

Data Management Services: Collecting, Organizing, and Curating Datasets and Raw Data

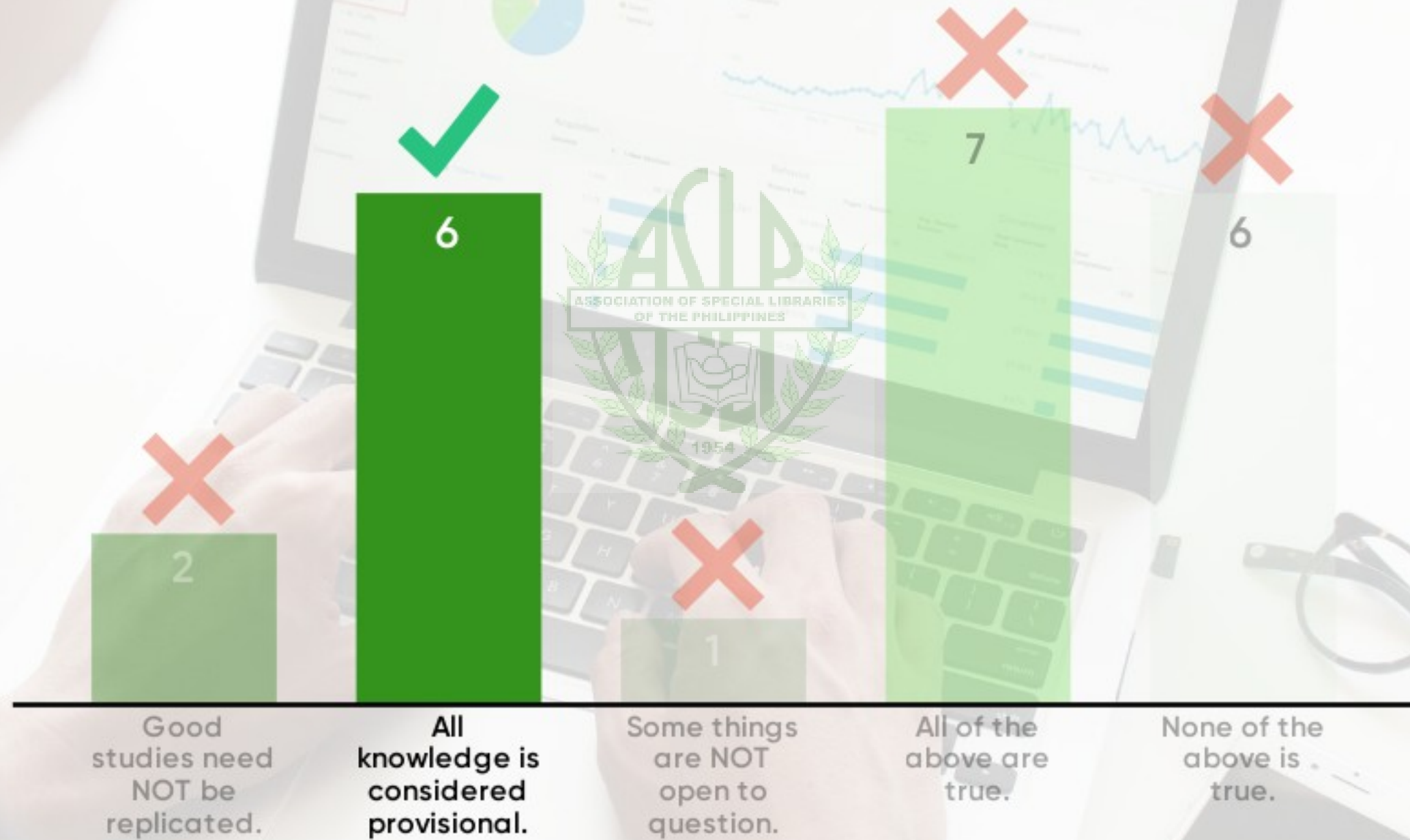
Asst. Prof. Dan Anthony Dorado

UP School of Library and Information Studies

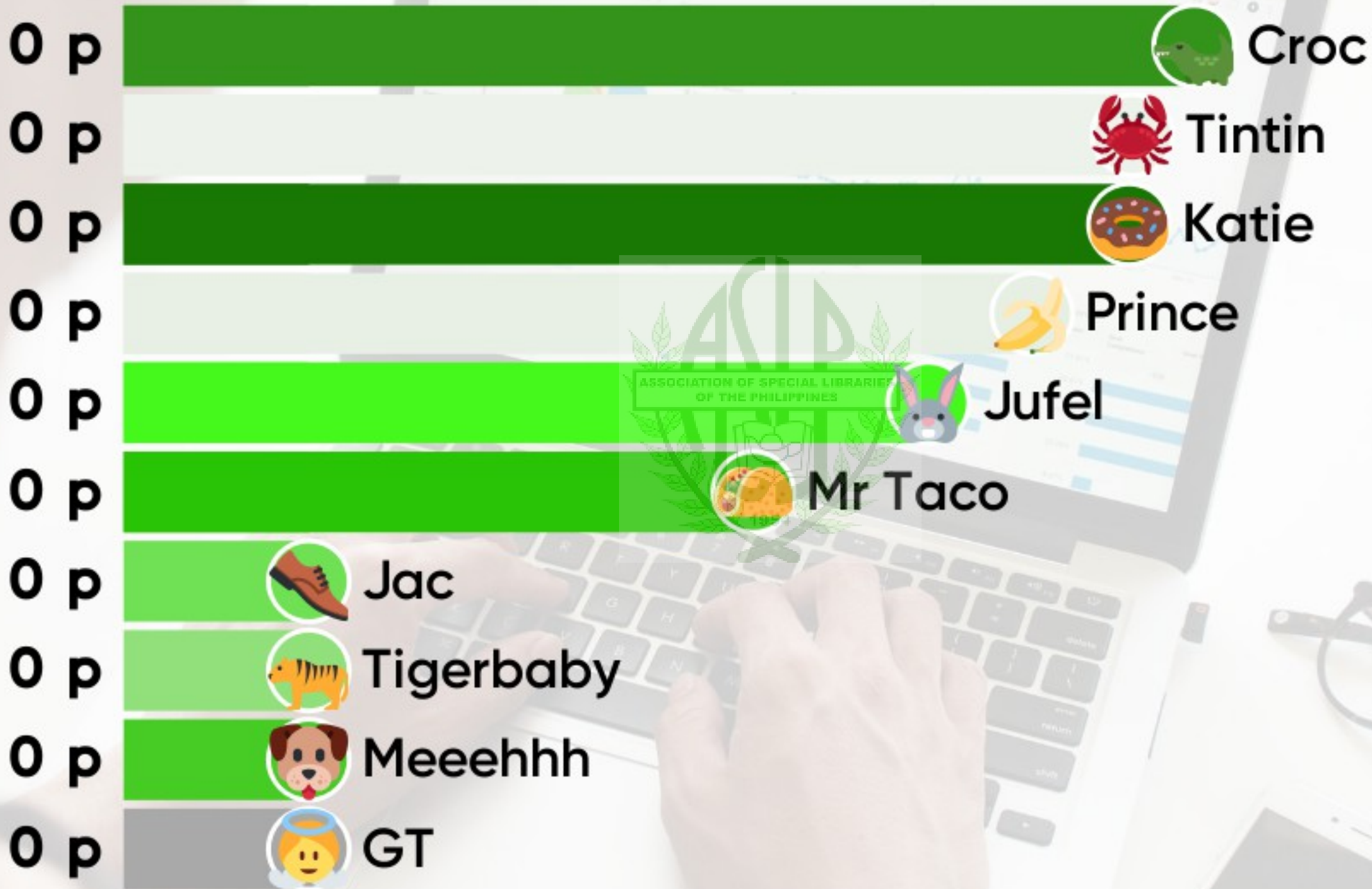
Bolstering Research: The Role of Librarian and Information Professionals
in the Ideation, Use, and Preservation of Outstanding Research

22-24 May 2016 | Bohol, Philippines

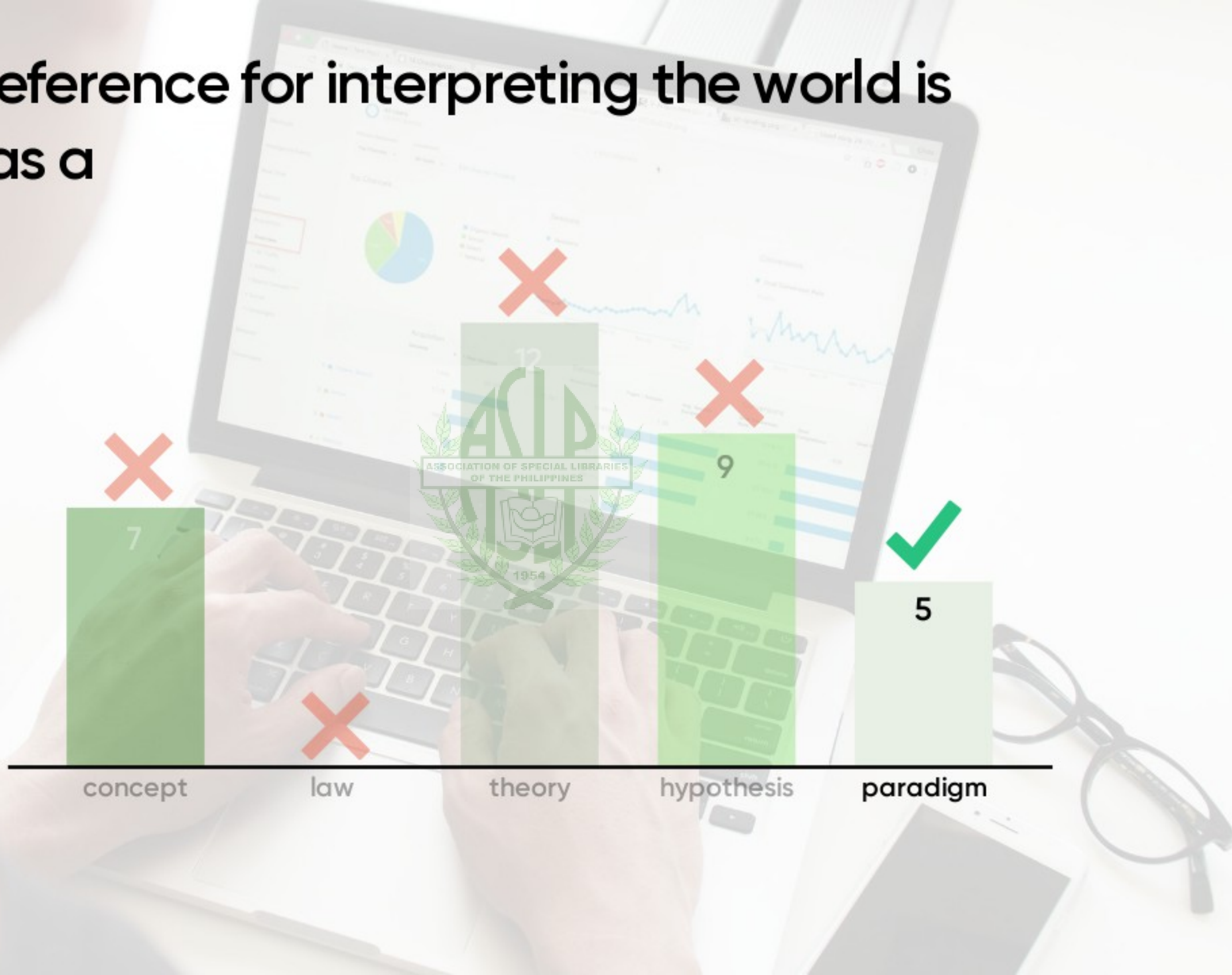
Which of the following statements is are true about the scientific method?



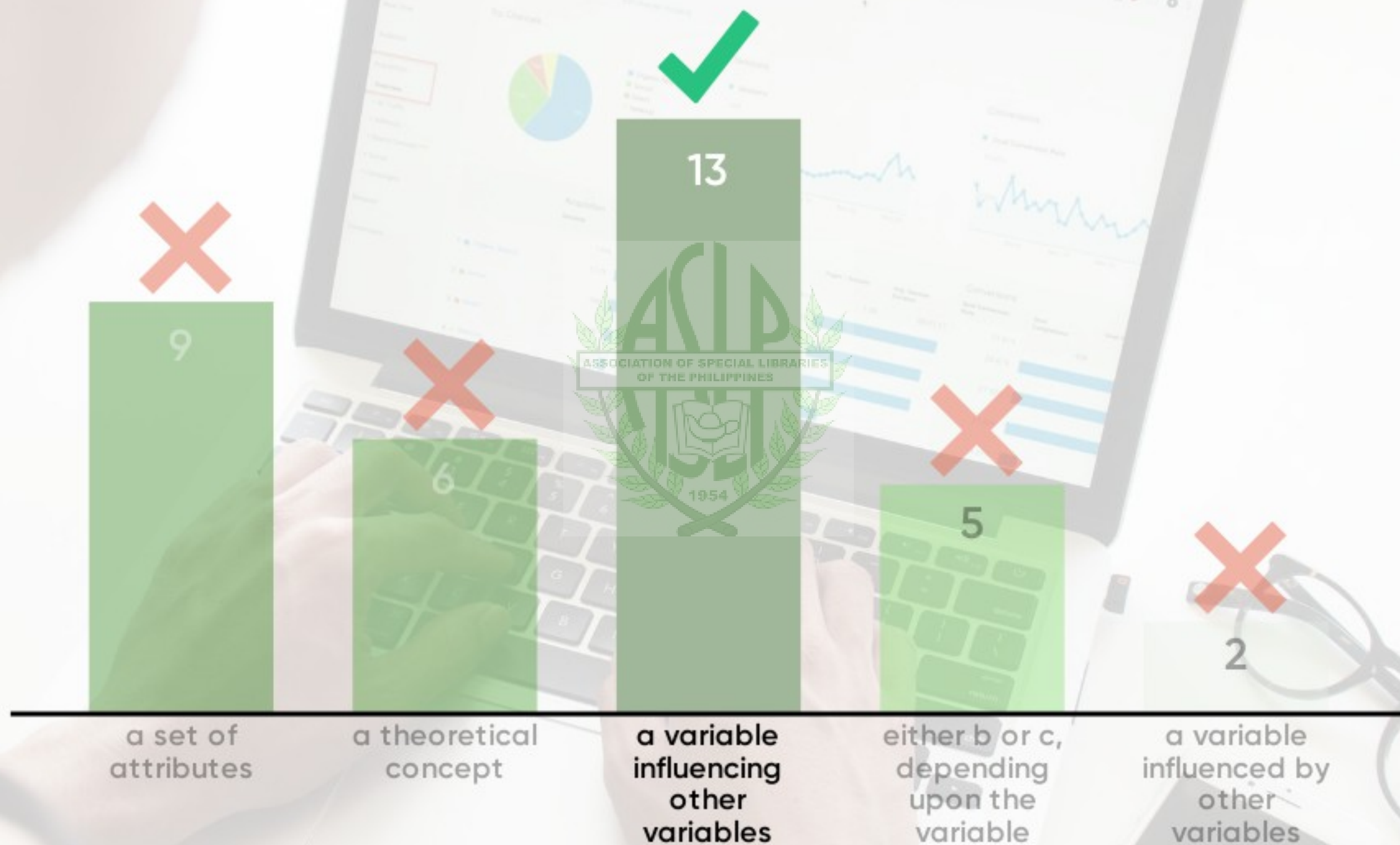
Leaderboard



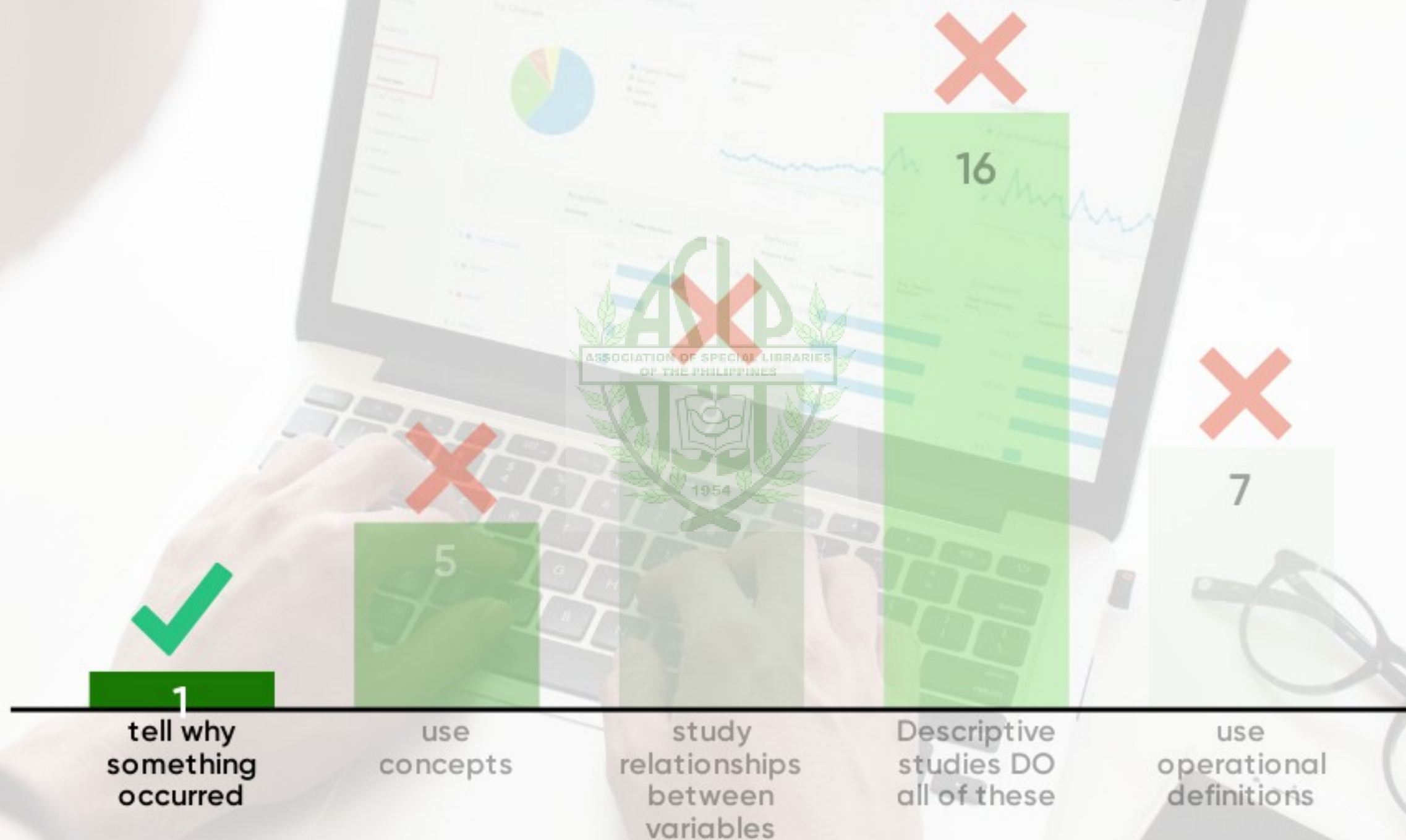
A frame of reference for interpreting the world is referred to as a



An independent variable is



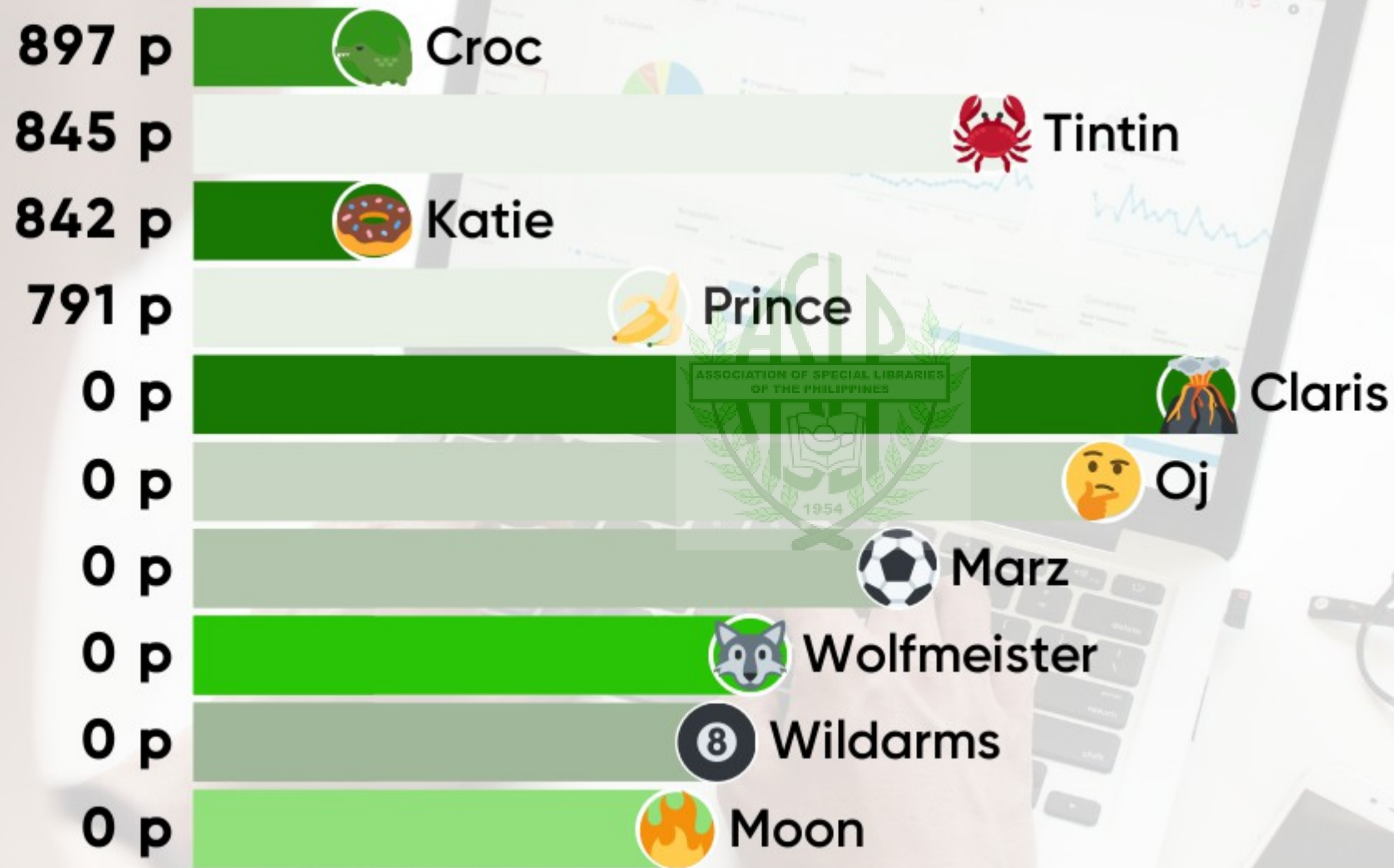
Descriptive studies do NOT



If a researcher wanted to test the hypothesis that female social workers are more empathic than male social workers, the researcher would design:



Leaderboard



Data Management Services: Collecting, Organizing, and Curating Datasets and Raw Data

Asst. Prof. Dan Anthony Dorado

UP School of Library and Information Studies

Bolstering Research: The Role of Librarian and Information Professionals
in the Ideation, Use, and Preservation of Outstanding Research

22-24 May 2016 | Bohol, Philippines

Agenda

- overview of data management services
- benefits in the promotion and use of research
- different frameworks and approaches



Digital Curation Centre

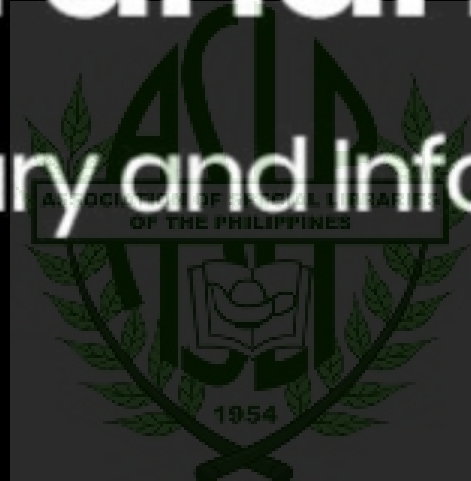
- internationally-recognised centre of expertise in digital curation
- "good research needs good data"
- <http://www.dcc.ac.uk>

Research Data

- Data sets: Observational, computational, simulated, or otherwise recorded output
- Digital collections: A grouping of digital objects, such as a photo archive or a vast text-based library of digitized books, can be interpreted as set
- Learning objects: Videos, digital online tutorials
- Multimedia: Recordings of film, music, and performance art
- Software: Applications including the code and documentation files

Philippine Librarian Census 2018

UP School of Library and Information Studies



<http://upslis.info/librariancensus/results/>

Repository

A repository (also referred to as a data repository or digital data repository) is a searchable and queryable interfacing entity that is able to store, manage, maintain and curate Data/Digital Objects. A repository is a managed location (destination, directory or 'bucket') where digital data objects are registered, permanently stored, made accessible and retrievable, and curated.

Do you have an Institutional Repository?



7

Yes



18

None



11

Sort of

What is Research Data Management?



Management

depository of all data as one

Management

Management and
administration of raw research
data

Organization and
management of research
data

Organizing of research data

Management of data
compiled

Organizing

Organizing and preservation

What is Research Data Management?



Repository

Managing research raw data

Storing, archiving data sets whilst making it accesible

It is managing the research data

Organization of data

Collection of information organized in a library

organization, storage, preservation, and sharing of data collected

Research purposes to discover and learn about data management.

Refers to on how to manage your research data

What is Research Data Management?

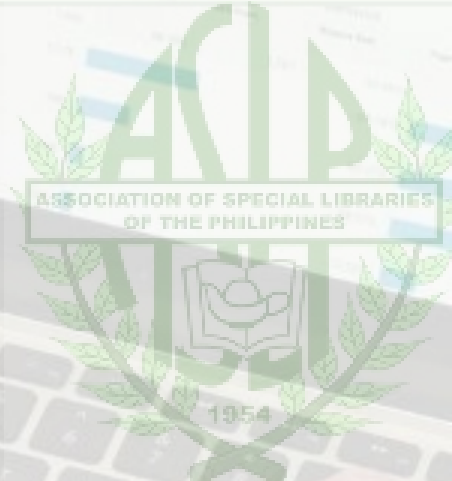


Practices and activities related to organizing your research data

Set of standards for storing organizing and making available data from research studies

Access

for easy retrieval



Why manage research data?

- Enable Reuse
- Research Integrity
- Research Impact
- Regulatory Requirements
- Controlling Cost
- Maximizing Value

Who are the main actors?

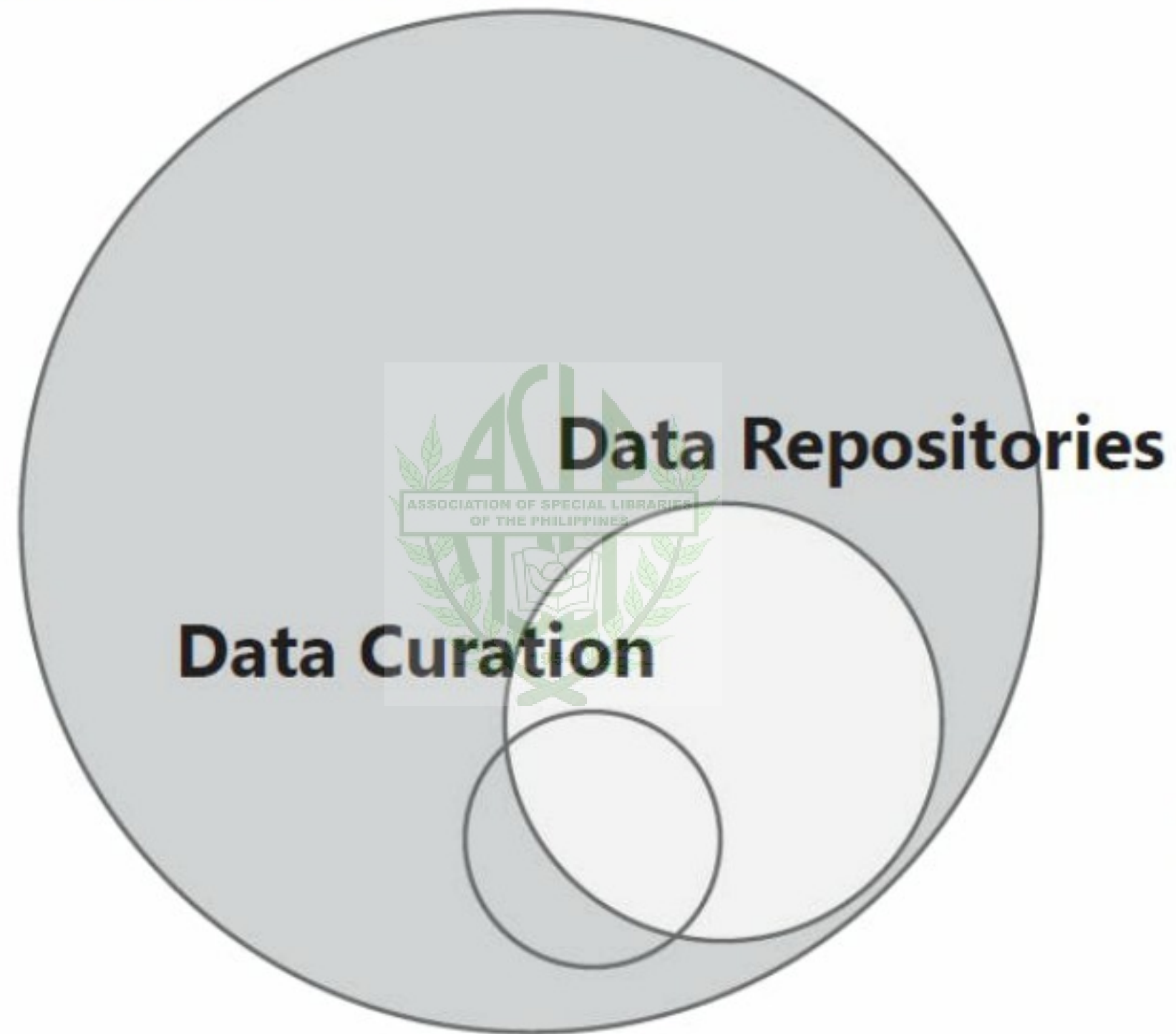
- Researchers
- Other data creators
- Other (re)users
- Funding bodies
- Data centers
- Computer Science Research
- Libraries / Archivist / Record managers

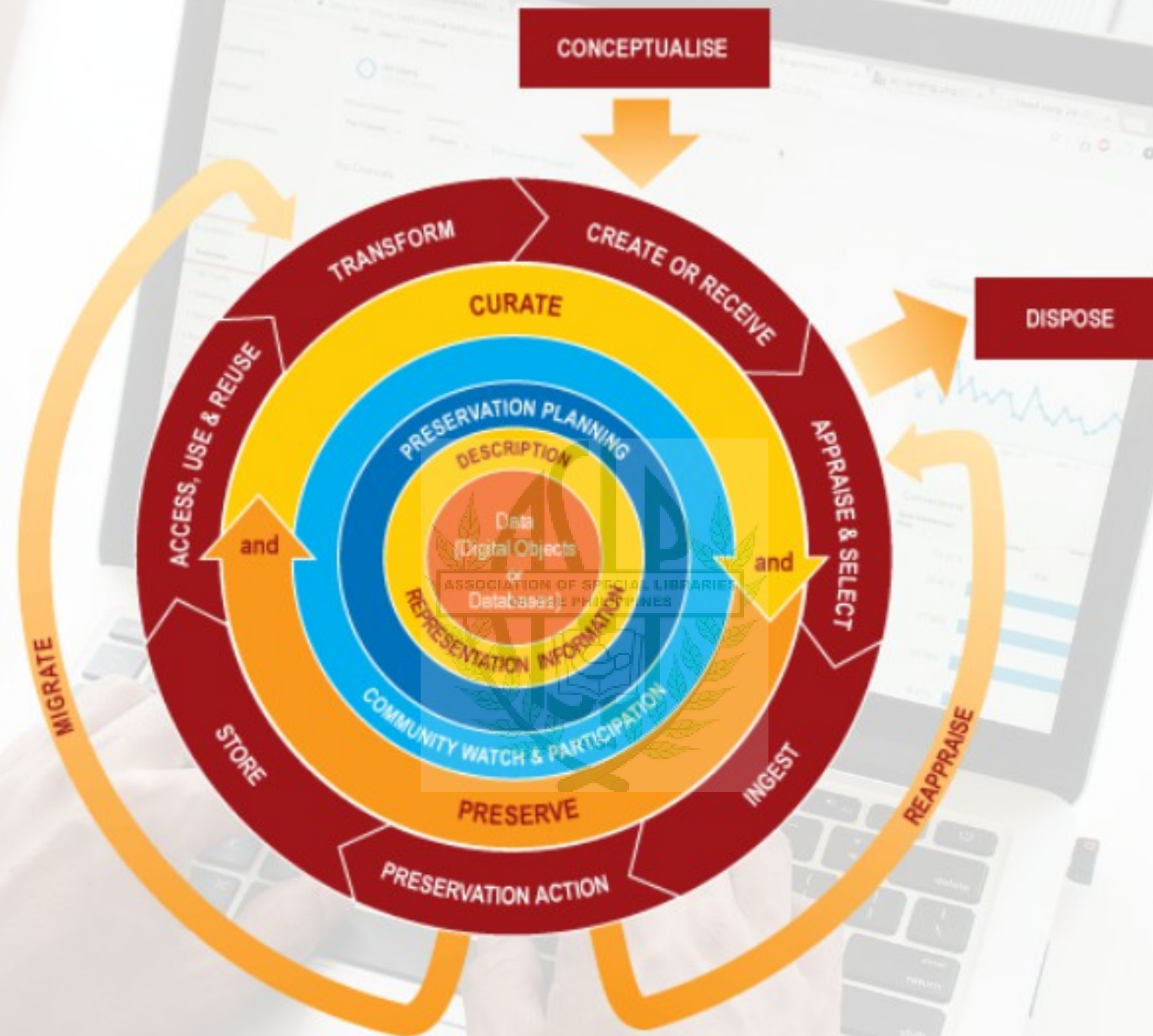
Requirements

- Technical Infrastructure
- Policy and Commitment
- Human Infrastructure

Research Data Frameworks

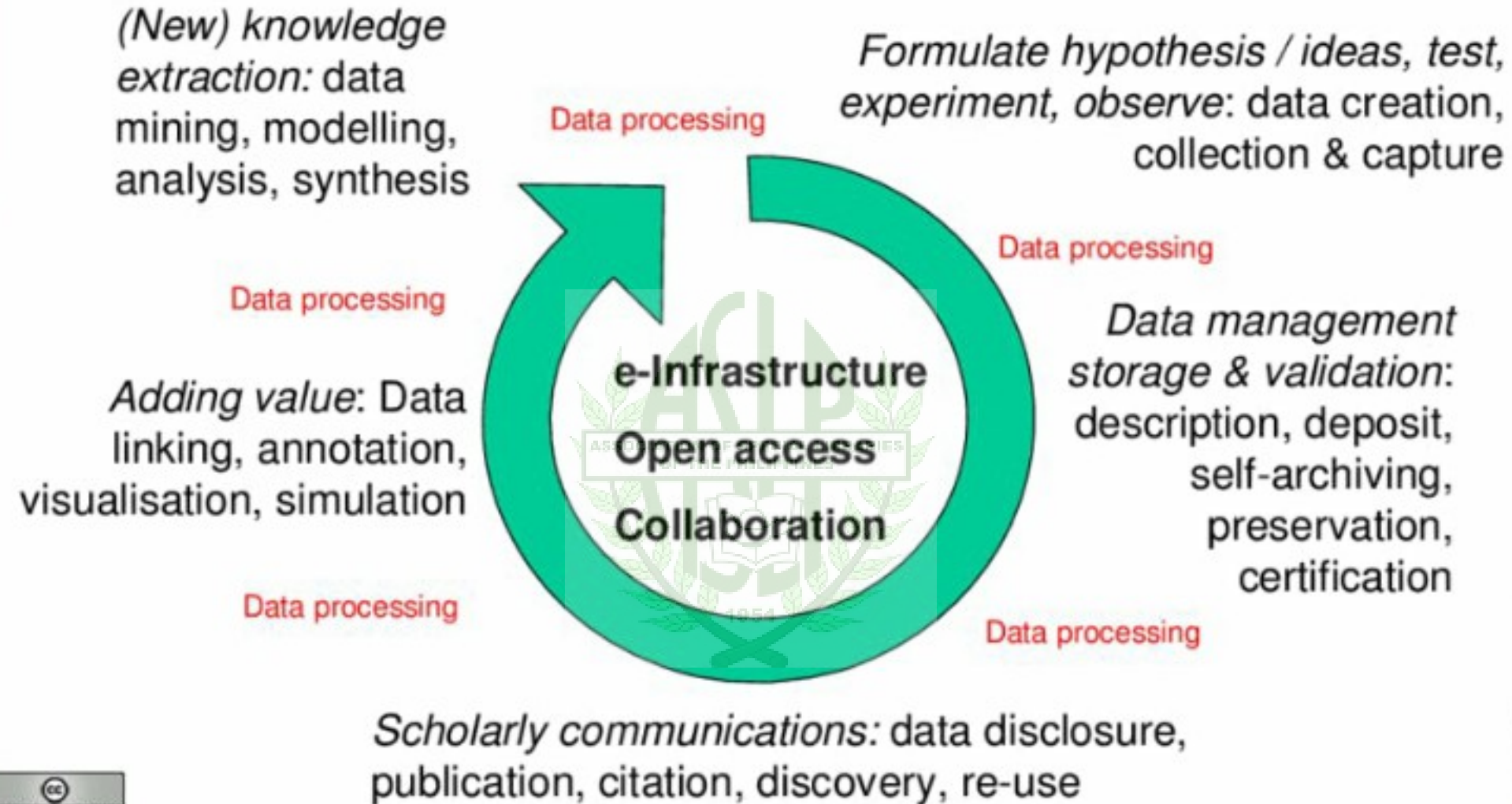
Research Data Services





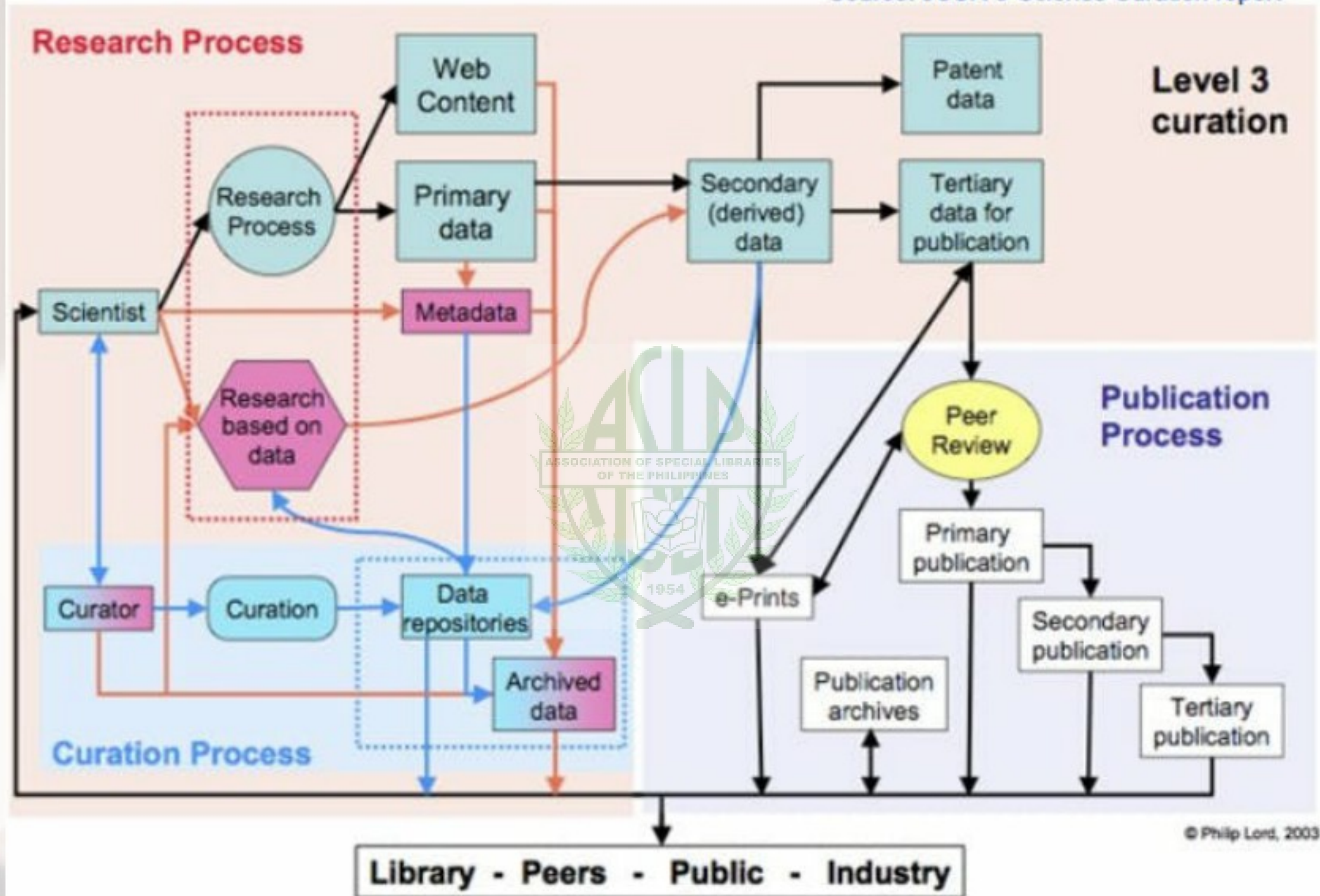
DCC Curation Lifecycle Model

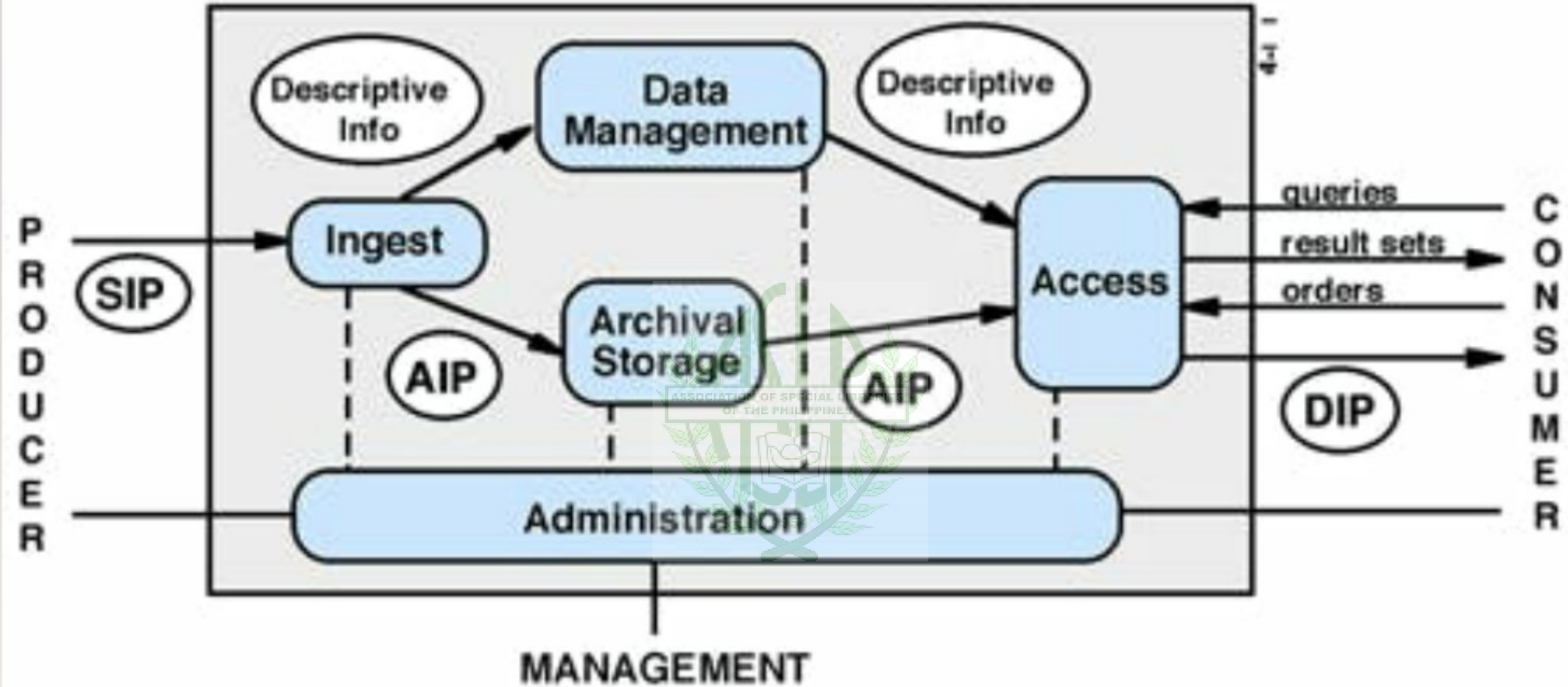
(e)-Research Life Cycle view of Data Curation?



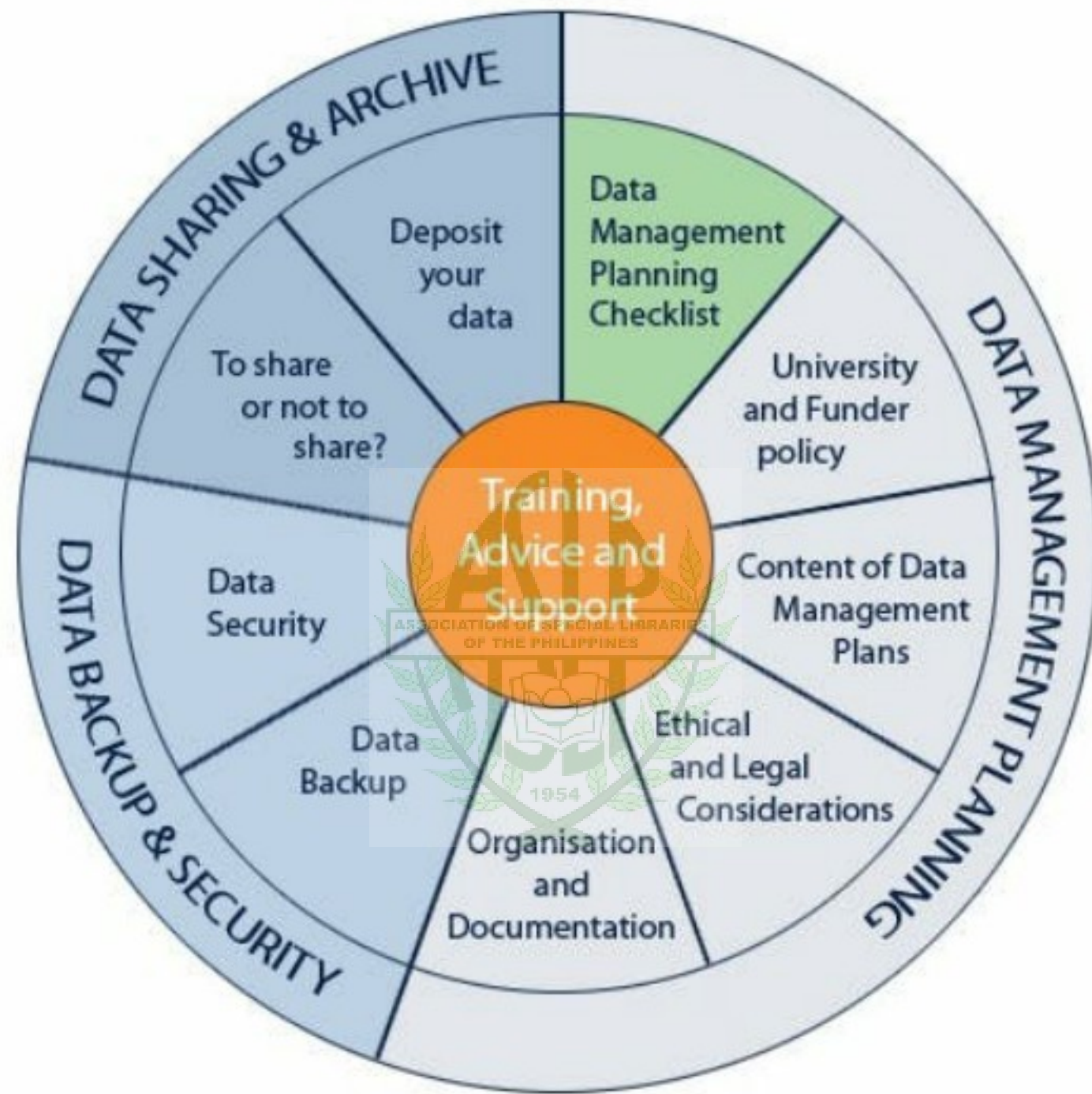
This work is licensed under a [Creative Commons License](https://creativecommons.org/licenses/by-sa/2.0/)
Attribution-ShareAlike 2.0

Data Curation





OAIS



Oxford RDM

Role of the library

- Leadership - coordination
- Audit
- Advice on access
- Preservation
- Promoting Data in Teaching
- Identifying Skills Gaps

Skill gaps of subject librarians (RLUK, 2012)

- Preserving research outputs
- Data management and curation
- Data mining and manipulation
- Metadata and Metadata Schema



Elaine Martin, Rong Tang, and Sanda Erdelez

Research Data Management

LIBRARIAN ACADEMY

A large, semi-transparent watermark of the ASIP logo is centered over the text "RDMLA".

RDMLA

RESEARCH DATA MANAGEMENT LIBRARIAN ACADEMY

<https://rdmla.github.io/home/>

What is RDMLA

- An online educational program offering individuals to learn about research data management practices
- Developed by a team of librarians who have expertise with RDM
- The Academy offers 8 units which can be taken individually or by group
- All units will be hosted online via the Canvas platform under CC-BY non commercial licensing

Partner Institutions

- Harvard
- Simmons
- Boston
- Brown
- MCPHS
- Northeastern
- Tufts
- Elsevier



RDMLA Curriculum

- Foundations of RDM
- Navigating Research Data Culture
- Advocating and Marketing the Value of RDM in Libraries
- Launching Data Services in Libraries
- Project management and Assessment
- Data Analysis and Visualization
- Coding Tools
- Platform Tools



Simmons University CE credits



OPEN DATA PHILIPPINES

search

Philippine Standard Time:
Thursday, May 23, 2019, 12:56:41 PM

Open Data for Filipinos everywhere

Looking for what the government spent on last year? Need statistics on public health services? Search for national government data using [GOV.PH/data](https://data.gov.ph/)

Search dataset...



Local data repository: <https://data.gov.ph/>



1

Developing a Research Data Management Service

Step 0: Establish Your Data Curation Service

In terms of establishing a service, using one word, describe this picture



Cardio

Mission impossible?

Focus in one direction

Fast

Direction

Running

Fast Speed

rush

Running

In terms of establishing a service, using one word, describe this picture

Speed

focus

Perseverance

Never ending

Focus

Goal

Speed

Vision

Mission

In terms of establishing a service, using one word, describe this picture



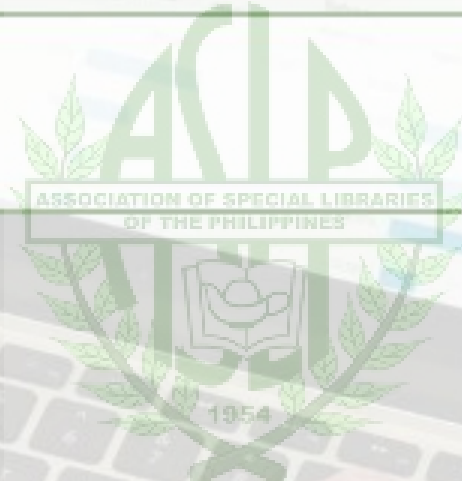
continues

Running

Fast

Mission

Fgg



Step 0.2



Policy

✓

Plannung

✗

Draft objectives

✗

Plan

✗

10x

Research

✗

List

✗

Objectives

✗

Objective

✗

undefined policy

write

✗

5x

Step 0.2



Concepts

✗

Method

✗

Record

✗

Goal

✗

Vision

✗

2x

Writing

✗

3x

undefined policy

Leaderboard



0 p		En2
0 p		Nicandro
0 p		Blink
0 p		Sg
0 p		reyes
0 p		Karen
0 p		Bohol
0 p		Dr Centipede
0 p		Rica
0 p		Pacman



Policies

Step 0.3



Audience

✓

5x

Appreciate

✗

Promotion

✗

Stakeholders

✗

2x

present

✗

2x

Correct

✗

Support

✗

3x

Communicate

✗

undefined audience

Promote

✗

Step 0.3



Accepted ✗

Performance ✗

Clients ✗

Dessiminate ✗

Discuss ✗

Presenting ✗

Approved ✗

approval ✗

undefined audience^{2x}

Dissemination ✗

Step 0.3



bob

Observation

Presentation



undefined audience



Target Audience

Step 0.4



Budget

✓

24x

Cost

✓

Save

✗

3x

Funding

✗

5x

Fund

✗

Fundinh

✗

Money

✗

Savings

✗

undefined cost

bob

✗

Step 0.4



Savings budget



Funds



invest



undefined cost



Cost

Step 0.5



Staff

3x

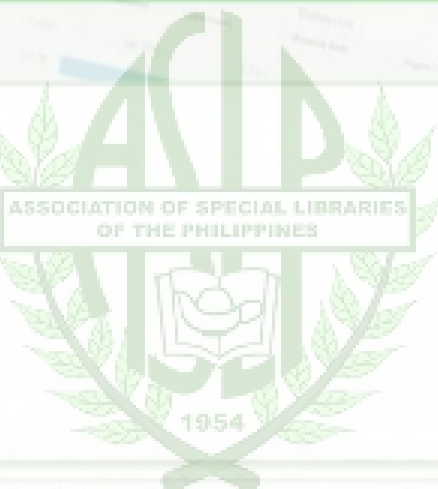
Staffing

2x

Personnel

Appreciate

Planning



4x

Leveling

Manpower

feedback

Deliberate

undefined staff

Step 0.5



Support

2x

Coordinate

Presentation

Agenda

People



2x

Decision

propose

Implementation

undefined staff

Plan

Step 0.5



Satisfaction



Collaboration



Human Resource



undefined staff



GREAT STAFF MEETING, THANKS.

Staffing

Step 0.6



Infrastructure

✓

3x

Storage space

✗

Repository

✗

2x

Filing

✗

Technology

✗

Storage

✗

5x

Server

✗

2x

Storing

✗

undefined infrastructure

Documentation

✗

Step 0.6



Facility

×

2x

Managing

×

Resources

×

Preservation

×

implementation

×

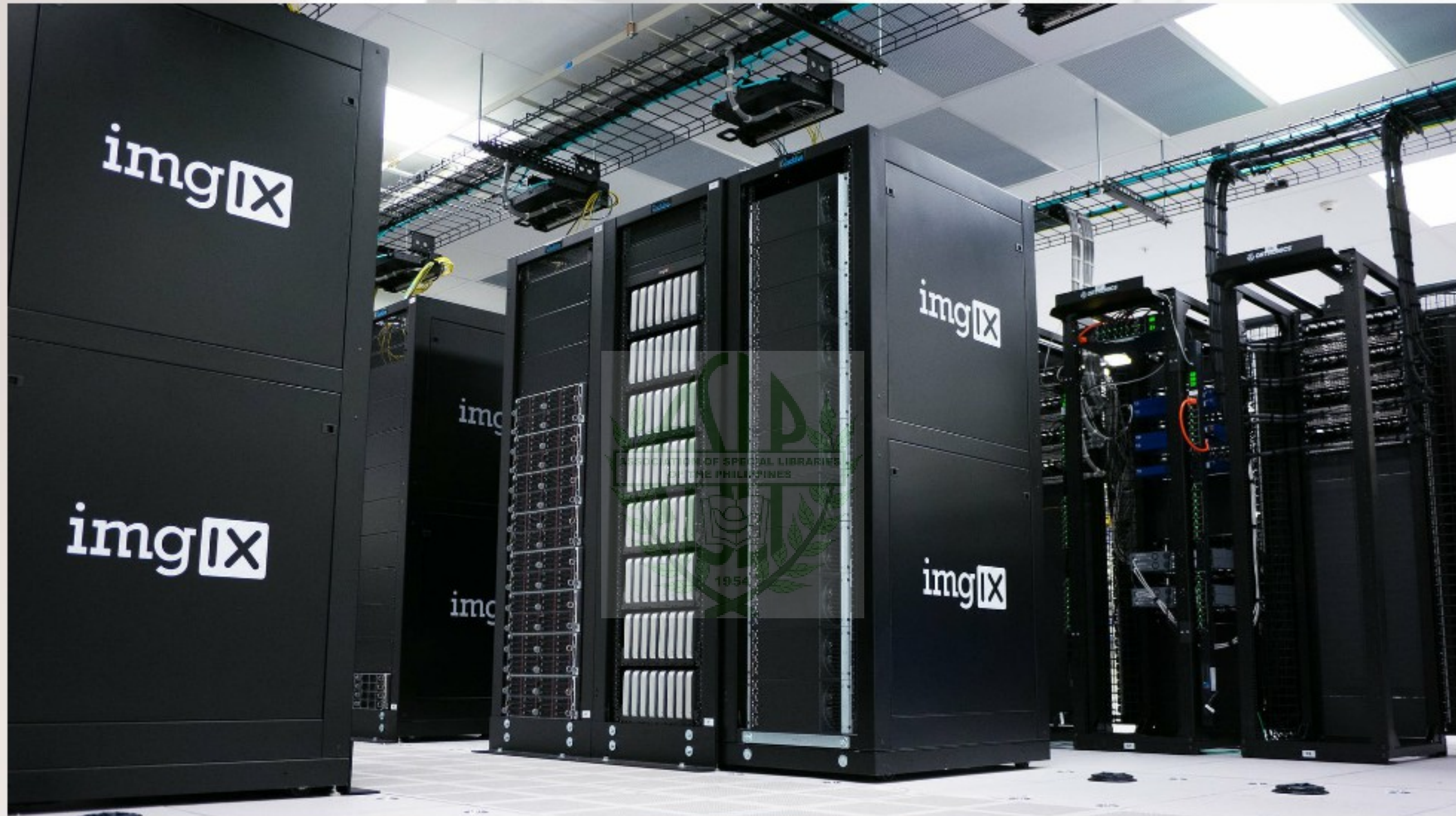
space

×

Network

×

undefined infrastructure



Technological Infrastructure

Leaderboard



Developing a Research Data Management Service

Step 1: Receive the Data

Step 1.1



Email

✕

Agreement

✕

Impression

✕

suggestions

✕

Asking

✕

2x

Ideas

✕

Comment

✕

Vission

✕

undefined recruit

Feedback

✕

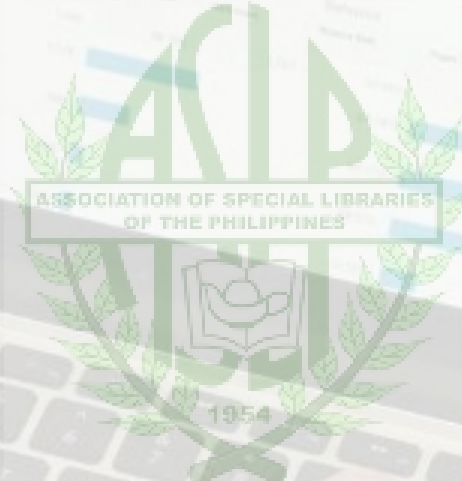
Step 1.1



W& A



select



undefined recruit



What Say You?

Recruit Data for Your Service



Step 1.2



Negotiate ✓

Benefit ✗

ask ✗

Funding ✗
2x

Salary ✗
4x

incentive ✗
3x

Budget ✗
2x

Reward ✗
undefined negotiate

bargaining ✗

Step 1.2



Priorities ✗

Offer ✗

2x

Procurement ✗

Buy ✗

Bargain ✗

Consultant ✗

Selecting ✗

undefined negotiate



Negotiate Deposit

Step 1.3



Agreement

✓

15x

Agree

✓

3x

Accept

✗

2x

Reward

✗

Deal

✗

6x

Collaborate

✗

Agreement/deal

✗

Coordinate

✗

undefined agree

Linkages

✗

Step 1.3



Partnership

acknowledging

Approval



undefined agree



Obtain Author Deposit Agreement

Step 1.4



Transfer

✓

Salary

✗

Payment

✗

5x

Payment terms

✗

Allocation

✗

2x

Budget

✗

2x

Communicate

✗

Implement

✗

undefined transfer

Funding

✗

3x

Step 1.4



Allocate



Send



Cost



Selling



Construction



contract

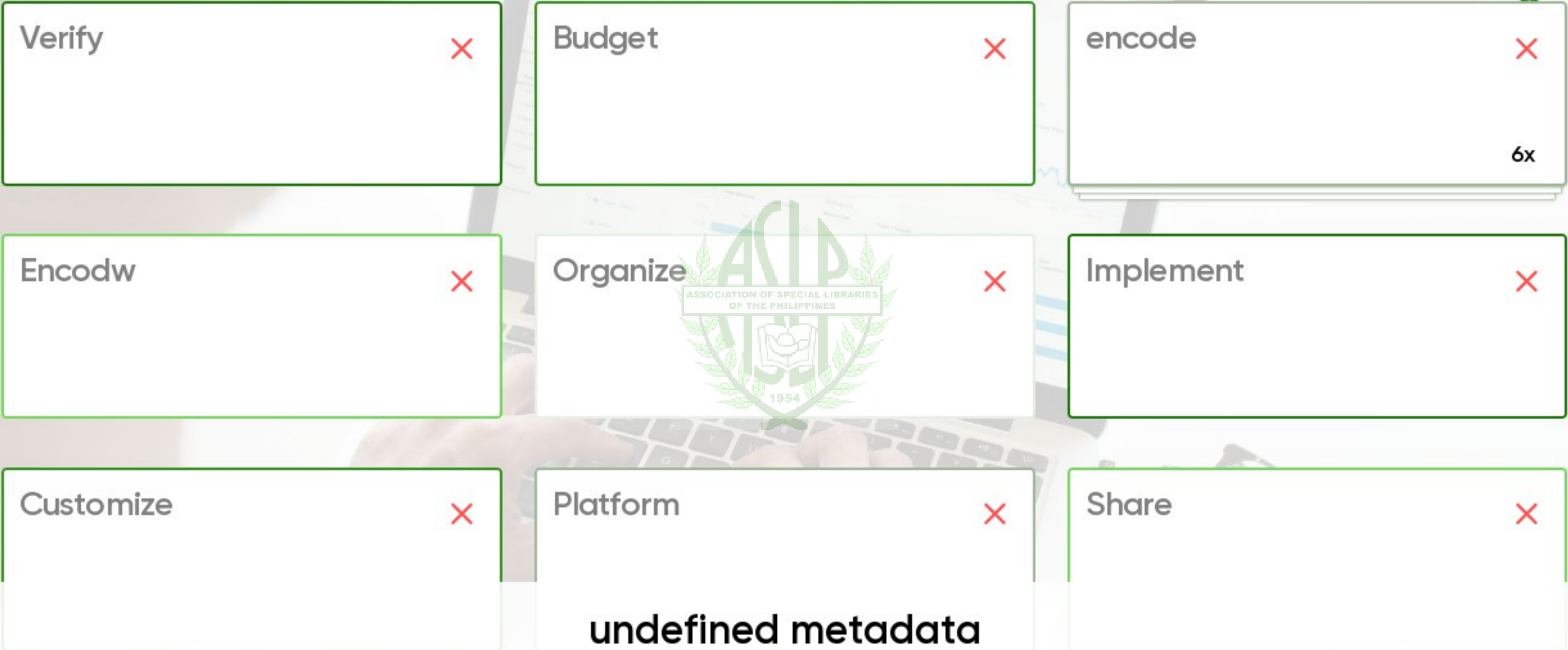


undefined transfer



Facilitate Transfer of Data

Step 1.5



Step 1.5



Format



2x

Filling



uploading



Translate



describe



undefined metadata

localhost:3000/#福

Hanzi decomposition

9373/9373

Metadata for 福 - 福

Definition: happiness, good fortune, blessings

Pinyin: fú

Strokes: 13

References: [Arch Chinese](#) • [MDBG](#) • [Wiktionary](#) • [YellowBridge](#)

Edit path

Choose a source for glyph path data for this character: [AR PL Kai GB](#) • [AR PL UKai](#)

Alternatively, use part of a glyph from another character: ?

- No path data.



Obtain Metadata and Documentation

Step 1.6



Assign

Search

Promote

Validate

payment

Dissiminate

Code

Circulate

Organize

undefined notify

Step 1.6



Process



Mail



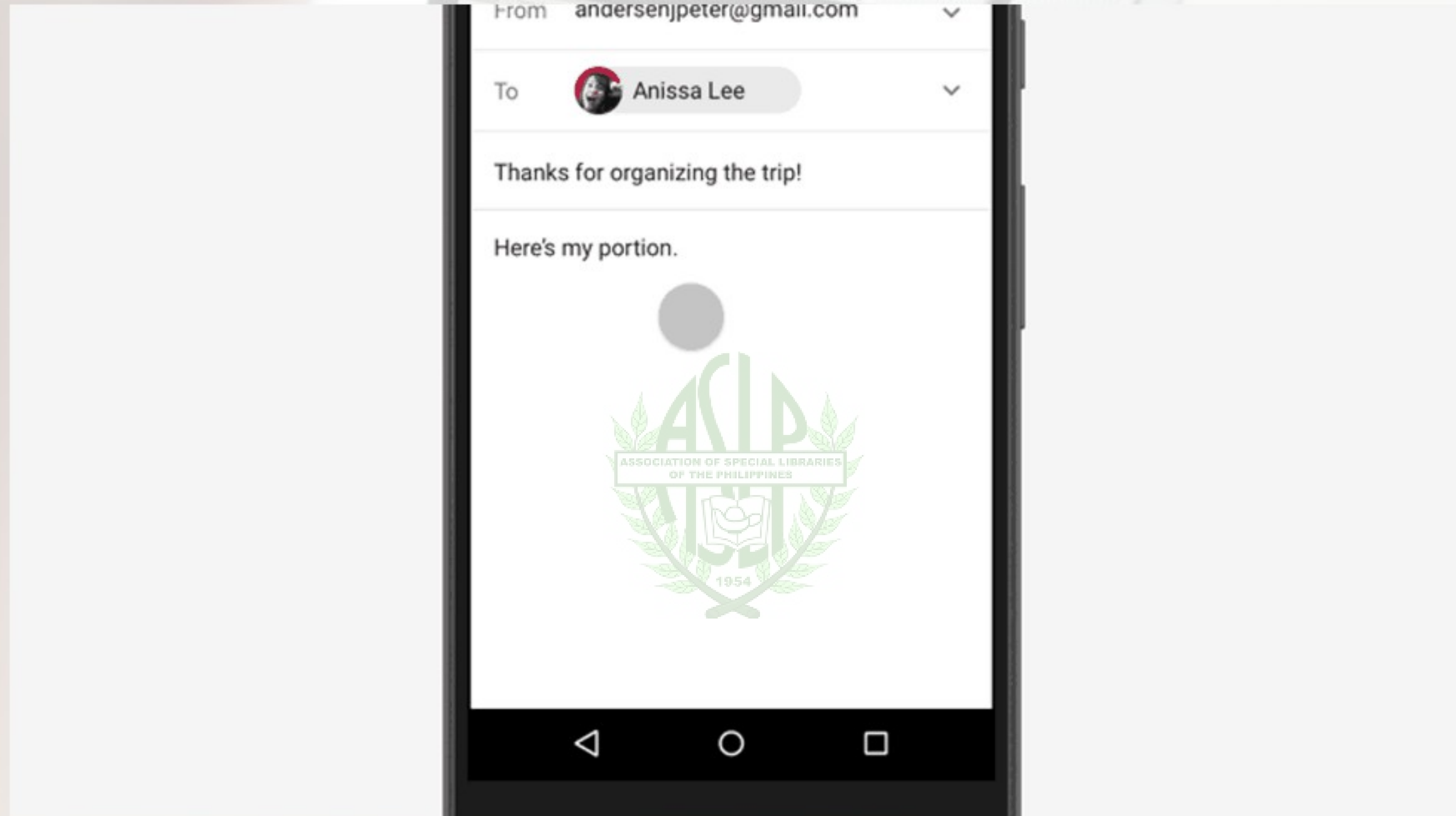
Calculate



Store













undefined notify



Receive Notification of Data Arrival

Leaderboard



3000 p		Blink
3000 p		59997
2000 p		En2
2000 p		Hey
1000 p		Karen
1000 p		Bohol
1000 p		Dr Centipede
1000 p		Rica
1000 p		Max
1000 p		Ee

Developing a Research Data Management Service

Step 2: Appraisal and Selection Techniques That Mitigate Risks Inherent to Data

Step 2.1



Appraise

✓

3x

Appraisal

✓

2x

Salute

✗

2x

Approve

✗

2x

Comment

✗

Select

✗

Evaluate

✗

judge

✗

undefined appraise



Appraise the Files

Step 2.2



Classify ✗

Circulate ✗

Review ✗

Virus ✗

Errors ✗

Security ✗

Evaluate ✗

Cycle ✗

undefined risk

process ✗

Step 2.2



Brand

branding

2x

Ask

Explode

undefined risk



Consider Any Risk Factors

Step 2.3



Inventory

11x

Classify

Record

Organize

Craft



2x

Mine

Strategic planning

Merge

undefined inventory

Evaluate

Step 2.3



Logic



sorting



undefined inventory



Inventory the Submission

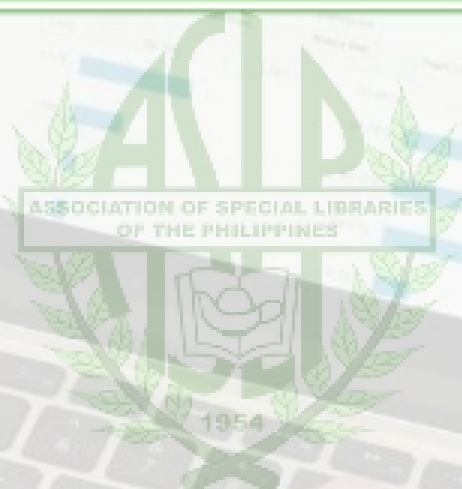
Step 2.4



Select✓
23x

Selection✗
6x

Selecting✗



undefined select



Select

Step 2.5



Assignment

✓

Time frame

✗

Duedate

✗

Contact

✗

Negotiate

✗

Deadline

✗

Extension

✗

2x

Discussion

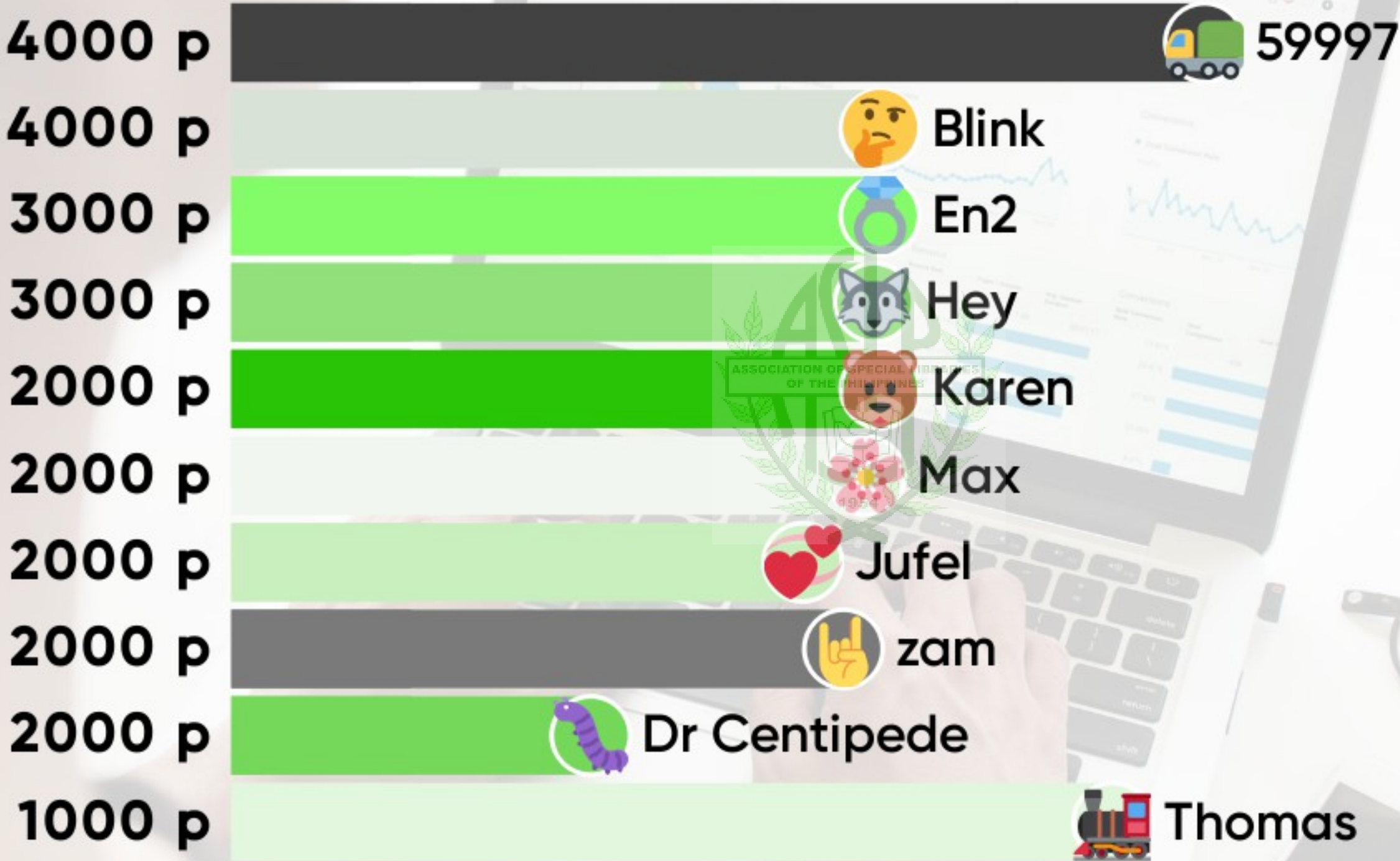
✗

undefined assign



Assign the Submission

Leaderboard



Developing a Research Data Management Service

Step 3: Processing and Treatment Actions for Data

Step 3.1



Code ✗

Security ✗
19x

Lock ✗

Secure ✗
11x

Cure ✗

Clean ✗

Secure/security ✗

Secured ✗
undefined

Key ✗

Step 3.1

Keys



undefined



Secure the Files

Step 3.2



Start

✓

14x

burn

✗

Upload

✗

Input

✗

Circulate

✗

2x

testing

✗

Initialize

✗

Distribution

✗

undefined start

Availability

✗

Step 3.2



Proof read ✗

Run ✗

Load ✗

Back up ✗

Organize ✗

begin ✗

Initiate ✗

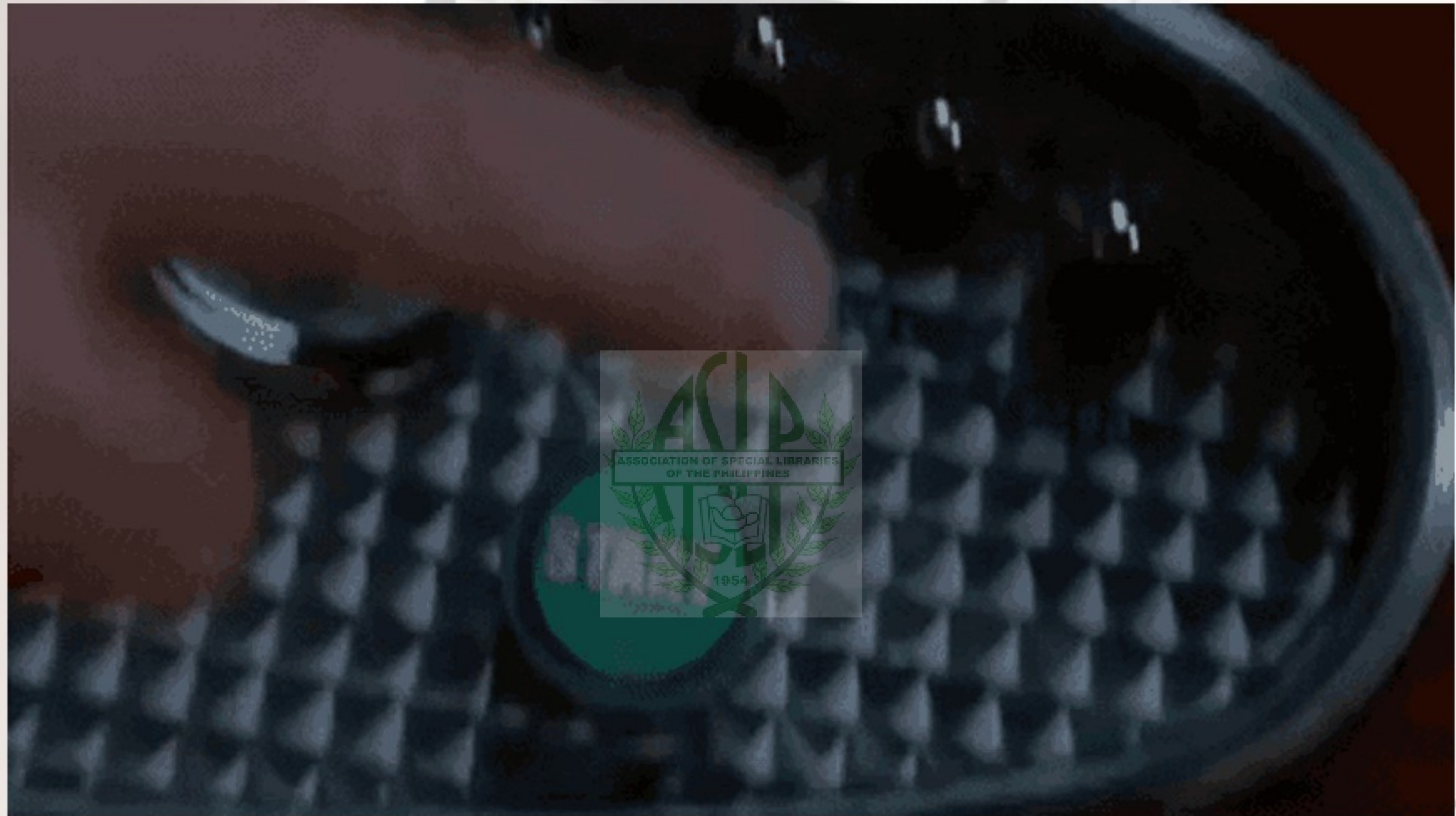
encode ✗

undefined start 2x

Stary ✗

Step 3.2





Start a Curation Log

Step 3.3



Assess

✗

2x

Rules

✗

Decide

✗

8x

Judge

✗

8x

Analize

✗

decisions

✗

2x

Decid3

✗

justify

✗

undefined order

Approve

✗

Step 3.3



Decision ✗

Judgment ✗

Agree ✗

evaluate ✗

2x

Judgement ✗

2x

Approval ✗

Justice ✗

undefined order



Inspect File Representation and Organization

Step 3.4



Inspect

✓

5x

Observe

✗

2x

verify

✗

2x

Examine

✗

6x

Readable

✗

Confirm

✗

Check

✗

Watch

✗

undefined inspect

Finding aid

✗

Step 3.4



Evaluate



2x

Define



Clarify



search



2x

Magnify



investigate



undefined inspect



Inspect the Data

Step 3.5



Recording

×

2x

Sign

×

5x

Record

×

10x

stamping

×

Labeling

×

draw

×

Process

×

Claim

×

undefined author

Log

×

Step 3.5



Launch



Circulate



Access



Drafting



Signature



Signatory



Collaring



undefined author



Work with the Author to Enhance the Data Submission

Step 3.6



File

✓

14x

Filing

✗

17x

Store

✗

3x

Filling

✗

Record

✗

Storing

✗

archiving

✗

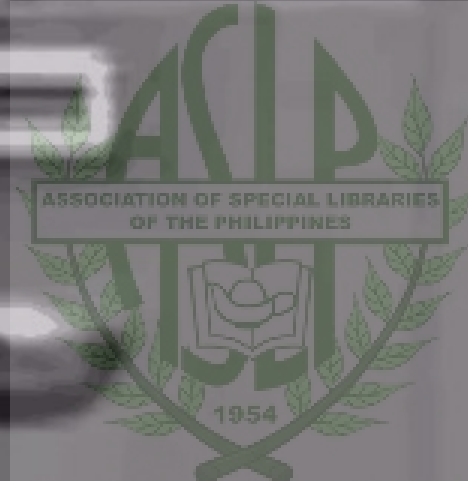
Storage

✗

undefined file

Step 3.6





Consider File Formats

Step 3.7



Describe

Interpret

Record

keep

Distribution

Measure

Size

Check

Monitor

undefined describe

Step 3.7



Ask



access



Analyze



Exam



meeting

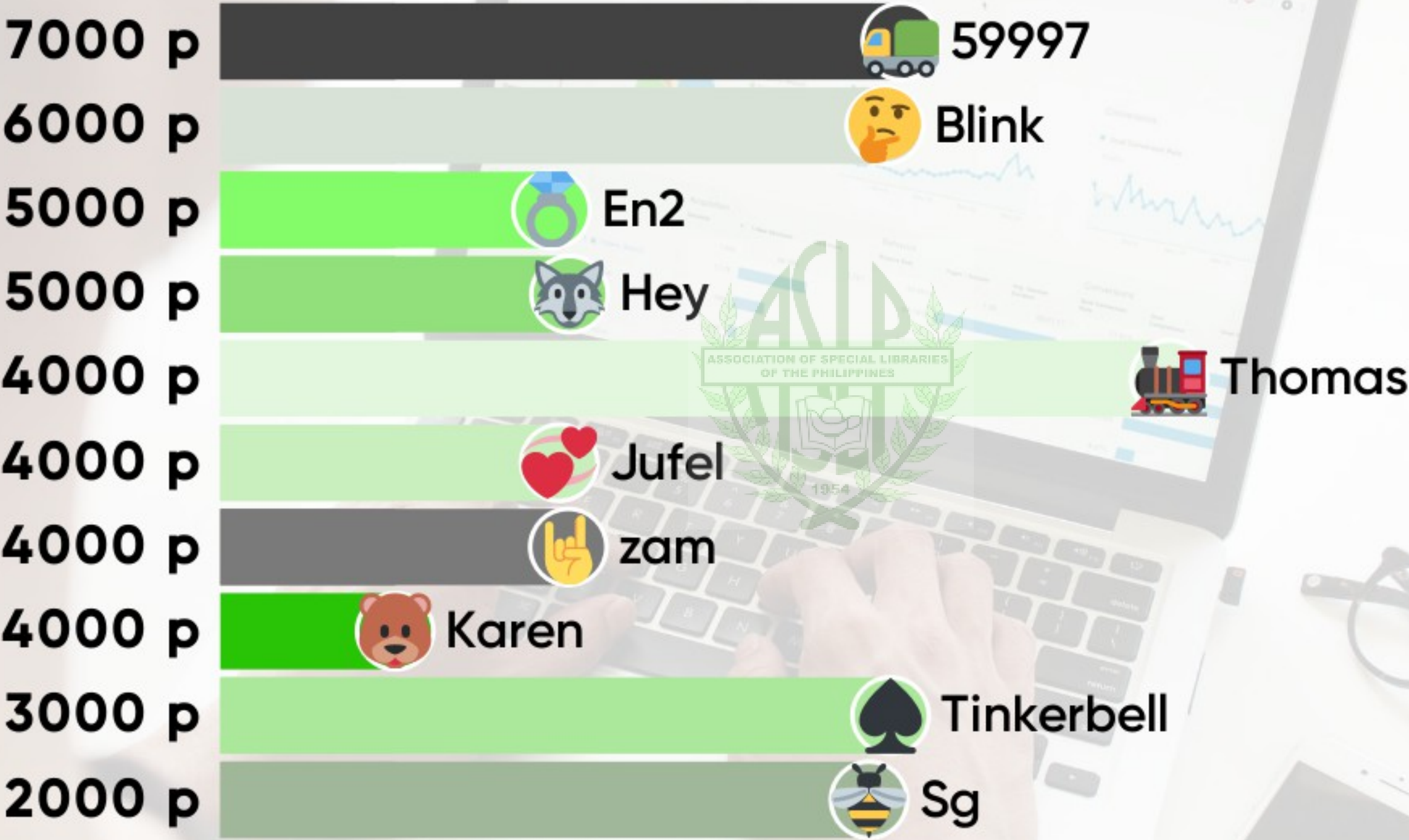


undefined describe



Arrangement and Description

Leaderboard



Developing a Research Data Management Service

Step 4: Ingest and Store Data in Your Repository

Step 4.1



Ingest

6x

Digest

4x

usage

Feed

3x

Access

2x

Eating

2x

Eat

3x

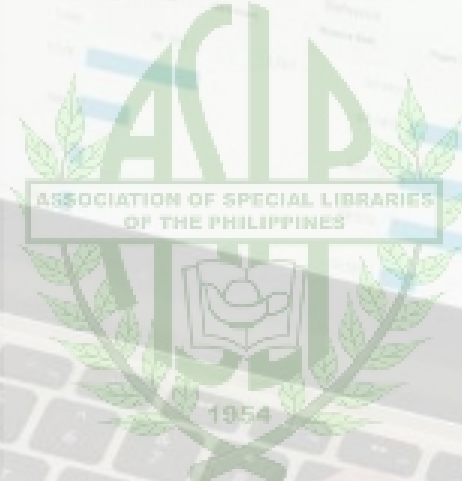
Consume

undefined ingest

Fees

Step 4.1

Use



undefined ingest



Ingest the Data Files

Step 4.2



Store

2x

Buy

2x

Shop

Choose

2x

Select

9x

gather

2x

Organize

Sort

undefined store

sorting

Step 4.2



Pick



Selecting



Market



undefined store



Store the Assets Securely

Step 4.3



Trust

✓

2x

Promise

✗

4x

store

✗

None

✗

Data

✗

Deal

✗

2x

Lock

✗

Fulfill

✗

undefined trust

Bind

✗

Step 4.3



secure



Entrust



Katie



agreement







undefined trust



Develop Trust in Your Repository

Leaderboard



9000 p		59997
8000 p		Blink
7000 p		Thomas
6000 p		En2
6000 p		Hey
5000 p		Tinkerbelle
5000 p		Jufel
5000 p		zam
4000 p		Karen
4000 p		Hello ♀

Developing a Research Data Management Service

Step 5: Descriptive Metadata

Step 5.1





Create and Apply Descriptive Metadata

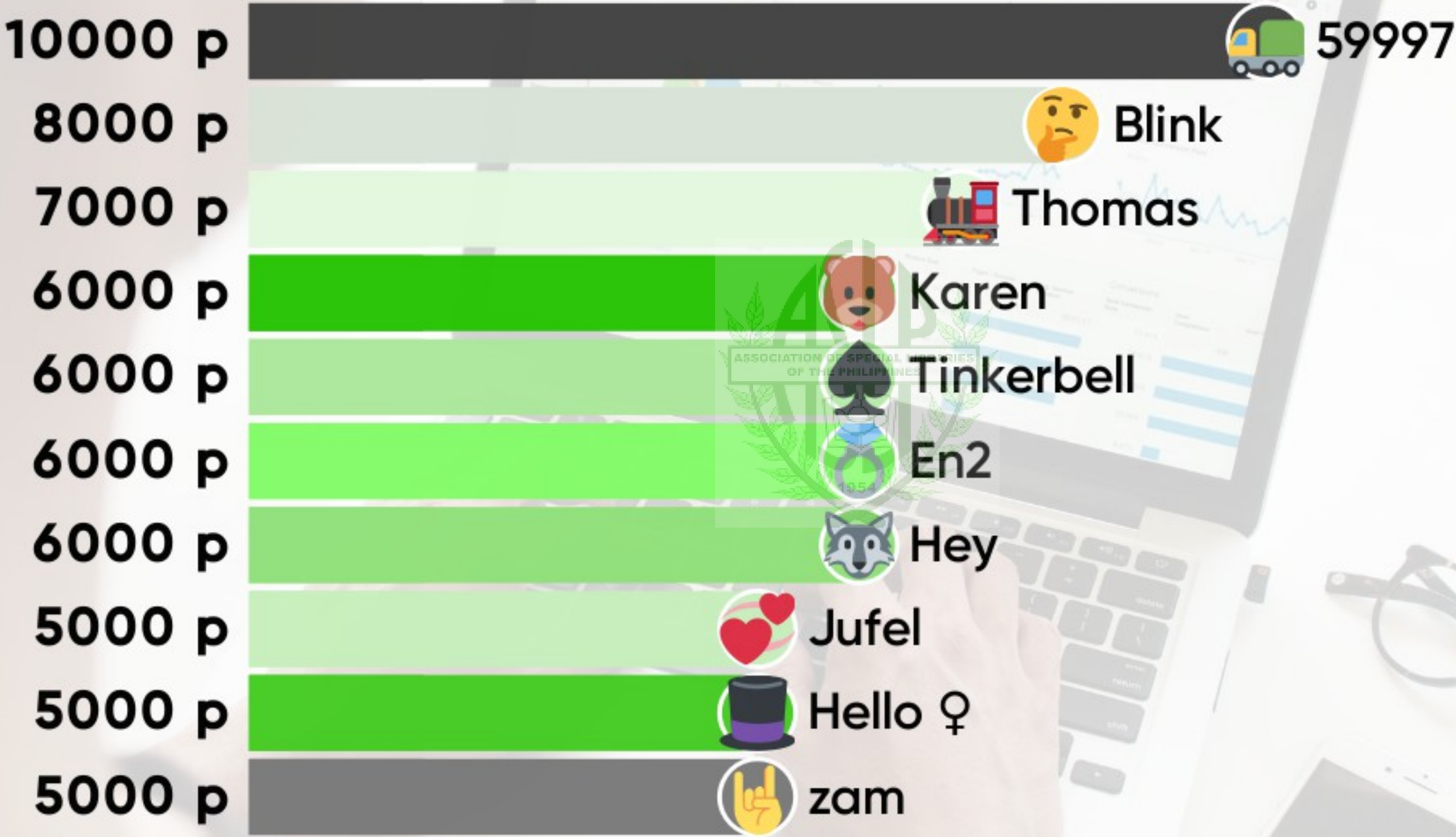
Step 5.2





Consider Metadata Standards for Disciplinary Data

Leaderboard



Developing a Research Data Management Service

Step 6: Access

Step 6.1





Determine Appropriate Access Conditions

Step 6.2

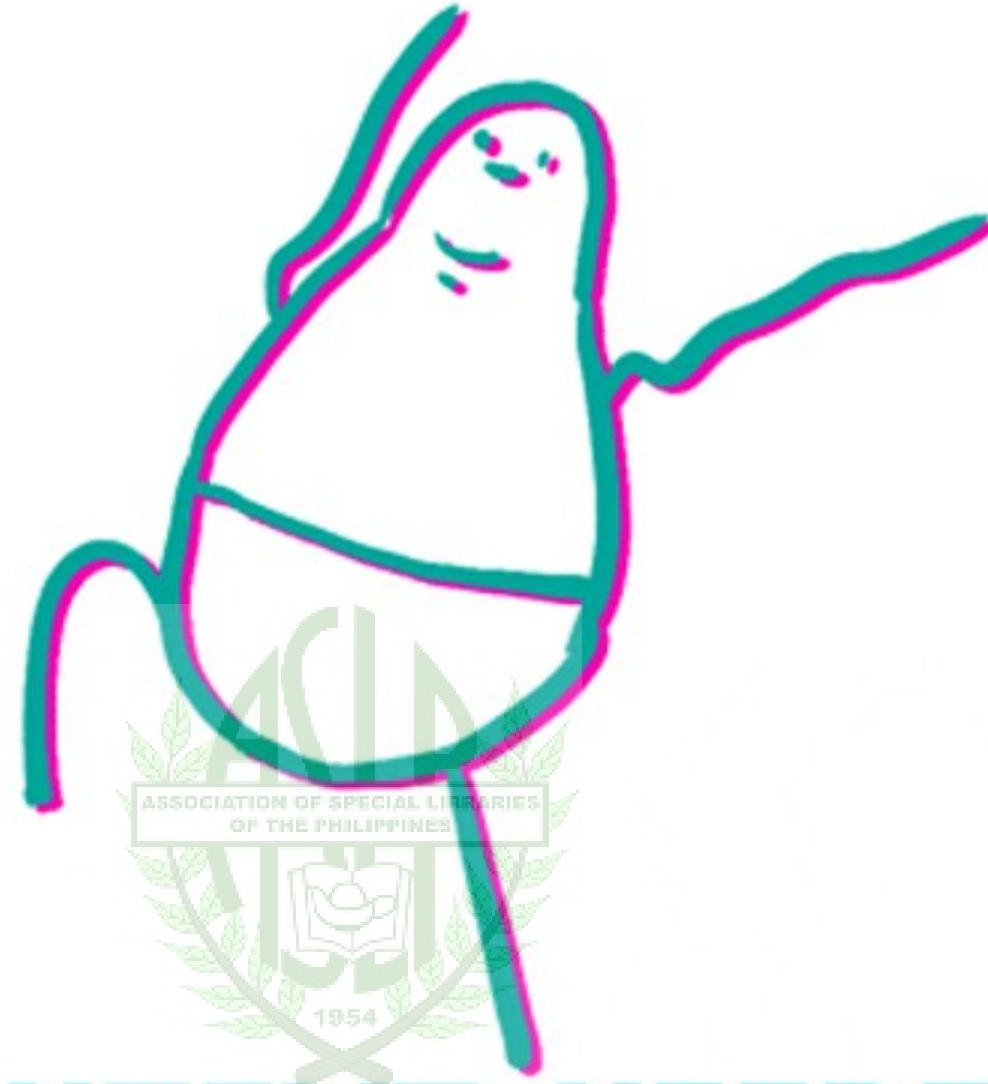




Apply the Terms of Use and Any Relevant Licenses and Copyright Notices for the Data

Step 6.3





NO CONTEXT NEEDED

LISA VERTUDACHES

Contextualize the Data

Step 6.4





Enhance the Submission to Increase Exposure and Discovery

Step 6.5





Apply Any Necessary Access Controls

Step 6.6





Ensure Persistent Access

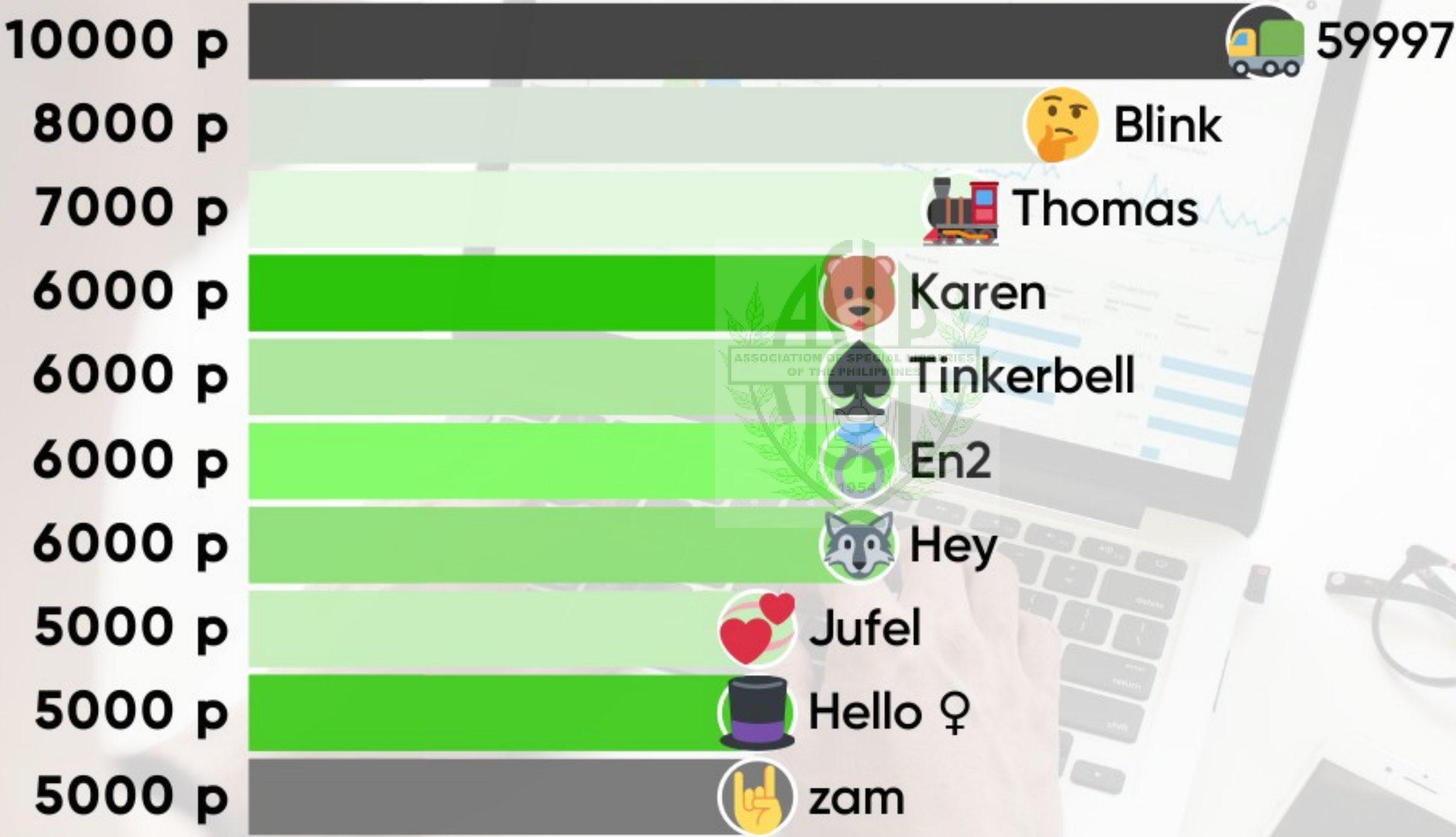
Step 6.7





Release Data for Access and Notify Author

Leaderboard



Developing a Research Data Management Service

Step 7: Preservation of Data for the Long Term

Step 7.1

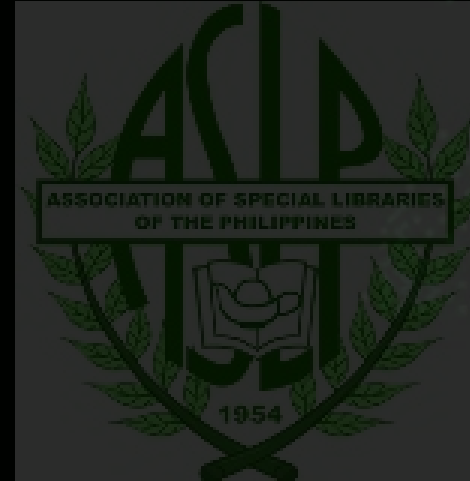




Plan for Long-Term Reuse

Step 7.2





Monitor Preservation Needs and Take Action

Developing a Research Data Management Service

Step 8: Reuse

Step 8.1





Consider Post-ingest Review Techniques

Step 8.2



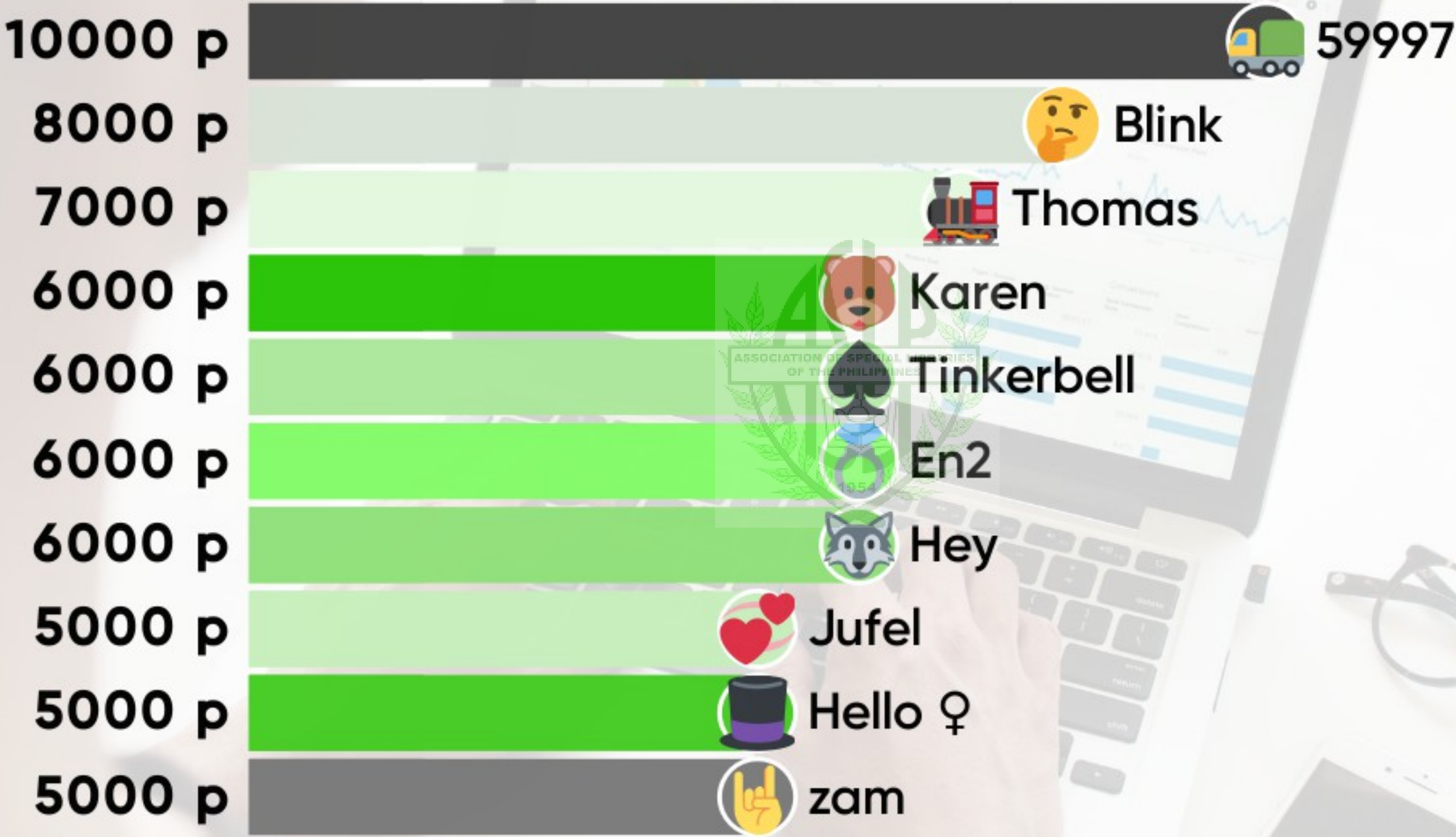


Provide Ongoing Support as Long as Necessary

Step 8.3



Leaderboard



twitch.tv/HyperRPG



STOP!

Cease Data Duration



Questions



0
questions



0
upvotes

