

# Stock Assessment Data Flow

## Status Update

Colin Millar, Arni Magnusson  
and the ICES Data Centre

WGCHAIRS  
24 Jan 2018



Science for sustainable seas

# The plan



Brief overview of TAF

**Diagram** - The TAF workflow

**R scripts** - Separation of concerns

Data flow in an ICES stock assessment

**Overview** - From national institutes to advice

**Catches** - Collation and raising to stock level

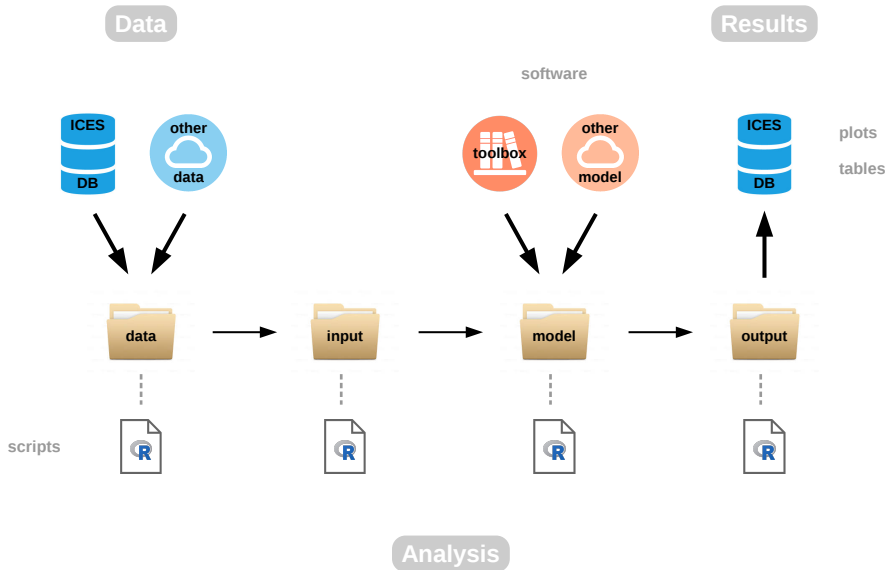
**Surveys** - Generation of survey indices

**Modelling** - Data processing and modelling

**Report** - Formatting and presenting results and data

**Advice** - High level summary, review, approval

# TAF core workflow overview



# Workflow scripts



## Initial

**upload.R** - Upload/Download raw data or model files

## Core

**data.R** - Preprocess data, write TAF data tables

**input.R** - Convert data to model format, write model input files

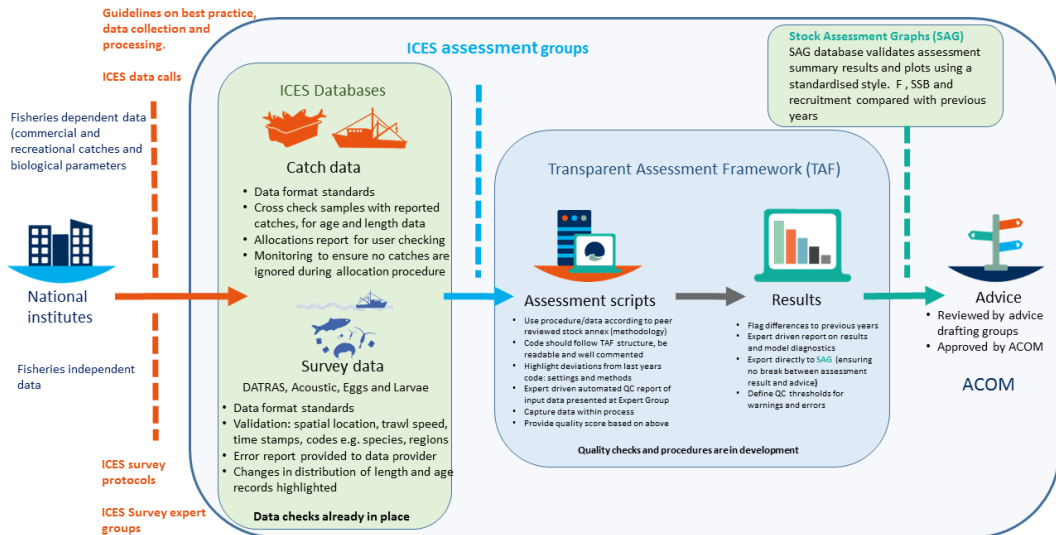
**model.R** - Run analysis, write model results

**output.R** - Extract results of interest, write TAF output tables

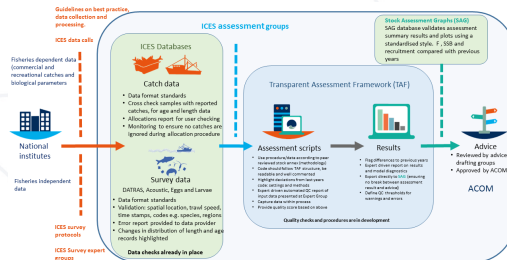
## Also

**report.R** - Prepare plots/tables for report

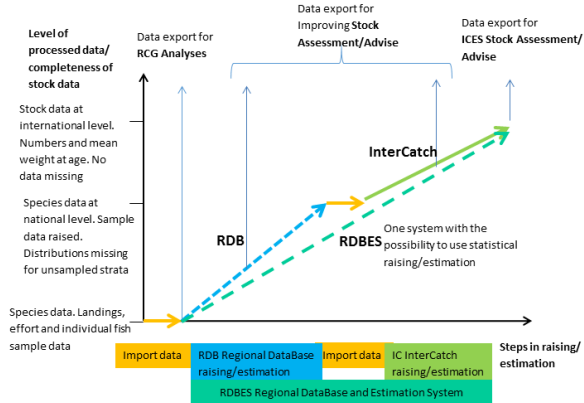
# From national institutes to advice



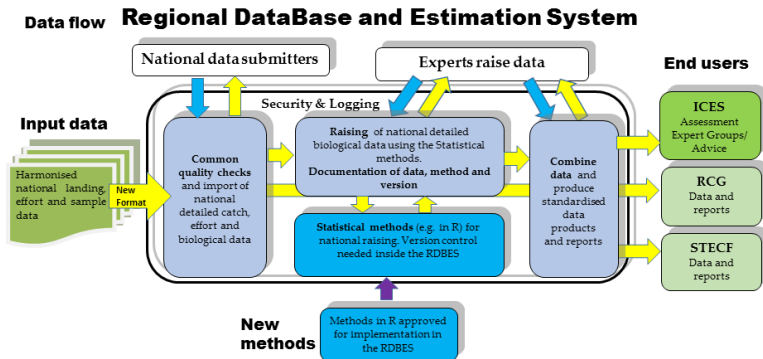
- ▶ Regional Database and Estimation System (RDBES)
- ▶ Database of Trawl Surveys (DATRAS)
- ▶ Acoustic Data Portal
- ▶ Stock Assessment Graphs (SAG)
- ▶ Eggs and Larvae, Stomach contents, Contaminants and VMS



# Data flow in RDB, IC and RDBES



# RDBES system overview





- ▶ Currently working with the RDB Development Support Group to specify the data model for the RDBES.
- ▶ When the RDBES data model is finalised it is sent to the countries for feedback. Can the countries use it? Is a change needed?
- ▶ The specifications of the RDBES should be ready in April.
- ▶ The plan is that the RDBES is ready in the spring 2019.

# DATRAS (Database of Trawl Surveys)



- DATRAS is an online database
- Standardised data extraction
- Updated continuously
- Standard input and quality assurance of all data
- Handles different survey designs (random and stratified)
- Outputs used for fisheries independent stock indices
- Marine litter and biodiversity indicators
- Access data product through WebAPI
- Data product calculation code on GitHub (On going..)

**DATRAS:  
FISH TRAWL  
SURVEYS**

ICES  
CIEM

Catch per unit effort, length, age, maturity, indices

Northeast Atlantic, Baltic Sea, North Sea, Irish Sea, Bay of Biscay

**Stations: 70 000**

**Years with data: 51**

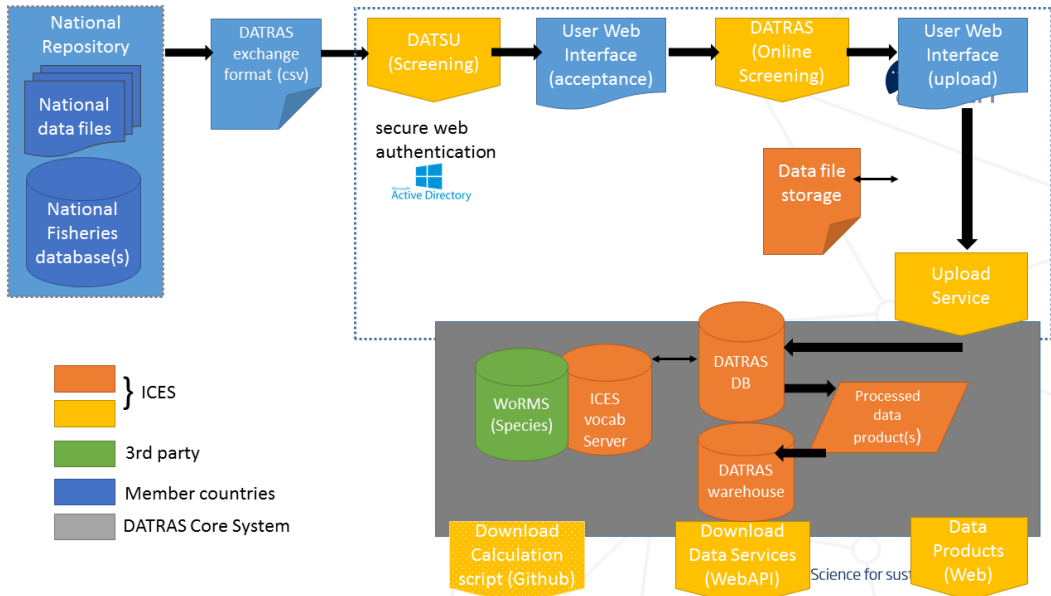
**First year: 1965**

**Latest year: Present**

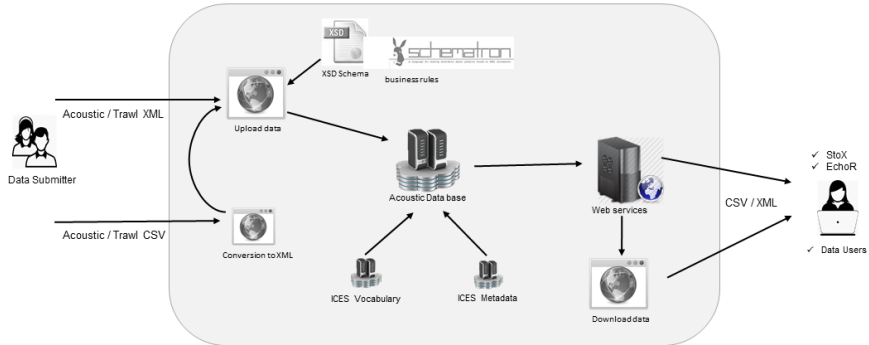
[datras.ices.dk](https://datras.ices.dk)

The poster features a background image of a large school of fish swimming in deep blue water. The text is arranged in a clean, modern layout with a teal footer bar.

# DATRAS dataflow

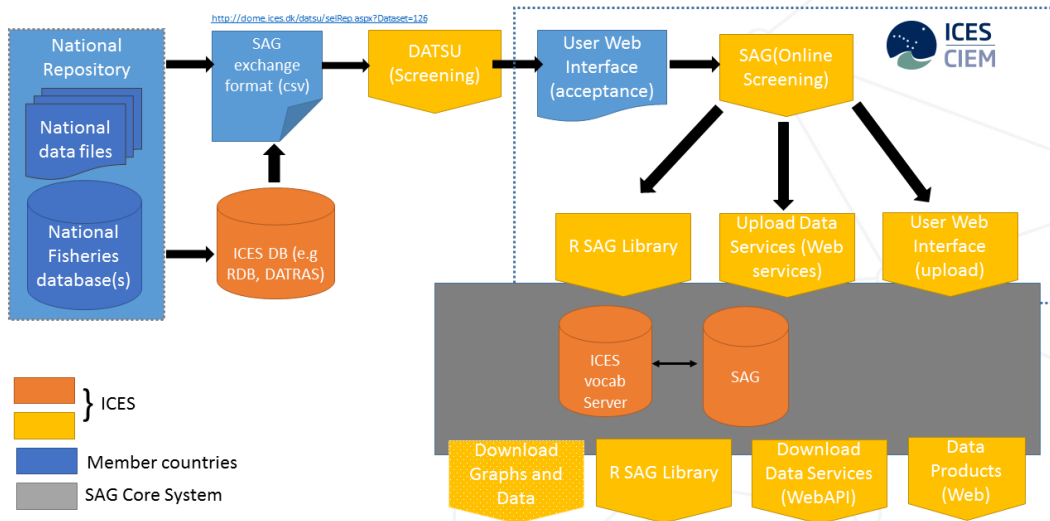


# Acoustic Data Portal

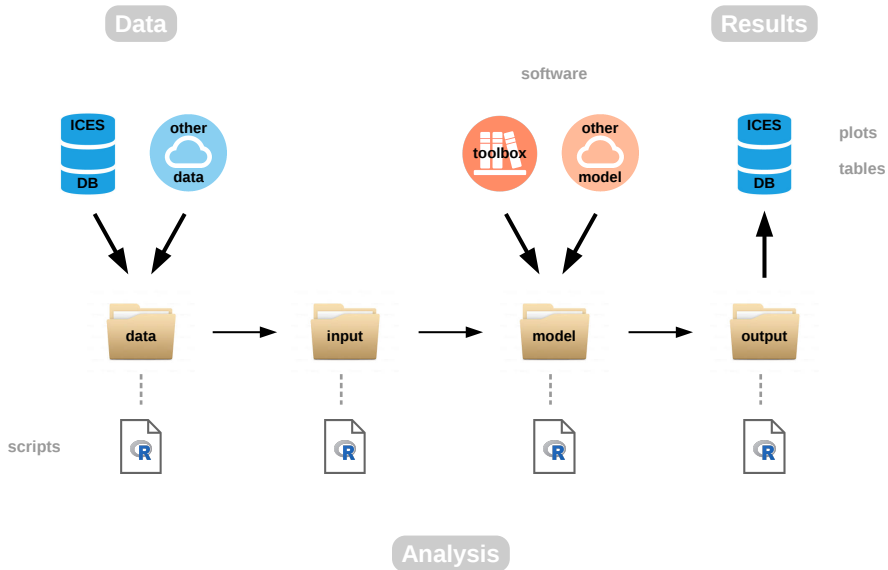


<http://acoustic.ices.dk>

# SAG dataflow



# TAF core workflow overview revisited



# Workflow scripts revisited



## Initial

**upload.R** - Upload/Download raw data or model files

## Core

**data.R** - Preprocess data, write TAF data tables

**input.R** - Convert data to model format, write model input files

**model.R** - Run analysis, write model results

**output.R** - Extract results of interest, write TAF output tables

## Also

**report.R** - Prepare plots/tables for report

## Uploading data from DATRAS Web service - upload.R

```
2 library(icesTAF)
3 library(icesDatras)
4
5 # download hh data from DATRAS web service
6 hh <- getDATRAS("HH", survey = "NS-IBTS", years = 2007:2017, quarter = 1)
7 # save locally to _raw folder
8 write.taf(hh, "_raw/hh.csv")
9
10 #upload to the 2017 North Sea Cod assessmet TAF repo: 2017_cod.27.2047d
11 upload("2017_cod.27.2047d", "raw", "_raw/hh.csv")
12
```



## Getting data into the core workflow - data.R

```
5 library(icesTAF)
6
7 # set the download location
8 url <- "http://taf.ices.local/taf/fs/2017\_cod.27.2047d/raw/"
9
10 ## Download hh data
11 hh <- read.taf(paste0(url, "hh.csv"))
12
```

## Getting data into SAG - output.R

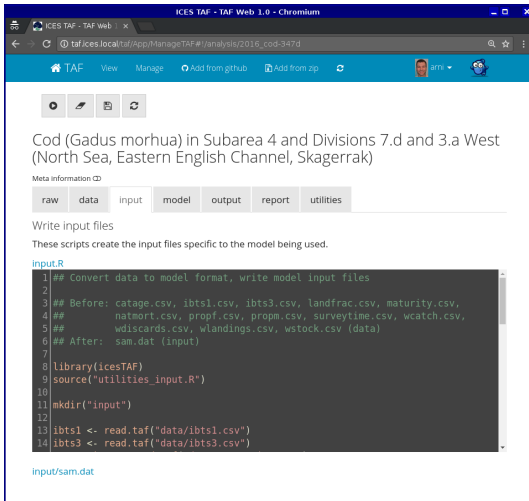
```
# upload to SAG (with some checking by DATSU on the way
options(icesSAG.use_token = TRUE)
options(icesSAG.messages = FALSE)
key <- uploadStock(info, fishdata)

## Converting to XML format ...
## Done
## Uploading ...
## Success: (200) OK
## Screening file ...
## Success: (200) OK
## Importing to database ...
## Done
## Upload complete! New assessmentKey is: 8651
## To check upload run (with 'options(icesSAG.use_token = TRUE)'):
## findAssessmentKey('whb.27.1-91214', 2010, full = TRUE)

key

## [1] 8651
```

# Viewing, editing and running R code online



ICES TAF - TAF Web 1.0 - Chromium

taf.ices.local/taf/App/ManageTAF#/analysis/2016\_cod-347d

TAF View Manage Add from github Add from zip artv

Play Edit Save Refresh

## Cod (Gadus morhua) in Subarea 4 and Divisions 7.d and 3.a West (North Sea, Eastern English Channel, Skagerrak)

Meta information

raw data **input** model output report utilities

Write input files

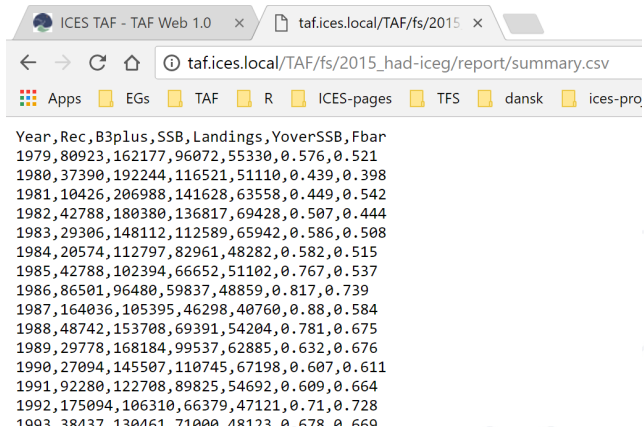
These scripts create the input files specific to the model being used.

[input.R](#)

```
1 ## Convert data to model format, write model input files
2
3 ## Before: catage.csv, ibts1.csv, ibts3.csv, landfrac.csv, maturity.csv,
4 ##         natmort.csv, propf.csv, propm.csv, surveytime.csv, wcatch.csv,
5 ##         wdiscards.csv, wlandings.csv, wstock.csv (data)
6 ## After:  sam.dat (input)
7
8 library(icesTAF)
9 source("utilities_input.R")
10
11 mkdir("input")
12
13 ibts1 <- read.taf("data/ibts1.csv")
14 ibts3 <- read.taf("data/ibts3.csv")
```

[input/sam.dat](#)

# Inputs and results available to view or download



Year	Rec	B3plus	SSB	Landings	YoverSSB	Fbar
1979	80923	162177	96072	55330	0.576	0.521
1980	37390	192244	116521	51110	0.439	0.398
1981	10426	206988	141628	63558	0.449	0.542
1982	42788	180380	136817	69428	0.507	0.444
1983	29306	148112	112589	65942	0.586	0.508
1984	20574	112797	82961	48282	0.582	0.515
1985	42788	102394	66652	51102	0.767	0.537
1986	86501	96480	59837	48859	0.817	0.739
1987	164036	105395	46298	40760	0.88	0.584
1988	48742	153708	69391	54204	0.781	0.675
1989	29778	168184	99537	62885	0.632	0.676
1990	27094	145507	110745	67198	0.607	0.611
1991	92280	122708	89825	54692	0.609	0.664
1992	175094	106310	66379	47121	0.71	0.728
1993	38137	130161	71000	48123	0.678	0.660

# And can be easily read into R

R Console (32-bit)

File Edit Misc Packages Windows Help

R is a collaborative project with many contributors.  
Type 'contributors()' for more information and  
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or  
'help.start()' for an HTML browser interface to help.  
Type 'q()' to quit R.

```
> read.csv("http://taf.ices.local/TAF/fs/2015_had-iceg/report/summary.csv")
```

	Year	Rec	B3plus	SSB	Landings	YoverSSB	Fbar
1	1979	80923	162177	96072	55330	0.576	0.521
2	1980	37390	192244	116521	51110	0.439	0.398
3	1981	10426	206988	141628	63558	0.449	0.542
4	1982	42788	180380	136817	69428	0.507	0.444
5	1983	29306	148112	112589	65942	0.586	0.508
6	1984	20574	112797	82961	48282	0.582	0.515
7	1985	42788	102394	66652	51102	0.767	0.537
8	1986	86501	96480	59837	48859	0.817	0.739
9	1987	164036	105395	46298	40760	0.880	0.584
10	1988	48742	153708	69391	54204	0.781	0.675
11	1989	29778	168184	99537	62885	0.632	0.676
12	1990	27094	145507	110745	67198	0.607	0.611

# Published assessments accessible on GitHub



ices-taf / 2015\_had-iceg

Unwatch 2 Star 0 Fork 0

Code Issues 0 Pull requests 0 Projects 0 Wiki Insights Settings

Icelandic haddock [Add topics](#) [Edit](#)

48 commits 2 branches 3 releases 2 contributors

Branch: master New pull request Create new file Upload files Find file Clone or download

colinpmillar test change for webhook post Latest commit 985241e 15 days ago

_model	Download model executable from TAF database	3 months ago
_raw	Download catageysa.dat from TAF database -> data, later moved to input	3 months ago
data.R	Use setwd("data") to shorten paths	a month ago
input.R	Sort header files alphabetically	2 months ago
model.R	Sort header files alphabetically	2 months ago
output.R	Extract model results -> Extract results	29 days ago

# List code changes between stages or between years



ices-taf / 2015\_had-iceg

Unwatch 2

Star 0

Fork 0

Code

Issues 0

Pull requests 0

Projects 0

Wiki

Insights

Settings

## Comparing changes

Choose two branches to see what's changed or to start a new pull request. If you need to, you can also [compare across forks](#).



base: 2.0

...

compare: 2.1

7 commits

5 files changed

0 commit comments

1 contributor



Commits on Sep 13, 2017



arnima-github

Rename report.R -> report\_tables.R

21c7d3f



arnima-github

Add report\_plots.R to demonstrate TAF

fb62492



Commits on Sep 15, 2017



arnima-github

Download raw data and model executable from [github.com/ices-taf/ftp](https://github.com/ices-taf/ftp)

e5e7093



arnima-github

Use new download() function

941d8ab

# View code changes between stages or between years

Showing 5 changed files with 86 additions and 10 deletions.

Unified

Split

2 data.R

View



```
@@ -9,7 +9,7 @@ source("utilities.R")
9      9
10     10  mkdir("data")
11     11
12     12  -url <- "http://taf.ices.local/taf/fs/2015_had-iceg/raw/"
12     12  +url <- "https://raw.githubusercontent.com/ices-taf/ftp/master/nwwg/2015/had-iceg/raw/"
13     13
14     14  ## Download data, this file will later be moved to input
15     15  download.file(paste0(url,"catageysa.dat"), "data/catageysa.dat", quiet=TRUE)
```

5 model.R

View



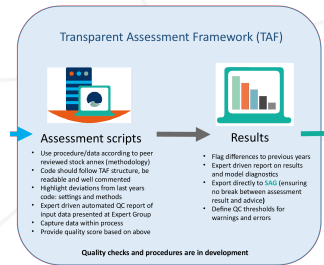
```
@@ -7,12 +7,11 @@ library(icesTAF)
7      7
8      8  mkdir("model")
9      9
10     10  -url <- "http://taf.ices.local/taf/fs/2015_had-iceg/model/"
```



# Quality Control

- ▶ Code should follow TAF **structure**, be readable and well commented
- ▶ Highlight **changes** from last year: code, settings, methods, data and results
- ▶ Expert driven **QC reports** covering inputs, results and model diagnostics
- ▶ Define **QC thresholds** to flag warnings
- ▶ **Directly connecting** DBs → assessment → SAG (advice)
- ▶ Provide **quality score**

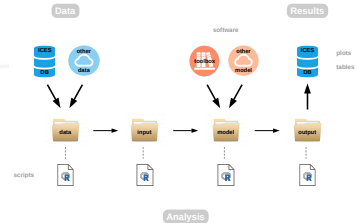
TIC-TAQ



# Benefits



- ▶ Easy to **find data and results** from final assessment
- ▶ **Open** and reproducible science, improved **quality control**
- ▶ Easy for scientists around the world to **get ICES data**
- ▶ Easy to run an **update assessment** next year
- ▶ If scientist **changes jobs**, next person can take over
- ▶ Existing and **future tools** can use TAF services



# Follow ongoing development



**[taf.ices.dk](http://taf.ices.dk)**

Main landing page

**[ices-taf-dev.github.io](https://ices-taf-dev.github.io)**

Technical overview and design comments

**[ices-taf-dev.github.io/pdf/taf-flyer.pdf](https://ices-taf-dev.github.io/pdf/taf-flyer.pdf)**

TAF outline and diagram

**[github.com/ices-taf](https://github.com/ices-taf)**

Example assessments (R scripts)

Time for a demo :)

**Thanks!**



**Questions?**