

# Update on 2025 Research Informatics

From the Office of the Chief Research Information Officer

David Dorr, MD, MS, CRIO, Professor, co-lead 4.3

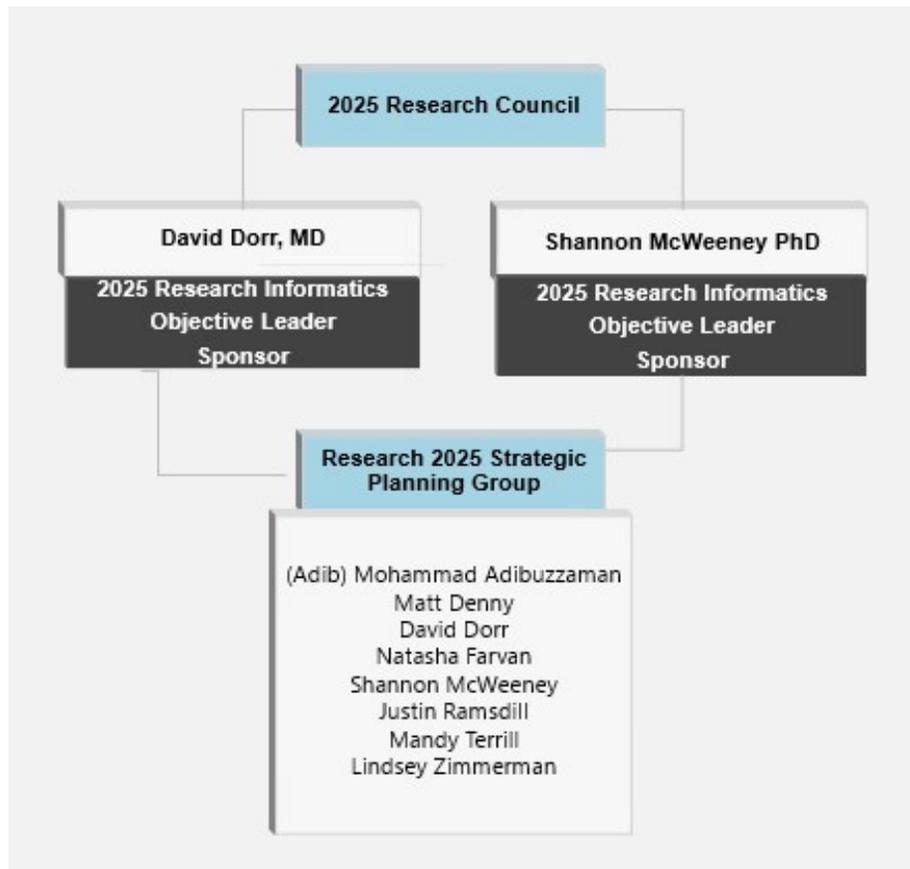
Shannon McWeeney, PhD, Professor, co-lead 4.3

Town Hall 4/4/2022

## Overview of 2025 Research Informatics

| Number            | Tactic   |
|-------------------|--|
| 4.3.1             | Develop our current HPC environment (Exacloud) into a University-wide flexible computing resource in the service of research needs to address critical risks resulting from aging hardware and provide researchers with additional HPC resources for their growing need. |
| 4.3.2             | Design and construct the next generation OHSU Research Data Warehouse to enable rapid and cost-effective data integration prioritized by research, clinical care and education.  |
| Expense           | Budget FY 2022 (ongoing funding 2025)  |
| Operating Expense | \$1,899,555  |
| Capital Expense   | \$527,000  |
| Total             | \$2,426,555  |

# 2025 Research Informatics Charter - structure



Charter created

(See draft)

Kick off held

Process and Procedure defined

- For gaps, submit SBAR
- For projects, submit Project intake forms
- **New projects reviewed at committees: Research Data Governance, Research Computing / Exacloud Executive Committee and Task Force**
- **Informatics Governance Group review as needed**

## Research Informatics 4.3 Tactics

| Number  | Tactics  | Achievement Indicator  | Leaders                                     |
|---------|--|--|---|
| 4.3.1   | Develop our current HPC environment (Exacloud) into a University-wide flexible computing resource in the service of research needs to address critical risks resulting from aging hardware and provide researchers with additional HPC resources for their growing need. | Increase in service utilization<br>User satisfaction<br>Decrease user costs and usability issues.<br>Leverage modern virtualization & "DevOps" techniques to provide users with on-demand computing resources (e.g., web/application/storage servers, etc.). | David Dorr, Shannon McWeeney, Mandy Terrill |
| 4.3.1.1 | Provide secure and compliant open source, flexible cloud computing platform with rapid deployment to support multi-site data sharing and analysis that will support the collaboration needs of research.   | Reduce dependency on aging hardware, Stabilize and expand computing infrastructure in dynamic manner   | Mandy Terrill, Marion Hakanson              |
| 4.3.1.2 | Offer centrally supported open-source (bioinformatics and other research) software and workflow management tools to reduce administrative burden on individual labs  | Reduce informatics systems administration costs for individual labs and duplication of effort. Offer new expanded research computing services  | Mandy Terrill, Marion Hakanson              |
| 4.3.1.3 | Develop training evaluation partition within Exacloud, that will provide a safe environment for new users and decrease their impact on production partitions   | Implementation of the training environment, as well as the supportive educational documentation that is readily available.   | Mandy Terrill, Marion Hakanson              |
| 4.3.1.4 | Improve storage options and value for researchers  |  |   |
| 4.3.2   | Design and construct the next generation OHSU Research Data Warehouse to enable rapid and cost-effective data integration prioritized by research, clinical care and education.  | Improved volume of use; number of users; satisfaction of users; revenue from grants using data resources   | David Dorr, Shannon McWeeney                |
| 4.3.2.1 | Fully staff the RDW engineering team to stabilize operations and establish capacity for new datatype ingestion   | Staff recruitment complete   | David Dorr, Shannon McWeeney, Matt Denny    |

## Research Informatics 4.3 Tactics

| Number  | Tactics   | Achievement Indicator                                       | Leaders                            |
|---------|---|---|------------------------------------|
| 4.3.2.2 | Establish RDW Data Enhancement pipeline for inferred (Natural Language Processing) and curated data, including social determinants of health, patient-reported outcomes and other data  | Pipeline in place and data populated                        | David Dorr,<br>Shannon<br>McWeeney |
| 4.3.2.3 | Using the Research Data Governance Committee, identify data integration priorities and begin integrating data; Initial integration examples include national and state mortality data, clinical genomics, organizational metadata, cost/utilization outcomes data, patient reported outcomes, community-level social determinants of health, and geographic data on health factors and outcomes | Satisfaction of users                                       | David Dorr,<br>Shannon<br>McWeeney |
| 4.3.3   | Provide a mechanism to allow interactive search for data and metadata including quality and provenance.   | Improved volume of use; improved productivity from research | David Dorr,<br>Shannon<br>McWeeney |
| 4.3.3.1 | With the Library, build a Data Inventory Tool that allows data repository holders to store information about their data assets and researchers to search for data.  |   |                                    |
| 4.3.4   | Improve the fitness for use for clinical and research questions and evaluation of integrated data by building standard data quality rules, fixing data quality problems and/or reporting them to data governance, and incorporating them into the relevant repositories.  |   | David Dorr,<br>Shannon<br>McWeeney |
| 4.3.4.1 | Deploy analytic platform and establish Data Quality Dashboard (DQD) and Data Characterization tool (Achilles)   | DQD and Achilles deployed and accessible to all             | David Dorr,<br>Shannon<br>McWeeney |

## Research Informatics 4.3 Tactics

| Number  | Tactics   | Achievement Indicator                               | Leaders                      |
|---------|---|---|------------------------------|
| 4.3.5   | Democratize data access and stimulate innovation by constructing de-identified data repository for education, testing, and observational studies  |   | David Dorr, Shannon McWeeney |
| 4.3.5.1 | Identify options for de-identified data repository that are secure, useful, and standardized.   | Data repository available and used.                 | David Dorr, Shannon McWeeney |
| 4.3.6   | Create a centralized data concierge service, including researcher support and metadata development to support data discovery.   |   | David Dorr, Shannon McWeeney |
| 4.3.6.1 | Hire Data Concierge   | Position Filled                                     | David Dorr, Shannon McWeeney |
| 4.3.6.2 | Data Concierge offers education, guidance, and 1:1 research support   | Number of projects supported; Satisfaction of users |                              |
| 4.3.7   | Facilitate access to relevant data resources and data integration capabilities by establishing policies and procedures for certified users; provide workforce development to enable certification for advanced research data users. |   |                              |
| 4.3.7.1 | With an advanced data scientist corps, use extant curricula and standards to build a certification program.   | Certification offered to advanced users.            | David Dorr, Shannon McWeeney |

# Major accomplishments and ongoing issues

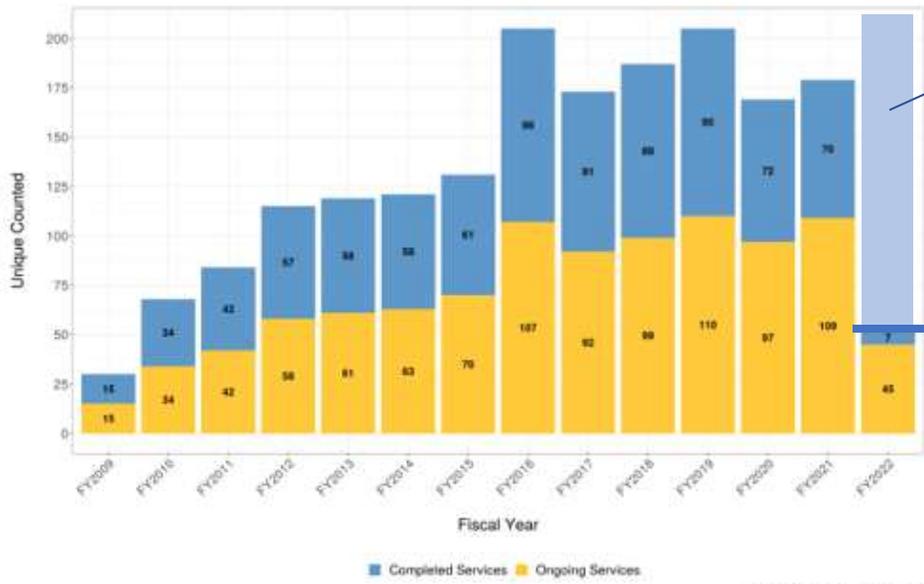
- Metrics – initial improvements, continued needs
- Organization – complex set of services and innovation
  - Staffing roles need reconsideration : Strategic Informatics Program Manager and Research Data Concierge
  - Research Computing Strategy
- Enhancements
  - New data – mortality, social determinants of health, patient-reported outcomes – enhance capacity
  - New ways to find and access – research data catalog, data tools
  - Partners – long term sustainability plan
- Innovation
  - Requirements to use new techniques to access interoperable data – FHIR, Common Data models – need additional support, given changed requirements

## Progress made- Overall metrics

| Objective 4.3 KPI   |                    |  |  |                      |                       |                         |
|---|--------------------|--|--|----------------------|-----------------------|-------------------------|
| Summary   | Service            | Baseline estimate (FY2021)                         | FY2022 Goal                                    | FY -2021             | FQ1 - 2022 (7/1-9/30) | FQ2 - 2022 (10/1-12/31) |
| Total RDW Requests  | RDW                | 179 Total; 70 completed; 109 ongoing               | Increase 5 - 10% = 188 – 197 (~48 / Q)         | 179                  | 63                    | 103                     |
| Unique RDW Requestors: Distinct count of PI per specified time period | RDW                | 96   | Increase 5 - 10% = 101 – 106                   | 96                   | 54                    | 74                      |
| Paid Users  | Research Computing | 199  | 15 - 20% increase = 230-240                    | 199                  | 240                   | 259                     |
| Queue Time Average  | Research Computing | 41min  | 10% decrease                                   | 41                   | 30                    | 57                      |
| Total Requests: sum of engagement requests                            | RDC                | 0  | New service                                    | 18                   | 28                    | 56                      |
| Unique Requestors:  | RDC                | 0  | New service                                    | 18                   | 27                    | 53                      |
| % Very Satisfied  | RDW                | 64.50%   | 80%  | 64.5                 | pending               | 100%                    |
| % Knowledgeable of Exacloud   | Research Computing | (28 out of 39 responses) 72% know what exacloud is | 80%  | 72%                  | annual collection     |                         |
| Grant Revenue (Direct/Indirect)                                       | RDW                | Direct = 119,475,153<br>Indirect = 27,382,137      | Maintain for 2022 (10 - 15 % increase by 2025) | \$119.5M/<br>\$27.4M | \$21.6M/\$5.8M        | \$41.9M/ \$11.6M        |
| Grant Revenue (Direct/Indirect)                                       | Research Computing | Total = \$90,567,121.79                            | Maintain for 2022                              | \$90.5 M             | pending               | pending                 |

# Metrics: RDW over time

Figure 1: RDW Volume FY2008-FY2022\*



On track for 206

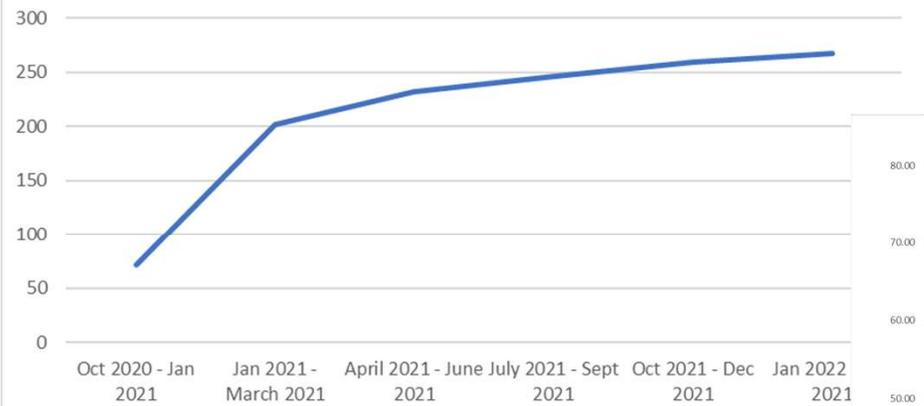
FY 2022 Goal

Increase 5-10%

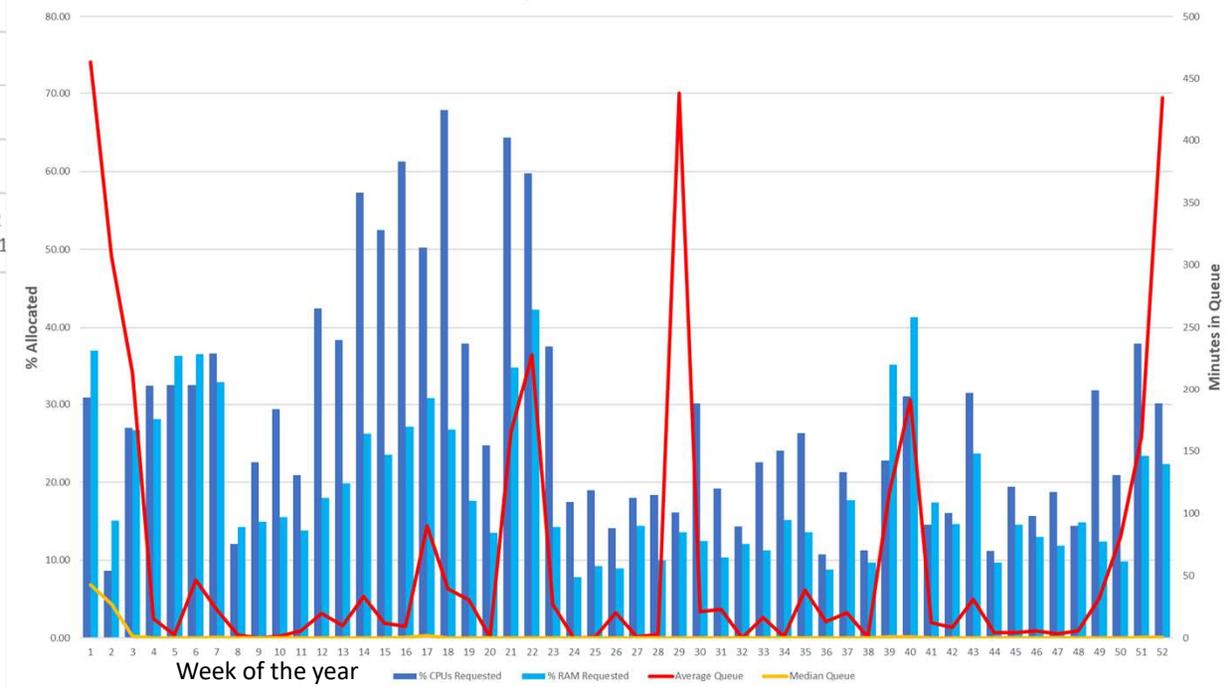
\*FY2022 Currently in Progress.

# Research Computing : Paid Users and Queue Times

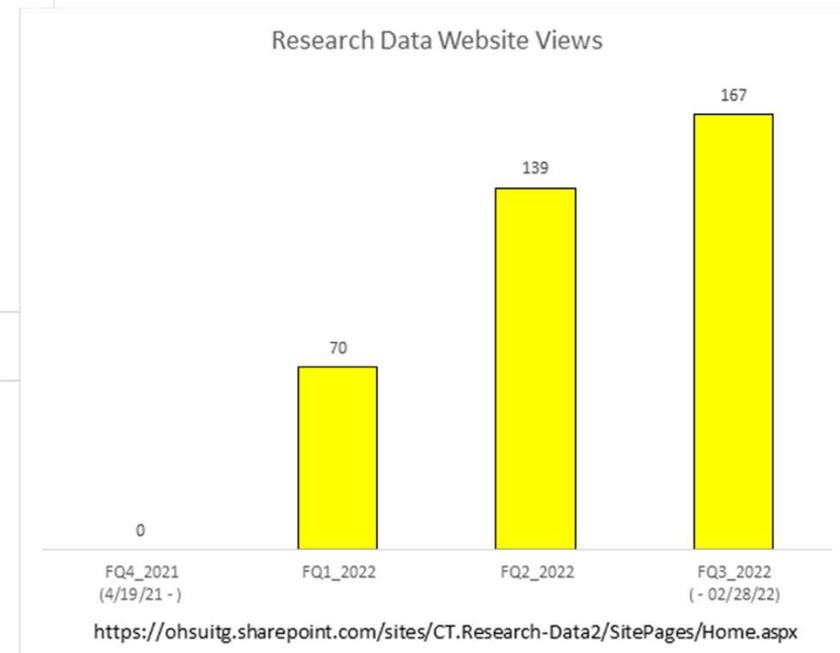
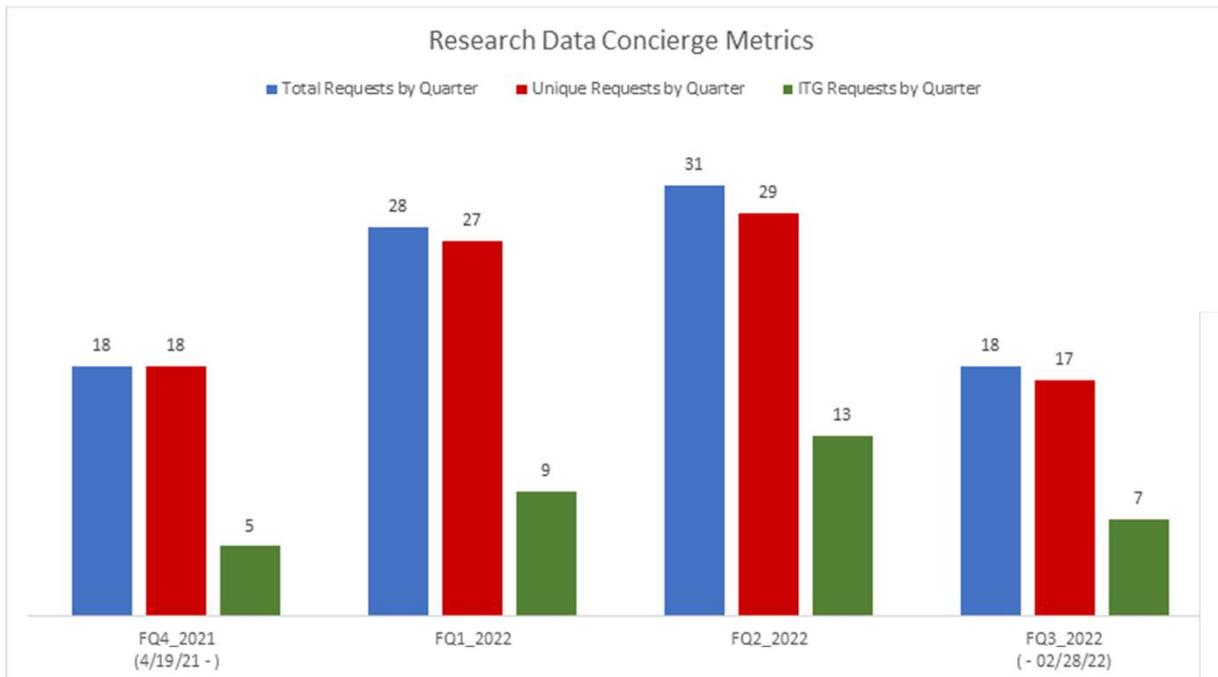
Paid Users ACC



%Busy and Queue Time – FY2021



# Progress made – Research Data Concierge (RDC) Metrics



# Research Data Concierge changes

- Policy / procedure for use of Slicer/Dicer and Reporting Data Warehouse (and related) in Epic
- Policy for quality improvement style support with research components
- Data expansion – death data, social determinants of health data
- Available for complex needs!
- [researchdata@ohsu.edu](mailto:researchdata@ohsu.edu)

## What's next and what's needed

- Research Computing- build roadmap and toolkit
- Research Computing – 3-5 year for hardware replacement
- RDW – multiyear transformation of data capabilities
- Request for ongoing support/feedback
- Communication opportunities – share your stories!
- Strategic Plan for Partner and Affiliate Research
- Innovation support

# Quick guide to services and engagement

| Service   | Link / Contact   |
|---|--|
| OCTRI Navigator ( <b>unsure? start here</b> )   | <a href="https://www.ohsu.edu/octri/your-one-stop-shop-questions-about-clinical-and-translational-research">https://www.ohsu.edu/octri/your-one-stop-shop-questions-about-clinical-and-translational-research</a>    |
| Epic for Research Team  | <a href="mailto:EpicResearchTeam@ohsu.edu">EpicResearchTeam@ohsu.edu</a>   |
| Clinical / Translational Research Informatics <ul style="list-style-type: none"><li>- Research Data Warehouse</li><li>- Databases</li><li>- Informatics tools</li></ul> | <a href="https://www.ohsu.edu/octri/powering-innovation-state-art-informatics">https://www.ohsu.edu/octri/powering-innovation-state-art-informatics</a> ; <a href="mailto:octrihelp@ohsu.edu">octrihelp@ohsu.edu</a> |
| Research Computing <ul style="list-style-type: none"><li>- Cloud or on premises</li><li>- Both PHI and non-PHI</li></ul>  | <a href="mailto:acc@ohsu.edu">acc@ohsu.edu</a>   |
| Clinical reporting <ul style="list-style-type: none"><li>- Quality improvement</li></ul>  | <a href="https://bridge.ohsu.edu/community/bioc/SitePages/Request.aspx">https://bridge.ohsu.edu/community/bioc/SitePages/Request.aspx</a>  |
| Research Data Network / Research applications   | <a href="mailto:risxsupport@ohsu.edu">risxsupport@ohsu.edu</a>   |
| Research Data Concierge   | <a href="mailto:researchdata@ohsu.edu">researchdata@ohsu.edu</a>   |

Thank You