# Tufts Data Science Singularity image

*Kyle Monahan, Statistics and Research Technology Specialist*

## Background

There is a great need to enable environment independent research workflows at Tufts. Right now, programs work on local computers, but transitioning these to distributed, high-performance infrastructure can be wrought with issues.

The software **Docker**1, an enterprise container platform aims to *containerize*, by placing inside small Linux based virtual machines, the software components of a research workflow. This reduces the reliance of researchers on IT support infrastructure, as these images can be deployed on an as-needed basis. The software **Singularity**2 sits on top of Docker images, allowing for users to access them without write permissions on the machine. This is useful for shared computer environments such as the Tufts High-Performance Compute cluster.

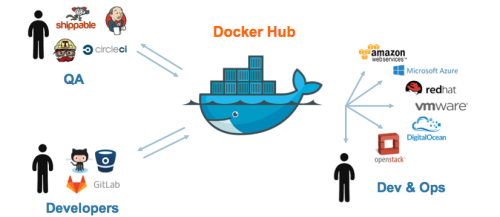


Figure 1 Docker enables operating-system level virtualization between developers, QA/QC, DevOps and researchers.

## Goal

The goal of this document is to document and produce an image for Delilah M. such that she can download (docker pull) and update (docker push) an image to a shared repository. This will solve ongoing issues with deploying across environments on the cluster.

## Steps

1. Use the vagrant subsystem to create an image where we can work with singularity and docker without issue. For development, we use the vagrant subsystem from Hashicorp to work with Virtualbox, but there are many image management techniques. More information and install instructions here: <https://www.vagrantup.com/docs/cli/init.html>
2. Navigate to the folder you want to share the image with, or work with docker files. Install *vagrant* as above. Once installed, type:
   1. *vagrant init hashicorp/precise64*
   2. This provides a minimal image for vagrant to use.
   3. Note that vagrant automatically mounts files, e.g. /vagrant [in the image we created] => current working directory [where we ran vagrant init]
3. Now we want to access some … [paused on 11-15-2018]

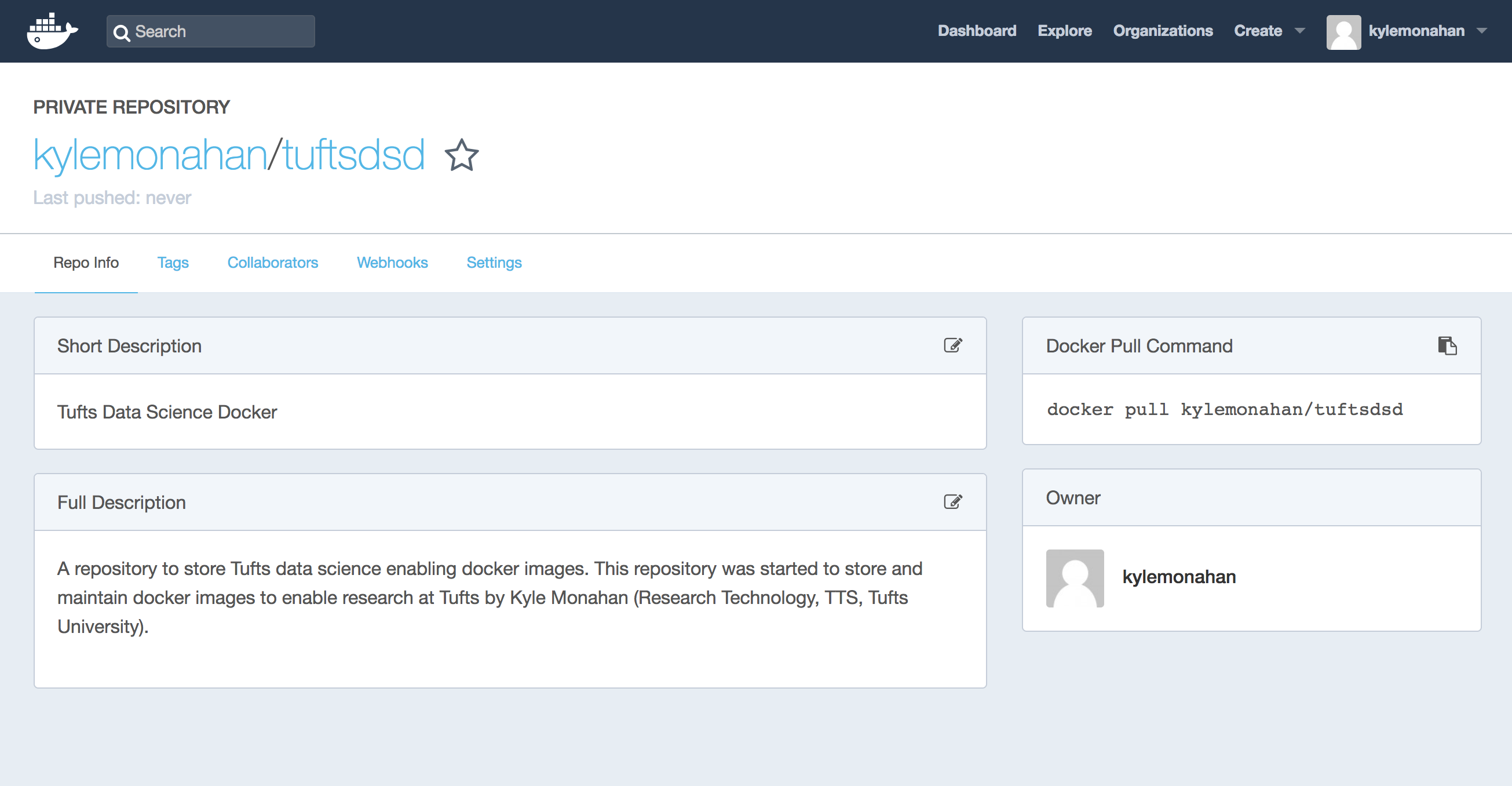


Figure 2 The Tufts Data Science Docker repository.

Sources (accessed 11-15-2018):

1. Docker Hub Blog Post. <https://blog.docker.com/2015/09/docker-hub-2-0/>
2. Singularity. https://singularity.lbl.gov/