

# Agenda V: Overall HPD Data Release Process

*Jasmine Neeley, Research Data Analyst II, HCAI*

*James Yi, Attorney, HCAI*

*Terrence Nolan, Healthcare Data Request Unit Supervisor, HCAI*

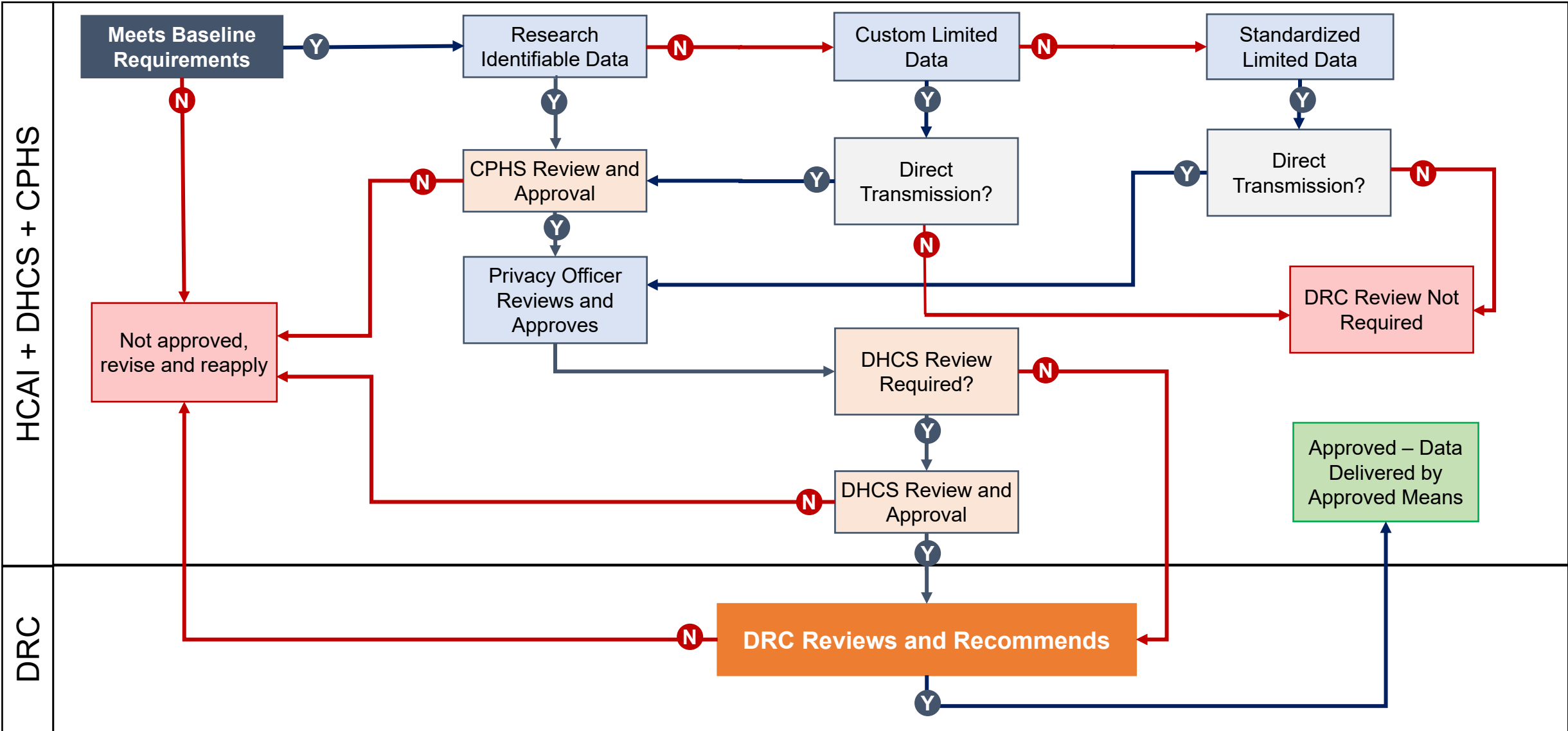
*Wade Luele, HPD Consultant, HCAI*

# Workflow

# Standard HCAI Review

- Analyst
- Supervisor
- Privacy Officer
- Final Approver
- Science Advisor
  - As needed

# Data Request Decision Tree and Flow Diagram



# Request Form

# Application Sections (high-level)

## Data Security

- Method of Data Access – Enclave or Direct Transmission
- Data User information – Contractors?
- Data Security Plan
- Description of how results and key findings will be shared
- Small cells

## Additional Information

- Does the project offer significant opportunities to achieve HPD goals?
- All records vs. Subset of Records
- Minimum Data necessary
- Planned Data Products
- Applicant History, including with Breaches/Convictions/Civil Actions

# Application Sections (high-level)

## Overview

- Data Applicant information, qualifications, and prior experience
- Alignment of request with HPD purpose, goals, requirements
- Submit data to HPD program?
- Requesting a Fee Reduction?
- New or Supplemental Request?

## Project Specific

- Description of project purpose, goals, and data use/methods
- Requested HPD data product, timeframe of data, and desired delivery format
- List of data elements with justification/reason based on project purpose
- DHCS data requested?
- Proposed Linkages to other sources with justification/reason

## HPD Data Request

### \*Request Type ⓘ

Select *New* if this is a Research Data Request that has not been previously submitted in this Data Request Portal. ✕  
If you have submitted a Research Data Request before, and are requesting additional years, or are adding additional information to your existing request, select *Supplemental*.

New  Supplemental

### \*Data Request Type

Limited Data  Research Identifiable Data

### \*Requested Limited Data Set

Standard Limited  Custom Limited

### \*Which method of data access are you requesting?

Secure Data Enclave Environment  Direct Transmission

### Data Applicant Information

#### \*First Name

#### \*Last Name

#### \*Title

#### \*Entity Type

#### \*Organization Name

#### \*Department

#### \*Address

#### \*Phone number

# HPD Data Request Form

(external view)

### Research Identifiable Data Product Selection Menu

#### \*Requested Data Sources

- Claims Medical
- Claims Pharmacy
- Member Enrollment
- Provider Information

#### \*Please indicate the timeframe of requested data for Claims Medical

ex. 2019, 2021

#### \*Please indicate the timeframe of requested data for Claims Pharmacy

ex. 2019, 2021

#### \*Are you requesting access to Medi-Cal variables?

Yes  No



# HPD Data Request Form

\* Describe how use of the requested data aligns with the goals of the HPD Program. Identify the specific goals your request aligns with below.

\* HPD Program Goals

- Provide public benefit for Californians and the state while protecting individual privacy.
- Increase transparency about health care costs, utilization, quality, and equity.
- Inform policy decisions on topics including the provision of quality health care, improving public health, reducing disparities, advancing health coverage, reducing health care costs, and oversight of the health care system and health care companies.
- Support the development of approaches, services and programs that deliver health care that is cost effective, responsive to the needs of Californians, and recognizes the diversity of California and the impacts of social determinants of health.
- Support a sustainable health care system and more equitable access to affordable and quality health care for all.

### Data Access

\* Please list all individuals who are anticipated to observe, use, or access confidential data, including yourself

**Add**Remove All

Actions	Data User Name	Data User Organization	Data User Email Address	Data User Phone Number	Data
No data to display					

\* Will a contractor be used for this project?

- Yes
- No

### Project Security

\* Have you previously requested HPD data?

- Yes
- No

\* Has the data applicant been involved in a data breach, or other similar issue, in which Personally Identifiable Information (PII) and/or Protected Health Information (PHI) was misused or improperly disclosed, within the last 7 years?

- Yes
- No

\* Has the data applicant been involved in any criminal convictions or substantiated violations of California, or other state, or federal laws regarding fraud, theft, data breach, data misuse or related offenses in the past 7 years? This includes civil or administrative penalties, civil judgements, or disciplinary actions.

- Yes
- No

\* Please describe your experience with data security and the protection of large sets of confidential information.

# Additional Documents

## **During Review**

- Variable Justification Grids
- CPHS Protocol
- Data Requestor CV\*
- Contract\*

## **Final Approval**

- CPHS Approval letter
- Data Use Agreement (DUA)
- Signed Request Form

# Data Request Portal

# Data Request Portal - Dashboard

☰ ☐ Tasks Dashboard ▾

Tasks assigned to me						
Number ▲	Subject	Short description	Name	Company	Assignment group	Assigned to
<a href="#">CSTASK0017532</a>	DRC First Review	HPD Data Request	Chris Martin	HCAI	HCAI-HPD-DRC	Nuriel Moghavem

« ‹ 1 to 1 of 1 › »

Unassigned Tasks in My Group						
Number ▲	Subject	Short description	Name	Company	Assignment group	Assigned to
<a href="#">CSTASK0017533</a>	DRC First Review	HPD Data Request	Chris Martin	HCAI	HCAI-HPD-DRC	(empty)

« ‹ 1 to 1 of 1 › »

Task - CSTASK0017784

State: Open

Assignment group: HCAI-HPD-DRC

Assigned to: Nuriel Moghavam

Parents: CS0003120 > CSTASK0017784

Number: CSTASK0017784

DRC Recommends Approval: -- None --

Due date:

Subject: DRC First Review

Description: The Data Release Committee (DRC) will review the data request and provide a recommendation.

# DRC Request Task

Work Notes

Work notes: Work notes

Activities: 1

System

Assigned to: Nuriel Moghavam

Field changes • 2023-10-11 09:18:08

State: Open

Assignment group: Open, Closed and Reviewed, Closed and Rejected

Assigned to: Nuriel Moghavam

SLAs | **Tasks (1)** | Approvers | Phone | Appointments | Emails (6) | Related Data Requests (6) | Problems | Attached Knowledge | Knowledge Gaps | Child Data Requests | Contracts (1) | HPD Data Sources

Tasks New Search Number Search 1 to 1 of 1

Parent = CS0003120

	Number	Subject	State	Created	Assigned to	Business duration	Duration
	CSTASK0017784	DRC First Review	Open	2023-10-11 09:18:08	Nuriel Moghavam		

Actions on selected rows...

1 to 1 of 1

# DRC Request View

Case CS0003120 [Case view]

Manage Attachments (2): Charles Darwin CV.docx [rename] [download] CPHS Protocol.docx [rename] [download]

Timeline - Began 6 days ago - Updated 6 days ago

6 days With Agent  
0 seconds With Customer

Number	CS0003120		
* Contact	My Test Account	🔍	ℹ️
Account	HCAI	🔍	ℹ️
Product	HCAI Research Direct Transmission	🔍	ℹ️
Asset		🔍	
Data Due Date		📅	
Description			

Opened	2023-10-11 08:46:52
State	Open
Stage	DRC Review
Supplemental Request	<input type="checkbox"/>
Requires Privacy Officer Review	<input type="checkbox"/>
* Assignment group	HCAI-HPD-Analysts
Assigned to	Jasmine Neeley

# DRC Request View

< Case - CS0003120

Variables Notes Resolution Information

### Variables

- \* Request Type
  - More information
  - New
- \* Data Request Type
  - Research Identifiable Data
- \* Which method of data access are you requesting?
  - Direct Transmission

### Principal Investigator Information

Principal Investigator First Name

Charles

Principal Investigator Last Name

Darwin

< Case - CS0003120

### Principal Investigator Information

Principal Investigator First Name

Charles

Principal Investigator Last Name

Darwin

Title

Head Researcher

PI CV/Resume

Charles Darwin CV.docx

\* Organization Type

Research organization

Organization

Darwin Institute for Health Research

Department

Claims Data Research Group

Address

1234 Galapagos Way

# Request Notes and Communications

The screenshot displays a web interface with the following components:

- Navigation Tabs:** Variables, Notes, Resolution Information.
- Watch list:** Includes a lock icon and a refresh icon.
- Work notes list:** Includes a lock icon and a refresh icon.
- Work notes:** A text input field with a yellow vertical bar on the left, a scroll bar on the right, and a list icon on the far right.
- Additional comments:** A checkbox and a "Post" button.
- Activities: 5:** A list of messages:
  - Message 1:** From "MA My Test Account" (Additional comments • 2023-10-17 08:51:08). Content: "Hello, Thank you for reviewing my application. Please let me know if you have any questions. Thank you, My Test Account".
  - Message 2:** From "JN Jasmine Neeley" (Additional comments • 2023-10-11 09:31:11). Content: "Hello applicant, Thank you for applying. Thank you, Assigned Analyst".
  - Message 3:** From "JN Jasmine Neeley" (Work notes • 2023-10-11 09:30:41). Content: "This applicant will need xyz".



# Data Use Agreement (DUA)

# DUA Process

Currently, it is planned to have a DUA form consisting of a cover sheet and exhibits with standard provisions based on the type of data request.

The approved application will be incorporated as an exhibit into the DUA. The DUA will have additional provisions about termination, amending/supplementing a data request, data destruction, liability, and enforcement provisions.

HCAI anticipates requiring some DUA provisions by regulation. Other standard provisions may be altered based on the circumstances of the data applicant or data use.

# Pricing Schedule

# Pricing Schedule Overview

- Pricing schedule is based on market research and historical experience
- Pricing schedule for **Data Enclave** use is based on cost to the state, not value of the data
- Pricing schedule for **Direct Transmission** is based on (a) value of the data, (b) incentivizing data enclave use, and (c) providing funding contribution to ongoing program operations
  - Ongoing HPD Funding Report contemplated as much as 5% (\$1m) of program funding coming from data user fees
- HCAI will need to monitor actual demand, but *rough* estimates are:
  - Up to 50 project requests per year
  - **Data Enclave** projects will be at least \$17,000 and up to \$25,000 per year
  - **Direct Transmission** projects can range from \$75,000 to \$200,000 or more (one-time fee)
- HCAI plans to charge a nominal (\$100) non-refundable application fee, which can be applied to the final data costs

# Price Reduction Overview

- HCAI will plan to have a minimum of \$600k available annually for **Data Enclave** price reductions
  - This amount may increase or decrease in the future, based on program funding
- **Data Enclave** price reductions will be granted on a first come, first served basis.
- At between \$17k and \$25k for a typical **Data Enclave** project, HCAI could fully subsidize between 24 and 35 projects annually.
  - A larger number of projects could be subsidized for partial pricing reductions.
- Any **Data Enclave** project can apply for consideration for a full price reduction but will need to provide sufficient justification for the full reduction and may be asked to provide additional supporting documentation.
- Proposed price reduction categories:
  - Financial Hardship
  - Data Innovation
  - High Priority Areas

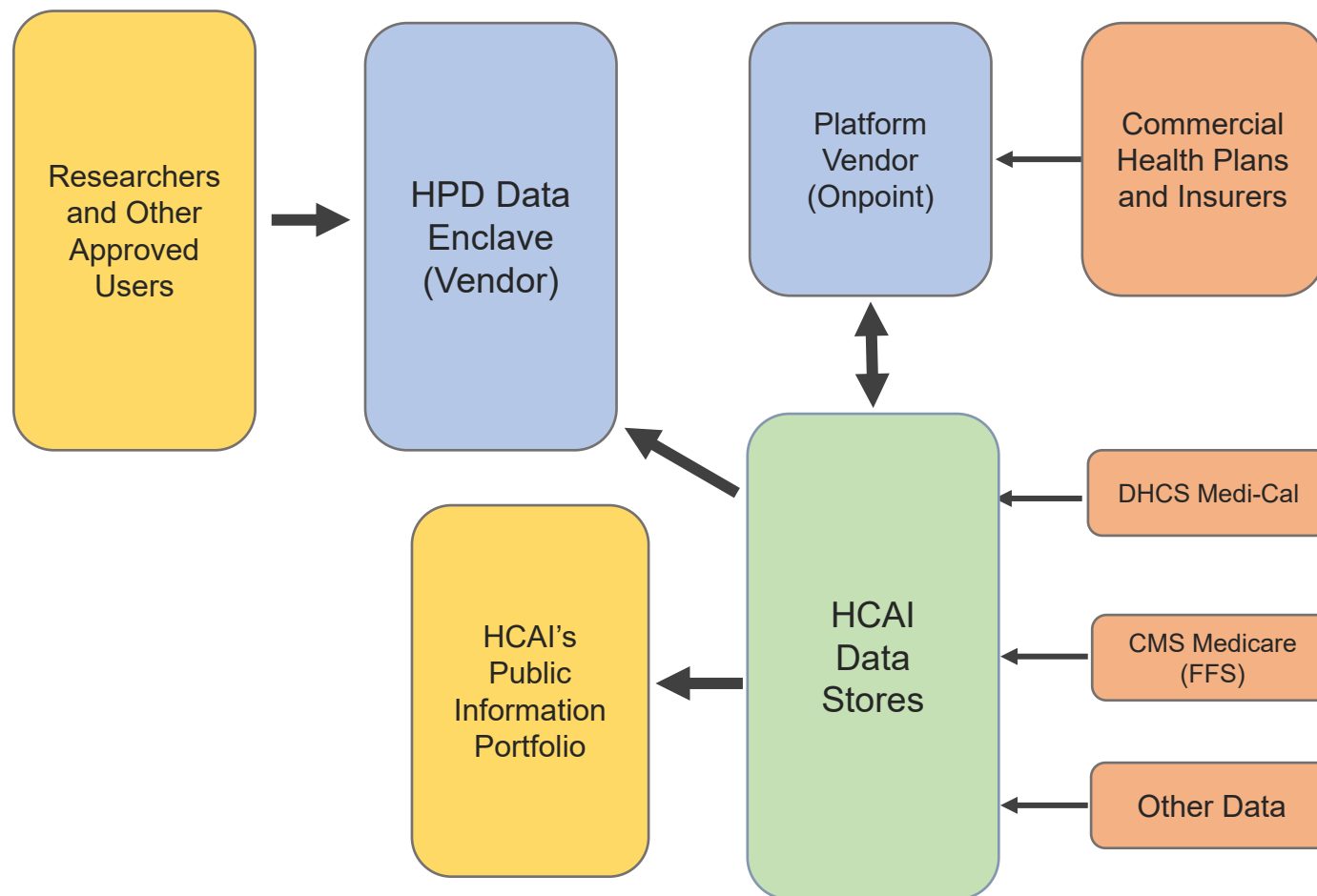
# Price Reductions by Category

Category	Sub-Category	Proposed Reduction Amount
<b>Financial Hardship</b>	Student Requestors	High
	Academic Fellows	Low
	Government Organizations	Low
	Consumer Group	High
	Other non-profit organization	Medium
	Data Provider/Submitter (experiencing financial hardship)	High
<b>Data Innovation</b>	Open-sourced code or code sharing with HCAI	Low
<b>High Priority Area</b>	Equity, Workforce, Behavioral Health	High

# Data Enclave

# HPD Enclave

- Secure environment for accessing potentially identifiable information
- Has controls so users can only access the data they are approved to see
- Protects against unapproved use of personal information





# What is an Enclave?

- Centralized service to remotely access sensitive data
- Houses data in a secure and scalable environment
- Protects data in accordance with state and federal security and privacy rules
- Data is accessed via a virtual machine launched from the user's computer
- Virtual machine operates on a remote server and is controlled by Enclave administrator
- Data products can only be downloaded to user's computer with administrator authorization

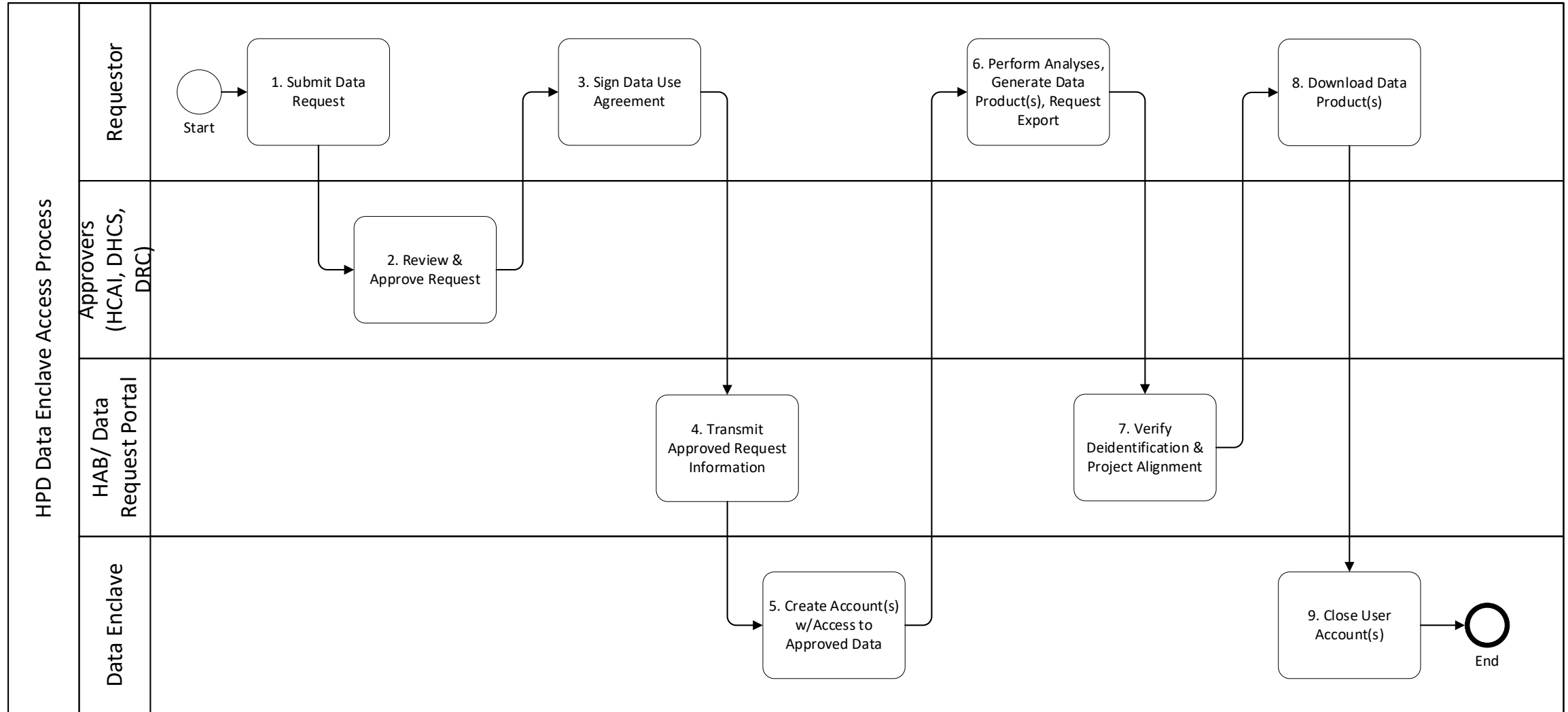
# How is an Enclave Used?

- Makes approved data available to approved users for approved projects
  - Each project's data are segregated from other projects
  - Will allow users to upload and link their own datasets
- Will include access to:
  - Standard tools (Python, SQL, R, Microsoft Office)
  - Additional tools (SAS, Stata, Tableau)
- Provides scalable computing power to perform analysis on a large volume of data

# Additional Enclave Info

- Will support single-user and multiple-user projects
- Multi-user projects include access to shared drive space to exchange code, reports, and tables with other team members
- Users will be able to create tables within the project space that contain data they import, and/or the output from their queries
- Users will receive training and technical assistance from the Enclave vendor, including information on system enhancements and any changes to the HPD datasets

# HPD Enclave Access Process

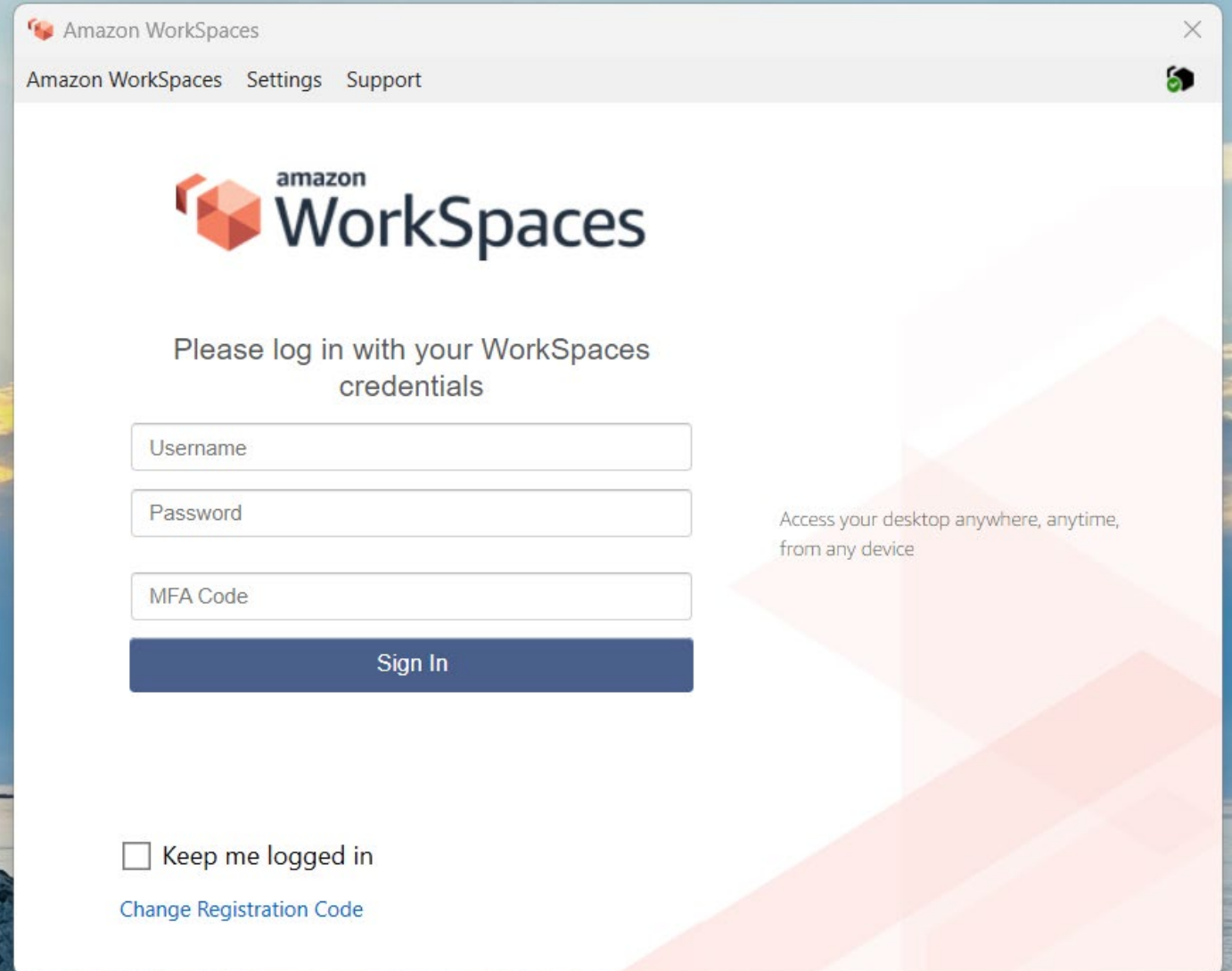


# Draft Enclave Screenshots

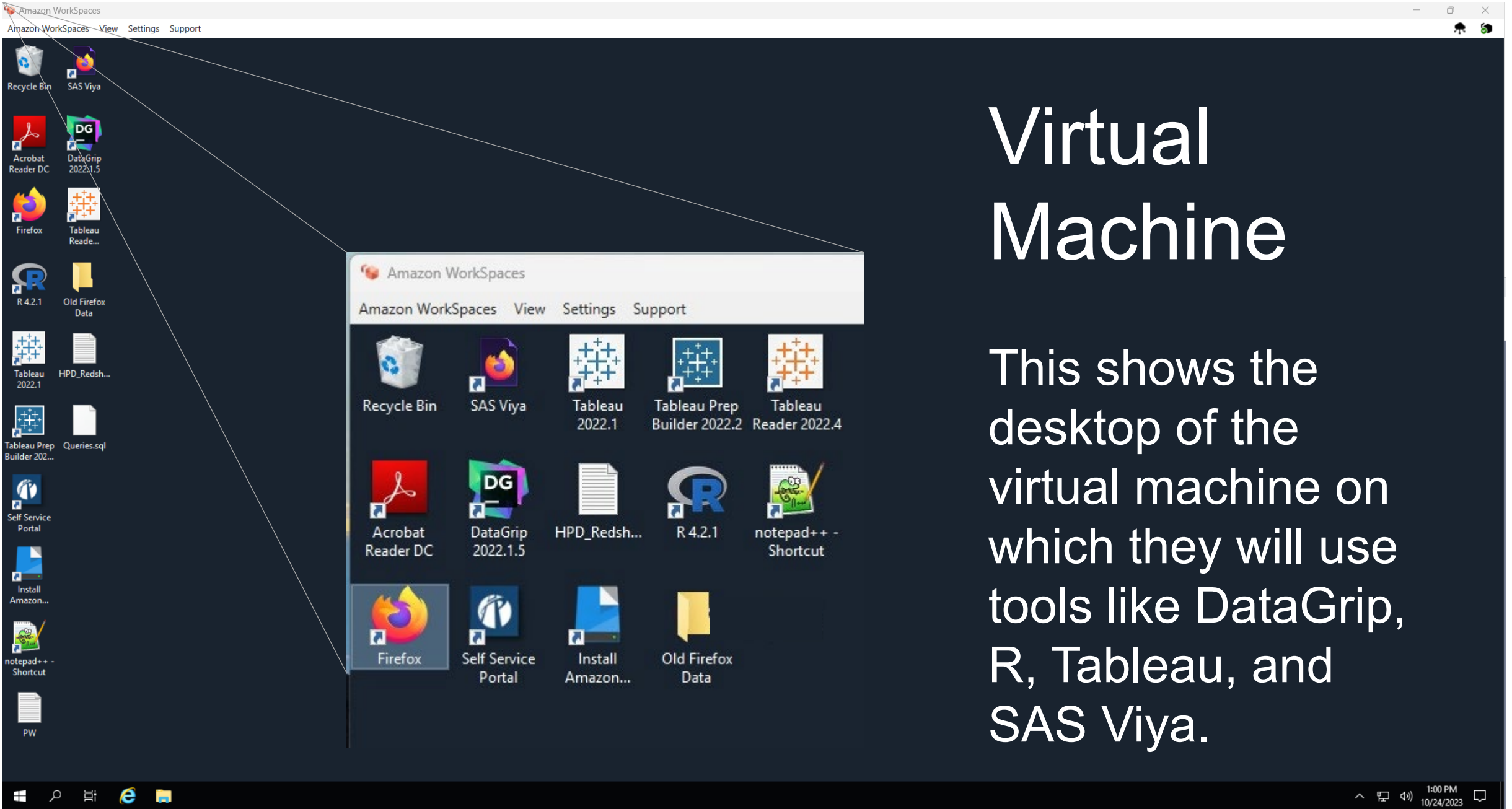
Currently in development and testing

# Logging In

This shows the user logging into the virtual machine they will use to access the HPD Enclave.



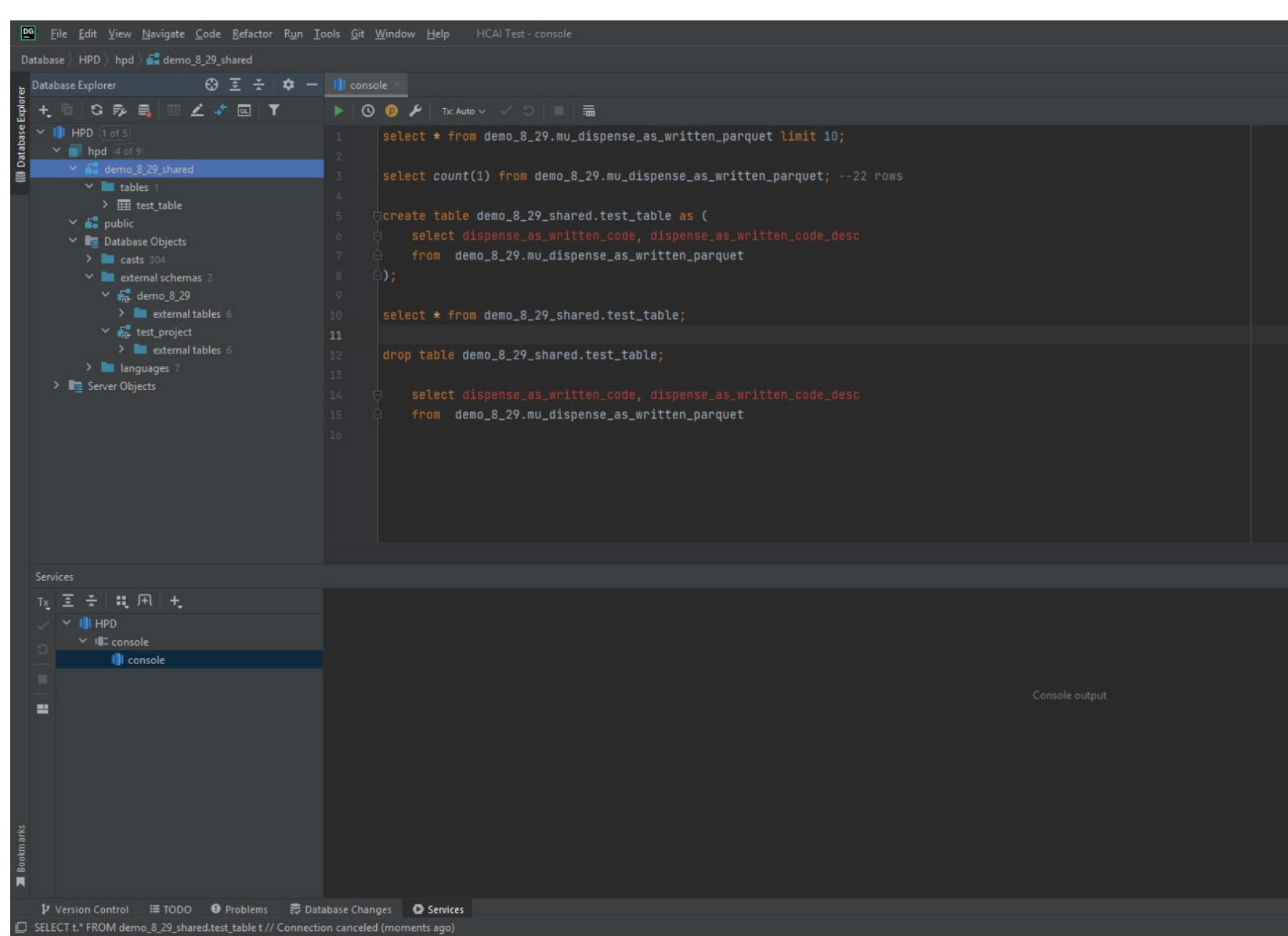
The screenshot shows the Amazon WorkSpaces login page. At the top, there is a navigation bar with "Amazon WorkSpaces", "Settings", and "Support" links. The main content area features the Amazon WorkSpaces logo and the instruction "Please log in with your WorkSpaces credentials". Below this, there are three input fields: "Username", "Password", and "MFA Code". A dark blue "Sign In" button is positioned below the input fields. At the bottom left, there is a checkbox labeled "Keep me logged in" and a link for "Change Registration Code". On the right side, a light blue box contains the text "Access your desktop anywhere, anytime, from any device". The background of the page is a light blue and white geometric pattern.



# Virtual Machine

This shows the desktop of the virtual machine on which they will use tools like DataGrip, R, Tableau, and SAS Viya.

Note: Browser applications in the Enclave cannot connect to the internet.



# Writing Code

This is an example of DataGrip, where users can write SQL to analyze the data. Data sources are on the left. Queries are written in the top-center box.