cessda TRAINING

metadata technology NORTH AMERICA

RICH DATA SERVICES

May 2021

https://www.richdataservices.com

AGENDA

- Welcome (5')
- RDS Overview (15')
- RDS Explorer / TabEngine demo (20')
- RDS Q&A (10')
- RDS Nesstar / DataVerse / DDI integration (15')
- Using RDS with R (10')
- Q&A (15')

This work is licensed under a
Creative Commons Attribution 4.0 International License

About Us

OUR MISSION

Modernization of infrastructure, tools, and practices surrounding official statistics and research data management, by leveraging information technology and standards.

OUR WORK

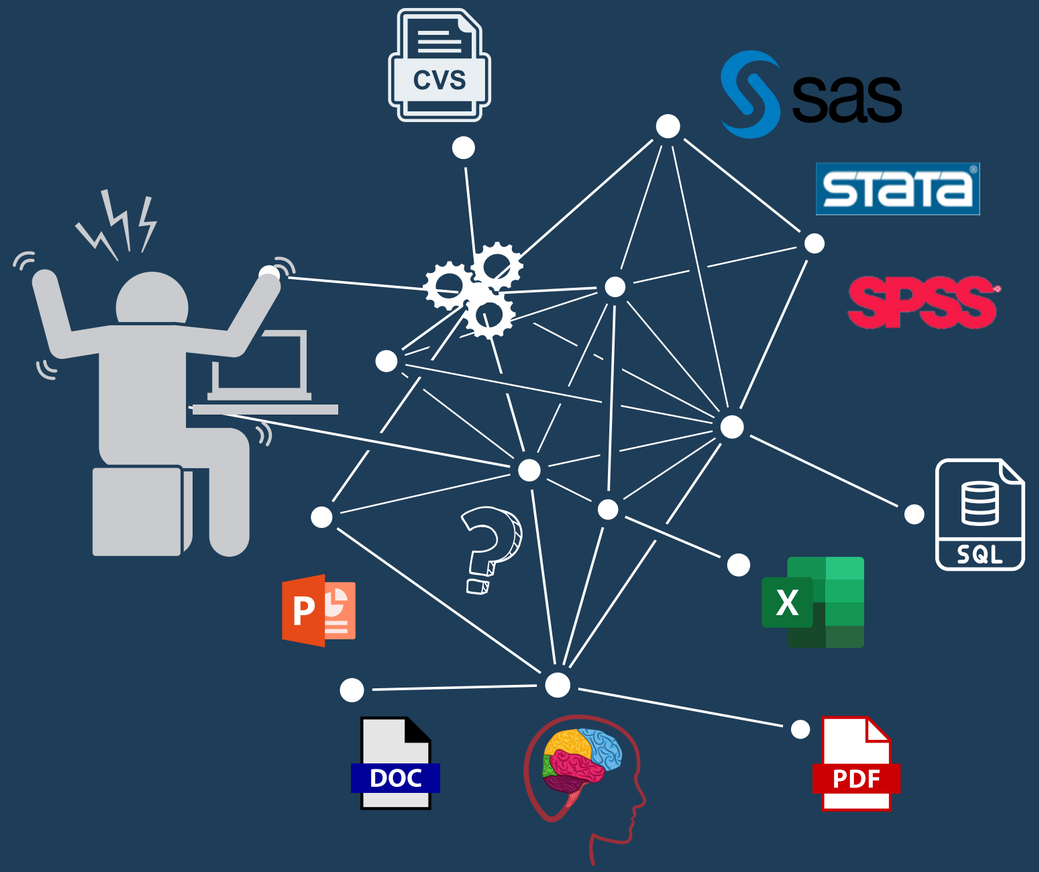
Collaborating since late 90's with national statistical offices, government and international organizations, data archives, research centers, academic institutions, and standards initiatives,

OUR SKILLS

Team with unique expertise in both information technology and data science

Without RDS

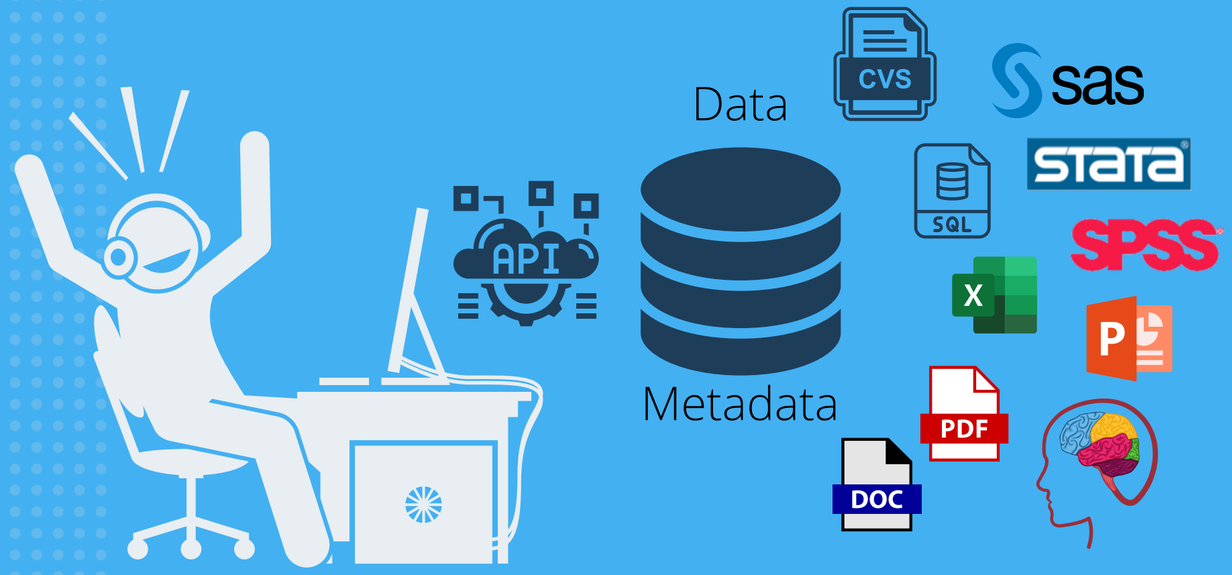
Users spend countless hours discovering, gathering, converting, data and documentation, with limited access to resources or support



STOP **Data Wrangling**

With Rich Data SERVICES

Users get immediate and concurrent access to data and metadata over industry standard web services or for offline use



Modernize your data management and publication infrastructure

https://www.richdataservices.com

GO FAIR DDI GSIM metadata technology NORTH AMERICA

WHAT IS RDS SOLVING?

In traditional environments, researchers and developers struggle to gain access to data and documentation

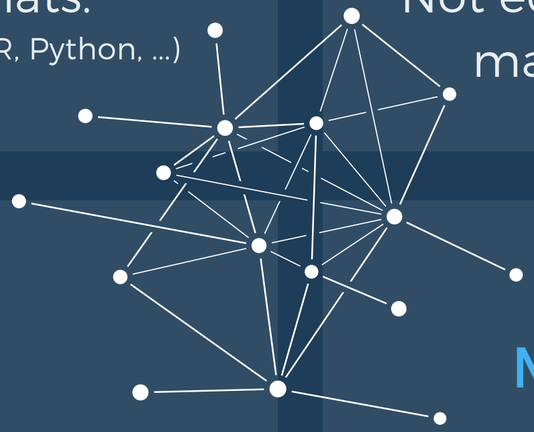
DATA WRANGLING

STATISTICIANS & DATA SCIENTISTS

Use tools / applications designed to work with specific formats.
(Excel, SAS, Stata, SPSS, R, Python, ...)

WEB / APPS DEVELOPPERS

Constrained by web and development platforms. Not equipped to read the many data formats.



DATA

Lives in a variety of text formats, proprietary files, and databases. Often needs to be converted, subsetted, or transformed.

METADATA

Knowledged stored in unstructured documents, emails, brain cells. Often hard to find, and not machine friendly

WITH RICH DATA SERVICES

RDS connects users and applications to data and metadata through an industry standard REST based API

STOP WRANGLING

STATISTICIANS & DATA SCIENTISTS

Can leverage API and applications to interact with data and metadata, or download for offline use

WEB & APPS DEVELOPPERS

Have access to web services to integrate data and metadata in portals, dashboards, and apps.

DATA

Stored in database, ready for on demand access, querying, sub-setting, or packaging in open formats for delivery

METADATA

Knowledge around the data converted to machine friendly formats in a standards driven metadata repository

Why APIs?

IT Standard: Modernization and data as a service

MODERNIZATION

Upgrade data publication and management infrastruture to 21st century

AUTOMATION

Enable discovery, access, and use of data by computer systems and applications

CATER TO NEXTGEN

Meet the needs of next generation of data scientists and developers

SELF-SERVICE

Let users answer their needs, reduce burden for publishers, facilitate exchange.

ALIGN ON STANDARDS

Foster adoption of *IT industry* standards and best practices

Why Metadata?

IT and Domain Standards: Data APIs are not enough

Metadata is simply knowledge converted to machine friendly formats

COMMON SENSE

No data should *ever* be delivered without documentation

MACHINE ACTIONABLE DATA

Empowers information system with necessary knowledge to process the data

AUTOMATION

Enables software driven management, analysys, and knowldge inference

DATA & RESEARCH QUALITY

Improves overall data quality and ensure preservation of knowledge

ALIGN ON STANDARDS

Foster the use and adoption of *domain standards* and related bestt practices

Why RDS?

API for concurrent access to data and metadata

NOT JUST A DATA API

Addresses the needs for access to both data and metadata in efficient ways

USER ECOSYSTEM

Caters to a wide variety of audiences on both the publisher and consumer sides

STOP WRANGLING

Designed to reduce data and code wrangling

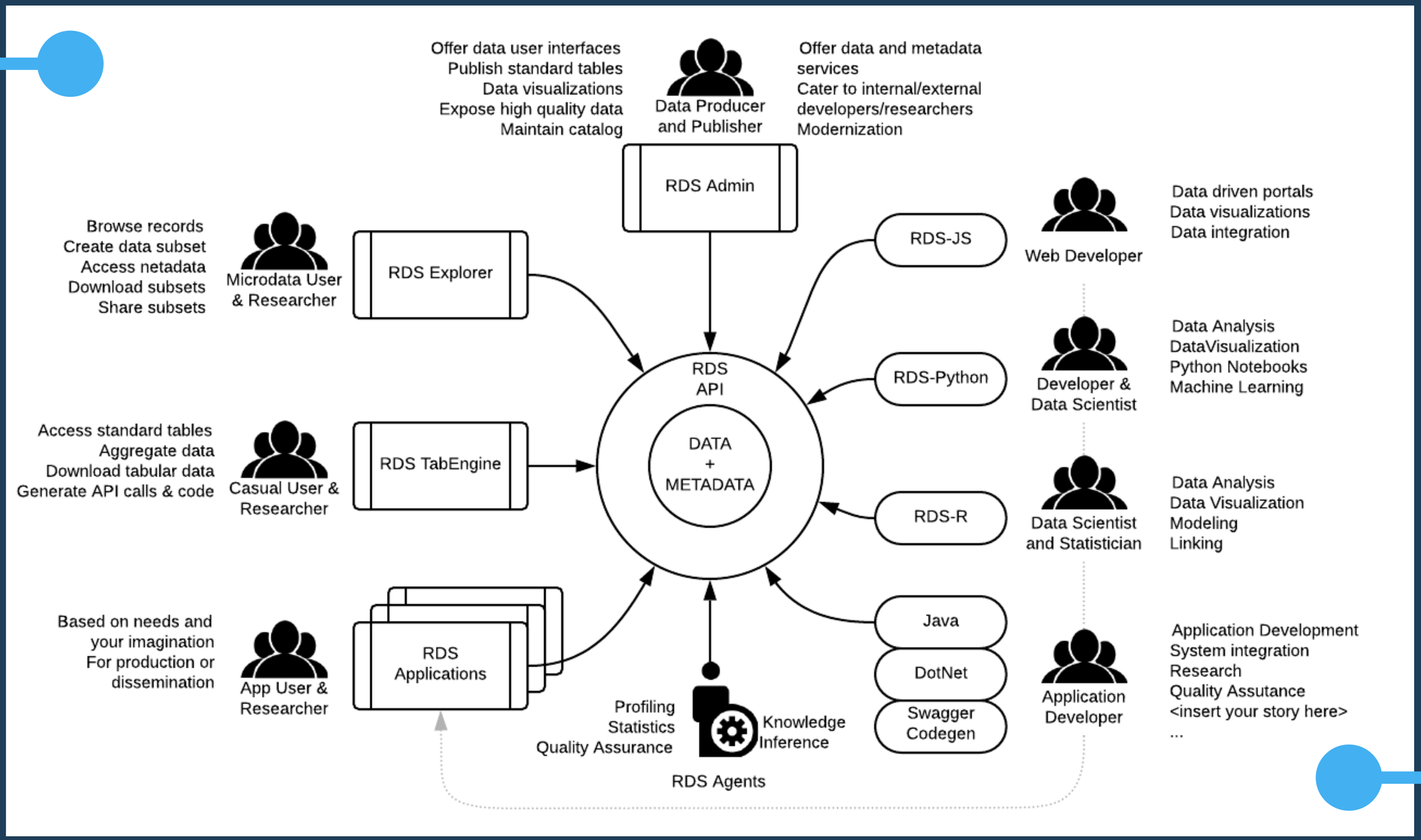
PLATFORM SUPPORT

Industry standard documentation, data driven Postman workspaces, tech support.

MTNA DATA & IT EXPERTISE

Backed by a team that understands both data and technology

RDS USERS ECOSYSTEM



RDS WEB SERVICES

CATALOGS AND DATA PRODUCTS

Access to information about the datasets, variables, classifications, and related properties

DATA QUERYING SERVICES

Interactive access to record level or tabulate/aggregated data, with injected metadata

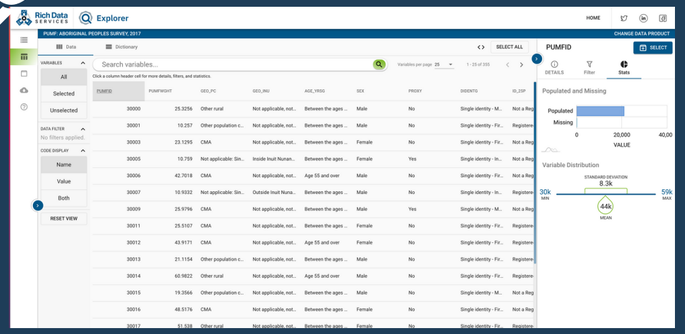
OPEN DATA PACKAGING

Download data / metadata in open formats, with scripts for immediate reuse in apps or SQL

ADMIN AND UTILITIES

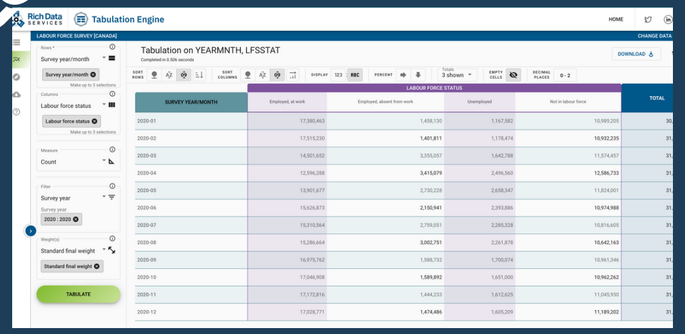
Content management, usage statistics, agents, and other administrative functions

RDS USER INTERFACES



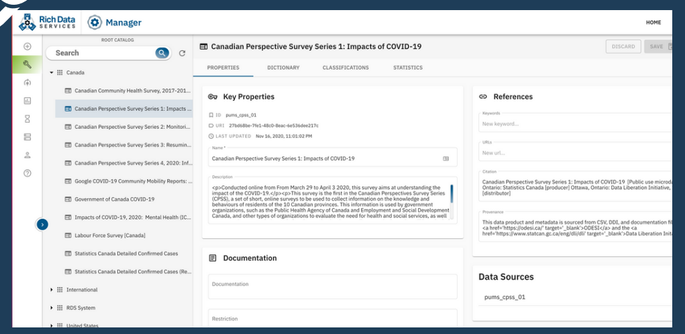
RDS EXPLORER

Browse and discover record level data, create subsets using filter and variables selection, download open data package



RDS TABENGINE

Quickly create multi-dimensional aggregated data rollups to produce tables and other findings around yor data.



RDS MANAGER

Maintain RDS catalogs and data products, monitor platform usage and statistics, and other management features

BUILD YOUR SOLUTIONS ON THE API

(OR LET US KNOW WHAT YOU NEED)

1

Apps & Visualizations

Create data driven applications and UIs, charts, graphs, or maps using popular frameworks, or develop story telling dashboards.

2

Integrations

Integrate data and metadata in web portals mobile applications, or social networks.

3

Direct Access

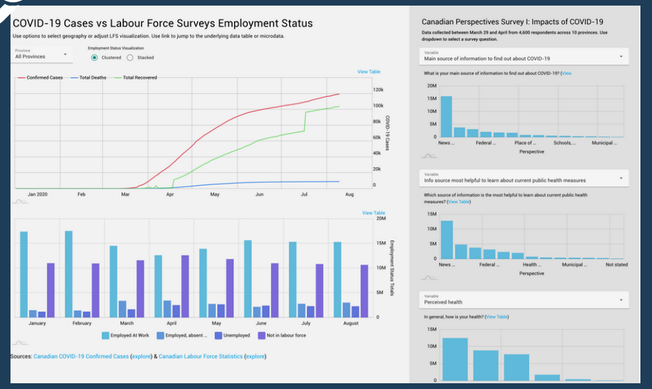
Interact with the data and metadata from Python, R, Java, C#, or any REST API aware environment, for analysis, visualization, machine learning, or any other purposes

4

<Insert here>

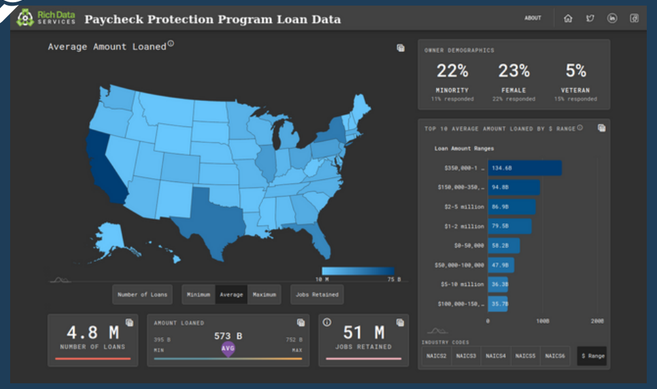
Creativity and imagination are your friends. (Meta)Data as a service opens the door to unlimited capabilities.

RDS SHOWCASE



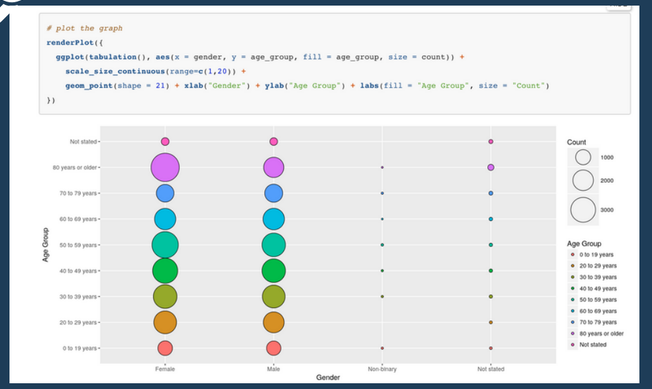
CANADA COVID-19

A dashboard bringing together data from multiple sources to explore COVID-19 and the Labor Force in Canada



PPP DASHBOARD

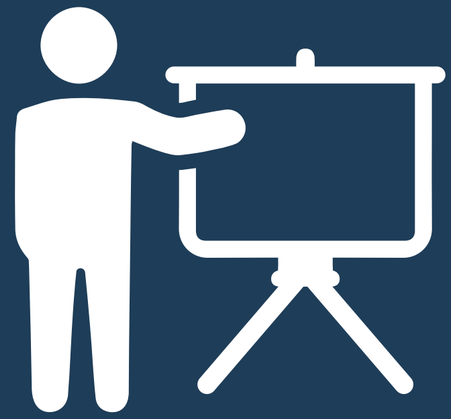
An interactive dashboard to explore the U.S. Paycheck Protection Program Loan data from the Small Business Administration



R-SHINY APP

Using R-Shiny to create visualizations around COVID-19 time series and record level data from Johns Hopking University

DEMO / STORIES



RDS EXPLORER

How a researcher can use the RDS Explorer to browse, subset, and download microdata for offline analysis using its favorite statistical package or database.

RDS TABENGINE

How a public or traditional, non-expert user can use the RDS Tabulation Engine to create aggregated data tables for analysis of visualization.

USING RDS FROM R

How a R user can leverage the RDS API, open source libraries, and code snippet generator to, in a matter of minutes, read and reuse data in R Studio, removing the burden of data and code wrangling.

Frequently Asked

- ***Availability***: Officially launched and open to early adopter
- ***Hosting***: Cloud / On-Premises / Docker
- ***Usage***: Public or Private, Publication or Production
- ***Scalability***: back-end SQL, vertical / horizontal (r/o), tasks queues
- ***Licensing***: per server, annual/perpetual. Flexible pricing. Free trial program/pilot for approved use cases.
- ***Services***: on-boarding, tech support, hosting, integration, data curation, metadata, training, institutional strategy, custom apps, ...
- ***Metadata***: DDI, JSON, Metasheet, RDS Agents, migration tools, Nesstar Publisher, manual capture
- ***RDS Suite***: Aria, Data ingest / QA / Harmonization / ETL, Publication
- ***Integration***: with other systems through SOA and standards

MIGRATION / INTEGRATION

Nesstar / DataVerse

Automated or manual sync

Data

Metadata

For more information

https://www.mtna.us
mtna@mtna.us

RDS HOME & SHOWCASE

https://www.richdataservices.com
https://www.richdataservices.com/showcase

PUBLIC DATA CENTERS

https://covid19.richdataservices.com
https://public.richdataservices.com

DOCUMENTATION

OpenAPI / Swagger
Postman Public Workspaces

GITHUB

Open sources libraries and examples
https://github.com/mtna

SUPPORT

Knowledge Base
Help Desk