

# CBRAIN as a platform for Open Science



Andrew Doyle<sup>1</sup>, Pierre Rioux<sup>1</sup>, Natacha Beck<sup>1</sup>, Tristan Aumentado-Armstrong<sup>1</sup>, Samir Das<sup>1</sup>, Marc-Etienne Rousseau<sup>1</sup>, Reza Adalat<sup>1</sup>, Tristan Glatard<sup>2</sup>, Alan C. Evans<sup>1</sup>

<sup>1</sup>McGill Centre for Integrative Neuroscience, Montreal Neurological Institute, McGill University

<sup>2</sup>Department of Computer Science and Software Engineering, Concordia University



## Introduction

CBRAIN [1] is a generic web framework to process and visualize data. The tools currently available perform mainly structural and functional brain image processing, and new developments presented in this work have allowed the rapid adaptation of new tools for genetics data analysis. With the Healthy Brain for Healthy Lives initiative [2], CBRAIN will process data for an increasing number of research studies, requiring more tools for brain analysis. A summary of CBRAIN vital statistics is presented in Figure 1.

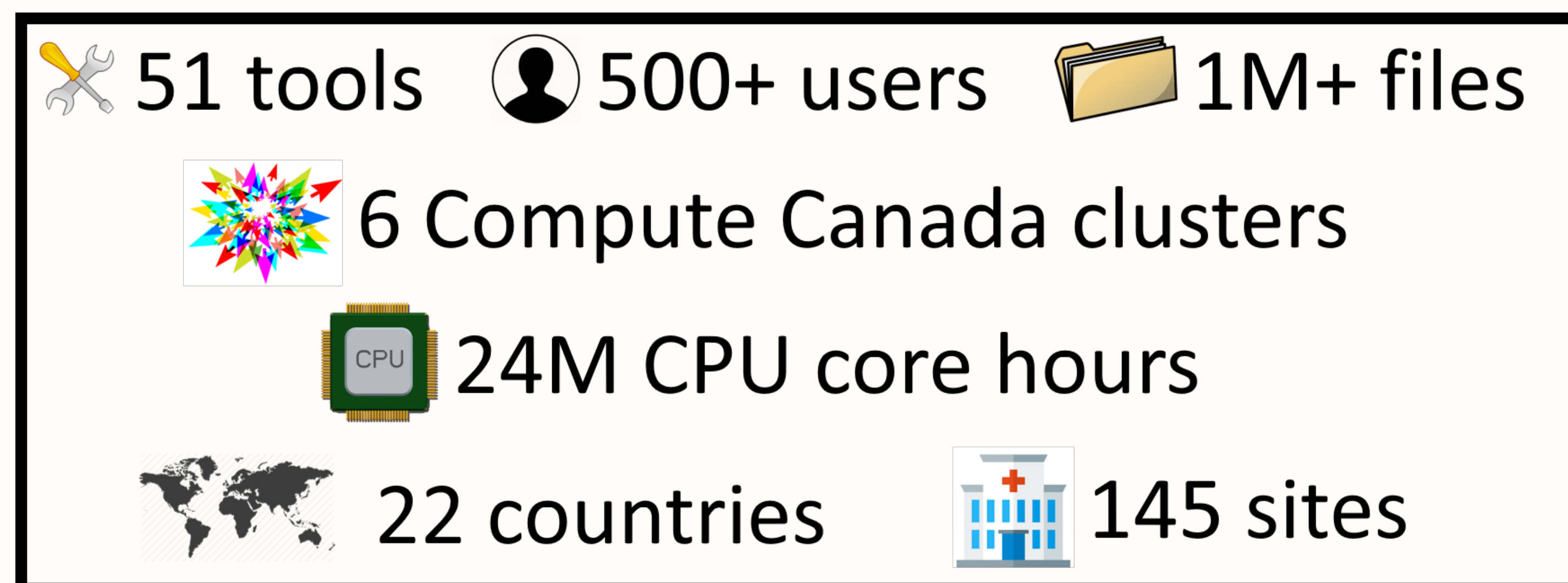


Figure 1: Statistics on CBRAIN's adoption and use.

The latest technical developments include:

- 1 Improved ways to add new tools
- 2 User account registration
- 3 External software integration through a web Application Programming Interface

These improvements allow CBRAIN to be easily extended to a wider variety of scientific applications, open the platform to new users, and support the Montreal Neurological Institute's mission of Open Science [3].

## Methods

Open Science will allow many new users to access CBRAIN, and a registration web form has been added to CBRAIN. Canadian scientists and their collaborators may register at <https://portal.cbrain.mcgill.ca/signup/new> (or scan Figure 2).

New users will require that their software tools be made available in CBRAIN. Tool integration into CBRAIN has until now required that one software developer has detailed knowledge of: (1) The new tool, (2) CBRAIN, and (3) Ruby. Tools had to be installed on each execution server, and a description for the tool had to be created in CBRAIN. To avoid frequent installation of tools on many distributed servers, the Docker [4] and Singularity [5] container technologies were adopted. Docker containers can be retrieved by execution servers from the DockerHub cloud container registry when needed.

In addition, Boutiques [6] is a JSON specification that allows the command-line parameters of scientific tools to be described in a portable and standard manner, and includes the expected input files, parameter names and their allowed values, and the output result files that the tool generates. Using Boutiques descriptors, CBRAIN can now automatically generate web pages and Ruby code for tools based on the parameter descriptions. The new tool integration features are outlined in Figure 3.



Figure 2: CBRAIN registration page

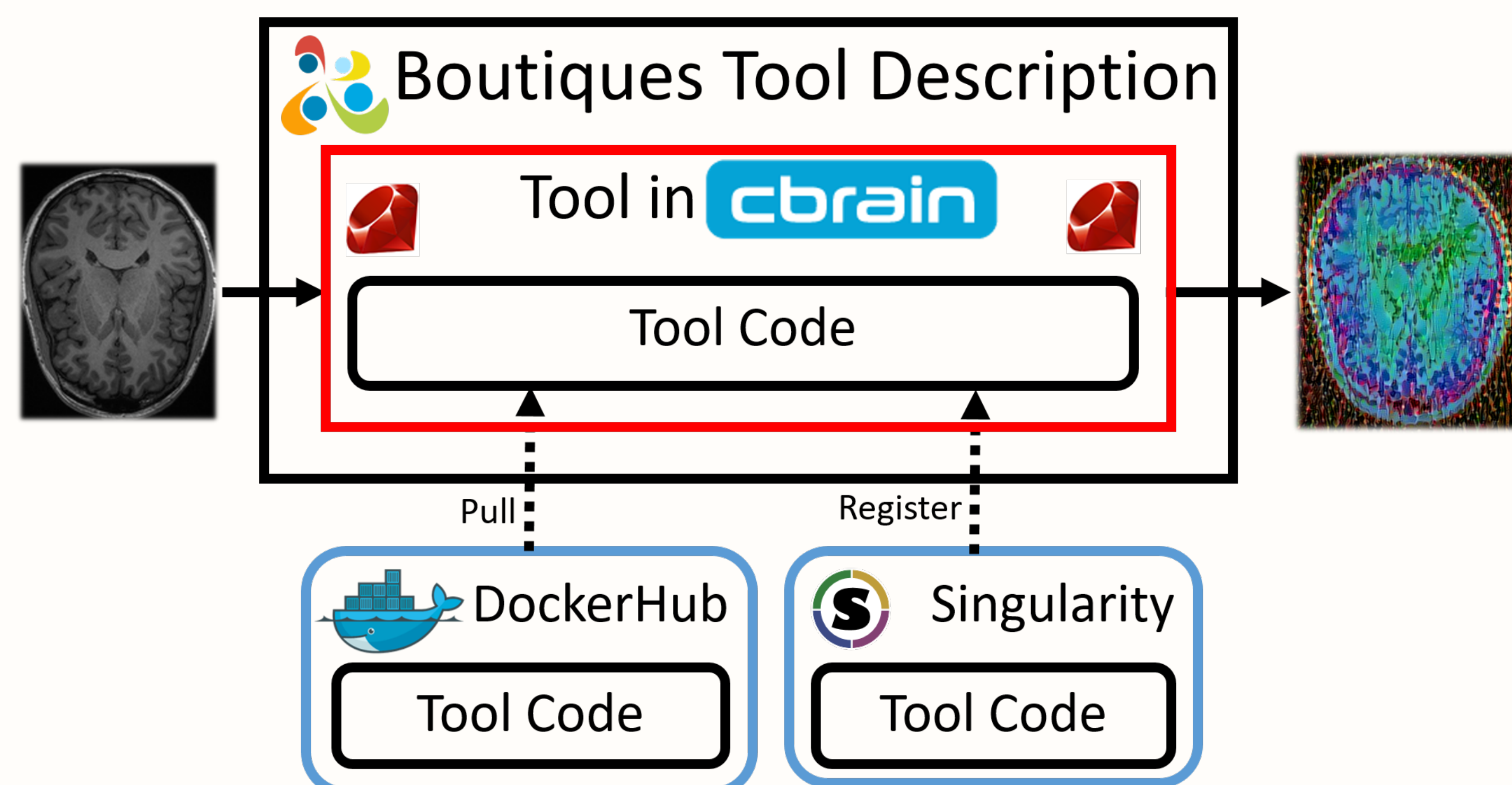


Figure 3: Boutiques descriptors allow the automatic generation of Ruby code and Docker/Singularity container integration allows easier distribution and reproducibility of results.

Open Science will also require that more systems connect to CBRAIN, such as databases or other processing clients. The CBRAIN RESTful API allows external systems to harness CBRAIN's computational power. The API has been simplified and described in the OpenAPI framework

specification, which fully defines the API in a JSON format. Swagger automatically generates dynamic documentation, including a web page to test each action, and allows the automatic generation of Software Development Kits (SDKs) for over 25 programming languages, allowing the easy integration of CBRAIN into any external system. Figure 4 shows how CBRAIN's API is described an example of the automatically-generated interactive documentation.

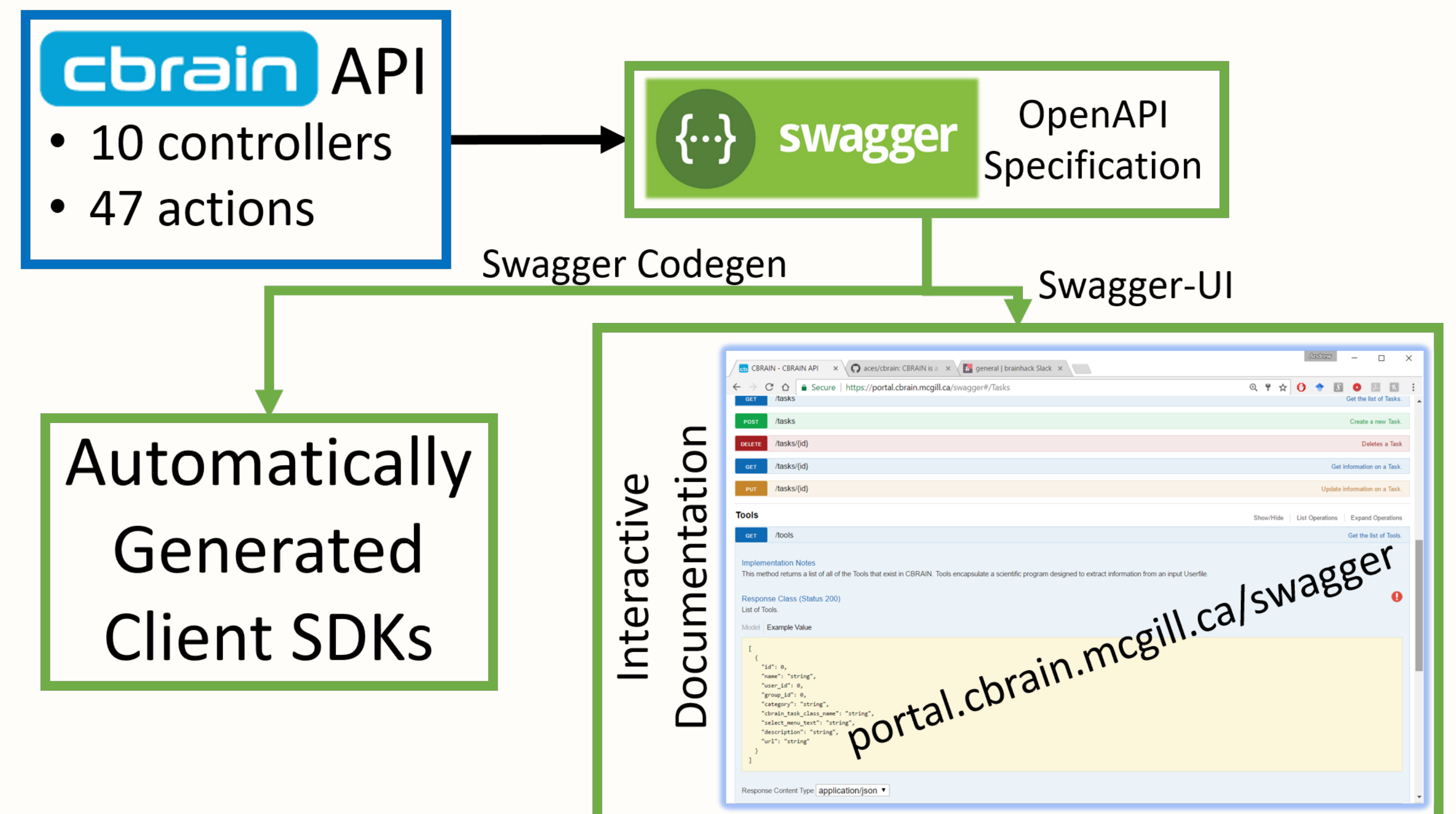


Figure 4: CBRAIN's description in the OpenAPI specification allows Swagger to generate interactive documentation and software development kits for many different programming languages.

## Results

Tool incorporation has been greatly accelerated, allowing many new tools to be added (see Figure 5 (a)). For neuroimaging, 10 new tools have been described by Boutiques, including many FSL tools and NIAK. Code for NIAK, ICA-Aroma and FSL's ProbtrackX is run in Docker containers. Two new tools for genetics data normalization and analysis have been rapidly integrated using Boutiques, Docker and Singularity.

So far 99 new users have used the registration system since January 2017 (see Figure 5 (b)). The API has been fully documented and simplified (see Figure 5 (c), is in use by LORIS [7](an internal project critical to Open Science) and has been sent to external collaborators.

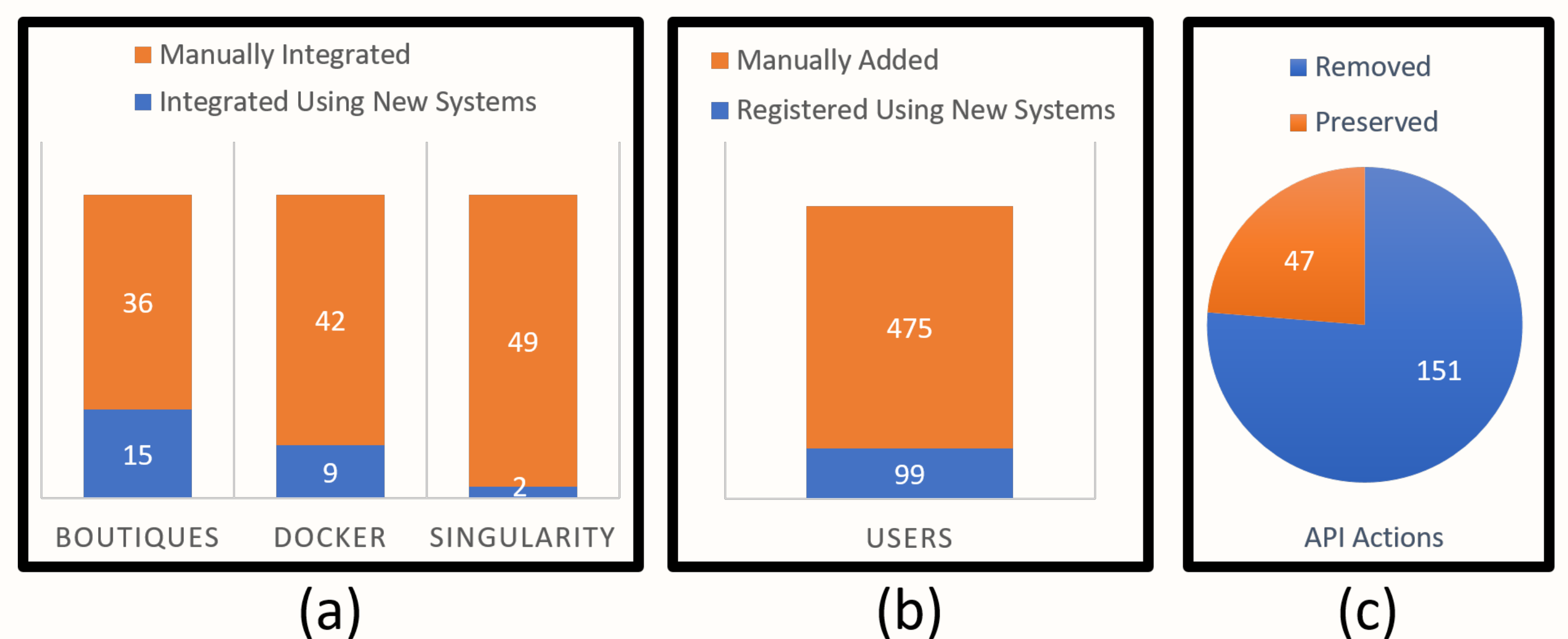


Figure 5: (a) Results for tools using the new systems (b) Users who have registered using the new system (c) API Endpoints preserved/documentated vs. superfluous actions that were removed

## Conclusion

Open Science will unlock CBRAIN to be used by a cohort of new users who will require new tools and want to connect their existing software systems. The new systems for seamless tool integration and external system connections will position CBRAIN to accelerate neuroscience research for everyone.

## References

- [1] Sherif, T, et al. "CBRAIN: a web-based, distributed computing platform for collaborative neuroimaging research." Recent Advances and the Future Generation of Neuroinformatics Infrastructure (2014): 102.
- [2] McGill University. "Healthy Brains for Healthy Lives". Canada First Research Excellence Fund (2016). <https://www.mcgill.ca/hbhl/>
- [3] Das, S, et al. "The MNI data-sharing and processing ecosystem." NeuroImage 124 (2016): 1188-1195.
- [4] Merkel, D. "Docker: lightweight linux containers for consistent development and deployment." Linux Journal 2014.239 (2014): 2.
- [5] Kurtzer, G, et al. "Singularity: Scientific containers for mobility of compute." PLoS one 12.5 (2017): e0177459.
- [6] Glatard, T, et al. "Boutiques: an application-sharing system based on Linux containers." Neuroinformatics (2015).
- [7] Das, S, et al. "Cyberinfrastructure for Open Science at the Montreal Neurological Institute." Frontiers in neuroinformatics 10 (2016).