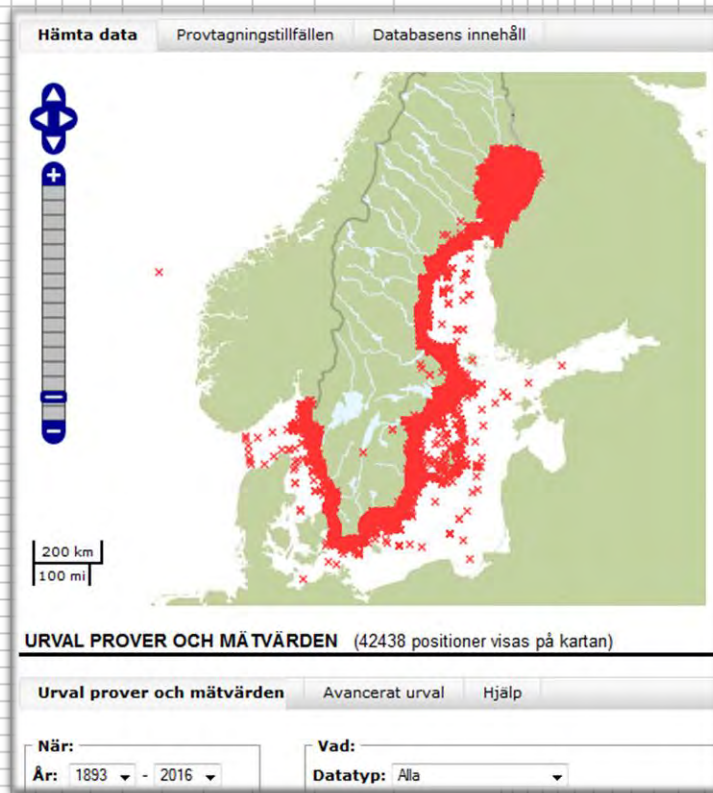


Swedish Meteorological and Hydrological Institute

SMHI

Patrik Strömberg Arnold Andreasson Nils Nexelius

Data harvesting



International conference on Marine Data and Information Systems 2016 Gdansk

Swedish Meteorological and Hydrological Institute

The end of reporting deadlines? Optimising dataflow by harvesting – a case study using open source REST- server technology from SMHI in corporation with ICES



Patrik Strömberg

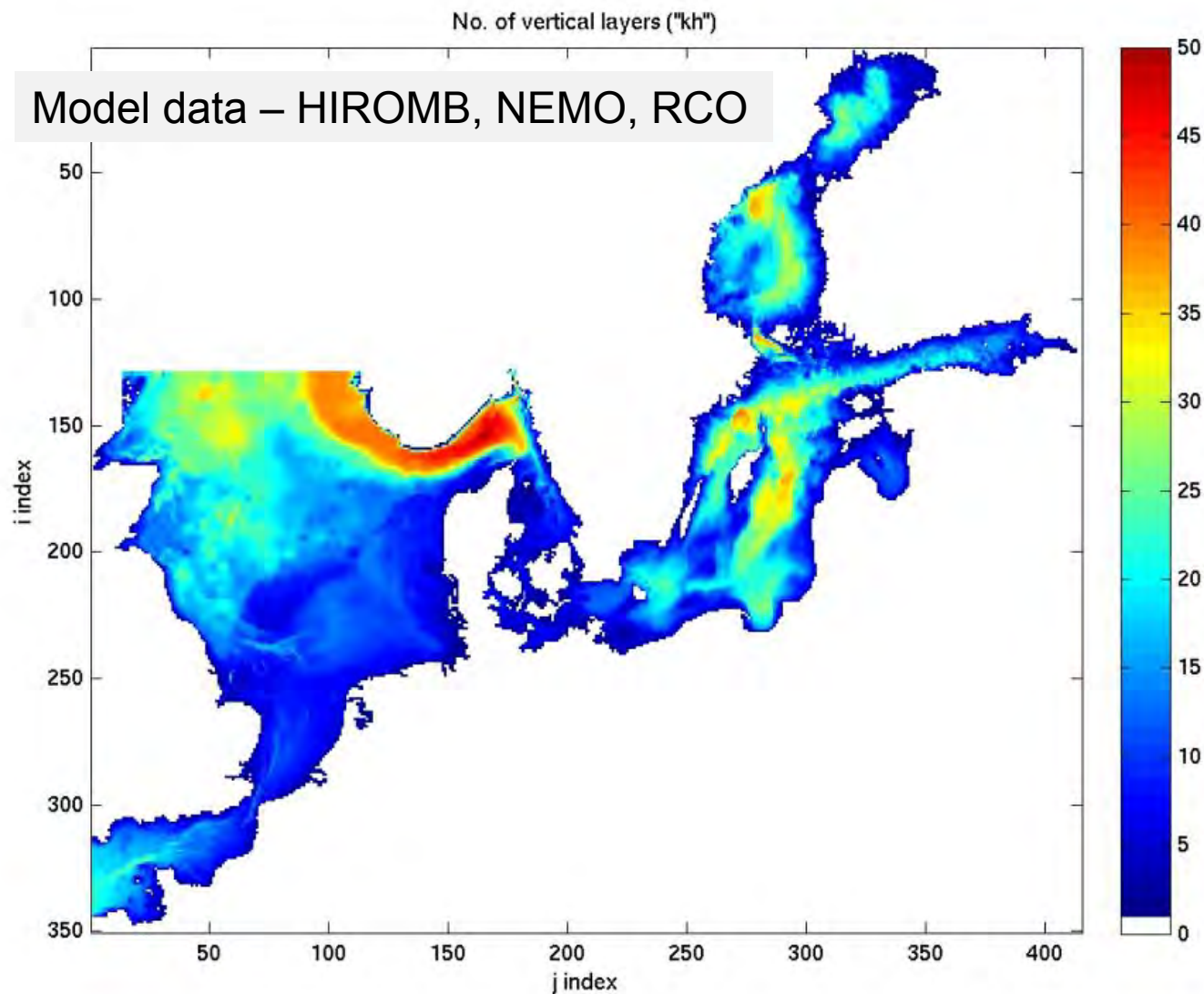
International conference on Marine Data and Information Systems 2016 Gdansk



From hunters and collectors to farmers

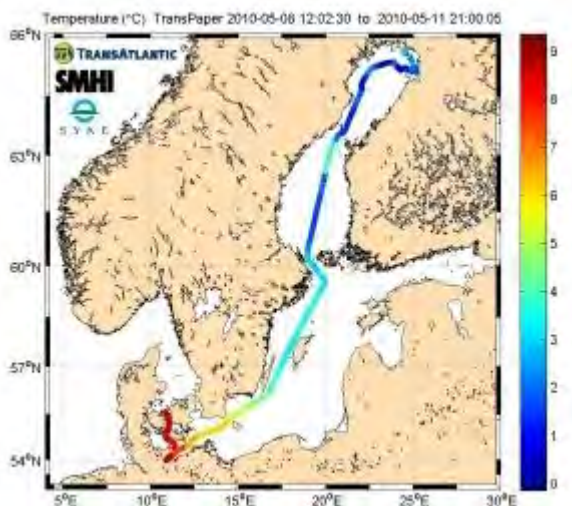


SMHI has huge amounts of data for (mostly) the Baltic sea area



SMHI has huge amounts of data for the Baltic sea area

Ferrybox



Buoys



2014-03-25

Klicka på en station för tidsdiagram

Mätplats: Väderöarna

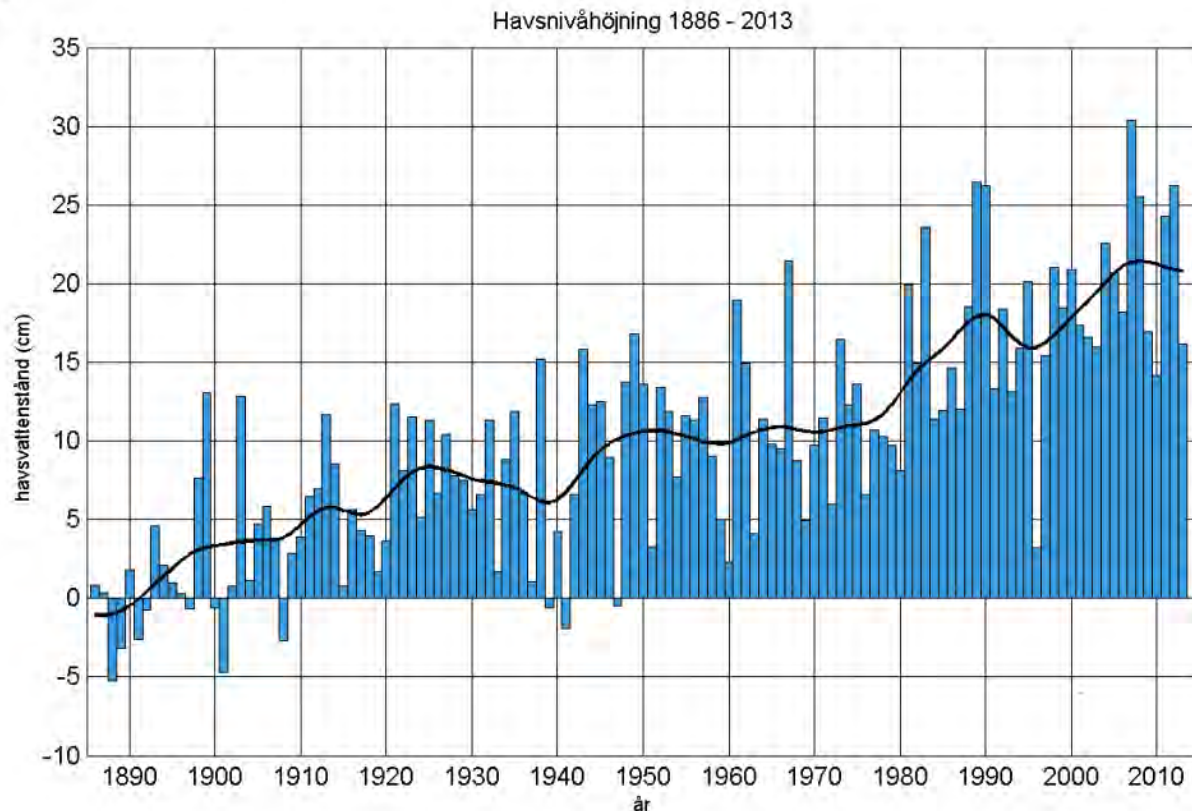
Lokal svensk tid Djup	Klorofyll-fluorescens µg/l 1m	Turbiditet NTU 1m
12:00	0.6	4.2
13:00	0.6	4.2
14:00	0.6	4.1
15:00	0.6	4.1

- Observation saknas

[Senaste 24 timmarnas marinbiologi](#)

SMHI has huge amounts of data for the Baltic sea area

Sea level



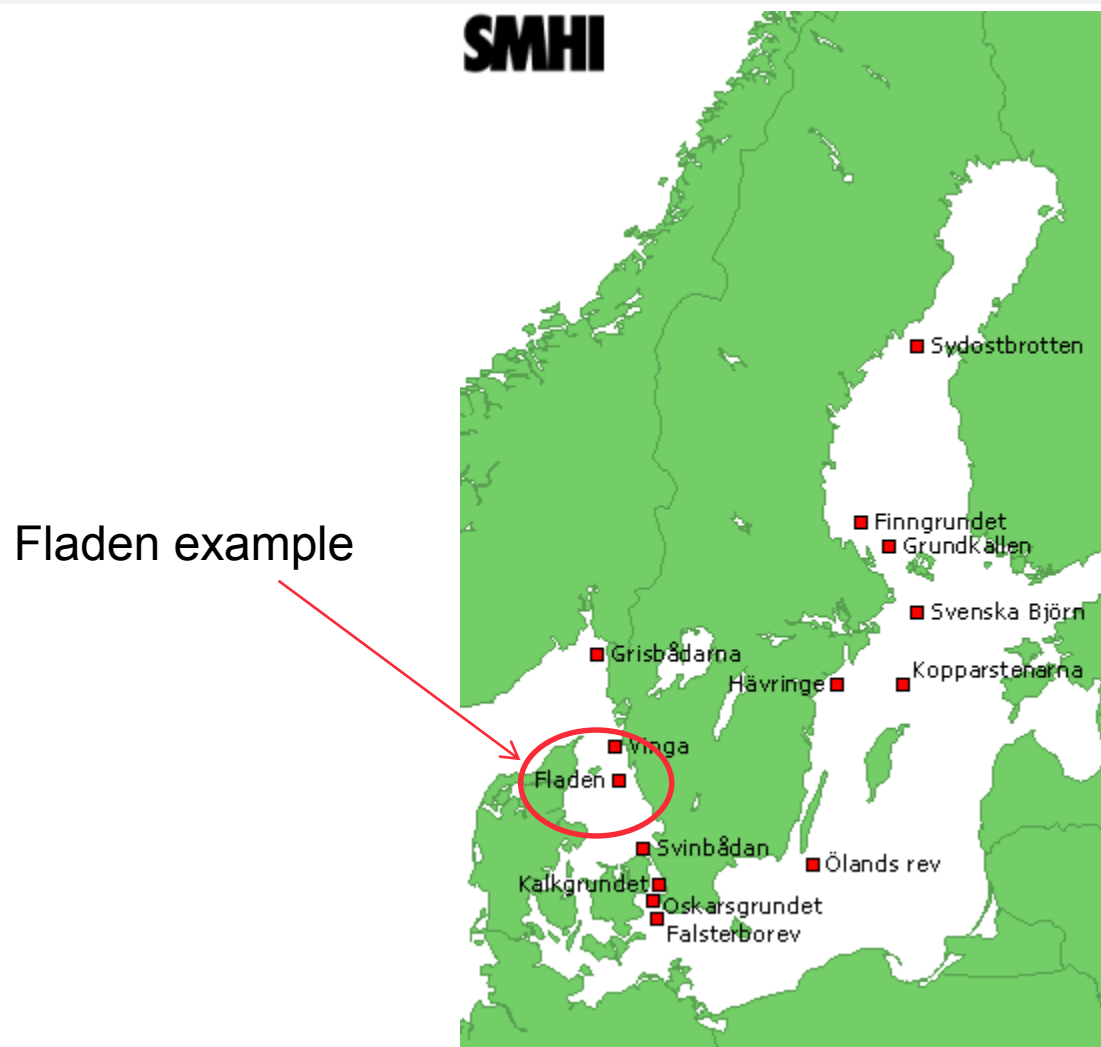
SMHI has historical data for the Baltic sea area

"Lighthouse vessel measurements" from 1860-1972

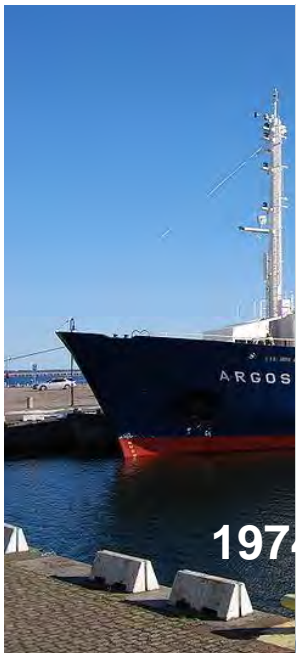


SMHI has historical data for the Baltic sea area

"Lighthouse vessel measurements" from 1860-1972



Marine environmental monitoring data

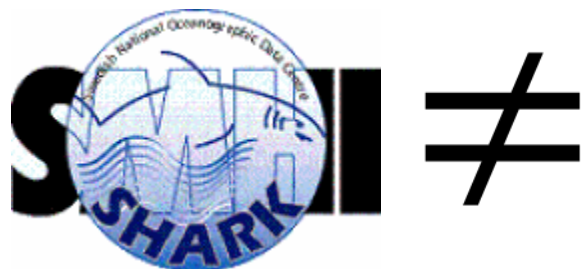


1974-2011

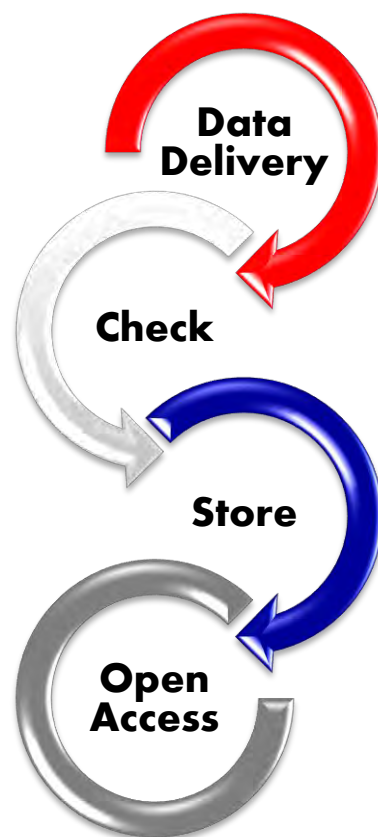


The hunters and collectors!

SHARK: "Svenskt HavsARKiv" - "Swedish Ocean Archive"



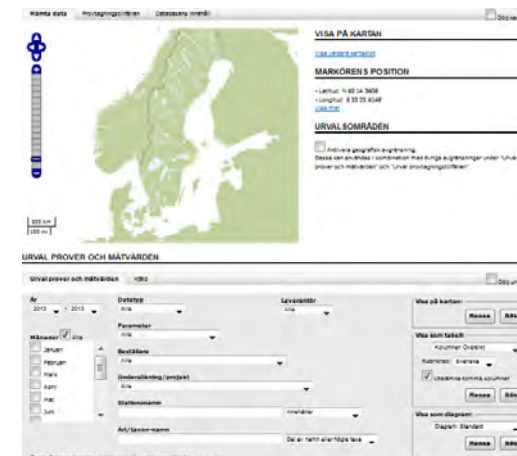
Workflow at the Swedish National Oceanographic Data Centre



Human interface SHARKweb.smhi.se

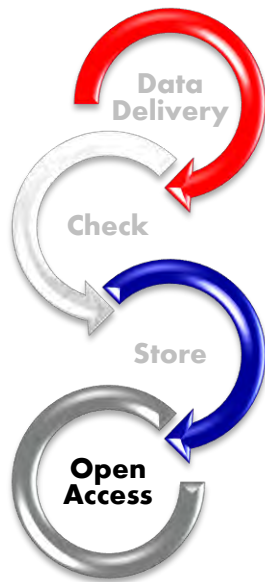


→ Manual admin.



SHARKweb.smhi.se

- marine environmental monitoring data - the seas surrounding Sweden.
- Search, preview and download
- Swedish only SORRY. But do send email to shark@smhi.se



Hämta data | Provtagningstillfällen | Databasens innehåll | Dölj karta

200 km
100 mi

VISA PÅ KARTAN

[Visa valbara kartskikt](#)

MARKÖRENS POSITION

- Latitud: N 52 51.0111'
- Longitud: E 1 10.1600'
[Visa mer](#)

URVAL SOMRÅDEN

Aktivera geografisk avgränsning.
Dessa kan användas i kombination med övriga avgränsningar under "Urval prover och mätvärden" och "Urval provtagningstillfällen".

URVAL PROVER OCH MÄTVÄRDEN

Urval prover och mätvärden | Hjälp | Dölj urval

År: 1971 - 2013

Datatyp: Zooplankton

Leverantör: Alla

Parameter: Alla

Beställare: Alla

Undersökning/projekt: Alla

Stationsnamn: Innehåller

Art/taxon-namn: Del av namn eller högre taxa

Visa på kartan:

Visa som tabell:

Kolumner: Översikt

Rubrikrad: Svenska

Utelämnna tomma kolumner

Visa som diagram:

Diagram: Standard

Geografisk avgränsning ej aktiverad. Kan aktiveras till höger om kartan.

SULTAT PROVER OCH MÄTVÄRDEN (20000 VÄRDEN AV 257734 VISAS)

Välj parameter/enhet:

= counted - ind/analysed sample fraction

Abundance - ind/m2

Abundance - ind/m3

Status SHARKweb 2016-10-09

Number of data deliveries: 2126

Sampling sites: 48 136

Sampling events: 225 049

Samples: 1 096 261

Physical/Chemical data values: 7 294 376

Biology data values: 3 006 928



RISE OF THE **FARMERS**

THE WAR BEGINS JULY 2

TM AND © 2008 COLUMBIA TRISTAR ENTERTAINMENT. ALL RIGHTS RESERVED. PHOTOFEST INC. SALE, DUPLICATION OR OTHER TRANSFER OF THIS MATERIAL IS STRICTLY PROHIBITED.

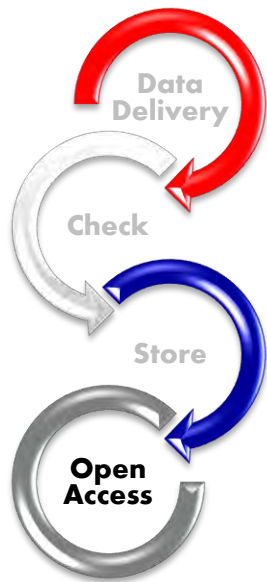
SHARKdata.se

- same as data as SHARKweb.smhi.se but mainly machine-machine
- harvest data or publish the data in other systems, portals, etc.

I.e. an REST API enabling REST-calls e.g.

http://sharkdata.se/datasets/SHARK_Speciesobs_2010_Zooplankton/data.txt

to get data from a dataset in text



Datasets

Available datasets						
Datatype	Dataset name	Metadata	Data	Version	Dataset	
Phytoplankton	SHARK_Phytoplankton_2010_SMHI	<input type="button" value="Metadata"/>	<input type="button" value="Data"/>	2013-09-09	SHARK_	
Phytoplankton	SHARK_Phytoplankton_2011_SMHI	<input type="button" value="Metadata"/>	<input type="button" value="Data"/>	2012-11-20	SHARK_	
Speciesobs	SHARK_Speciesobs_2010_Zooplankton	<input type="button" value="Metadata"/>	<input type="button" value="Data"/>	TEST2014-02-14	SHARK_	
Speciesobs	SHARK_Speciesobs_2011_Zooplankton	<input type="button" value="Metadata"/>	<input type="button" value="Data"/>	TEST2014-02-14	SHARK_	
Speciesobs	SHARK_Speciesobs_2012_Zooplankton	<input type="button" value="Metadata"/>	<input type="button" value="Data"/>	TEST2014-02-14	SHARK_	

- Datasets
- Resources
- Species observations
- Documentation
- About

Datasets

SHARKdata.se

Available datasets

Datatype	Dataset name	Metadata	Data	Version	Dataset
Phytoplankton	SHARK_Phytoplankton_2010_SMHI	Metadata	Data	2013-09-09	SHARK
Phytoplankton	SHARK_Phytoplankton_2011_SMHI	Metadata	Data	2012-11-20	SHARK
Speciesobs	SHARK_Speciesobs_2010_Zooplankton	Metadata	Data	TEST2014-02-14	SHARK
Speciesobs	SHARK_Speciesobs_2011_Zooplankton	Metadata	Data	TEST2014-02-14	SHARK
Speciesobs	SHARK_Speciesobs_2012_Zooplankton	Metadata	Data	TEST2014-02-14	SHARK

- Datasets
- Resources
- Species observations
- Documentation
- About

json / text

EMODnet
European Marine Observation and Data Network

Home Map Viewer Data Catalog Data Download Product Gallery Project Contribute Contact Blog

Search Active layers All Layers

Google Satellite

Copepoda, including child taxa in EurOBIS

records/grid

- ≤ 1
- ≤ 10
- ≤ 100
- ≤ 1000
- > 1000

Lat: 59.66 Lon:19.09

Google

Bilder ©2016 NASA, TerraMetrics | Användarvillkor | Rapportera ett kartfel



OCEAN BIOGEOGRAPHIC
INFORMATION SYSTEM

Datasets

SHARKdata.se

- Datasets
- Resources
- Observations
- Documentation
- About

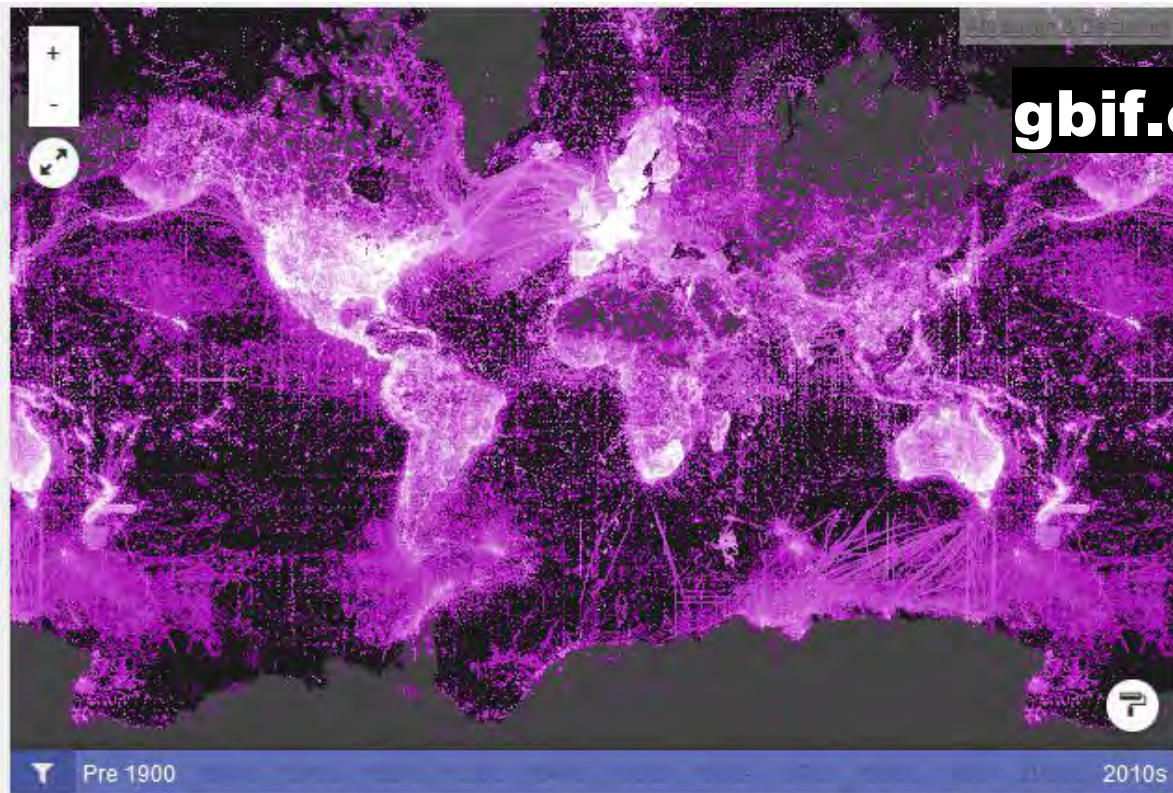
Available datasets

Datatype	Dataset name	Metadata	Data	Version	Dataset
Phytoplankton	SHARK_Phytoplankton_2010_SMHI	Metadata	Data	2013-09-09	SHARK
Phytoplankton	SHARK_Phytoplankton_2011_SMHI	Metadata	Data	2012-11-20	SHARK
Speciesobs	SHARK_Speciesobs_2010_Zooplankton	Metadata	Data	TEST2014-02-14	SHARK
Speciesobs	SHARK_Speciesobs_2011_Zooplankton	Metadata	Data	TEST2014-02-14	SHARK
Speciesobs	SHARK_Speciesobs_2012_Zooplankton	Metadata	Data	TEST2014-02-14	SHARK



<http://www.svenskalifewatch.se/en/>

gbif.org



Georeferenced data

[VIEW RECORDS](#)

[All records](#) | [In viewable area](#)

ABOUT

This map shows the density of all 546,525,317 georeferenced occurrence records published through the GBIF network.

To explore the records, zoom into the map or click on the links above and add further filters to customize search results.

SHARKdata geeky stuff

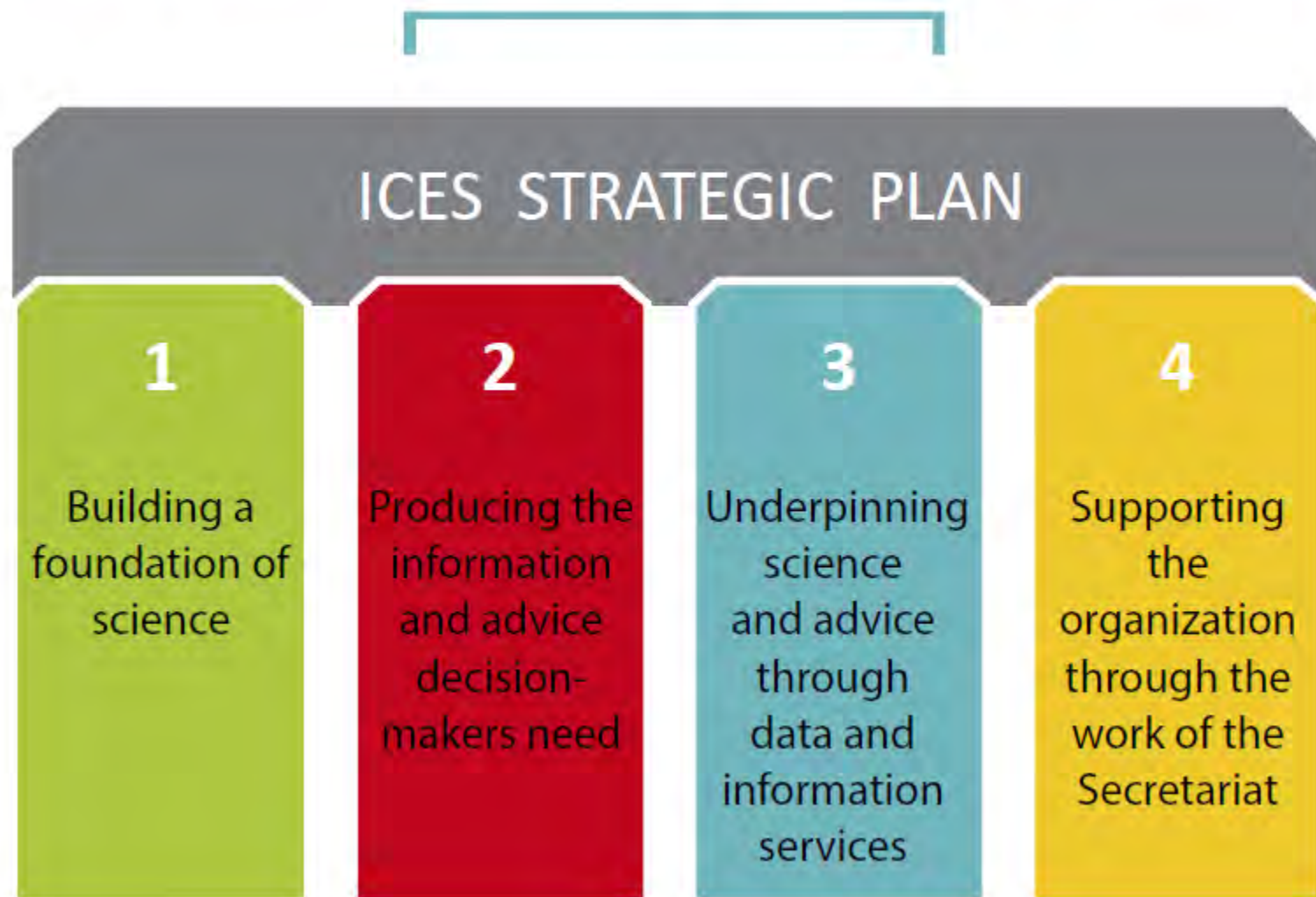
- Runs on ubuntu server on a cloud VPS
- Django framework (Python) + postgresSQL
- Developed by one consultant
- Maintained by one manager
- Uptime 100% (uptime robot: 2268 hrs, 58 mins)

```
library('jsonlite')
library('httr')
library(RCurl)
# Obtain list of all datasets
list_datasets<-function(type=TRUE){
  datasets<-fromJSON('http://sharkdata.se/datasets/list.json')
  if(type){
    return(datasets$dataset_name)
  }else{
    dtypes<-unique(datasets$datatype)
    return(dtypes)
  }
}
```

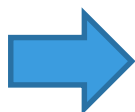
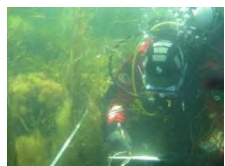


The end of reporting deadlines?

The ICES Strategic Plan is based on four pillars:



The future is happening as we speak



SHARKdata Home Documentation

SHARKdata.se

Resources
Species observations
Documentation
About

Import new dataset Delete all Reload all

Available datasets

Datatype	Dataset name	Metadata	Data	Version	Dataset
Phytoplankton	SHARK_Phytoplankton_2010_SMHI	Metadata	Data	2013-09-09	SHARK_
Phytoplankton	SHARK_Phytoplankton_2011_SMHI	Metadata	Data	2012-11-20	SHARK_
Speciesobs	SHARK_Speciesobs_2010_Zooplankton	Metadata	Data	TEST2014-02-14	SHARK_
Speciesobs	SHARK_Speciesobs_2011_Zooplankton	Metadata	Data	TEST2014-02-14	SHARK_
Speciesobs	SHARK_Speciesobs_2012_Zooplankton	Metadata	Data	TEST2014-02-14	SHARK_

ICES xml



ICES
CIEM

DOME

- I. Reduce the amount of manual work (human error)
- II. Ensure the most recent version in the ICES DOME (Marine Environment) database
- III. Allow the quality checks performed by ICES to be called by the SHARKdata.se preprocessing systems directly, which collectively leads to
- IV. Higher quality of the data.

SHARKdata.se/about

code repository etc: dev.sharkdata.se

