BioDialog Project: Scientific Biodiversity Data Management

Prof. Dr. Taysir Hassan A. Soliman
Vice Dean for Graduate Studies & Research
Faculty of Computers & Information, Assiut University
Assiut University BIODIALOG PI

Dr. Ibrahim Fathy Moawad
Associate professor, Ain Shams University
Decision support systems dept. manger, supreme council of universities
Ain Shams BIODIALOG PI

Agenda

- BioDialog Project
- BioDialog Summer School 2017
- Scientific Biodiversity Data Management
- The BEXIS System

weblink: http://fusion.cs.uni-jena.de/bdg/

BioDialog Motivation

Climate Change

weblink: http://fusion.cs.uni-jena.de/bdg/

Environmental Risks on Biodiversity

- The global risks report published in 2016 by the world economic forum

weblink: http://fusion.cs.uni-jena.de/bdg/

BioDialog Motivation

- Biodiversity Informatics recent study for countries ranking.

1. The biodiversity potential in partner countries.
2. The capacity to generate biodiversity data: raw data with high quality.
3. The availability of technical infrastructure: hosting, managing and sharing biodiversity data records.

    - Germany: 12, Egypt: 88 and Tunisia: 103.

SO, we need more education & research on biodiversity and biodiversity informatics.

weblink: http://fusion.cs.uni-jena.de/bdg/

Data Conservation

Aswan Botanich garden

Egypt's other endemic butterfly is the Sinai Hairstreak. Like the Sinai Baton Blue, it only occurs in and around Mt. Sinai.

BioDialog

BioDiversity Without Boundaries

BioDialog Objectives

1. Understanding biodiversity and biodiversity informatics practices in local context.
2. Constructing new regional research networks.
3. Contributing in the development of a knowledge-based society.
4. Bridging the gap between data management techniques and biodiversity research.

BY exchanging staff members & PhD/MSc students for training, attending workshops, doing research, etc.

BioDialog Activities

Assuit Uni. on November 2015

AinShams Uni. on September 206

Kick-off Meeting March (2016)

Sfax Uni. on November 2016.

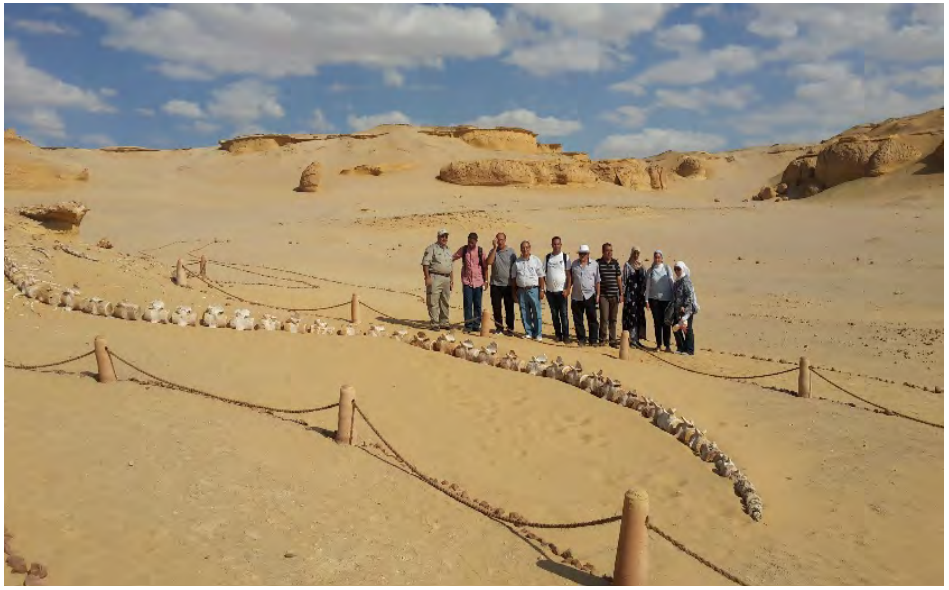
BioDialog Activities

- Staff training in Germany.
- Ph.D. and M.Sc. Students Studying at Jena Uni. for one Semester: Two Semesters (2015-2016), (2016-2017)

- Sharing Projects Development.
- Get Training on Intercultural Communication Skills.

BioDialog Activities (Cont…)

- Visiting Biodiversity Experiments & Reserves: Jena, Wale Valley, and Petrified Forest, etc.



BIODIALOG UPCOMING EVENT

YOU ARE INVITED

Motivation ...

There is a growing importance of data particularly obvious in biodiversity research. This discipline studies the totality and variability of organisms, their morphology and genetics, life history and habitats, and geographical ranges. Biodiversity Informatics is the application of informatics techniques to biodiversity data for improved management, presentation, discovery, exploration, and analysis. Biodiversity informatics is a relatively new discipline extending computer science in the context of biodiversity data. For the success of any biodiversity informatics project, close collaboration with scientists from other domains, including Biology, Ecology, Earth Sciences, and/or Economics is crucial. This communication poses significant intercultural challenges. This is why our summerschool provides training on aspects of biodiversity, scientific data management and intercultural communication.

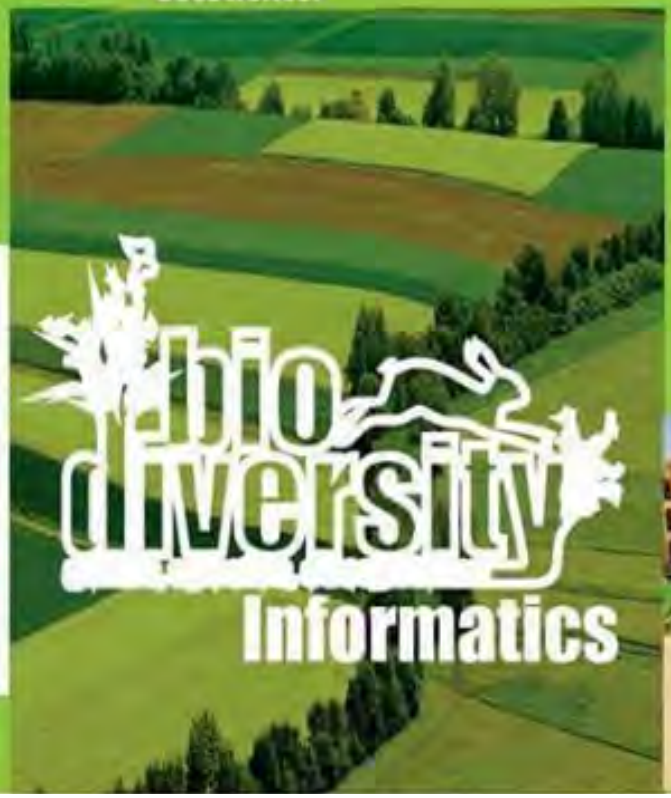
http://fusion.cs.uni-jena.de/bdg/

Hochschuldialog mit der islamischen Welt

Intended Audience

Scientists from :

- Computer Engineering ,
- Computer Science,
- Ecology,
- Biology,
- Agriculture,
- Geoscience.



http://fusion.cs.uni-jena.de/bdg/

DAAD BioDialog Project

The First Biodiversity Informatics Summer School Hurghada, Egypt August 22-28, 2017

Do You Want To ...

Obtain theoretical knowledge about **scientific data management?**

Acquire hands-on experience with a state of the art **biodiversity data management system?**

Learn how to submit your data to **GBIF and make it internationally visible?**

Better understand the possibilities of **Big Data Analytics?**

Improve your intercultural **communication skills?**

Important Dates

Submission deadline
June 01, 2017
Notification date
June 10, 2017
Summer school date
August 22-28, 2017
Place
Hurghada, Egypt

How to Apply

https://goo.gl/l7445K

Costs

Attendance is free of charge, but participants are required to pay for their own travel and accommodation.
Limited financial support for these costs will be available.
For more details, please refer to the project website.

Contact

M.SC. Dina Sharafeldeen
dina.sharafeldeen@uni-jena.de

Sponsored by DELL EMC

Organizing Team

- Prof. Dr. Taysir Hassan, Assiut University
- Dr. Ibrahim Moawad, Ainshams University
- Dr. Bassem Bouaziz, ISIMS -Sfax University
- Dr. Rania Elgohary, Ainshams University
- Dr. Ing. Alsayed Algergawy, FSU - Jena
- Prof. Dr. Birgitta König-Ries, FSU - Jena

Past Events

- 2015 Workshop
Jena University, Germany

- 2015 Workshop
Assiuit University, Egypt

- 2016 Summer visit
Jena University, Germany

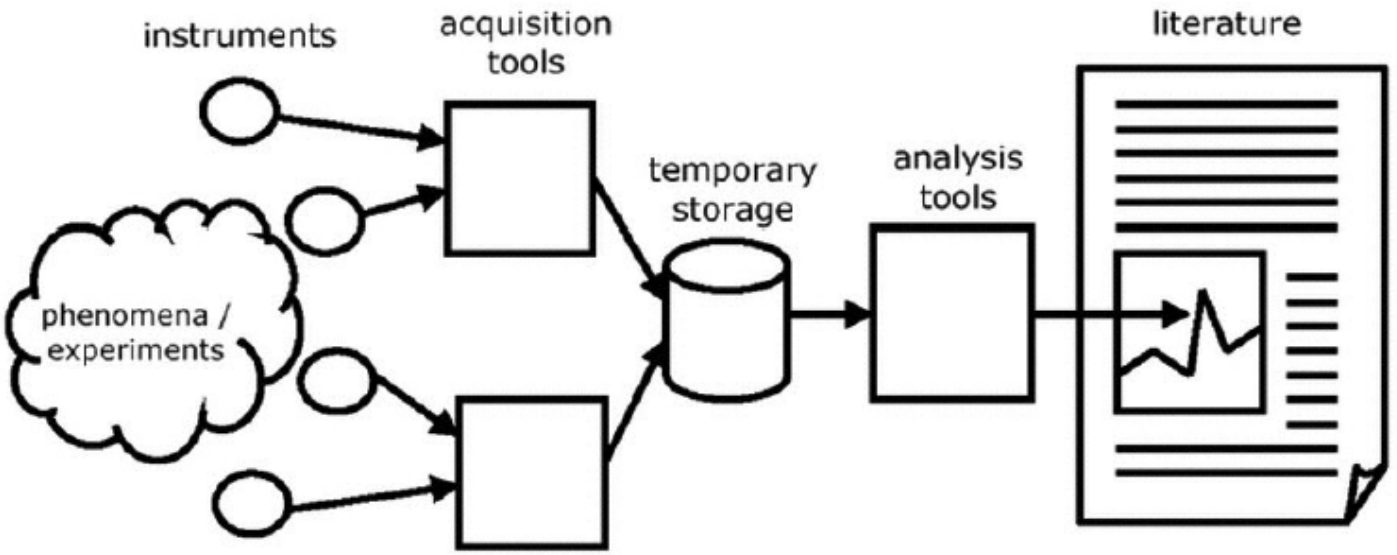
- 2016 Summer training
Ain-shams University, Egypt

- 2016 Workshop
University of Sfax, Tunisia

Scientific Biodiversity Data Management

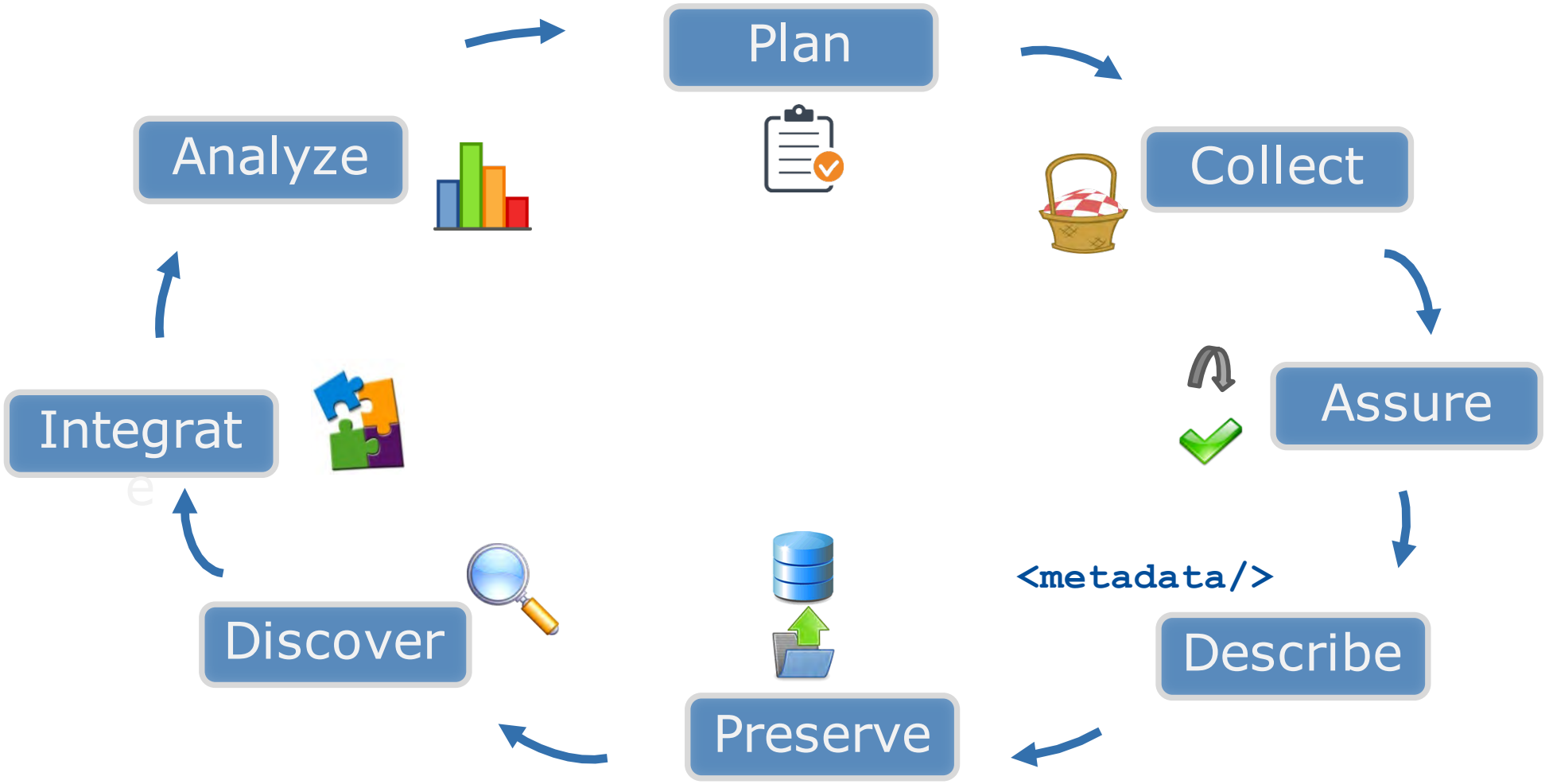
Scientific Biodiversity Data Management

- Data is Gold especially the Scientific Biodiversity Data.
  - Do we manage data well?
  - Do we plan how to collect data?
  - Do we have data with high quality?
  - Do we describe data?
  - Do we preserve our data for later use even by ourselves?
  - The answer NO.
  - Why? Because we follow the publish & forget model

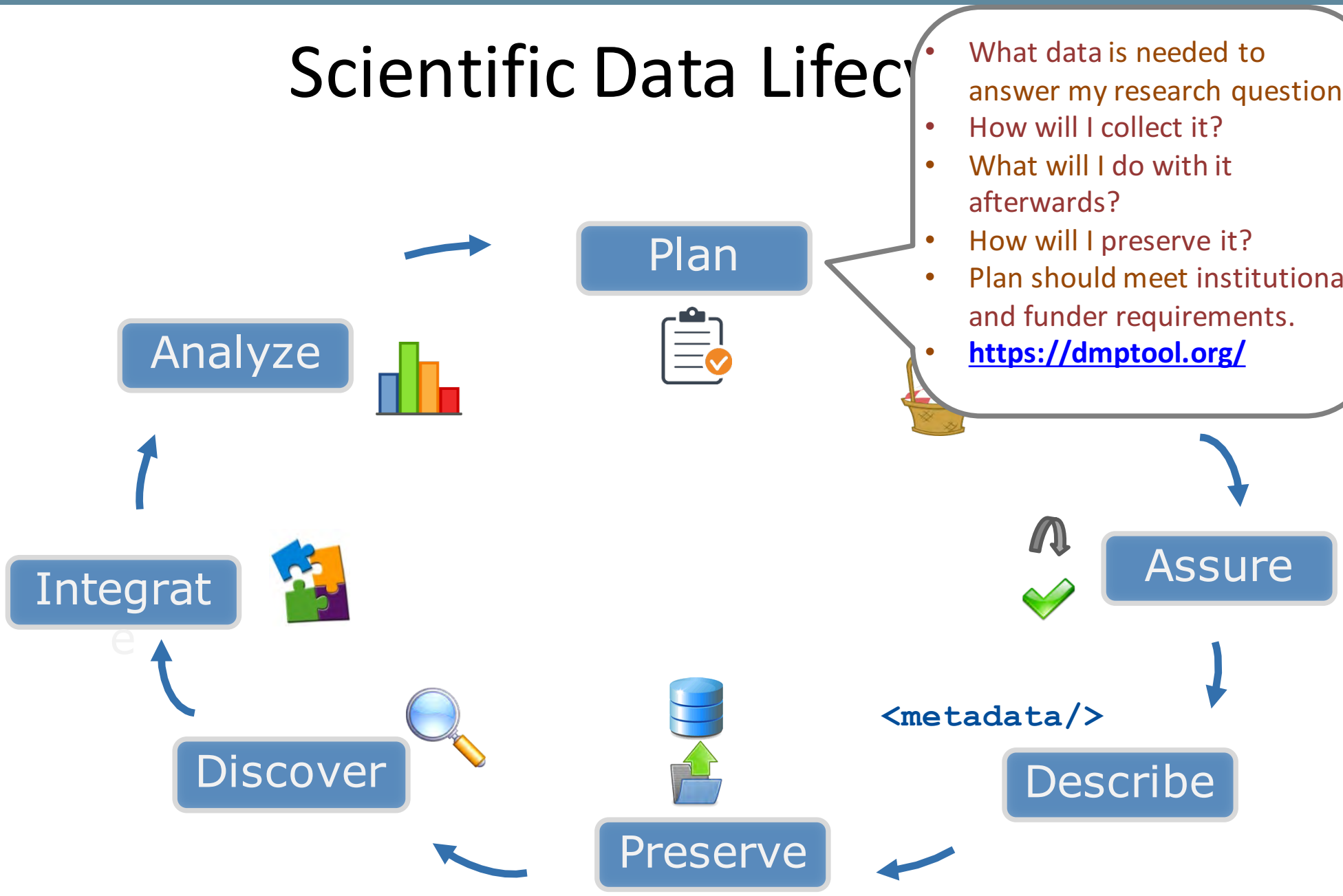

instruments
acquisition tools
literature
temporary storage
analysis tools
phenomena / experiments


Plan
Collect
Assure
<metadata/>
Describe
Preserve
Discover
Integrat
Analyze

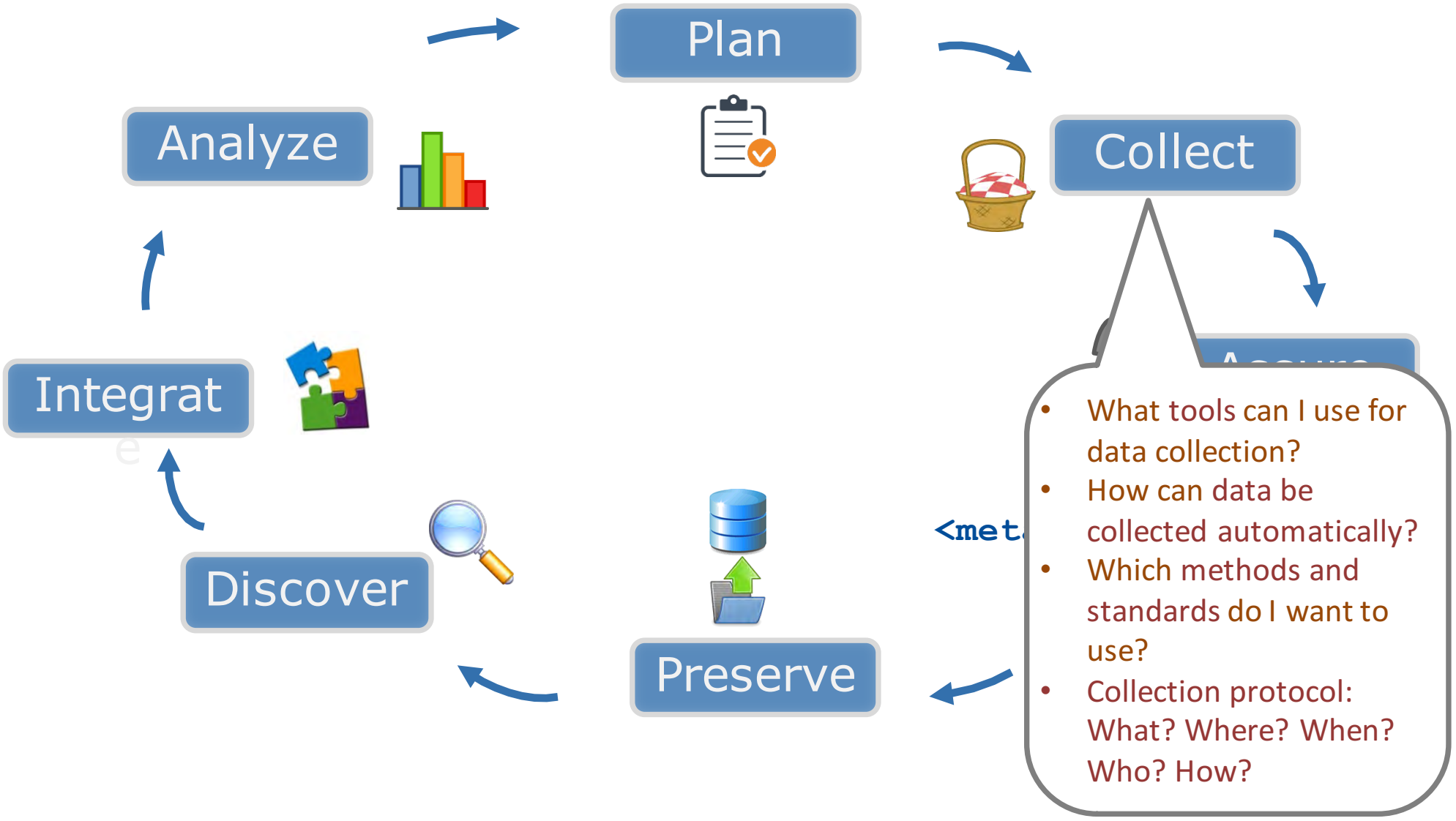
Scientific Data Lifecycle is a conceptual tool which helps to understand the different steps that data follow from data generation to knowledge creation



Scientific Data Lifecy



Scientific Data Lifecycle



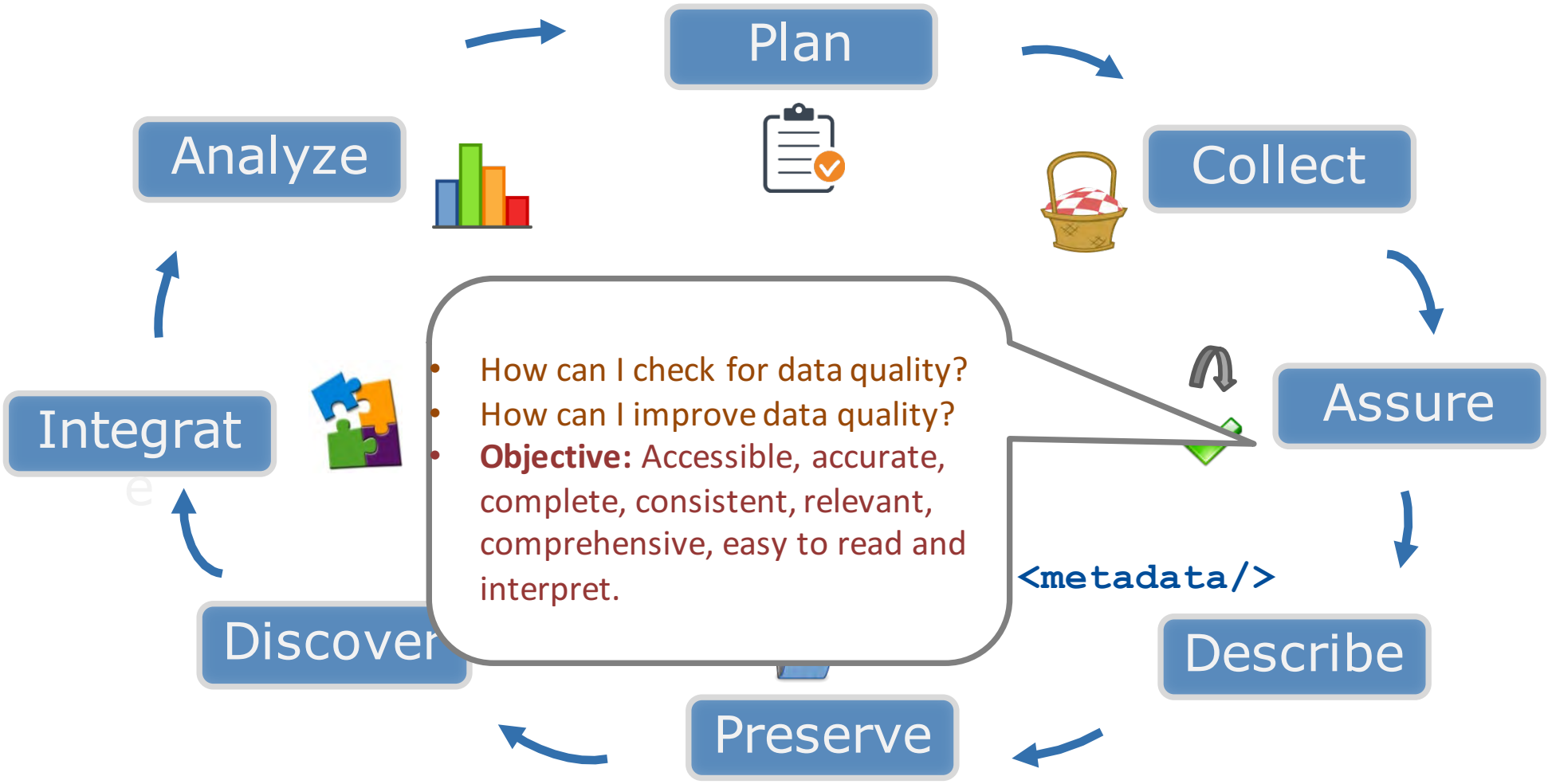
Biodiversity data collection

Humans
analog & manual

Sensors
digital & autonomous

Crowds + Sensors
digital & manual

Scientific Data Lifecycle



Scientific Data Lifecycle - Assure

- Just Error Examples

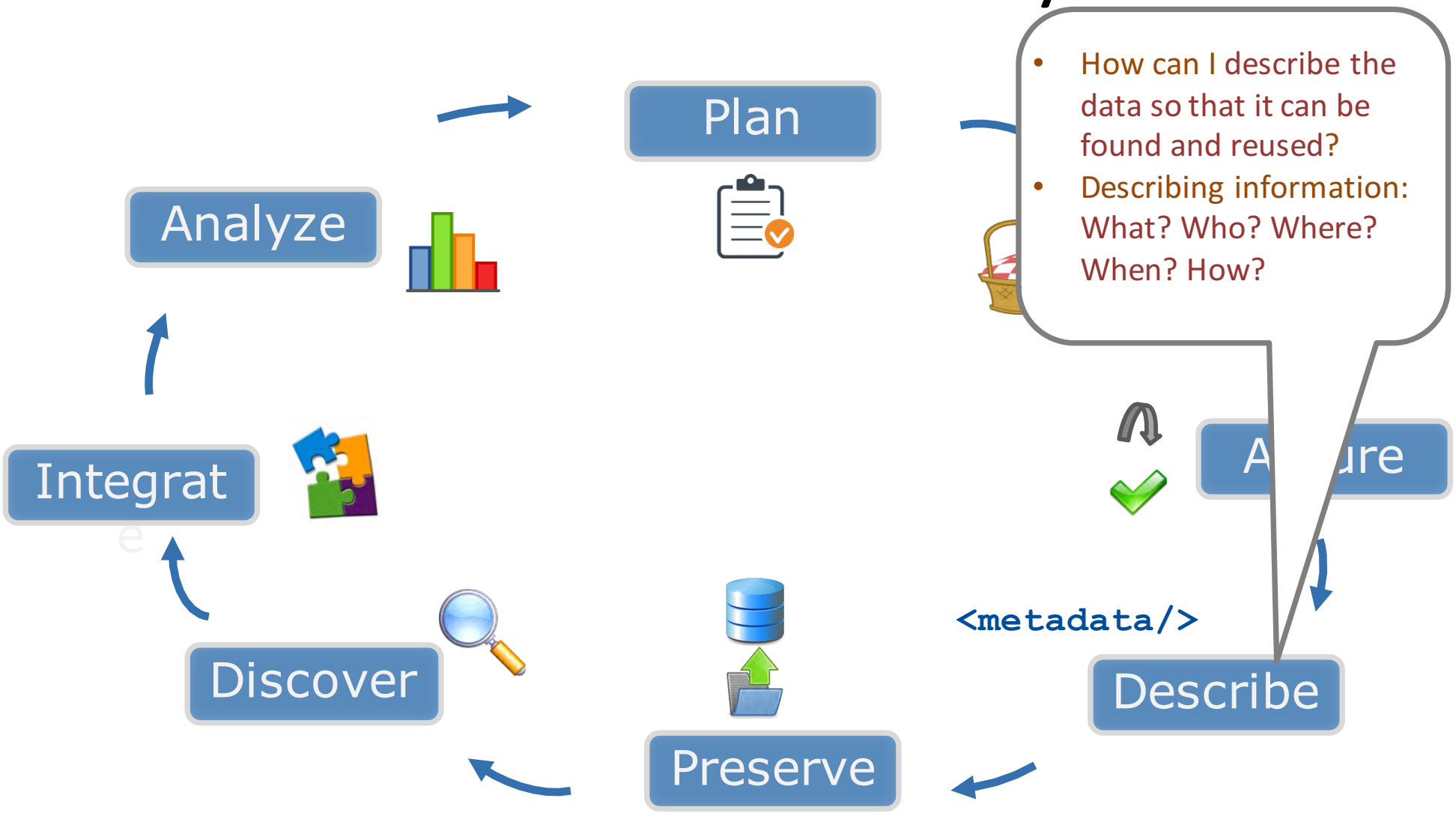
data.xls

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Site	Date	Plot	Species	Weight	Acult		Rodent Trapping 3/15/2010						
2	DeepWell	2/13/2010	1	DIPO	12.1	j		Site	Plot	Adult	RodentSp	Weight		
3	Deep Well	Feb-10	2	Pero	13.22	j		DW	1	y	Pero	12		
4	rioSalado	2/13/2010	1a	pero	16	N		RS	2	j	PERO	escaped <15		
5	rioSlado	"	1"	CleGap	18.92	got away		RS	3	n	Clegap	91		
6				Mean1	15.06									
7														
8														
9														
10														
11														
12	Rodent Trapping		MJK & ALN	10-Apr-10										
13	Site	Plot	Adult	Species	grams	Ccmments								
14	deep well	1	y	woodrat	13									
15	riosalado	2	y	PERO	24.5									
16	riosalado	3	y	Clegap	91									
17														
18														
19														

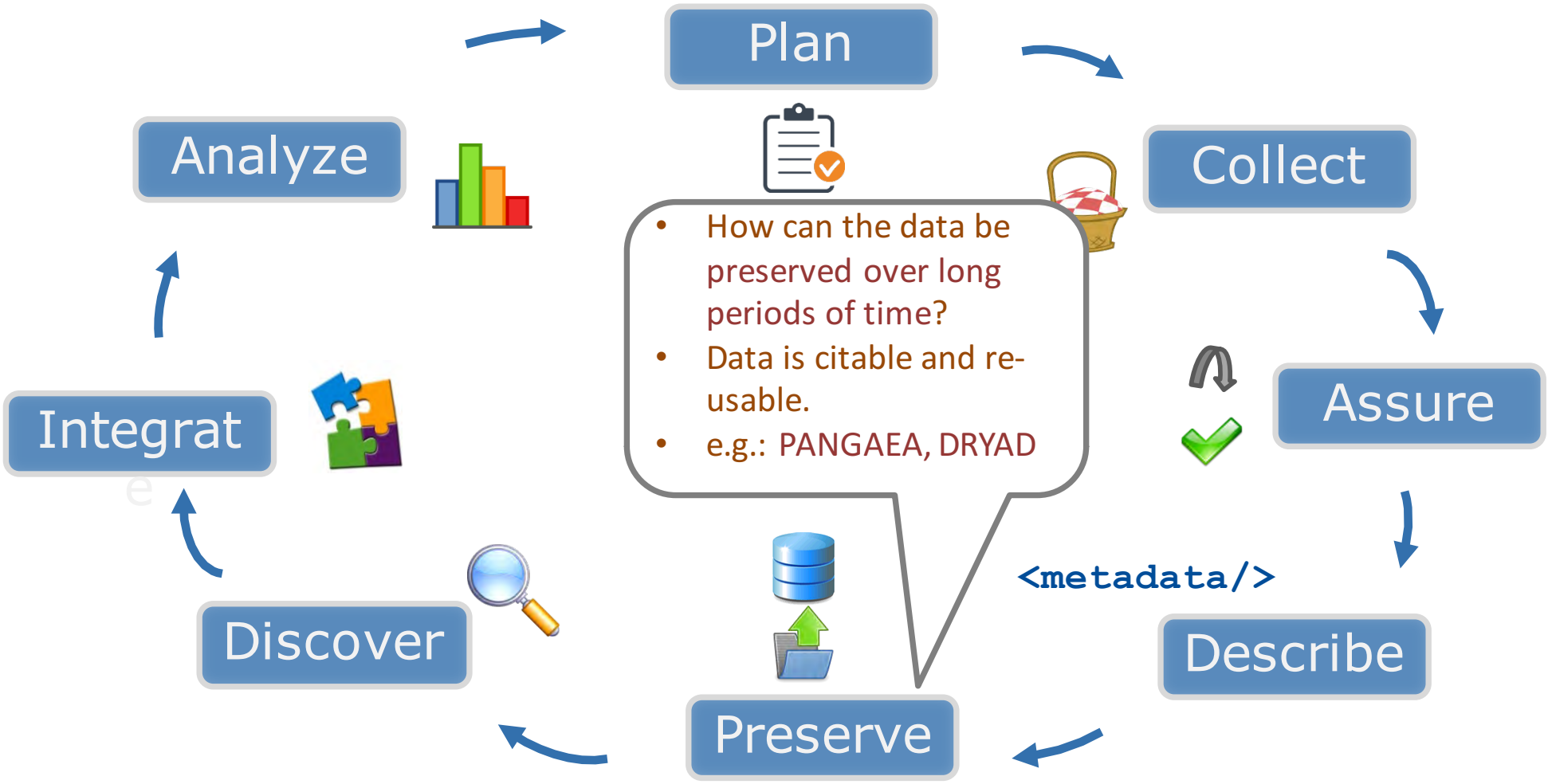
Sheet1

- Inconsistent data format
- Column names
- Order of columns
- Different spelling, capitalization,
- Spaces in site names
- Code used for sites names but spelled out for others
- Text and numbers in same column

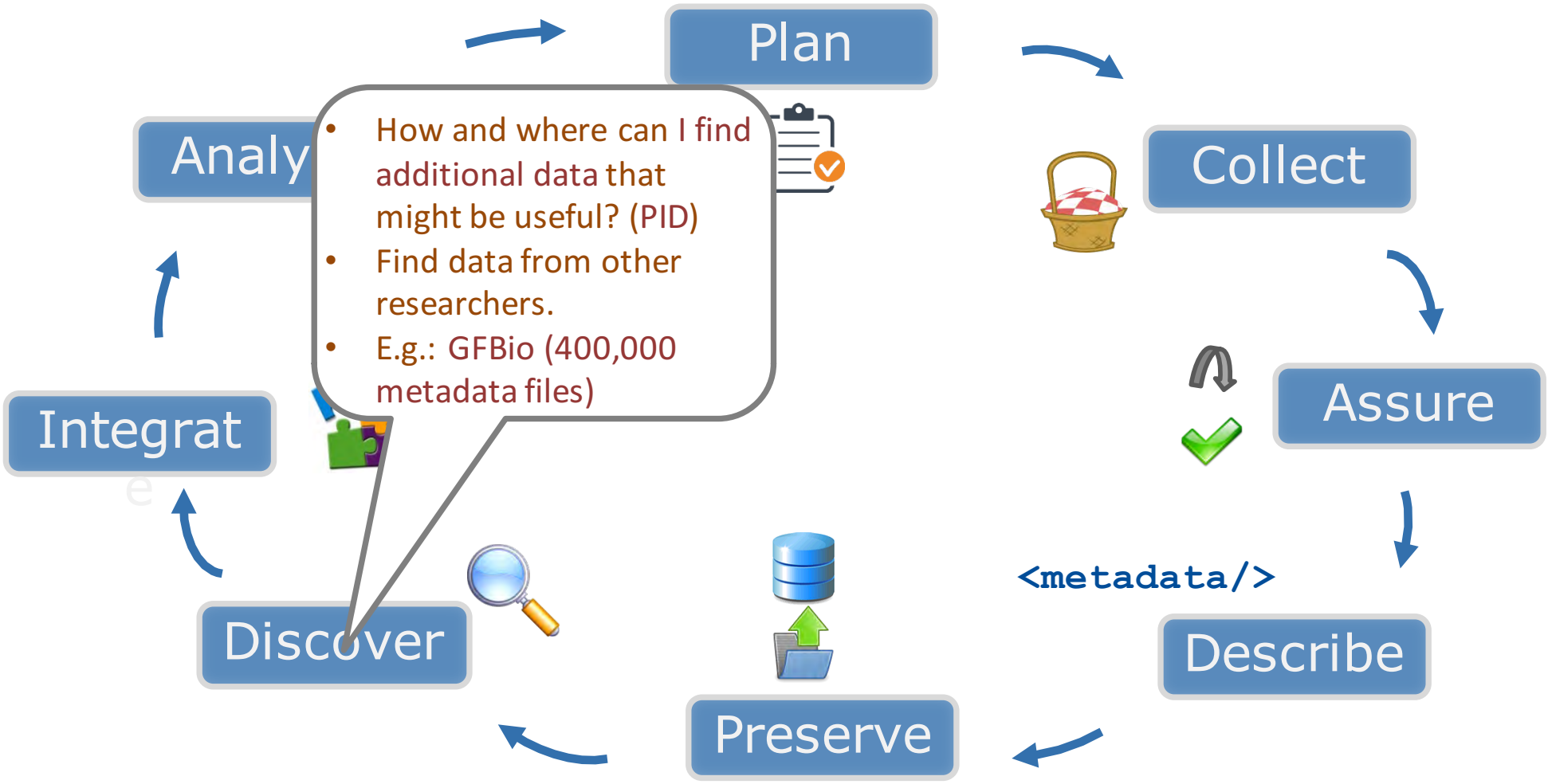
Scientific Data Lifecycle



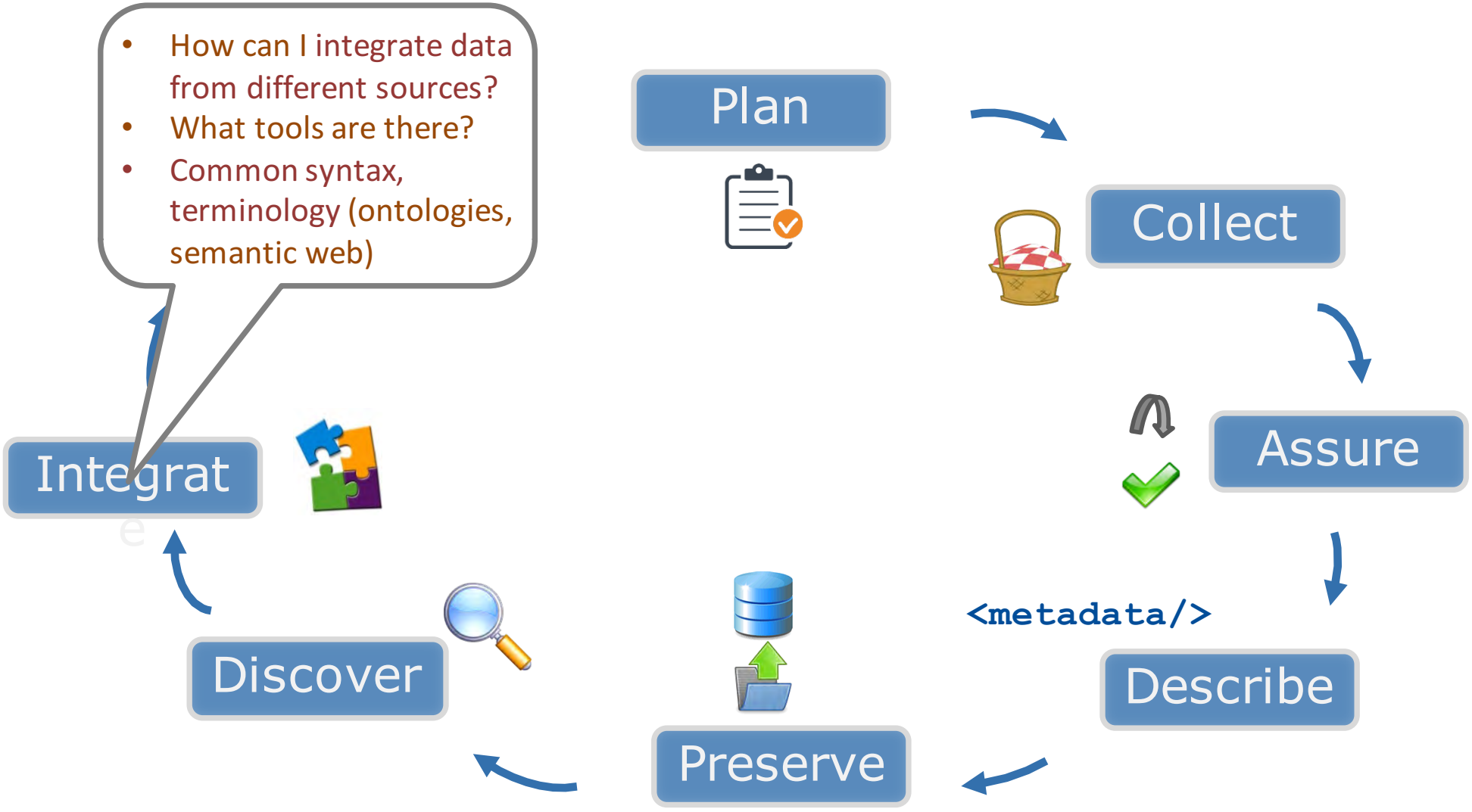
Scientific Data Lifecycle



Scientific Data Lifecycle

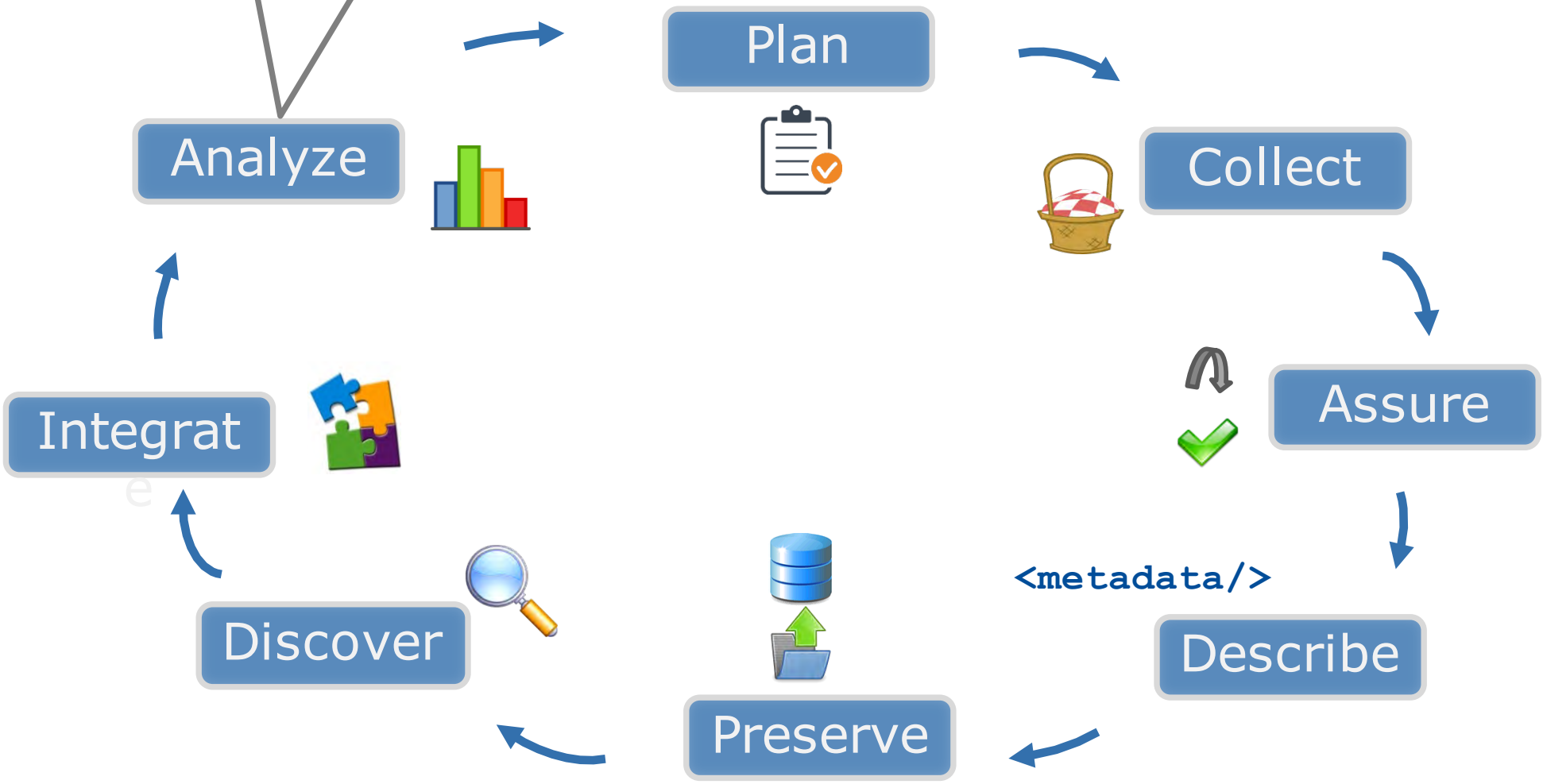


Scientific Data Lifecycle



Scientific Data Lifecycle

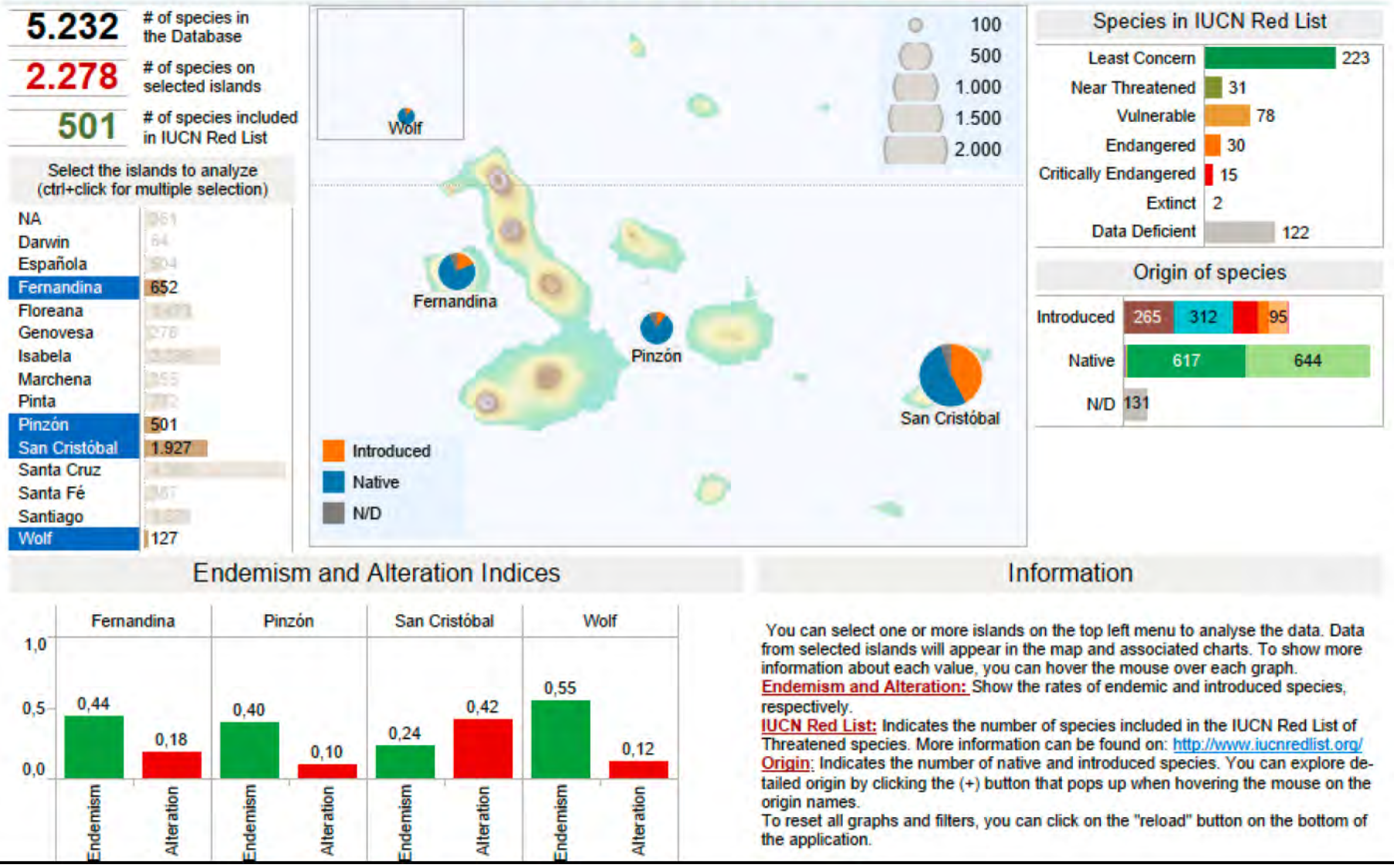
- How can I analyze data?
- What tools are there?
- What should I keep in mind?



Scientific Data Lifecycle

- Statistical analysis through visualization
- Analytical technics: machine learning, text mining, data mining, regression analysis
- Tools:

VISUAL ANALYTICS



www.galapagos.org/conservation/galapagos-climate-change-initiative

The BEXIS System

Scientific Data Management System

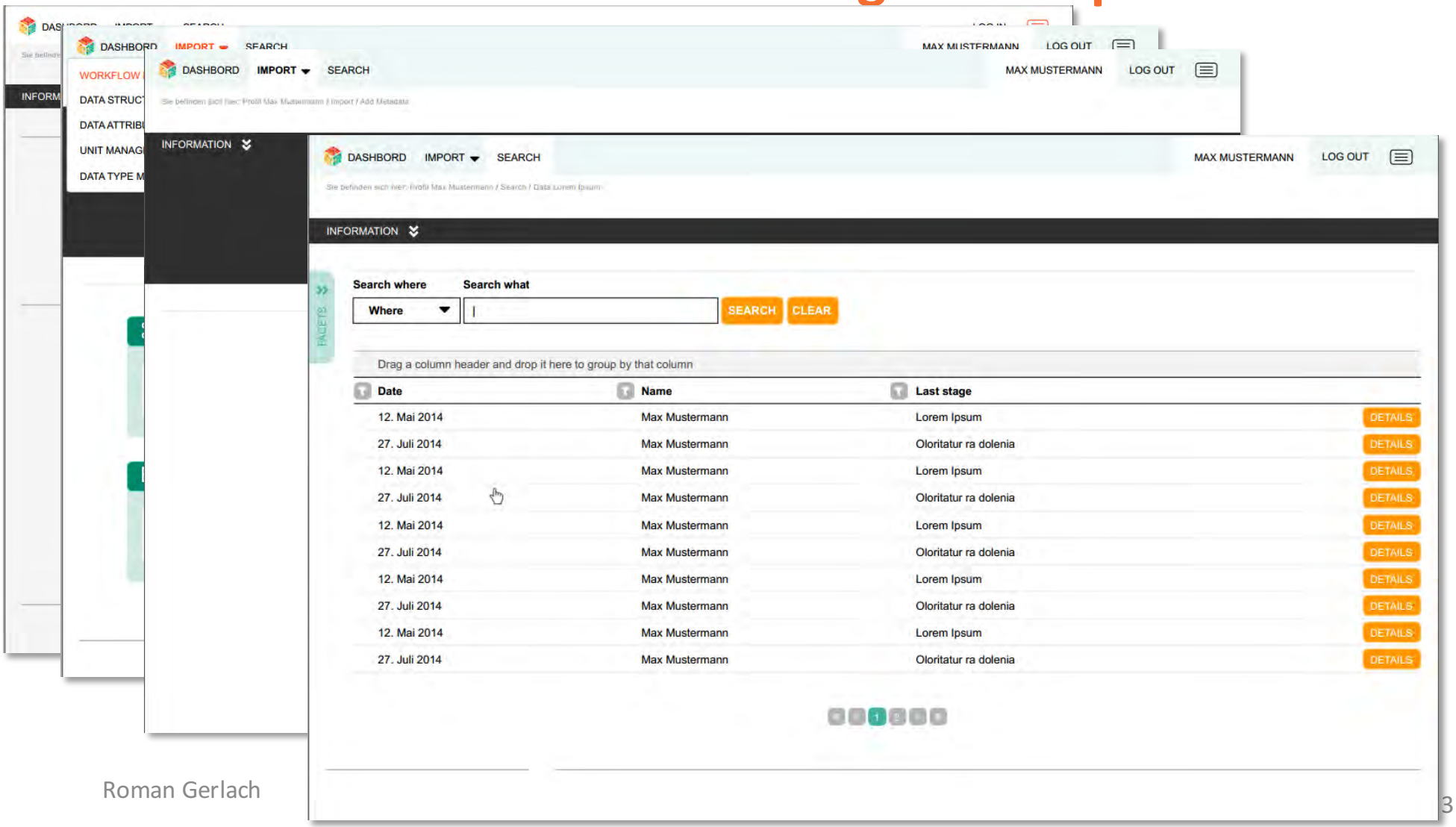
Key Features

- Modular, flexible, scalable, free and open source
- Focus on tabular data, but not limited to.
- Focus on data integration and re-use.
- User defined data structures.
- Re-use of data structures and data attributes
- Dataset versioning & views

Currently, we are hosting BEXIS on EUN to be Exploited by the Egyptian Universities.

Usability, Usability, Usability

New user interface following user requests



DAAD

Thank You

BiODialog