



# DISCOVERY: A Photo-Identification Data Management System



South African  
NATIONAL PARKS

14<sup>th</sup> Savanna  
Science Network  
Meeting

STEPHEN C.Y. CHAN<sup>1</sup>

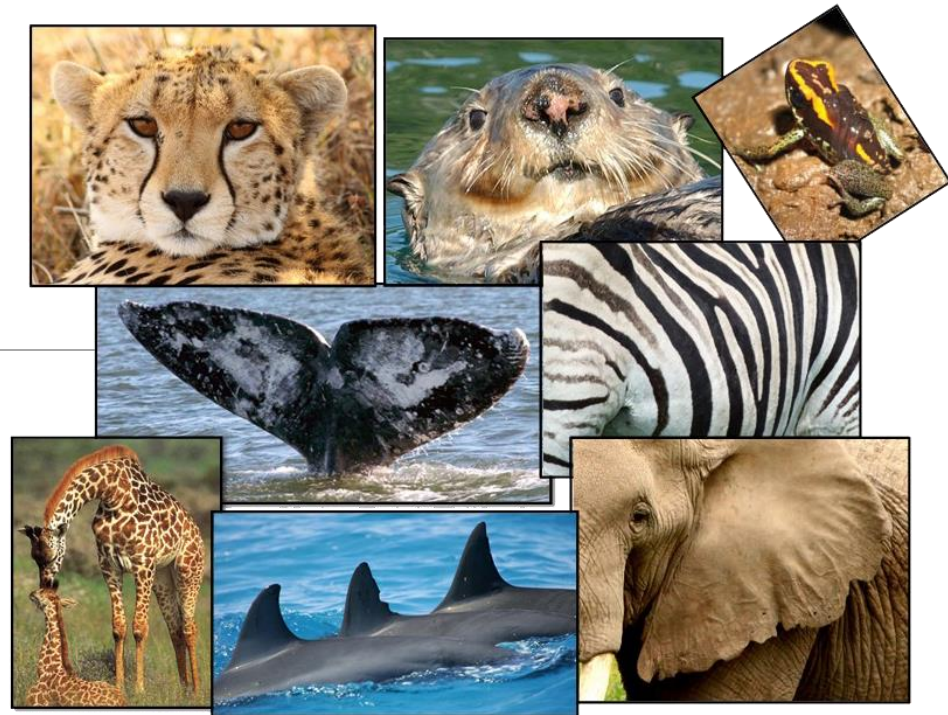
LESZEK KARZMARSKI<sup>1</sup>

GLENN GAILEY<sup>2</sup>

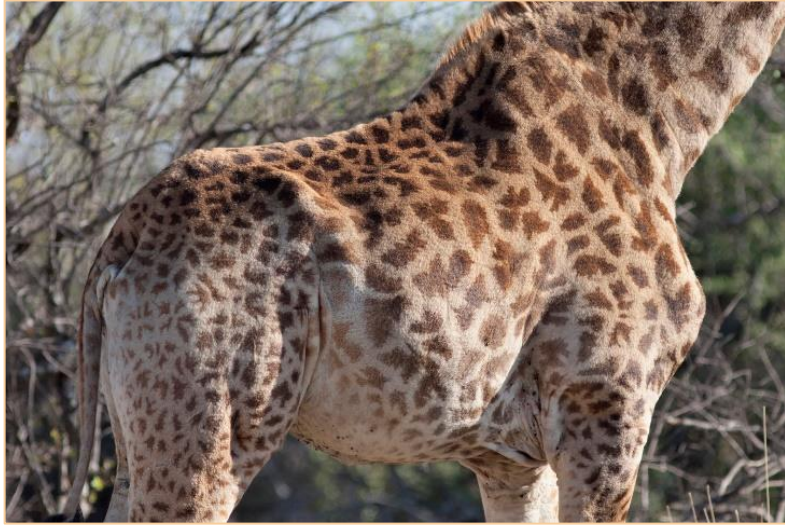
---

<sup>1</sup> THE SWIRE INSTITUTE OF MARINE  
SCIENCE  
THE UNIVERSITY OF HONG KONG

<sup>2</sup> CASCADIA RESEARCH COLLECTIVE

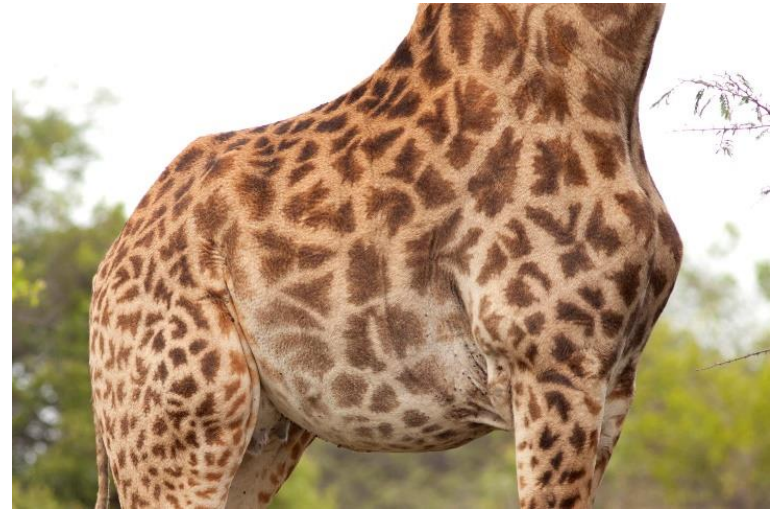


# Photo-Identification



???

=



- Time consuming
- Visually fatiguing
- Cross comparisons between databases



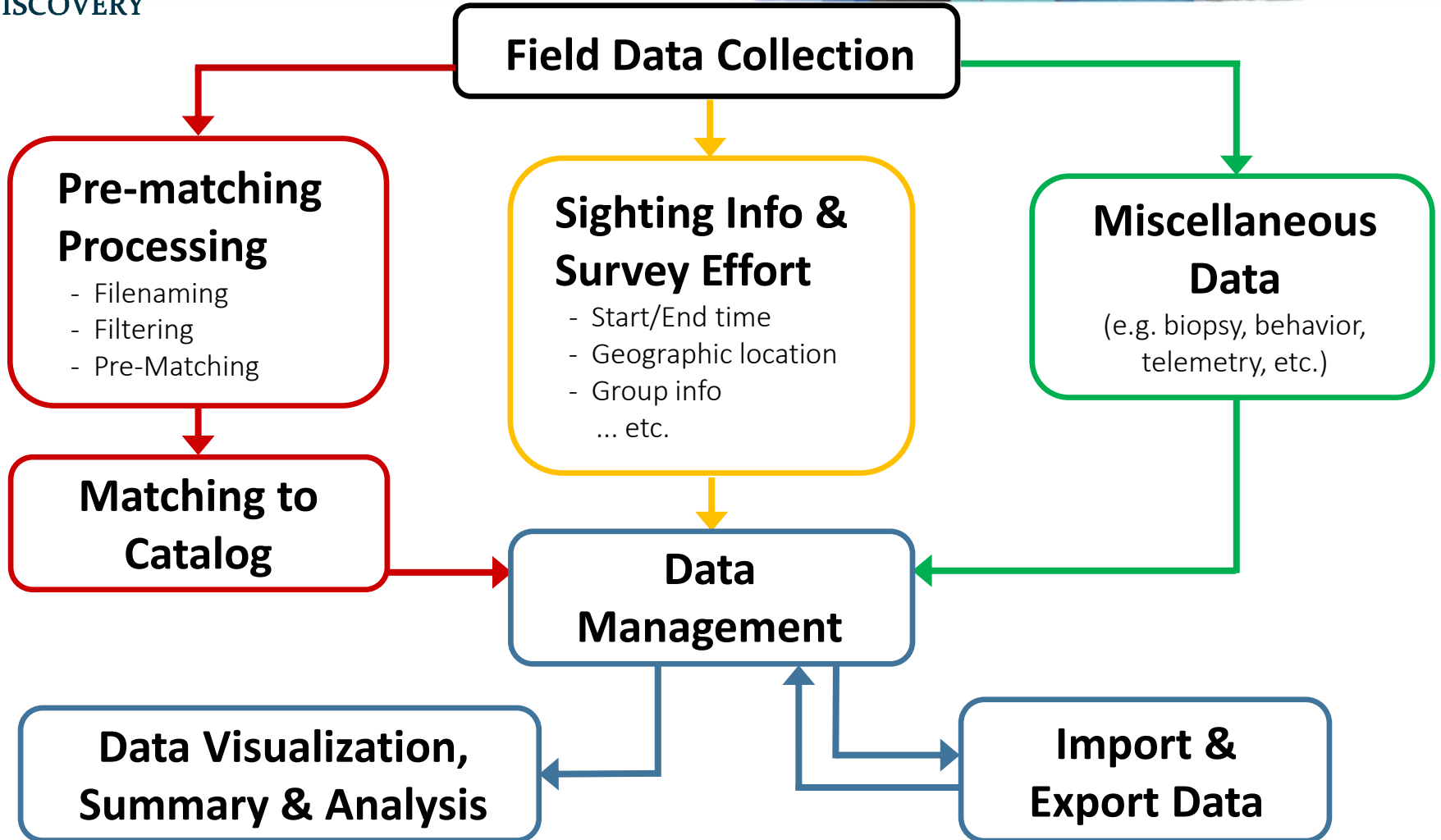
# Approach



- **Integrative system**
  - Store, visualize, manage and analyze photo-ID/associated data
- **Dynamic setting**
  - Meet various needs of research projects and user preferences
- **Inclusive** of other tools (*e.g. program R*)
- **Compatible** with other software (*e.g. MARK, SocProg, ArcGIS*)
- Efficient to maintain **long-term, multi-team** datasets



# Approach



# Pre-Matching Procedures

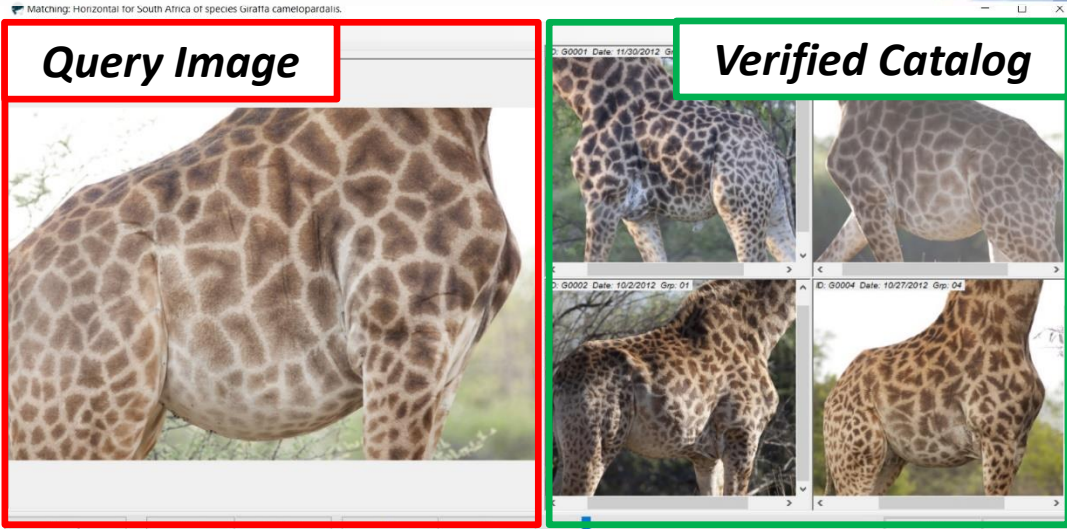


- Image processing
- Image filtering

***User-defined  
folders from  
dynamic setup***

A screenshot of an image processing software interface. The interface includes a file explorer on the left showing a directory structure with folders like 'Miscellaneous (2)', 'Quality1 (11)', 'Quality2 (4)', 'Quality3 (1)', and 'Trash (2)'. A red box highlights these folders with the text 'User-defined folders from dynamic setup'. The main window displays an image of an elephant with a green selection box around it. A toolbar at the top right contains buttons for 'Quality1', 'Quality2', 'Quality3', 'Trash', 'Unmark', 'Mother and Calf', and 'Miscellaneous'. A bottom toolbar contains buttons for 'Zoom', 'Select Area', 'Pan', 'Crop', and 'Save Image', along with sliders for 'Contrast', 'Brightness', and 'Sharpen'. A red arrow points from the 'Quality1' button in the top toolbar to the 'Quality1' folder in the file explorer.

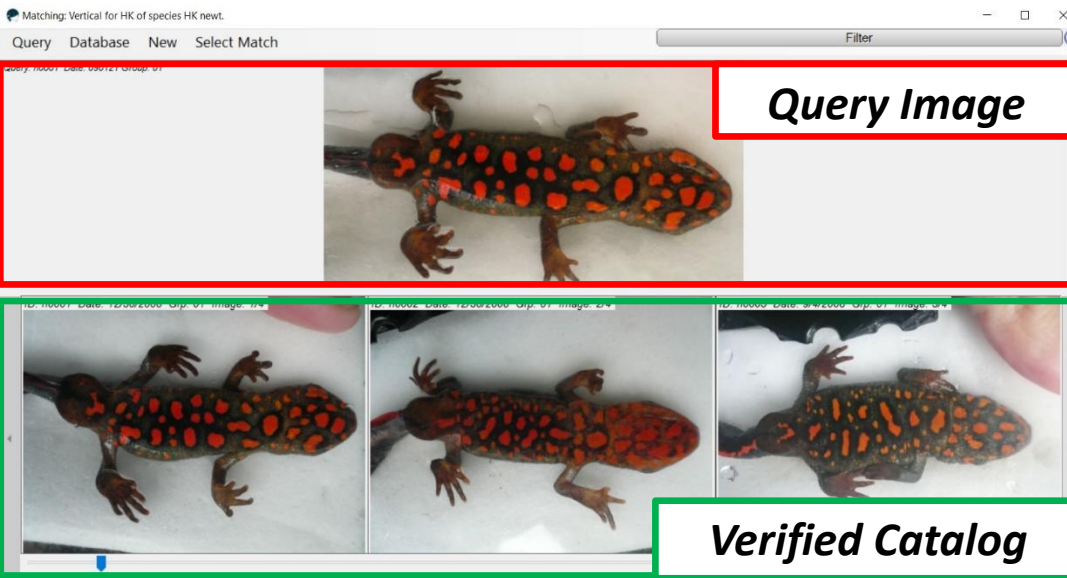
# Photo-ID Matching



*Horizontal View Mode*

✓ Enhanced searching of IDs by **categorizing database**

✓ **User-defined settings** to optimize efficiency



*Vertical View Mode*

# Survey Associated Info.



Sighting Information

Sighting ID: [Blank]

Study Area: Study:

Date:

Start Time:

End Time:

Survey:

Species:

Group ID:

Behavior:

Grp Size:

Group Info | Photographs | Position | Environment | **Comments** | Biopsy

Time:

#	Date	Time	Comment
1	02-Feb-2013	13:33:00	This is my comment
2	02-Feb-2013	13:55:00	This is my second comment

Depth:

# Data Management - Individual Catalog



Individuals

ID	# Images	# Sightings	First Year	Last Year	Sex
a0001	4	3	2005	2014	C
a0002	14	8	2005	2014	C
a0003	8	4	2004	2014	C
a0004	5	5	2003	2014	C
a0005	9	5	2005	2014	C
a0006	6	4	2006	2014	C
a0007	11	5	2004	2014	C
a0008	7	5	2005	2014	C
a0009	6	4	2006	2014	C
a0010	11	5	2004	2014	B
a0011	9	5	2004	2014	B
a0012	8	4	2006	2014	B
a0013	11	5	2005	2014	B
a0014	1	1	2014	2014	B
a0015	1	1	2014	2014	B
a0016	1	1	2016	2016	B

*Catalog of  
Individuals*



Left

Right

Front

Images of a0016

Image	GROUP_IMG	Date	Aspect
a0016_2016031...	01	10-Mar-16	Right

*Individual Images*



# Data Management - Sighting Records



**Sightings**

HongKong  
Zhuhai

**Study Site**

Sousa chinensis  
Neophocaena phocaenoides

**Species**

2010  
2011  
2012  
2013  
2014  
1/1/2014  
1/2/2014  
1/3/2014  
1/17/2014  
1/20/2014  
1/26/2014  
2/1/2014  
2/26/2014  
2/27/2014  
3/1/2014

Add Sighting

	Study Area	Date	Group	Start	Stop	Species	Survey	Behavior	Group Size	Ind	Environment	Comments	Geo	Photo
1	HongKong	26-Jan-2014	01	11:02:00	11:22:00	Sousa chinensis	PhotoID	Foraging	8	Ind	Env Data	Comments	Geo	Photo
2	HongKong	26-Jan-2014	02	11:35:00	11:45:00	Sousa chinensis	PhotoID	Travelling	2	Ind	Env Data	Comments	Geo	Photo
3	HongKong	26-Jan-2014	03	11:47:00	12:00:00	Sousa chinensis	PhotoID	Travelling	3	Ind	Env Data	Comments	Geo	Photo
4	HongKong	26-Jan-2014	04	12:48:00	13:10:00	Sousa chinensis	PhotoID	Foraging	2	Ind	Env Data	Comments	Geo	Photo

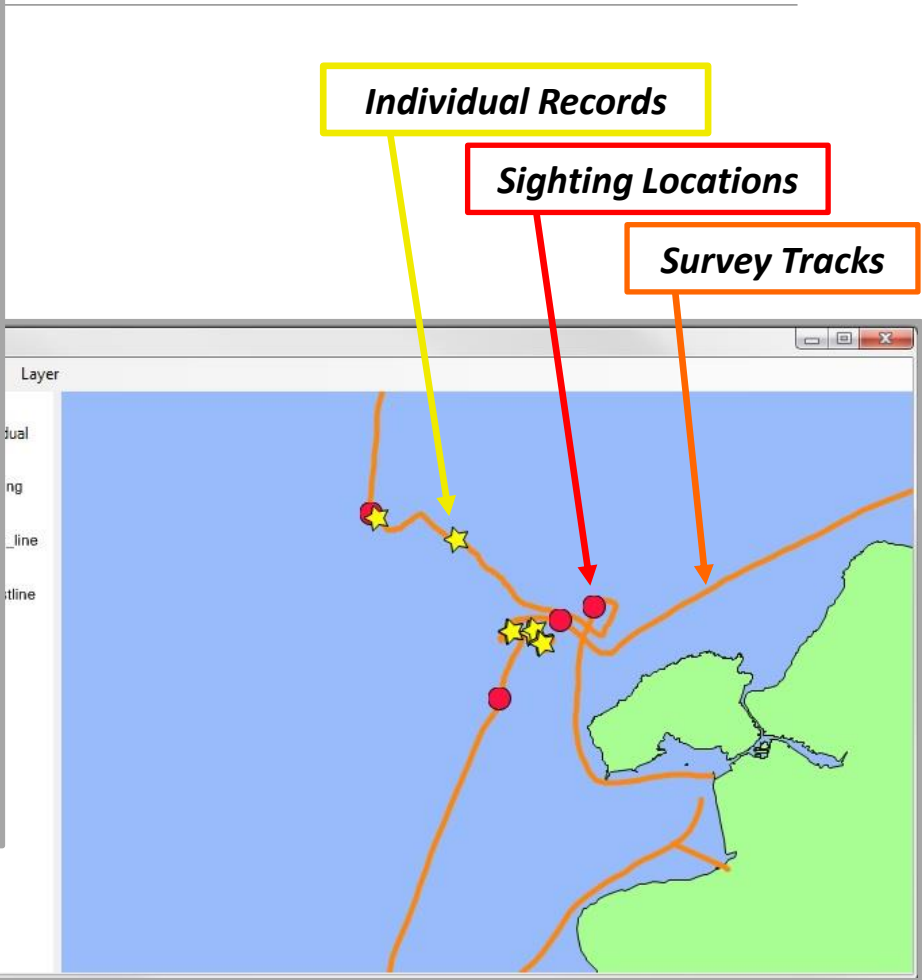
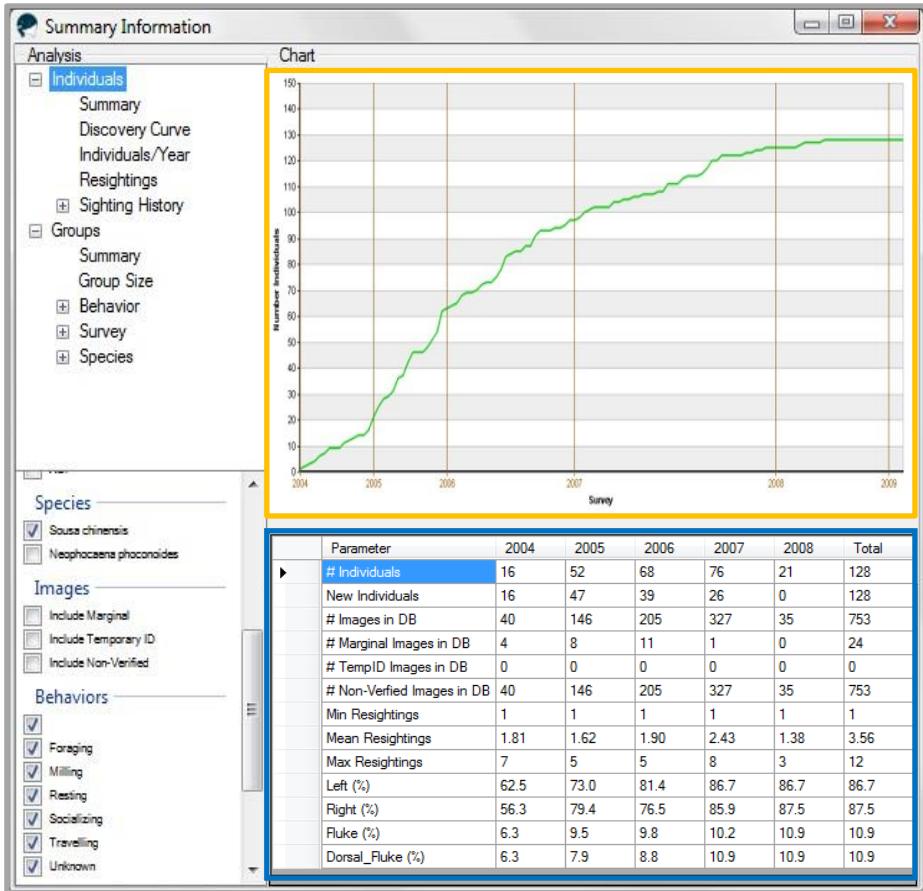
**Environment**

#	Time	Depth	SST	Offshore	Tide	Seastate	Habitat
1	11:02:00	6.1	20	3000	Ebb	1	open sea
2	11:12:00	5.7	19	5000	Ebb	1	
3	11:22:00	5.9	19	5000	Ebb	1	open sea
4	11:22:00	5.9	19	6000	Ebb	1	

**Geographic**

#	Time	Latitude	Longitude
1	11:31:46	22.35864	113.87870
2	11:47:24	22.36547	113.87570

# Data Visualization, Summary & Analysis



# Discovery R



- Wide range of analytical and display functions (open source)
- Built-in interface for users not familiar with R

**Database**

File Send  
Dataframe  
Images  
Sightings  
Vessel  
Group\_or\_Agg  
NN  
Foraging\_Tactics  
Gen\_Comments

Variables

PKey  
STUDY\_SITE  
INDIVIDUAL  
DIR\_LOCATION  
FILENAME  
DATE  
TIME  
GROUP\_IMG  
LATITUDE  
LONGITUDE  
TYPE\_SPECIMEN  
SPECIES  
VERIFIED  
TEMPORARYID  
QUALITY  
Plot

X  
Y  
Fill

Plot

**User-defined Plot**

**R Functions**

Behavior\_PieChart  
Discover\_Curve  
GoogleMap  
Histo\_Distinct  
ID\_YEAR  
Individuals\_Year  
Individuals\_Per\_Variable  
KdeAllFirst  
Resightings  
Sighting\_History  
TEST\_ME

**R Functions**

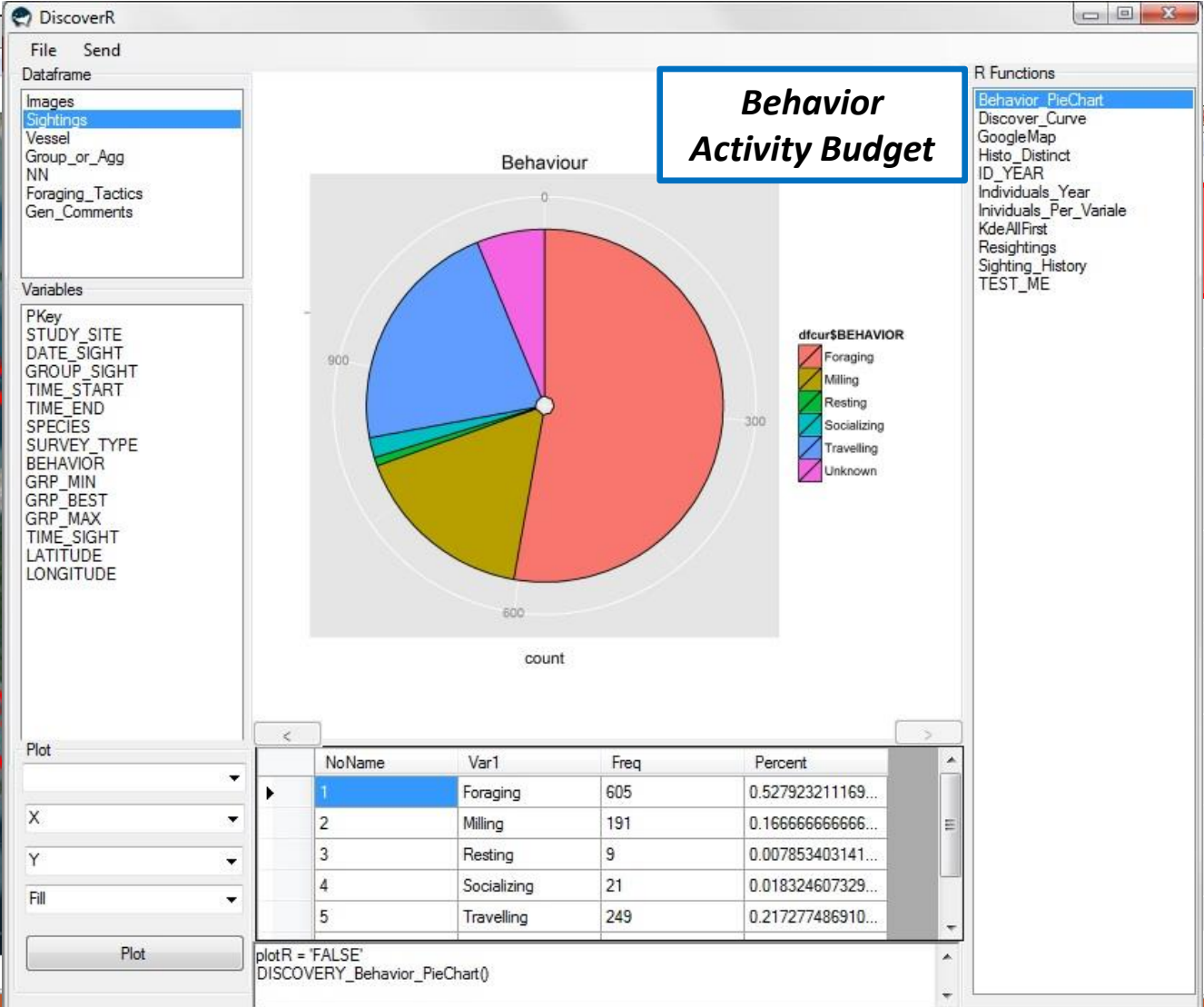
**Graphic & Table Output**

The screenshot shows the Discovery R interface. On the left, there is a 'Database' panel with a list of variables. In the center, a bar chart displays the 'Number of Individuals' for the years 2010 through 2014. Below the chart is a data table with columns for 'Year', 'NoName', and 'Individuals'. The 'Individuals' column has values 189, 226, 1, 0, 0, 0, 0 for the years 2010 through 2016. On the right, there is a panel for 'R Functions' with a list of functions, including 'Individuals\_Year' which is highlighted. Below the functions panel is a 'User-defined Plot' panel with a 'Plot' button. At the bottom right, there is a 'Graphic & Table Output' panel showing the bar chart and the data table.

Year	NoName	Individuals
2010	189	1
2011	226	0
2012	1	0
2013	0	0
2014	0	0
2015	0	0
2016	0	0

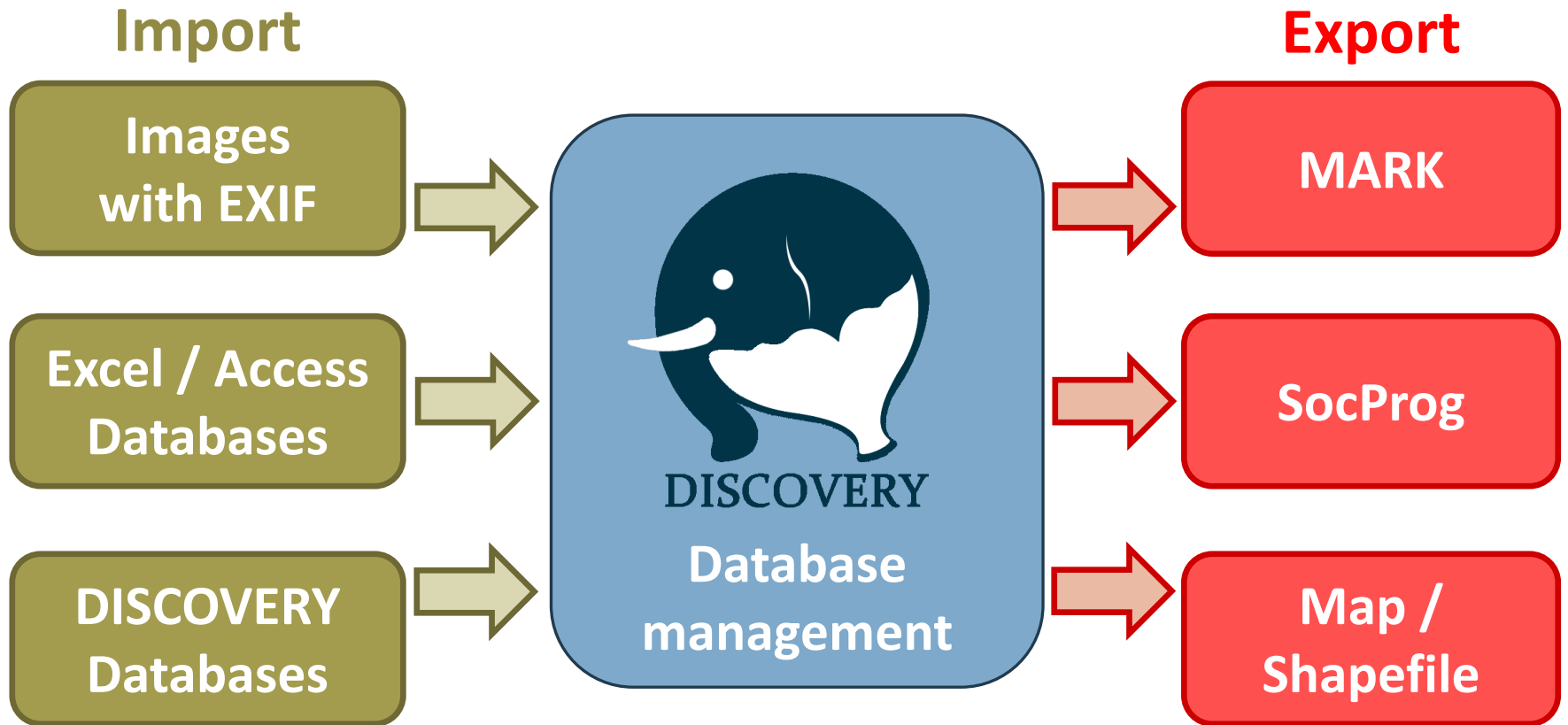
```
plotR = FALSE  
DISCOVERY_Individuals_Year()
```

# Discovery R



# Data Management

## - Import & Export



# Website (program download and manual):

<http://www.biosch.hku.hk/ecology/staffhp/lk/Discovery/>

*(New version out soon! Stay tuned!)*

OR

**Skukuza Unit 225 (14<sup>th</sup> – 17<sup>th</sup> March)**



**DISCOVERY: Photo-Identification Data-Management System for Individually Recognizable Animals**

## About the Software

- DISCOVERY Development Team
- Download DISCOVERY
- DISCOVERY Manual
- FAQ
- Related Links
- Forum

## About the Software

Individual photographic identification (photo-ID) represents a powerful technique to study behavioural and population ecology of free-ranging animals. This approach has been applied across species and habitats, both aquatic and terrestrial, gathering a large variety of data.

All photo-ID studies require many hours of intensive field surveys and even longer hours of subsequent, labour-intensive processing of photographic material. Thanks to the recent advances in digital photography, high quality digital images can be obtained in a short space of time and the photo-ID data can be processed immediately upon the completion of a field day, even in remote locations where processing of traditional photographic material would not have been possible. However, such fast accumulation of data can pose an obvious and often considerable obstacle to data management. This is where DISCOVERY comes handy; it provides a dynamic, user-friendly platform to assist researchers not only with the matching of individual photo-ID data, but also at the multitude of steps of field data collection and the complex data management and analyses that follow after individual matching is completed.



The DISCOVERY system assists with filtering of raw data and all levels of individual-ID matching; it assists with processing, storing and managing digital images; it provides file naming routines and links sighting information with environmental, geographic, and numerous user-defined parameters; it provides graphic displays of data and basic analytical tools. DISCOVERY can be used to centralize a database for multiple species and multiple study areas; it is particularly useful for maintaining a single database for research projects collecting data at large geographical scales and between multiple research teams working on different databases. DISCOVERY also provides a means of linking the new system with traditional datasets based on film photography, to form continuous complete datasets. The DISCOVERY system has been designed so that it can easily facilitate integration of all collected and stored data to and from other tools; with a multitude of dynamic functions it was designed to meet project-specific requirements and user-specific needs.

© 2012 Glenn Gailey and Leszek Karczmarski



# Acknowledgements



- **Principal Investigators**

- Glenn Gailey
- Leszek Karczmarski

- **DISCOVERY Development Team**

- Stephen C.Y. Chan, Simon W.H. Wong, Olga Sychenko, Carmen K.M. Or, Scott Y.S. Chui, Derek Y.W. Ho

- **Home Institutions**

- The Swire Institute of Marine Science, University of Hong Kong
- Mammal Research Institute, University of Pretoria, South Africa
- Cascadia Research Collective

- **Funding Organizations**

- National Research Foundation (NRF), South Africa
- Research Grants Council (RGC), Hong Kong
- Ocean Park Conservation Foundation, Hong Kong (OPCFHK)
- The University of Hong Kong



**National  
Research  
Foundation**

**Research Grants Council  
of Hong Kong**  
香港 研究資助局

