WHY AND HOW TO SCORE TAXONOMIC DATA IN AN INTERNATIONALLY RECOGNIZED FORMAT ?

The ABCD Schema of the TDWG as an example

museum

Topics

- What are taxonomic data ?
- What is the minimum quantity and the minimum quality of taxonomic data that must be captured?

in the field / in the lab / in the permanent voucher collection / in a local database

- Why and how to score taxonomic data in an internationally format : the ABCD Schema
- Practical implementations of the ABCD Schema
- Conclusions

Topics

- **What are taxonomic data ?**
- What is the minimum quantity and the minimum quality of taxonomic data that must be captured?

in the field / in the lab / in the permanent voucher collection / in a local database

- Why and how to score taxonomic data in an internationally format : the ABCD Schema
- Practical implementations of the ABCD Schema
- Conclusions

taxonomic data

Entire history of a collected specimen :

- Collecting data
- Identification data
- Extra information (details)

Collecting data

- Who : collector’s name
- Where : locality description (field number, station number, …) When recording data in the field, whether from a map or when using a GPS, it is important to record locality information as well as the geo-references, so that later validation can take place if necessary
- When : collecting data / period (start date, end date)
- How : collecting method

Identification data

Who ; Where ; When ; How +

What :

- (Scientific) name (taxonomy data)
- Authority => name and date of descriptor(s)

Some scientists assume that they are the only people able fully to understand the data that they have collected !

But what will happen if those scientists would be kidnapped by aliens ?

Topics

- What are taxonomic data ?
- **What is the minimum quantity and the minimum quality of taxonomic data that must be captured?**

in the field / in the lab / in the permanent voucher collection / in a local database

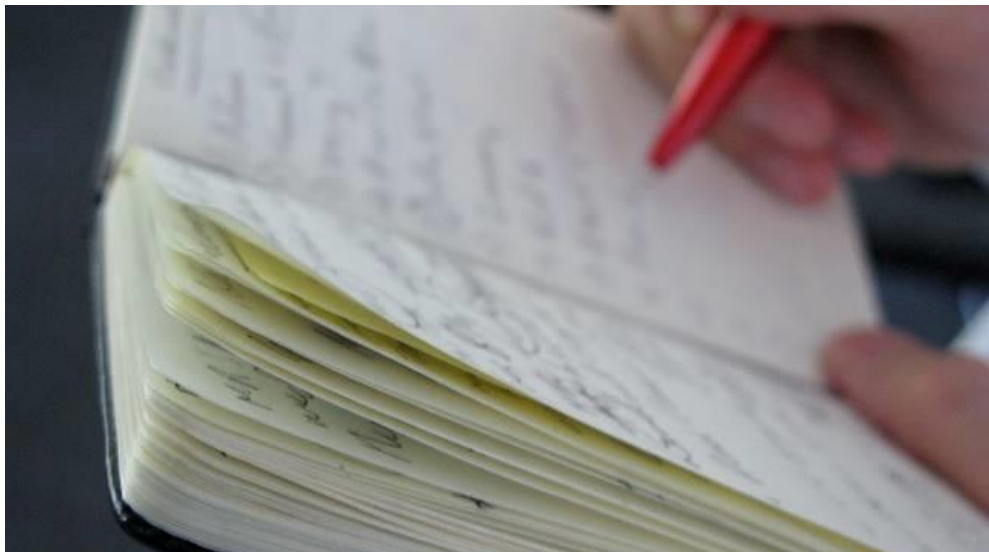
- Why and how to score taxonomic data in an internationally format : the ABCD Schema
- Practical implementations of the ABCD Schema
- Conclusions

In the field : field notes / logbook

Field notes contains scientific observations : locations, date and time, physical characteristics, various observations (weather, water temperature, …) sketches, ethology, …

- It’s a draft of the future labels
- **Minimum quantity** : who, where, when, how
- Field notes is a **back up** of your observations
- **Minimum quality** :

  Legible, durable
- **<u>Never underestimate it</u>**



Example : field book

EATSA Project RSAKZN 2016							
Date:	Time:	Latitude:		Location:			Duration of sampling:
Name of collector(s):		Longitude:		Sampling methodology:			Sampling time:
Field number	**Preliminary Taxonomy**	**Specimen count**	**Min-max depth**	**Substrate type**	**Picture no**	**DNA sample**	**Comments**

In the lab

Transposition of the data to labels

- It depends on your field book
- **Minimum quantity** = who, where, when, how
- **Minimum quality** = Primary labels : some specimens could stay hundred years before being identified

Durable labels

Good quality of paper, good ink (pencil), correct information

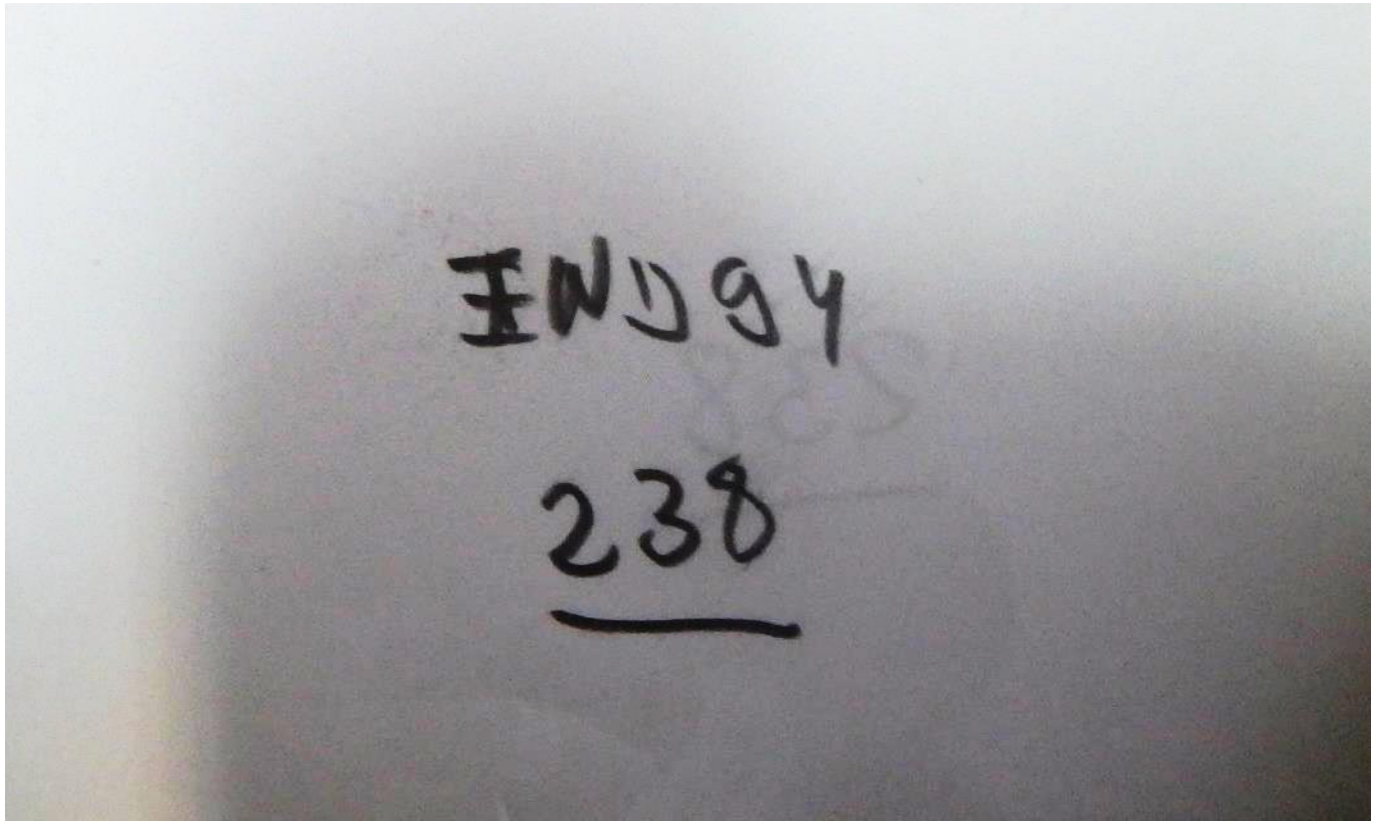
Capital letters

In the permanent voucher collection

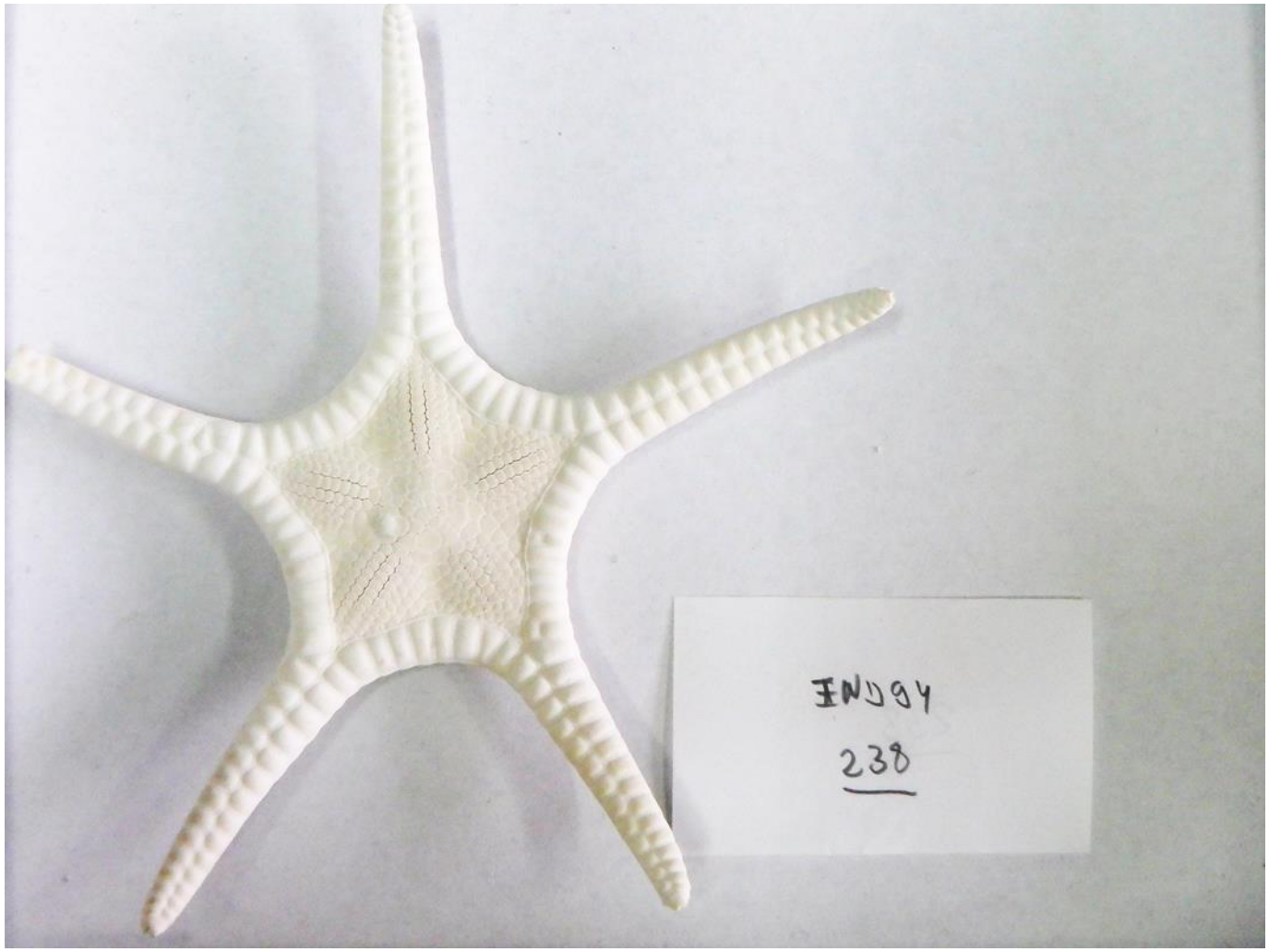
- Transposition of the data on a durable label :
- It depends on the primary label
- Minimum quantity : the same as above
- Minimum quality : durable means good quality of paper, ink…

Sometimes it's like a police investigation
Especially when the scientist who had participated on the expedition is not there anymore

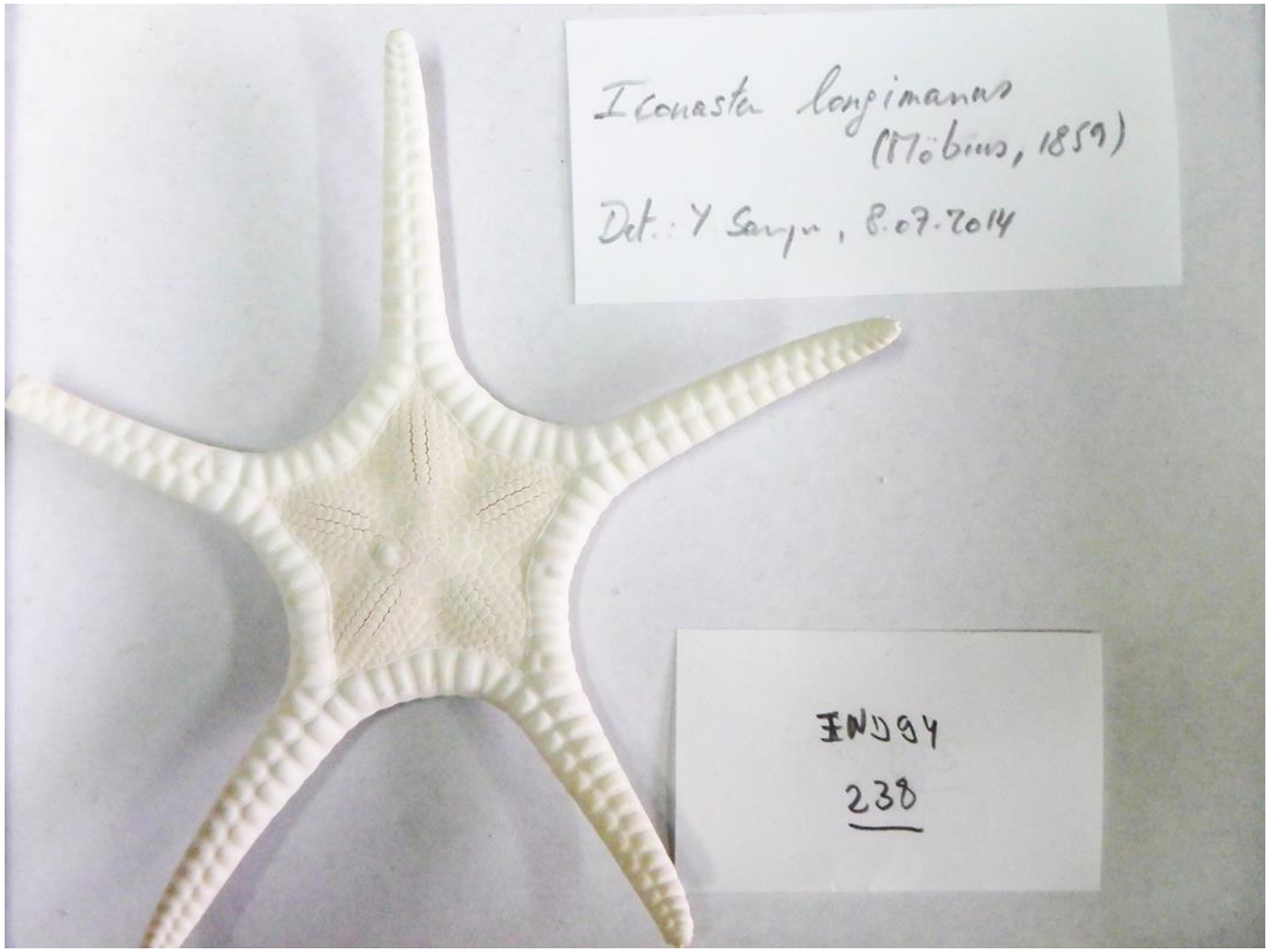
Minimum quantity / quality ?



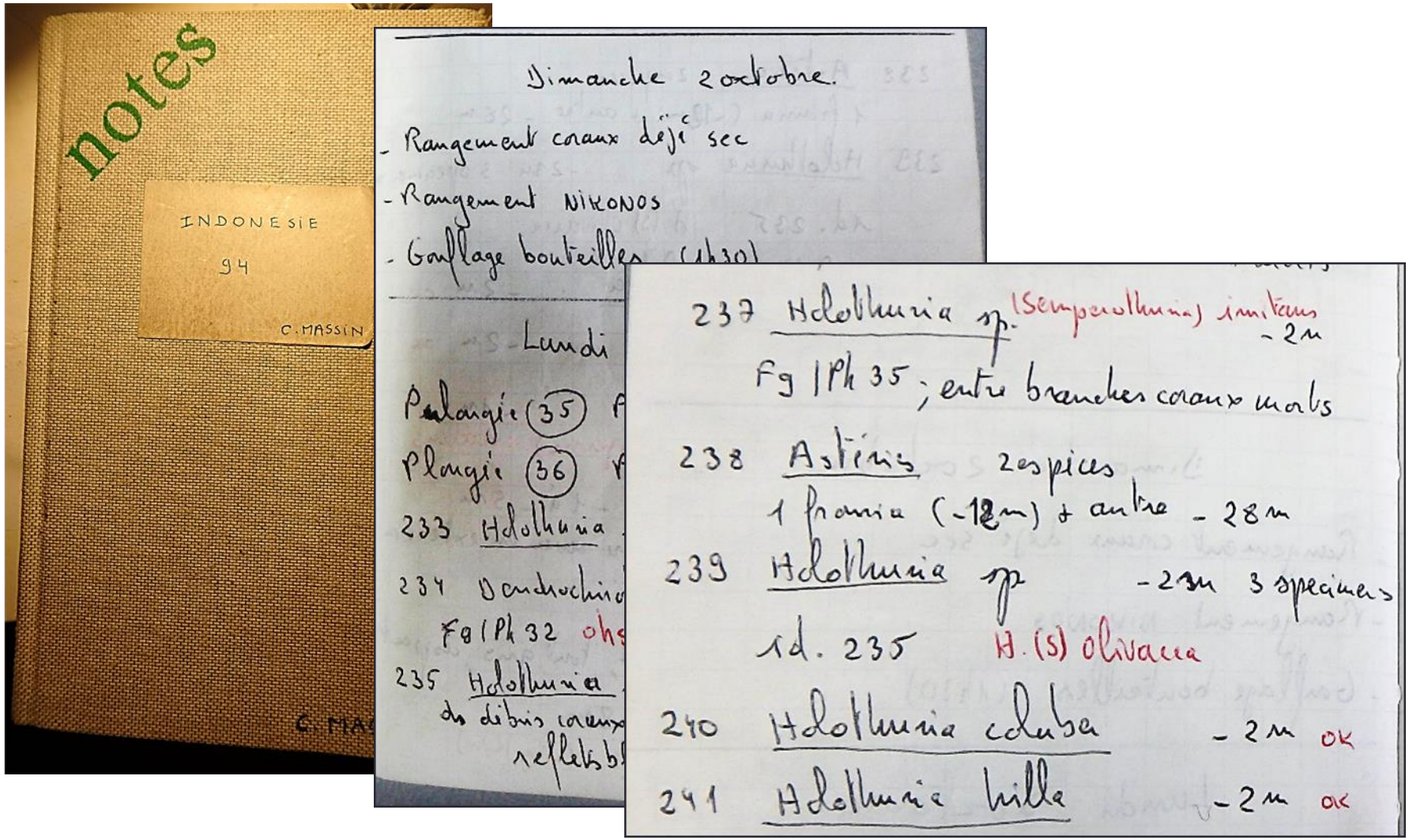
Labeling in practice



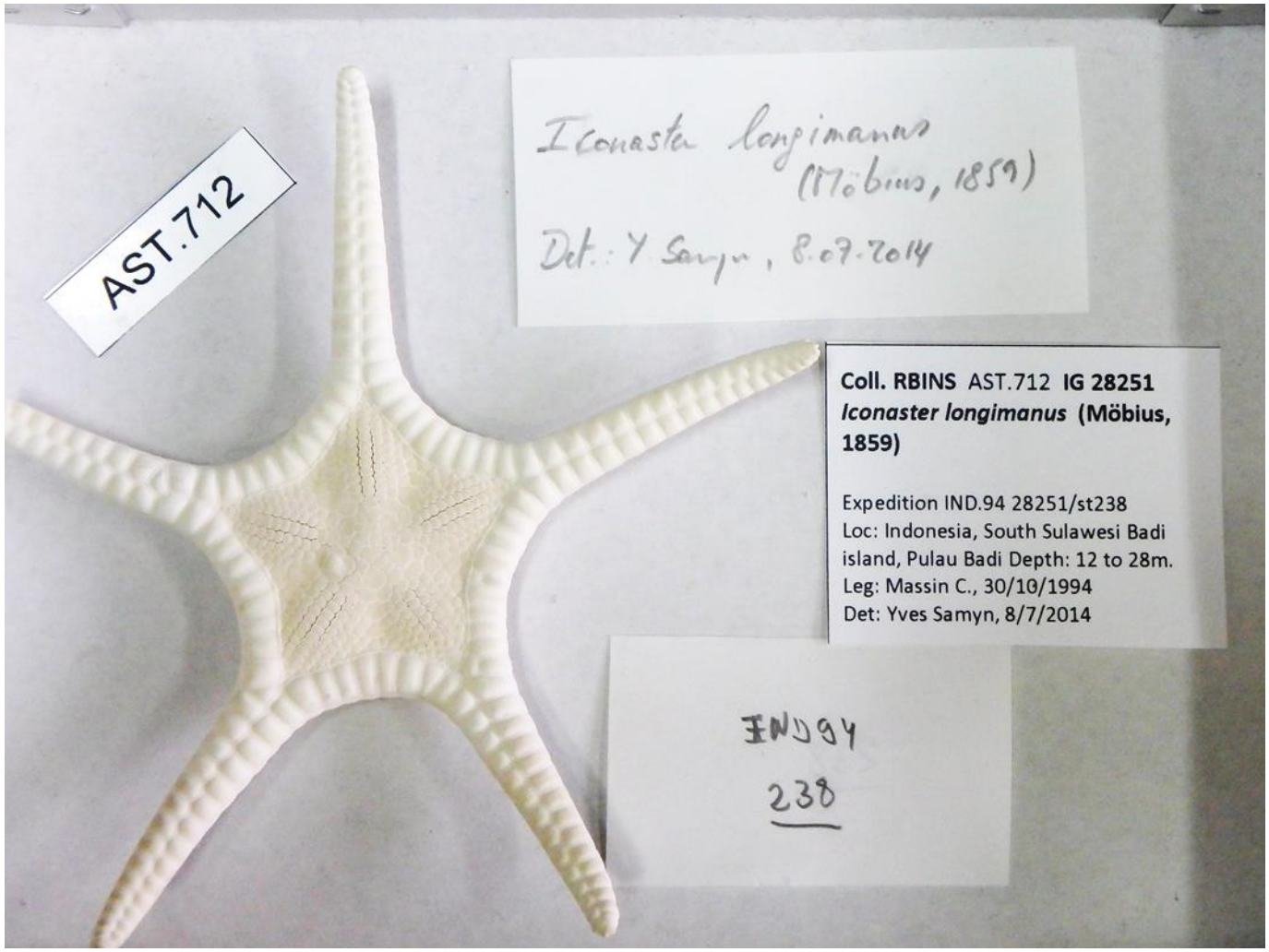
Labeling in practice



Labeling in practice



Labeling in practice



Labeling in practice

Coll. RBINS AST.712 **IG 28251**

Iconaster longimanus **(Möbius, 1859)**

Expedition IND.94 28251/st238
Loc: Indonesia, South Sulawesi Badi island, Pulau Badi Depth: 12 to 28m.
Leg: Massin C., 30/10/1994
Det: Yves Samyn, 8/7/2014

In a local database

Each database are home-made and different :

a simple excel-sheet -> sophisticated database

Transposition of data from the labels to computers :

Minimum quantity :

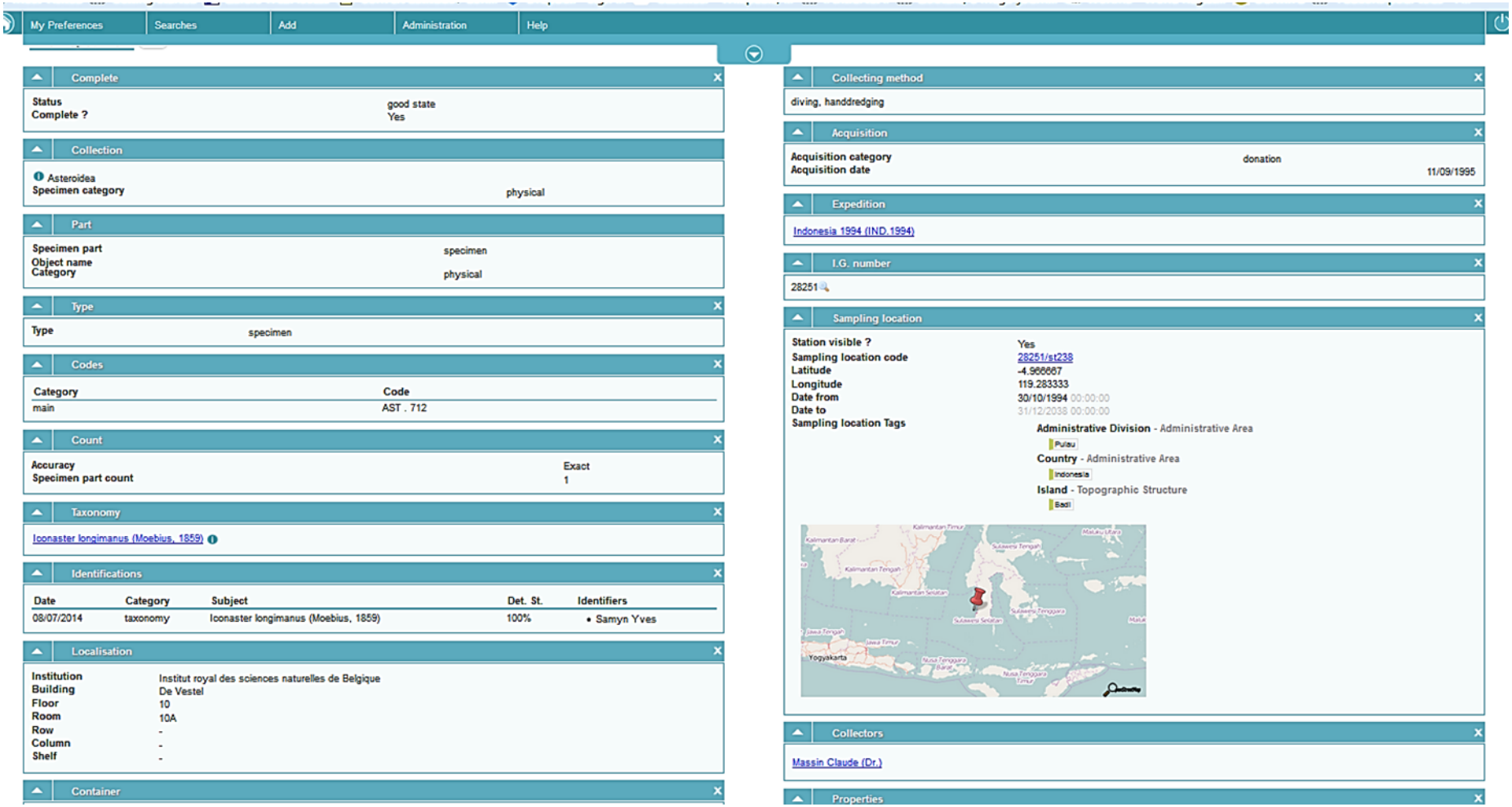
who, where, when, how

+ what, where (location inside the collection)

Minimum quality :

usable in time

Darwin : database of RBINS/RMCA



Importance of the minimum quality/quantity data ?

Most expeditions are extremely expensive

Sometimes specimens can no longer be found anymore (loss of habitat, extinction of the species…)

Each voucher specimen optimally becomes a permanent occurrence record

Topics

- What are taxonomic data ?
- What is the minimum quantity and the minimum quality of taxonomic data that must be captured?

in the field / in the lab / in the permanent voucher collection / in a local database

- **Why and how to score taxonomic data in an internationally format : the ABCD Schema**
- Practical implementations of the ABCD Schema
- Conclusions

Why ?

- The scientific impact of the inventory clearly depends on the **cooperation** between participants and the sharing of the data collected
- **Every day** probably more than 100,000 scientific biological records (observations, collected specimens) are recorded (cf Abc Taxa vol 8)
- **Today, there is no single global inventory or directory of just all known species available (even in protected areas)**
- Making available a vast amount of crossed information

How ?

One standard to allow inter-compatibility and ease of data consultation:

ABCD (& ABCD-DNA & ABCDEFG) Schema

ABCD

Access to **B**iological **C**ollection **D**ata

- **accessibility** of existing **biological collection data** banks at the international level by developing and maintaining a comprehensive and commented **schema** for biological collection records (ABCD Schema).
- ABCD Schema is a standard for the **access** to and **exchange** of data.
- In the process, it promotes **standardization** of the terminology used to model biological collection information and provides a **general format** for data exchange and retrieval for biological collections.

- ABCD is a common data specification for biological collection units, including living and preserved specimens
- The design goal of the data specification is to be both **comprehensive and general**
- ABCD has been developed by **TDWG** and is supported by GBIF

TDWG

Taxonomic **D**atabases **W**orking **G**roup (1995)

=> Biodiversity Information Standards

www.tdwg.org

- a non profit scientific and educational association.
- The organization was formed to establish an **international collaboration** to promote the wider and more effective **dissemination of information** about the world's heritage of biological organisms
- It aims to provide an international **forum** for biological data projects
- It should facilitate data exchange

Biodiversity Information Standards

TDWG

GBIF

Global **B**iological **I**nformation **F**acility

is a mega-science project with the aim to make the world's primary data on biodiversity freely and universally available via the Internet = network

Free and Open Access to Biodiversity Data

It should sharing biodiversity data

It might provide evidence for research

and decisions

www.gbif.org

ABCD Schema in practice

- **comprehensive** and highly structured, supporting data from a wide variety of databases.
- **compatible** with several existing data standards.
- in use with the **GBIF**
- **XML** data

ABCD Schema

ABCD Biological collection

ABCD DNA

ABCDEFG Extended For Geosciences

700 elements for the Biological collection

76 elements for the DNA section

504 elements for the EFG section

———

1280 elements

RBINS/RMCA use a simplified ABCD with 160 elements

Topics

- What are taxonomic data ?
- What is the minimum quantity and the minimum quality of taxonomic data that must be captured?

in the field / in the lab / in the permanent voucher collection / in a local database

- Why and how to score taxonomic data in an internationally format : the ABCD Schema
- **Practical implementations of the ABCD Schema**
- Conclusions

« our » ABCD Schema

IDENTIFICATION				ACCESSION				ACQUISITION		
specimenI	additionalI	datasetNa	isPhysical	accession	creationDa	creationMo	creationYe	acquisition	acquiredFr	acquisition
AST.625			yes	26080				Gift		11
AST.626			yes	28251				Gift		11
AST.627			yes	28251				Gift		11
AST.628			yes	25715				Gift		20
AST.629			yes	26086				Gift		12
AST.630			yes	26086				Gift		12
AST.631			yes	26373				Gift		10
AST.632			yes	25715				Gift		20
AST.633			yes	26080				Gift		11
AST.634			yes	26086				Gift		12
AST.635			yes	26373				Gift		10
AST.636			yes	26700				Gift		19

« our » ABCD Schema

IDENTIFICATION	specimenID	AST.625
	additionalID	
	datasetName	
	isPhysical	yes
ACCESSION	accessionNumber	26080
	creationDay	
	creationMonth	
	creationYear	
ACQUISITION	acquisitionType	Gift
	acquiredFrom	
	acquisitionDay	11
	acquisitionMonth	2
	acquisitionYear	1980

« our » ABCD Schema

TAXONOMY	phylum	Echinodermata
	classis	Asteroidea
	ordo	Valvatida
	superfamilia	
	familia	Ophidiasteridae
	subfamilia	
	genus	Gomophia
	subgenus	
	species	egyptiaca
	author_year	Gray, 1840
	subspecies	
	variety_form	
	taxonFullName	
	oldGenus	
	oldSubgenus	
	identifiedBy	Yves Samyn
	identificationDay	10
	identificationMonth	1
	identificationYear	2014
	identificationNotes	
	referenceString	
	publicationString	
	urlPicture	

« our » ABCD Schema

COLLECTING EVENT	samplingCode	26080/st.49
	ocean	
	continent	
	sea	
	country	Papua New Guinea
	state_territory	
	province	Madang Province
	region	Milliat
	archipelago	
	district	
	county	
	department	
	island	
	city	
	municipality	
	populatedPlace	
	naturalSite	
	exactSite	large reef at East side of Wongat

« our » ABCD Schema

COLLECTING EVENT	elevationInMeters	
	depthInMeters	3 to 5 m
	latitude	
	longitude	
	ecology	in and under coral
	samplingMethod	handdredging
	expedition_project	Papua New Guinea 1979
	collectedBy	Pierret Jean
	collectionStartDay	
	collectionStartMonth	3
	collectionStartYear	1979
	collectionStartTimeH	
	collectionStartTimeM	
	collectionEndDay	
	collectionEndMonth	3
	collectionEndYear	1979
	collectionEndTimeH	
	collectionEndTimeM	
	localityNotes	

« our » ABCD Schema

PROPERTIES	kindOfUnit	Animal, specimen
	sex	
	lifeStage	
	statusType	Holotype, non type
	socialStatus	
	totalNumber	
	adultCount	
	larvaCount	
	pupaCount	
	juvenileCount	
	immatureCount	
	stageUnknownCount	
	maleCount	
	femaleCount	
	sexUnknownCount	
	protonymphCount	
	deutonymphCount	
	tritonymphCount	
	hostName	

« our » ABCD Schema

ASSOCIATION	associatedUnitInstitution	
	associatedUnitCollection	
	associatedUnitID	
	associationType	
STORAGE	fixation	formaldéhyde
	conservation	dry
	institutionStorage	Royal Belgian Institute of Natural Sciences
	buildingStorage	De Vestel
	floorStorage	10
	roomStorage	10A
	laneStorage	16
	columnStorage	F
	shelfStorage	4
	container	
	containerType	minigrip
	containerStorage	dry
	subcontainer	
	subcontainerType	
	subcontainerStorage	

Topics

- What are taxonomic data ?
- What is the minimum quantity and the minimum quality of taxonomic data that must be captured?

in the field / in the lab / in the permanent voucher collection / in a local database

- Why and how to score taxonomic data in an internationally format : the ABCD Schema
- Practical implementations of the ABCD Schema
- **Conclusions**

Conclusions

- **Having data** in a spreadsheet is fine but can quickly become useless (new format, crash disk, …)
- **Having data** in a local database is better
- **Having data** in a share-database exposed to the world by international standards validates the taxonomic data optimally and frames science in a data-richer environment.

- Multiple taxa in a joint and standardized database brings additional value in terms of conservation and for understanding biodiversity

Conclusions

- The structure of a database with taxonomic content is very important to ensure **compatibility** with other database systems.

- For the **exchange** of taxonomic information it is necessary to have **standards** and **protocols** to permit the presentation, e.g. on a web system like GBIF, of species data from different database sources.

Thank you for your attention…