PROJECT DETAILS

Project Name: Data Enclave Implementation Sub Project
Project Type: New Service
Related Projects: Iowa Health Data Resource (IHDR) Project

PROJECT OVERVIEW– WHAT ARE WE DOING

The University of Iowa P3 Board has approved a project to implement a Iowa Health Data Resource (IHDR). This three-year project will establish the IHDR on which future health informatics research initiatives can be built. This foundation will allow us to create and use transformative datasets that will establish the University of Iowa as a national competitor for building interdisciplinary programs for next generation research, curriculum, and student engaged research development. This foundation has 3 parts:

1. An infrastructure of faculty experts that can bridge the gap between investigator-initiated projects and the data gathering and analysis process.
2. A set of new, expansive, and accessible health-related datasets that can be immediately mined for a wide variety of research projects across campus.
3. A data enclave that encompasses the hardware and other infrastructure to keep health information protected while also making it far more accessible.

This sub project will focus on item 3, implementing "a data enclave that encompasses the hardware and other infrastructure to keep health information protected while also making it far more accessible." A data enclave is a secure network through which confidential data, such as identifiable information from the patient electronic medical record, can be processed and analyzed while maintaining patient confidentiality.

The development of the data enclave is intended to establish a data storage platform that allows for utilizing exiting UI advanced computational resources such as Argon and the Interactive Data Analytics Service while aligning with UI Health Care business rules for management of UIHC protected health information (PHI). The storage of data from non-UIHC sources utilizing existing services such as R Drive, RDSS, LSS, and OneDrive will not be impacted by the creation of this service.

Implementing the data enclave will require resources from across ITS. Research Services will be leading the ITS component of the project in collaboration with IT resources from CCOM, ICTS and HCIS.

PROJECT SCOPE

In Scope: Implementation of a Data Enclave for the storage of data from UIHC sources, that align with UI HealthCare business rules for management of UIHC protected health information (PHI).

Out of scope: The storage of data from non-UIHC sources utilizing existing services such as R Drive, RDSS, LSS, and OneDrive will not be impacted by the creation of this service.

Project Requirements / Deliverable

The deliverable for this project is a data enclave that encompasses the hardware and other infrastructure to keep health information protected while also making it far more accessible. To accomplish this the following requirements will need to be met:

End Users Requirements

- Authorized users should be able to mount directories connected with their Hawk IDs from computational systems (such as IDAS and ARGON) in ITS. They should be enabled to process and analyze data on these systems. Authorization will be enabled by placing their Hawk ID in AD groups that will be managed by HCIS/ICTS utilizing IAM access management tools (similar to LSS/HPC queue management today).
- Future access – Because of the amount of storage available in the data enclave, it could become storage area for accessing datasets from researcher PCs. This could be either for additional data exploration or for using low-computational tools to process the dataset.
- Downloading of the source data should be restricted. This will be addressed by policy. If there is technology available to restrict this, that should be added to the requirements. Existing services can be used for retrieving analytics results.
- When users leave the University or when their project ends, their access should be revoked.
- Backups – the data enclave should be treated as scratch space and will not be backed up. However, the potential for backups in the future should be considered. These backups would be disk to disk backups, where data would be periodically mirrored in a different facility.
- Tools – tools used on this data will initially be driven by the environments available in ITS, especially IDAS and Argon. The project will need to document these tools, test and eventually communicate about these tools. If new/additional tools or applications are needed in these environments, they will need to go through the service change management process for approval and prioritization.

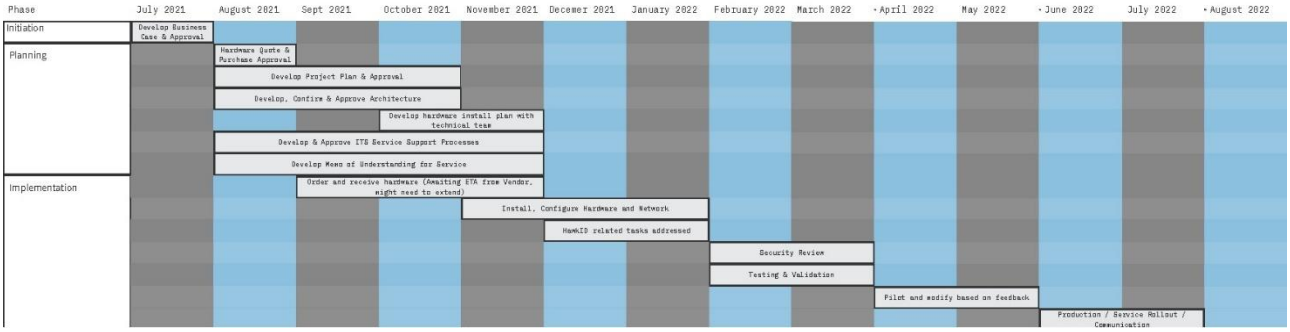
Administrative / Technical Requirements

- Times of operation: The data enclave should operate 24x7, with service levels that address most issues during business hours. (This should be similar to what Argon, IDAS and LSS service levels are). This information will be included in the Memo of Understanding that will be developed.
- Disaster recovery (similar to LSS): Since this is scratch space, there is risk of losing data. For data extracts – ICTS BMI will maintain the queries used to create datasets.
- For computational results – tools will be made available to copy results from the data enclave to secure storage on other systems. Identifying sensitivity and ownership of output data will allow us to determine where data can be stored and how we maintain compliance
- Support: Service requests and service incidents should be managed through the Cherwell suite of tools. Service management will be provided by ICTS BMI. Infrastructure management will be provided by ITS (for LSS operations) and ISO (for network operations). ICTS/HCIS will assist with this and will help trouble shoot infrastructure problems.
- User fees: User fees for the data enclave should mimic what LSS fees are. This will reduce overuse of the resource and provide a funding stream for refreshing the hardware. Initially will mimic LSS fees as service evolves will revisit fee model.
- What data sources will the enclave be utilized for?
    - Data extracts from the ICTS Enterprise Data Warehouse for Research (EDW4R) that includes a subset of the UIHC Epic patient record which may contain PHI. There is a growing number of extant data that are linked to the EDW4R including cancer registry data, patient reported outcomes, external de-identified patient data, dental patient data and UIHC imaging data managed by the IIBI.

- Enclave will undergo a risk assessment and security scan before entering production.
- Connectivity to ITS computational environments will be via existing ITS storage networks.

Project Schedule (When will the work be started/completed)

The project to implement and the Data Enclave hardware and associated service is estimated to take 12 months. The milestones and timeline below outline high level tasks that will be completed.



Project Staffing (Who will perform the work)

There will be two teams working on the Data Enclave implementation. The Core Team meets weekly and is composed of representatives from.

It will be responsible for developing the project plan, monitoring progress, making decisions and consulting with the technical team as needed.

The technical team will meet on an ad hoc basis composed and is composed of a broad group of subject matter experts who will provide technical expertise, guidance, consultation, and recommendations to the core team.

Who	Skill Set/ Team
Heath Davis	Core Team – Institute for Clinical and Translational Science
Mike Frangi	Core Team – Project Management
Joe Hetrick	Core Team - Research Services
Boyd Knosp	Core Team – Carver College of Medicine
Joe Wagner	Core Team – Health Care Info Systems
Nick Bell	Technical Team – Physical Infrastructure
Brian Beninga	Technical Team – ISPO / HCIS Data Center
Kevin Davison	Technical Team – Health Care Information Systems
Stephen Hoffman	Technical Team – ITS Data Center
Brendel Krueger	Technical Team – Research Services
Jordan O'Konek	Technical Team – Identity & Access
Scott Stiegelmeyer	Technical Team – ISPO / HCIS Data Center
AJ Klopp	Technical Team – ITS Network Engineering Services
Mike Ryan	Technical Team – ISPO / HCIS Data Center

Project Stakeholders (Consult with)

Name	Department
Ben Rogers	ITS Enterprise Infrastructure
Robin Springer	ITS Office of CIO
Zach Furst	ITS Information Security and Policy
Steve Fleagle	ITS Office of CIO
Tim Evans	ITS Office of CIO
Mike Noel	ITS Administrative Information Systems
Jerry Protheroe	ITS Enterprise Infrastructure
George Stumpf	ITS Enterprise Infrastructure
Jeff Gillitzer	ITS Enterprise Infrastructure
Genevieve Johnson	RS LSS Service Owner

Project Budget

The P3 project budget for the overall IHDR implementation that was approved by the P3 Board is:

	Year 1	Year 2	Year 3
Faculty Liaisons	$145,442	$150,172	$152,332
Existing Staff	$235,158	$207,318	$153,050
New Staff	$182,453	$374,906	$344,610
Data Enclave	$75,000		
Software & Training	$90,000	$60,000	$50,000
Yearly Total	$728,053	$792,396	$699,992
3 Year Total			$2,220,441

The estimated Data Enclave sub project budget for the (Aim 3 of IHDR project):

Item	Year 1	Year 2	Year 3
*Hardware	$75,000	000	000
*Software & Training	$90,000	$60,000	$50,000
**ITS Labor	$54,375	$1,875	$1,875
**ITS Hardware	$3,000 parts & labor		
Total	$222,375	$61,875	$51,875

*Hardware budget has been approved and funded by the University P3 Board.
*Software and Licensing budget has been approved and funded by the University P3 Board.
Note - Software budget is not specifically for the data enclave. It could be software used in the enclave but could also include data license and software used in other contexts of the IHDR. *ITS Labor=$75/hour

Estimated ITS Labor Hours

Team	Project Labor Estimate	Ongoing Annual Support Hours
RS	~200 hours	~25 hours
PMO	~ 150 hours	0 hours
*EI	~200 hours	0 hours
ISPO	~150 hours	0 hours
IAM	~25 hours	0 hours
Total Hours	~ 725 hours	~25 hours
Total Dollars	~ $54,375	~ $1,875

Communications Plan (How will information be communicated)

Communication related to the implementation of this project will be distributed to the following audiences.

Target Audience	Primary Contact	Purpose/Description of Communication	Communication Method (Push, Pull, Interactive)
OneIT Operations Team	Rachel Napoli	Engage ops team for input and approval of the project plan	Zoom meeting (Interactive), eMail (push)
OneIT Leaders	Rachel Napoli	To inform community about the project	eMail (push)
OneIT Community	Rachel Napoli	To inform community about the project	eMail (push)
ITAdmin Community	Nick Carino-Marek	To update this community about the project and receive feedback	Zoom meeting (Interactive), eMail (push)
IHDR Steering Committee	Boyd Knosp	To inform most of the Research Deans about the project and provide updates on project progress	Zoom meeting (Interactive), eMail (push)
Research Computing Council	Joe Hetrick	To inform the council about the project and provide updates on project progress	Zoom meeting (Interactive), eMail (push)
Stakeholders	Core Team	To provide update to and receive feedback from Stakeholders	Zoom meeting (Interactive), eMail (push)
Technical Team	Mike Frangi	To provide update to and receive feedback from Technical Team	Zoom meeting (Interactive), eMail (push)
IHDR IAI Team (Intercollegiate Advisory and Implementation)	Heath Davis	To provide update to and receive feedback from the IHDR IAI Team	Zoom meeting (Interactive), eMail (push)
ICTS Stakeholder Group	Justin Kahler	To provide update to and receive feedback from the ICTS Stakeholder Group	Zoom meeting (Interactive), eMail (push)
Healthcare Data Governance Team	Mary Jo Duffy	To provide update to and receive feedback from the HC Data Governance Team	Zoom meeting (Interactive), eMail (push)

Issue Tracking and Resolution Plan

A Microsoft Teams site has been created for this project and members of the project team have been added to the site. This will be the primary collaboration and communication mechanism for the Core and Technical Project Team. As issues arise, they will be identified tracked and resolved utilizing the planner and lists components of the Data Enclave Teams site.

Risk Management Plan

The project team has identified an initial list of risks and mitigation strategies. If additional risks are identified they will be identified, tracked, and resolved utilizing the Data Enclave Teams site.

Risk Number	Risk Description	Likelihood (H,M,L)	Impact (H,M,L)	Mitigation Strategy
001	*High lead times for hardware*	*High*	*High*	*Shift timeline to accommodate longer lead times and communicate impact*
002	*Overall workload affecting timeframe and priorities*	*Medium*	*Medium*	*Address in the core team meeting. Escalate if needed to higher level leadership.*
003	*Need to order additional fiber (Up to 52 wk lead times)*	*Medium*	*Medium*	*Brainstorm ways with technical team to accommodate with existing Fiber.*

Change Control Plan (What is the process for managing change)

Requests for change in project scope will be reviewed and assessed by the Core Project Team. If a change is deemed to be substantial in nature it will be reviewed by ITS and HCIS leadership followed by the IHDR Steering Committee. Minor changes that do not have an impact on scope or timeline will be reviewed by the Core Project Team.

Metrics / Key Performance Indicators

P3 IHDR Metrics that impact Data Enclave

- Progress: Being built and in progress with a Target Date
    - Access to computational hardware
    - Adequate hardware for the analyses
- Scholarship
    - # Grants Supported
    - # of MFKs
    - # Publications using resource
        - citing IHDR / P3 Funding (language to be developed)
- Number of researchers using the system be able to breakout by:
    - Role (fac/staff/student) Job title
    - college/org
        - department/unit
    - type of research/ research domain
    - Initial #
    - Annual growth
- # of transformational data sets using this resource (AIM2)
- Reduction in backlog

ITS Data Enclave Metrics

- Reduction in wait times for data so they can compute faster
- Number of researchers using the system initially and annual growth by:
    - College,
    - Department,
    - Type of research they are doing.