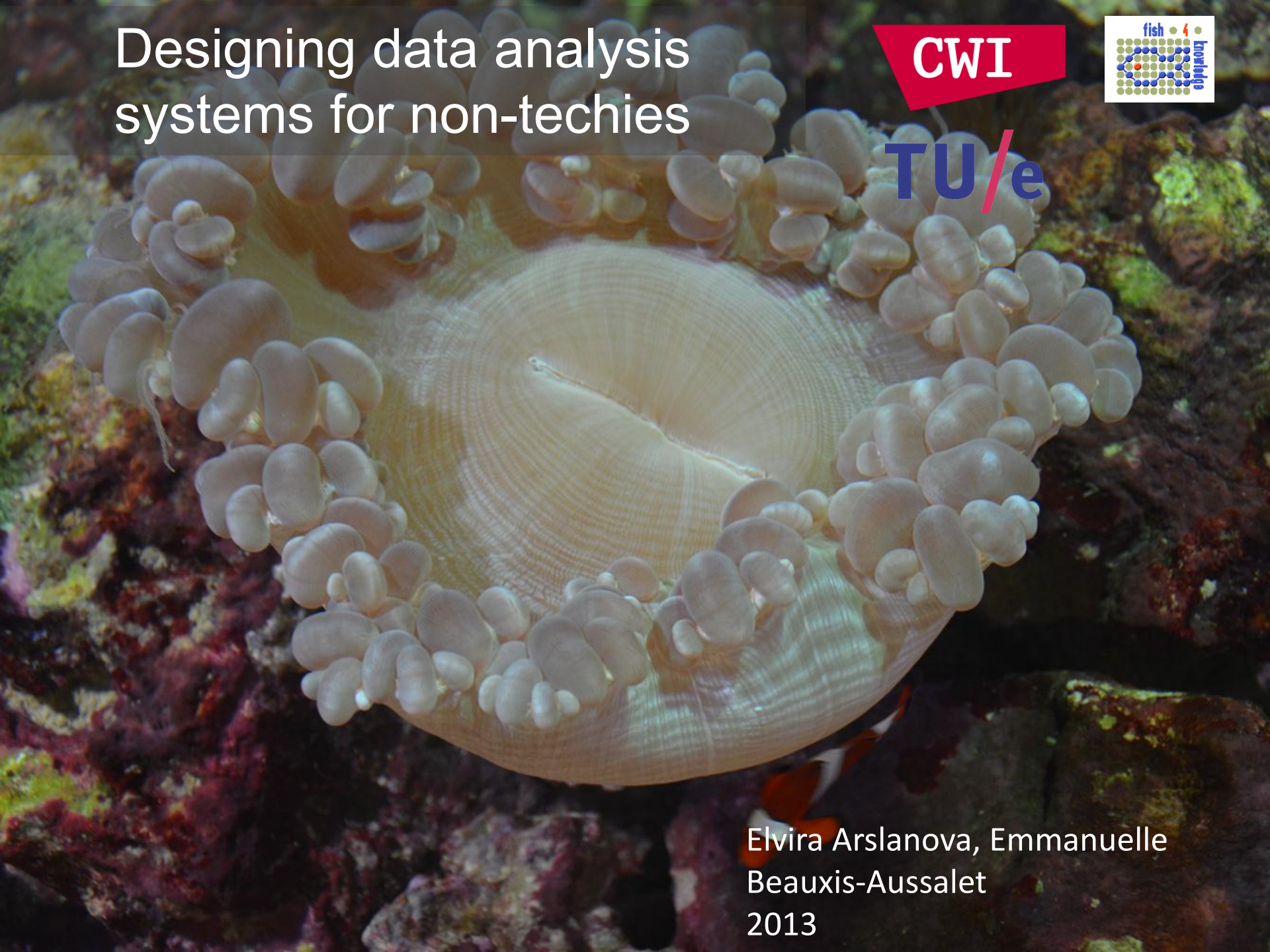
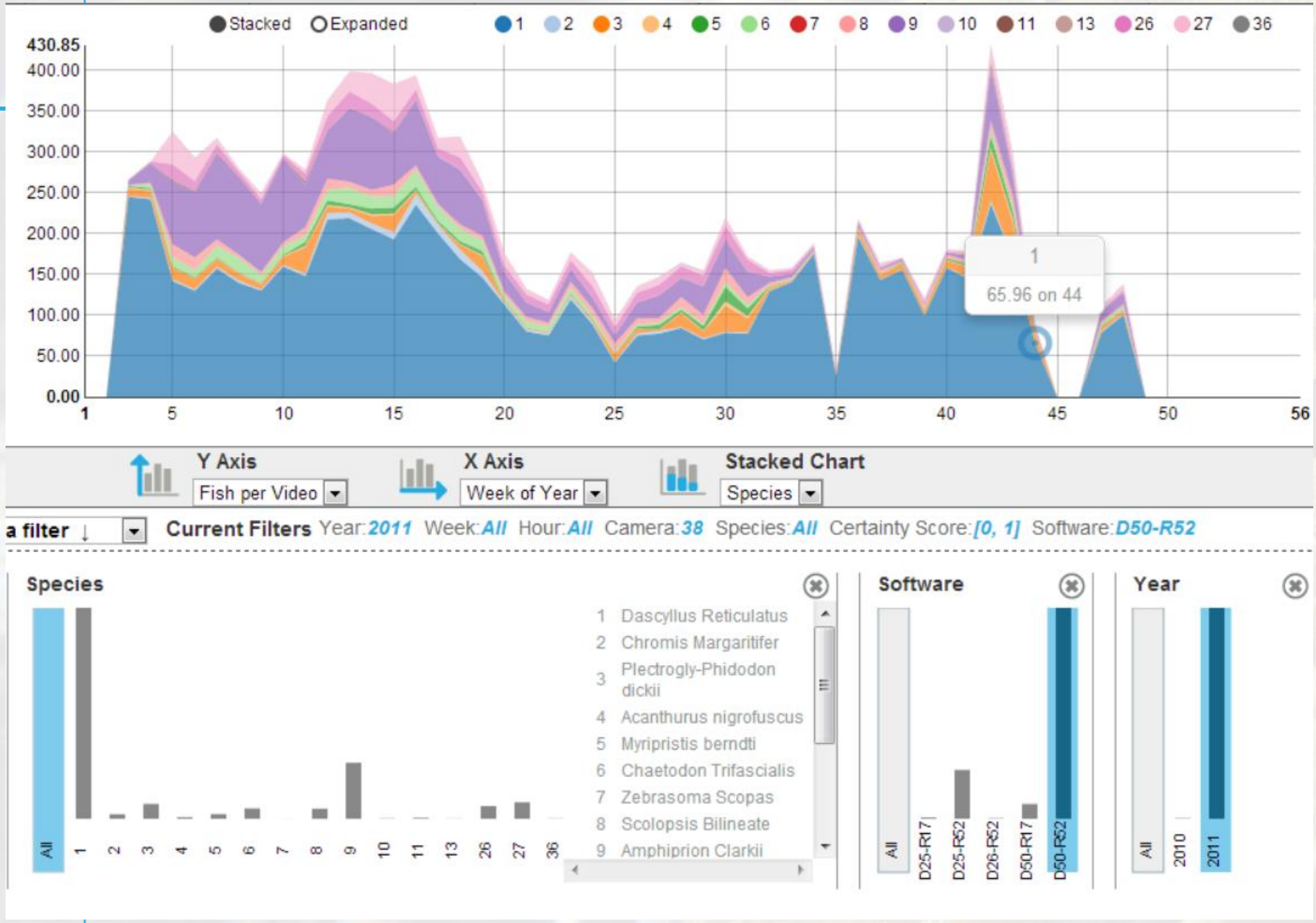
Designing data analysis systems for non-techies



Elvira Arslanova, Emmanuelle Beauxis-Aussalet
2013


by Navid Baraty
Fish4Knowledge


Stacked
Expanded
1
2
3
4
5
6
7
8
9
10
11
13
26
27
36
430.85
400.00
350.00
300.00
250.00
200.00
150.00
100.00
50.00
0.00
1
5
10
15
20
25
30
35
40
45
50
56
1
65.96 on 44
Y Axis
Fish per Video
X Axis
Week of Year
Stacked Chart
Species
a filter
Current Filters Year:2011 Week:All Hour:All Camera:38 Species:All Certainty Score:[0, 1] Software:D50-R52
Species
All
1
2
3
4
5
6
7
8
9
10
11
13
26
27
36
1 Dascyllus Reticulatus
2 Chromis Margaritifer
3 Plectrogly-Phidodon dickii
4 Acanthurus nigrofuscus
5 Myripristis berndti
6 Chaetodon Trifascialis
7 Zebrasoma Scopas
8 Scolopsis Bilineate
9 Amphiprion Clarkii
Software
All
D25-R17
D25-R52
D26-R52
D50-R17
D50-R52
Year
All
2010
2011

Challenges

- Technology Push
- Uncertainty in data and the tool
- Situation Awareness

Technology Push

- Involving users
- Explicit feedback
- Implicit feedback

User Studies: Trust Issues

Goal

What is the optimal amount of information about the algorithms to reveal in order to increase trust and acceptance to the software?

Method

Experimental interfaces with 3 levels of complexity of the explanation

Test questions

Interviews, questionnaire measuring trust

Participants

20 Dutch and Taiwanese biologists: fishery(morphology, taxonomy) coral biologists

Analysis

Mainly qualitative of the answers + quantitative

JANUARY FEBRUARY MARCH APRIL MAY JUNE JULY AUGUST SEPTEMBER

User Studies: Trust Issues

User Trust (Technical Competence and Reliability of the Software)

Perceived Understanding of Technical Concepts

Acceptance of the Tool

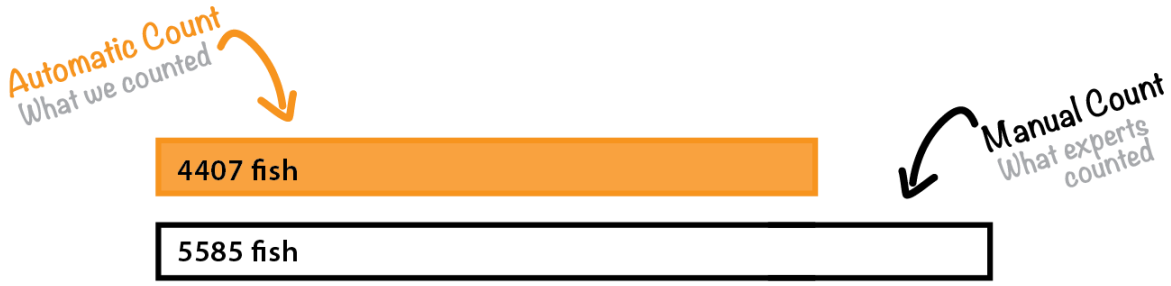
Satisfying User Information Needs

THE ACCURACY OF OUR FISH COUNTS

To evaluate the quality of our automatic count of fish, we asked marine biology experts to manually count the fish that appear in the set of Videos for Evaluation. In the 102 videos for evaluation, experts found 5585 fish, whereas our automatic count is of 4407 fish.

How many fish in the videos?

THE ACCURACY OF AUTOMATIC FISH COUNT



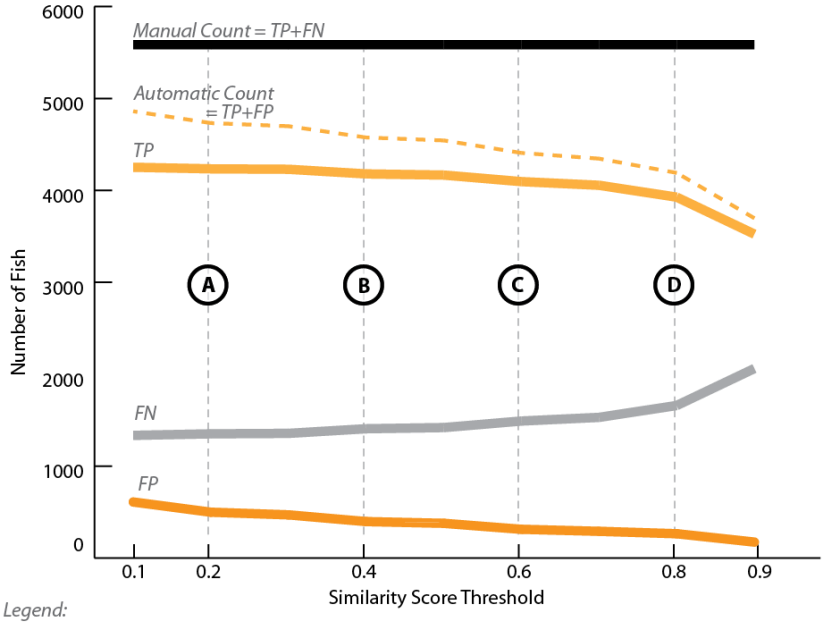
Legend:

- **Automatic Count**: the number of fish detected by our video analysis software (4407 fish)
- **Manual Count**: the number of fish detected by experts (5585 fish)

THE SIMILARITY SCORES

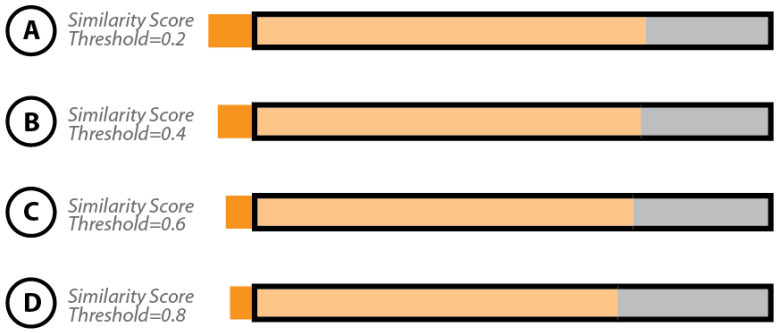
To further improve our automatic counts, we can calculate a Similarity Score that indicates how a fish image is similar to our fish model.
We give a Similarity Score to all detected fish. And we use a Similarity Score threshold to discard the fish that are not similar enough to our model.
The figures below show the fish counts and their accuracy at various thresholds.

FISH COUNTS, TP, FP AND FN OVER SIMILARITY SCORE THRESHOLDS



Legend:
Manual Count = TP+FN
TP **FP** **FN** **Automatic Count = TP+FP**

ACCURACY OF FISH COUNTS OVER SIMILARITY SCORE THRESHOLDS



Legend:
False Positives (FP): the non-fish objects we incorrectly detected as fish
True Positives (TP): the fish we correctly detected
False Negatives (FN): the fish we did not detect
Manual Count = TP+FP: all the fish that experts manually detected

JANUARY FEBRUARY MARCH APRIL MAY JUNE JULY AUGUST SEPTEMBER

User Studies: Trust Issues

User Trust

- Most didn't trust
- Most confident participants showed worse understanding
- Slightly improved after 3rd interface

Understanding of Technical Concepts

- Big cognitive effort
- Misunderstandings
- Extra explanations

Acceptance of the Tool

- High
- Slightly improved after 3rd interface
- Chose the 3rd version

Satisfying User Information Needs

- Need for more information

Do they want to Know?

- No

Do they need to Know?

- Yes

Do they understand?

- They say “yes”
- The more confident they were, the less correct the results were

What can WE do Accept & Trust the tool?

- Data Provenance
- Be honest about possible errors and provide the full overview
- Defaults approved by people in their field
- Validate the results using the well established methods in their field

Goal

What situation awareness issues arise with different levels of task complexity?

Method

Users are exposed to 3 predefined UI states with increasing level of complexity

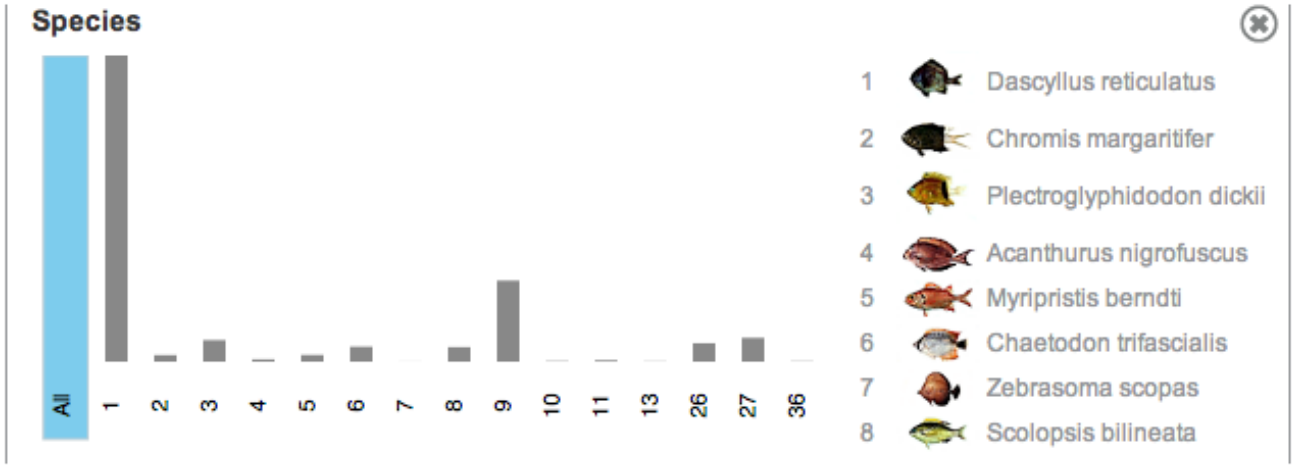
They are asked to accomplish tasks related to 3 levels of Situation Awareness processes:

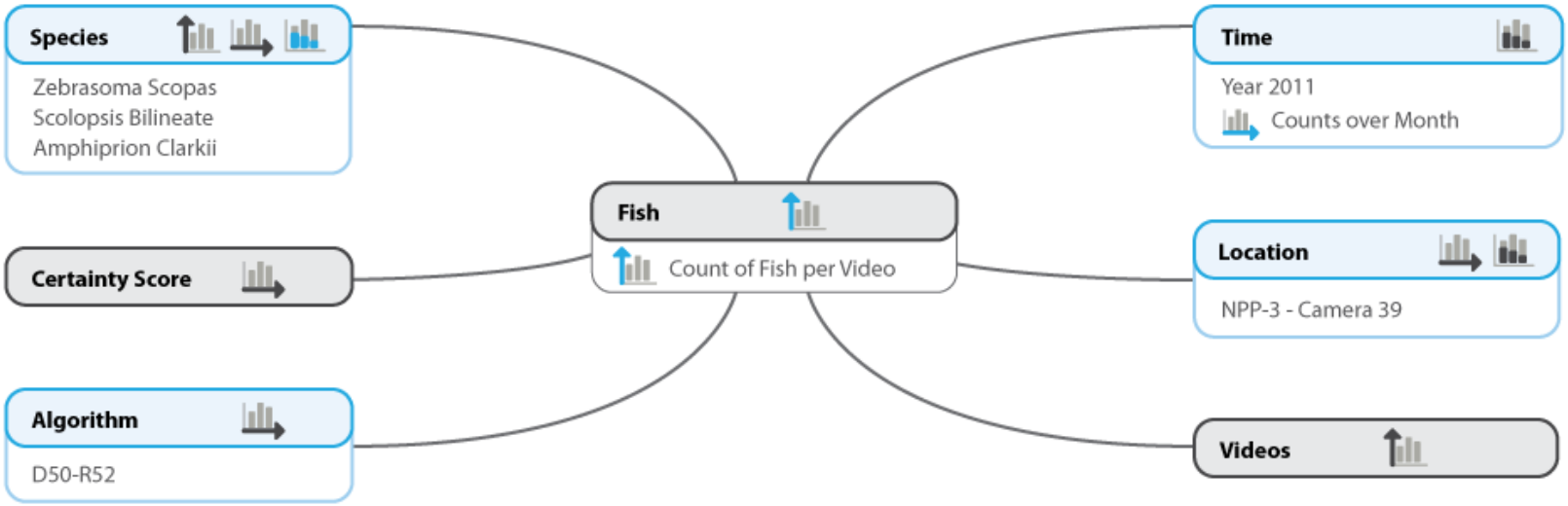
- Perception (reading the displayed state)
- Comprehension (reading and correlatiing facts)
- Projection (high-level interpretation of data)

Participants

13 Taiwanese biologists: fishery(morphology, taxonomy) coral biologists, plankton, toxicology

Overview


Species
All
1
2
3
4
5
6
7
8
9
10
11
13
26
27
36
1 Dascyllus reticulatus
2 Chromis margaritifer
3 Plectroglyphidodon dickii
4 Acanthurus nigrofuscus
5 Myripristis berndti
6 Chaetodon trifascialis
7 Zebrasoma scopas
8 Scolopsis bilineata


Species
Zebrasoma Scopas
Scolopsis Bilineate
Amphiprion Clarkii
Certainty Score
Algorithm
D50-R52
Fish
Count of Fish per Video
Time
Year 2011
Counts over Month
Location
NPP-3 - Camera 39
Videos

Complexity of task

- The more complex the task is the more vivid should be changes made to the dataset

Visualizations

- Careful with choice of the diagrams: type of data

Defaults

- They know that they don’t need to use this function, however, they don’t tend to check if the default is correct

Questions? Suggestions?

www.fish4knowledge.eu

Team:

E. Beauxis-Aussalet, M.Sc.

Prof.dr. L. Hardman

Dr.J.He

Dr. J.R. van Ossenbruggen

TU/e CWI