



# R SHINY APPLICATION FOR PHUSE OPEN DATA REPOSITORY

---

PHUSE Virtual US 2021 – The Clinical Data Science Conference June 14<sup>th</sup> – 18<sup>th</sup>

Hanming Tu, VP, CIT & DBA, Frontage Laboratories, Inc



# ABSTRACT

---

- This demonstration will show a R Shiny application developed for providing a web interface to easily connect, access and review datasets in Pharmaceutical User Software Exchange ('PHUSE') open data repository ('PODR'). The R Shiny application included in this package allows users to connect to 'PODR' and access over 260 open data sets. The open-source packages used to create this R Shiny application will be described and the technique used to access a backend database and develop this R Shiny application will be demonstrated.
- Here are related links:
  - ✓ Package podr info: <https://cran.r-project.org/web/packages/podr/index.html>
  - ✓ Package vignette: <https://cran.r-project.org/web/packages/podr/vignettes/about-podr.html>
  - ✓ R Shiny app hosted at ShinyApps website: [https://geotiger.shinyapps.io/01\\_podr/](https://geotiger.shinyapps.io/01_podr/)



# INDUSTRY OPEN-SOURCE INITIATIVES

---

## ➤ Open Data Initiatives:

- ✓ Big tech companies like Microsoft started the Open Data Initiative a few years back in addressing business data<sup>1</sup>
- ✓ The openFDA provides an elastic search API to enable the public to access FDA data about products like drugs, devices, and foods <sup>2</sup>.
- ✓ The Test Data Factory in Data Visualization and Open Source Technology (DVOST) working group in PHUSE works on automated generation of synthetic SDTM and ADaM data to further standards development.

1. <https://www.microsoft.com/en-us/open-data-initiative>

2. <https://open.fda.gov/apis/>



# PHUSE OPEN DATA REPOSITORY (PODR)

---

## ➤ What is PODR:

- ✓ It gathers over 260 sets of publicly available health-related raw data distributed by academic and government sources<sup>1</sup>.
- ✓ It converts various data formats (CSV, Excel, XML) and different file layouts from these raw data sets into one format and stored in a relational database
- ✓ It integrates all those disparate datasets from multiple sources into an easy-to-search system
- ✓ It gives users a single interface to search across all available datasets.

1. <https://github.com/phuse-org/PODR>



# R `podr` PACKAGE and APPLICATION

- The R `podr` package was initially developed during PHUSE hackathon in 2020
- The following are the main packages used to develop R Shiny application:
  - ✓ Package **shiny**: is an R package that makes it easy to build interactive web apps straight from R. You can also use Shiny server to set up a web server.
  - ✓ Package **shinydashboard**: provides an easy way to create dashboard.
  - ✓ Package **shinyjs**: lets you perform common useful JavaScript operations in Shiny apps that will greatly improve your apps without having to know any JavaScript
  - ✓ Package **shinyBS**: Twitter Bootstrap Components for Shiny to add additional functionality and interactivity to your Shiny applications
  - ✓ Package **Hmisc**: contains many functions useful for data analysis, high-level graphics, utility operations, functions for computing sample size and power, importing and annotating datasets, variable clustering, etc.
  - ✓ Package **rhandsonable**: is a `htmlwidget` based on the `handsontable.js` library. Handsontable is a data grid component with an Excel-like appearance and with powerful features like data validation, sorting, grouping, etc.
  - ✓ Package **DT**: provides an R interface to the JavaScript library `DataTables`, providing filtering, pagination, sorting, and many other features in the tables
  - ✓ Package **stringr**: is simple, consistent wrappers for common string operations.
  - ✓ Package **tibble** is the central data structure for the set of packages known as the `tidyverse`, including `dplyr`, `ggplot2`, `tidyr`, and `readr` and provides stricter checking and better formatting than the traditional data frame.
  - ✓ Package **Rpostgres** is fully 'DBI'-compliant 'Rcpp'-backed interface to [PostgreSQL](#), an open-source relational database.
  - ✓ Package **assertthat** is an extension to `stopifnot()` that makes it easy to declare the pre and post conditions that your code should satisfy, while also producing friendly error messages so that your users know what's gone wrong.



# R podr Shiny App: WEB INTERFACE

- A web-based R application usually has three components: user interface, server and run-time.

```
ui <- dashboardPage( ... )
server <- function(input, output, session) { .. }
shinyApp(ui, server)
```

- The shinydashboard package provides an easy and simple way of laying out the header, sidebar and page of a dashboard

```
header <- dashboardHeader(title = "Display PODR Datasets " )
sidebar <- dashboardSidebar(
  sidebarMenu(id = "tab1",
    menuItem("PODR", icon = icon("cog"),
      menuSubItem("PODR in GitHub", href = 'https://github.com/phuse-org/PODR', newtab = TRUE)
      , menuSubItem("About this Package", href = 'https://github.com/TuCai/podr', newtab = TRUE)
      , menuSubItem('Source
Code', href='https://github.com/TuCai/podr/blob/master/inst/apps/01_podr/app.R', newtab = TRUE)
    )
    , style = "background-color: blue; "
  ))
```



# R podr Shiny App: CONTROL INPUTS

- R provides three functions to interact with user's inputs
  - ✓ Function **reactive**: allows a user to monitor the status of an input or other changing variable and return the value to be used elsewhere in the code
  - ✓ Function **observe**: is similar reactive, the main difference is that it does not return any values to any other environment besides its own, and it is not lazy.
  - ✓ Function **observeEvent**: is an event trigger. It watches one event or change in a variable, and then fires when the event happens.

```
observeEvent(input$tab4, {  
  updateTextInput(session, "col", value = NA)  
})  
observe({  
  if(input$showpanel == TRUE) {  
    js$showSidebar()  
  } else {  
    js$hideSidebar()  
  }  
})
```



# R podr Shiny App: DISPLAY DATA

- Two functions in `podr` package were used:
  - ✓ Function `conn_podr`: to make a connection to PHUSE Open Data Repository (PODR)
  - ✓ Function `read_podr`: to read a dataset from PODR
  - ✓ Function `datatable`: from DT package is used to display the context

```
get_conn <- reactive ({
  validate(
    need(input$username != "", "Please provide Database
User Name.")
  )
  validate(
    need(input$userpwd != "", "Please provide Database
User Password.")
  )
  req(input$username)
  req(input$userpwd)
  conn_podr(username = input$username, userpwd =
input$userpwd)
})
```

```
output$tabLogin <- renderUI({
  tabPanel("Login"
    , div(id = "form"
      , style="display:inline-block"
      , textInput("username", "Database User Name **", value = "phuse_su67e99huj" )
      , bsAlert("alert")
      , passwordInput("userpwd", "Database User Password **" )
      , submitButton("Show", icon("refresh"))
    )
    , hr()
    , h1("Public Tables")
    , DT::dataTableOutput("DT1")
  )
})
```



# R *podr* Shiny App: DOWNLOAD and ACCESS

- The *podr* package is hosted in GitHub and accepted by Comprehensive R Archive Network (CRAN).
- The *podr* R Shiny app is published to [https://geotiger.shinyapps.io/01\\_podr/](https://geotiger.shinyapps.io/01_podr/)

## How to get R *podr* Package

In R or RStudio

```
install.packages("devtools")  
library(devtools)  
install_github("TuCai/podr")
```

Or directly install from CRAN

```
install.packages("podr")
```

## How to run R *podr* Application

Run the application in RStudio

```
library(podr)  
start_app(n=1)
```

Run the application from hosted website

- [https://geotiger.shinyapps.io/01\\_podr/](https://geotiger.shinyapps.io/01_podr/)
- It is a stateless app and does not retain any information on the website.



# R podr Shiny App: Login Page

- You need to obtain a user name and password for accessing the dataset
- With the podr R Shiny app, you can use the default user to access all the datasets

Login ReadMe Table Show

Database User Name \* phuse\_su67e99huj

Database User Password \* .....

Show

## Public Tables

Show 10 entries

Search:

|    | table_name                    | libname              | dataset      |
|----|-------------------------------|----------------------|--------------|
| 1  | cms_npi                       | cms                  | npi          |
| 2  | cms_npi_provider_taxonomy     | cms_npi_provider     | taxonomy     |
| 3  | covid19_beat19_daily_survey   | covid19_beat19_daily | survey       |
| 4  | covid19_beat19_enrollment     | covid19_beat19       | enrollment   |
| 5  | covid19_codr_embeddings       | covid19_codr         | embeddings   |
| 6  | covid19_codr_metadata         | covid19_codr         | metadata     |
| 7  | ct_aact__readme_first         | ct_aact__readme      | first        |
| 8  | ct_aact_baseline_counts       | ct_aact_baseline     | counts       |
| 9  | ct_aact_baseline_measurements | ct_aact_baseline     | measurements |
| 10 | ct_aact_brief_summaries       | ct_aact_brief        | summaries    |

Showing 1 to 10 of 263 entries

Previous 1 2 3 4 5 ... 27 Next



# R podr Shiny App: ReadMe Page

- ReadMe page allows you to display readme file for each dataset
- You can search the readme content

[Login](#) [ReadMe](#) [Table](#) [Show](#)

Library Name:

FDA\_DRUGSATFDA\_\_README\_FIRST

Show  entries Search:

|                    | V1                                                                                                                                                                                                                                                 |
|--------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| contact_us         | Jose C. Lacal - Jose.Lacal@NIHPO.com                                                                                                                                                                                                               |
| copyright          | Compilation and value-added processing (c) 2007-2020 NIHPO, Inc.                                                                                                                                                                                   |
| credits            | US Food and Drug Administration                                                                                                                                                                                                                    |
| description        | Drugs@FDA contains information about the following FDA-approved products for human use: * Prescription brand-name drug products, generic drug products, and many therapeutic biological products; * Over-the-counter brand-name and generic drugs. |
| disclaimer         | NIHPO only loads data sourced from government and international agency sources. Thus, NIHPO is not liable for neither the accuracy nor correctness of this data.                                                                                   |
| documentation      | Documentation                                                                                                                                                                                                                                      |
| source             | <a href="https://www.fda.gov/drugs/drug-approvals-and-databases/about-drugsfda">https://www.fda.gov/drugs/drug-approvals-and-databases/about-drugsfda</a>                                                                                          |
| support            | Jose.Lacal@PHUSE.eu; Priority support and personalized training are available for a fee.                                                                                                                                                           |
| usage_restrictions | For non-commercial use only.                                                                                                                                                                                                                       |

Showing 1 to 9 of 9 entries Previous 1 Next



# R podr Shiny App: Table Page

- The Table page provides you access to dataset library and its data sets.
- You can define how many rows in each page

Login ReadMe Table Show

Library Name: virtual\_css\_2020\_adam

Dataset Name: adadas

ADADAS from VIRTUAL\_CSS\_2020\_ADAM

Show 10 entries Search:

|    | ordinal_position | column_name | is_nullable | data_type        |
|----|------------------|-------------|-------------|------------------|
| 1  | 1                | STUDYID     | YES         | text             |
| 2  | 2                | SITEID      | YES         | bigint           |
| 3  | 3                | SITEGR1     | YES         | bigint           |
| 4  | 4                | USUBJID     | YES         | text             |
| 5  | 5                | TRTSDT      | YES         | double precision |
| 6  | 6                | TRTEDT      | YES         | double precision |
| 7  | 7                | TRTP        | YES         | text             |
| 8  | 8                | TRTPN       | YES         | double precision |
| 9  | 9                | AGE         | YES         | double precision |
| 10 | 10               | AGEGR1      | YES         | text             |

Showing 1 to 10 of 40 entries

Previous 1 2 3 4 Next



# R podr Shiny App: Show Page

- You can select columns/variables to be displayed
- You can filter number of records to be displayed by value
- You can set number of records per page
- You can transpose the columns into rows

The screenshot displays the 'Show Page' of the R podr Shiny App. The interface includes tabs for 'Login', 'ReadMe', 'Table', and 'Show'. The 'Show' tab is selected. The form contains the following elements:

- Library Name:** A text input field containing 'virtual\_css\_2020\_sdtm'.
- Dataset Name:** A dropdown menu with 'ae' selected.
- Select Columns:** An empty text input field with a dropdown menu open, showing a list of column names: AEACN, AEBDSYCD, AEBODSYS, AEDECOD, AEDTC, AEDY, AEENDTC, and AEENDV.
- Column Name:** A dropdown menu with 'AEACN' selected.
- = Value:** A dropdown menu with '\_\_\_ALL\_\_\_' selected.
- Record Limit:** A dropdown menu with '1' selected.
- Transpose:** A checkbox labeled 'Tranpose' (note the typo in the image).



# R podr Shiny App: Demonstration

---

- [https://geotiger.shinyapps.io/01\\_podr/](https://geotiger.shinyapps.io/01_podr/)





# QUESTION AND ANSWER



## Hanming Tu, VP of Clinical IT & Database Administration

Phone: 484-202-6479; C: 484-881-2384

Email: [htu@frontagelab.com](mailto:htu@frontagelab.com)

Address: 700 Pennsylvania Drive  
Exton, PA 19341, USA

Web: [www.frontagelab.com](http://www.frontagelab.com)

Fax: 610-232-0101