen D, Maes F. MICCAI 2012 Workshop on Multi-Atlas Labeling. CreateSpace Independent Publishing Platform; 2012.

Lüders E, Steinmetz H, Jäncke L. Brain size and grey matter volume in the healthy human brain. NeuroReport. 2002 Dec; 13(17):2371–2374. https://journals.lww.com/neuroreport/Abstract/2002/12030/Brain_size_and_grey_matter_volume_in_the_healthy.40.aspx.

MacNicol E, Ciric R, Kim E, Censo DD, Cash D, Poldrack R, Esteban O. Atlas-based brain extraction is robust across rat MRI studies. In: *IEEE 19th International Symposium on Biomedical Imaging (ISBI 2021)* Nice, France; 2021. p. (accepted). doi:10.1109/ISBI48211.2021.9.

Marcus DS, Wang TH, Parker J, Csernansky JG, Morris JC, Buckner RL. Open Access Series of Imaging Studies (OASIS): Cross-sectional MRI Data in Young, Middle Aged, Nondemented, and Demented Older Adults. J Cogn Neurosci. 2007 Aug; 19(9):1498–1507. doi:10.1162/jocn.2007.19.9.1498.

Mars RB, Passingham RE, Jbabdi S. Connectivity Fingerprints: From Areal Descriptions to Abstract Spaces. Trends Cogn Sci. 2018 Nov; 22(11):1026–1037. doi:10.1016/j.tics.2018.08.009.

Martin RF, Bowden DM. Primate Brain Maps: Structure of the Macaque Brain. Elsevier; 2000.

Matsuzawa J, Matsui M, Konishi T, Noguchi K, Gur RC, Bilker W, Miyawaki T. Age-related Volumetric Changes of Brain Gray and White Matter in Healthy Infants and Children. Cereb Cortex. 2001 Apr; 11(4):335–342. doi:10.1093/cercor/11.4.335.

Mazziotta J, Toga A, Evans A, Fox P, Lancaster J, Zilles K, Woods R, Paus T, Simpson G, Pike B, Holmes C, Collins L, Thompson P, MacDonald D, Iacoboni M, Schormann T, Amunts K, Palomero-Gallagher N, Geyer S, Parsons L, et al. A Four-Dimensional Probabilistic Atlas of the Human Brain. J Am Med Inform Assoc. 2001 Sep; 8(5):401–430. doi:10.1136/jamia.2001.0080401.

Mazziotta JC, Toga AW, Evans A, Fox P, Lancaster J. A Probabilistic Atlas of the Human Brain: Theory and Rationale for Its Development: The International Consortium for Brain Mapping (ICBM). NeuroImage. 1995 Jun; 2(2, Part A):89–101. doi:10.1006/nimg.1995.1012.

MNI, NeuroImaging & Surgical Technologies Lab (Montreal Neurological Institute) – MNI Atlases; n.d. [Online] Available: http://nist.mni.mcgill.ca/category/atlas/ (Accessed: 2021-07-01).

Nooner KB, Colcombe S, Tobe R, Mennes M, Benedict M, Moreno A, Panek L, Brown S, Zavitz S, Li Q, Sikka S, Gutman D, Bangaru S, Schlachter RT, Kamiel S, Anwar A, Hinz C, Kaplan M, Rachlin A, Adelsberg S, et al. The NKI-Rockland Sample: A Model for Accelerating the Pace of Discovery Science in Psychiatry. Front Neurosci. 2012; 6. doi:10.3389/fnins.2012.00152.

Osen KK, Imad J, Wennberg AE, Papp EA, Leergaard TB. Waxholm Space atlas of the rat brain auditory system: Three-dimensional delineations based on structural and diffusion tensor magnetic resonance imaging. NeuroImage. 2019 Oct; 199:38–56. doi:10.1016/j.neuroimage.2019.05.016.

Papp EA, Leergaard TB, Calabrese E, Allan Johnson G, Bjaalie JG. Addendum to "Waxholm Space atlas of the Sprague Dawley rat brain" [NeuroImage 97 (2014) 374-386]. NeuroImage. 2015 Jan; 105:561–562. doi:10.1016/j.neuroimage.2014.10.017.

Papp EA, Leergaard TB, Calabrese E, Johnson GA, Bjaalie JG. Waxholm Space atlas of the Sprague Dawley rat brain. NeuroImage. 2014 Aug; 97:374–386. doi:10.1016/j.neuroimage.2014.04.001.

Paxinos G, Watson C. The rat brain in stereotaxic coordinates. Elsevier Academic Press; 1997.

Pisner D, Hammonds R. PyNets: A Reproducible Workflow for Structural and Functional Connectome Ensemble Learning. In: *Annual Meeting of the Organization for Human Brain Mapping*, vol. 26 Online Event; 2020. https://github.com/dPys/PyNets/.

Robinson EC, Jbabdi S, Glasser MF, Andersson J, Burgess GC, Harms MP, Smith SM, Van Essen DC, Jenkinson M. MSM: A new flexible framework for Multimodal Surface Matching. NeuroImage. 2014; 100:414–426. doi:10.1016/j.neuroimage.2014.05.069.

Rohlfing T. Incorrect ICBM-DTI-81 atlas orientation and white matter labels. Front Neurosci. 2013; 7. doi:10.3389/fnins.2013.00004.

Satterthwaite TD, Connolly JJ, Ruparel K, Calkins ME, Jackson C, Elliott MA, Roalf DR, Hopson R, Prabhakaran K, Behr M, Qiu H, Mentch FD, Chiavacci R, Sleiman PMA, Gur RC, Hakonarson H, Gur RE. The Philadelphia Neurodevelopmental Cohort: A publicly available resource for the study of normal and abnormal brain development in youth. NeuroImage. 2016 Jan; 124:1115–1119. doi:10.1016/j.neuroimage.2015.03.056.

Sawiak SJ, Wood NI, Williams GB, Morton AJ, Carpenter TA. SPMMouse: A new toolbox for SPM in the animal brain. In: *Proc. Intl. Soc. Mag. Reson. Med.*, vol. 17 Hawaii, USA; 2009. p. 1086.

Schaefer A, Kong R, Gordon EM, Laumann TO, Zuo XN, Holmes AJ, Eickhoff SB, Yeo BTT. Local-Global Parcellation of the Human Cerebral Cortex from Intrinsic Functional Connectivity MRI. Cereb Cortex. 2018 Sep; 28(9):3095–3114. doi:10.1093/cercor/bhx179.

Schurr PH, Merrington WR. The Horsley–Clarke stereotaxic apparatus. Br J Surg. 1978; 65(1):33–36. doi:10.1002/bjs.1800650110.

Shi F, Yap PT, Wu G, Jia H, Gilmore JH, Lin W, Shen D. Infant Brain Atlases from Neonates to 1- and 2-Year-Olds. PLoS One. 2011 Apr; 6(4):1–11. doi:10.1371/journal.pone.0018746.

Sowell ER, Peterson BS, Thompson PM, Welcome SE, Henkenius AL, Toga AW. Mapping cortical change across the human life span. Nat Neurosci. 2003 Mar; 6(3):309–315. doi:10.1038/nn1008.

Szulc KU, Lerch JP, Nieman BJ, Bartelle BB, Friedel M, Suero-Abreu GA, Watson C, Joyner AL, Turnbull DH. 4D MEMRI atlas of neonatal FVB/N mouse brain development. NeuroImage. 2015 Sep; 118:49–62. doi:10.1016/j.neuroimage.2015.05.029.

Talairach J, Tournoux P. Co-planar stereotaxic atlas of the human brain. Stuttgart New York: Georg Thieme Verlag/Thieme Medical Publishers; 1988.

Talairach J, David M, Tournoux P, Corredor H, Kvasina J. Atlas d'anatomie stéréotaxique: repérage radiologique indirect des noyaux gris centraux des régions mésencéphalo-sous-optique et hypothalamique de l'homme. Masson; 1957.

Tavor I, Jones OP, Mars RB, Smith SM, Behrens TE, Jbabdi S. Task-free MRI predicts individual differences in brain activity during task performance. Science. 2016 Apr; 352(6282):216–220. doi:10.1126/science.aad8127.

Thompson WH, Fanton S, Open Source Software: NetPlotBrain; 2021. [Online] Available: https://zenodo.org/record/4593837 (Accessed: 2021-07-08).

Tosun D, Siddarth P, Levitt J, Caplan R. Cortical thickness and sulcal depth: insights on development and psychopathology in paediatric epilepsy. BJPsych Open. 2015 Oct; 1(2):129–135. doi:10.1192/bjpo.bp.115.001719.

Van Essen DC. Windows on the brain: the emerging role of atlases and databases in neuroscience. Curr Opin Neurobiol. 2002 Oct; 12(5):574–579. doi:10.1016/S0959-4388(02)00361-6.

Van Essen DC, Glasser MF, Dierker DL, Harwell J, Coalson T. Parcellations and Hemispheric Asymmetries of Human Cerebral Cortex Analyzed on Surface-Based Atlases. Cerebral Cortex. 2012 Oct; 22(10):2241–2262. doi:10.1093/cercor/bhr291.

Von Economo CF, Koskinas GN. Atlas of cytoarchitectonics of the adult human cerebral cortex, vol. 10. Karger Basel; 2008.

Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, Blomberg N, Boiten JW, da Silva Santos LB, Bourne PE, Bouwman J, Brookes AJ, Clark T, Crosas M, Dillo I, Dumon O, Edmunds S, Evelo CT, Finkers R, Gonzalez-Beltran A, et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data. 2016 Mar; 3(1):160018. doi:10.1038/sdata.2016.18.

Yarkoni T, Markiewicz C, de la Vega A, Gorgolewski K, Salo T, Halchenko Y, McNamara Q, DeStasio K, Poline JB, Petrov D, Hayot-Sasson V, Nielson D, Carlin J, Kiar G, Whitaker K, DuPre E, Wagner A, Tirrell L, Jas M, Hanke M, et al. PyBIDS: Python tools for BIDS datasets. J Open Source Softw. 2019 Aug; 4:1294. doi:10.21105/joss.01294.

Yeo BT, Krienen FM, Sepulcre J, Sabuncu MR, Lashkari D, Hollinshead M, Roffman JL, Smoller JW, Zöllei L, Polimeni JR, Fischl B, Liu H, Buckner RL. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. J Neurophysiol. 2011 Jun; 106(3):1125–1165. doi:10.1152/jn.00338.2011.

Yoon U, Fonov VS, Perusse D, Evans AC. The effect of template choice on morphometric analysis of pediatric brain data. NeuroImage. 2009 Apr; 45(3):769–777. doi:10.1016/j.neuroimage.2008.12.046.

Online methods

Design and architecture

TemplateFlow comprises four cardinal components: (i) a cloud-based archive, (ii) a Python client for programmatically querying the archive, (iii) automated systems for synchronizing and updating archive data, and (iv) inter-template registration workflows. Here, we discuss the details of each component's implementation in turn, as well as the manner in which they interact with one another to form a cohesive whole.

The *TemplateFlow Archive*.

The archive itself comprises directories of template data in cloud storage. For redundancy, the data are stored on both Google Cloud using the Open Science Framework (OSF) and on Amazon's Simple Storage Service (S3). Prior to storage, all template data must be named and organized in directories conforming to a data structure adapted from the Brain Imaging Data Structure (BIDS) standard (***Gorgolewski et al., 2016***). The precise implementation of this data structure is a living document and is detailed on the *TemplateFlow* homepage (http://www.templateflow.org). We detail several critical features here.

The archive is organized hierarchically, and descriptive metadata follow a principle of inheritance: any metadata that apply to a particular level of the archive also apply to all deeper levels. At the top level of the hierarchy are directories corresponding to each archived template. If applicable, within each template directory are directories corresponding to sub-cohort templates. Names of directories and resource files constitute a hierarchically ordered series of key-value pairs terminated by a suffix denoting the datatype. For instance, `tpl-MNIPediatricAsym_cohort-3_res-high_T1w.nii.gz` denotes a T1-weighted template image file for resolution "high" of cohort "3" in the "MNIPediatricAsym" template (where the definitions of each resolution and cohort are specified in the template metadata file, *TemplateFlow Archive*). The most common *TemplateFlow* datatypes are indexed in Suppl. Table S2; an exhaustive list is available in the most current version of the BIDS standard (https://bids.neuroimaging.io/).

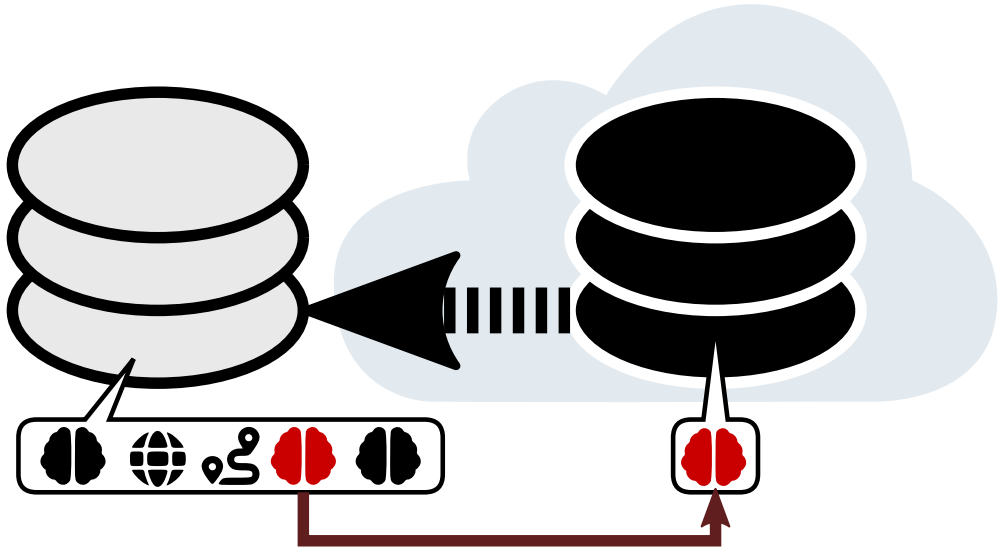
Within each directory, template resources include image data, atlas and template metadata, transform files, licenses, and curation scripts. All image data are stored in gzipped NIfTI-1 format and are conformed to RAS+ orientation (i.e., left-to-right, posterior-to-anterior, inferior-to-superior, with the affine qform and sform matrices corresponding to a cardinal basis scaled to the resolution of the image). Template metadata are stored in a JavaScript Object Notation (JSON) file called `template_description.json`; an overview of metadata specifications is provided in Figure 3 of the main manuscript. In brief, template metadata files contain general template metadata (e.g., authors and curators, references), cohort-specific metadata (e.g., ages of subjects included in each cohort), and resolution-specific metadata (e.g., dimensions of images associated with each resolution). Atlas metadata are often stored in TSV format and specify the region name corresponding to each atlas label. Transform files are stored in HDF5 format and are generated as a diffeomorphic composition of ITK-formatted transforms mapping between each pair of templates.

The archive has a number of client-facing access points to facilitate browsing of resources. Key among these is the archive browser on the *TemplateFlow* homepage, which indexes all archived resources and provides a means for researchers to take inventory of possible templates to use for their study.

The Python client.

TemplateFlow is distributed with a Python client that can submit queries to the archive and download any resources as they are requested by a user or program. Valid query options correspond approximately to BIDS key-value pairs and datatypes. A compendium of common query arguments is provided in Suppl. Table S2, and comprehensive documentation is available on the *TemplateFlow* homepage.

When a query is submitted to the *TemplateFlow* client, the client begins by identifying any files



```
>>> from templateflow import api as tflow
>>> tflow.get('MNI152NLin6Asym', desc=None, resolution=1,
...           suffix='T1w', extension='nii.gz')
PosixPath('/templateflow_home/tpl-MNI152NLin6Asym/tpl-
          MNI152NLin6Asym_res-01_T1w.nii.gz')
```

Figure 5. Example usage of the Python-based *TemplateFlow Client*. After importing the API, the user submits a query for the T1-weighted FSL version of the MNI template at 1 mm resolution. The client first filters through the archive, identifies any files that match the query, and finds their counterparts in cloud storage. It then downloads the requested files and returns their paths in the local *TemplateFlow* installation directory. Future queries for the same resource can be completed without any re-downloading.

in the archive that match the query. To do so, it uses *PyBIDS* (***Yarkoni et al., 2019***), which exploits the BIDS-like architecture of the *TemplateFlow Archive* to efficiently scan all directories and filter any matching files. Next, the client assesses whether queried files exist as data in local storage. When a user locally installs *TemplateFlow*, the local installation initially contains only lightweight pointers to files in OSF cloud storage. These pointers are implemented using *DataLad* (***Halchenko et al., 2021***), a data management tool that extends git and git-annex. *TemplateFlow* uses *DataLad* principally to synchronize datasets across machines and to perform version control by tracking updates made to a dataset.

If the queried files are not yet synchronized locally (i.e., they exist only as pointers to their counterparts in the cloud), the client instructs *DataLad* to retrieve them from cloud storage. In the event that *DataLad* fails or returns an error, the client falls back on redundancy in storage and downloads the file directly from Amazon's S3. When the client is next queried for the same file, it will detect that the file has already been cached in the local filesystem. The use of resource pointers with the client thus enables lazy loading of template resources. Finally, the client confirms that the file has been downloaded successfully. If the client detects a successful download, it returns the result of the query; in the event that it detects a synchronization failure, it displays a warning for each queried file that encountered a failure.

Continued functionality and operability of the client is ensured through an emphasis on maximizing code coverage with unit tests. Updating the client requires successful completion of all unit tests, which are automatically executed by continuous integration (CI) and continuous delivery (CD) services connected to GitHub. CI and CD also keep the web-based archive browser up to date by automatically indexing data files.

Ancillary and managerial systems.

TemplateFlow includes a number of additional systems and programs that serve to automate stages of the archive update process, for instance addition of a new template or revision of current template resources. To facilitate the update and extension process, *TemplateFlow* uses GitHub actions

Box 1. Integration of *TemplateFlow* in processing workflows

TemplateFlow maximizes the accessibility and reuse potential of templates and atlases. For example, let's reuse the base configuration file for FSL FEAT we proposed in our paper (***Esteban et al., 2019***). The design file `design.fsf` specifies a simple preprocessing workflow with FSL tools. The simplified code listing below shows that, just to make non-default templates available to FSL using the graphical user interface (GUI), at least four steps are necessary:

```
# 1. User determines two nondefault templates they want to spatially normalize into
# 2. User manually download templates, extract the required files from packages
$ curl -sSL <url> | tar zxv --no-same-owner -C /data/templates/
# 3. User opens FSL's GUI, edits the target template box content pointing to the appropriate files
# 4. User generates FSL configuration files to permit batch execution on the command line
# 5. For the default and the two nondefault templates, execute FSL's feat:
$ feat design_<template>.fsf
```

The outputs of each `feat design_<template>.fsf` call will follow the pre-specified patterns of FSL, with whatever customization the user has introduced into the design file. The user, therefore, must then adapt the downstream analysis tools to correctly interpret the derived dataset, in each standard space, or reformat the output dataset according to the expectations of the analysis tools.
The user is also responsible for all aspects of provenance tracking and adequately reporting them in their communications. Information such as version of the template (or download date), citations to relevant papers, and other metadata (e.g., RRIDs) must be accounted for manually throughout the research process.

In contrast, tools using *TemplateFlow* dramatically simplify the whole process (note that `MNI152NLin2009cAsym` and `OASIS30Ants` are the two templates not found within the FSL distribution, and `MNI152NLin6Asym` denotes *FSL's MNI space* (i.e., the default FSL template):

```
$ fmriprep /data /derivatives participant --output-spaces MNI152NLin2009cAsym MNI152NLin6Asym OASIS30Ants
```

fMRIPrep generates the results with BIDS-Derivatives organization for the three templates. The tool also leverages *TemplateFlow* to generate a *boilerplate citation text* that includes the full names, versions and references to credit the template's authors for each of the templates involved.
fMRIPrep internally stages one spatial normalization workflow for each of the output spaces. Each of these normalization sub-workflows uses a simple line of Python code to retrieve the necessary resources from *TemplateFlow* using the *TemplateFlow Client* interface (Figure 5):

```
>>> from templateflow.api import get
>>> tpl_ref_file = get("MNI152NLin6Asym", desc=None, resolution=1, suffix="T1w", extension="nii.gz")
```

One detail overseen in the FSL example is that, for a robust spatial normalization process, a precise binary mask of the brain is generally used. While FSL would require the user to manually set this mask up in the GUI, in the case of *TemplateFlow*, it requires a second minimal call:

```
>>> msk_ref_file = get("MNI152NLin6Asym", desc="brain", resolution=1, suffix="mask", extension="nii.gz")
```

These examples are extreme simplifications of what a pipeline developer can automate and make more robust by integrating *TemplateFlow* in their workflows.
For further examples on how *TemplateFlow* can be leveraged, *PyNets* is a package for ensemble learning of functional and structural connectomes (***Pisner and Hammonds, 2020***), and *NetPlotBrain* for visualization (***Thompson and Fanton, 2021***).

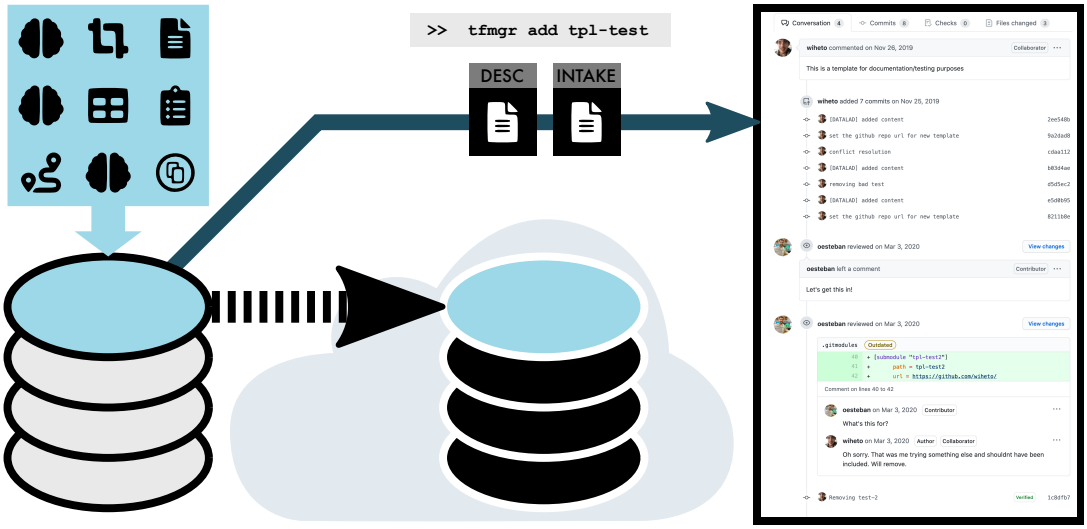


Figure 6. Contributing a new resource with the *TemplateFlow Manager*. To contribute a new template to *TemplateFlow*, members of the community first organize template resources to conform to the BIDS-like *TemplateFlow* structure. Next, `tfmgr` (the *TemplateFlow Manager*, see Suppl. Table S3) synchronizes the resources to OSF cloud storage and opens a new pull request proposing the addition of the new template. A subsequent peer-review process ensures that all data are conformant with the *TemplateFlow* standardd. Finally, *TemplateFlow* curators merge the pull request, thereby adding the template into the archive.

to automatically synchronize dataset information so that all references remain up-to-date with the current dataset. These actions are triggered whenever a pull request to *TemplateFlow* is accepted. For example, GitHub actions are used to update the browser of the *TemplateFlow Archive* so that it displays all template resources as they are uploaded to the archive.

Whereas the *TemplateFlow Client* synchronizes data from cloud storage to the local filesystem, a complementary *TemplateFlow Manager* handles the automated synchronization of data from the local filesystem to cloud storage. The Python-based manager is also used for template intake, i.e., to propose the addition of new templates to the archive. To propose adding a new template, a user first runs the *TemplateFlow Manager* using the `tfmgr add <template_id> --osf-project <project_id>` command.

The manager begins by using the *TemplateFlow* client to query the archive and verify that the proposed template does not already exist. After verifying that the proposed template is new, the manager synchronizes all specified template resources to OSF cloud storage. It then creates a fork of the `tpl-intake` branch of the *TemplateFlow* GitHub repository and generates an intake file in Tom's Obvious Minimal Language (TOML) markup format; this intake file contains a reference to the OSF project where the manager has stored template resources. The *TemplateFlow Manager* commits the TOML intake file to the fork and pushes to the user's GitHub account. Finally, it retrieves template metadata from `template_description.json` and uses the metadata to compose a pull request on the `tpl-intake` branch. This pull request provides a venue for discussion and vetting of the proposed addition of a new template.

Inter-template registration workflow.

To enable the flow of knowledge across template spaces, *TemplateFlow* includes a workflow for computing robust transformations between any pair of adult human template spaces. To compute a transformation between two template spaces, the inter-template registration workflow makes use of 10 of the high-quality T1-weighted adult human brain images used in the creation of the MNI 152 template portfolio. In the first step of the workflow, these 10 images are registered to both template spaces using the symmetric normalization (SyN) algorithm (***Avants et al., 2008***). Next, a 10-channel registration is performed in ANTs using the SyN algorithm. Thus, the workflow com-

putes a single transformation that simultaneously optimizes the alignment between all 10 images in both coordinate spaces.

MNI space text mining analysis.

To investigate the use of the term "MNI" in the neuroimaging literature, we conducted an exploratory text mining analysis. For this, we used the Elsevier API to download the entire corpus of two leading journals of neuroimaging methodology, *NeuroImage* and *NeuroImage: Clinical*. In this way, we retrieved a total 16,812 full-text articles that were subsequently segmented into lists of sentences. A scan of these sentences revealed 14,870 sentences across 6,048 articles that contained the word "MNI". Sentences were cleaned (i.e., removing punctuation, single letters, accents, numbers) and tokenized into words, which were subsequently lemmatized (i.e., converted to base form) using the NLTK wordnet lemmatizer. From the lemmatized words, we filtered out stopwords (i.e., NLTK stopwords and a custom list) and included words with a frequency above 10 as part of our "dictionary"; this yielded a dictionary size of 2,324 words.

Next, we computed a sparse dictionary by article count matrix (i.e., 2,324 $\times$ 6,048), on which we performed topic modelling with latent Dirichlet allocation (LDA; ***Blei et al., 2003***, implementation from scikit-learn with the learning decay hyperparameter set to 0.7). The number of topics (k=15) was selected by identifying the LDA model yielding the lowest perplexity (***Blei et al., 2003***) sweeping the interval [8-16] for the parameter. The 20 words from the dictionary that loaded the highest on the 15 topics were visualized using word clouds. Topics were sorted by descending dominance, and the dominance fraction (number of articles where the topic is the most loaded with respect to the total 6,048 documents) was included above the corresponding topic's word cloud (Figure 4, and Suppl. Figure S1).

Supp. Box S1. The FAIR Guiding Principles

(Reproduced from ***Wilkinson et al., 2016***, Box 2)

To be **Findable**:

- F1. (meta)data are assigned a globally unique and persistent identifier
- F2. data are described with rich metadata (defined by R1 below)
- F3. metadata clearly and explicitly include the identifier of the data it describes
- F4. (meta)data are registered or indexed in a searchable resource

To be **Accessible**:

- A1. (meta)data are retrievable by their identifier using a standardized communications protocol
  - A1.1. the protocol is open, free, and universally implementable
  - A1.2. the protocol allows for an authentication and authorization procedure, where necessary
- A2. metadata are accessible, even when the data are no longer available

To be **Interoperable**:

- I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
- I2. (meta)data use vocabularies that follow FAIR principles
- I3. (meta)data include qualified references to other (meta)data

To be **Reusable**:

- R1. (meta)data are richly described with a plurality of accurate and relevant attributes
  - R1.1. (meta)data are released with a clear and accessible data usage license
  - R1.2. (meta)data are associated with detailed provenance
  - R1.3. (meta)data meet domain-relevant community standards

Supp. Box S2. Quick start with the *TemplateFlow Client* API

Finding templates

At the time of writing there are 15 templates available within the resource, and all the unique identifiers can be accessed with the `templates()` method:

```
>>> from templateflow.api import templates
>>> f"{', '.join(templates())} ({len(templates())} templates)"
"""Fischer344, MNI152Lin, MNI152NLin2009cAsym, MNI152NLin2009cSym, MNI152NLin6Asym, MNI152NLin6Sym,
MNIInfant, MNIPediatricAsym, NKI, OASIS30ANTs, PNC, UNCInfant, WHS, fsLR, fsaverage (15 templates)"""
```

Accessing Metadata

We can query metadata associated to individual data files (e.g., a volume or a surface) or general metadata of the template. For example, the `get_metadata(<template_id>)` returns the general metadata as a dictionary. Hence, consulting the full name corresponding to some template identifiers yields:

```
>>> from templateflow.api import get_metadata
>>> [(tpl, get_metadata(tpl)["Name"]) for tpl in templates()[:3]]
[("Fischer344", "MRI-Derived Neuroanatomical Atlas of the Fischer 344 Rat Brain"),
 ("MNI152Lin", "Linear ICBM Average Brain (ICBM152) Stereotaxic Registration Model"),
 ("MNI152NLin2009cAsym", "ICBM 152 Nonlinear Asymmetrical template version 2009c")]
```

Similarly, we can check the license of a given template:

```
>>> get_metadata("UNCInfant")["License"]
"CC-BY"
```

Or the proper citations (please note that the output in this example has been manipulated for demonstration purposes):

```
>>> print(get_citations("UNCInfant", bibtex=True)[0])
@article{uncinfant1,
      doi = {10.1371/journal.pone.0018746},
      year = 2011,
      volume = {6},
      number = {4},
      pages = {e18746},
      author = {Feng Shi and Pew-Thian Yap and Guorong Wu and Hongjun Jia and John H. Gilmore and
                Weili Lin and Dinggang Shen},
      title = {Infant Brain Atlases from Neonates to 1- and 2-Year-Olds},
      journal = {{PLoS ONE} {ONE}}
}
```

```
>>> from templateflow.api import get
>>> print(get("MNI152NLin6Asym", desc=None, resolution=1, suffix="T1w"))
Downloading https://templateflow.s3.amazonaws.com/tpl-MNI152NLin6Asym/tpl-MNI152NLin6Asym_res-01_T1w.nii.gz
100%|##################################################################| 3.29M/3.29M [00:01<00:00, 2.11MB/s]
/home/oesteban/.cache/templateflow/tpl-MNI152NLin6Asym/tpl-MNI152NLin6Asym_res-01_T1w.nii.gz
```

Supp. Table S1. Common terms and definitions.

Atlas	A map between spatial coordinates and knowledge-based annotations, such as neuroanatomical structures or functional brain networks.
Template	A digital image that is representative of a population of interest (for instance an average of individual images from the population) and can be used as a spatial reference for that population.
Volumes and surfaces	Two common schemes for spatially organising brain data. Volumetric data are situated in the Euclidean topology of physical three-dimensional space, while surface data make use of the native topology of the cortex's sulcal curvature.
Stereotaxic coordinate system	A three-dimensional coordinate system that can be referenced to unambiguously localise a position in the brain, defined using an origin together with a basis set of axes.
Standard space	A coordinate system engendered by a template, which can be used as a reference when studying a population. A volumetric *template* engenders a *stereotaxic coordinate system*, which can be annotated in different ways to define brain *atlases*.
Transform	A diffeomorphic function that maps a set of spatial coordinates to their analogues or homologues in another coordinate space. For example, transforms can map from structures in the brains of individual subjects to analogous structures in a population-average template.

Supp. Table S2. *TemplateFlow* data entities.

Data entity	API query example	Description
Template	`"MNI152Lin"`	The template dataset to which an image or other data file belongs.
Resolution	`res=1`	The image resolution. Each resolution is assigned a key, which is defined in the `res` field of `template_description.json`.
Mask	`desc="brain", suffix="mask"`	Indicates that the image is a binary-valued annotation, where voxels labelled 1 are part of the mask.
Discrete segmentation	`desc="malf", suffix="dseg"`	Indicates that the image is an integer-valued annotation. Each segmentation image file (`.nii.gz` format) is paired with a dictionary of segment names (`.tsv` format).
Probabilistic segmentation	`label="CSF", suffix="probseg"`	Indicates that the image is a probabilistic annotation, wherein the value of each voxel indicates the probability of that voxel belonging to the specified label.
Atlas	`atlas="Schaefer", desc="7Network"`	The atlas to which a segmentation file belongs.
Transformation	`from="MNI152Lin", suffix="xfm"`	File containing a mapping between 2 stereotaxic coordinate spaces. The source space is defined in the `from` field, while the target space is defined in the `tpl` field.
Image modality	`suffix="T1w"`	For non-annotation brain images, the suffix indicates whether the image is T1-weighted (`T1w`), T2-weighted (`T2w`), proton density-weighted (`PD`), or T2*-weighted (`T2star`).
Template cohort	`cohort=1`	Subsample of a dataset used to generate an average template.

Supp. Table S3. Command-line interface for *TemplateFlow Manager*.

Argument	Environment variable	Specifications
`template_id`		Identifier of the template. This is the value of the `tpl` field in all file names.
`--osf-project`	`OSF_PROJECT`	The OSF project where the template data are to be stored. The project must be writable by the user account whose credentials are specified in the `--osf-user` and `--osf-password` arguments.
`--osf-user`	`OSF_USERNAME`	Account username or identifier for OSF cloud storage.
`--osf-password`	`OSF_PASSWORD`	Account password for OSF cloud storage.
`--osf-overwrite`		Flag that indicates that the OSF client should force the overwrite of any existing files in the OSF project that have names conflicting with those of new files.
`--gh-user`	`GITHUB_USER`	Account username for GitHub. The user account whose credentials are provided must have a fork of the TemplateFlow repo.
`--gh-password`	`GITHUB_PASSWORD`	Account password for GitHub.
`--path`	`OSF_PROJECT`	Path to a local directory where template resources are located. The path must either be a directory whose name is `tpl-<template_id>` or contain such a directory.
`--nprocs`		Maximum number of parallel processes to run when uploading to or fetching from OSF.

Supp. Figure S1. Full report for the topic modeling exploration. All 15 topics modeled are represented —sorted by dominance— in this figure, completing the partial results of Figure 4.