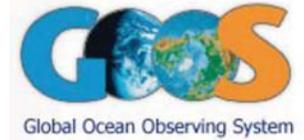




*Biological Survey of Canada
Halifax Nova Scotia, November 8, 2011*



Take Care of the Zeros and the Numbers Will Take Care of Themselves

Robert M. Branton

Ocean Tracking Network, Dalhousie University, Halifax, Canada.

bob.branton@dal.ca

www.oceantrackingnetwork.org



**NSERC
CRSNG**

SSHRC  CRSH

*Acknowledgements to Ocean Tracking Network Data Centre Staff:
Lenore Bajona, Susan Dufault, Brian Jones, Marta Mihoff.*

Bugs are everywhere!



In 1946, Grace Hopper joined the Harvard Faculty at the Computation Laboratory where she work on the [Mark II](#) and [Mark III](#). Operators traced an error in the Mark II to a [moth](#) trapped in a relay, coining the term *bug*. This bug was carefully removed and taped to the log book. Stemming from the first bug, today we call errors or glitch's [*sic*] in a program a *bug*... wikipedia

Today's Objective

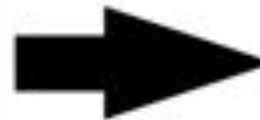
Provide insight into how modern computer systems can help biologists improve the way that they manage their data and in doing so advance the quality of their research.

Why do databases generally not include zero observations?

- Not recorded
- Better design
- Save space
- etc ...

Worksheet

taxon	1	2	3	4	5
site 1	0	1	0	0	0
site 2	0	0	0	0	0
site 3	0	4	0	0	0
site 4	0	0	0	2	0
Sum	0	5	0	2	0
Avg	0	1.25	0	0.5	0



Database

site	taxon	count
1	2	1
3	2	4
4	4	2

Is this a good thing?

Not if users leave zeroes out of their summary statistic calculations ...

```
> select * from bsc;
```

site	taxon	counted
1	2	1
3	2	4
4	4	2

```
> select taxon, sum(counted) as counted  
from bsc group by taxon;
```

taxon	counted	** correct result?
2	5	
4	2	

```
> select taxon, round(avg(counted),2) as counted  
from bsc group by taxon;
```

taxon	counted	** incorrect result!
2	5.00	
4	2.00	

```
> select * from sites;
```

site	** collection metadata
1	
2	

... same for taxa

How PostgreSQL can be used to generate missing zeroes ...

```
> select select taxon, round(avg(counted),2)  
from (  
select md.site, md.taxon,  
       coalesce(bsc.counted,0) as counted  
from bsc  
right outer join  
  (select * from sites, taxa) md  
on bsc.site=md.site and bsc.taxon=md.taxon  
) bsc_view  
group by taxon order by 1;
```

taxon	counted	** correct result!
1	0.00	
2	1.25	
3	0.00	
4	0.50	
5	0.00	

Cartesian Product of set X and set Y is the set that contains all ordered pairs (x, y) for which x belongs to X and y belongs to Y.

Outer Join does not require each record in the two joined tables to have a matching record.

Coalesce function accepts a list of parameters, returning the first non-Null value from a list

Likewise using R rep and match functions

```
> x<-data.frame(
+   site=c('s1','s3','s4'),
+   taxon=c('t2','t2','t4'),
+   count=c(1,4,2),
+   stringsAsFactors=FALSE)
> x
  site taxon count
1  s1    t2     1
2  s3    t2     4
3  s4    t4     2
>
> tapply(x$count,list(x$taxon),mean)

 t2  t4 ** incorrect result!
2.5 2.0
>
> sites<-c('s1','s2','s3','s4')
> taxa<-c('t1','t2','t3','t4','t5')
>
> x2<-data.frame(
+   cbind(
+     site=rep(sites,each=length(taxa)),
+     taxon=rep(taxa,length(sites)),
+     count=0),
+   stringsAsFactors=FALSE)
>
```

```
>x2[match(paste(x$site, '-',x$taxon),
          paste(x2$site, '-',x2$taxon)),
     'count']<-x$count
> x2
  site taxon count
1  s1    t1     0
2  s1    t2     1
3  s1    t3     0
...
12 s3    t2     4
...
19 s4    t4     2
20 s4    t5     0
>
> tapply(as.numeric(x2$count),list
(x2$taxon),mean)

 t1  t2  t3  t4  t5 ** correct result!
0.00 1.25 0.00 0.50 0.00
```

rep function replicates the values in x.

match function returns a vector of the positions of (first) matches of its first argument in its second.

What is the Ocean Tracking Network?



The Ocean Tracking Network is a Collaboration



Data policy...

A) **DEFINITIONS** - “Deployment Collaborators”, “Tracking Collaborators”, “Restricted Data” “Unrestricted Data” ...

B) **DATA SUBMISSION** - Deployment Collaborators are responsible for ensuring that data are uploaded in accordance with predefined OTN standards ...

C) **RESTRICTED DATA SUBMISSION** - Tracking Collaborators shall be entitled to require that access to the data which they upload to OTN initially be restricted ...

D) **ACCESS TO OTN DATA**

1. **Unrestricted Data** - Unrestricted public access to the OTN Portal is generally limited to static information products ...

2. **Restricted Data** - Those who wish to access OTN data other than what is provided via www.oceantrackingnetwork.org are required to become a registered member/user ...

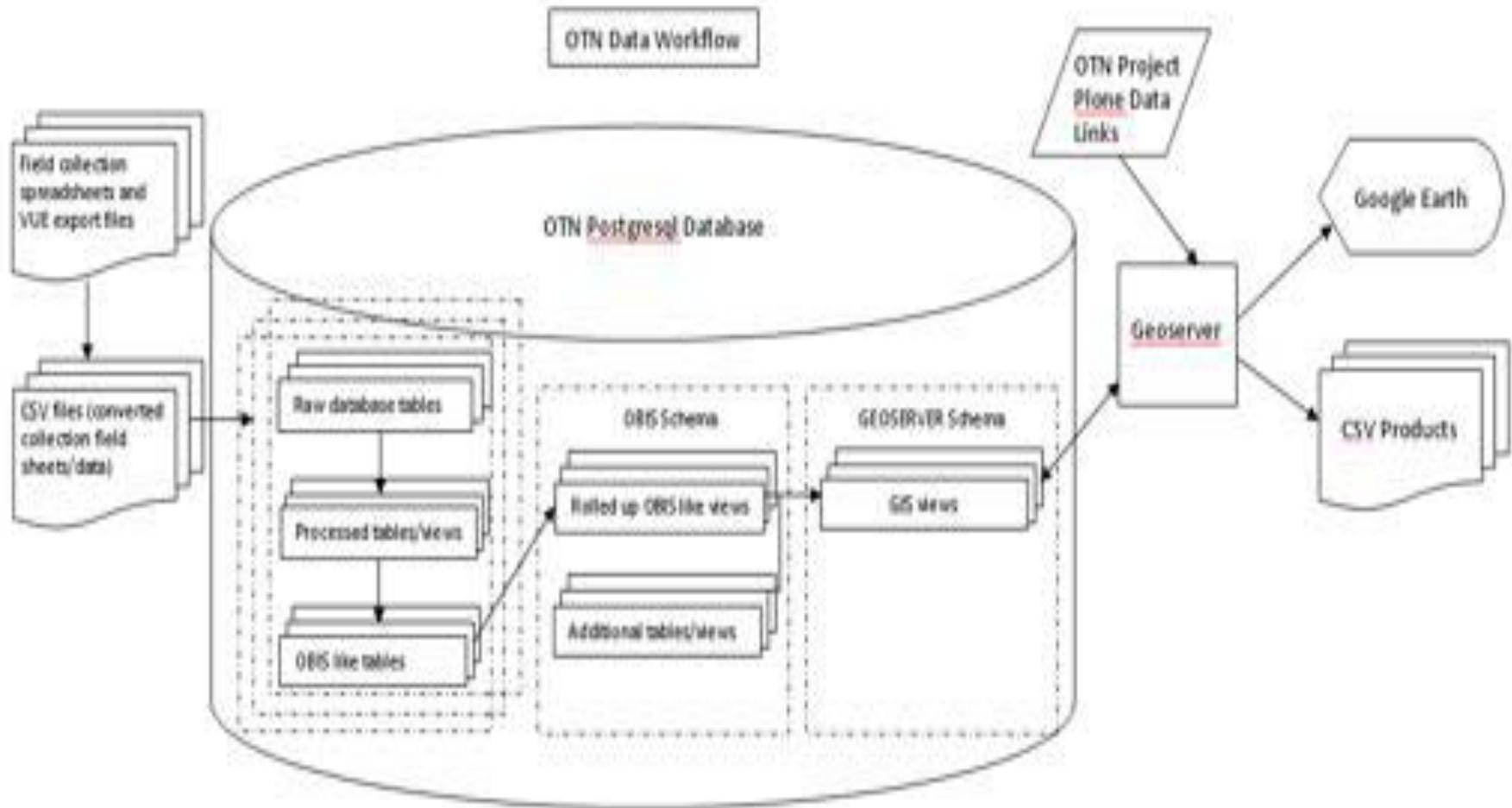
E) **USE OF OTN DATA** - All individuals who access OTN Data shall provide proper attribution to all Data Providers ...

F) **SUSTAINABILITY AND AVAILABILITY** - The OTN International Project Coordination Committee will be responsible for conducting periodic reviews of data management activities in each OTN region ...

G) **ACKNOWLEDGMENTS** - Australian Acoustic Tracking And Monitoring System (AATAMS), Canadian Department of Fisheries and Oceans (DFO), Pacific Ocean Shelf Tracking (POST) Project, Ocean Yearbook 22

<http://www.marinebiodiversity.ca/OTN/policies/OTN-Data-Policy>

OTN Information Factory



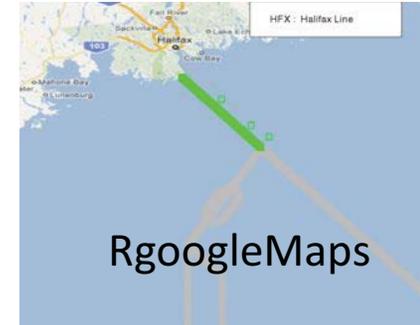
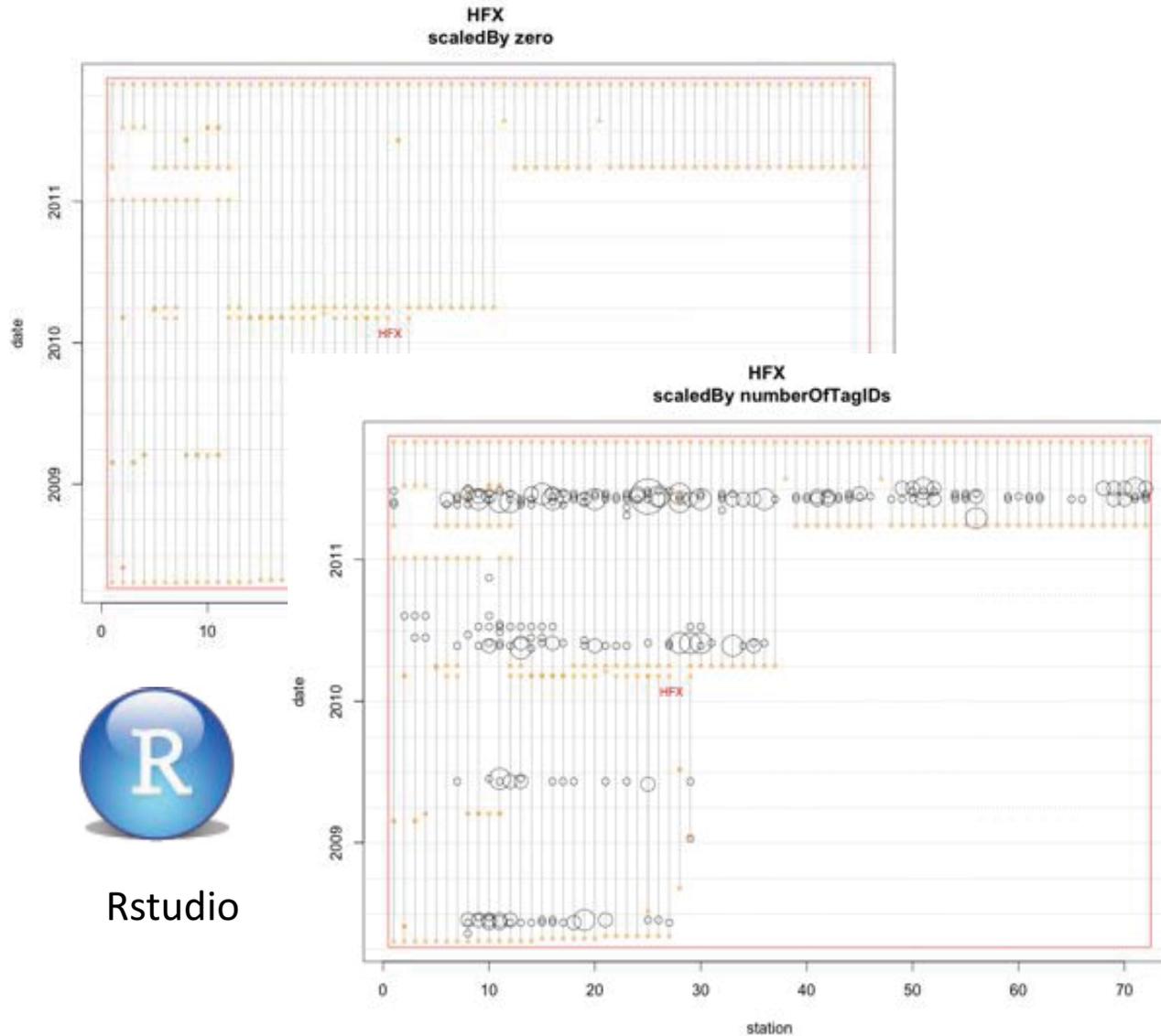
Individual collaborator's data are kept in their own password protected PostgreSQL schema.

Courtesy – Lenore Bajona

Overview of Methods

- Tracker and deployment operator metadata forms are prefilled with vendor instrument metadata
- Metadata and instrument logs are stored in a controlled access data management system
 - Public products like the GoogleEarth Flyover and Discovery Metadata are generated from the controlled access data
 - Trackers must post tag metadata to access detection data and are routinely informed when posted tags are detected
 - Trackers and deployment operators otherwise have full control on who may access on their data
- False and or mystery detections are flagged and retained for ongoing review and reprocessing.

OTN Collection Metadata



Rstudio

A screenshot of the "Manipulate" interface showing filter options for the data collection. The options are:

- oceanCode: all
- collectionCode: HFX
- trackerCode: all
- scientificName: all
- sensorType: all
- colorCodeBy: combined
- scaleBy: zero
- showLegend: bottomright
- showCollectionCodeLabels

Potential Strategies for Revitalizing the Biological Survey of Canada

Outreach

<http://oceantrackingnetwork.org/>

Public Portal ...

OCEAN TRACKING NETWORK
Headquartered at Dalhousie University, Canada

A new world standard for ocean research

Our climate is changing — of this we are sure. Marine life survival is becoming uncertain due to overfishing and changing migration patterns. Animals such as polar bears are becoming visibly anxious as their habitats begin to melt. Oceans are becoming warmer; the polar ice caps are melting.

The alarming thing is that we don't really know why. Information from beneath the sea's surface is very limited, despite the fact that continued human survival is directly linked to stable oceanic life.

The Ocean Tracking Network (OTN), a \$168-million conservation project, will put an end to this knowledge void. With it, thousands of marine animals around the world — from fish to birds to polar bears — will be tracked using acoustic telemetry technology. At the same time, we will be building a record of climate change — data that can be analyzed and then applied.

Headquartered at Dalhousie University, OTN unites leading ocean scientists around the globe. OTN is conducting the world's most comprehensive and revolutionary examination of marine life and ocean conditions, and how they are changing to a global standard for ocean management in a way

Members login

OTN Funding Partners:

- CFI
- NSERC
- SSHRC

OTN is a pilot project of:

- The Global Ocean Observing System

OTN is an affiliated project of:

- CENSUS OF MARINE LIFE

Copyright 2005-2009 Dalhousie University. All rights reserved. Disclaimer

Members portal ...

OCEAN TRACKING NETWORK

Home Data News Log in

You are here: Home

Login Name:
Password:

forgot your password? If you have forgotten your password

The Flex® Open Source CMS

Google Earth

RSS News Feed

About OTN

Other pages: About, Collaborators, Data, DataTypes, Centers, DataLastModified, Institutions, Keywords, Orphans, Regions, Species, Status, Taxonomy

About Ocean Tracking Network Data

Ocean Regions: 13
DateLastModified: 2011-10-28
TotalRecords: 5,014,323

Cumulative Record Count by BasisOfRecord and Time

MachineObservations
HumanObservations
MooringsRecords

Cumulative Record Count log10

TIME yyyy-mm

RSS News Feed

OTNGlobal: OTN Global

Click links to view species and/or related regions and projects descriptions at which point you may be asked to login. Request userids, report problems and/or make suggestions to otnd@dal.ca.

NW ATLANTIC: NW ATLANTIC Ocean

RSS feed Send this

Name World Only Data Report

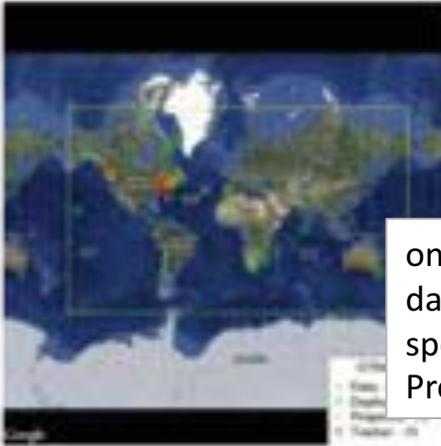
OTN project headquarters metadata (descriptive info, keywords, species etc)

Google

National / Regional Rollups

Title: OTNGlobal: OTN Global / Ocean Tracking Network Global Data Atlas
Citation: O'Dor, R., Alverson, K., Whoriskey, F., Branton, R. 2008. Ocean Tracking Network Global Metadata and Data Atlas. Retrieved: October 30, 2011 from www.oceantrackingnetwork.org
Status: Ongoing

Taxonomic Coverage: *Acipenser oxyrinchus oxyrinchus* / Atlantic sturgeon, *Dermodochelys coriacea* / leatherback turtle, *Gadus* / Atlantic cod, *Myoxocephalus scorpius* / shortfin mako, *Salvelinus alpinus* / Arctic char, *Salvelinus fontinalis* / brook trout, *Stenobothrus* / American eel, *Boreogadus saida* / arctic cod, *Phoca vitulina* / grey seal, *Morone saxatilis* / Atlantic salmon, *Renhardtius hippoglossoides* / halibut, *Thunnus maccoyii* / southern bluefin tuna

Geographic Coverage:  **byregion.htm**
 by Robert M. Branton -- last modified May 24, 2011 04:00 PM
 Other pages: [About](#), [Collaborators](#), [Collections](#), [CollaborationTypes](#), [Countries](#), [DataLastModified](#), [Institutions](#), [Species](#)
Ocean Region:
 OTNGlobal
 1. ARCTIC
 1. ACS: Cumberland Strait
 2. ALS: Lancaster Sound
 3. BLS: Bellefleur Strait
 4. BRS: Barrow Strait
 5. DSB: Disco Bay Area

Temporal Coverage: Start/End Dates: 1999-09-03 to 2011-09-25 - **bydatelastmodified.htm**
 by Robert M. Branton -- last modified May 24, 2011 04:00 PM
 Other pages: [About](#), [Collaborators](#), [Collections](#), [CollaborationTypes](#), [Countries](#), [DataLastModified](#), [Institutions](#), [Species](#)
DateLastModified:
 2011-05-20 HFX: Halifax Line
 2011-05-13 GMC: Gulf of Maine GoMOOS Buoys

Data Summary:

Activity Summary	Released	Detected	Total
Moorings - With Detectors / Total Deployed	1 / 304	4 / 3806	16 / 16430
Acoustic Tags - Released			
Satellite Tags - Released			
Qualified Acoustic Tags - Detected / Number of Deployments			
Sentinel Tags - Detected / Number of Deployments			
Orphaned Acoustic Tags - Detected / Number of Deployments			
Satellite Tags - Detected / Number of Deployments			

OTNGlobal Tag Detection Summary: Released=5379, Detected=1044, TotalDetections=3786
Basis Of Record Summary: HISTORY=801, MISSION=158, STATION=1157, RECEIVER=1492, DOWNLOAD=1315

byorphans.htm
 Other pages: [About](#), [Collaborators](#), [Collections](#), [CollaborationTypes](#), [Countries](#), [DataLastModified](#), [Institutions](#), [Species](#), [Status](#), [TotalRecords](#)
Orphans:
 185 - WRS: NS Southern Upland Salmon Tagging
 21 - HFX: Halifax Line
 21 - MPD: Minas Passage
 20 - JDE: U.S. Laysan American Eel tag releases
 20 - MPS: Minas Passage
 17 - ALS: Lancaster Sound: tag release and receiver...
 6 - BRT: Perth Line
 4 - ERQ: Frobisher Bay, Canada

Records:
 Total records:
 820055 - WRS: NS Southern Upland Salmon Tagging
 211572 - SGS: Sable Island Grey Seals
 205745 - ALS
 90696 - CSTM
 80625 - HFX
 74173 - GEORG
 67441 - JDE: U.S. Laysan American Eel tag releases
 31130 - CRO: Canada
 27834 - MPD
 17088 - ERQ: Frobisher Bay, Canada
 14480 - MPS
 5428 - BRT: Perth Line
 3714 - VP: Viti Levu
 2139 - JRC: JRC

ongoing projects – 40
data records – 5,014,323
species - 24
Proposed/planned projects – 25

The Ocean Tracking Network (OTN) is a global project headquartered at Dalhousie University in Halifax, Nova Scotia, Canada. OTN researchers worldwide are pressing for standardization of biological and physical oceanographic information. Where animals are, what conditions they experience, how they interact, and how individual behaviours change.

Individual Project Reporting

Contact to the designated project leader*

Links to other projects with tags released by and or detected by this projects .

Links to data portal and various other project websites*

Don Bowen

If you want to contact this author, fill in the form below.

Feedback for author

Subject *

Message *

Send

Latest content created by this user

May 13, 2011 SCS: Sable Island Grey Seals

All content created by Don Bowen.

Other Functionality

- Descriptive information as given by principal investigator are static, while all other details are routinely derived automatically from OTN's PostgreSQL database.
 - Citations, rows, columns & lists all expand and contract depending on type and volume contained in the particular project and series ...
- Appears in same form on GoogleEarth, Plone CMS & PDF whereas styling adapts accordingly.
- Can be displayed singly for ad-hoc viewing or extracted en-mass for periodic reporting to sponsor agencies

*Access to data is controlled by the designated project leader.

Standards Based Sustainability and Availability

The screenshot shows the Fisheries and Oceans Canada website. The main heading is "Fisheries and Oceans Canada" with the URL "www.dfo-mpo.gc.ca". Below this is a navigation menu with "Home", "Contact Us", "Help", "Search", and "canada.gc.ca". The main content area is titled "Integrated Science Data Management (ISDM)" and includes a "Data Request Form" link. A sidebar on the left lists various topics and regions. The main text describes ISDM's mandate to manage and archive data collected by DFO or acquired through international programmes.



www.iode.org

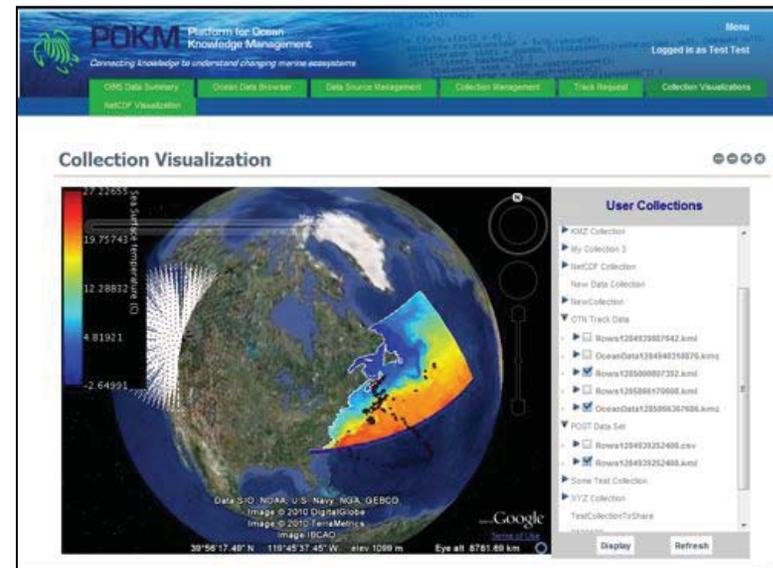
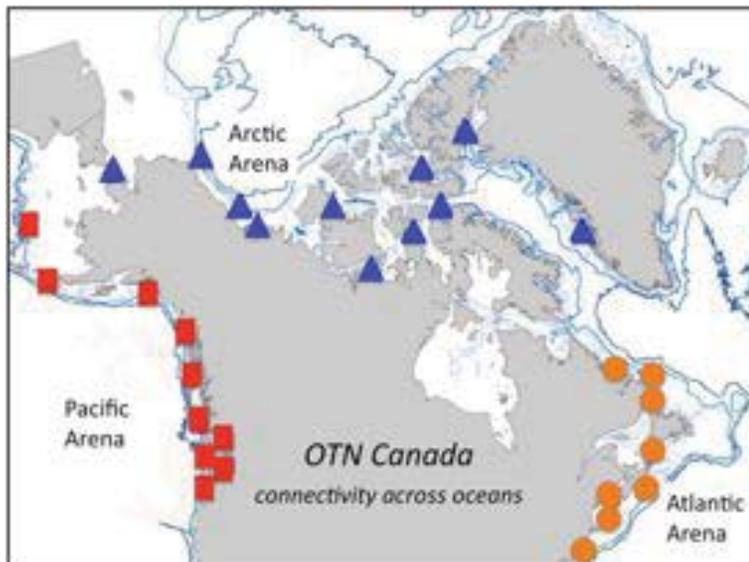
The screenshot shows the OBIS website interface. The main heading is "OBIS OCEAN BIOGEOGRAPHIC INFORMATION SYSTEM". The page includes a search bar, navigation links, and a map of the Bay of Fundy region. A red box highlights a specific data entry for "GTN/DFO Bay of Fundy Spiny Dogfish". The entry includes details such as "GTN/DFO Marine Spiny Dogfish Tagging" and "GTN/DFO Marine Spiny Dogfish Tagging (OBIS Canada)".

www.meds-sdmm.dfo-mpo.gc.ca/isdm-gdsi/index-eng.html

http://iobis.org/mapper/?resource_id=2308

Collaboration with related initiatives ...

OTN Canada - Understanding Species Movements, Interactions, and Environmental Variability Across Canada's Three Oceans (NSERC/SSHRC)



POKM – Platform for Ocean Knowledge Management (CANARIE)

POST – Pacific Ocean Shelf Tracking (CoML)

Neptune - North-East Pacific Time-series Undersea Network Experiments (CFI)

CHONe – Canadian Healthy Ocean Network (NSERC)

FSRS – Fishermen and Scientist Research Society (NSERC)





Discussion