

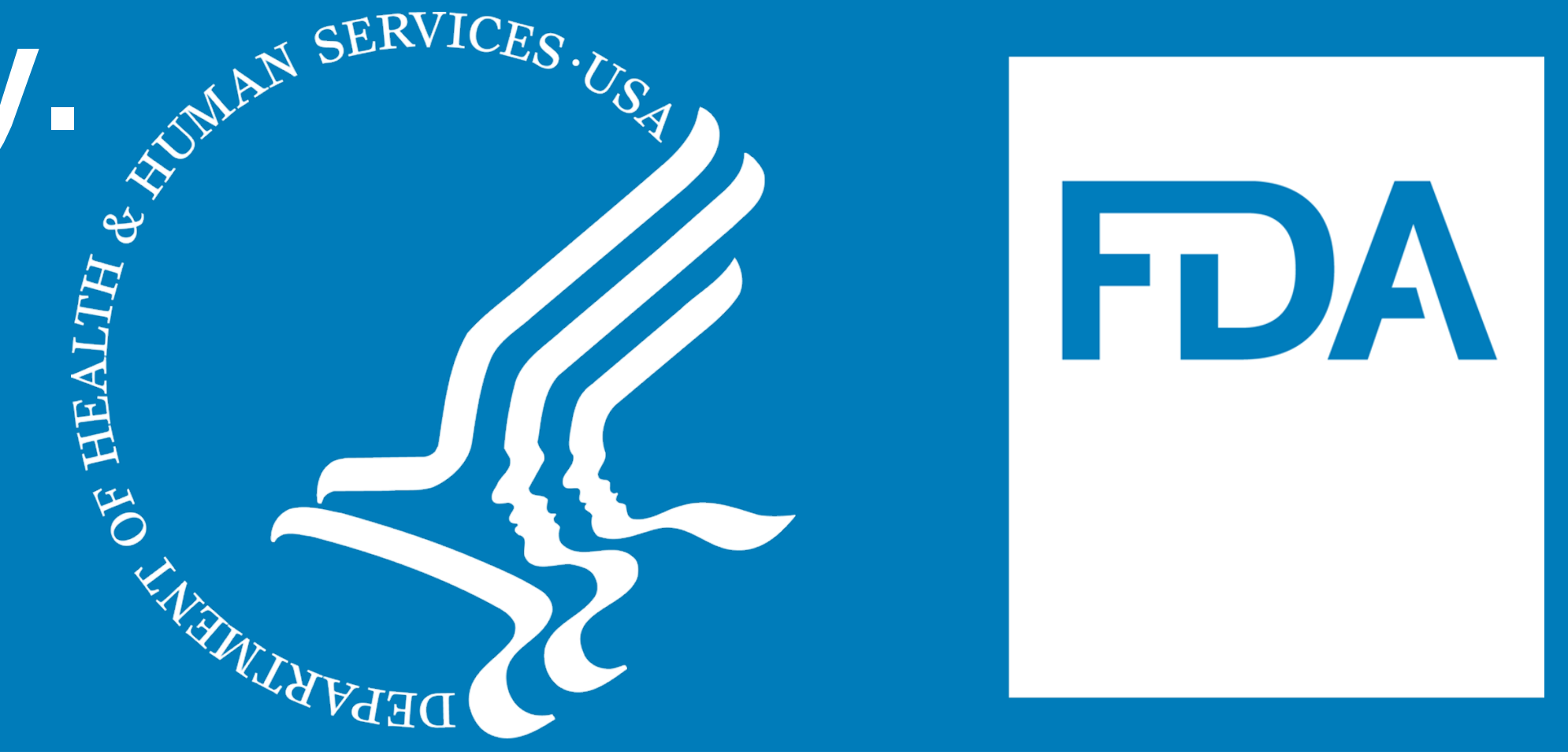
# Communication Methods and Workflows to Develop and Implement High Throughput Truthing of Pathologist Annotations as a Reference Standard for Validating Artificial Intelligence in Digital Pathology.

Qian, Phoebe<sup>1,2</sup>, Setty, James K<sup>1,3</sup>, Elfer, Katherine N PhD, MPH<sup>1</sup>, Gallas, Brandon D PhD<sup>1</sup>

1. FDA/CDRH/OSEL/Division of Imaging, Diagnostics and Software Reliability

2. Mclean High School, VA, USA

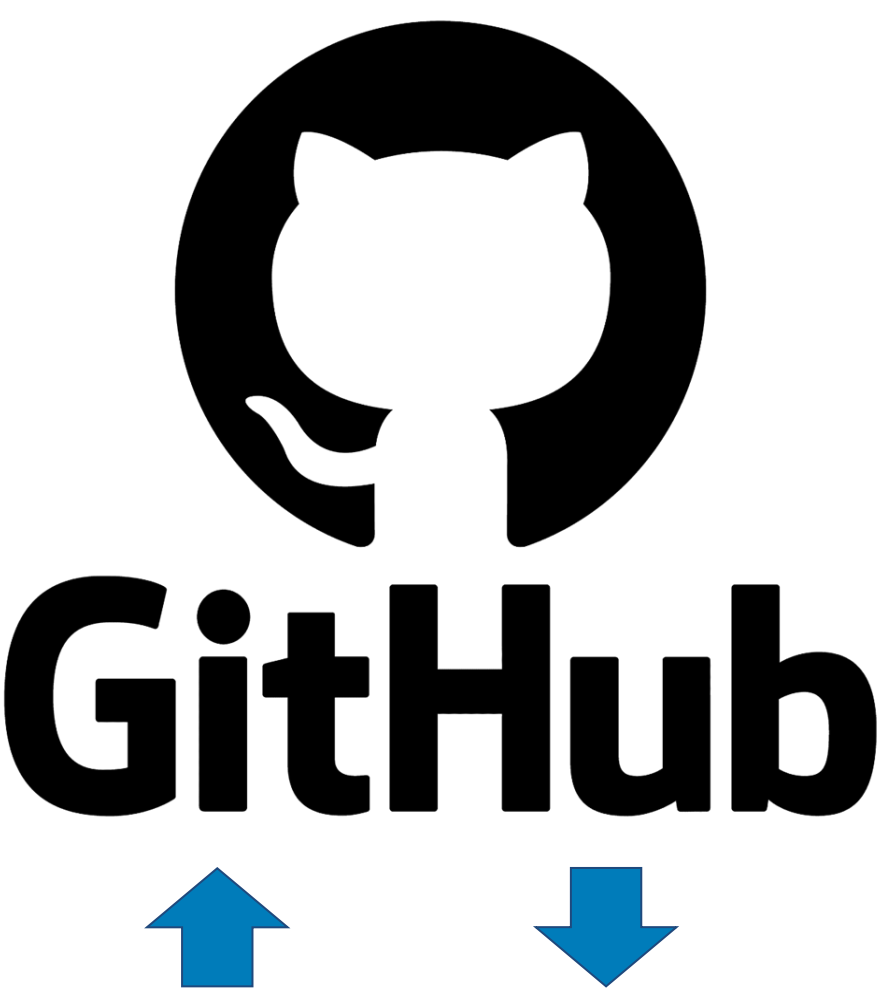
3. School Without Walls High School, Washington DC, USA



## Abstract

We are crowdsourcing pathologists to collect data (images + pathologist annotations) to qualify for the FDA/CDRH medical device development tool program (MDDT). If successful, the MDDT qualified data along with a statistical software package for data analysis would be available to any algorithm developer to be used to validate algorithm performance in a submission to the FDA/CDRH. This poster will include visual descriptions of how we assist in ongoing data collection and project communication efforts. We will introduce our communication platforms (GitHub, NCI Hub), methods (documentation, updates), and programs (Git, RStudio & R scripts) to improve the project workflow with collaborators and team members.

## Tools and Workflows



### GitHub HTT Repository

“Fork”, or copy, the main High-Throughput Truthing (HTT) repository to make changes without disrupting the master branch.

### SmartGit

Collaborative tool that is used to pull/push the forked repository from/to GitHub into/out of our local machine

### RStudio

RStudio is an integrated development environment for R, a language for statistical computing and graphics.

### Push & Pull Requests

Use pull requests to keep our forks up to date and push requests to send changes to the main HTT master branch

## HTT R Package

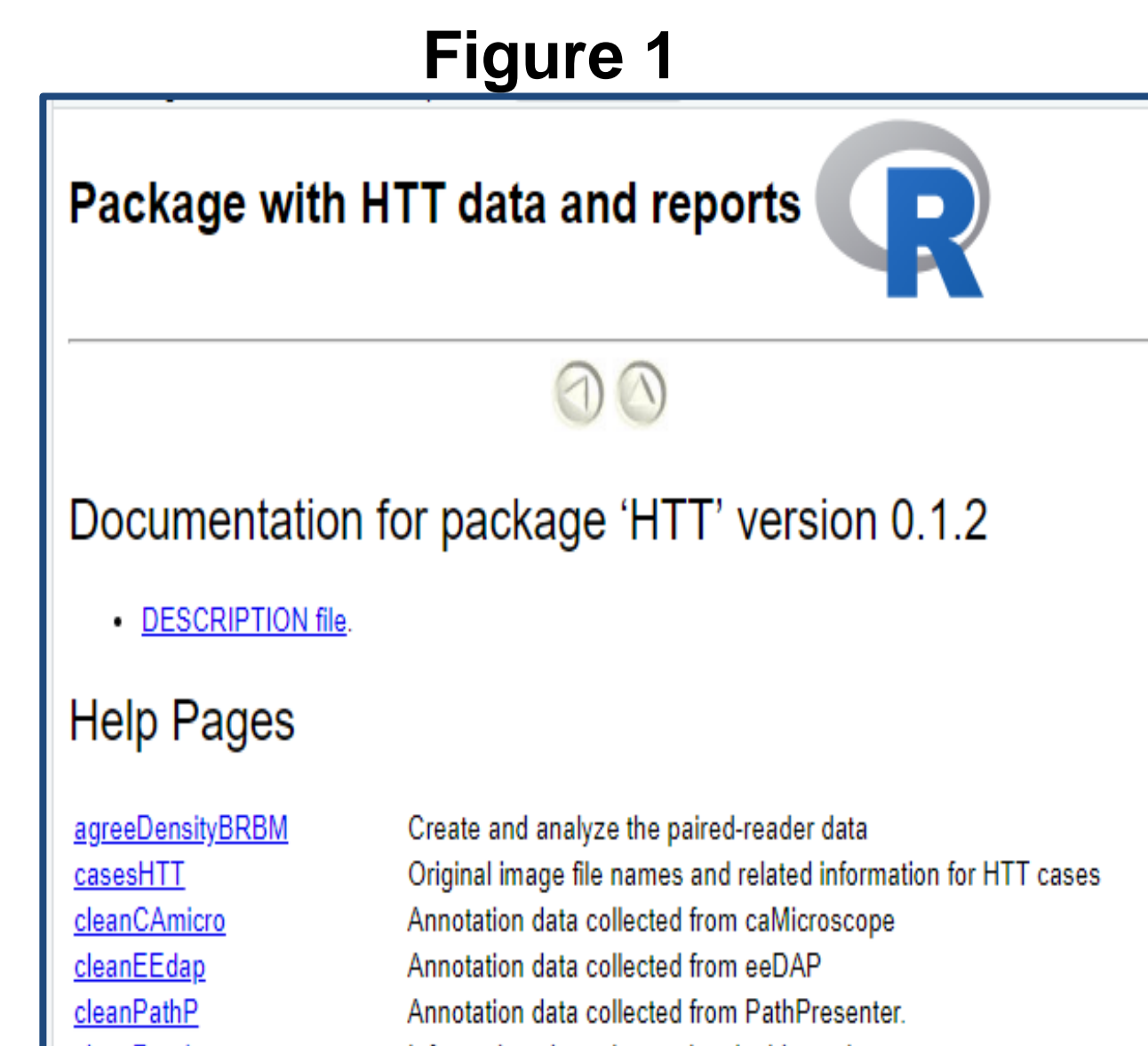


Figure 1: List of help pages in the HTT R Package.

- An R package for data and data analyses
- Documentation pages were updated for other scientists & collaborators to read, understand, and utilize (Fig. 1)
- Shiny Web Apps are included to let users interact with the data and analysis (Fig 2.a, 2.b).
- **HTT Package GitHub:** <https://github.com/DIDSR/HTT>

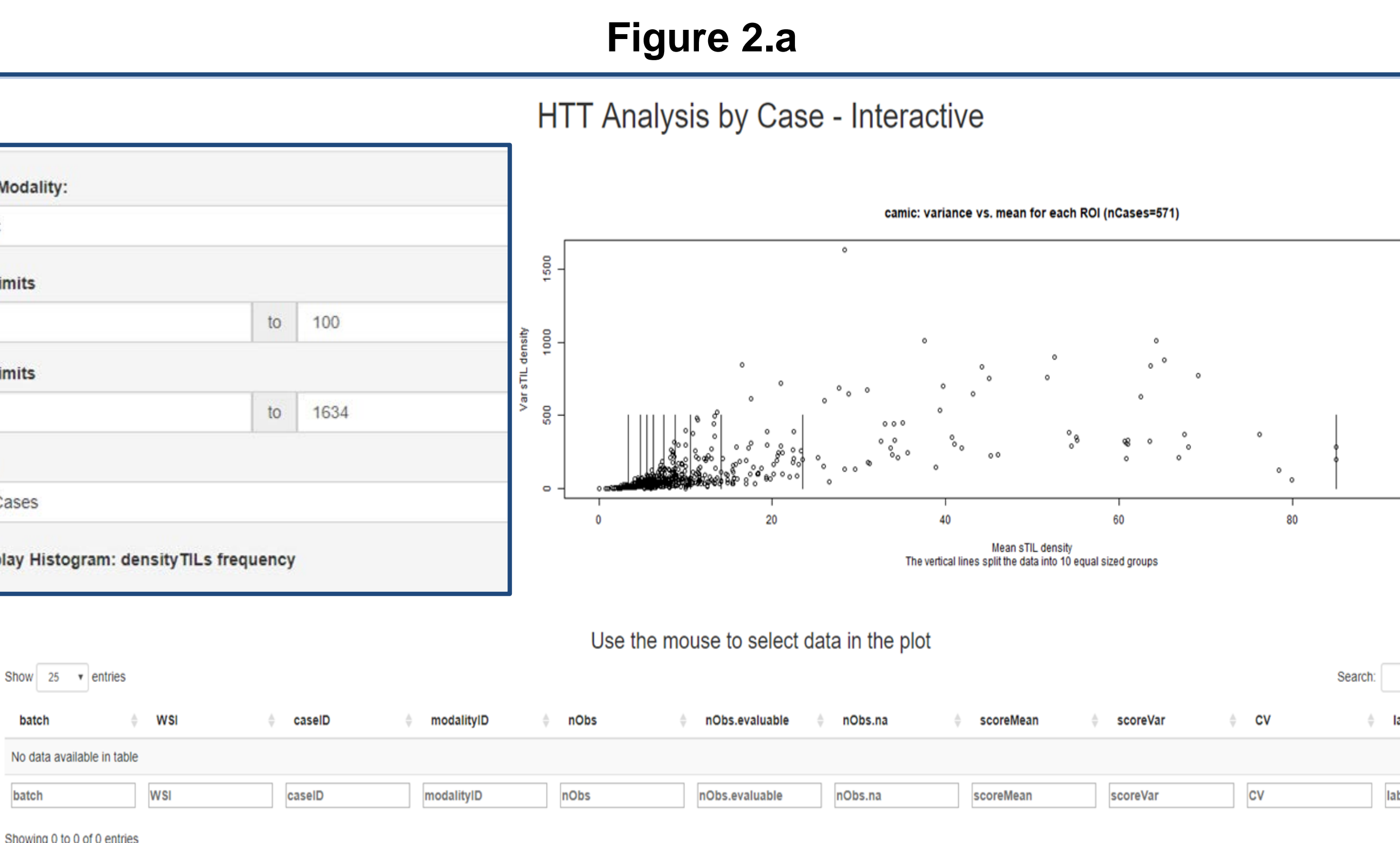


Figure 2.a: HTTAnalysisByCase-interactive.R Shiny (interactive) application showing variance vs. mean TILs density scatterplot. The user is able to select different data points on the plot to see the observations being summarized.

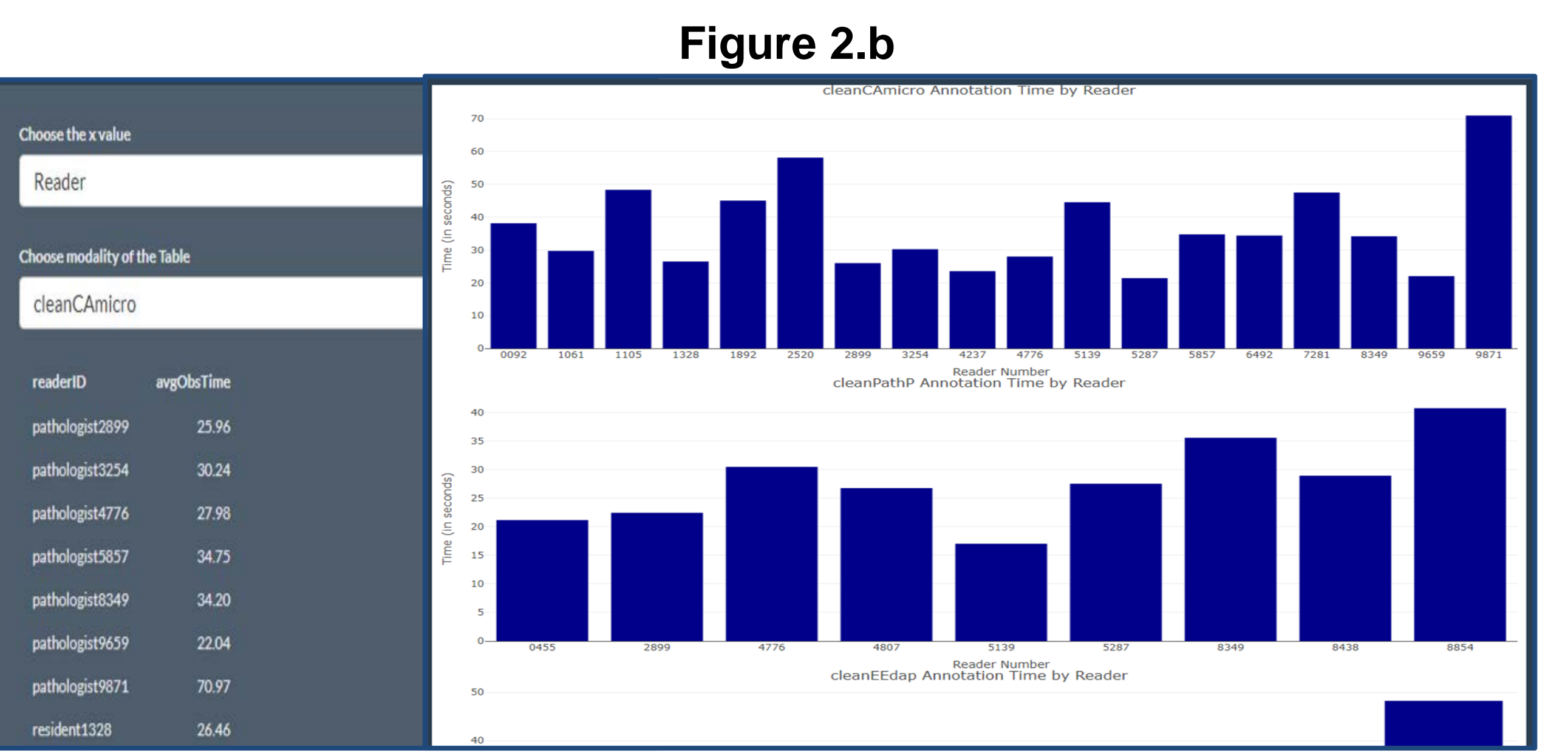


Figure 2.b: averageTimeShinyApp.R. Shiny (interactive) application reflecting the average time it took readers to do one annotation. Annotations summarized by reader and platform (modality).

## NCI Hub Group & Wiki Pages

- **eeDAP Studies group on NCI Hub:**
  - <https://ncihub.org/groups/eedapstudies>
  - Project wiki space for internal and external communications
  - Points users to online data collection, project overviews, updates, and resources.
  - Newly formatted and revised wiki pages improve communication

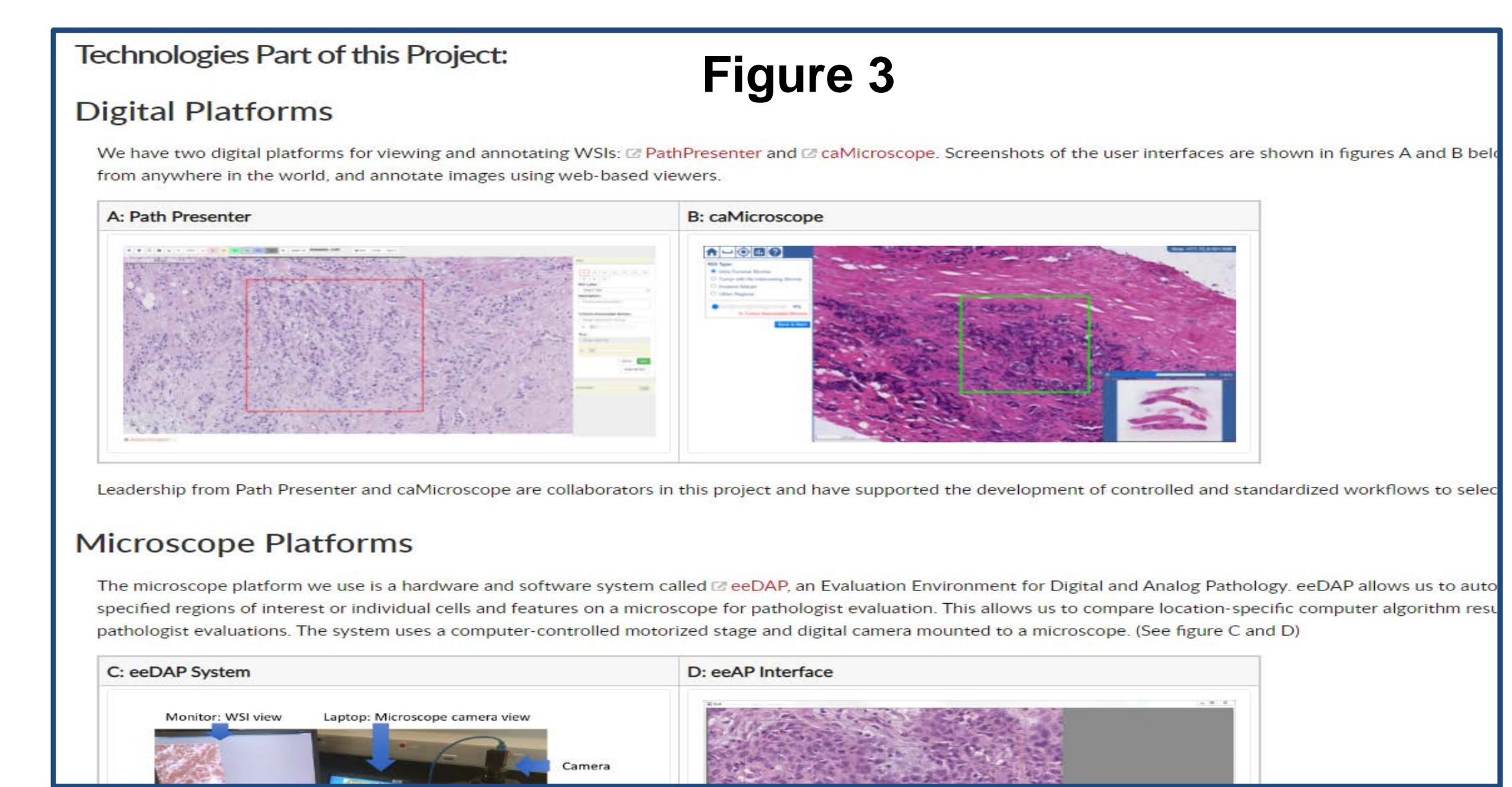


Figure 3: Example of an updated wiki page: eeDAP Studies NCI Hub introduction to the HTT project page.

## caMicroscope & PathPresenter

- Web-based interfaces that collect data by letting pathologists view and annotate whole slide images (WSIs).
- **Training videos** are part of the data collection training process on our NCI Hub. The new videos help guide pathologists to understand how to utilize these tools, how they work, and where to ask for help.

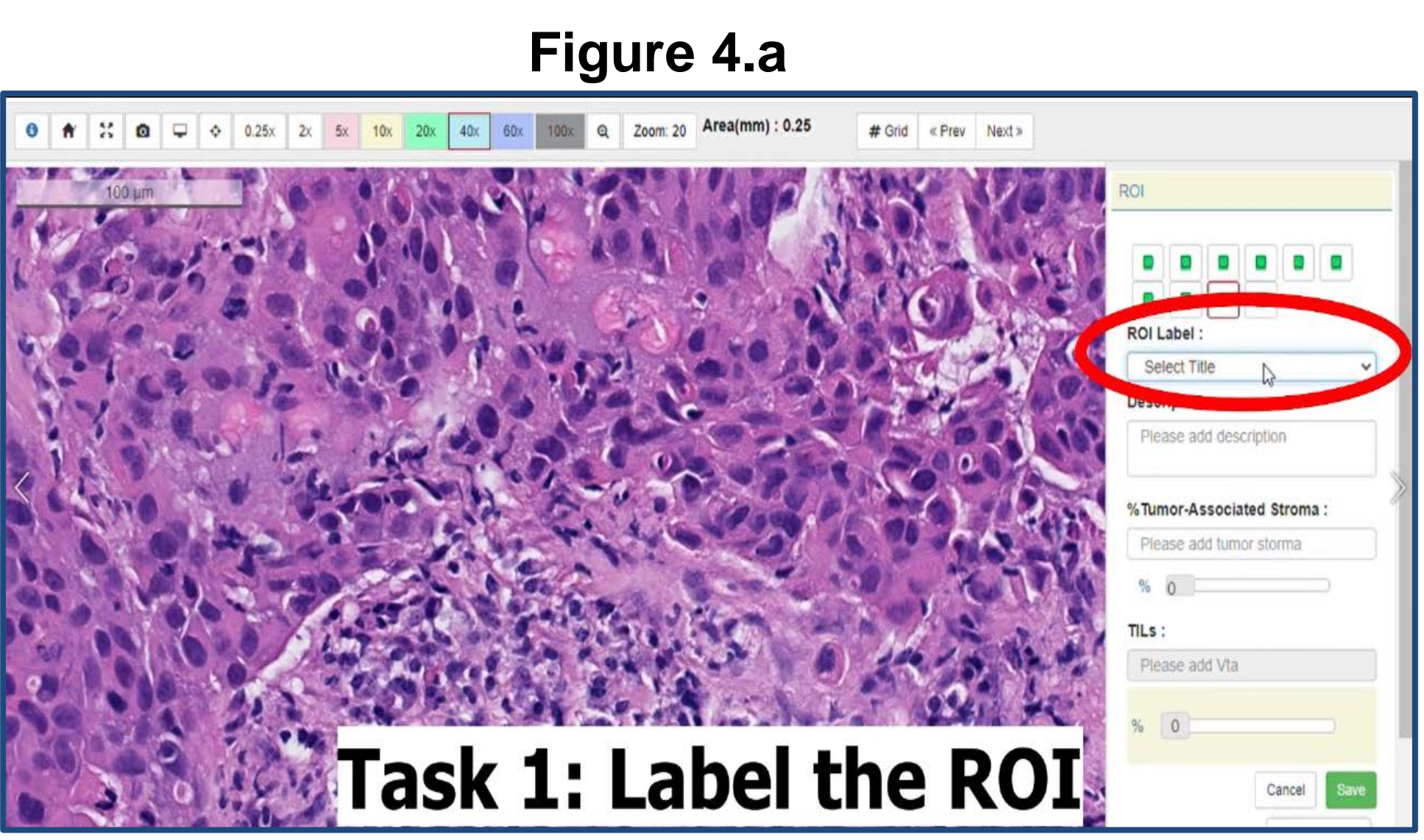
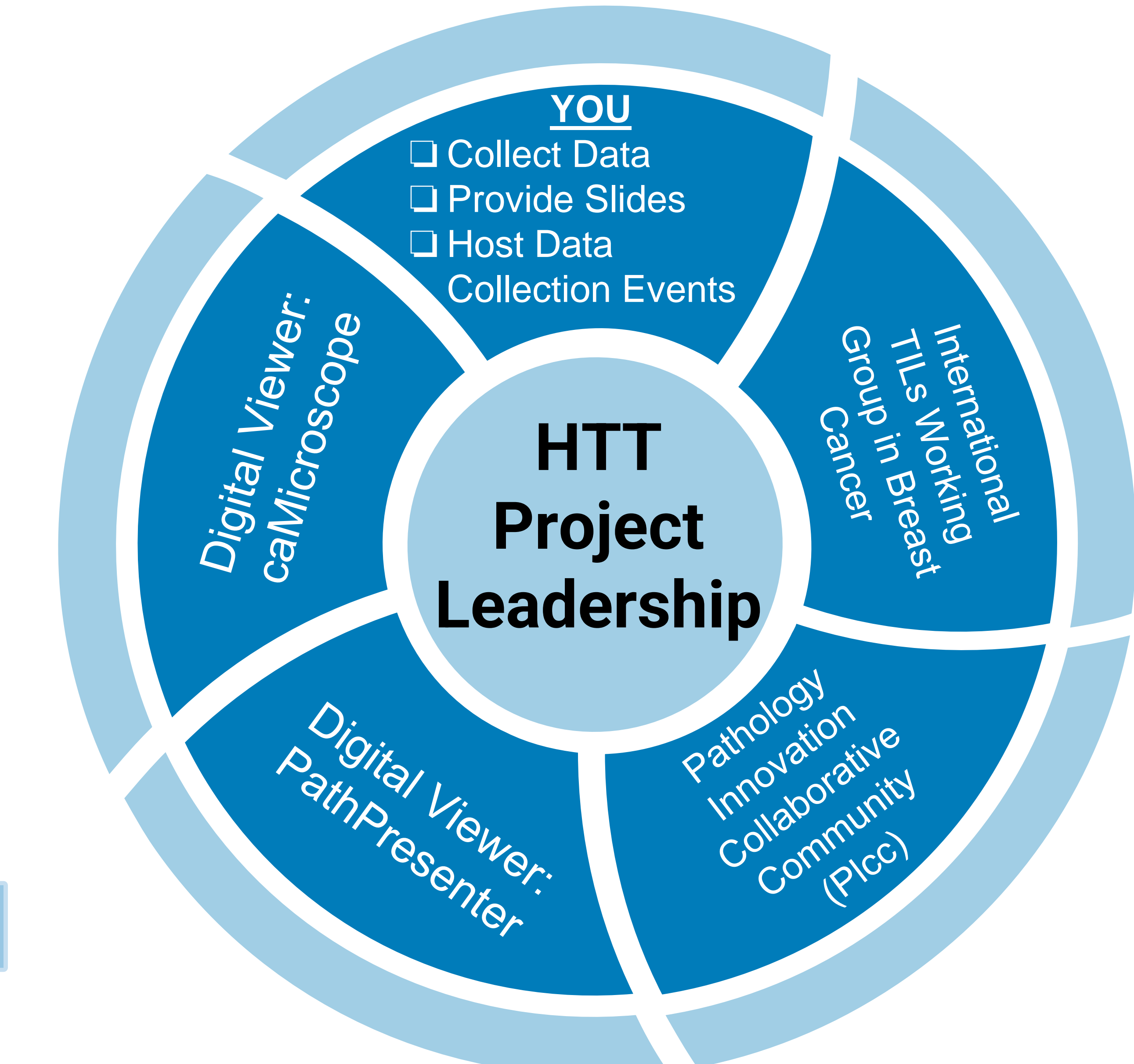


Figure 4.a: Screenshot from the new PathPresenter training video.



Figure 4.b: Screenshot from the new caMicroscope training video.

## Who Contributes?



- YOU: <https://ncihub.org/groups/eedapstudies>
- International TILs Working Group: <https://www.tilsinbreastcancer.org>
- PICC: <https://www.digitalpathologyalliance.org/>
- caMicroscope: <https://wolf.cci.emory.edu/camic/htt/login.html>
- PathPresenter: <https://ai.pathpresenter.net/Account/User/Register>

## Conclusion

- Accomplishments:**
- Completed pilot study with annotations on caMicroscope and PathPresenter and began statistical analyses
  - Constructed Communication Portal: NCI Hub eeDAP Studies
  - Prepared Documentation for HTT R Package

- Future Work:**
- Need to recruit many pathologists to collect data (images + pathologist annotations) for a pivotal study.
  - Submit the HTT R package for internal review.
    - Release R Package for public use after review and approval.

- References:**
- Dudgeon SN, et. al “A Pathologist-Annotated Dataset for Validating Artificial Intelligence: A Project Description and Pilot Study.” *J Pathol Inform.* In press.