

TRUSTED CI

THE NSF CYBERSECURITY
CENTER OF EXCELLENCE

UC Berkeley Secure Research Data and Compute Platform

Jan 1, 2020 - June 30, 2020

6/12/2020

v1.1

Distribution: Release to Public after June 30, 2020

Trusted CI Engagement Team:

Anurag Shankar, Andrew Adams, John Zage, Sean Peisert

This work is made available under the terms of the Creative Commons Attribution 3.0 Unported License. Please visit the following URL for details: http://creativecommons.org/licenses/by/3.0/deed.en_US

Cite this work using the following information: Andrew K. Adams, Sean Peisert, Anurag Shankar, John Zage, "Trusted CI UC Berkeley Engagement: Final Report", July 2020.

About Trusted CI

The mission of Trusted CI is to provide the NSF community with a coherent understanding of cybersecurity, its importance to computational science, and what is needed to achieve and maintain an appropriate cybersecurity program.

Acknowledgments

Trusted CI's engagements are inherently collaborative; this report would not be possible without the collaborative effort that the UCB staff provided, including: Chris Hoffman, Ken Lutz, Jason Christopher, Setareh Sarrafan, and Gary Jung.

This document is a product of the NSF Cybersecurity Center of Excellence (Trusted CI). Trusted CI is supported by the National Science Foundation under grant #1920430. For more information about Trusted CI please visit: <http://trustedci.org/>. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Table of Contents

Executive Summary	4
1. Engagement Background & Process	4
2. Overviews	5
2.1. Stakeholders	5
2.2. Existing Environment	6
2.3. Proposed Environment	6
2.4. SRDC Implementation Plan	8
2.5. Initial Questions Regarding the SRDC	9
3. Findings & Recommendations	9
3.1 Review of the SRDC Plan	9
3.2 Review of SRDC Compliance Status	11
3.3 Recommendations	12
4. Future Engagement Opportunities	14
4.1 Review of completed governance & procedure documents	14
4.2 Developing use cases	14
4.3 Future alignment process	15
5. References and Artifacts	16
Appendix A. Example Use Cases	17
Appendix B. Compliance Attestation Boilerplate Example	19

Executive Summary

Trusted CI engaged with University of California, Berkeley (UCB), for the purpose of reviewing its strategic vision to provide a Secure Research Data and Compute (SRDC hereafter) platform for protected health information (PHI) and possibly other types of personally identifiable information (PII). The SRDC seeks to leverage existing resources and expertise to serve the compliance needs of a wide range of scientists at UC Berkeley, representing fields including biology, engineering, computer science, and a broad spectrum of social sciences and professional schools such as business, public health, and law. We reviewed documents that highlighted various aspects of the UCB vision. Additionally, on multiple occasions we met with UCB personnel to supplement our understanding of both existing practices and the proposed approach to SRDC. Finally, we developed a set of “recommendations” using our experiences in this area. This report itemizes those recommendations, including areas that we believe could benefit from strengthening, based on our understanding of the UCB vision, and describes a few key findings.

We note upfront that it is our belief that the UCB is ahead of peers in using a sound approach to coping with the growing needs for handling regulated research data within the university environment. Leveraging our experiences, we examined both UCB’s progress thus far and what it proposes, and are impressed with what we reviewed.

This report explains the purpose, scope, and process of the engagement (Section 1); expands on the current and desired workflows that describe the SRDC (Section 2); details observations and recommendations for improving UCB’s vision (Section 3); and identifies potential areas for future Trusted CI - UCB engagements (Section 4). Finally, the list of all UCB documents (artifacts) reviewed and potentially referenced throughout this report are appended (see Artifacts).

1. Engagement Background & Process

UCB is building the SRDC Platform to handle restricted research data on campus. SRDC is funded by UCB executive leadership as a “condo-style” research computing service. The institutionally-supported foundation for restricted data research will be professionally managed and supported by Research IT staff from UCB and Lawrence Berkeley National Lab, and researchers will contribute computation and storage hardware to the platform using their research funds. The SRDC Platform will bring together high performance computing (HPC) nodes, virtual machines, and “big data”

storage for researchers working with highly sensitive data (e.g., PHI and PII) across a range of domains, many of which are NSF-funded.

At the beginning of the engagement, Trusted CI and UCB did not have a clear scope of work. Discussions revealed that UCB had also engaged a commercial third party consultant to help with the SRDC project. This allowed us to narrow the scope of work of Trusted CI's engagement by ensuring that effort was not duplicated. Trusted CI and UCB decided upon and agreed to the engagement's primary goal, namely to review UCB's design of the SRDC environment, including third party contributions, and recommend strategies to protect sensitive data from an academia/research perspective. Our strategy to accomplish this goal was comprised of the following objectives:

- a. Trusted CI reviewed UCB's SRDC plan to understand in detail its design, goals, and their future vision (contained in documents which they provided).
- b. Trusted CI compared SRDC's plan to existing institutional practices (e.g., IU), and identified missing components/concepts, requesting UCB develop/provide omissions, if necessary.
- c. Trusted CI recommended changes/enhancements to SRDC's existing plan based on gaps analyzed during the comparisons/best practices.
- d. Trusted CI and UCB explored ways to optimize UCB's overall restricted research data approach by incorporating lessons learned at peer institutions.

2. Overviews

This section describes the current environment being used for UCB's "moderately sensitive" data¹, the proposed SRDC environment, and the plan to implement the SRDC. We specify the stakeholders next, and then describe the workflows in detail.

2.1. Stakeholders

1. Research Information Technology (RIT) - UCB's research computing organization
2. Information Services Technology (IST) - UCB's central IT organization
3. Lawrence Berkeley National Lab (LBNL) - RIT partner

¹ UCB classifies its institutional data into four categories based on adverse business impact: P1 - Minimal, P2 - Low, P3 - Moderate, and P4 - High.

4. Information Security Office (ISO) - manages UCB's institutional cybersecurity operations.
5. Vice Chancellor for Research
6. Chief Information Officer
7. Chief Information Security Officer
8. Researchers

The SRDC service will be provided by Research IT at UC Berkeley. Technical support will be provided by staff from UC Berkeley's central IT organization (IST) as well as from LBNL's Scientific Computing Group who have been retained by Research IT to help administer the SRDC Platform.

2.2. Existing Environment

UCB provides two research computing environments today to researchers - an HPC cluster and a virtual machine (VM) environment.

1. HPC Cluster - Called Savio. Originally implemented for data classified at a P1 level (low sensitivity), but hardened one year ago for P3 (moderately sensitive). Savio is maintained by staff from LBNL's Scientific Computing Group who have been retained by Research IT.
2. Virtual Machine (VM) Environment - Called Analytics Environment on Demand (AEoD). A VMware/Citrix service. Originally implemented for data classified at a P1 level, but hardened one year ago for P3 (Moderately sensitive). Maintained by IST.

2.3. Proposed Environment

UCB plans to build a brand new HPC and VM environment for SRDC. Figures 1 and 2 illustrate the proposed architectures for the two environments.

SRDC HPC Architecture and Network Diagram

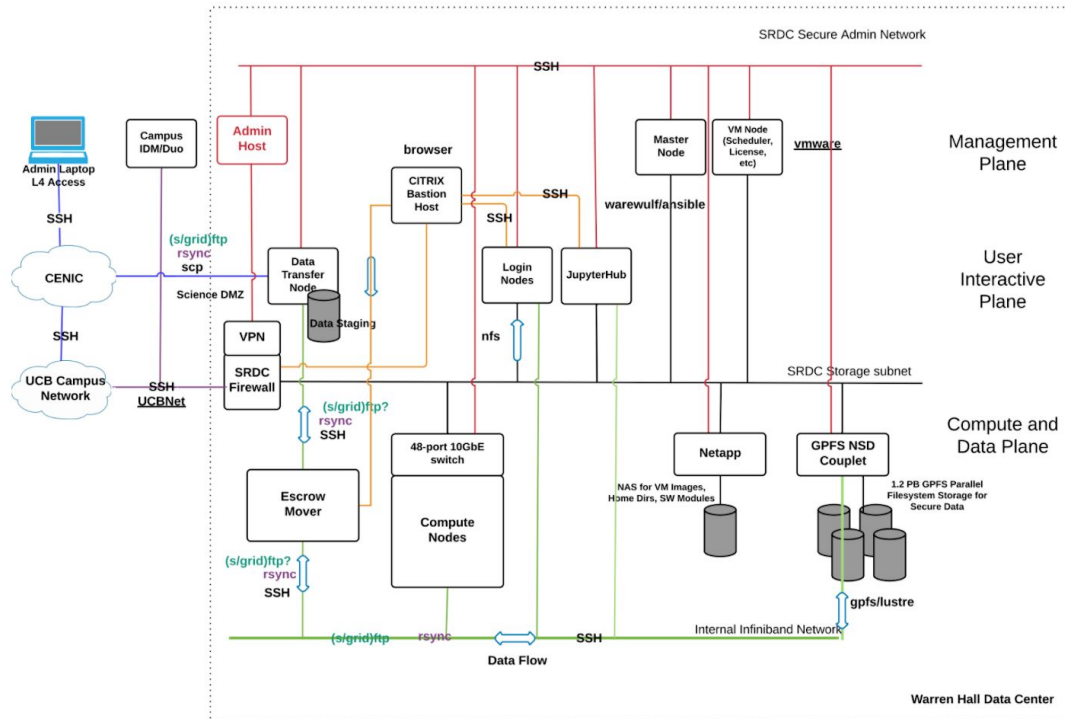


Figure 1: The SRDC HPC Architecture

The HPC environment will consist of compute nodes, login nodes, and Jupyterhub nodes. Storage will be provided by GPFS storage servers. All these components will be located within a Science DMZ. The details of user access are still under discussion with researchers.

DRAFT
SRDC VM Phase 1 Architecture
and Network Diagram
Version 0.1
2020-06-12

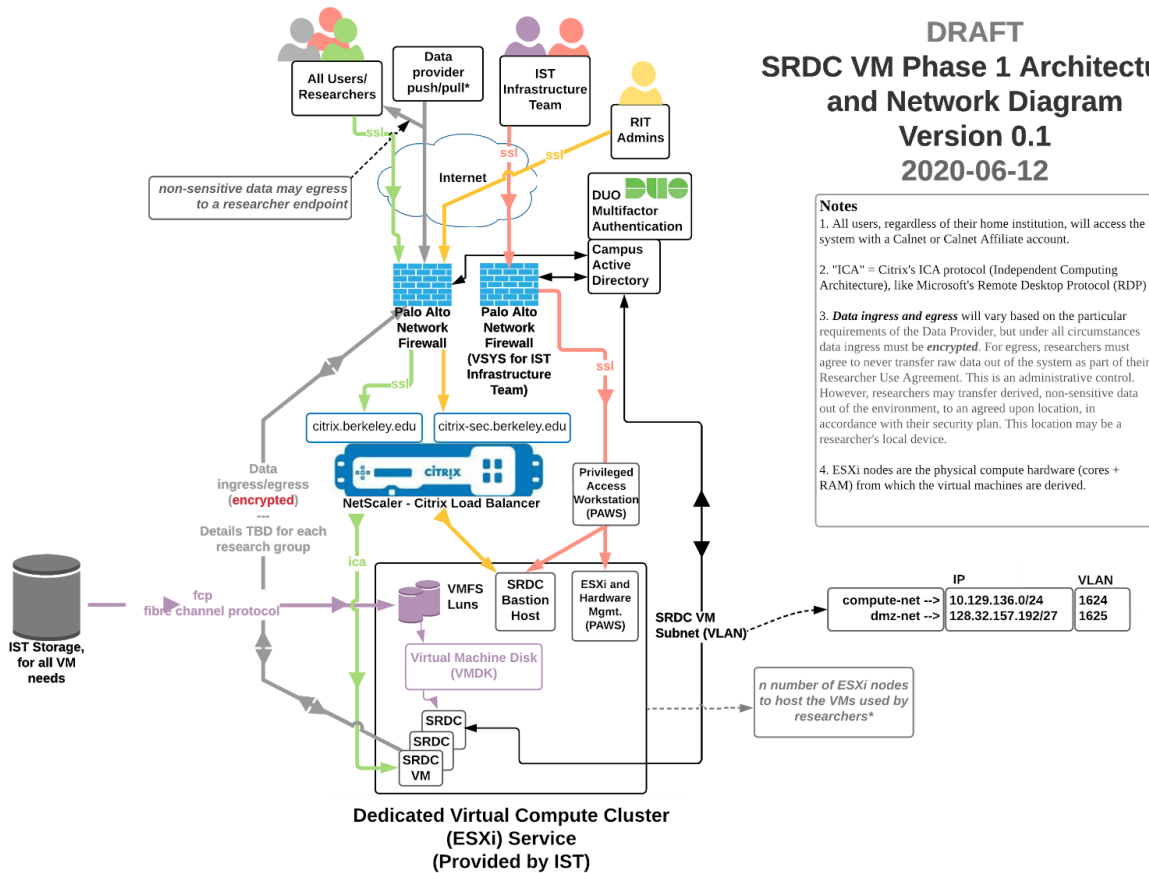


Figure 2: SRDC VM Architecture

The SRDC virtualization environment will use dedicated hardware running the VMware ESXi hypervisor. Storage will be shared with the HPC environment.

2.4. SRDC Implementation Plan

The SRDC rollout has been hampered due to Data Center power issues, creating the need for an interim VM environment to meet urgent needs. UCB is building this environment using existing resources. This “SRDC VM Phase 1” environment will consist of VMs approved for data at P4 level (highly sensitive data). The project will have three initial users, expanding to an expected six next year.

The final SRDC implementation is using a phased approach with three distinct phases, each with distinct goals.

1. Phase I - Select security framework, review existing documentation
2. Phase II - Develop security plan, controls, policies, and procedures
3. Phase III - Confirm that everything is in place

To accomplish this, UCB engaged a commercial third party called GuidePoint. At the end of the Trusted CI engagement (June 30, 2020), UCB and GuidePoint had completed Phase I and were mid-way through Phase II of the SRDC project. It resulted in UCB deciding to adopt a GuidePoint recommended cybersecurity framework that combines controls from two different security frameworks. By the end of Phase II, expected to finish later this year, GuidePoint will help UCB develop SRDC policies and procedures, implement controls, and document them in a SRDC system security plan.

2.5. Initial Questions Regarding the SRDC

The proposed plan that UCB desires to implement contained unanswered questions at the start of the engagement:

1. Should UCB proceed with GuidePoint for the remainder of Phase II?
2. Which cybersecurity framework is appropriate to meet UCB's compliance needs?
3. How do regulated data 'use cases' map to resources?
4. How best to institute a security perimeter (authorization boundary)?
5. What storage subsystem controls are appropriate, for instance encryption at rest/transit, file system, performance/usability?
6. How are peers implementing infrastructures for regulated data?
7. How should UCB approach compliance from an academic research perspective?
8. What controls process should be required of users to access the system?

3. Findings & Recommendations

3.1 Review of the SRDC Plan

3.1.1 Interim plan. Being forced to build an interim solution due to power issues will impact the SRDC, but also provides an opportunity to pilot the VM architecture and socialize it with users.

3.1.2 Engaging a third party. We endorse UCB's decision to engage a commercial third party consultant to help with SRDC's security program. (UCB is fortunate to have the resources; many peers struggling with regulated data do not.) We believe it is a prudent choice for a number of reasons:

1. It adds valuable resources to the project, reducing the time to completion.
2. It provides a checkpoint to gauge the current state of security, in particular institutional policies and procedures.
3. It greatly mitigates the lack of regulatory expertise on campus. While it is possible for an institution to develop a regulated data program from scratch, the path is long and arduous (as the decade-long experience at Indiana University, IU, shows).
4. It jumpstarts UCB's regulated data program and will inject regulatory expertise locally. This will allow UCB to self-manage its regulated data program going forward.
5. The sponsored research funding enabled by SRDC over the long run will more than make up for the initial investment.

We also observe that UCB has done its homework well and wisely chosen a vendor with experience serving academia. Many commercial vendors interact largely with other commercial entities, such as healthcare organizations. This gives them a very narrow perspective on compliance, which is typically not ideally suited for a research environment.

3.1.3. Trusted CI Engagement. Engaging Trusted CI was a commendable strategy by UCB. It provided the much needed academic research perspective on compliance as well as an independent checks and balances mechanism for the GuidePoint engagement.

3.1.4 Framework selection. Trusted CI found GuidePoint's recommended framework that combines NIST 800-53 Moderate and High Baseline, ISO 27002, HIPAA, and GDPR controls to address compliance to be sound. It matches the approach some peers and other organizations have taken. However, we find two issues with this choice: (a) GuidePoint has decided to choose NIST 800-53 rev 4 despite the fact that rev 5 is imminent, and (b) Cherry-picking from multiple frameworks is likely to prove challenging going forward as components in each framework evolve over time. Tracking changes for just one framework and making the documentation reflect it is difficult enough; doing it for two will be thorny.

3.1.5 Decision to enclave. UCB has chosen to create an enclave for regulated data. This is sound from a security perspective, but places constraints on the variety of use cases that it can address. Researchers have a diverse set of needs (see [Appendix A](#) for a sample of use cases encountered at IU), and enclaves are not ideally suited to meet them all.

3.1.6 Security Architecture.

HPC: While still somewhat a work in progress, the architecture and process in place for HPC system use appears solid. There are no direct connections between HPC nodes and the outside world, the compute nodes are never directly accessed at all, and even the login nodes, where the batch schedule runs, are accessible only via a Citrix-based bastion host in which only pixels are actually moving between the end user and the host, as even copy out of the environment is planned to be disabled, along with file transfer. The bastion host itself is only accessible through VPN. We note that the devil is in the details, however. As we understand, it is the intent of the system architects to implement the system in this fashion, but the actual architectural diagrams have a number of lines that do not entirely match up with this description, such as the direct connection between the compute nodes and switch and the SRDC Firewall (see Figure 1).

VM: The virtual machine architecture is set up to enable management of two different types of VMs. The first includes Windows-based VMs that are used for administrative tasks. The second are VMs that are actually used by end users. As with the HPC nodes, access to the VMs is only through remote desktop from VPNS, and not from the Internet. The VMs on the SRDC are used by researchers for data management, analysis, and de-identification of the highly confidential data. They have been designed for highly sensitive data that is classified by the University of California as Protection Level 4 (P4).

Both the HPC and VM components of the SRDC meet many of the physical, administrative, and technical safeguards of the HIPAA security rule and most UC Berkeley information security policies. These safeguards will be documented in the SRDC System Security Plan. As such, the environment meets or exceeds many of the security requirements of the service provider.

3.2 Review of SRDC Compliance Status

3.2.1 Documentation. UCB is completing documentation required by its ISO to approve SRDC to host P4 level data. Trusted CI greatly commends the UCB for requiring these “Minimum Security Standards for Electronic Information” (MSSEI) documents that describe RIT’s implementation of the security controls required by the UCB ISO. They will prove critical as support tools in UCB’s current and future

compliance efforts, providing a seed for the inventory of controls the SRDC and other systems will need in order to show UCB's regulated data due diligence.

3.2.2 Controls. Since UCB has only recently selected a framework and controls, and the final SRDC plan is still in execution, Trusted CI focused on reviewing the recommended controls. Additionally, we reviewed the MSSEI documents for the existing systems and the proposed SRDC. Based on this review, the following key points were identified:

1. Many of MSSEI controls already overlap with HIPAA security rule requirements.
2. The GuidePoint proposed controls will fill in the gaps.

3.2.3 Gaps. A cybersecurity consulting firm is slated to help UCB with a risk assessment in Phase III of the engagement which should provide a more accurate picture of gaps and risk, but the following observations on HIPAA gaps, some of which are obvious, can already be made.

1. Documentation. There is no maturity compliance without documentation and the documentation as it exists is insufficient to address HIPAA. The UCB team was beginning this work as this report was published.
2. Controls. A number of HIPAA security rule safeguards (controls) are not currently covered by MSSEI. These include:
 - a. Risk analysis
 - b. Risk management plan
 - c. Assigned security responsibility
 - d. Awareness and training
 - e. Contingency plan
 - f. Evaluation
 - g. Business associate contracts
 - h. Integrity controls
 - i. Documentation (already alluded to in 1 above)
 - j. Time limit (retention)

3.3 Recommendations

3.3.1 Do mission-centric security and compliance. Be cognizant of the enclave pitfalls. Ensure researcher workflows are not unduly restricted by security and compliance. Since compliance, especially HIPAA, is a local interpretation of the statute, interpret it wisely. Accept some risk to keep the mission front and center. As an

example, if user workflows are severely compromised by security impediments that limit direct SSH connections to the HPC cluster, allow them, but require, e.g. a VPN and 2-factor authentication.

3.3.2 Build security in. Don't overburden the researchers with cybersecurity. Let them concentrate on research. Anticipate insecure workflows that researchers are likely to execute, explore the reasons that drive their behavior, secure those workflows by providing alternate solutions, and weave security into these solutions from the get go.

3.3.3 Inventory use cases. To ensure researcher needs will be met, understand those needs by surveying current and potential use cases by asking the researchers what they are doing now and plan to do in the future. Appendix A provides a number of use cases collected by IU over a decade of HIPAA compliant research computing to the School of Medicine. Another resource that may be useful is the IU SecureMyResearch Cookbook (<https://go.iu.edu/coobook>).

3.3.4 Provide secure research facilitation. We cannot overstate the importance of providing one-on-one support to researchers to help secure research data. No amount of policies and procedures or controls will yield better security than a few minutes of one on one time. This support needs to exist during the entire lifecycle of a grant, i.e., before submission and after upon grant approval. UCB should make this a priority and expand its already commendable effort. (See 3.3.7 below.)

3.3.5 Use existing resources. Both UCB's contract with GuidePoint and engagement with Trusted CI are excellent examples of using available resources wisely. There are other good resources, too. Regulatory pressure in recent years has produced peers and resources within the academic research community that are able to provide research cybersecurity and compliance help and guidance, for instance REN-ISAC, Trusted CI's Large Facilities Working Group, the HighEdCUI Slack channel, etc.

3.3.6 Join the community. The SRDC implementation will place UCB among a select group in academia (e.g. IU, Purdue, U Chicago, U Florida, U Connecticut, UCSD/SDSC, and Duke) that have research computing solutions and expertise to address regulatory compliance centrally (instead of selectively, e.g., compliance being limited to departments or a medical school). Due to the pressing need for this experience, UCB (and the rest of us) should also strive to help peers struggling with compliance.

3.3.7 Create Boilerplates. Create short passages of text describing the SRDC security and make them available to researchers to include in grants, contracts, and data use agreements as needed. An example template is provided in Appendix B.

3.3.8 Security Architecture. Work with the UCB ISO and GuidePoint to vet the architecture for compliance with UCB policies and the HIPAA Security Rule, respectively.

In a security diagram use a legend to describe the meaning of each line color, what type of connection they are and whether it is a physical network or logical communication pathways through the same network.

Access to the SRDC for administration and management purposes are connected through a secure system, either a bastion host or a privileged access workstation. This is a good security control for the elevated level of privilege needed by these users.

In an ideal security architecture it is a good idea to control traffic through one central pathway rather than have it distributed through many different methods, thus making monitoring easier. There seem to be 3 different pathways to access the login nodes and JupyterHub from the UCB Campus Network.

Finally, in the case each different color line represents a different network, nodes that connect to different networks are ideal targets for attackers. These nodes, such as the Login Nodes and JupyterHub should be considered a larger risk, since they bridge between networks. At a minimum, centralized logging, host integrity checking and network monitoring should be implemented.

3.3.9. Address HIPAA gaps. Phase II of the SRDC project/GuidePoint is likely to address many of the following, but Trusted CI's gap analysis shows that full HIPAA compliance for SRDC will require additional controls, especially when adding completely new components or systems to SRDC in the future.

1. Risk analysis.
2. Risk management plan.
3. Assigned security responsibility. Designate a formal SRDC security lead.
4. Awareness and training. Ensure all staff members and researchers with access to PHI take HIPAA training annually.
5. Contingency plan. Create a documented disaster recovery (DR) plan.

6. Evaluation. Perform annual reviews of documentation, test to ensure controls are doing what they are supposed to, document changes to systems since the last review and assess risk the changes have introduced or mitigated, and respond appropriately.
7. Business associate contracts. Ensure BAAs with vendors with access to PHI on SRDC (e.g. hardware support vendors).
8. Integrity controls. Use a tool like AIDE to monitor file integrity.
9. Policies and procedures. Create SRDC policies and procedures. Leverage institutional policies and procedures in the SRDC SSP.
10. Documentation. Document due diligence, secure the documents, keep them for six years, perform annual updates, and make sure everyone who should have access to it does.

3.3.10. Implement the GuidePoint guidance sanely.

1. Do not attempt to implement everything in the guidance column of the SSP. Customize implementation such that the controls are reasonable and proportionate to the needs of the project and the researchers.
2. Do not create a policy for each control. Have only a few high level, briefly articulated policies that will stay relatively immutable in time. A possible approach would be to have 11 policies, one per domain. Create procedures for the rest. Use the Trusted CI and the IU NIST policies templates for guidance.
3. It may also be useful to also create a simpler document with only the HIPAA safeguards (using the provided template) and indicate in brief how each is being addressed. This will make it easier for an auditor specifically auditing against HIPAA.
4. Describe briefly how the applicable controls are implemented in the SSP and add a column that refers to source documents (such as documented procedures) and URLs that provide further details. This makes it easy for an auditor to quickly review the SSP without needing to dig deep down.
5. Use the RACI tab in the SSP as evidence of an internal “Authority to Operate” (ATO) when required by contracts.
6. Review the SSP, policies, and procedure at least annually.

4. Future Engagement Opportunities

The scope of Trusted CI’s work in this engagement was limited to reviewing the UCB MSSEI documents and exploring issues encountered by UCB as they defined their

workflows. In part due to both the engagement's scope and upon developing recommendations, a few areas emerged which could be the basis for future work. The topics follow:

4.1 Review of completed governance & procedure documents

As alluded to above, it was not possible to review the entire collection of policy & procedural documents that would define the framework, since SRDC was still in development. A future review, however, of the eventual set of complete policy & procedure documents may prove beneficial to both UCB and Trusted CI.

4.2 Developing use cases

A practical tool that departments associated with the UCB framework could leverage would be a list of 'use cases.' The use-cases would allow those responsible for identifying proposals that could/would satisfy inclusion into the UCB framework, as well as perhaps even guiding them so far as identifying which platforms address the needs of the proposal. Developing a database of these use-cases could benefit more than just UCB, and seems worthy of an additional engagement.

4.3 Future alignment process

Aligning new CI for inclusion into the UCB framework may be non-trivial and require guidance once the GuidePoint engagement is complete. This provides another potential opportunity for UCB to engage with Trusted CI.

5. References and Artifacts

- [1] 2019_02 UCBProgram Status, <https://docs.google.com/presentation/d/1144c6Z7UwuFv5QL5YK9R9DNjc7QzizROBWYldfk83fs/>.
- [2] Box_REED_Folder_DSP_Template, <https://app.box.com/file/398511303944>.
- [3] CUI Framework Outline, <https://app.box.com/file/397844893002>.
- [4] EXRC3, <https://app.box.com/file/397136165316>.
- [5] Globus, <https://app.box.com/file/397128711092>.
- [6] Pre Assessment - What the survey looks like, <https://app.box.com/file/435868730095>.
- [7] Process Overview, <https://app.box.com/file/431632143461>.
- [8] Readme, <https://app.box.com/file/397068684418>
- [9] RedCap, <https://app.box.com/file/397123039505>.
- [10] UCB, <https://app.box.com/file/397135902346>.
- [11] UCB(Current), <https://app.box.com/file/397131235229>.
- [12] Research Framework, <https://app.box.com/folder/66325033308>.
- [13] Research Security Model - consulting group, <https://app.box.com/file/397844938777>.
- [14] Task Status, <https://app.box.com/file/397850571484>.
- [15] TrustedCI_01162019 Master consolidated_4. Identify ITaP gap assessment, <https://app.box.com/file/397076800214>.

Appendix A. Example Use Cases

The following enumerates the most common technology use cases in research that need securing. The list is not meant to be exhaustive; it simply documents use cases encountered during the provision of HIPAA compliant research computing solutions to the Indiana University (IU) School of Medicine over a decade.

It is implicitly assumed that the data typically originates on an instrument, a researcher desktop, or a departmental system, and the use cases require scale not achievable by departments.

Finally, the IU approach to PHI on central research systems was based on the following tenets:

- The #1 priority is to make sure researchers can get work done, not compliance.
- Security measures that slow down research workflow will be bypassed by researchers.
- Research use cases are too broad to be enclaved.
- Manage risk on resources researchers are using already.
- Interpret the regulation as HHS intended² in choosing *reasonable* and *appropriate* safeguards.

Data

1. Share files with internal and/or external collaborators. (The most common use case)
2. Store massive volumes of data for an extended period and share it with internal and/or external collaborators. (Unlimited deep store use case)
3. Publish a massive data repository via the web. (Online museum use case)
4. Workstation backups.
5. Access data using a desktop application without the application being installed on individual desktops. (Desktop virtualization use case)
6. Manage data stored in Excel or CSV files by easily moving to a database.

² What is reasonable and appropriate "... will depend on a variety of factors, such as, among others, the entity's risk analysis, risk mitigation strategy, what security measures are already in place, and the cost of implementation". (<https://www.hhs.gov/hipaa/for-professionals/faq/2020/what-is-the-difference-between-addressable-and-required-implementation-specifications/index.html>)

- a. Make them available to internal and/or external collaborators. (REDCap use case).
7. Manage surveys and collected data (REDCap use case).
 - a. A built-in survey component suitable for longitudinal studies that can export in a format ingestible by statistical packages such as SAS, SPSS, etc.
8. Store data in a database (relational or otherwise) and access it from a custom or vended web application. (Web hosting use case)
9. Store structured data in a RDBMS with the ability to quickly develop a web application to access/serve this data. (Oracle Apex use case)
10. Moving very large amounts of data quickly. (The high UPS bandwidth use case)
11. Data storage/analysis in the field (no internet connectivity). (REDCap use case)
12. Extract, transform, load (ETL) data.

Compute

11. Analyze data requiring high end CPUs and GPUs using a custom or vended application. (HPC use case)
12. Data may reside on a file system or a database (relational or otherwise).
13. Complete server environment administered by the researcher. (VM use case)

Mobile

14. Push and pull data to/from a repository via a mobile app. (Collect data from patients use case)

Cloud

15. A custom desktop application processing/accessing data stored in the cloud via a cloud API.
16. Store and process data in a SaaS application in the cloud.
17. Cloud HPC.
18. Data storage in the cloud. (S3 use case)
19. Backups to the cloud. (Glacier use case)

Appendix B. Compliance Attestation Boilerplate Example

To Whom it May Concern

The University Information Technology Services (UITs) at Indiana University (IU) is committed to the strictest cybersecurity and risk management standards to protect its systems and comply with applicable local, state, and federal rules and regulations.

1. We have instituted the NIST Risk Management Framework (RMF) based on NIST Special Publication (SP) 800-37 "Guide to Applying the Risk Management Framework to Federal Information Systems" to manage cybersecurity risk to the information systems in question. Our RMF leverages applicable controls from NIST SP 800-53 "Security and Privacy Controls for Federal Information Systems and Organizations", a control catalog used by the US government comprising nearly a thousand controls in seventeen different control families.

2. We follow guidance provided by NIST SP 800-30 "Guide to Conducting Risk Assessments" and SP 800-39 "Managing Information Security Risk".

3. All relevant systems are subject to a rigorous process that includes:

- System and software inventory
- Risk assessment and response
- A FISMA-style "System Security Plan" documenting applicable NIST 800-53 controls
- Institutional authority to operate by IU Data Stewards
- Annual reviews and updates
- Training
- Comprehensive documentation

4. All systems satisfy applicable HIPAA Security Rule requirements as per NIST SP 800-66, "An Introductory Resource Guide for Implementing the Health Insurance Portability and Accountability Act (HIPAA) Security Rule".

5. All personnel responsible for managing the systems take HIPAA Privacy Rule training and HIPAA Security Rule training annually as well as a one-time HIPAA and Mobile Device Security training. A record of training completion is maintained electronically.

Feel free to contact us for further information or questions.

Signed and dated

<Title>