

Collecting User Stories for FAIR Data Generation and Re-use

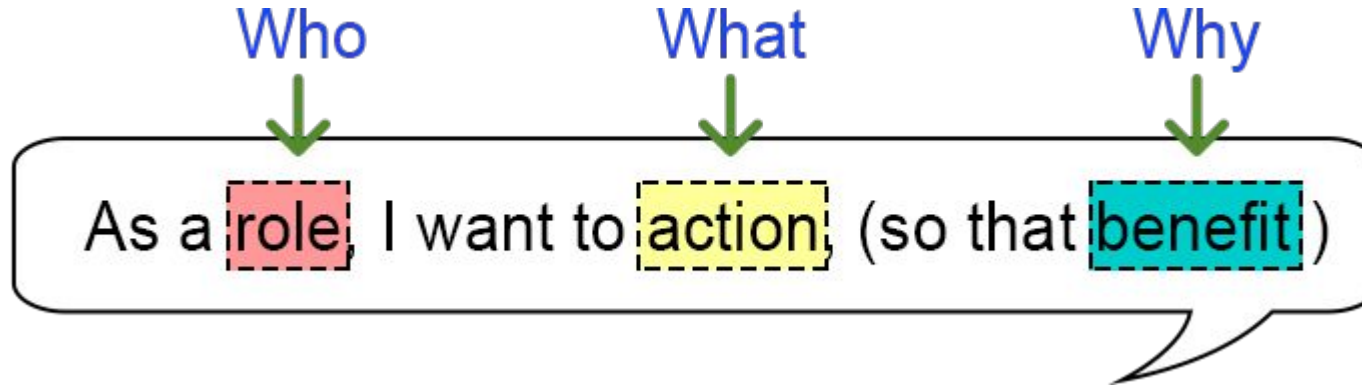
FAIRplus CMM Task Force

Oya Beyan, Nick Juty, Ebtisam Alharbi, Ibrahim Emam

15.04.2020

Scrum Agile Development: USER STORIES

- A **User Story** is a brief statement of intent that describes something the system needs to do for the user.



As a *<type of user>*, I want *<some goal>* so that *<some reason>*.

- The [primary actor, user role, or 'persona']
- A single [action or function that the actor wants to happen]
- That actor's main reason for doing it, or expected benefit

Who Writes User Stories ?

Anyone can write user stories ...



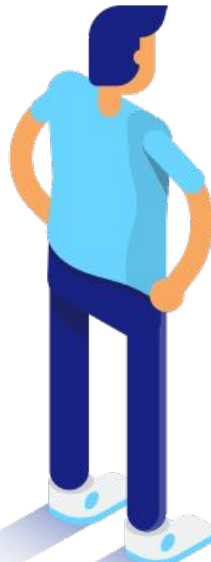
user stories != software system requirements

put actual end users at the center of the conversation

*why we are
building*



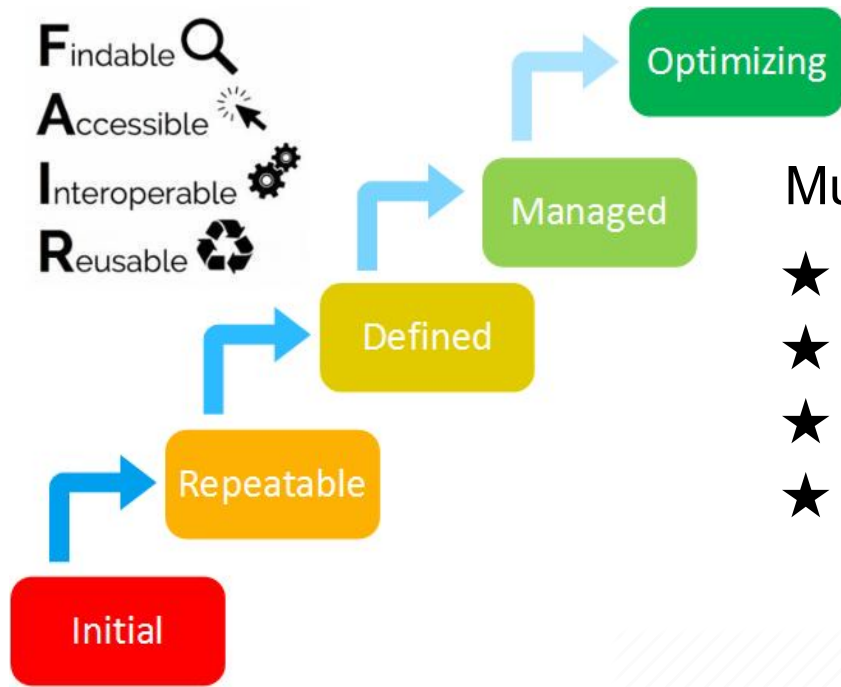
*what is the
value*



*use
non-technical
language*

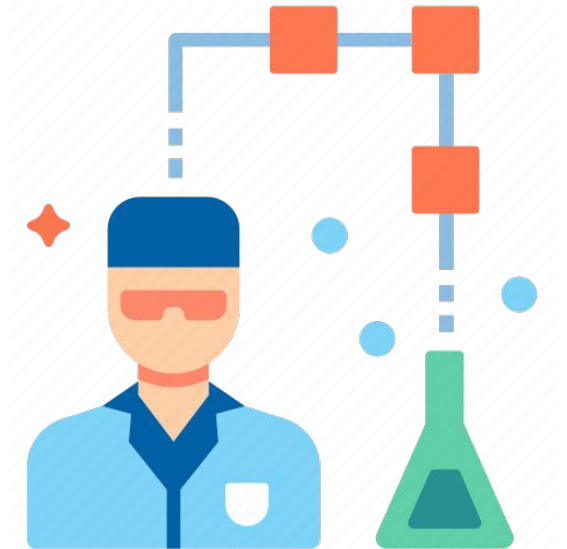
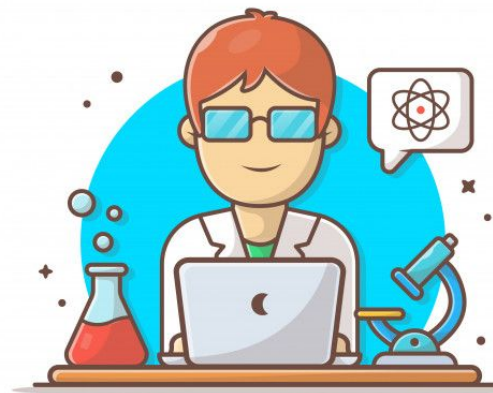
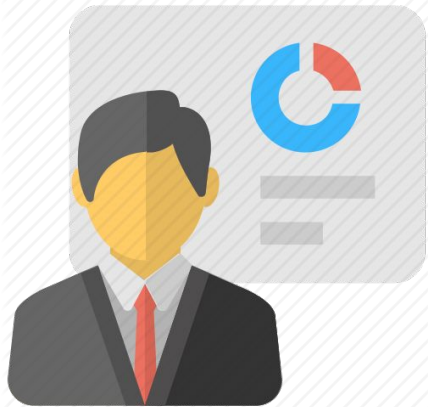
*provide context
for the
development
team*

Why do we need user stories in FAIR data maturation



Multiple parties with diverse motivations

- ★ Articulate the need
- ★ Identify the lack of certain capabilities
- ★ Focus on business value
- ★ Collect meaningful small steps of maturity improvement



WHO ?

AS A [actor/user role]

- Who wants this feature?
- who is performing the action ?
- who is receiving the value from the activity?
- The user could be
 - a generic actor (e.g. AS A user of the website),
 - a specific user role (e.g. AS A frequent business traveller),
 - another system (AS A BACS payment system).

Writing User Stories

<WHO> could describe roles in business units

Management

- CEO- Chief Executive Officer
- CIO-Chief Information Officer
- CDO-Chief Data Officer
- CISO/CSO-Chief Information Security Officer
- CAO -Chief Analytics Officer
- Director
- Head of Research Area
- Head of Department
- Head of Clinical Research and Development

Research and Development

- Principal Investigator (PI)
- Sub Investigator (Sub I)
- Regulatory Coordinator
- Subject Matter Expert
- Data Analyst (scientist)
- Scientist
- Senior Scientist
- Lab Head
- Team Lead
- Development Quality Assurance

Clinical Trials:

- Clinical Development Scientist
- Clinical Investigator
- Clinical Trials Project Manager
- Clinical Trial Statistician
- Clinical Research Nurse Coordinator (CRNC)
- Clinical Research Coordinator (CRC)
- Clinical Operations Manager
- Pharmacovigilance Manager
- Medical/Scientific Writer
- Clinical Data Manager
- Biostatistician

IT Department Lead

- IT Service Provider
- Data Custodian
- Data Steward
- IT Specialist

Data Governance

- Data Privacy Officer
- Data Protection Officer (Controller)
- Data Quality Officer

Writing User Stories

or could be the stakeholder of FAIR data production and reuse



Data Owner

- Data Protection
- Data Maintenance



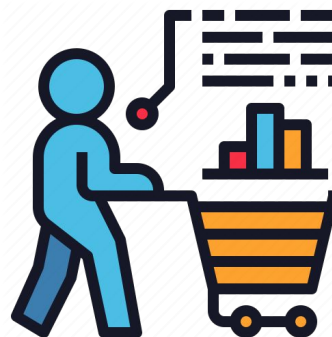
Decision Making/Planning

- Selecting Data Sources
- Calculating cost and benefit



Data Producer

- Data Annotation
- Data Storage
- Data Accessibility
- Legal Aspects
- Privacy and Security



Data Consumer

- Study Planning
- Scientific Discovery
- Business Intelligence
- Reporting / Management
- Data Search
- Data Analysis



IT Provider

- Central Vocabulary Services
- Master Metadata Management
- PID services
- Storage



Service Provider

- Catalogs
- Search Portal
- Data Governance

User Stories: Typical format

WHAT ?

I WANT [feature/action]

- What does the user want?
- Represents the action to be performed by the system.
- The user will typically want the system to perform a new behavior, e.g.
 - I WANT the ability to track an order
 - I WANT to pay for orders using an AMEX card
 - I WANT to cancel an order without any hassle

WHY

SO THAT [benefit]

- Why does a user want this functionality?
- This section provides the justification/benefit of the feature.

Writing User Stories

Who

What

Why

As a **role**, I want to **action** (so that **benefit**.)

As a researcher
working in the nuclear
receptor field

I want to access richly
annotated and systematically
archived datasets

so that I can reuse these data
sets for the routine generation
or validation of research
hypotheses



Writing User Stories

Who

What

Why

As a **role**, I want to **action** (so that **benefit**)

As a user of a public data repository

I want to identify and retrieve "like" datasets using free text queries in a certain accuracy

so that I can meaningfully interpret and compare Transcriptomic and ChIP-Seq datasets.



Writing User Stories

Who

What

Why

As a **role**, I want to **action** (so that **benefit**.)

A data manager

I want to choose relevant data & metadata standards for my project's generated data



So that my data can be semantically and structurally described in accordance to a community adopted standard, which would in turn enable systematic discoverability and interoperability for my project's data

Guidelines and Templates

Use the excel template:

https://docs.google.com/spreadsheets/d/1I_l3Zecd-O0fMv_8DTTKjznDuxbqEdtTVMhR9fyK7SM/edit#gid=1264957282

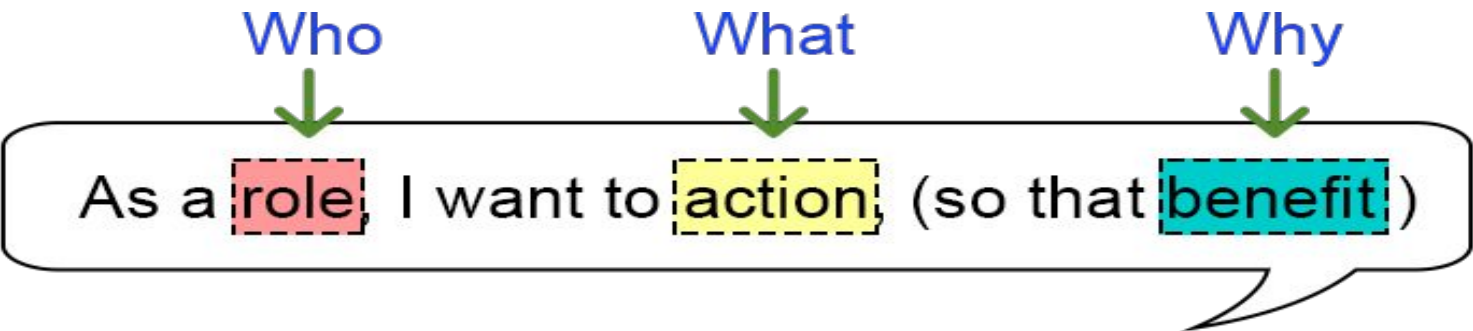
USE CASES					
			<WHO> Specify your role or the user's role (e.g. As a scientist, or as a data analyst, etc.)	<WHAT> Specify the action you or the user want to perform	<WHY> Specify why the actions needs to be performed, for what result, for what benefit
ID	Title	Owner	As a	I want	So that
6	choosing standards	Ibrahim Eman	A data manager	I want to choose relevant data & metadata standards for my project's generated data	So that my data can be semantically and structurally described in accordance to a community adopted standard, which would in turn enable systematics discoverability and interoperability for my project's data



optional:
provide a
title



Do NOT
forget to
write your
name



Guidelines and Templates

Annotate with Role and/or Stakeholder + Data Granularity + Related FAIR principles by using dropdown

	Role-Main	Role-Detail	Stakeholder	Stakeholder-Focus	Granularity	Related FAIR Principle
ally						
sted						
able						
ata	Information Technolog ▼	Data Steward ▼	FAIR Data Product ▼	Data Annotation ▼	Data Set Level ▼	I1 Metadata representation ▼

if you can not find desired label, feel free
to add to related tab

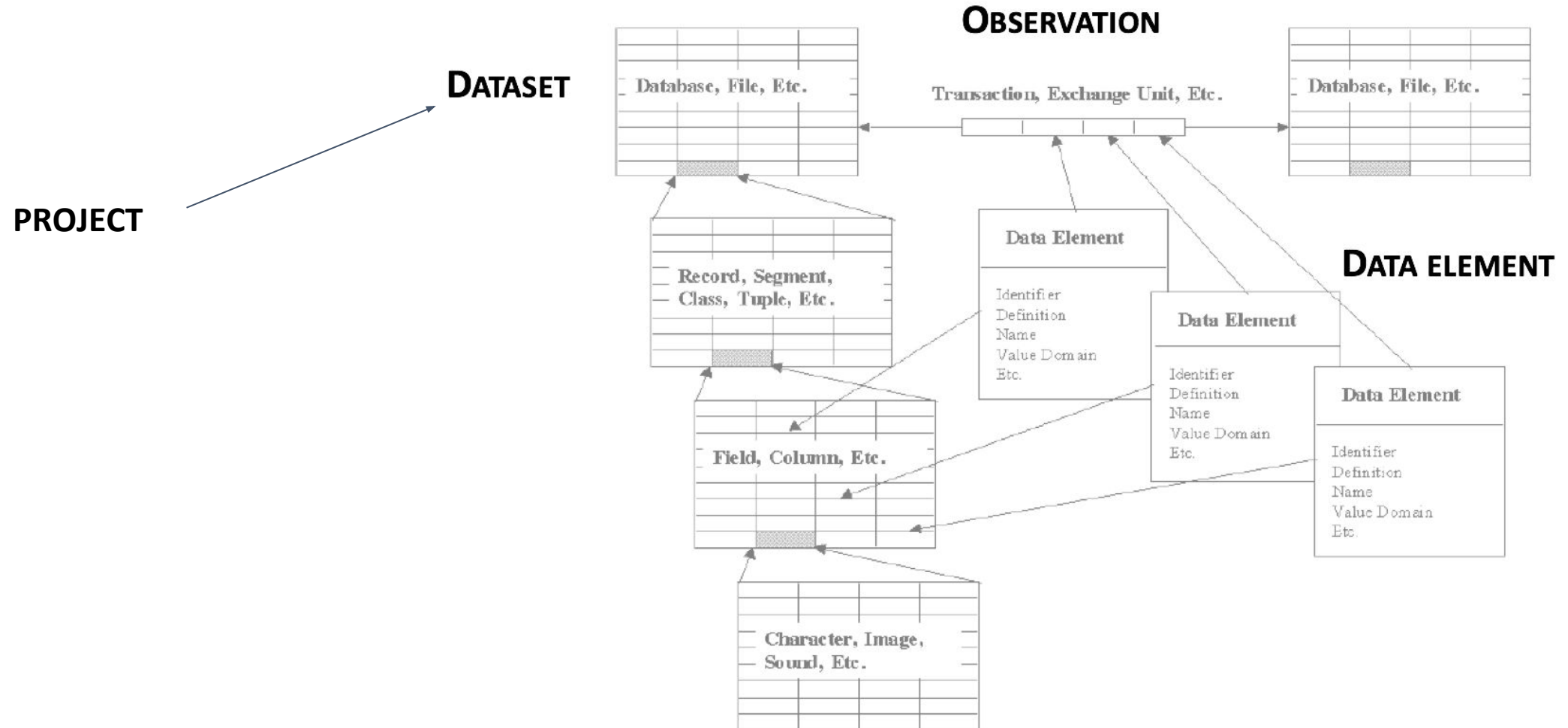
Annotate with Related FAIR Principles

Which FAIR principles is more related with your user story:

	Definition of individual FAIR principles				
Generic					
F Findability					
F1 PIDs	F1. (meta)data are assigned a globally unique and eternally persistent identifier.				
F2 Findability metadata	F2. data are described with rich metadata.				
F3 Searchable resource	F3. (meta)data are registered or indexed in a searchable resource.				
F4 Links between (meta)data	F4. metadata specify the data identifier.				
A Accessibility					
A1 Standard retrieving protocol	A1 (meta)data are retrievable by their identifier using a standardized communications protocol.				
A1.1 Open Free protocol	A1.1 the protocol is open, free, and universally implementable.				
A1.2 Authentication / Sensitive	A1.2 the protocol allows for an authentication and authorization procedure, where necessary.				
A2 Accessible metadata	A2 metadata are accessible, even when the data are no longer available.				
I Interoperability					
I1 Metadata representation	I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.				
I2 FAIR (meta)data vocabularies	I2. (meta)data use vocabularies that follow FAIR principles.				
I3 Qualified links between metadata	I3. (meta)data include qualified references to other (meta)data.				
R Reusability					
R1 Metadata	R1. meta(data) have a plurality of accurate and relevant attributes.				
R1.1 including license	R1.1. (meta)data are released with a clear and accessible data usage license.				
R1.2 including provenance	R1.2. (meta)data are associated with their provenance.				
R1.3 meeting community standards	R1.3. (meta)data meet domain-relevant community standards.				

Annotate with Data Granularity

- Data is hierarchical, hence granular



Annotate with Data Granularity

- A **project** / study generates a collection of **datasets**
 - Project metadata describe the project as a whole (*to give context to data within the datasets*)
 - e.g. A user story that involves managing / dealing with data about a project (i.e. project descriptive metadata) is at the granularity of a 'project level'
 - This is MORE than the FDP "data catalogue metadata" which is defined as the curated collection of metadata about *datasets*.
- A **Dataset** is a designed collection of variables / fields / **data elements** with observed/measured values
 - Dataset metadata includes descriptive, structural and contextual information about the dataset as a whole and its content
 - User stories that deal with the discoverability, accessibility, usability, interoperability, sharing of identifiable datasets require management (metadata) on the 'dataset' level
- A **data element** is the most granular form of 'observable data' that makes up the content of a dataset
 - e.g. 'age', 'drug dosage', 'log intensity'
 - Data harmonisation, integration, meta-analyses, cross-study data discovery are all user-stories that require data management on the 'data element' level

How can you contribute

- ★ if you are writing a recipe
- ★ if you are a stakeholder in FAIR data in your profession
 - reusing data,
 - generating data,
 - curating data,
 - developing solutions to support data management
- ★ if you are taking managerial and strategic decision by using data / about the data

Please provide us your user stories

https://docs.google.com/spreadsheets/d/1l_I3Zecd-O0fMv_8DTTKjznDuxbqEdtTVMhR9fyK7SM/edit#gid=1264957282

If you have any difficulty or suggestion let us know:

There is a Feedback / Suggestions Tab or sent an email

Please leave your comments and feedback related to the collection method of the user stories

User Stories vs Use Cases

- User Stories deliberately leave out a lot of important details.
- User Stories are meant to elicit conversations by asking questions during review.
- As the product owner presents the stories, the team will ask questions to further clarify the user story and the acceptance criteria.

User story vs Use case

