Docker does not control what packages are in the images, whether the image will deploy correctly or the images might have any security problem.

Daniel Walsh, Consulting Engineer at Red Hat said Docker is about running random crap from the internet as root on your host.

DockerPedia is an RDF linked dataset that stores the information about Docker images hosted in Docker Hub including 4.5 million of images, its layers and packages. We also provide vulnerability analysis of these packages obtained with Clair.

To extract the information from the Docker images hosted in Docker Hub, we performed a search over its free text box to obtain all the Docker images. In February 2018, this search returned 1,360,510 Docker repositories and 4,609,643 images composed of 4,593,602 community images and 14,841 official images. The total size of these images is 53.47 PB.

After that, we use the tool from the Clair project to detect vulnerabilities within each downloaded image.

The ontology first imports abstract classes and relations from the OWL-S standard, and then it imports the Docker Ontology for some of the Docker concepts and the WICUS ontology 

Our work

Our proposal allows creating a visualization tool, which uses the data available at DockerPedia, to help Docker users search and compare between different Docker images, allowing them to find software distributions which fit their needs and monitor the state of Docker repositories over time.

By first creating a descriptor file called Dockerfile.

Anyone has the chance to create and store images into the Docker Hub registry by first creating a descriptor file called Dockerfile. This descriptor describes what software packages will be within the image, builds the image and finally uploads it to Docker Hub.

Figure 4: Layers of a image

DockerHub

Docker is about running random crap from the internet as root on your host.

Anyone has the chance to create and store images into the Docker Hub registry by first creating a descriptor file called Dockerfile. This descriptor describes what software packages will be within the image, builds the image and finally uploads it to Docker Hub.

Figure 4: Risk of user images

To extract the information from the Docker images hosted in Docker Hub, we performed a search over its free text box to obtain all the Docker images. In February 2018, this search returned 1,360,510 Docker repositories and 4,609,643 images composed of 4,593,602 community images and 14,841 official images. The total size of these images is 53.47 PB.

After that, we use the tool from the Clair project to detect vulnerabilities within each downloaded image.