



**CoastGIS 2021**

Novia University, Raseborg Finland  
Sept 16-17, 2021



# **‘Advancing Sustainable Coastal Planning through Effective Use of Open-Access Biodiversity Information Systems’**

*Where getting names right always matters!*

bob.branton@gmail.com

[GitHub.com](https://github.com)

[OceanExpert.org](https://oceanexpert.org)

## Outline ...

- Present situation ...
  - Ecosystem change
  - Information infrastructure
  - Challenges
  - Getting names right
- Using data science ...
  - Data Preparation
  - Data Visualisations
  - Control panel / Dashboard
- Recap
  - Software components
  - Benefits
  - More Challenges
- Conclusion
  - What's next
  - Collaboration
- Demos

# Ecosystem Change

## Invasive Species

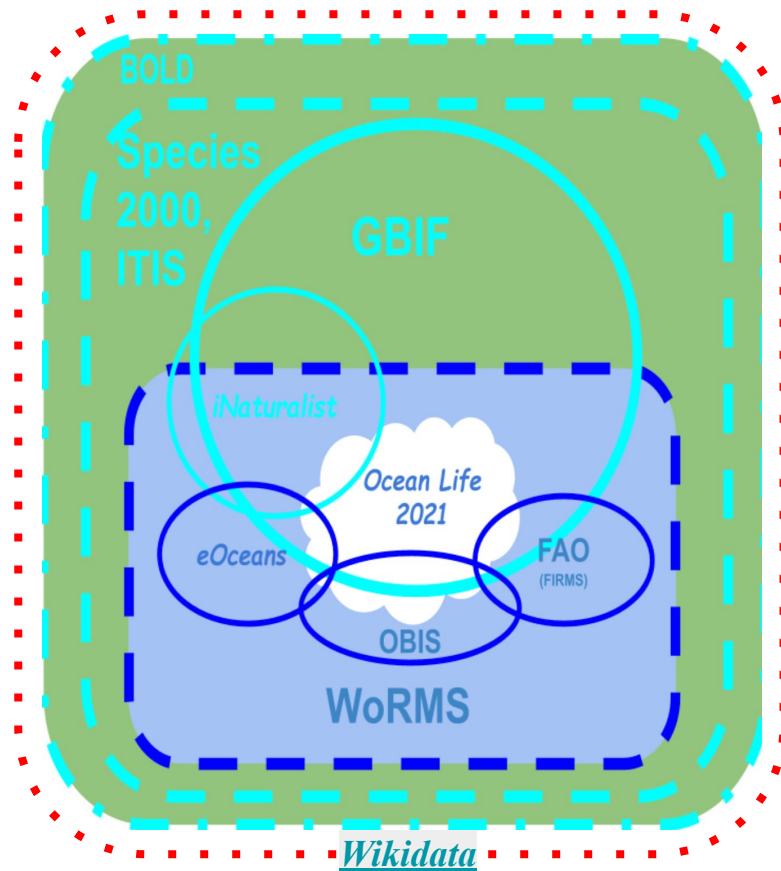
**"Removing green crab proves effective in controlling invasive species."**



HALIFAX, 2016 ... trap invented by a local fisherman was used to catch more than two million green crab [*Carcinus maenas* Linnaeus, 1758] from one estuary over a few years. Since then, eelgrass [*Zostera L.*] in that estuary has recovered by about 34 per cent and soft-shell clam [*Mya arenaria* Linnaeus] populations are on the rise ...

[Atlantic CTV News](#)

# Information Infrastructure



# Challenges

Scientific names change over time\*

- 55 of 73 Ocean Life 1859 scientificNames required updating ...

Data are not all in one place ..

- 63 of 78m OBIS obs (~80%) and 13 of 40m iNat Research Grade obs (~33%) not on GBIF

WoRMS/OBIS.authority not = GBIF.authority

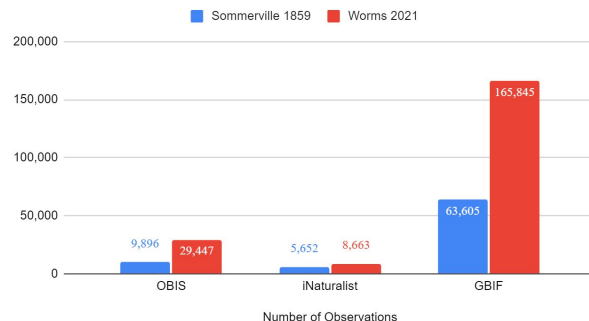
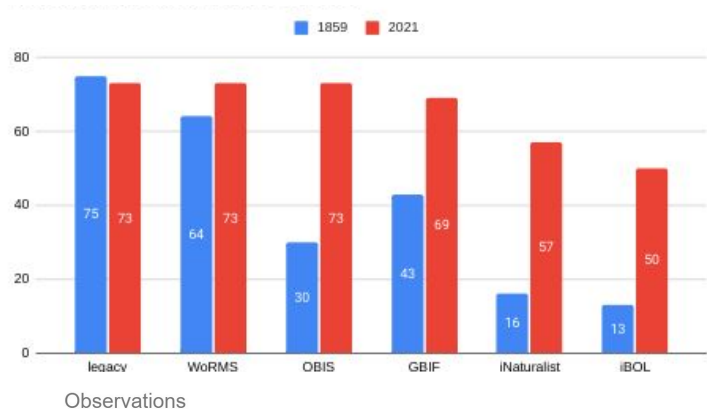
- 26 of 68 GBIF.authorities contained abbreviations, 14 of 68 GBIF.authorities did not include year . The GBIF search window rejected WoRMS/OBIS.scientificName that contained dates

Inconsistent approach between portals regarding data queries, downloads & citations

OBIS and GBIF downloads often too large for Google Sheets

# Getting Names Right

Using World Registry of Marine Species (WoRMS) accepted names greatly improve query results ...



[See Oceanlife-2021/worms](https://oceanlife-2021.worms.org/)

# Using Data Science ...

\* [ocean.si.edu/ocean-life/5-invasive-species-you-should-know](https://ocean.si.edu/ocean-life/5-invasive-species-you-should-know)



Smithsonian

## Species

[Carcinus maenas](#) [European Green Crab]\*

[Caulerpa taxifolia](#) [Killer Algae]\*

[Codium fragile](#) [Dead Man's Fingers]

[Dreissena polymorpha](#) [Zebra Mussel]\*

[Mnemiopsis leidyi](#) [Sea Walnut]\*

[Pterois volitans](#) [Lion Fish]

[Rapana venosa](#) [Veined Rapa Whelk]\*

## Primary data products ...

### Summary ...

	prov	species	year	date	OCCS	DUPS	longitude	latitude	eez	fao
	<chr>	<chr>	<int>	<date>	<dbl>	<dbl>	<dbl>	<dbl>	<chr>	<chr>
8687	UNIQUE	Carcinus maenas	2018	2018-09-28	1	3	-122.30957	37.86517	USA	99
11319	UNIQUE	Carcinus maenas	2017	2017-07-25	2	4	-127.76165	51.48631	CAN	67

See: [species\\_style.r on GitHub](#)

\* <https://docs.ropensci.org/spocc/>



## Providers

[ala](#) - Atlas of Living Australia [AUS]\*

[bison](#) - Biodiversity Information Serving Our Nation [USA]\*

[eBird](#) - Cornell Lab of Ornithology {Global}

[gbif](#) - Global Biodiversity Information System [Global]\*

[idigbio](#) - Integrated Digitized Biocollections [USA]\*

[inat](#) - iNaturalist [Global]\*

[obis](#) - Ocean Biodiversity Information System [Global]\*

[Vertnet](#) - Distributed databases with a backbone

### Expanded / hypercube ...

	species	prov	year	eez	fao	area
	<chr>	<chr>	<int>	<chr>	<int>	<chr>
1	Carcinus maenas	ala	1916	AUS	99	-38 145.1....
2	Carcinus maenas	ala	1916	AUS	99	-38 145.1....

See: [prepare\\_data.r on GitHub](#)

\* <https://www.marineregions.org/>



## FAO\*

Boundaries of the FAO Fishing Areas

See: <https://www.marineregions.org/sources.php#fao>

## EEZ\*

Marineregions: the intersect of the Exclusive Economic Zones and IHO ocean areas

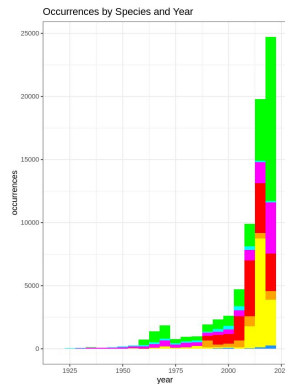
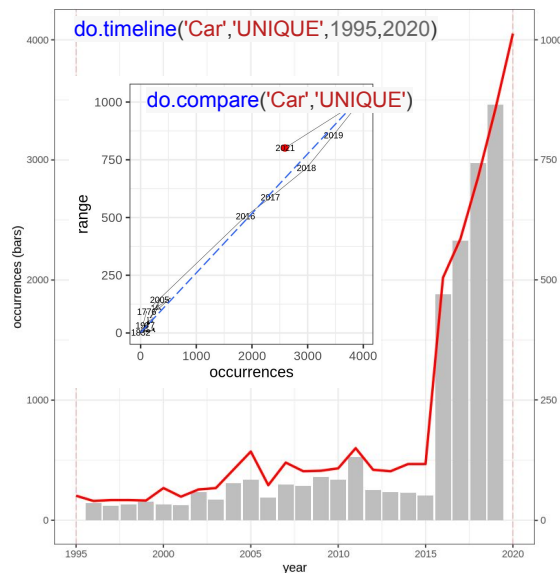
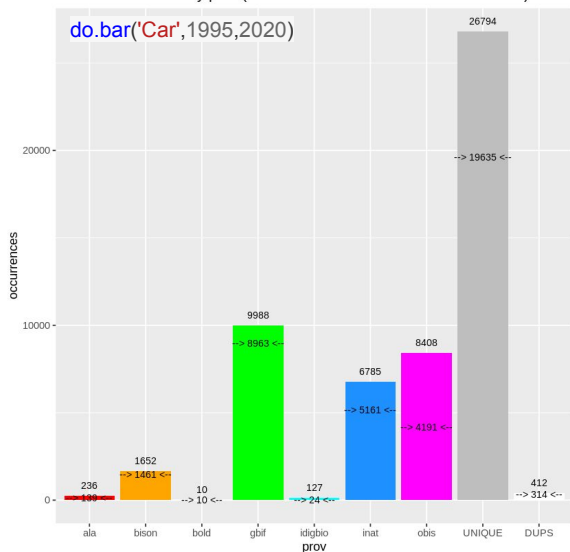
See: <https://www.marineregions.org/sources.php#ihoez>

# Visualizations ....

	ala	bison	bold	gbif	idigbio	inat	obis	UNIQUE	DUPS	total
<b>Carcinus maenas</b>	236	1652	10	9988	127	6785	8408	26794	412	54412
<b>Caulerpa taxifolia</b>	501	248	0	1355	403	55	637	2432	767	6398
<b>Codium fragile</b>	454	1268	0	6006	895	2341	759	10345	1378	23446
<b>Dreissena polymorpha</b>	0	4717	20	9979	514	1284	769	15185	2098	34566
<b>Mnemiopsis leidyi</b>	0	87	0	1166	12	262	2314	3479	362	7682
<b>Pterois volitans</b>	515	1213	0	9999	282	1523	2602	15175	959	32268
<b>Rapana venosa</b>	0	0	1	272	10	191	178	627	25	1304
<b>total</b>	1706	9185	31	38765	2243	12441	15667	74037	6001	160076

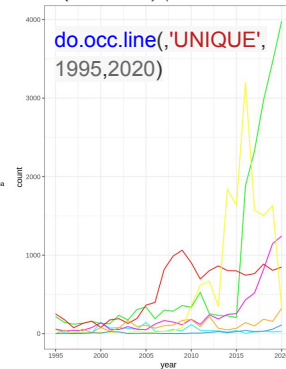
Carcinus maenas - by prov (all: 1776-2021 vs subset: -->1995-2020<--)

`do.bar('Car', 1995, 2020)`



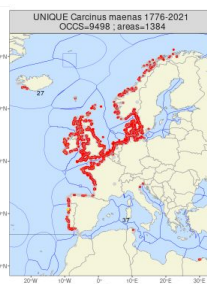
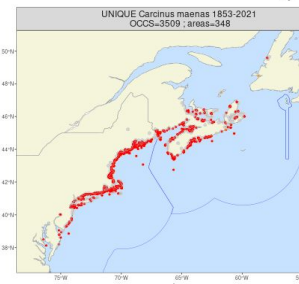
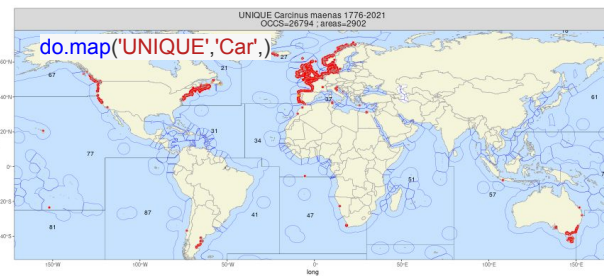
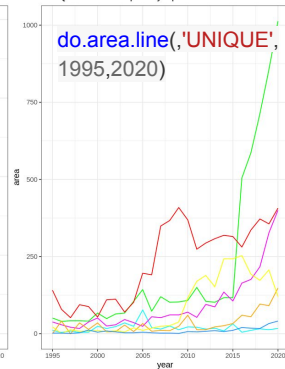
UNIQUE Occurrences by Species and Year

`do.occ.line('UNIQUE', 1995, 2020)`



UNIQUE Area Occupied by Species and Year

`do.area.line('UNIQUE', 1995, 2020)`



See: <https://ggplot2-book.org/introduction.html>

See: [Analytics.r on GitHub](#)

# Dashboard / Control Panel

Live Access: <https://rmbranto.shinyapps.io/invasives/>

## Invasive marine species ...

select species

Carcinus maenas

select provider

UNIQUE

year:



☒ zoom to data:

select fao

☒ fao labels:

select eez

☐ eez labels:

select data

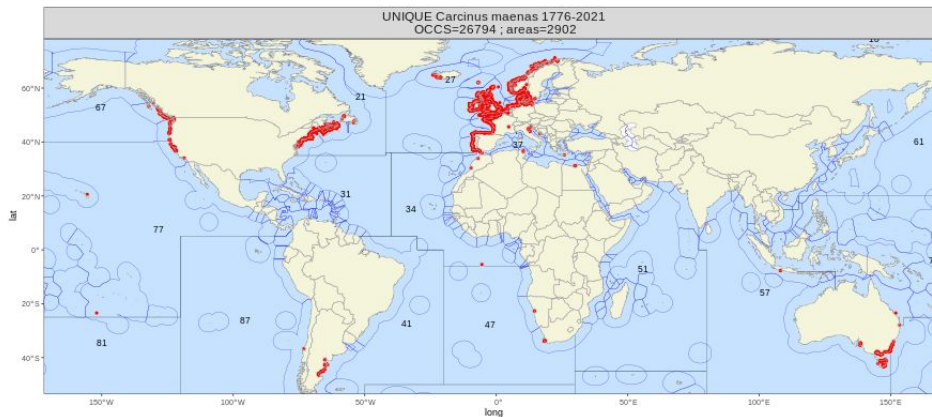
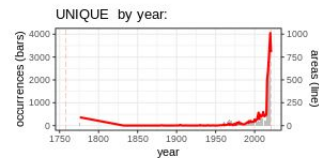
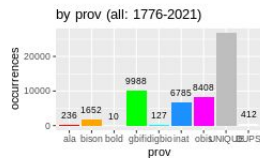
summarized

Download



\*\* See Wikidata TAXON INFO \*\*

spocc: *Carcinus maenas* [European Green Crab]



Provider home page links: [Worms](#) - [ala](#) - [bison](#) - [gbif](#) - [idigbio](#) - [inat](#) - [obis](#)



# Recap ...



## **Software Components:**

- [R](#): A highly programmable free computer environment under [GNU General Public License](#);
- [Jupyter](#): Combine software code, computational output & explanatory text into a single document;
- [Spocc](#) - (SPecies OCCurrence) query & collect species occurrence data from many sources;
- [ggplot](#) - a general scheme for creating graphs and maps;
- [ShinyApp](#): build interactive web apps straight from R;
- [GitHub](#); where the world builds software;

## **Features and benefits:**

- Queries species occurrence data from available providers using list of WoRMS accepted names
- Generates synthetic UNIQUE / DUPLICATE data objects from available provider data
- Generates provider specific Taxon IDs and Keys and so formulate the deep links into live provider portals
- Filter and zoom maps and graphs using list of Providers, date range, FAO fishing Area and EEZs

## **Challenges:**

- Spocc queries limited to 9999 occurrence records per species and provider
- bison and inigbio providers do not support deep linking by species and or taxon key/id
- using EEZ and FAO shapefiles for Pacific centred mapping

# Last slide

## Conclusion

Continue preparing for IOI 2022 ...

- More testing and refinement
- Add provider API links to dashboard
- Experiment with single species control panel
- Creation and use of defined area polygon queries
- Optional eez/fao and/or

Seek collaborations ...

- marine plants, trawl bycatch species lists
- Create vignette for OpenSci
- Create wikipedia content

## Demos

[Dashboard](#)

[GitHub](#)

[Hyperslide](#)

[Wikidata](#)

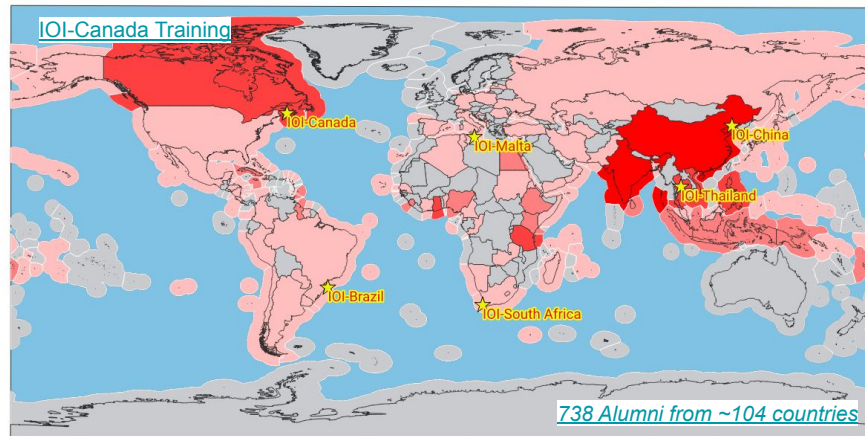
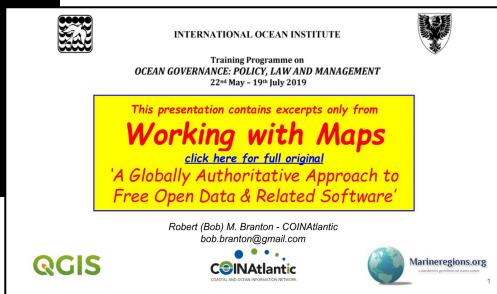


# Appendix

# History of Marine Biodiversity Information ...

- 1758** Linnaeus publishes 10th edition of [Systema Naturae](#)
- 1832-42** US Ex Ex and **1872-76** HMS Challenger global ocean expeditions;
- 1981** IOI-Canada initiates training program;
- 1982** International Law of the Sea declares exclusive economic zones (EEZs);
- 1992** Convention on Biological Diversity;
- 2000-10** Census of Marine Life creates Ocean Biogeographic Information System (OBIS);
- 2011** International Ocean Data Exchange (IODE) initiates the Steering Group for OBIS;
- 2020** OBIS celebrates 20 years & changes name to 'Ocean Biodiversity Information System'.

My recent IOI Canada training materials ....



# Synthetic Data Objects

species.style

A data.frame: 7 × 5

Names	cNames	sColors	wikidata	filimages
<chr>	<chr>	<chr>	<chr>	<chr>
Carcinus maenas	European Green Crab	green	Q27779	https://upload.wikimedia.org/wikipedia/commons/1/17/Carcinus_maenas.jpg
Caulerpa taxifolia	Killer Algae	cyan	Q310961	https://upload.wikimedia.org/wikipedia/commons/e/e7/CaulerpaTaxifolia.jpg
Codium fragile	Dead Mans Fingers	magenta	Q2712208	https://upload.wikimedia.org/wikipedia/commons/e/ed/Codiumfragile.jpg
Dreissena polymorpha	Zebra Mussel	red	Q752130	https://upload.wikimedia.org/wikipedia/commons/thumb/a/a9/Dreissena_polymorpha3.jpg/220px-Dreissena_polymorpha3.jpg
Mnemiopsis leidyi	Sea Walnut	orange	Q133630	https://upload.wikimedia.org/wikipedia/commons/thumb/1/1e/Sea_walnut%2C_Boston_Aquarium.jpg/220px-Sea_walnut%2C_Boston_Aquarium.jpg
Pterois volitans	Lion Fish	yellow	Q824672	https://upload.wikimedia.org/wikipedia/commons/thumb/b/bf/Pterois_volitans_Manado-e_edit.jpg/220px-Pterois_volitans_Manado-e_edit.jpg
Rapana venosa	Veined Rapa Whelk	dodgerblue	Q139053	https://upload.wikimedia.org/wikipedia/commons/thumb/f/fb/Rapana_Black_Sea_2008_G1.jpg/220px-Rapana_Black_Sea_2008_G1.jpg

head(prov.keys)

A data.frame: 6 × 7

	species	ala	bison	gbif	idigbio	inat	obis
	<chr>	<chr>	<chr>	<chr>	<chr>	<chr>	<chr>
1	Carcinus maenas	284294	98734	5178595	5178595	52523	107381
2	Caulerpa taxifolia	83052	6974	2643172	2643172	50919	144476
3	Codium fragile	84082	6897	5272096	5272096	67555	370562
4	Dreissena polymorpha	NA	81339	2287072	2287072	116340	181566
5	Mnemiopsis leidyi	NA	53917	2501248	2501248	180788	106401
6	Pterois volitans	189453	166883	2334438	2334438	47280	159559

head(df.prov[df.prov\$DUPS>2&df.prov\$species=='Carcinus maenas',])

A data.frame: 6 × 10

	prov	species	year	date	OCCS	DUPS	longitude	latitude	eez	fao
	<chr>	<chr>	<int>	<date>	<dbl>	<dbl>	<dbl>	<dbl>	<chr>	<chr>
8687	UNIQUE	Carcinus maenas	2018	2018-09-28	1	3	-122.30957	37.86517	USA	99
11319	UNIQUE	Carcinus maenas	2017	2017-07-25	2	4	-127.76165	51.48631	CAN	67
11320	DUPS	Carcinus maenas	2017	2017-07-25	6	3	-127.76165	51.48631	CAN	67
23906	UNIQUE	Carcinus maenas	2016	2016-08-18	1	3	-60.33113	46.80721	CAN	99
27236	UNIQUE	Carcinus maenas	2017	2017-10-13	1	3	-63.93700	44.51300	CAN	21
30240	UNIQUE	Carcinus maenas	2017	2017-05-31	1	3	-66.89060	45.03820	CAN	21

head(df.exp)

A data.frame: 6 × 6

	species	prov	year	eez	fao	area
	<chr>	<chr>	<int>	<chr>	<int>	<chr>
1	Carcinus maenas	ala	1916	AUS	99	-38.145.1....
2	Carcinus maenas	ala	1916	AUS	99	-38.145.1....
3	Carcinus maenas	ala	1955	AUS	99	-37.6.149.8..
4	Carcinus maenas	ala	1957	GBR	99	50.3.-4.1....
5	Carcinus maenas	ala	1958	AUS	57	-38.4.144.9..
6	Carcinus maenas	ala	1958	AUS	57	-38.7.145.7..

prov.style

A data.frame: 9 × 4

prov	id	color	order
<chr>	<chr>	<chr>	<dbl>
UNIQUE	UNI	grey	8
DUPS	DUP	white	9
ala	ala	red	1
bison	bis	orange	2
bold	bol	yellow	3
gbif	gbi	green	4
idigbio	idi	cyan	5
inat	ina	dodgerblue	6
obis	obi	magenta	7