

Standards Compliance - An automated way to measure the Interoperability of Clinical studies

MINT – Metadata INTeroperability

Paper SI11, PHUSE US Connect

Orlando Mar 11-14

Rammprasad Ganapathy,

Global Data Standards Manager

Data Standards & Governance

Genentech, A member of Roche Group.



Agenda



1. Introduction – What is 'I' in F.A.I.R
2. Why we need this app?
3. Roche Global Data Standards
4. Roche MDR
5. Metadata INTeroperability App Architecture
6. Technology
7. User Experience
8. Take away message

Introduction – What is ‘I’ in F.A.I.R

Data Standards, adherence to standards and Terminologies is key to the 'Interoperability' aspect of the FAIR principles.



Findability

Where are our data?

All data assets must be cataloged and annotated with rich descriptions that allow for intuitive search by anyone in the company to discover what is available and where data are located.



Accessibility

How can I get the data?

Data access should be seamless and automated when possible (e.g., by machine using code) with appropriate access restrictions by user, release timing, or legal compliance requirements.



Interoperability

How can I connect or integrate the data?

Use of standards and consistent terminologies enables data interoperability for faster and easier integration into analyzable datasets for generating insights.



Reusability

Can our data be easily shared and used again?

Making data F, A, and I, makes it easier to share and use for answering new scientific questions, for [reverse translation](#), and for achieving the scale required to support personalized healthcare.

What is 'I' in F.A.I.R

- The use of data standards & terminologies in the clinical studies **contribute to the I of FAIR** enabling interoperability.
- **Interoperability principles allows a consistent approach** to the design of clinical studies and also in combining of data from multiple studies.
- Interoperable clinical **studies supports the creation of standard processes, programs, macros and systems** to streamline workflows in data collection, tabulation, and analysis with the aim of getting medicines to patients faster.

Why we need this app?

Challenges from the study teams



Study Teams

Data Collection/Study Data Manager

- How Interoperable is my study?
- Have I used standards in all applicable cases?
- Can I quickly identify the non-standards?
- Are there any discrepancies in adopting standards?

Data Tabulation/Analysis

- How does the non-standards in the studies impact Data Tabulation & Analysis.
- Can I quickly identify the non-standards?
- Can the down stream Data Tabulation activities be automated?



Data Standards Team

- How can we facilitate implementation of sponsor's standards without discrepancies?
- Can I quickly identify the non-standards for future standards development?
- How can we maximize the use of standards in the clinical studies?

There is no easy way to do it

Underlying Problem - Interoperability

One of the key underlying problem in the Interoperability of Clinical studies is the implementation of standards.

- **Data Collection standards are not consistent** which leads to inconsistent study design and mapping to SDTM.
- **Inconsistent Controlled terminologies** across studies makes it challenging to pool.
- Inconsistent Data collection and Tabulation leads to **less Interoperability** of the studies that makes it challenging to pool data across studies, molecules & Indication. Thus limiting the further reuse of the clinical data.



Metadata INTeroperability App helps the study teams to overcome these challenges and maximize the use of sponsor's data standards in clinical studies.

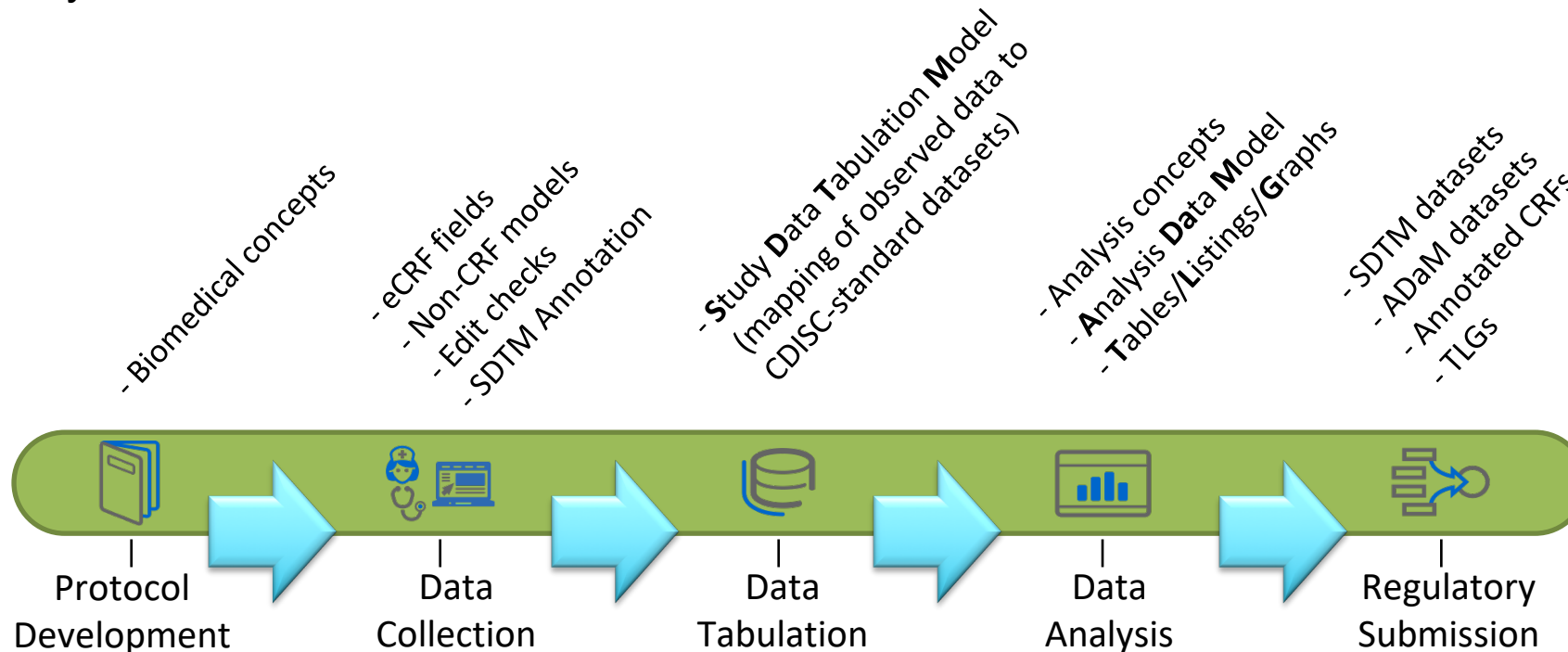
Driven by the Metadata, sematic design and latest technologies, the app measures the standards compliance in an automated way .

Roche Global Data Standards

Roche Global Data Standards

What are they?

- Define how Roche clinical trial data should be collected, tabulated, analyzed, and submitted to regulatory authorities.
- Include all the 'layers' required, from biomedical concepts, metadata, and controlled terminology.
- Fully aligned with CDISC submission data standards (SDTM & ADaM), extended to ensure consistency across Roche clinical trials.



A F.A.I.R MDR

FAIR Data Standards in the GDSR (Roche MDR)

- The Data standards available in the GDSR contribute to the I of the FAIR principles, enabling Interoperability.
- The GDSR is Roche and Genentech's MDR and is **the single source of truth for clinical data standards**, storing and making them available to users globally in the most business useful format.
- The data standards are **stored in semantic graphs** and the technology behind the GDSR is based on RDF, ontologies, and linked data principles aligned with Semantic Web (W3C) standards.

REST API

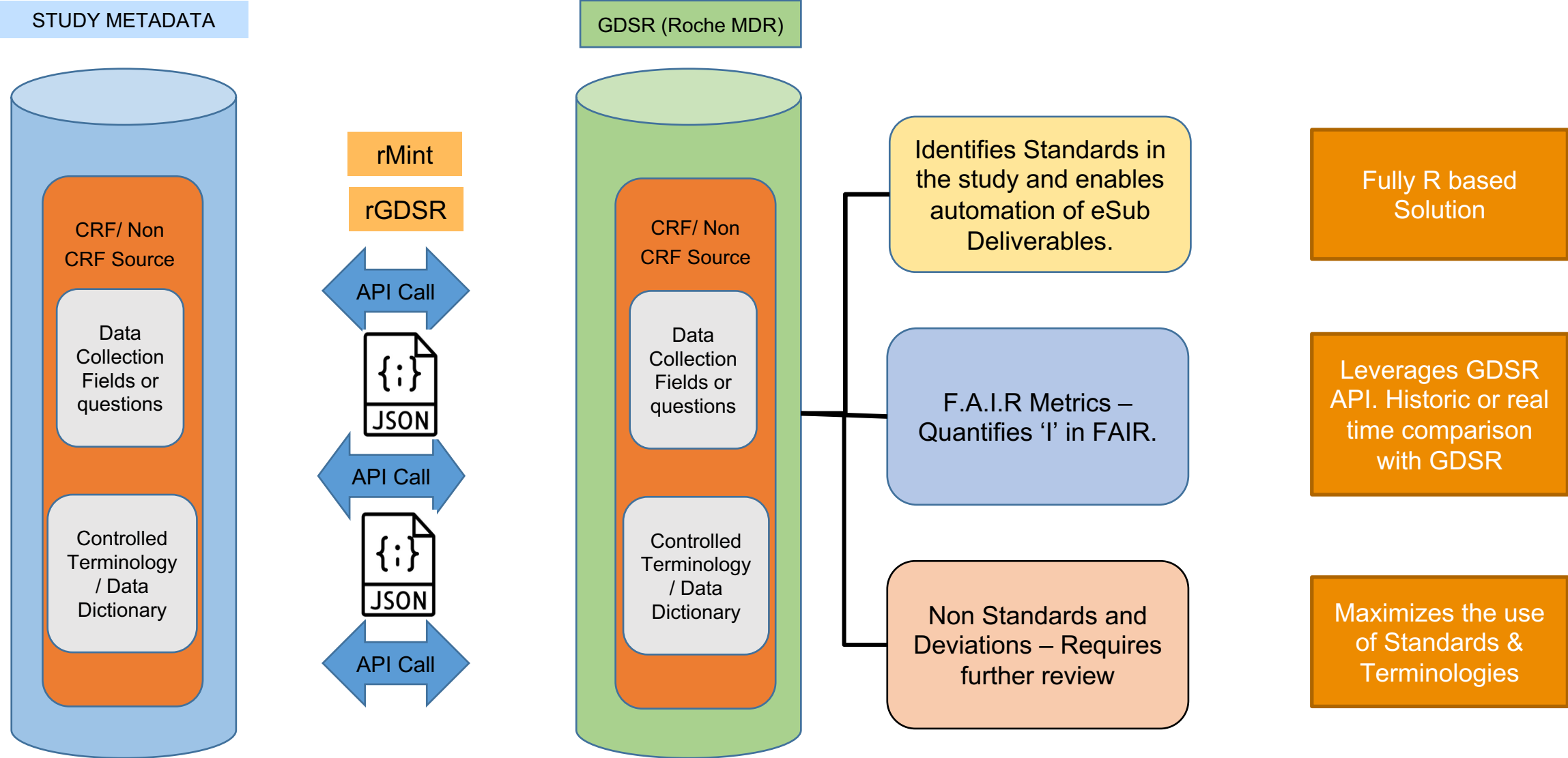
The GDSR makes content available through **a public REST API**. Users can programmatically access the data standards, catalog, publication information, and the archives, making automation of previously manual tasks possible.

PUBLICATION VERSIONING

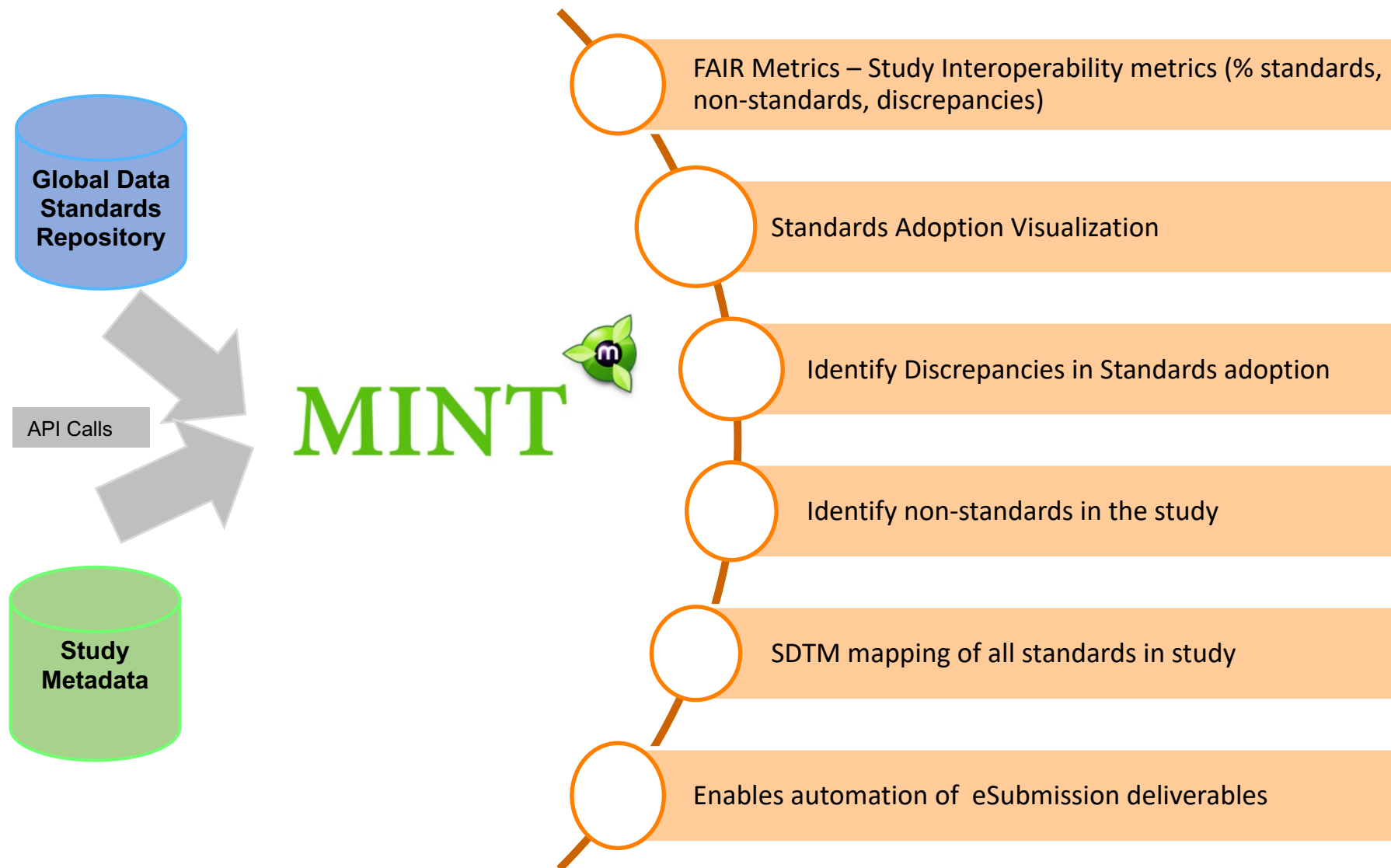
The GDSR uses **publication versioning** which allows users to access the archives of previous data standards versions. New, updated, and retired data standards are made available to users in a publication cycle and the previous publication is moved to an accessible archive.

Metadata INTeroperability – MINT App Architecture

MINT - Metadata INTERoperability App Architecture



MINT - Metadata INTeroperability App



Metadata INTeroperability Technology - R PACKAGE AND SHINY MODULE

- To enable faster processing, the core logic of the app resides in the R-shiny packages and the user interface is built in R shiny using Shiny dashboard.
- rGDSR and rMINT are the two home developed developed R packages for MINT.
- rGDSR parses the JSON files that is in nested tree like format to two-dimensional data frames for further processing by rMINT
- rMINT performs rest of the operations such as comparing the study to the standards and presenting them in a user-friendly way.

R Shiny Module – Shiny Dashboard

Home

Provide Input

F.A.I.R Dashboard

Non Standards

Discrepancies

Data Tabulation

- >> SDTMv Annotation
- >> Controlled Terminology

Download

Metadata INTERoperability App



MINT

Introduction
Use of standards and consistent terminologies enables data interoperability for faster and easier integration into analyzable datasets for the primary use of data and beyond. The **M**etadata **I**NTeroperability App (MINT App) is an R based solution that compares the study metadata to the GDSR metadata via API and provides useful information regarding study Data collection (Standards adherence, Discrepancies, etc) and Data Tabulation (SDTMv) standards to the study teams.

[GitHub Link to MINT app](#)

References

[Introduction & Road map](#)

[How do I use this app? - Video Tutorial](#)

[Test Data Download](#)

[App User Guide](#)

Feedback

For any question or feedback, please contact:

- Rammprasad Ganapathy

MINT Roadmap

FAIR Metrics – Study Interoperability
(% standards, non-standards, discrepancies)

Identify Discrepancies in Standards adoption

Identify non-standards in the study

SDTMv mapping of all CRF standards in study



MINT

Metadata
INTERoperability
App

Support Non-CRF Data (EntimICE & API access to Edison)

Generate eSUB ready SDTMv spec & 100% Controlled Terminology facilitating creation of DEFINE.XML.

Build Database & Store the app output to support a study from DB build until DB lock.

Support building Study Metadata Repository (SMR)

Compare Multiple studies with standards

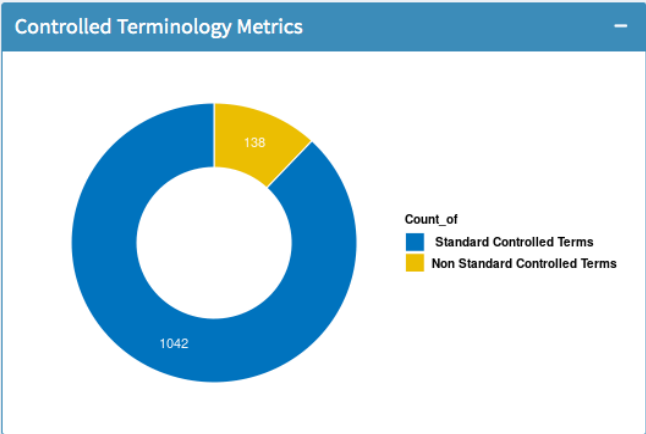
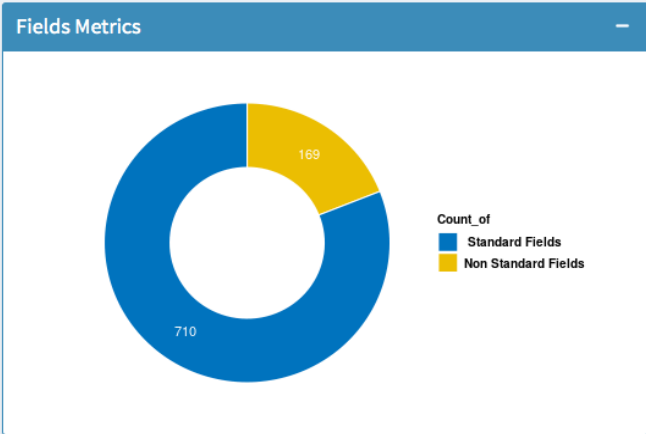
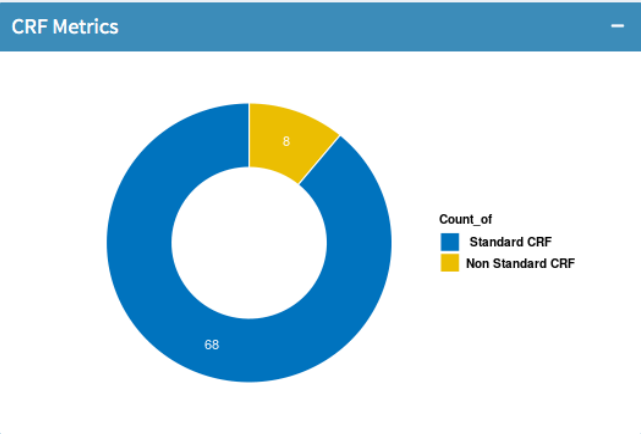
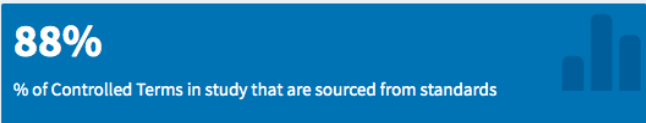
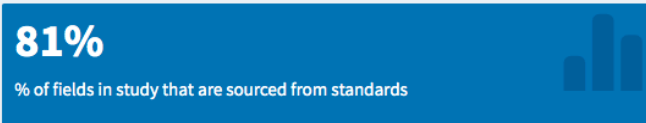
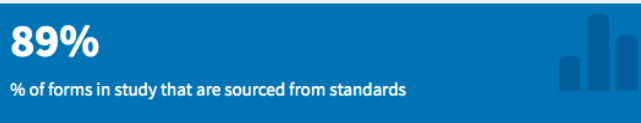
Standards Adoption Visualization



Study Interoperability Metrics

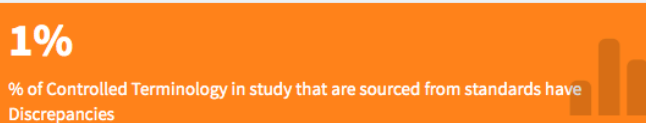
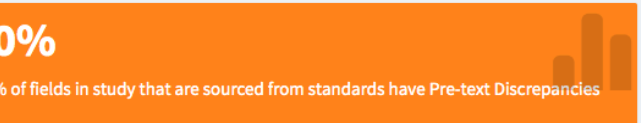
Metrics - Standards Utilization:

Metrics about the standards that are sourced at the study level. The metrics are presented for CRF pages, Fields and Controlled Terminologies.



Metrics - Discrepancy in Standards Adoption:

Metrics about the discrepancies in the standards that are sourced at the study level. The metrics are presented for Fields and Controlled Terminologies. Ideally there shouldn't be any discrepancies at study level.



How the study teams use it?



The study teams use the app to

- Review the standards compliance and maximize standards adoption.
- Review the non-standards and make sure it adheres to the Roche design principles and also to make sure that the non-standards are mapped to SDTM following CDISC principles.
- Review and resolve the discrepancies in standards adoption before the study goes LIVE.
- Get the SDTM annotations of the standards in the study and use it to create SDTM aCRF and mapping specifications. In future, the app will enable further automation of e-Submission deliverables.

The versioning in GDSR lets user to choose the version of standards used for study design to produce accurate results. Essentially the users upload the study metadata to the app and choose the version of standards used for study design and review the results.

User Experience

User Experience and What it means to the study teams.

The study teams found it extremely useful.

The app has been used more than 1500 times since it was launched Q3 of 2019.

Average Standards Compliance from a few Early Adopter studies – 90%



What it means to the study teams?

- Standards adoption and building Interoperable clinical studies enables seamless automation from Data Collection to Data Tabulation to Data Analysis tasks.
- Most of the programming tasks in Data collection (CRFs, Edit checks & Custom Functions) & Data Tabulation (SDTM spec, SAS template programs & macros) can be automated for standards in the study.
- e-Sub deliverables can be automated for standards in the study.

Take away message

Take away message

- Following FAIR principles and a Metadata driven approach has helped us create a solution that **enables the creation of Interoperable clinical studies**.
- Having the right infrastructure, well defined data standards, a FAIR MDR and efficient use of technology is key to the success in creation of a tool that **automates the measure of Interoperability of Clinical studies and creation of high quality eSubmission documents**.
- This idea could be **extended to any EDC and MDR** that meets the FAIR principles.

References



- <https://www.nature.com/articles/sdata201618>
- <https://www.force11.org/group/fairgroup>
- <https://www.pistoiaalliance.org/>
- Roche FAIR Data playbook
- GDSR FAIR – FAIR Data Standards in the GDSR (Global Data Standards Repository) Huw Mason, F Hoffmann - La Roche, Basel, Switzerland
- <https://www.lexjansen.com/phuse/2019/pp/PP06.pdf>
- R packages by Hadley Wickham
- Shiny Dashboard <https://rstudio.github.io/shinydashboard/index.html>

Acknowledgements



The author wishes to thank our colleagues who contributed to various aspects of this initiative: The Data Standards & Governance team, the Early-Phase/Late-Phase Harmonization Team, the Global Data Standards Implementation Team, the Enhanced Data and Insights Sharing Reference Data Services, FAIR Playbook team and the Enhanced Data and Insights Sharing FAIR Metrics Team, Agile eSubmission team.

The author also wishes to thank Chris Price/ Mercy Navarro and Jonathan Chainey for reviewing the content of this presentation.

***Doing now what patients need
next***