CITIZEN SCIENCE – LEARNING WITH AND FROM THE PUBLIC THROUGH CITIZEN SCIENCE: TOOLS FOR BUILDING CONNECTIONS

Yolanda Wiersma

Associate Professor, Biology

MEMORIAL UNIVERSITY

ACCESS 2018 Workshop & Meeting, St. John's, NL

Collaborators

- Dr. Jeffrey Parsons, Memorial University, Faculty of Business
- Dr. Roman Lukyanenko, U Sask, Edwards School of Business
- Dr. Max Liboiron, Memorial University, Department of Geography
- Dr. Gisela Wachinger, University of Stuttgart/DIALOGIK
- Matt McWilliams, MSc(Env) student

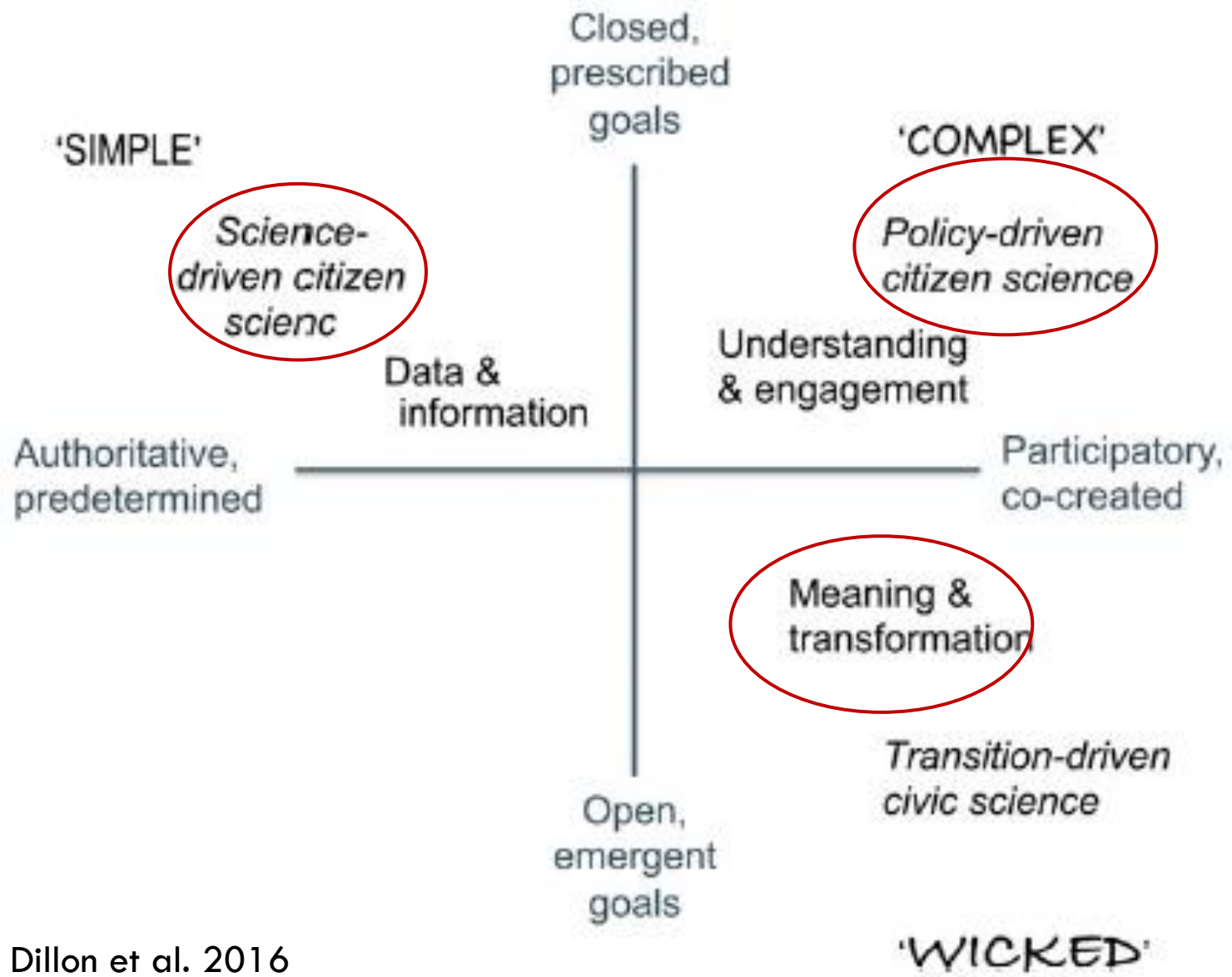
interdisciplinary
communication
natural science
lab biology
synchronous
fun field
issues
data
habitat
engage
hands on
biodiversity
experiment
problem solving
science education
conservation biology
environmental science
distance learning
migration
21st century
explore
questions
experiential
species
asynchronous
inquiry
investigate
creativity
ecosystems

citizen science

A new era of citizen science

Muki Haklay

Participatory Citizen Science



Dillon et al. 2016

My own research on CS

- NLNature.com
  - Collaboration with Information Systems Scientists (Dr. Jeff Parsons and Dr. Roman Lukyanenko) on "Crowd IQ" (Crowd Information Quality)

Easier citizen science is better

Non-scientists are now participating in research in ways that were previously impossible, thanks to more web-based projects to collect and analyse data. Here we suggest a way to encourage broader participation while increasing the quality of data.

Participation may be passive, as when someone donates their computer's 'downtime' to projects such as SETI@home, or active, as when someone uses eBird to log birds they have spotted. Unfortunately,

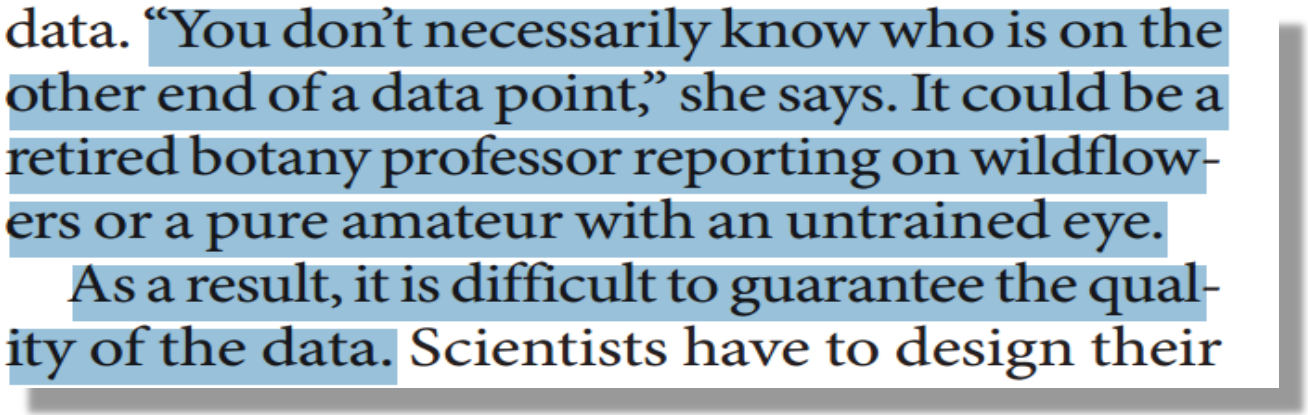
bird as oil-covered may be more valuable than asking them to guess what the species is. For such data to be used effectively, they need to be stored in a way that supports attributes rather than fixed, predetermined classes.

Jeffrey Parsons, Roman Lukyanenko and Yolanda Wiersma *Memorial University of Newfoundland, Canada.*
jeffreyp@mun.ca

Nature (2011) 471: 37

Research Problem

- Major challenge in making effective use of citizen data is crowd data quality (DQ)
  - *E.g., accuracy of a citizen science observation on eBird.org*



data. "You don't necessarily know who is on the other end of a data point," she says. It could be a retired botany professor reporting on wildflowers or a pure amateur with an untrained eye.

As a result, it is difficult to guarantee the quality of the data. Scientists have to design their

Gura 2013 *Nature*

Prevailing Perspective on Crowd DQ

- Popular approaches to ensure DQ
  - Educate and train online users
  - Provide data collection instructions
  - "Clean" data post-hoc
- Focus on data consumers (e.g., scientists)
  - Dissuade contributors from providing data
  - Prevent contributors from communicating important local knowledge

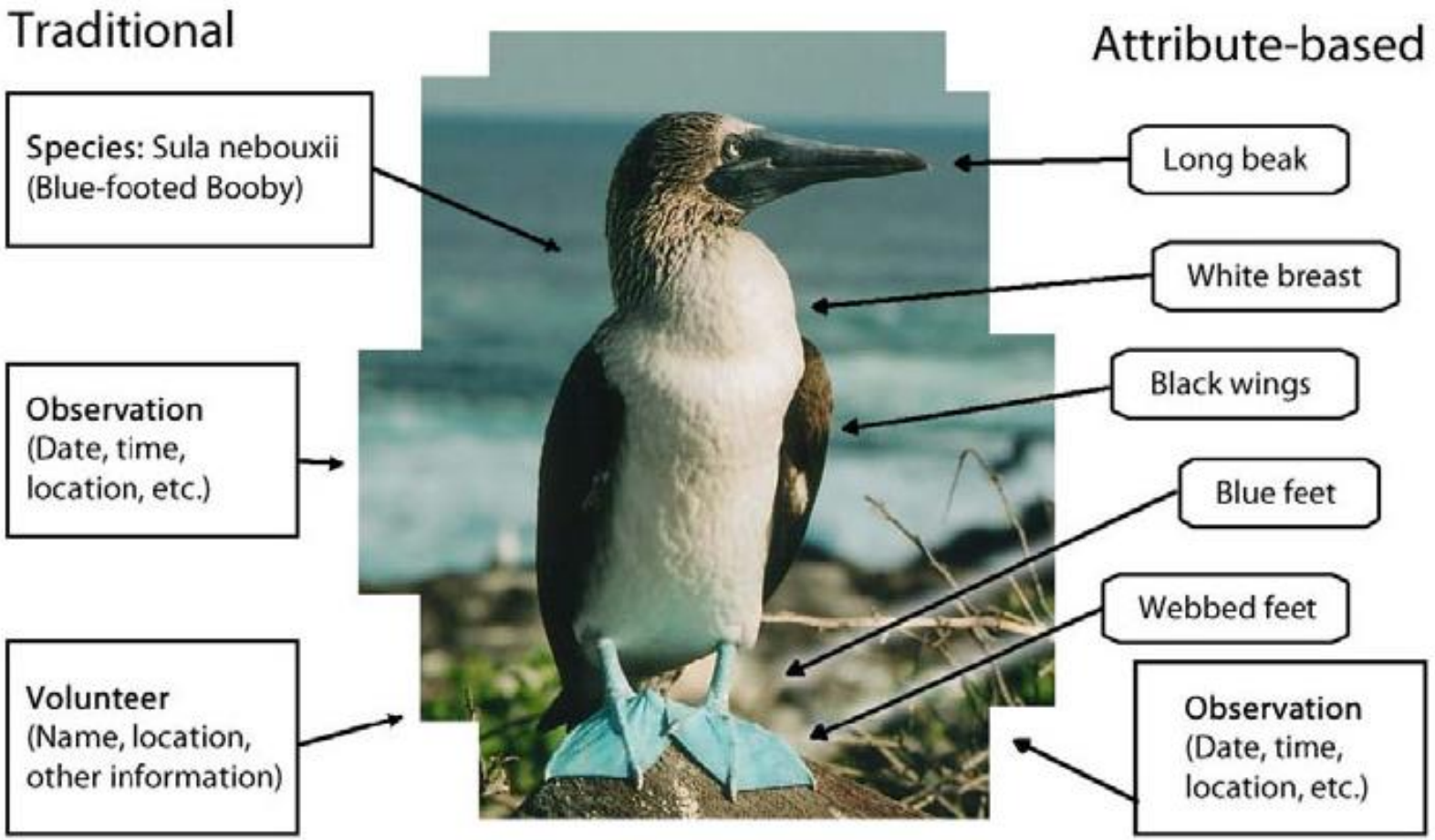
Alternative Approach to DQ

- DQ from contributors' perspective
  - "*the extent to which stored information represents the phenomena of interest to data consumers, as perceived by information contributors*"
    - Use-agnostic
    - Contributor-centric
    - Cognizant of data consumers

Our Research Focus

- Minimizing participation constraints
- Increasing the quantity and accuracy of data generated
- Capturing data on unexpected organisms

Class-based vs. instance-based



Lukyanenko, Parsons & Wiersma (2011) *Lecture Notes in Computer Science*

Illustration of the problem

?

Great Egret

Snowy Egret

White Ibis

Incorrect guess
↓ **accuracy**

Avoid participating
↓ **dataset completeness**

Any choice (incl. correct)
↓ **instance completeness**
(attribute loss)

Experiment 1: Free form

- N=247 non-experts (141 female, 106 male)
  - 24 full-color images of plants and animals




  - Task: ***What is it?***
  - Free-form responses

Experiment 1: Results

Useful classes

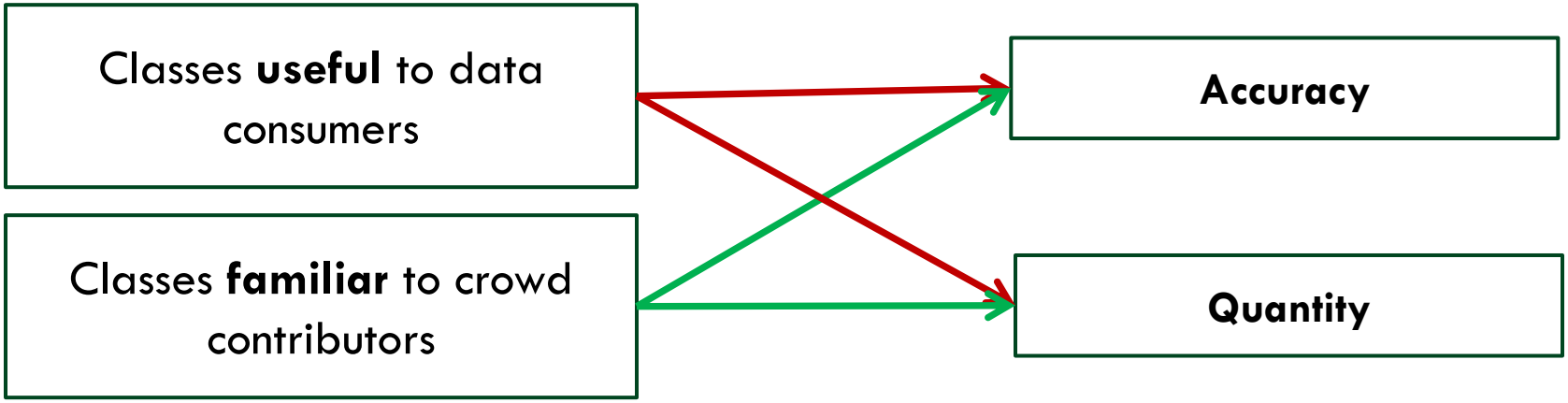
(e.g., great egret):

- 141 total
- 19.15% correct

Familiar classes

(e.g., bird)

- 1550 total
- 98.26% correct

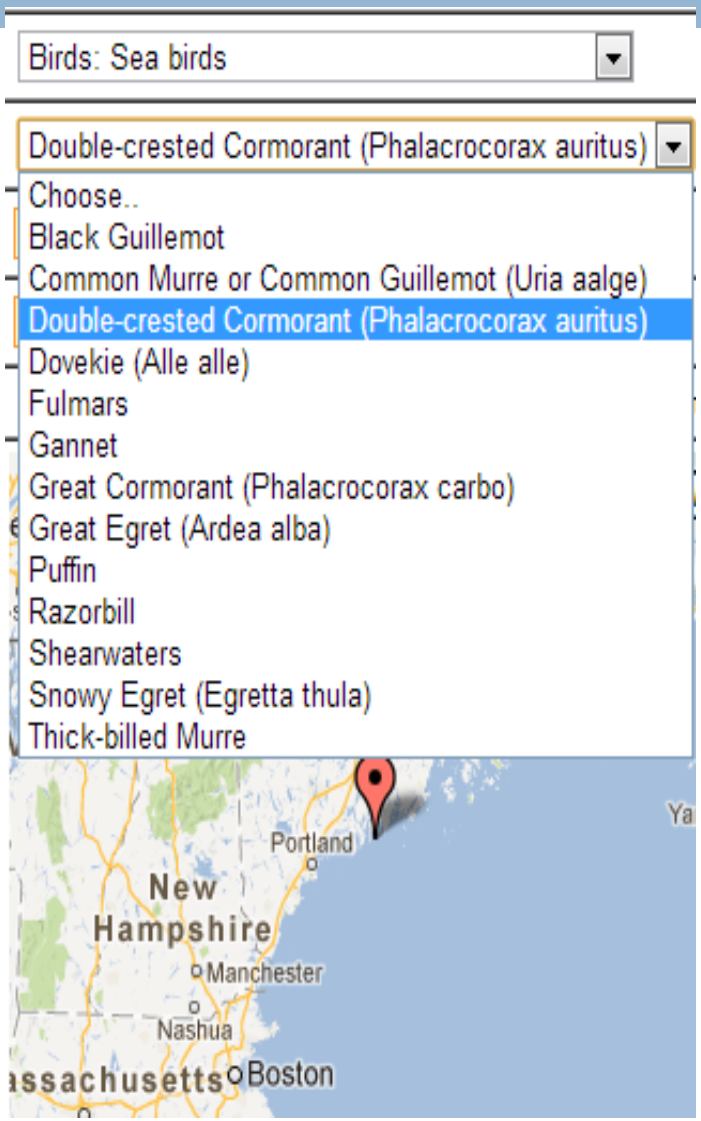


avg. $p < 0.001$*

* Based on Fisher's exact test; Sig with Bonferroni correction

Experiment 2: Fixed-choice

- With predefined classes
  - Species, useful to scientists
  - Generic, familiar to non-experts
- N=77 non-experts
- Task: select class from predefined list



Experiment 2: Materials

"Useful" Condition

What is it? Select one:

- *Arctic Tern*
- *Bonaparte's Gull*
- *Caspian Tern*
- *Common Tern*
- *Herring Gull*
- *Iceland Gull*
- *Killdeer*
- *Parasitic jaeger*
- *Pomarine jaeger*
- *I don't know*
- *Other ___*

Cognitive psychology

"Familiar" Condition

What is it? Select one:

- *Animal*
- *Common Tern*
- *Iceland Gull*
- *Killdeer*
- *Seagull*
- *Shorebird*
- *Tern*
- *Waterfowl*
- *Bird*
- *I don't know*
- *Other ___*

Experiment 2: Results

Useful classes

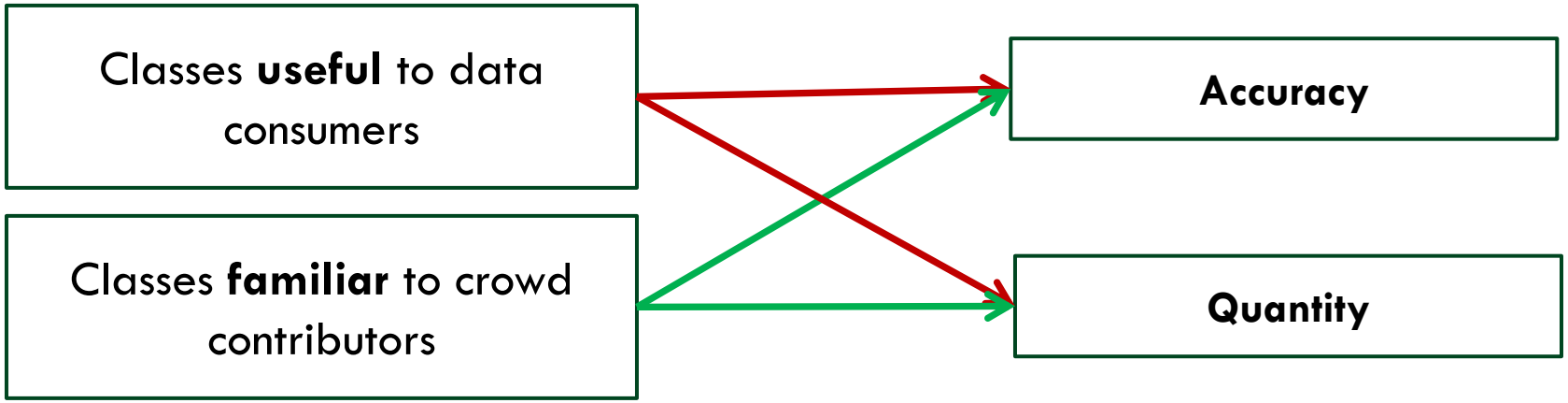
(e.g., great egret):

- 271 total
- 73 (24.84%) correct
- ***IDK, Other not used!***

Familiar classes

(e.g., bird)

- 375 total
- 277 (73.88%) correct



* Based on χ^2 test;
Sig with Bonferroni correction

Conclusions of initial experiments

- Accuracy, quantity of data declines when citizens are asked to comply with needs of scientists
- Accuracy does not necessarily increase when "familiar" options are included
  - Training – no panacea!

Problem

- Data is highly variable, inconsistent
- Low precision: (rarely at species level)

Sighting Info

Observed: July 24, 2013 @ 1:45 PM
Posted on: August 17, 2013 @ 9:09 AM (diff: 24 days)
Comments:

> I photographed this while I was on a Gatheralls boat tour of the Witless Bay islands.

Sighting's Identification

- **Bird** - **Puffin** - **Sea bird**

Sighting Info

Observed: July 10, 2013 @ 1:00 PM
Posted on: July 12, 2013 @ 9:41 AM (diff: 2 days)

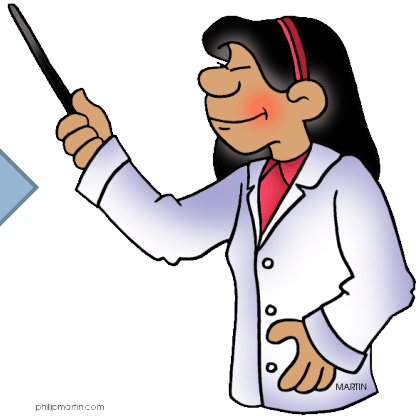
Sighting's Identification

- **Atlantic Puffin**
(Fratercula arctica)

Solution:

Apply AI / Analytics to IB Data

Using experts and AI to post-process data

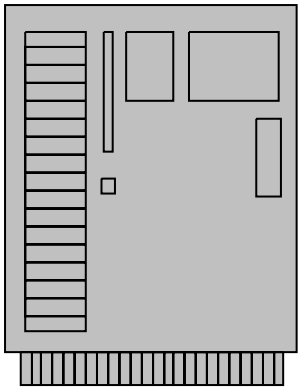


59.4%

American robin (Turdus migratorius)

Common tern (Sterna hirundo)

Lung lichen (Lobaria pulmonaria)



74.4%

Conclusions

- Use-agnostic, instance-based data has advantages for
  - Accuracy, Quantity of data, Ease of use, Participation, Discoveries
- **Augmented with Artificial Intelligence/ Analytics**
  - **Can be made more consistent**
  - **Can achieve the desired level of specificity (e.g., species)**

My own research on CS

- Fogo Island Fishermen Project (with Dr. Max Liboiron and Matt McWilliams)
  - Participatory citizen science

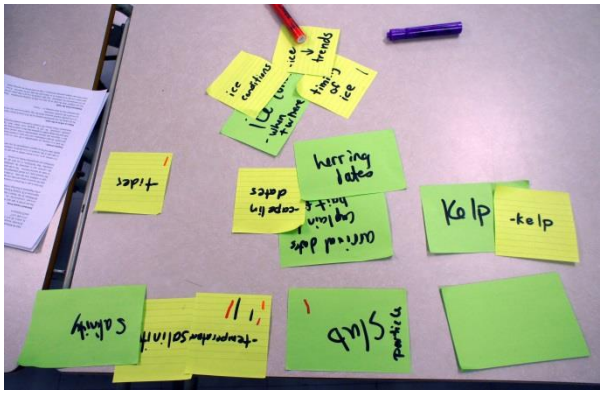


Photo credits: Phoebe Sengers

Acknowledgements

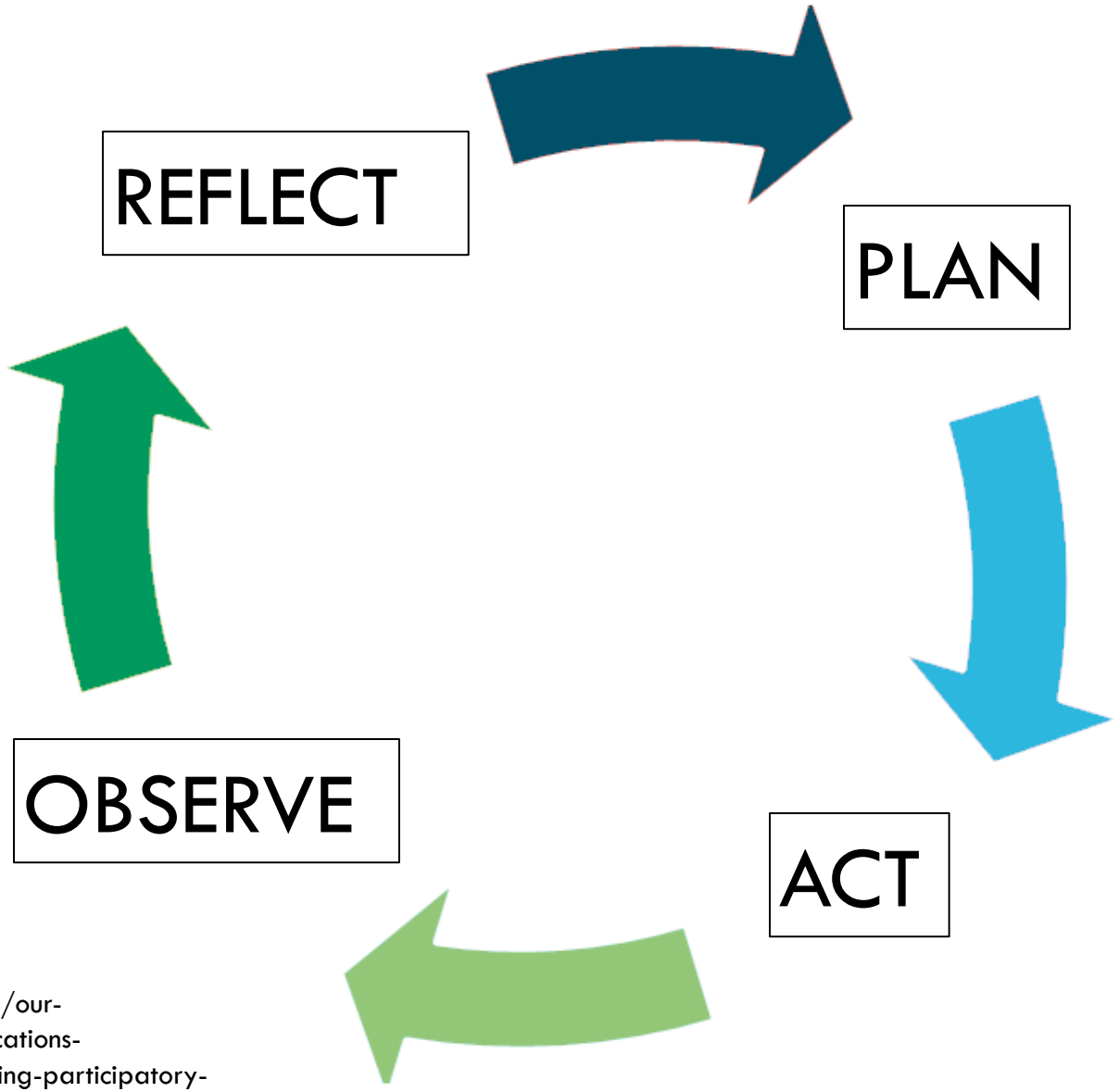
Fogo Island Research Team:

Don Best, Glen Best, Jerry Best, Rodney Budden, Bernadette and Gerard Dwyer, George Ford, Shawn Lynch, Aubrey and Marie Payne, Gordon Payne, Austin Reid

shorefast FOUNDATION

THE HARRIS CENTRE Memorial University

Participatory Action Research

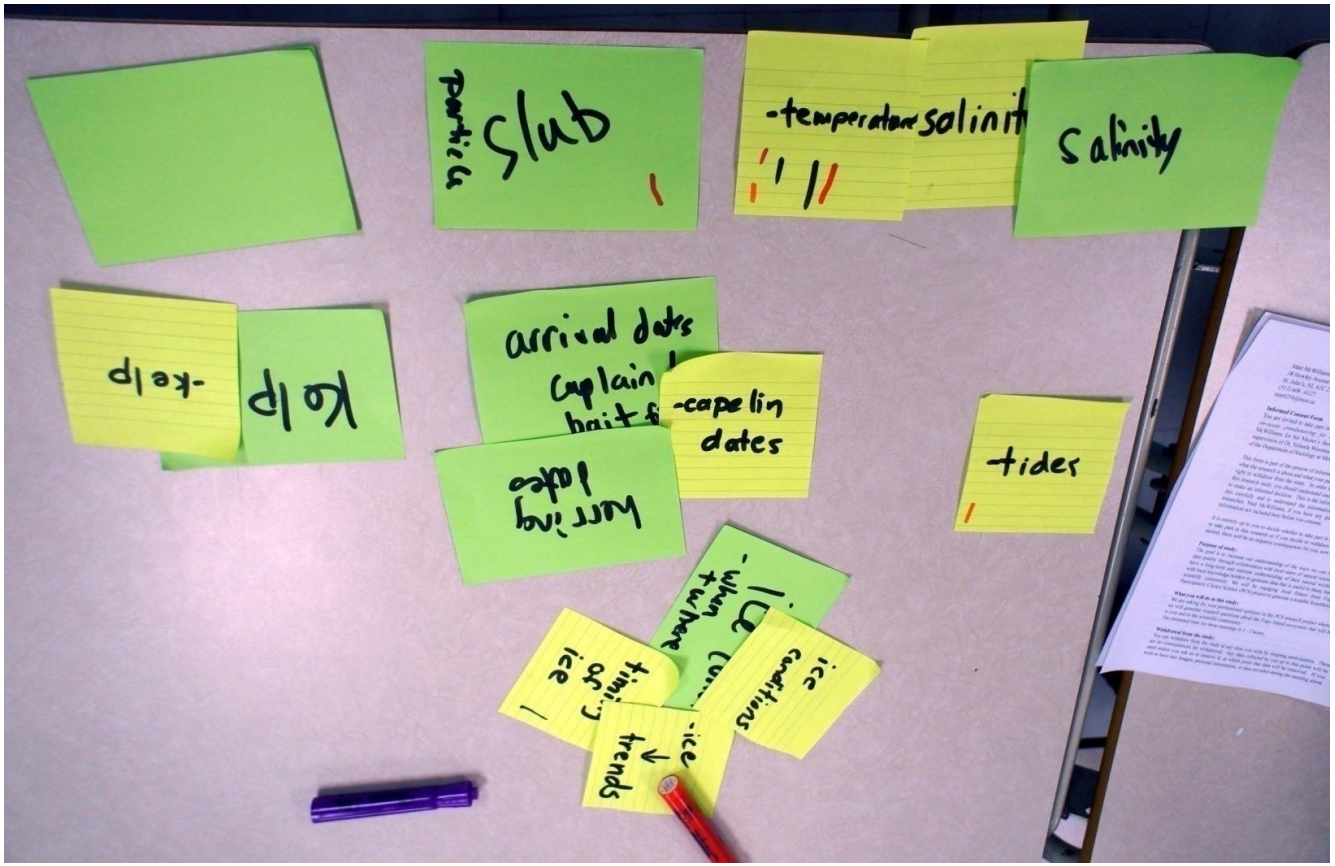


Adapted from https://www.dss.gov.au/our-responsibilities/housing-support/publications-articles/homelessness-youth/on-par-using-participatory-action-research-to-improve-early-intervention?HTML

Participatory Citizen Science

The Research Question

What effect does temperature have on different fisheries in terms of catch rates?



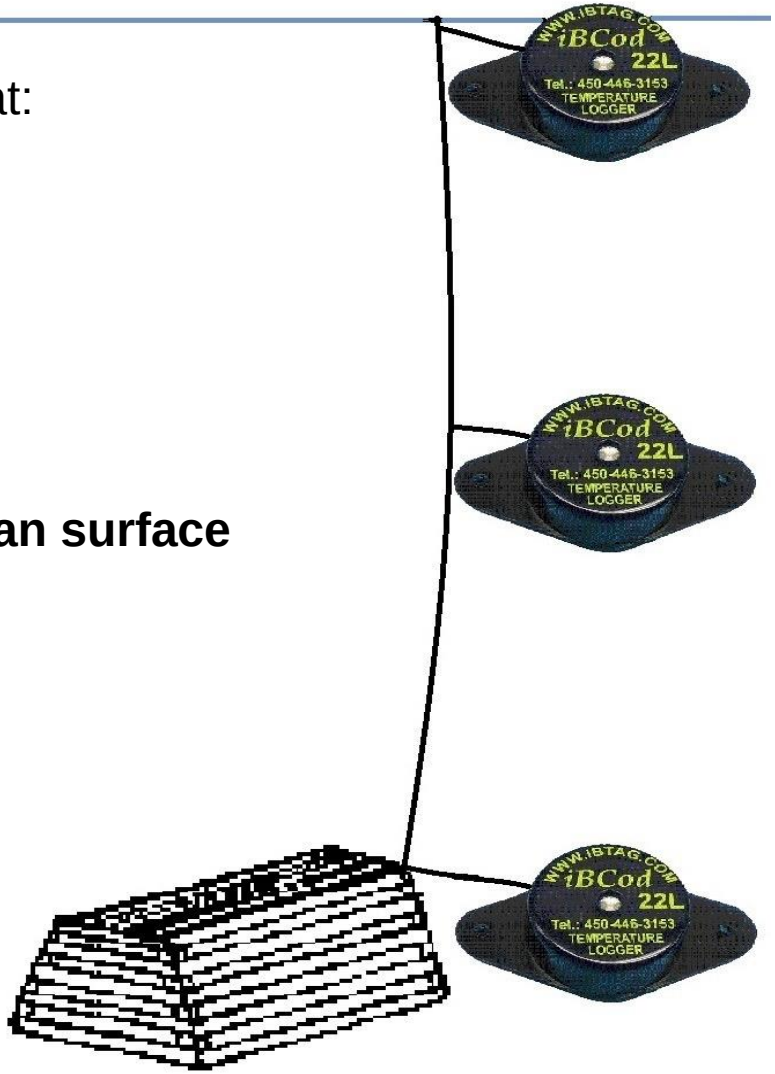
The Methods

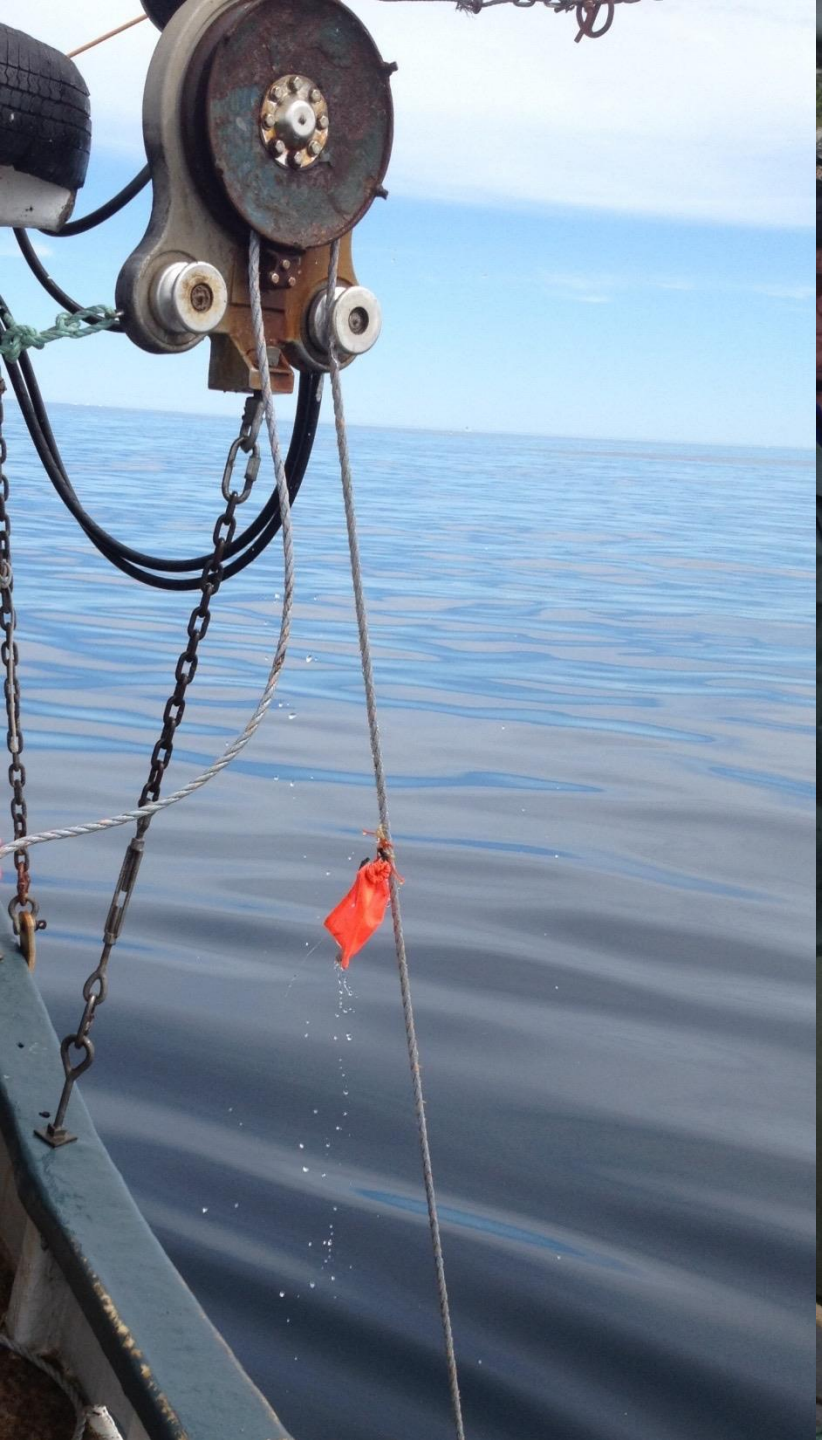
Arrange 3 iBCod temperature loggers placed at:

1. **The ocean surface**

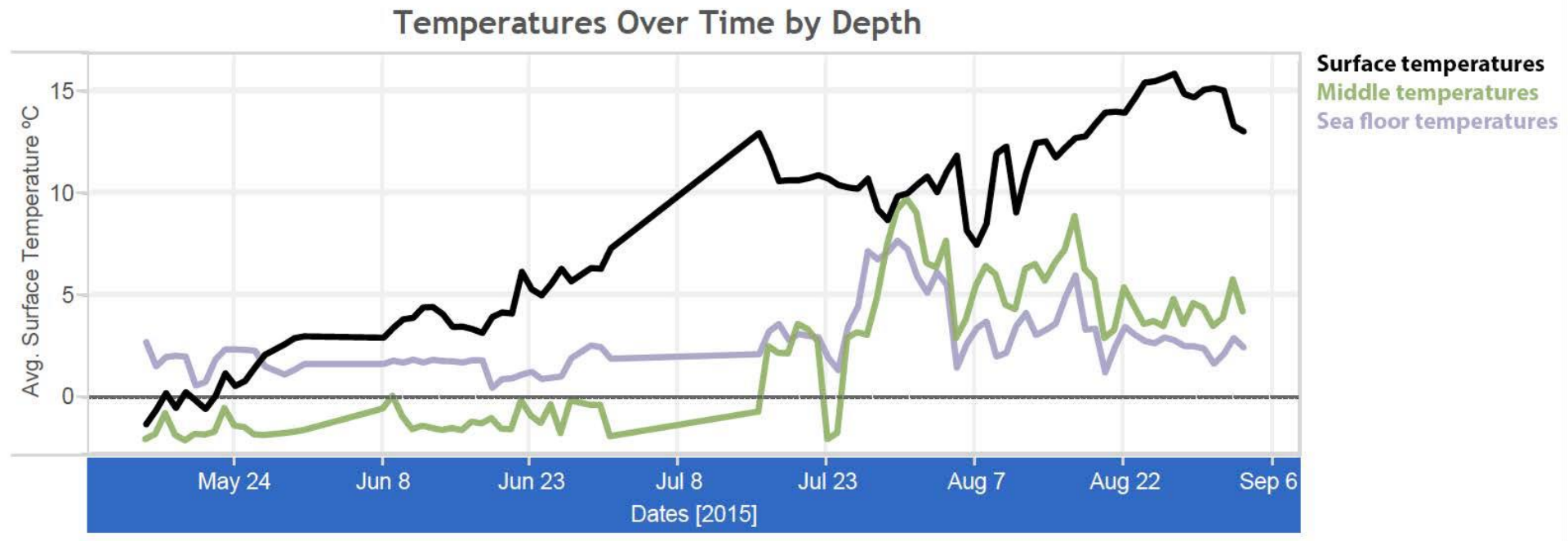
2. **Halfway between the seafloor and ocean surface**

3. **At the seafloor** attached to fishing gear





Participatory Analysis



Citizen Science?

What is the best use?

- what are the major needs that may be filled by citizen science?
- what can citizen science provide uniquely?

What is best practice?

- look to research ON citizen science

BioBlitz

BioBlitz Canada

BioBlitz

For more information

- Citizen Science Association (citizenscience.org)
- Cornell Lab
- citizenscientists.ca
- citizenscientistsleague.com
- citizensciencecenter.com

- I can share a link to a Dropbox folder
  - Key academic papers
  - List of resources
  - Extensive list of existing CS projects

Comments/Questions?

Thank you!

Yolanda Wiersma
Department of Biology, Memorial University

www.nllandscapeecology.com

ywiersma@mun.ca

@YolandaWiersma