VIRGINIA TECH.

VT Research Data Infrastructure

DCSS SPRING 2022

MARCH 30, 2022

Research Data Infrastructure Update

VIRGINIA TECH

Overview

- Current Problem

- Solution overview

- Researcher Driven Process

- Project Phases

"My sponsors expect me to have a secure environment to store, analyze and share data."

"The current option availability is haphazard and makes me uncomfortable as a PI."

"Remote access for research is an enormous challenge."

"I don't know where I can store my research data securely."
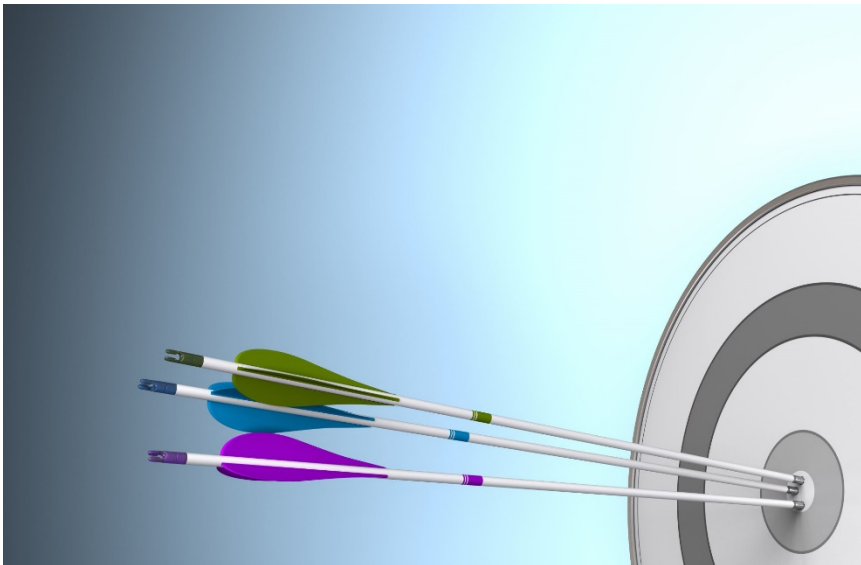
Researcher Need

Information Gathering

- Dr. Robin Queen, SIRC Faculty Fellow

- Town hall workshops

- Faculty Senate Discussions

- Individual interviews with concerned faculty

- Internal and External focus groups

Our Solution

- Storage and analysis environment

- Certified, secure computing resource

- Researcher-designed and directed

- Single or collaborative virtual workspaces

- Access via VT-owned/managed devices or personal devices

- External collaboration options

- Self-serve or supported options

- Videos, images, or large files

- Code repository

Designed with Accessibility and Usability in mind

Offering extensive Research Support

- Data use and protection guidance

- Guidance and templates for grant and sponsorships proposals

- Data management planning

- Open access support

- Costs kept low

- Protocol support

Bio/Description

Rachel is an assistant professor at a large university that conducts biomedical research. She collects her data through online surveys, in person consultations, and virtual consultations. She collects the data on paper and digitally. Her studies can last anywhere from a few months to up to several years. Her study participants are assigned ID's, some of which are generated from a local database. Any PII captured is kept separate from the study results. She saves her results in three different places.

Data she analyzes

- Survey data (in person and digital, REDCap, etc.)
- Longitudinal study data (in person and digital, video)
- fMRI results (imaging results)
- Local database (ID generator)
- Clinical records
- Biological measures
- Electronic health records

Data classifications

✓ PII
✓ PHI
✓ Sensitive information
- Public data

Data structural types

- Text
- Images (MRI, fMRI)
- Videos
- Sensors
- Biological samples

Rachel, the BioMedical Researcher

"I need a large and easily accessible, but protected, environment where I can conduct, share, and store my data in a simple way."
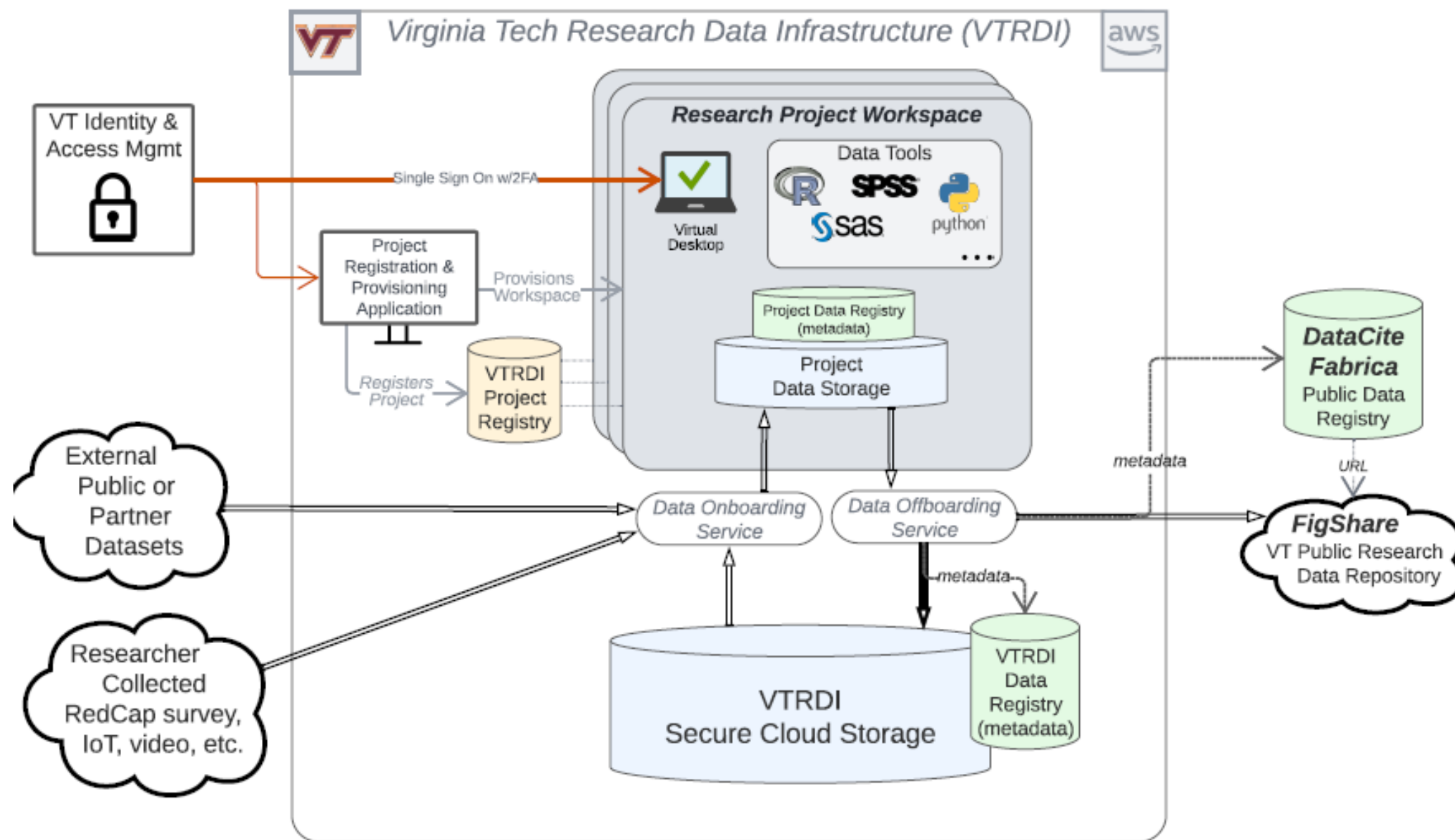
Goals (platform)

- Simple, easy way to set up new virtual research environment that has built in security based on the data classification.
- Zero tolerance for data loss.
- Ability to collaborate with others during research phase.
- Ability to use a variety of tools to analyze the data without having to install, set-up, acquire licenses, etc., for those tools.
- Ability to archive data, but with the ability to access it at a later time, if needed.
- Integration with external data and systems.
- Fully masked PII within research data.
- *Secure access to data by both* internal *and external resources*.
- Storage that is HIPPA and NIH compliant.

Challenges

- Ability to keep data accessible and connected for long periods of time (up to 5 years)
- Lack of time to scan paper forms/notes so it resides with the clinical data

We are currently collecting and studying these types of profiles, wishing to gain a thorough understanding of researcher needs across the university.

VT Research Data Infrastructure Development Timeline

Define and design

Document researcher needs and preferences and turn those into an actionable plan for development of the infrastructure.

Develop and test essential functionality

Pilot the infrastructure and governance processes with early adopters, gather feedback, and deploy the infrastructure to the Virginia Tech community.

Add functionality based on user feedback

Refine governance and resources and expand reach of the infrastructure to all researchers at Virginia Tech.

Questions?

Discussion