Frederick National Laboratory
for Cancer Research
sponsored by the National Cancer Institute

Resources for research on The Cancer Imaging Archive; a catalog of enhanced tools for downloading, visualizing and analyzing TCIA data

Justin Kirby

Technical Director, Cancer Imaging Informatics Lab

Email: justin.kirby@nih.gov

DEPARTMENT OF HEALTH AND HUMAN SERVICES • National Institutes of Health • National Cancer Institute

Frederick National Laboratory is a Federally Funded Research and Development Center operated by Leidos Biomedical Research, Inc., for the National Cancer Institute


Submit Your Data
Access The Data
Help
THE CANCER IMAGING ARCHIVE
About Us
Research Activities
News
NIH
NATIONAL CANCER INSTITUTE
DCTD Division of Cancer Treatment & Diagnosis
CIP Cancer Imaging Program
Frederick National Laboratory
for Cancer Research
sponsored by the National Cancer Institute
UAMS
UNIVERSITY OF ARKANSAS
FOR MEDICAL SCIENCES
Welcome to The Cancer Imaging Archive
The Cancer Imaging Archive (TCIA) is a service which de-identifies and hosts a large archive of medical images of cancer accessible for public download.
SUBMIT YOUR DATA
ACCESS THE DATA

TCIA components

Data Collection Centers

- Tools and staffing to support data collection, curation, and de-identification

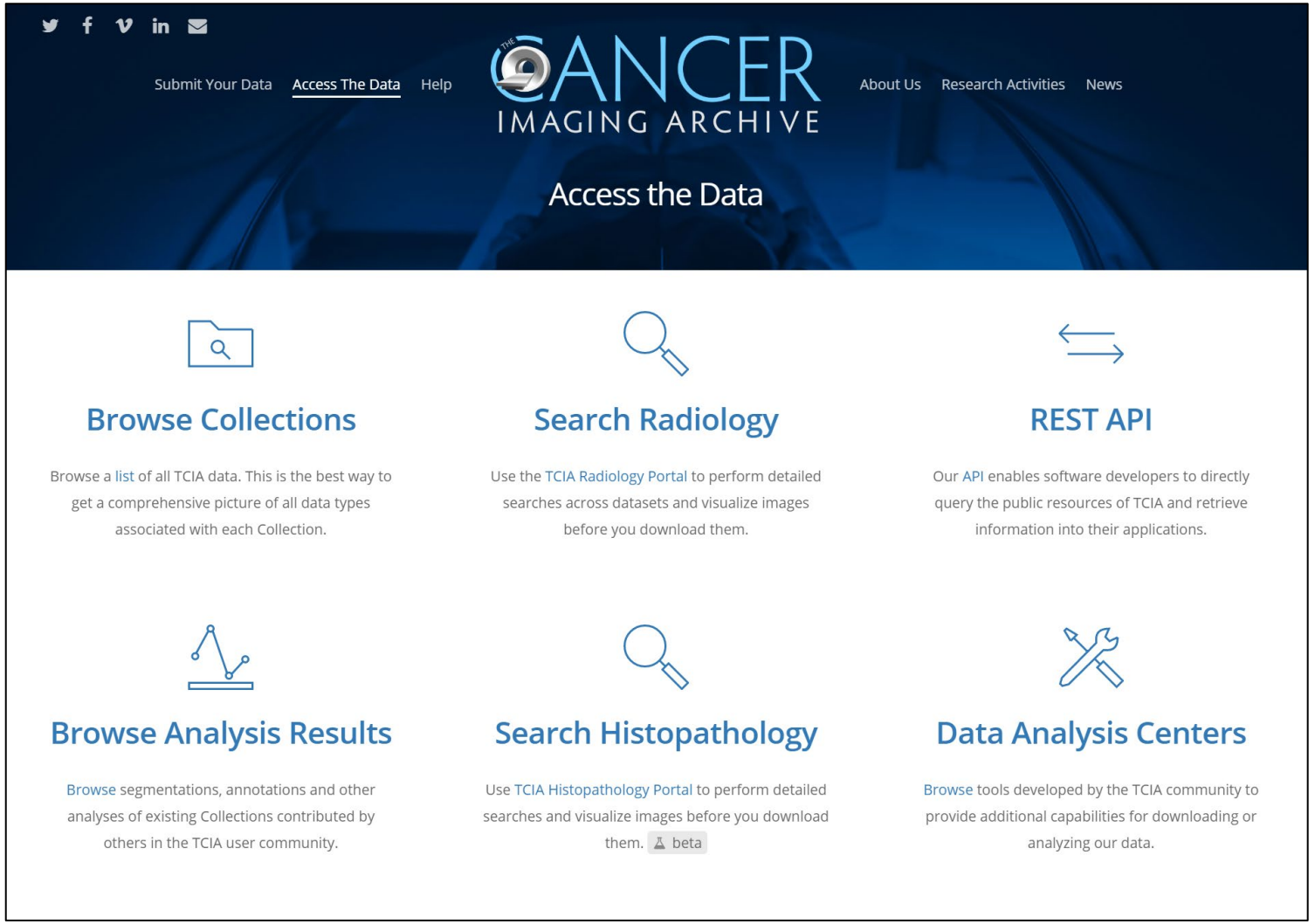
Data Sets and Access Tools

- Browse
- Filter/Search
- Visualize
- REST API

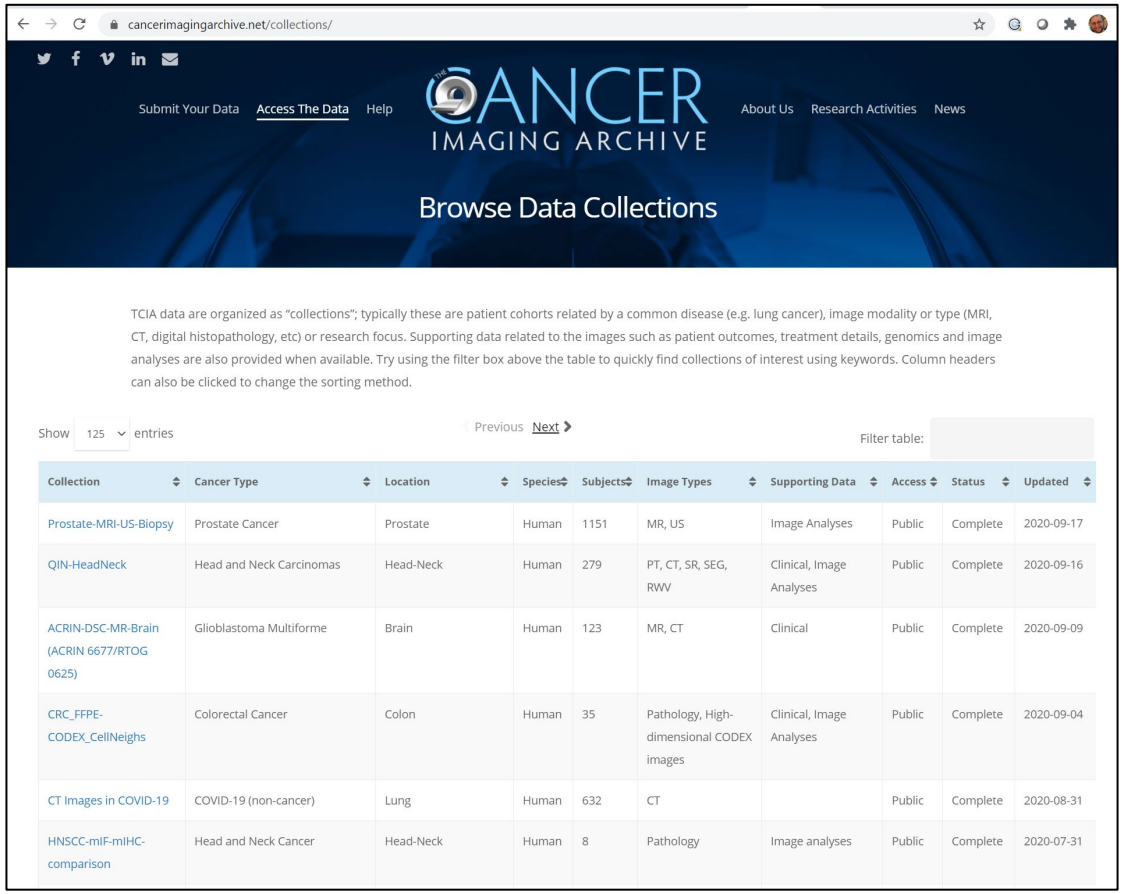
Data Analysis Centers

- 3rd party web sites or tools which use TCIA's API or mirror its data

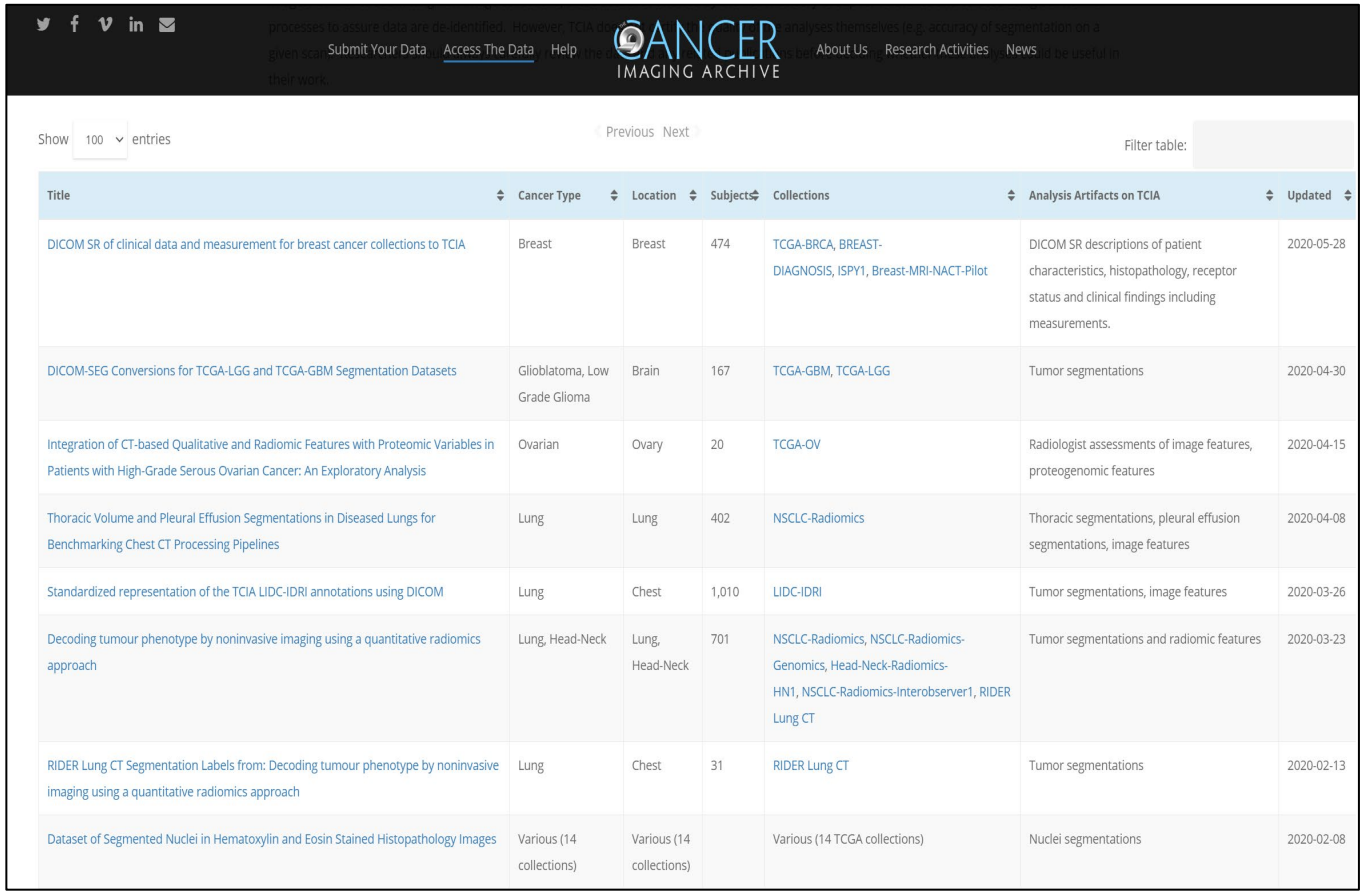
Accessing the Data



Browsing the data



Collection	Cancer Type	Location	Species	Subjects	Image Types	Supporting Data	Access	Status	Updated
Prostate-MRI-US-Biopsy	Prostate Cancer	Prostate	Human	1151	MR, US	Image Analyses	Public	Complete	2020-09-17
QIN-HeadNeck	Head and Neck Carcinomas	Head-Neck	Human	279	PT, CT, SR, SEG, RWV	Clinical, Image Analyses	Public	Complete	2020-09-16
ACRIN-DSC-MR-Brain (ACRIN 6677/RTOG 0625)	Glioblastoma Multiforme	Brain	Human	123	MR, CT	Clinical	Public	Complete	2020-09-09
CRC_FFPE-CODEX_CellNeighs	Colorectal Cancer	Colon	Human	35	Pathology, High-dimensional CODEX images	Clinical, Image Analyses	Public	Complete	2020-09-04
CT Images in COVID-19	COVID-19 (non-cancer)	Lung	Human	632	CT		Public	Complete	2020-08-31
HNSCC-mIF-mIHC-comparison	Head and Neck Cancer	Head-Neck	Human	8	Pathology	Image analyses	Public	Complete	2020-07-31



Title	Cancer Type	Location	Subjects	Collections	Analysis Artifacts on TCIA	Updated
DICOM SR of clinical data and measurement for breast cancer collections to TCIA	Breast	Breast	474	TCGA-BRCA, BREAST-DIAGNOSIS, ISPY1, Breast-MRI-NACT-Pilot	DICOM SR descriptions of patient characteristics, histopathology, receptor status and clinical findings including measurements.	2020-05-28
DICOM-SEG Conversions for TCGA-LGG and TCGA-GBM Segmentation Datasets	Glioblatoma, Low Grade Glioma	Brain	167	TCGA-GBM, TCGA-LGG	Tumor segmentations	2020-04-30
Integration of CT-based Qualitative and Radiomic Features with Proteomic Variables in Patients with High-Grade Serous Ovarian Cancer: An Exploratory Analysis	Ovarian	Ovary	20	TCGA-OV	Radiologist assessments of image features, proteogenomic features	2020-04-15
Thoracic Volume and Pleural Effusion Segmentations in Diseased Lungs for Benchmarking Chest CT Processing Pipelines	Lung	Lung	402	NSCLC-Radiomics	Thoracic segmentations, pleural effusion segmentations, image features	2020-04-08
Standardized representation of the TCIA LIDC-IDRI annotations using DICOM	Lung	Chest	1,010	LIDC-IDRI	Tumor segmentations, image features	2020-03-26
Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach	Lung, Head-Neck	Lung, Head-Neck	701	NSCLC-Radiomics, NSCLC-Radiomics-Genomics, Head-Neck-Radiomics-HN1, NSCLC-Radiomics-Interobserver1, RIDER Lung CT	Tumor segmentations and radiomic features	2020-03-23
RIDER Lung CT Segmentation Labels from: Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach	Lung	Chest	31	RIDER Lung CT	Tumor segmentations	2020-02-13
Dataset of Segmented Nuclei in Hematoxylin and Eosin Stained Histopathology Images	Various (14 collections)	Various (14 collections)		Various (14 TCGA collections)	Nuclei segmentations	2020-02-08

Radiology Downloads & Data Portal

CPTAC-GBM

Created by Tracy Nolan, last modified by natasha honomichl on Feb 14, 2020

Summary

This collection contains subjects from the National Cancer Institute's Clinical Proteomic Tumor Analysis Consortium Glioblastoma Multiforme (CPTAC-GBI molecular basis of cancer through the application of large-scale proteome and genome analysis, or proteogenomics. Radiology and pathology images fr by The Cancer Imaging Archive to enable researchers to investigate cancer phenotypes which may correlate to corresponding proteomic, genomic and cl

CPTAC Phase 3 collects data from ten cancer types. In TCIA, imaging from each cancer type will be contained in its own TCIA Collection, with the collectic available on TCIA each quarter as it is collected. A summary of CPTAC Phase 3 imaging efforts can be found on the CPTAC Imaging Proteomics page.

Radiology imaging is collected from standard of care imaging performed on patients immediately before the pathological diagnosis, and from follow-up heterogeneous in terms of scanner modalities, manufacturers and acquisition protocols. Pathology imaging is collected as part of the CPTAC qualificatior

CPTAC Imaging Special Interest Group

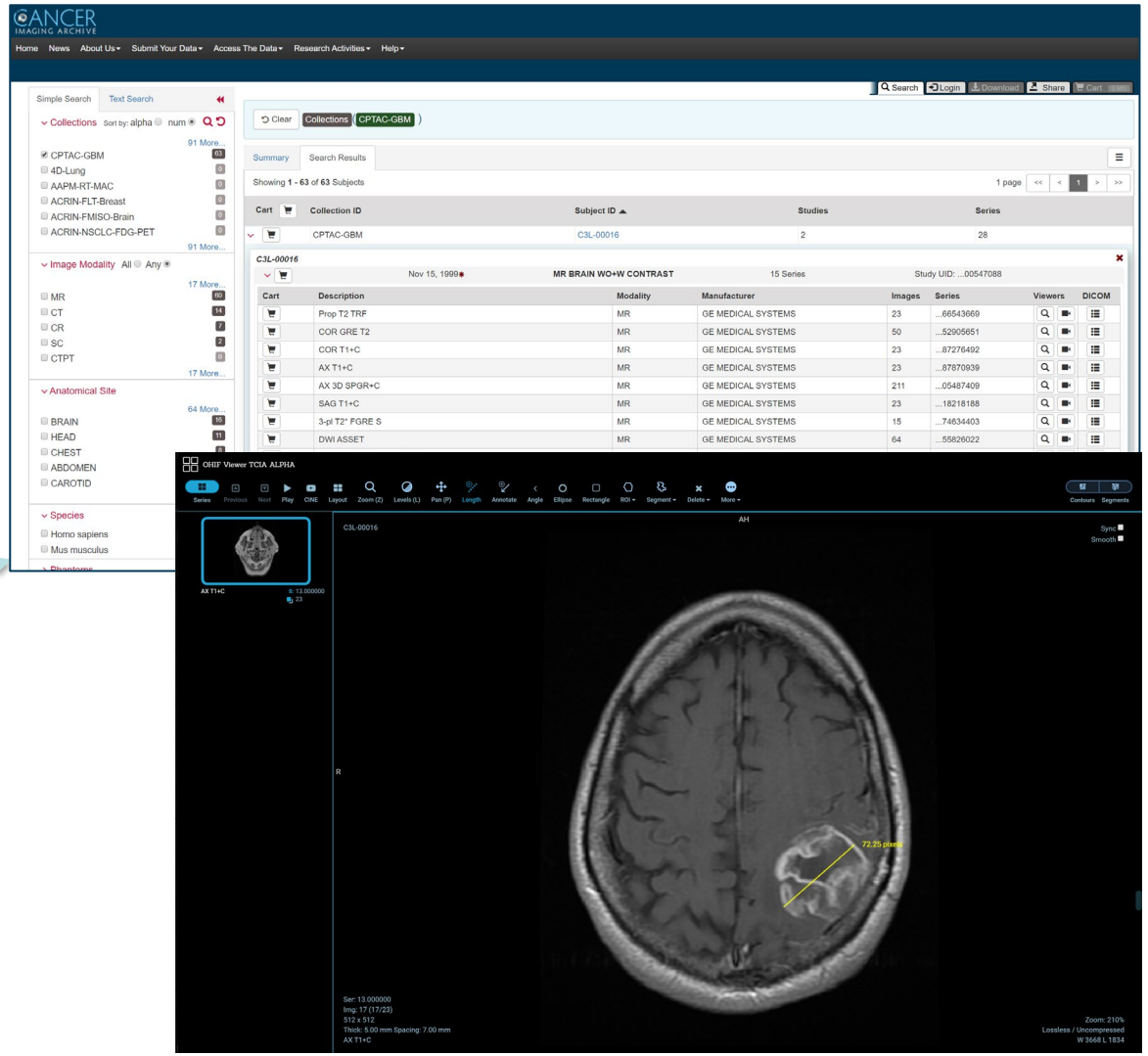
You can join the CPTAC Imaging Special Interest Group to be notified of webinars & data releases, collaborate on common data wrangling tasks and seek webinars such as slide decks and video recordings can be found on the CPTAC SIG Webinars page.

Data Access | Detailed Description | Citations & Data Usage Policy | Versions

Data Access

Click the **Download** button to save a ".tcia" manifest file to your computer, which you must open with the NBIA Data Retriever. Click the **Search** button t download a subset of its contents.

Data Type	Download all or Query/Filter
Images (DICOM, 39.4 GB)	Download Search
Tissue Slide Images (SVS, 87 GB)	Download Search
Clinical Data API (JSON - more info)	Download
Discovery Study Proteomics/Clinical Data (external)	• CPTAC Data Portal (Georgetown) • Proteomic Data Commons
Genomics/Clinical Data (External)	Genomic Data Commons



Pathology Downloads & Data Portal

CPTAC-GBM

Created by Tracy Nolan, last modified by natasha honomichl on Feb 14, 2020

Summary

This collection contains subjects from the National Cancer Institute's Clinical Proteomic Tumor Analysis Consortium Glioblastoma Multiforme (CPTAC-GBI molecular basis of cancer through the application of large-scale proteome and genome analysis, or proteogenomics. Radiology and pathology images fr by The Cancer Imaging Archive to enable researchers to investigate cancer phenotypes which may correlate to corresponding proteomic, genomic and cl

CPTAC Phase 3 collects data from ten cancer types. In TCIA, imaging from each cancer type will be contained in its own TCIA Collection, with the collectic available on TCIA each quarter as it is collected. A summary of CPTAC Phase 3 imaging efforts can be found on the CPTAC Imaging Proteomics page.

Radiology imaging is collected from standard of care imaging performed on patients immediately before the pathological diagnosis, and from follow-up heterogeneous in terms of scanner modalities, manufacturers and acquisition protocols. Pathology imaging is collected as part of the CPTAC qualificatior

CPTAC Imaging Special Interest Group

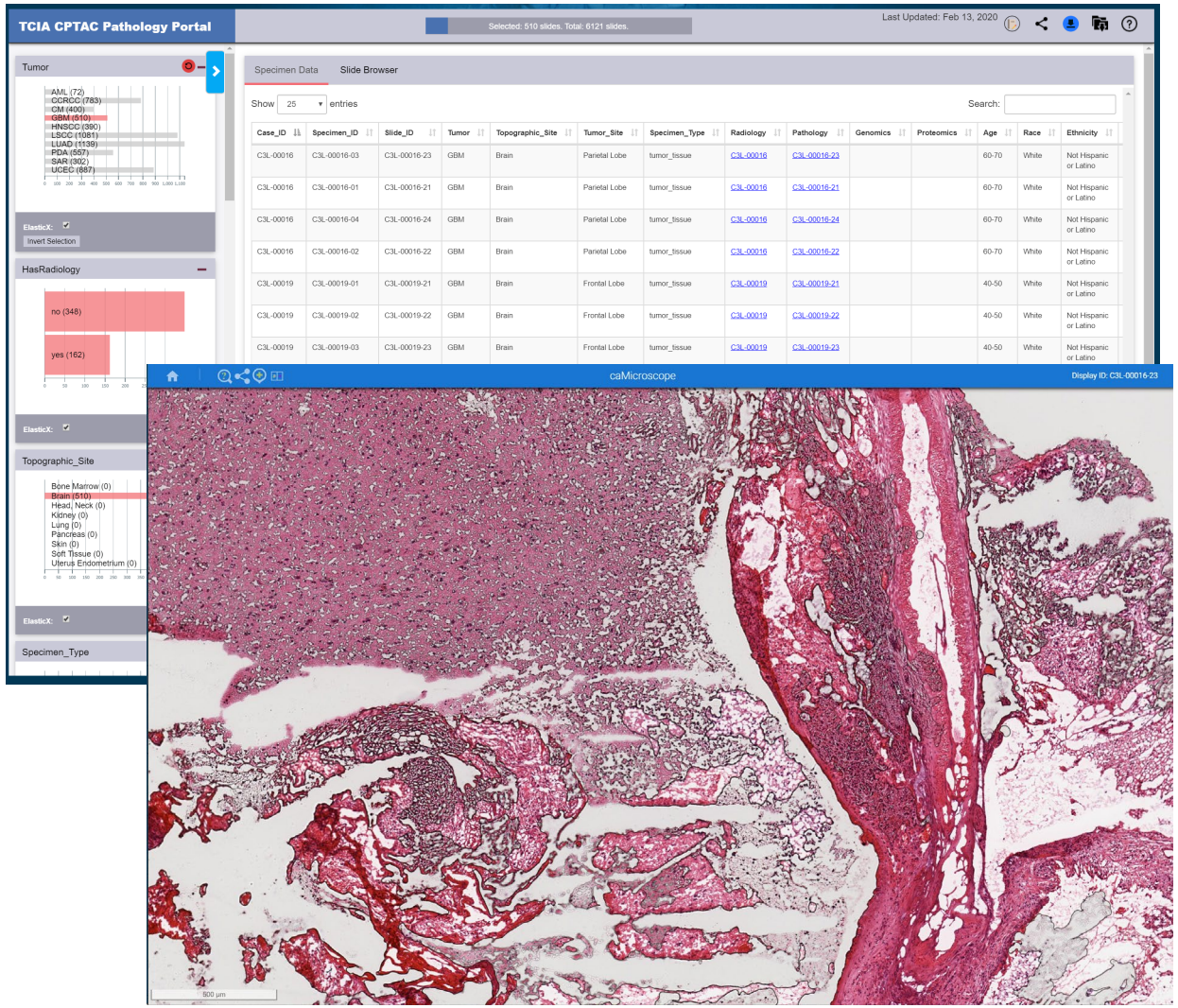
You can join the CPTAC Imaging Special Interest Group to be notified of webinars & data releases, collaborate on common data wrangling tasks and seek webinars such as slide decks and video recordings can be found on the CPTAC SIG Webinars page.

Data Access | Detailed Description | Citations & Data Usage Policy | Versions

Data Access

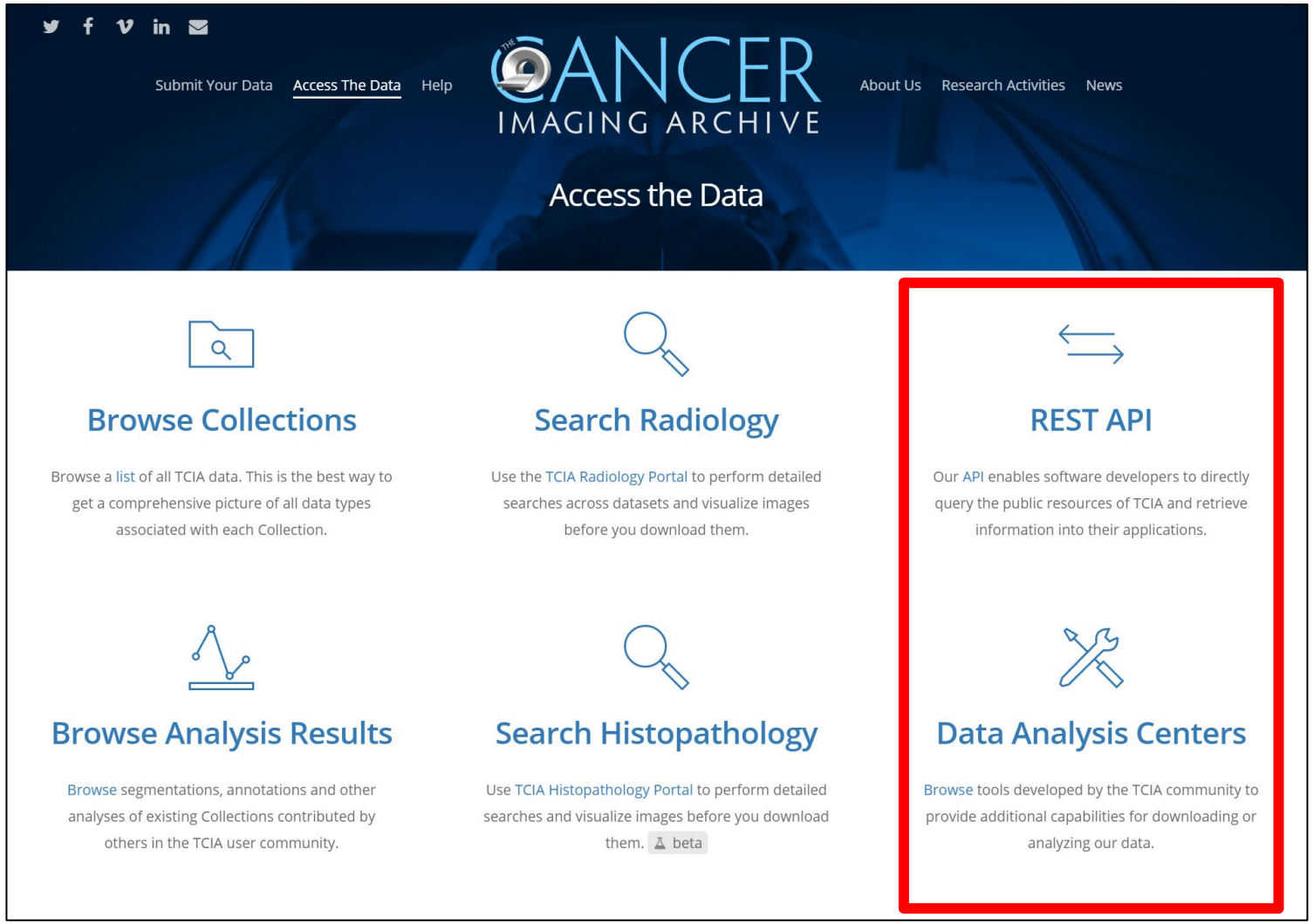
Click the **Download** button to save a ".tcia" manifest file to your computer, which you must open with the NBIA Data Retriever. Click the **Search** button t download a subset of its contents.

Data Type	Download all or Query/Filter
Images (DICOM, 39.4 GB)	Download Search
Tissue Slide Images (SVS, 87 GB)	Download Search
Clinical Data API (JSON - more info)	Download
Discovery Study Proteomics/Clinical Data (external)	• CPTAC Data Portal (Georgetown) • Proteomic Data Commons
Genomics/Clinical Data (External)	Genomic Data Commons



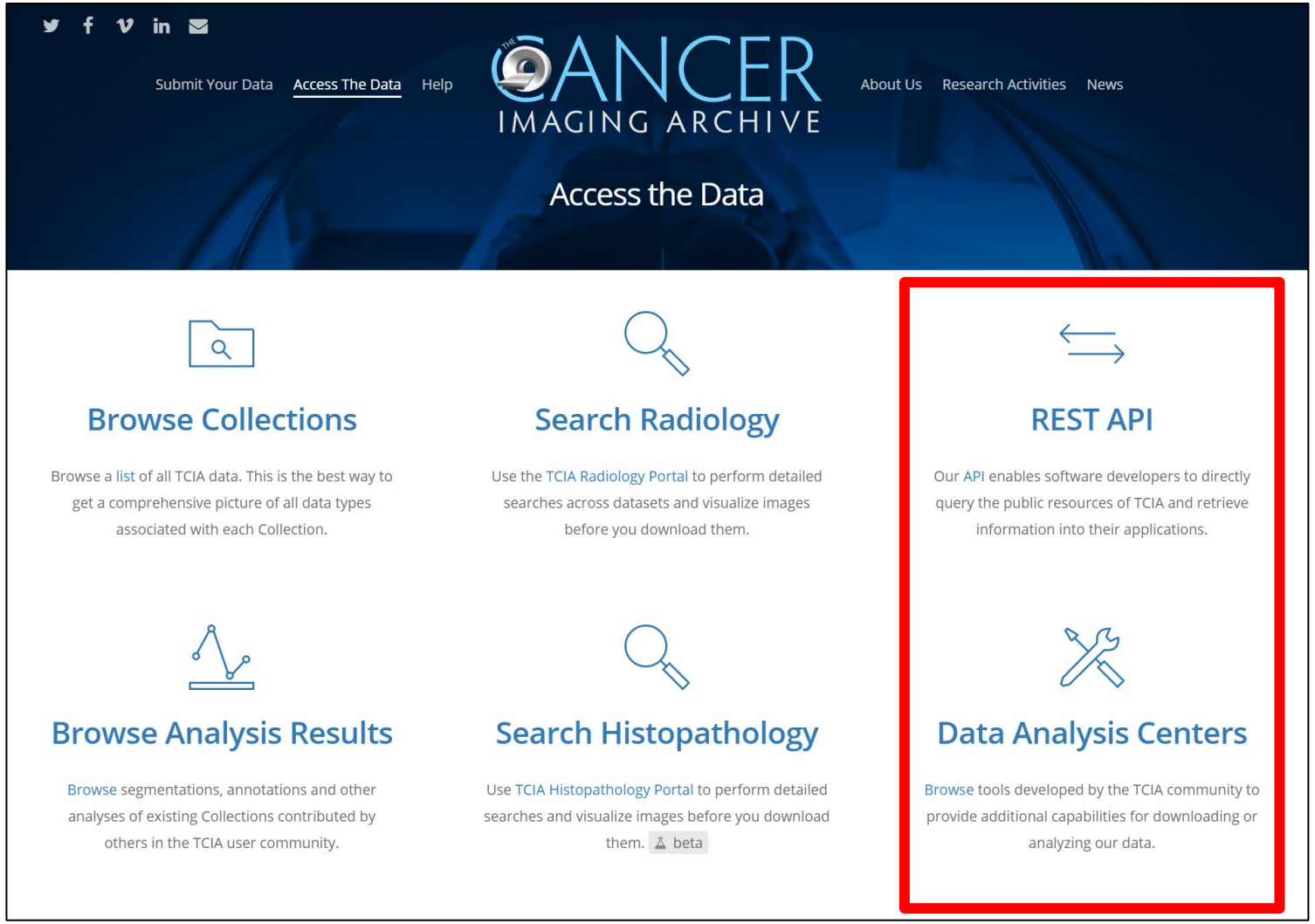
Extending the utility of TCIA data

- API enables software developers to access data directly without a web browser
- Data Analysis Centers provide a mechanism to alert others about tools/sites that provide additional ways to access TCIA data
  - Tools that use the API
  - Sites mirroring TCIA datasets



Data Analysis Center Categories

- Tools you can install yourself
  - Desktop applications
  - Web applications
- Code sharing
  - Github repositories
  - Coding tutorials/notebooks
- 3rd Party Sites/Services
  - Groups hosting copies of our datasets their own websites



Data Analysis Center Catalog

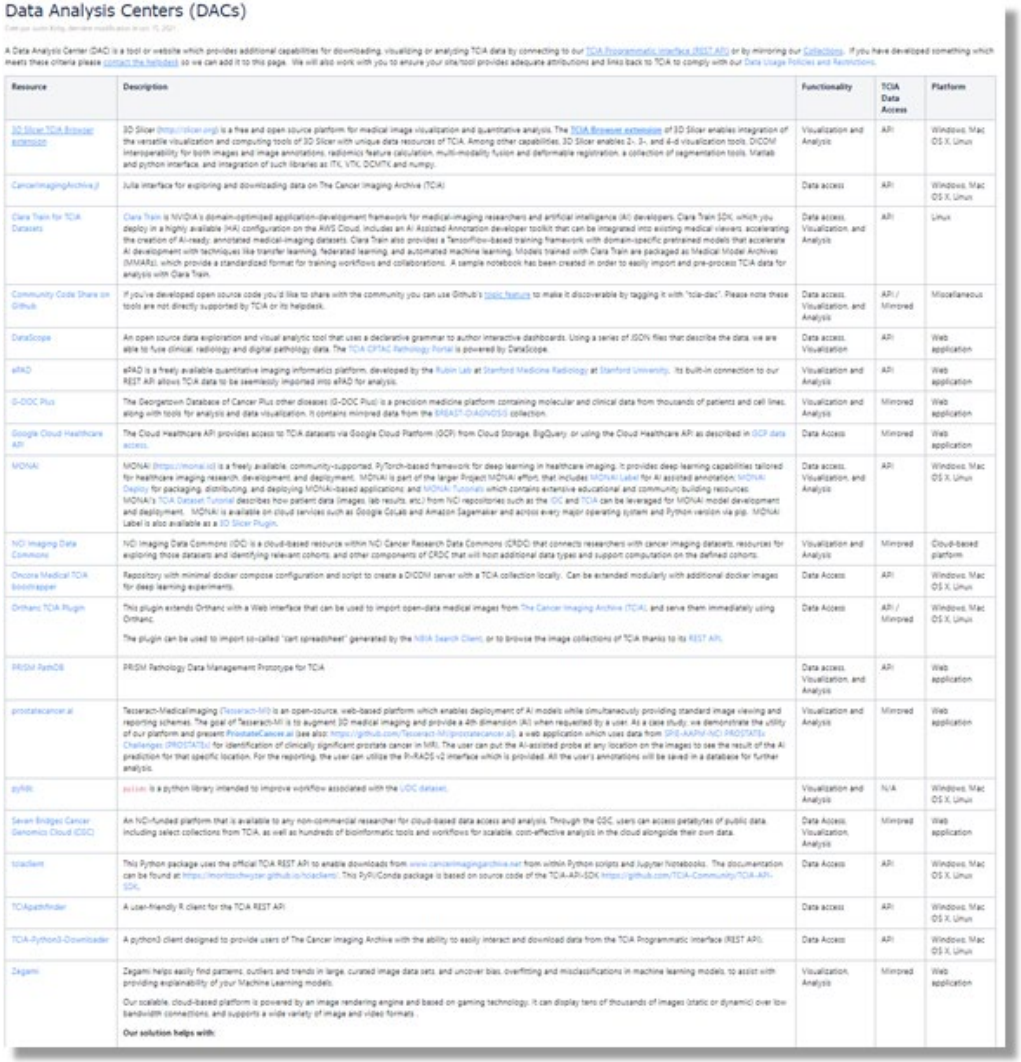
Data Analysis Centers (DACs)

Created by Justin Kirby, last modified on Oct 15, 2021

A Data Analysis Center (DAC) is a tool or website which provides additional capabilities for downloading, visualizing or analyzing TCIA data by connecting to our TCIA Programmatic Interface (REST API) or by mirroring our Collections. If you have developed something which meets these criteria please contact the helpdesk so we can add it to this page. We will also work with you to ensure your site/tool provides adequate attributions and links back to TCIA to comply with our Data Usage Policies and Restrictions.

Resource	Description	Functionality	TCIA Data Access	Platform
3D Slicer TCIA Browser extension	3D Slicer (http://slicer.org) is a free and open source platform for medical image visualization and quantitative analysis. The **TCIA Browser extension** of 3D Slicer enables integration of the versatile visualization and computing tools of 3D Slicer with unique data resources of TCIA. Among other capabilities, 3D Slicer enables 2-, 3-, and 4-d visualization tools, DICOM interoperability for both images and image annotations, radiomics feature calculation, multi-modality fusion and deformable registration, a collection of segmentation tools, Matlab and python interface, and integration of such libraries as ITK, VTK, DCMTK and numpy.	Visualization and Analysis	API	Windows, Mac OS X, Linux
CancerImagingArchive.jl	Julia interface for exploring and downloading data on The Cancer Imaging Archive (TCIA)	Data access	API	Windows, Mac OS X, Linux
Clara Train for TCIA Datasets	Clara Train is NVIDIA's domain-optimized application-development framework for medical-imaging researchers and artificial intelligence (AI) developers. Clara Train SDK, which you deploy in a highly available (HA) configuration on the AWS Cloud, includes an AI Assisted Annotation developer toolkit that can be integrated into existing medical viewers, accelerating the creation of AI-ready, annotated medical-imaging datasets. Clara Train also provides a TensorFlow-based training framework with domain-specific pretrained models that accelerate AI development with techniques like transfer learning, federated learning, and automated machine	Data access, Visualization, and Analysis	API	Linux

DAC Catalog – Over 20 projects & Code Repositories



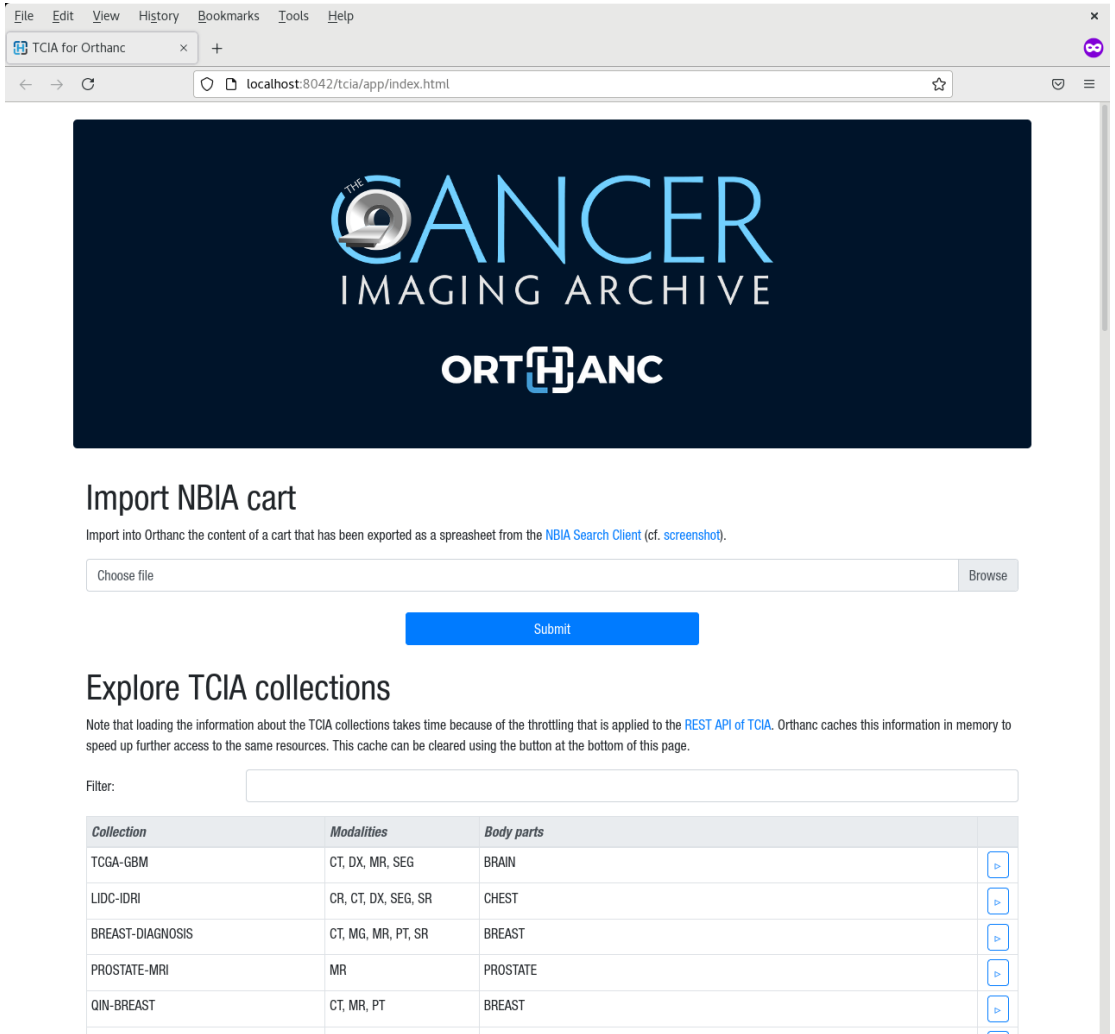
Data Analysis Centers (DACs)

A Data Analysis Center (DAC) is a tool or website which provides additional capabilities for downloading, visualizing or analyzing TCIA data by connecting to our TCIA Programmatic Interface (REST API) or by mirroring our Collections. If you have developed something which meets these criteria please contact the helpdesk so we can add it to this page. We will also work with you to ensure your site/tool provides adequate attributions and links back to TCIA to comply with our Data Usage Policies and Restrictions.

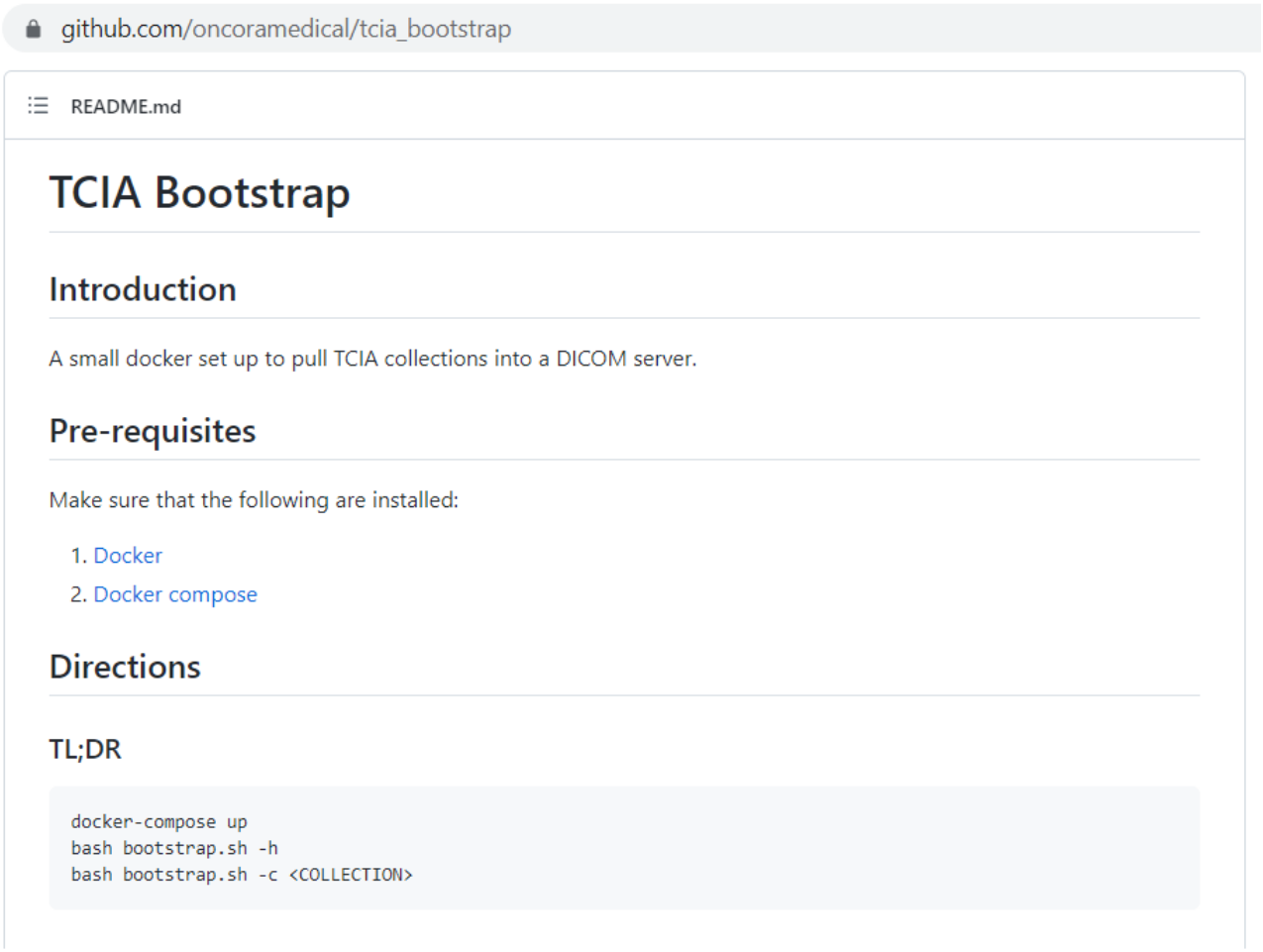
Resource	Description	Functionality	TCIA Data Access	Platform
3D Slicer TCIA Browser extension	3D Slicer (http://slicer.org) is a free and open source platform for medical image visualization and quantitative analysis. The TCIA Browser extension of 3D Slicer enables integration of the versatile visualization and computing tools of 3D Slicer with unique data resources of TCIA. Among other capabilities, 3D Slicer enables 2-, 3-, and 4-d visualization tools. DICOM interoperability for both images and image annotations, radiomics feature calculation, multi-modality fusion and deformable registration, a collection of segmentation tools. Matlab and python interface, and integration of such libraries as ITK, VTK, DCMTK and numpy.	Visualization and Analysis	API	Windows, Mac OS X, Linux
CancerImagingArchive.jl	Julia interface for exploring and downloading data on The Cancer Imaging Archive (TCIA)	Data access	API	Windows, Mac OS X, Linux
Clara Train for TCIA Datasets	Clara Train is NVIDIA's domain-optimized application-development framework for medical-imaging researchers and artificial intelligence (AI) developers. Clara Train SDK, which you deploy in a highly available (HA) configuration on the AWS Cloud, includes an AI-Assisted Annotation developer toolkit that can be integrated into existing medical viewers, accelerating the creation of AI-ready, annotated medical-imaging datasets. Clara Train also provides a TensorFlow-based training framework with domain-specific pretrained models that accelerate AI development with techniques like transfer learning, federated learning, and automated machine learning. Models trained with Clara Train are packaged as Medical Model Archives (MMARs), which provide a standardized format for training workflows and collaborations. A sample notebook has been created in order to easily import and pre-process TCIA data for analysis with Clara Train.	Data access, Visualization, and Analysis	API	Linux
Community Code Share on Github	If you've developed open source code you'd like to share with the community you can use Github's topic feature to make it discoverable by tagging it with "tcia-dac". Please note these tools are not directly supported by TCIA or its helpdesk.	Data access, Visualization, and Analysis	API / Mirrored	Miscellaneous
DataScope	An open source data exploration and visual analytic tool that uses a declarative grammar to author interactive dashboards. Using a series of JSON files that describe the data, we are able to fuse clinical, radiology and digital pathology data. The TCIA CPTAC Pathology Portal is powered by DataScope.	Data access, Visualization	API	Web application
eRAD	eRAD is a freely available quantitative imaging informatics platform, developed by the Rubin Lab at Stanford Medicine Radiology at Stanford University. Its built-in connection to our REST API allows TCIA data to be seamlessly imported into eRAD for analysis.	Visualization and Analysis	API	Web application
G-DOC Plus	The Georgetown Database of Cancer Plus other diseases (G-DOC Plus) is a precision medicine platform containing molecular and clinical data from thousands of patients and cell lines, along with tools for analysis and data visualization. It contains mirrored data from the BREAST-DIAGNOSIS collection.	Visualization and Analysis	Mirrored	Web application
Google Cloud Healthcare API	The Cloud Healthcare API provides access to TCIA datasets via Google Cloud Platform (GCP) from Cloud Storage, BigQuery, or using the Cloud Healthcare API as described in GCP data access.	Data Access	Mirrored	Web application
MONAI	MONAI (https://monai.io) is a freely available, community-supported, PyTorch-based framework for deep learning in healthcare imaging. It provides deep learning capabilities tailored for healthcare imaging research, development, and deployment. MONAI is part of the larger Project MONAI effort, that includes MONAI Label for AI assisted annotation, MONAI Deploy for packaging, distributing, and deploying MONAI-based applications, and MONAI Tutorials which contains extensive educational and community building resources. MONAI's TCIA Dataset Tutorial describes how patient data (images, lab results, etc.) from NCI repositories such as the IDC and TCIA can be leveraged for MONAI model development and deployment. MONAI is available on cloud services such as Google Colab and Amazon Sagemaker and across every major operating system and Python version via pip. MONAI Label is also available as a 3D Slicer Plugin.	Data access, Visualization, and Analysis	API	Windows, Mac OS X, Linux
NCI Imaging Data Commons	NCI Imaging Data Commons (IDC) is a cloud-based resource within NCI Cancer Research Data Commons (CRDC) that connects researchers with cancer imaging datasets, resources for exploring those datasets and identifying relevant cohorts, and other components of CRDC that will host additional data types and support computation on the defined cohorts.	Visualization and Analysis	Mirrored	Cloud-based platform
Oncora Medical TCIA bootstrapper	Repository with minimal docker compose configuration and script to create a DICOM server with a TCIA collection locally. Can be extended modularly with additional docker images for deep learning experiments.	Data Access	API	Windows, Mac OS X, Linux
Orthanc TCIA Plugin	This plugin extends Orthanc with a Web interface that can be used to import open-data medical images from The Cancer Imaging Archive (TCIA), and serve them immediately using Orthanc. The plugin can be used to import so-called "cart spreadsheet" generated by the NBIA Search Client, or to browse the image collections of TCIA thanks to its REST API.	Data Access	API / Mirrored	Windows, Mac OS X, Linux
PRISM PathDB	PRISM Pathology Data Management Prototype for TCIA	Data access, Visualization, and Analysis	API	Web application
prostatecancer.ai	Tesseract-MedicalImaging (Tesseract-MI) is an open-source, web-based platform which enables deployment of AI models while simultaneously providing standard image viewing and reporting schemes. The goal of Tesseract-MI is to augment 3D medical imaging and provide a 4th dimension (AI) when requested by a user. As a case study, we demonstrate the utility of our platform and present ProstateCancer.ai (see also: https://github.com/Tesseract-MI/prostatecancer.ai), a web application which uses data from SPIE-AAPM-NCI PROSTATEx Challenges (PROSTATEx) for identification of clinically significant prostate cancer in MRI. The user can put the AI-assisted lesion at any location on the images to see the result of the AI prediction for that specific location. For the reporting, the user can utilize the PI-RADS v2 interface which is provided. All the user's annotations will be saved in a database for further analysis.	Visualization and Analysis	Mirrored	Web application
pylidc	pylidc is a python library intended to improve workflow associated with the LIDC dataset.	Visualization and Analysis	N/A	Windows, Mac OS X, Linux
Seven Bridges Cancer Genomics Cloud (CGC)	An NCI-funded platform that is available to any non-commercial researcher for cloud-based data access and analysis. Through the CGC, users can access petabytes of public data, including select collections from TCIA, as well as hundreds of bioinformatic tools and workflows for scalable, cost-effective analysis in the cloud alongside their own data.	Data Access, Visualization, Analysis	Mirrored	Web application
tciaclient	This Python package uses the official TCIA REST API to enable downloads from www.cancerimagingarchive.net from within Python scripts and Jupyter Notebooks. The documentation can be found at https://kirbyju.github.io/tciaclient/. This PyPI/Conda package is based on source code of the TCIA-API-SDK https://github.com/TCIA-Community/TCIA-API-SDK.	Data Access	API	Windows, Mac OS X, Linux
TCIApathfinder	A user-friendly R client for the TCIA REST API	Data access	API	Windows, Mac OS X, Linux
TCIA-Python3-Downloader	A python3 client designed to provide users of The Cancer Imaging Archive with the ability to easily interact and download data from the TCIA Programmatic Interface (REST API).	Data Access	API	Windows, Mac OS X, Linux
Zegami	Zegami helps easily find patterns, outliers and trends in large, curated image data sets, and uncover bias, overfitting and misclassifications in machine learning models, to assist with providing explainability of your Machine Learning models. Our scalable, cloud-based platform is powered by an image rendering engine and based on gaming technology. It can display tens of thousands of images (static or dynamic) over low bandwidth connections, and supports a wide variety of image and video formats. **Our solution helps with:**	Visualization, Analysis	Mirrored	Web application

ORTHANC - PACS Integration

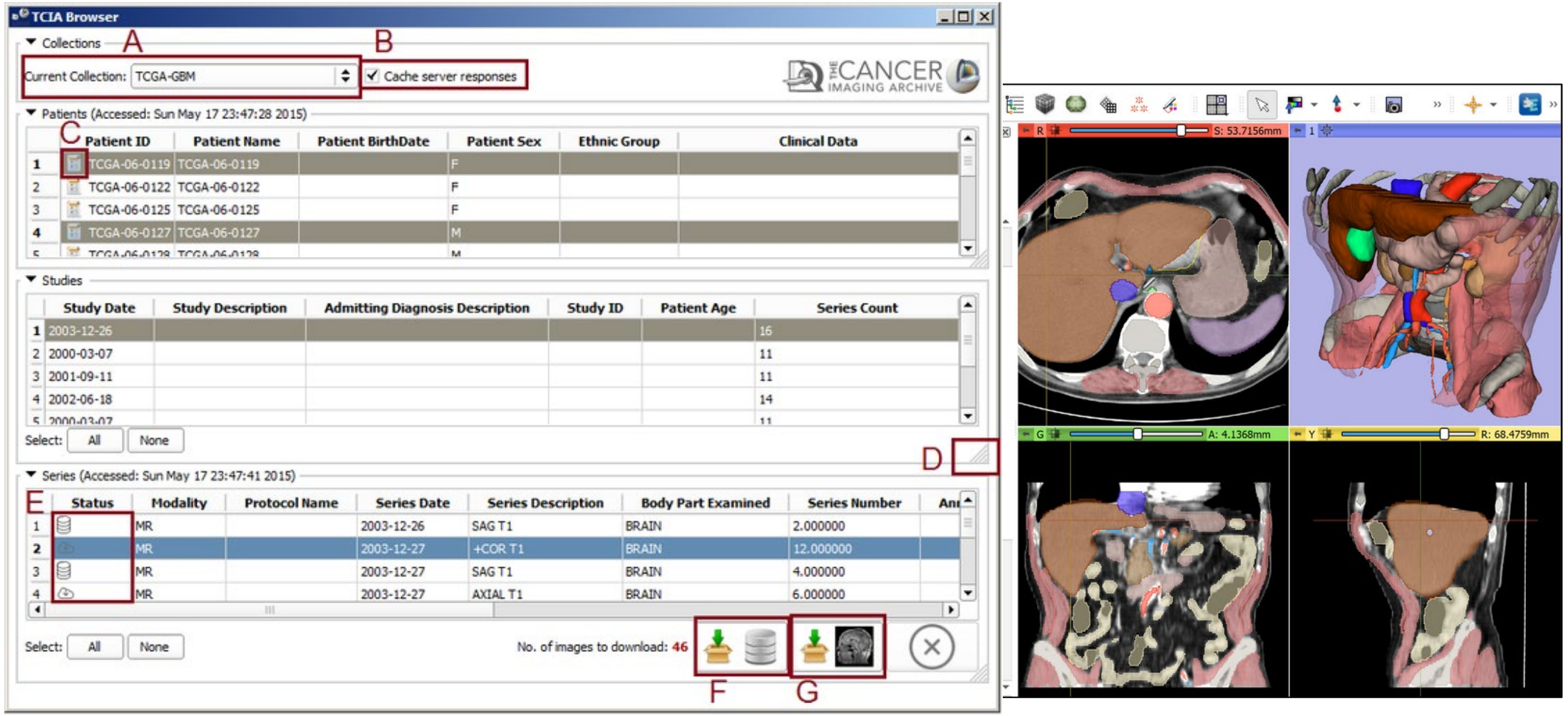
ORTHANC Official Plugin



Oncora Medical "TCIA Bootstrap" w/ ORTHANC

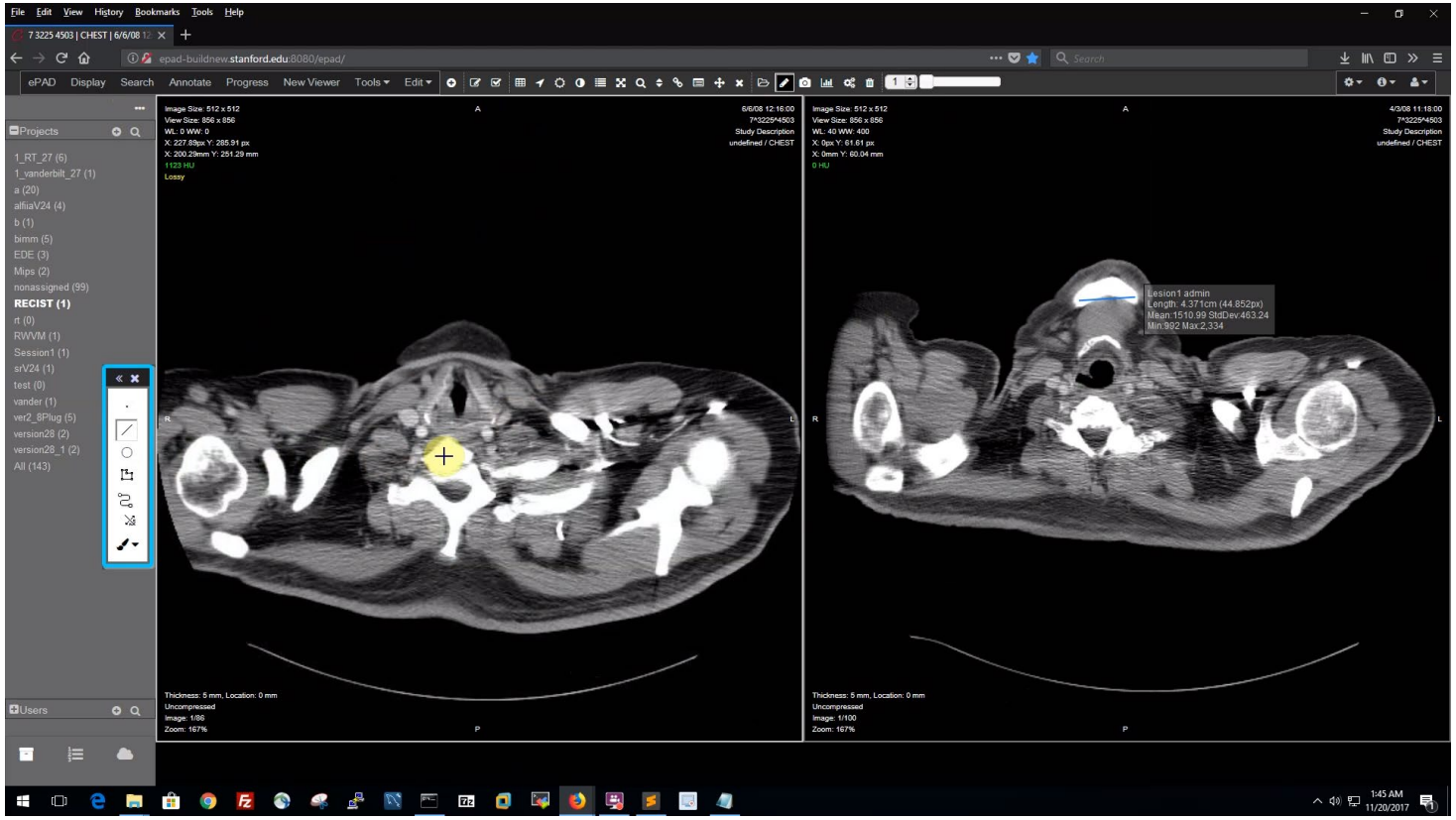


3D Slicer - Image Analysis Integration



Epad - PACS and Image Analysis Integration

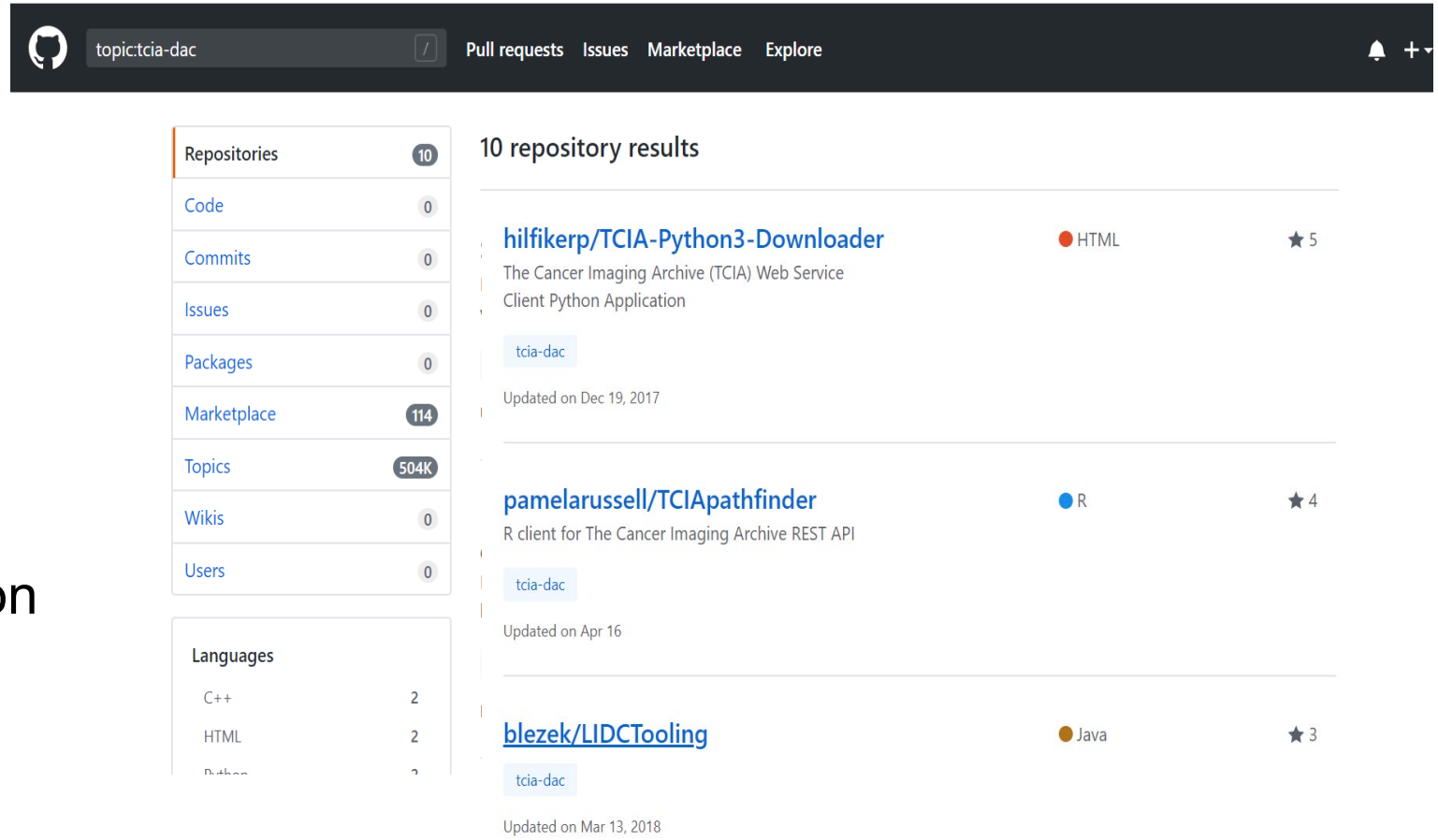
- Uses API to easily import images from TCIA into a PACS back end
- Allows visualization, annotation/segmentation, and structured data capture with case report forms and controlled ontologies



TCIA community code sharing

- API wrappers for popular programming languages
  - Python
  - R
  - Julia
- Specialized code for working with specific datasets
  - Visualizing custom annotation formats

Tag your Github repo with “tcia-dac” topic tag to appear in the list



Code Sharing – Reproducible Research

README.rst

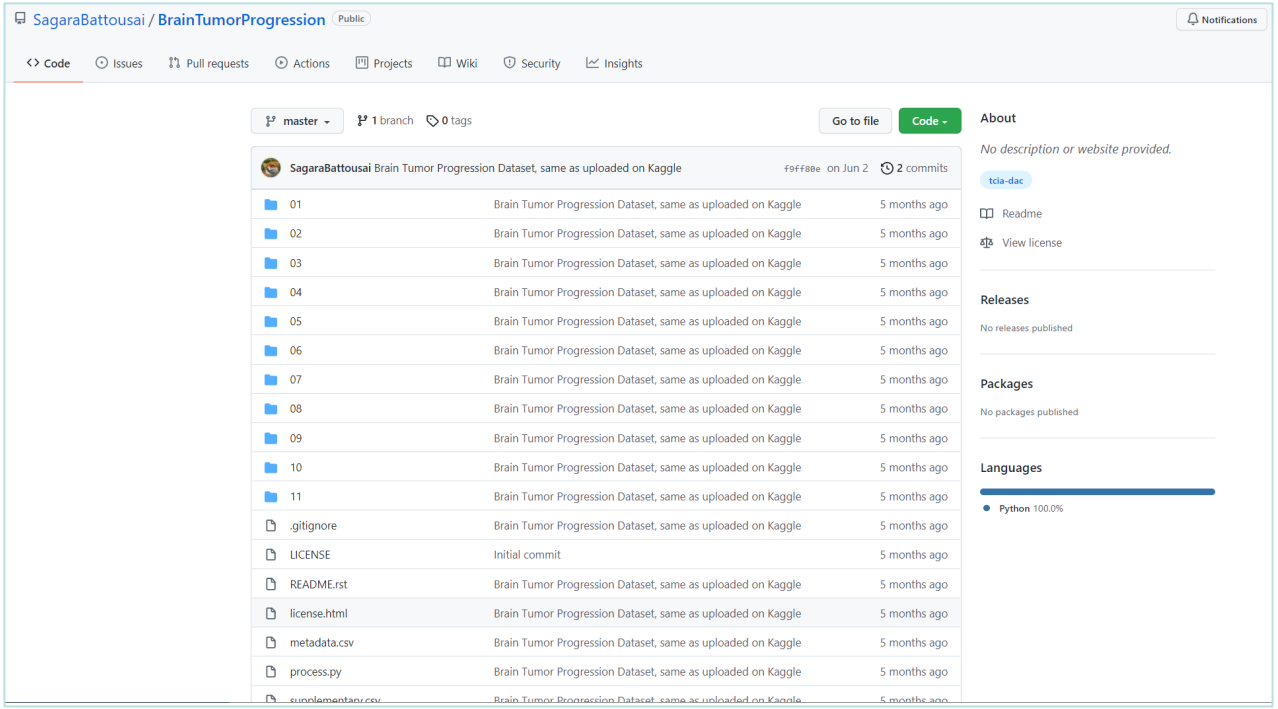
Brain Tumor Progression Dataset

Files

Directories:	Directories 01 - 11 include pairs of Pre and Post T1 MRI Scans and Masks (four per directory)
supplementary.csv:	CSV file that gives additional information for the data and includes details of what slices were removed in order to transform the original data into this dataset. This file is used by `process.py` (explained below) to modify the original dataset into this dataset.
process.py:	Python script to modify original dataset into this dataset, contains useful helper functions to use in the interpreter or script to make further modifications. This file is messy and unneeded but gives a technical description of the changes made to supplement the explanation below.
metadata.csv:	The metadata file from the TCIA downloader showing the raw data this dataset is based on/downloaded from.
licence.html:	The licence file from the TCIA downloader. This is the original licence file that is included when the data is downloaded from the source.
README.rst:	This document

Data References and Usage

This repository includes data from the Brain-Tumor-Progression collection published by The Cancer Imaging Archive. The original data has been modified in the following ways: Firstly all DICM image slices per set have been combined into one Gzipped NIfTI file, renaming the resulting NIfTI images Pre.... and Post... for the first and second scan respectivly, for both the MRI Scans and associated tumor masks resulting in four .nii.gz images. The images have all been resized to 256 X 256 using Nearest Neighbors interpolation and the number of slices was first reduces to 22, for some sets that required a few slides to be dropped and resulted in a closer relationship between Pre and Post slices for the same slice number, for others no change was required. Following that the slices were then padded with blank slides to bring the final count to 24 slices each; this was done so that three subsample convolutional layers could be composed without ending up with decimal dimensions (24 % 2^3 = 0).



Jupyter Notebooks - Coding Tutorials

NVIDIA Clara SDK Tutorial

github.com/NVIDIA/clara-train-examples/blob/master/PyTorch/NoteBooks/Data/TCIA/TCIADownloader.ipynb

Download dataset form NIH cancer imaging archive

This cancer imaging archive is a great source for medical data sets. This notebook will walk you through: 1- Selecting a dataset 2- Down
data dicom images to nifti 4- Converting data dicom seg images to nifti 5- Creating your dataset json to begin training your AI models

0. Prerequisite

0.1 Clone code

Clone code from tcia_downloader repo, move it to this directory

In [ ]:
```
!git clone https://github.com/lescientifik/tcia_downloader /claraDevDay/Data/TCIA/tmp
!mv -vn /claraDevDay/Data/TCIA/tmp/* /claraDevDay/Data/TCIA/
!mv -vn /claraDevDay/Data/TCIA/tmp/src/* /claraDevDay/Data/TCIA/src/
```

In [ ]:
```
DataRoot="/claraDevDay/Data/"
CodeRoot="/claraDevDay/Data/TCIA"
%cd $CodeRoot
!pwd
```

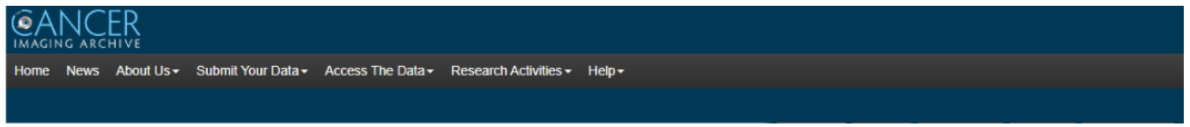
0.2 Download needed conversion tool

In [ ]:
```
! wget https://github.com/QIICR/dcmqi/releases/download/v1.2.2/dcmqi-1.2.2-linux.tar.gz && \
    tar xf dcmqi-1.2.2-linux.tar.gz && \
    cp dcmqi-1.2.2-linux/bin/segimage2itkimage /usr/local/bin/ && \
    rm -rf dcmqi-1.2.2-linux*
```

In [ ]:
```
! curl -fLO https://github.com/rordenlab/dcm2niix/releases/latest/download/dcm2niix_lnx.zip
! unzip dcm2niix_lnx.zip -d /usr/local/bin/
! chmod +x /usr/local/bin/dcm2niix
```

1. Select images to download

You can find out different studies from their site. Using the online tool, you can download list of images you would like to download.



github.com/Project-MONAI/tutorials/blob/82e1e623c2cfaad3b3dd94db537bb743dce523a6/modules/tcia_csv_processing.ipynb

Download raw DICOM series at runtime and execute transform

In this demo, we extract the series UID from loaded TCIA table and try to download the DICOM raw images with REST API, more details: https://wiki.cancerimagingarchive.net/display/Public/TCIA+REST+API+Guide

In [6]:
```
data = dataset[0]
print(f"Series UID is: {data['Series UID']}")
```

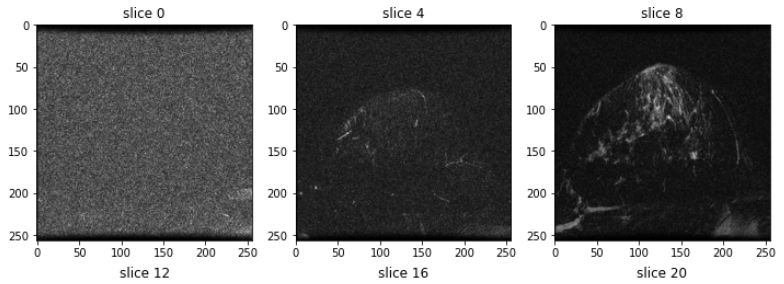
```
1.3.6.1.4.1.14519.5.2.1.7695.1700.506244151085816331158034893215.zip: 1.74MB [00:04, 379kB/s]
Downloaded: /workspace/data/medical/tcia_images/1.3.6.1.4.1.14519.5.2.1.7695.1700.506244151085816331158034893215.zip
Expected md5 is None, skip md5 check for file /workspace/data/medical/tcia_images/1.3.6.1.4.1.14519.5.2.1.7695.1700.506244151085816331158
034893215.zip.
Writing into directory: /workspace/data/medical/tcia_images/1.3.6.1.4.1.14519.5.2.1.7695.1700.506244151085816331158034893215.
Series UID is: 1.3.6.1.4.1.14519.5.2.1.7695.1700.506244151085816331158034893215
```

Plot the loaded DICOM series content

Here we access the loaded DICOM image by key `image`, then plot several slices.

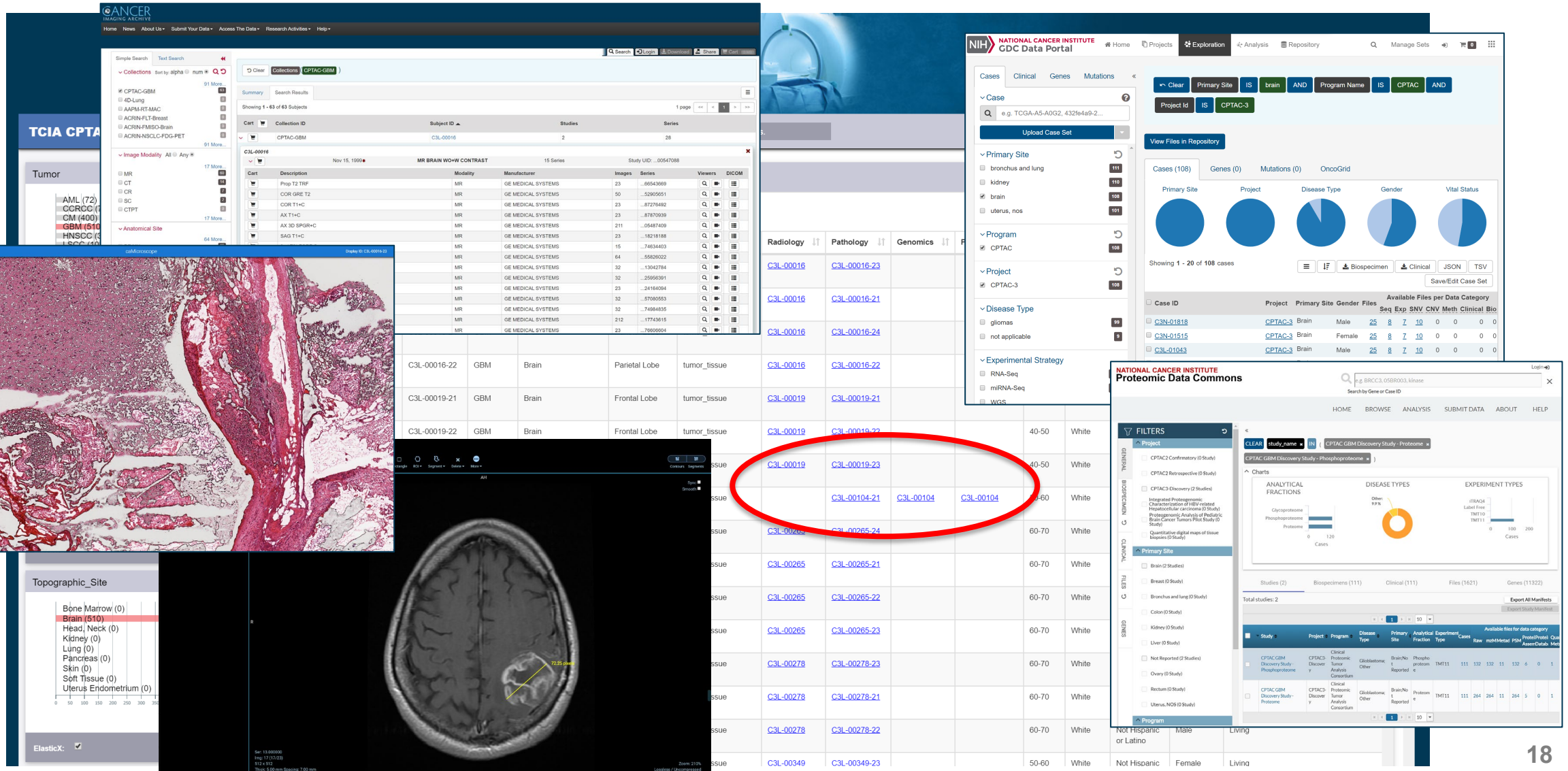
In [7]:
```
img = data["image"]
print(f"image shape: {img.shape}, min value: {img.min()}, max value: {img.max()}")

plt.figure("check", (12, 8))
for i in range(6):
    plt.subplot(2, 3, i + 1)
    plt.title(f"slice {i * 4}")
    plt.imshow(img[:, :, i * 4], cmap="gray")
plt.show()
```

```
image shape: (256, 256, 24), min value: 0.0, max value: 1624.0
```



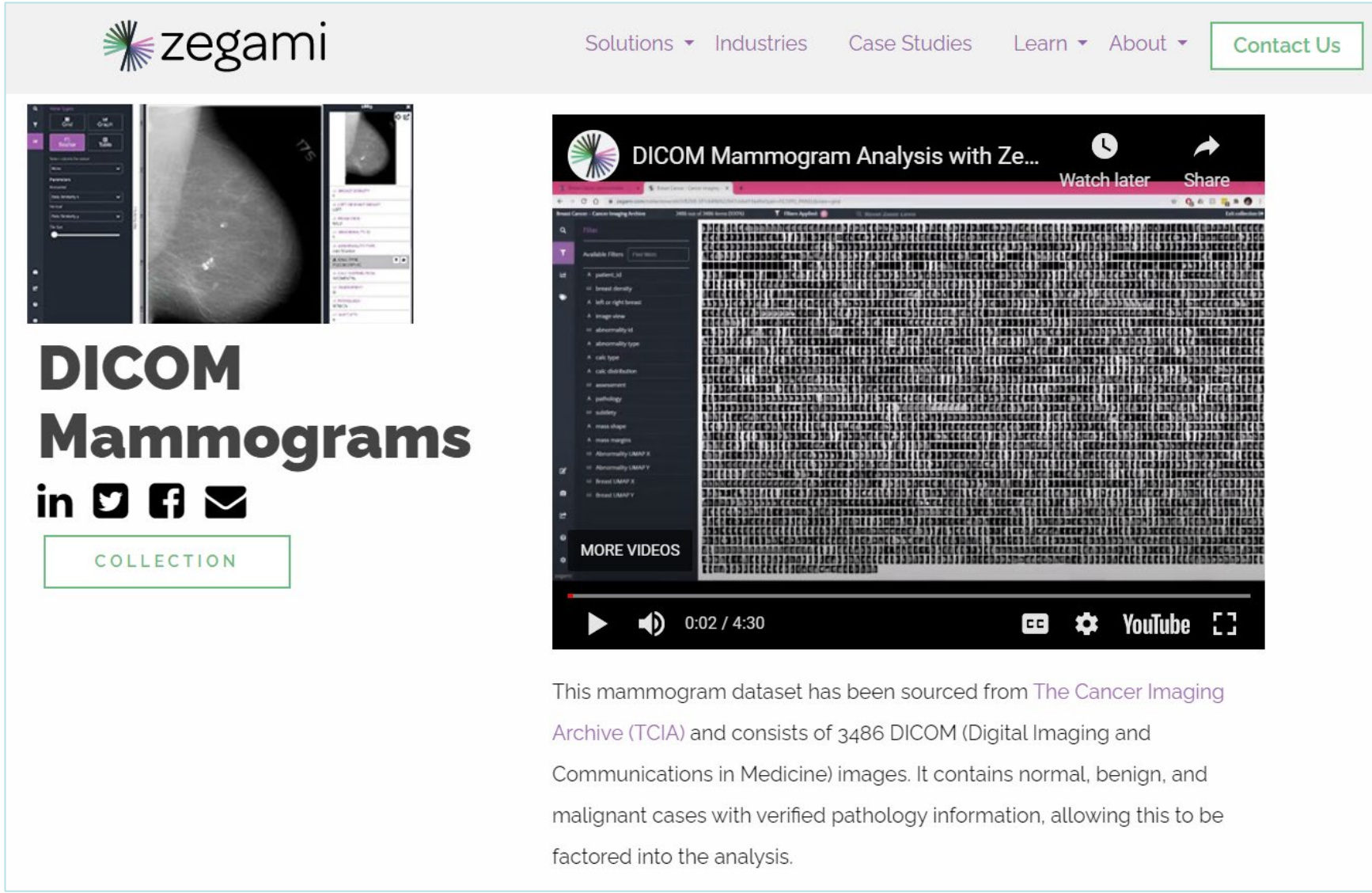
MONAI Tutorial

Datascope – Multi-modal Data Connectivity



Zegami

- Quickly visualize the data based on image characteristics (e.g. similarity) or supporting data (e.g. diagnosis) to discover patterns and outliers
- Watch their demo with the TCIA CBIS-DDSM mammogram dataset at https://youtu.be/xKJRI_Hbm50

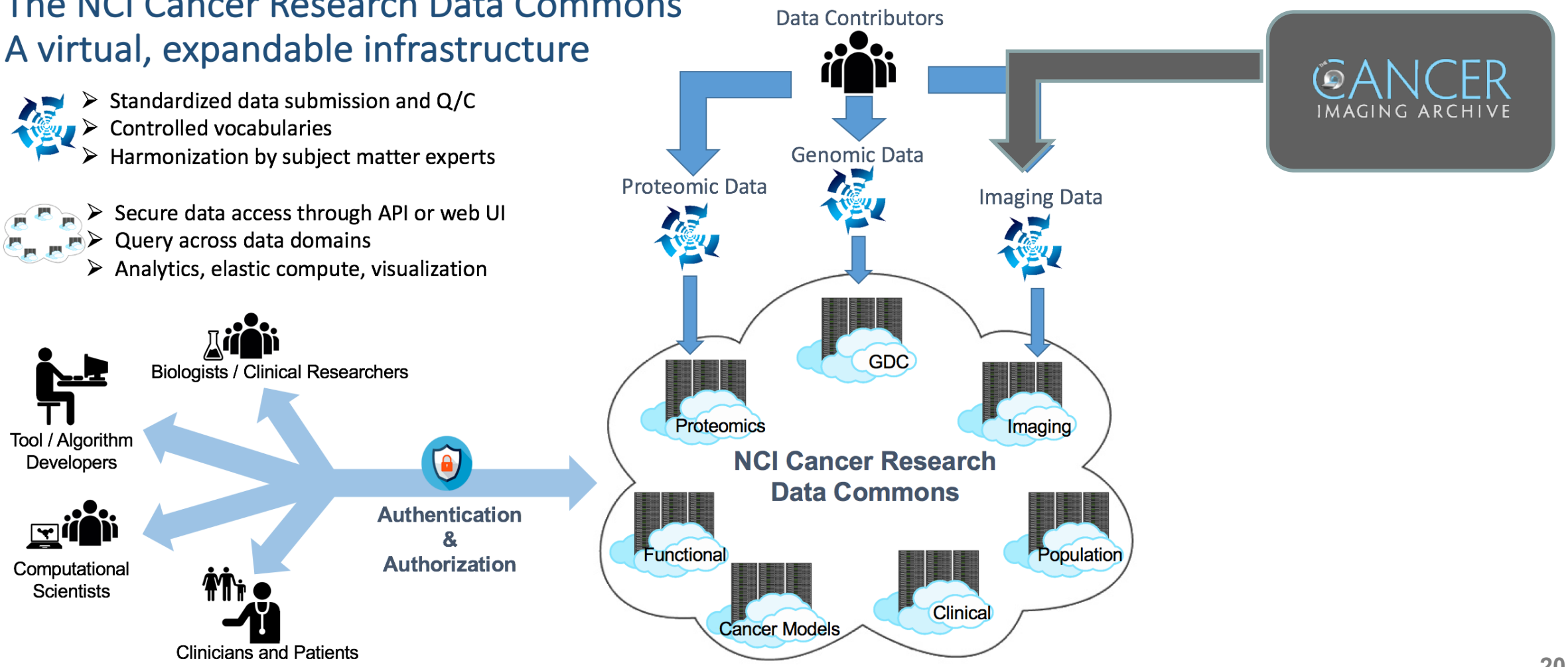


TCIA data in the NCI Data Commons cloud

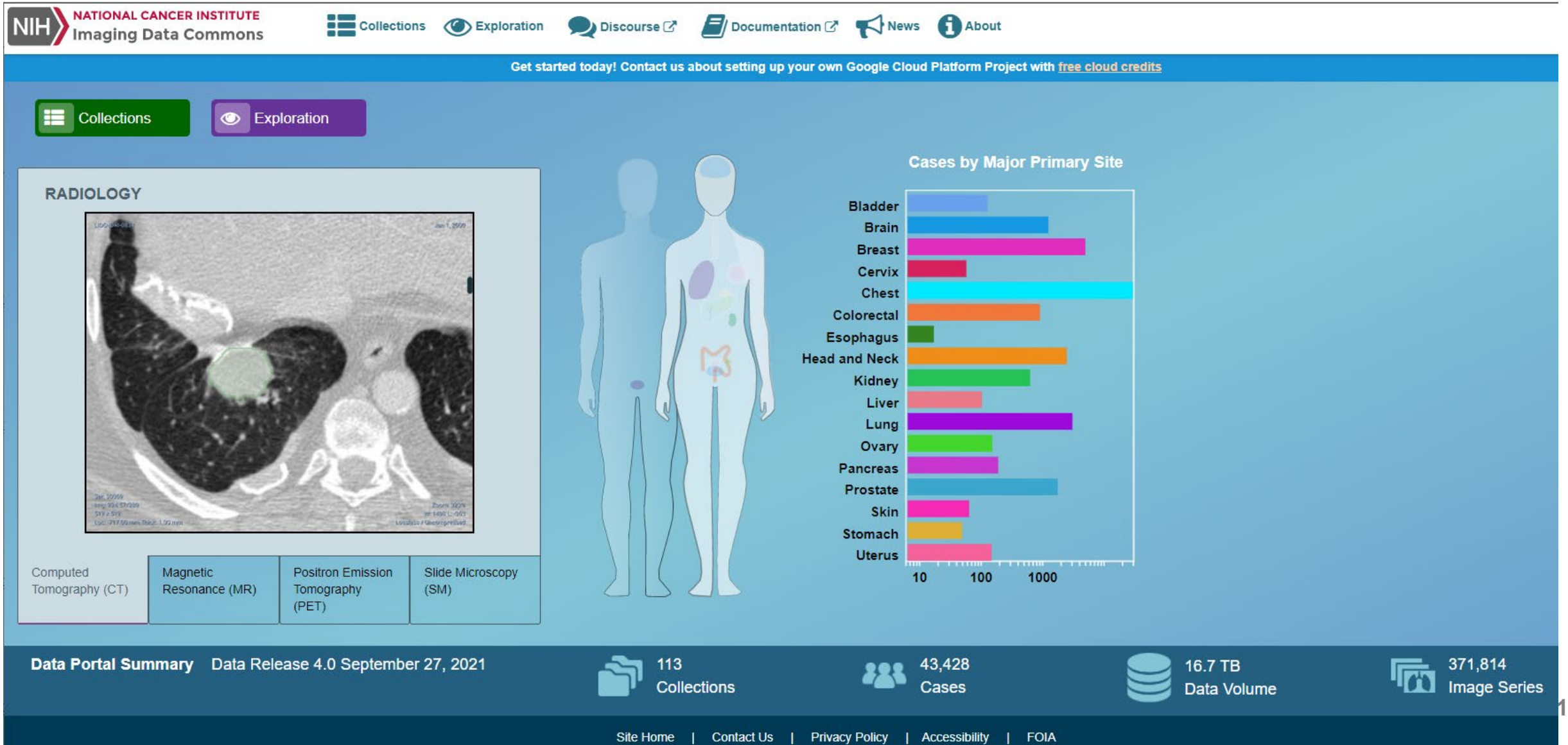
The NCI Cancer Research Data Commons
A virtual, expandable infrastructure

- Standardized data submission and Q/C
- Controlled vocabularies
- Harmonization by subject matter experts

- Secure data access through API or web UI
- Query across data domains
- Analytics, elastic compute, visualization



Imaging Data Commons

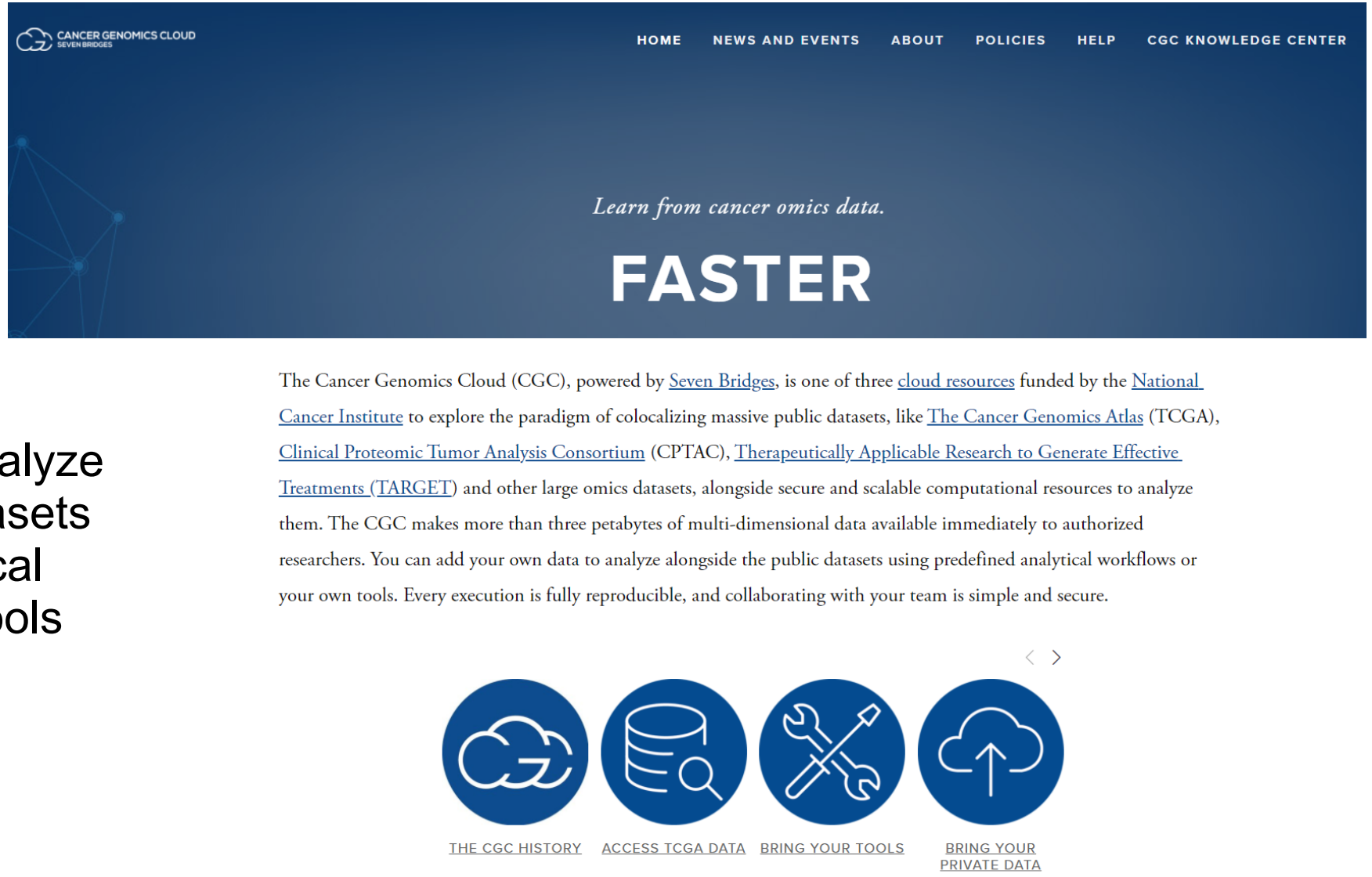


Imaging Data Commons – Use Cases

- Use the IDC portal to build a cohort of cases that are needed for your use case: see demo video - Introduction to IDC portal
- Interrogate the DICOM metadata directly from a Colab notebook (or BigQuery Console): see demo notebook - Exploration of LIDC collection
- Use Google Colab notebooks to operate on the cohort: see demo notebook - Working with IDC cohorts and the accompanying tutorial video
- Use Google DataStudio to build interactive data dashboards for your cohort: see example of using DataStudio to build a custom dashboard of IDC content
- Deploy tools on a GCP Virtual Machine to support visualization and analysis of your cohort: learn how you can launch a desktop-like VM running 3D Slicer
- Develop reproducible analysis workflows operating on public data that can be easily shared with your colleagues and referenced from your paper: see the growing list of the Colab Notebooks implementing use cases relevant in cancer imaging (such as the use of nnU-Net for segmenting lung nodules)

Seven Bridges Cancer Genomics Cloud

- Combined imaging and ‘omic data analyses for various NIH projects
- Bring your own tools to analyze CGC public datasets
- Add your own data to analyze alongside the public datasets using predefined analytical workflows or your own tools



Social Media to push new data notifications

Twitter: @TCIA_News

Facebook: The Cancer Imaging Archive (page)

LinkedIn: The Cancer Imaging Archive (group)

TCIA Acknowledgements

Fredrick National Laboratory for Cancer Research
John Freymann
Justin Kirby
Brenda Fevrier-Sullivan
Pam Angelus
Luis Cordeiro
Michelle Tacconelli

NBIA Team (NCI/FNLCR/Ellumen)
Ulli Wagner
Scott Gustafson
Qinyan Pan
Russ Rieling
Carolyn Klinger
Martin Lerner
Tin Tran
Keyvan Farahani

University of Arkansas
Dr. Fred Prior
Kirk Smith
Lawrence Tarbox
Dwayne Dobbins
Jeff Tobler
Sonya Utecht
Julie Frund
Bill Bennett
Tracy Nolan
Quasar Jarosz
Brittney Camp
Erica Bilello
Claren Freeman
Natasha Honomichl
Geri Blake
Diana Stockton
Erica Bilello
Robert Brown

Emory University
Dr. Ashish Sharma
Ryan Birmingham

Contractors
John Perry
Betty Levine

Imaging and Radiation Oncology Core
Dr. Michael Knopp
Connor Nealer