



Modern Data Management Principles

Øystein Godøy

<https://www.youtube.com/watch?v=N2zK3sAtr-4>

Making Your Research Easier and Cheaper

The 5 P's matter!

Prior

Planning

Prevents

Poor

Performance!

In other words

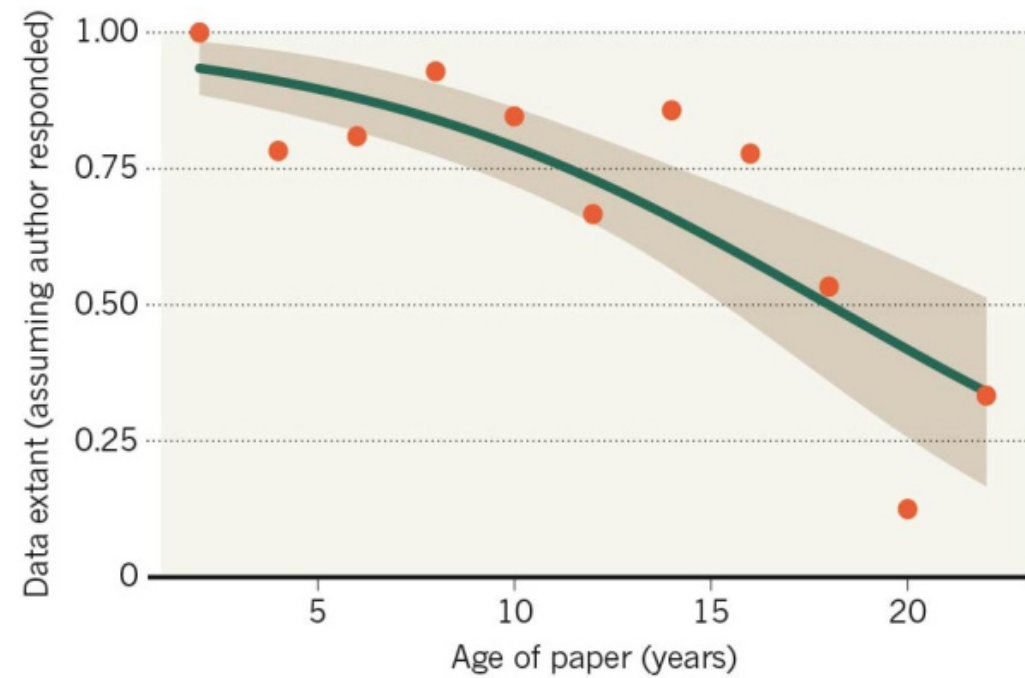
Learn form the mistakes of other...

Loosing scientific data

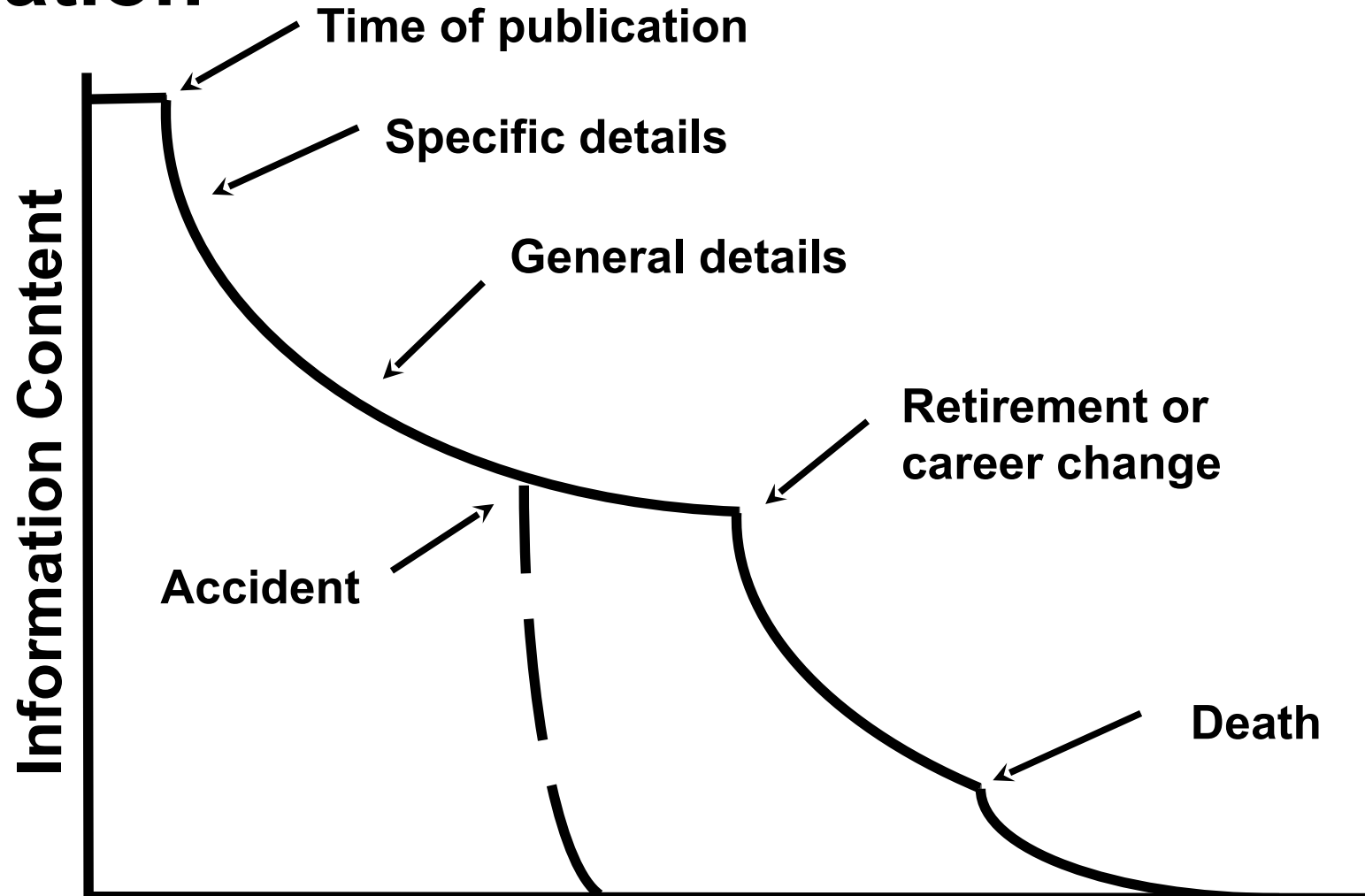
- Decline can mean 80% of data are unavailable after 20 years.
 - Gibney and Van Noorden (2013), Nature

MISSING DATA

As research articles age, the odds of their raw data being extant drop dramatically.



Poor data practice results in loss of information



(Michener et al. 1997)

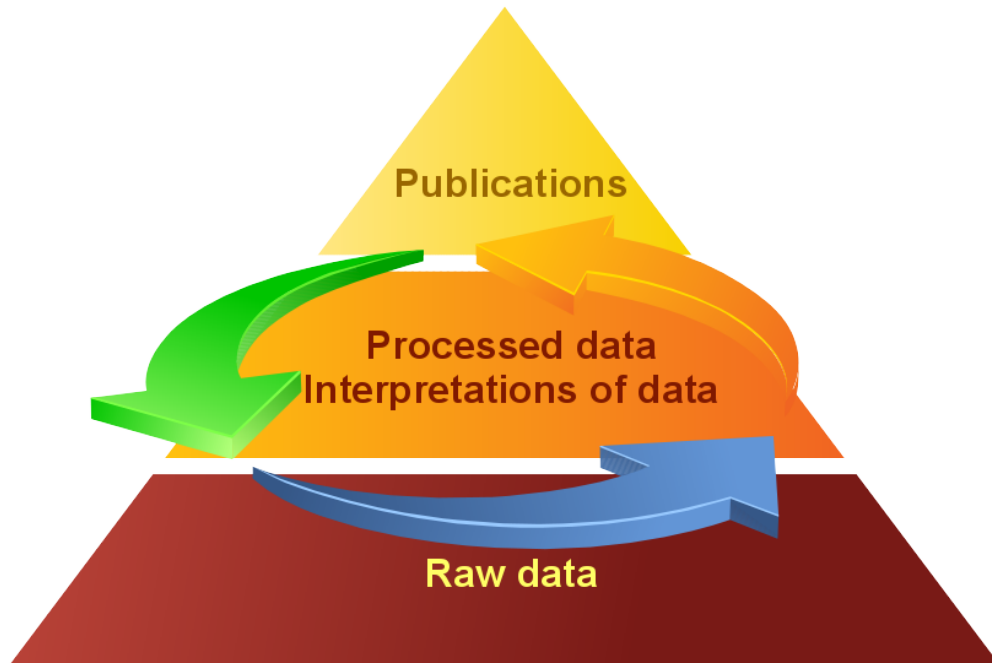
Why bother with structured data management?

- Science paradigms
 - according to Jim Gray
 - empirical science
 - theoretical science
 - computational science
 - data exploration science
- Maximise public investment in data collection and production
- Promote scientific collaboration
- Promote interdisciplinary science
- Promote scientific transparency
- Leave a legacy



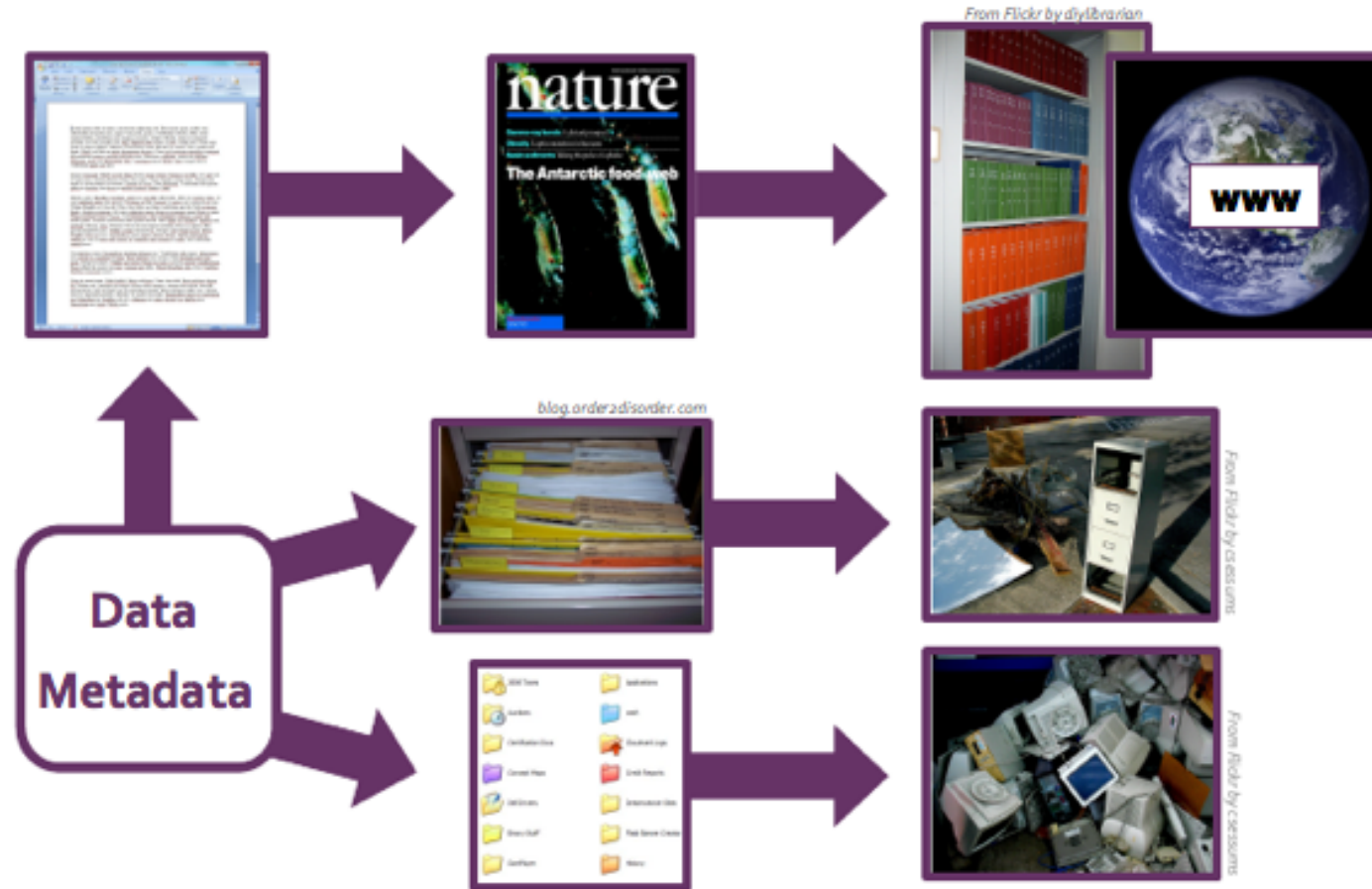
All scientific data online

Source: Jim Gray on eScience:
A Transformed Scientific Method

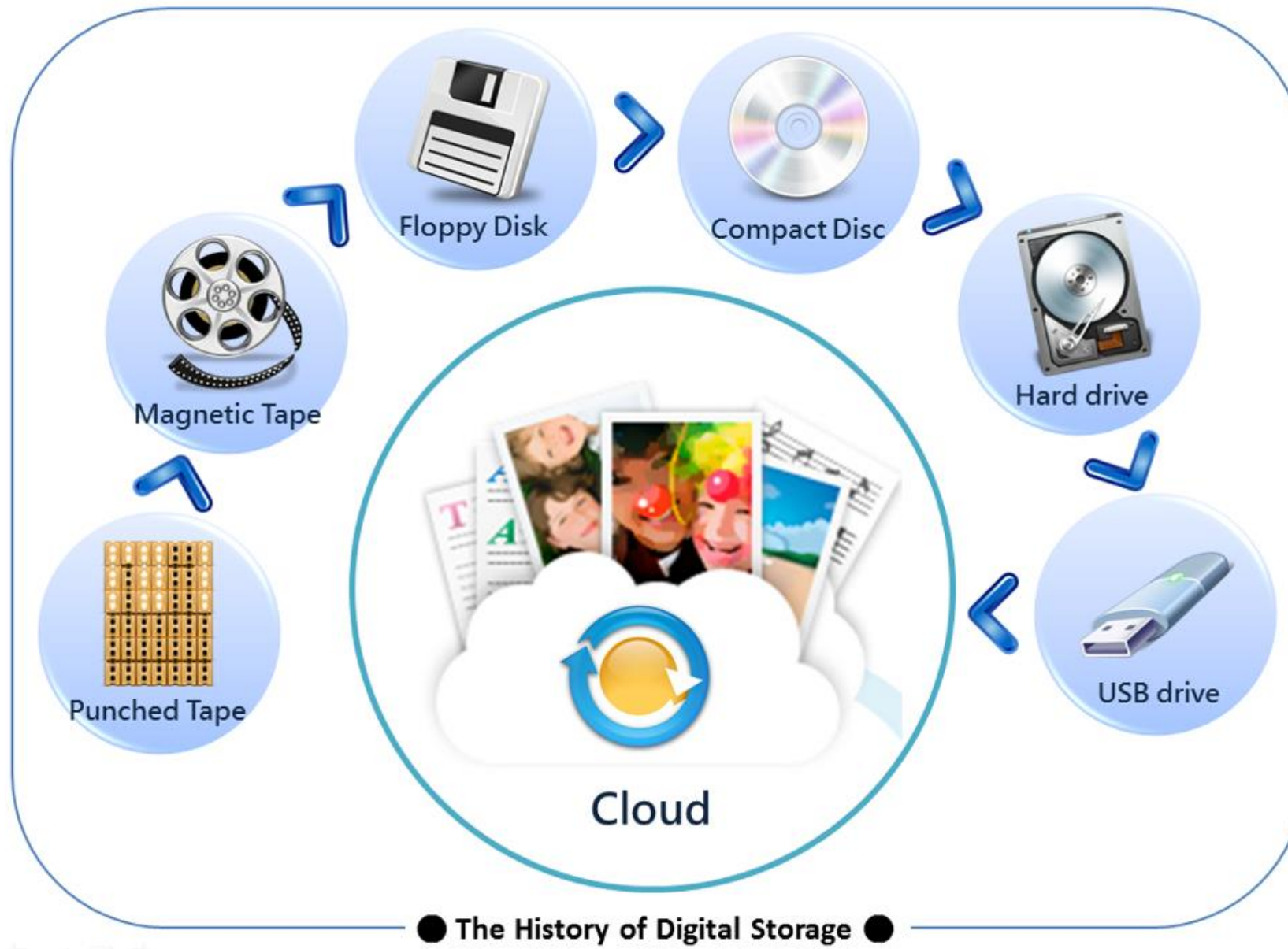


- Many disciplines overlap and use data from other sciences
- Science, government agencies and companies get a broader data background
- Internet can unify data, software and literature
- Go from literature to computation to data back to literature
- Information is at your fingertips for everyone and everywhere
- Potentially Increased Scientific Information Velocity
- Potentially Huge increase in Science Productivity

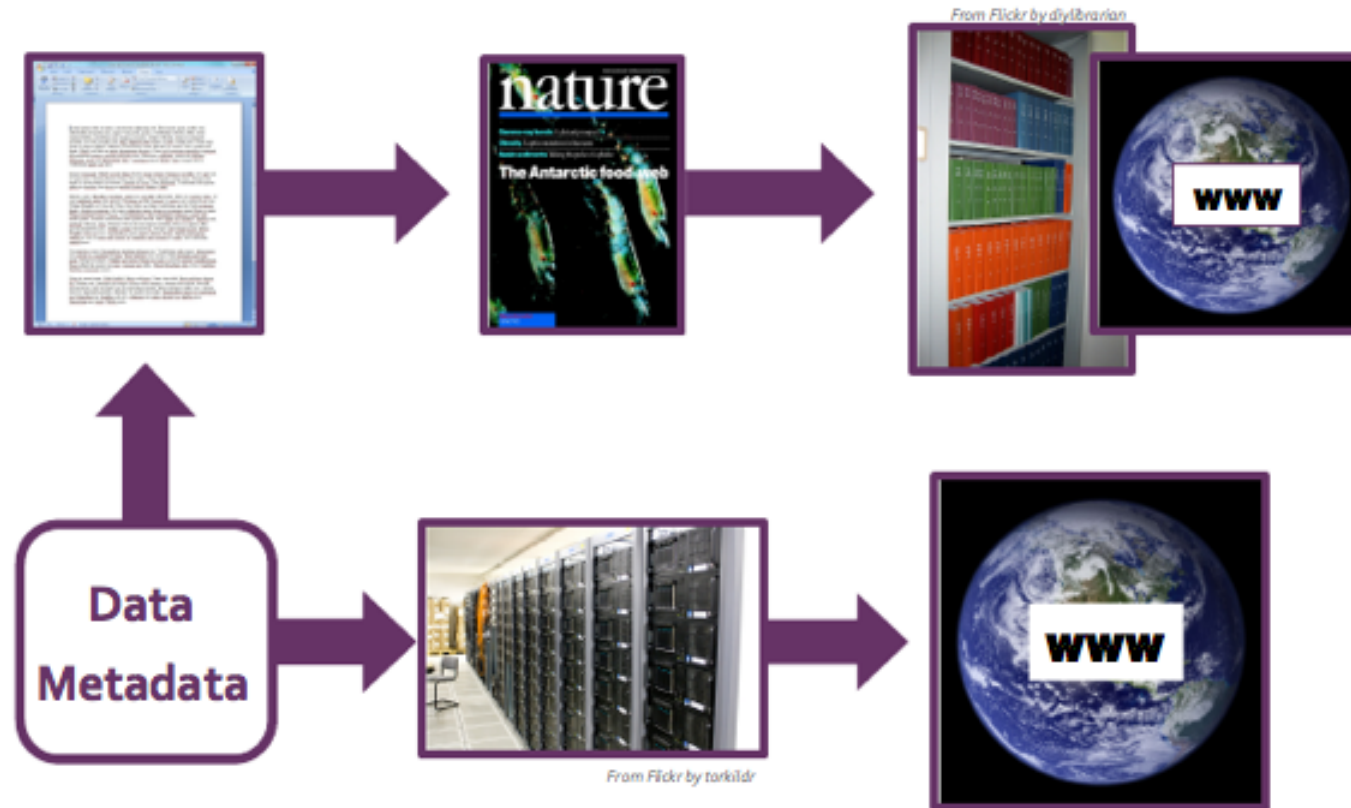
The reality today



Recreated from Klump et al. 2006

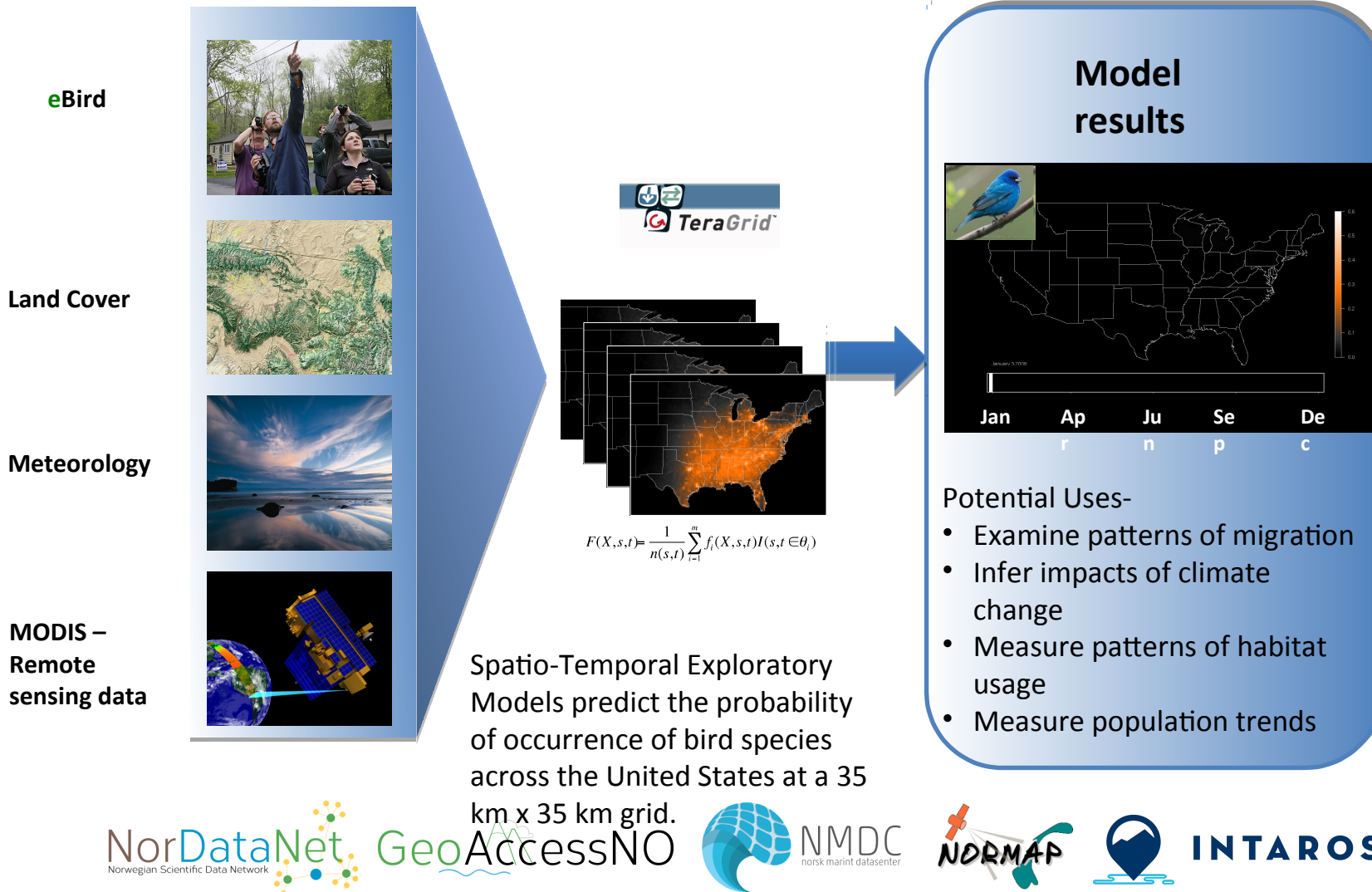


The vision for the future

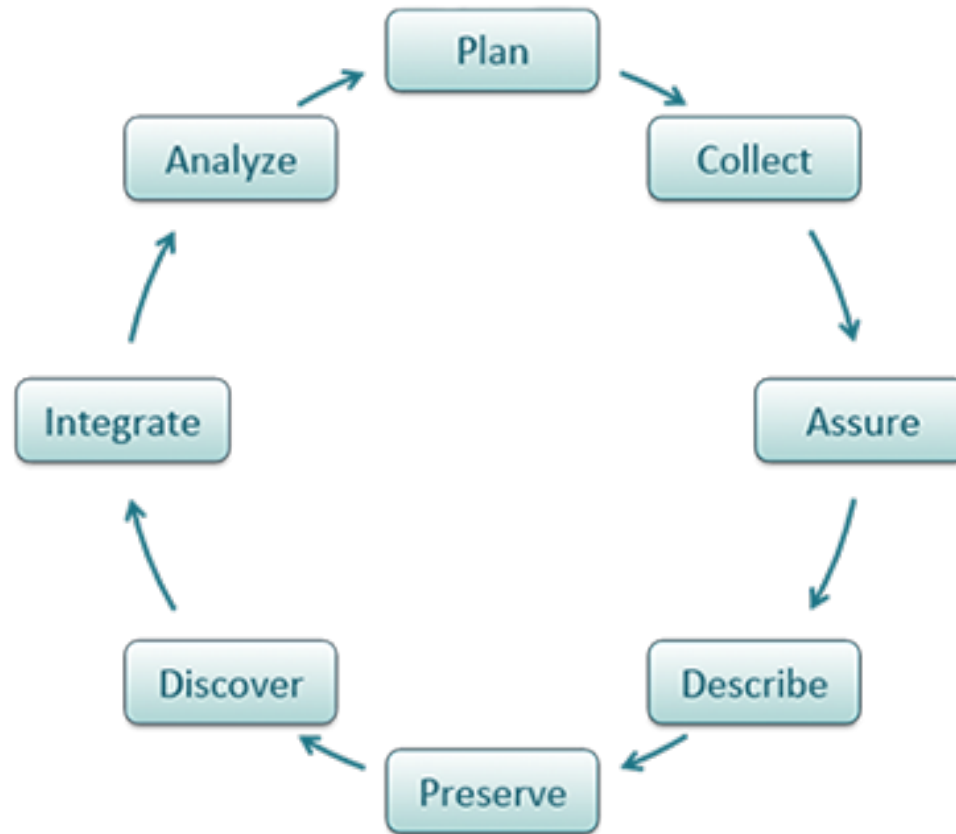


Recreated from Klump et al. 2006

New science



By re-using data collected from a variety of sources – eBird database, land cover data, meteorology, and remotely sensed by NASA – this project was able to compile and process the data using supercomputing to determine bird migration routes for particular species.



Why share data?

- Research sponsor require it
 - recognition as an authoritative source and wise investment
- Quality control
 - improved data quality due to expanded use, field checks, and feedback
- Improved visibility
 - improved connections to scientific network, peers, and potential collaborators
- Journals require it
- Far upstream sponsors require it



CC image by SLU Madrid Campus
on Flickr

Ownership of data

- This depends on the contract between the funding agency and the institution affiliated with the scientist
- For RCN and H2020, ownership lies with the institution
 - Not the individual scientist



A large, stylized illustration of an iceberg floating in dark blue water. The iceberg is light blue and has a jagged, geometric shape. The water is a deep blue, and the sky is a light grey-blue. The iceberg is the central focus of the slide.

Documentation of data Various types of metadata

Øystein Godøy

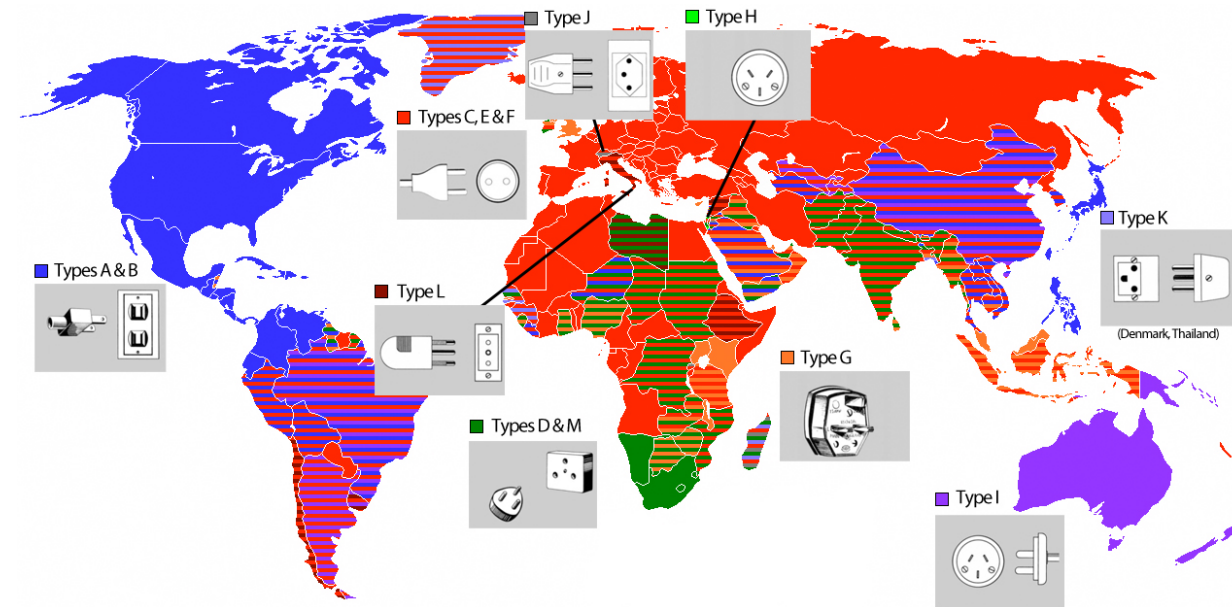
Interoperability

- At the discovery level
 - Documentation
 - Exchange of documentation
- At the data level
 - Documentation
 - Exchange of documentation



Benefits of standardisation

- Makes life easier
 - Promotes reuse of efforts
- Promotes a common understanding of content
 - Improves performance
 - Reduce risk
 - Promotes sustainability
 - Encourage innovation
 - Reduce cost
 - Improve quality



Benefits of standardised documentation

- Why not use the “Google” approach?
- Science is based on a shared terminology
 - There will never be only one proper way of documenting
 - There will always be a need for brokering
- Data and metadata must be connected
 - To find data
 - To use data
- Standardised documentation and formatting
 - enables the possibility to filter datasets
 - enables the possibility to link datasets
 - enables standardised applications to analyse data
 - enables users to use the data
- Need to be pragmatic...
 - And let computers do the boring part
 - But humans need to instruct computers

Types of metadata

- Discovery metadata
 - **who** measured, simulated or analysed **what**, **where**, and **when** as well as **conditions for reuse** and **access mechanisms** for the data
 - to enable users to find appropriate data for the task
- Use metadata
 - **identification of the variables/parameters** generated, **units of variables/parameters**, how **missing values** are encoded, definition of **grid and map projections** for gridded data, **methodology applied in space or time** to achieve the values in a dataset etc
 - to enable users to properly understand the data found
- Site metadata
- Configuration metadata



Eksempel på merking

Betegnelse **1**

Frivillige opplysninger **14**

Nettoinnhold **5**

Kjøttpølse

Kjølevarer: 0 + 4 °C **7**

Holdbarhet **6** **Siste forbruksdag:** 11.10.15 **Pakkedato:** 18.09.15 **Nettovekt:** 520 g

10 **Må gjennomvarmes før servering**

Næringsdeklarasjon **12**

Næringsinnhold	per 100 g:
Energi	963 kJ/ 232 kcal
Fett	18 g
hvorav mettede fettsyrer	6,6 g
Karbohydrater	51 g
hvorav sukkerarter	0,8 g
Protein	11 g
Salt	1,8 g

2 **Ingredienser**

4 **Mengdeangivelse av ingredienser**

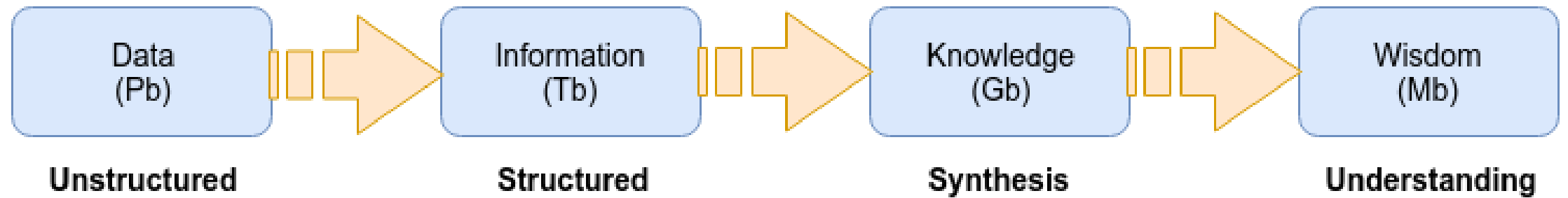
3 **Stoffer som kan fremkalle allergi eller intoleranse**

13 **Særskilte merke-
regler for visse
matvarer.
For eksempel god-
kjenningsnummer.**

8 **Ansvarlig næringsmiddel-
virksomhet** **Produsent:** Ensen kjøtt, Virksomhetsvegen 1, 1111 Nyby
Forbrukerkontakt. www.ensen.no

Ingredienser: Kjøtt (55 %) av storfe og svin, **melk**, potetmel, salt, krydder, **selleri**, antioksidant E 300, konserveringsmiddel E 250.

DIKW chain



Data in context

Lynn Yarmey, NSIDC, 2013 ESIP Material



Has anyone checked
the quality of this value?

For what purpose?

Collected when?

Of what?

Precision/accuracy?

Location?

Temperature

In what units?

31.5

According
to whom?

AKA – T, Temp, degC, C, °F... lots of different names

When was the sensor last
cleaned/calibrated?

Is this value
averaged? Calculated?

Collected how?

- Make data “talk”
 - Standardised documentation
 - e.g. Climate and Forecast convention
- What is a number?
 - Use metadata is required to efficiently reuse data
 - Describe the context of observations using e.g. WIGOS metadata


```

File Edit View Bookmarks Settings Help
netcdf radflux_bjornoya {
dimensions:
    time = UNLIMITED ; // (3847970 currently)
    strlen25 = 25 ;
variables:
    double time(time) ;
        time:long_name = "time of the observation" ;
        time:short_name = "time" ;
        time:standard_name = "time" ;
        time:units = "seconds since 1970-01-01 00:00:00 UTC" ;
        time:axis = "T" ;
    char stationid(strlen25) ;
        stationid:long_name = "name and/or stationnumber used as identifier" ;
    float latitude ;
        latitude:long_name = "latitude" ;
        latitude:short_name = "latitude" ;
        latitude:standard_name = "latitude" ;
        latitude:units = "degree_north" ;
        latitude:valid_min = -90.f ;
        latitude:valid_max = 90.f ;
    float longitude ;
        longitude:long_name = "longitude" ;
        longitude:short_name = "longitude" ;
        longitude:standard_name = "longitude" ;
        longitude:units = "degree_east" ;
        longitude:valid_min = -180.f ;
        longitude:valid_max = 180.f ;
    float ssi(time) ;
        ssi:long_name = "shortwave irradiation at the surface" ;
        ssi:short_name = "ssi" ;
        ssi:standard_name = "surface_downwelling_shortwave_flux" ;
        ssi:FillValue = -999.f ;
        ssi:units = "watts/meter2" ;
        ssi:cell_method = "time: mean (last minute)" ;
    float ssisenstemp(time) ;
        ssisenstemp:long_name = "temperature of the surface shortwave irradiation sensor" ;
        ssisenstemp:short_name = "ssisenstemp" ;
        ssisenstemp:FillValue = -999.f ;
        ssisenstemp:units = "degC" ;
        ssisenstemp:cell_method = "time: mean (last minute)" ;
    float dli(time) ;
        dli:long_name = "difference between downward atmospheric longwave irradiation and emitted CGR4 irradiance" ;
        dli:short_name = "dli" ;
        dli:standard_name = "surface_net_downward_longwave_flux" ;
        dli:FillValue = -999.f ;
        dli:units = "watts/meter2" ;
        dli:cell_method = "time: mean (last minute)" ;
    float dlsenstemp(time) ;
        dlsenstemp:long_name = "temperature of the surface longwave irradiation sensor" ;
        dlsenstemp:short_name = "dlsenstemp" ;
        dlsenstemp:FillValue = -999.f ;
        dlsenstemp:units = "degC" ;
        dlsenstemp:cell_method = "time: mean (last minute)" ;
    float battery(time) ;
        battery:long_name = "minimum battery voltage" ;
        battery:short_name = "battery" ;
        battery:FillValue = -999.f ;
        battery:units = "V" ;
        battery:cell_method = "time: min (last minute)" ;
}

```

Bjørnøya : ncdump

```

File Edit View Bookmarks Settings Help
// global attributes:
    :Conventions = "CF-1.0" ;
    :history = "2008-10-23 creation\n",
        "2016-01-01 revision" ;
    :title = "Downwelling surface radiative fluxes at Bear Island" ;
    :abstract = "Downwelling surface radiative fluxes observed at the meteorological station at Bear Island in the Barents Sea. Measurements are made using Kipp and Zonen CMP21 and CGR4 pyranometers and pyrgeometers. Daily maintenance is performed by the meteorological personnel at the station. Data are averaged over the last minute and the time is set to UTC. This data set has been collected with support from the Norwegian Research Council. The quality control focuses on the radiative parameters, thus sensor temperatures may contain errors." ;
    :topiccategory = "ClimatologyMeteorologyAtmosphere" ;
    :keywords = "Radiative Flux" ;
    :gcmd_keywords = "Atmosphere > Atmospheric Radiation > Shortwave Radiation\n",
        "Atmosphere > Atmospheric Radiation > Longwave Radiation" ;
    :area = "Barents Sea" ;
    :activity_type = "Land station" ;
    :PI_name = "<D8>ystein God<F8>y" ;
    :contact = "o.godoy@met.no" ;
    :institution = "Norwegian Meteorological Institute" ;
    :url = "http://www.met.no/" ;
    :product_name = "radiative fluxes" ;
    :Platform_name = "Bj<F8>rn<F8>ya" ;
    :project_name = "iA00S-Norway/IPY-THORPEX" ;
    :start_date = "2008-04-01 13:14 UTC" ;
    :stop_date = "2015-12-16 12:50 UTC" ;
    :distribution_statement = "Restricted to iA00S-Norway" ;
    :southernmost_latitude = 74.5166667 ;
    :northernmost_latitude = 74.5166667 ;
    :westernmost_longitude = 19.0166667 ;
    :easternmost_longitude = 19.0166667 ;
    :quality_statement = "Quality controlled" ;
    :nco_openmp_thread_number = 1 ;

data:
    time = 1207055640, 1207055700, 1207055760, 1207055820, 1207055880,
    1207055940, 1207056000, 1207056060, 1207056120, 1207056180, 1207056240,
    1207056300, 1207056360, 1207056420, 1207056480, 1207056540, 1207056600,
    1207056660, 1207056720, 1207056780, 1207056840, 1207056900, 1207056960,
    1207057020, 1207057080, 1207057140, 1207057200, 1207057260, 1207057320,
    1207057380, 1207057440, 1207057500, 1207057560, 1207057620, 1207057680,
    1207057740, 1207057800, 1207057860, 1207057920, 1207057980, 1207058040,
    1207058100, 1207058160, 1207058220, 1207058280, 1207058340, 1207058400,
    1207058460, 1207058520, 1207058580, 1207058640, 1207058700, 1207058760,
    1207058820, 1207058880, 1207058940, 1207059000, 1207059060, 1207059120,
    1207059180, 1207059240, 1207059300, 1207059360, 1207059420, 1207059480,
    1207059540, 1207059600, 1207059660, 1207059720, 1207059780, 1207059840,
    1207059900, 1207059960, 1207060020, 1207060080, 1207060140, 1207060200,
    1207060260, 1207060320, 1207060380, 1207060440, 1207060500, 1207060560,
    1207060620, 1207060680, 1207060740, 1207060800, 1207060860, 1207060920,
    1207060980, 1207061040, 1207061100, 1207061160, 1207061220, 1207061280,
    1207061340, 1207061400, 1207061460, 1207061520, 1207061580, 1207061640,
    1207061700, 1207061760, 1207061820, 1207061880, 1207061940, 1207062000,
    1207062060, 1207062120, 1207062180, 1207062240, 1207062300, 1207062360,
    1207062420, 1207062480, 1207062540, 1207062600, 1207062660, 1207062720,
    1207062780, 1207062840, 1207062900, 1207062960, 1207063020, 1207063080,
    1207063140, 1207063200, 1207063260, 1207063320, 1207063380, 1207063440,
    1207063500, 1207063560, 1207063620, 1207063680, 1207063740, 1207063800,
    1207063860, 1207063920, 1207063980, 1207064040, 1207064100, 1207064160,
    1207064220, 1207064280, 1207064340, 1207064400, 1207064460, 1207064520,

```

Bjørnøya : ncdump


```
Bjørnøya : ncview — Konsole
File Edit View Bookmarks Settings Help
-rw-rw-r-- 1 steingod steingod 2,6M mai 1 2015 radflux_bjornoya-201504.dat
-rw-rw-r-- 1 steingod steingod 1,2M mai 1 2015 radflux_bjornoya-201504.nc
-rw-rw-r-- 1 steingod steingod 2,6M juni 1 2015 radflux_bjornoya-201505.dat
-rw-rw-r-- 1 steingod steingod 1,2M juni 1 2015 radflux_bjornoya-201505.nc
-rw-rw-r-- 1 steingod steingod 2,4M juli 1 2015 radflux_bjornoya-201506.dat
-rw-rw-r-- 1 steingod steingod 1,1M juli 1 2015 radflux_bjornoya-201506.nc
-rw-rw-r-- 1 steingod steingod 2,4M aug. 1 2015 radflux_bjornoya-201507.dat
-rw-rw-r-- 1 steingod steingod 1,1M aug. 1 2015 radflux_bjornoya-201507.nc
-rw-rw-r-- 1 steingod steingod 2,5M sep. 1 2015 radflux_bjornoya-201508.dat
-rw-rw-r-- 1 steingod steingod 1,2M sep. 1 2015 radflux_bjornoya-201508.nc
-rw-rw-r-- 1 steingod steingod 2,4M okt. 1 2015 radflux_bjornoya-201509.dat
-rw-rw-r-- 1 steingod steingod 2,2M nov. 1 2015 radflux_bjornoya-201510.dat
-rw-rw-r-- 1 steingod steingod 2,6M des. 1 2015 radflux_bjornoya-201511.dat
-rw-rw-r-- 1 steingod steingod 1,2M des. 1 2015 radflux_bjornoya-201511.nc
-rw-rw-r-- 1 steingod steingod 1,4M jan. 1 2016 radflux_bjornoya-201512.dat
-rw-rw-r-- 1 steingod steingod 616K jan. 1 2016 radflux_bjornoya-201512.nc
-rw-rw-r-- 1 steingod steingod 15M aug. 16 2016 radflux_bjornoya-2016.cdl
-rw-rw-r-- 1 steingod steingod 21M aug. 16 2016 radflux_bjornoya-2016.dat
-rw-rw-r-- 1 steingod steingod 9,4M aug. 16 2016 radflux_bjornoya-2016.nc
-rw-rw-r-- 1 steingod steingod 103M okt. 14 2016 radflux_bjornoya.nc
-rw-rw-r-- 1 steingod steingod 73M aug. 29 2013 radflux_Bjornoya.nc
-rw-rw-r-- 1 steingod steingod 32M aug. 29 2013 radflux_Bjornoya.tgz
-rwxr-xr-x 1 steingod steingod 322 juni 6 2008 radobs_collection_status.txt*
drwxr-xr-x 2 steingod steingod 4,0K juni 25 2013 tmp/
drwxr-xr-x 2 steingod steingod 4,0K juni 25 2013 tmp2/
steingod@tuba:/disk1/data/radflux/Bjornoya$ ncdump radflux_bjornoya.nc | m
steingod@tuba:/disk1/data/radflux/Bjornoya$ ncview radflux_bjornoya.nc
Ncview 2.1.6 David W. Pierce 29 Oct 2015
http://meteora.ucsd.edu:80/~pierce/ncview_home_page.html
Copyright (C) 1993 through 2015, David W. Pierce
Ncview comes with ABSOLUTELY NO WARRANTY; for details type `ncview -w'.
This is free software licensed under the Gnu General Public License version 3; t
ype `ncview -c' for redistribution details.

Warning: Cannot convert string "--helvetica*-r*-14*-14*-14*-14*" to type F
ontStruct
Note: 43208 missing values were eliminated along axis "time"; index= 1 2 3 4 5
6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 4294967298..
```

no variable selected

Ncview 2.1.6 David W. Pierce 29 Oct 2015

*** SELECT A VARIABLE TO START ***

Current: x=20-May-2008 10:10:08, y=904.478

Quit ->1 << < || > >> Edit ? Delay: ☐ Opts

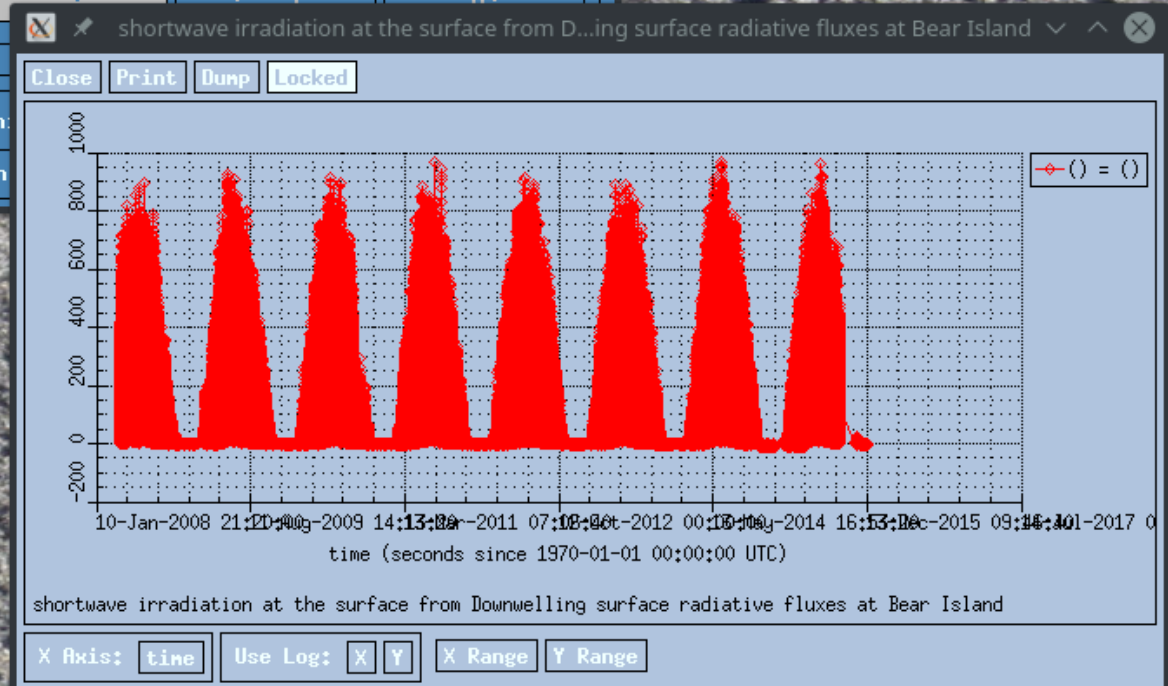
3gauss Inv P Inv C Mag X1 Linear Axes Range Bi-lin Print

Var: stationid

dlsenstemp

Din: Name: Min

strlen25 Min





How to publish data

Øystein Godøy

The FAIR Guiding Principles for scientific data management and stewardship

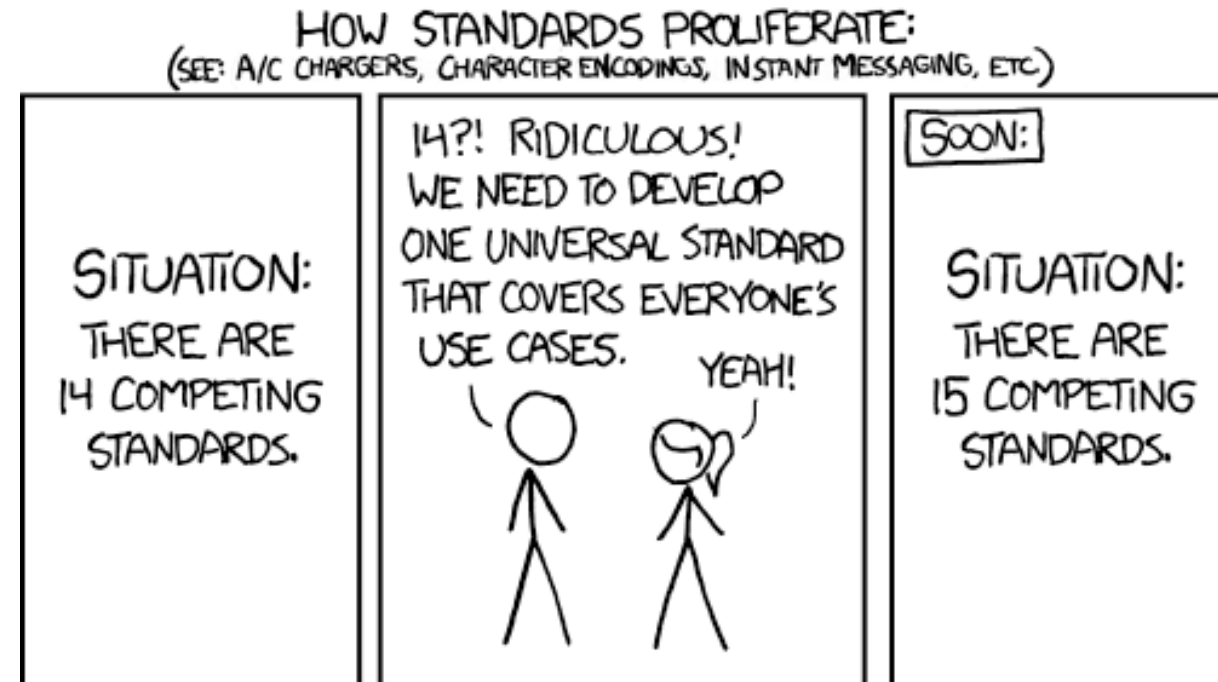
- To be Findable:
 - F1. (meta)data are assigned a globally unique and persistent identifier
 - F2. data are described with rich metadata (defined by R1 below)
 - F3. metadata clearly and explicitly include the identifier of the data it describes
 - F4. (meta)data are registered or indexed in a searchable resource
- To be Accessible:
 - A1. (meta)data are retrievable by their identifier using a standardized communications protocol
 - A1.1 the protocol is open, free, and universally implementable
 - A1.2 the protocol allows for an authentication and authorization procedure, where necessary
 - A2. metadata are accessible, even when the data are no longer available
- To be Interoperable:
 - I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
 - I2. (meta)data use vocabularies that follow FAIR principles
 - I3. (meta)data include qualified references to other (meta)data
- To be Reusable:
 - R1. meta(data) are richly described with a plurality of accurate and relevant attributes
 - R1.1. (meta)data are released with a clear and accessible data usage license
 - R1.2. (meta)data are associated with detailed provenance
 - R1.3. (meta)data meet domain-relevant community standards

Data Formats: Choosing and Adopting Community Accepted Standards

- Most projects (rightly so) focus on the content of their data files, you need to consider the format as well.
- Since you captured or created the data, and stored them in your own files, you know
 - how the data are organized,
 - how to read them,
 - how to use them,
 - characteristics of the data that could constrain their use.
- The goal of a good data format is to make it easier for others to read the data too.
- Many hours have gone into developing standards for formats – try to learn from them.

Why use community standards

- If you try to develop your data format from scratch, you will forget something.
- Build on the experience and improvements built into the community standards over years of use.
- Tools and analysis software natively support reading community standard data.
- Reduce development effort and support reuse.
- Positive feedback – they are more likely to be adopted by others.



<http://xkcd.com/927/>

Use self describing data formats

- Self-describing data formats have become a well accepted way of archiving and disseminating scientific data.
- Before self-describing data formats became widely used, each project often invented their own data formats, often raw binary or even ASCII.
- These approaches had a number of problems:
 - Machine dependent byte ordering or floating point organizations
 - Required a 'key' to be able to open the file and read the right data.
 - A new custom reader is needed for each different data organization. Working in a new language could be very difficult since you have to redevelop the reader anew.

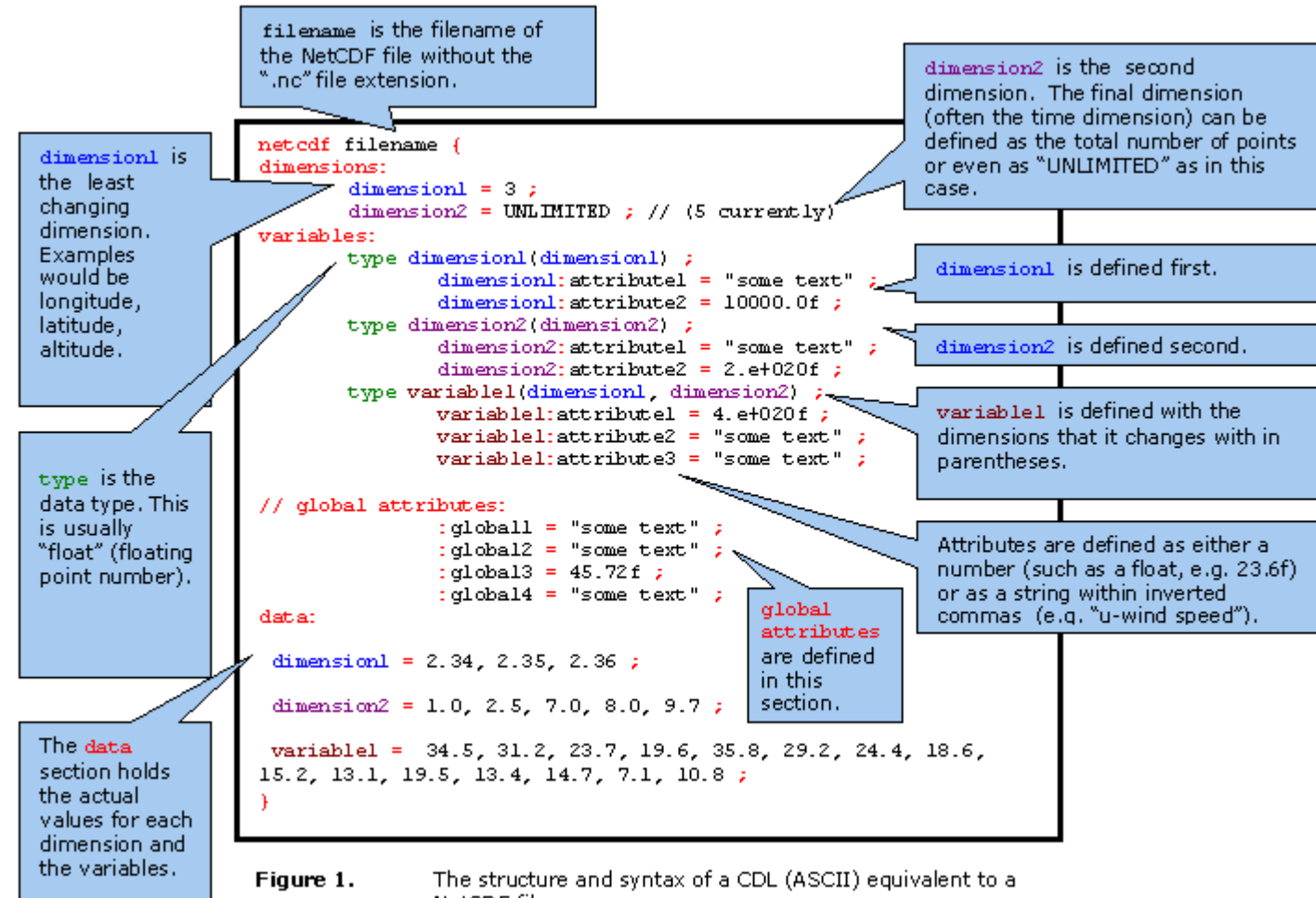


Figure 1.

The structure and syntax of a CDL (ASCII) equivalent to a NetCDF file.

http://artefacts.ceda.ac.uk/formats/netcdf/index_cf.html

FileEditViewHistoryBookmarksToolsHelp

Dataset Validation | NorDataNet X

←→↺🏠

https://www.nordatanet.no/dataset_validation/form

...🔒🌟🔍 Search

⚙️ Most Visited🔗 Getting Started🌐 NorDataNet | Arctic ...📄 XPath and XSLT with I...🌐 Parsing XML and HTM...🌐 xslt.md.html

🏠 DashboardContentStructureAppearancePeopleModulesConfigurationReportsHelp

👤 Hello steingod🔒 Log out

Add contentFind contentPerformanceBlocksEdit shortcuts

HomeAboutSearch for dataSubmit dataSupportMy account

NorDataNet

Norwegian Scientific Data Network

Dataset Validation

Webform for validation of netCDF files based on the [IOOS compliance checker](#)

Select the test you want to run *

☐ CF-1.6☐ ACDD

Upload Your File *

Browse...

No file selected.

Upload

You can only upload a single netCDF file with ".nc" extension, with a maximum size of 1500M. You need to upload a bigger file, take contact with the website support directly.

Submit

NorData
Norwegian Scientific Data Network

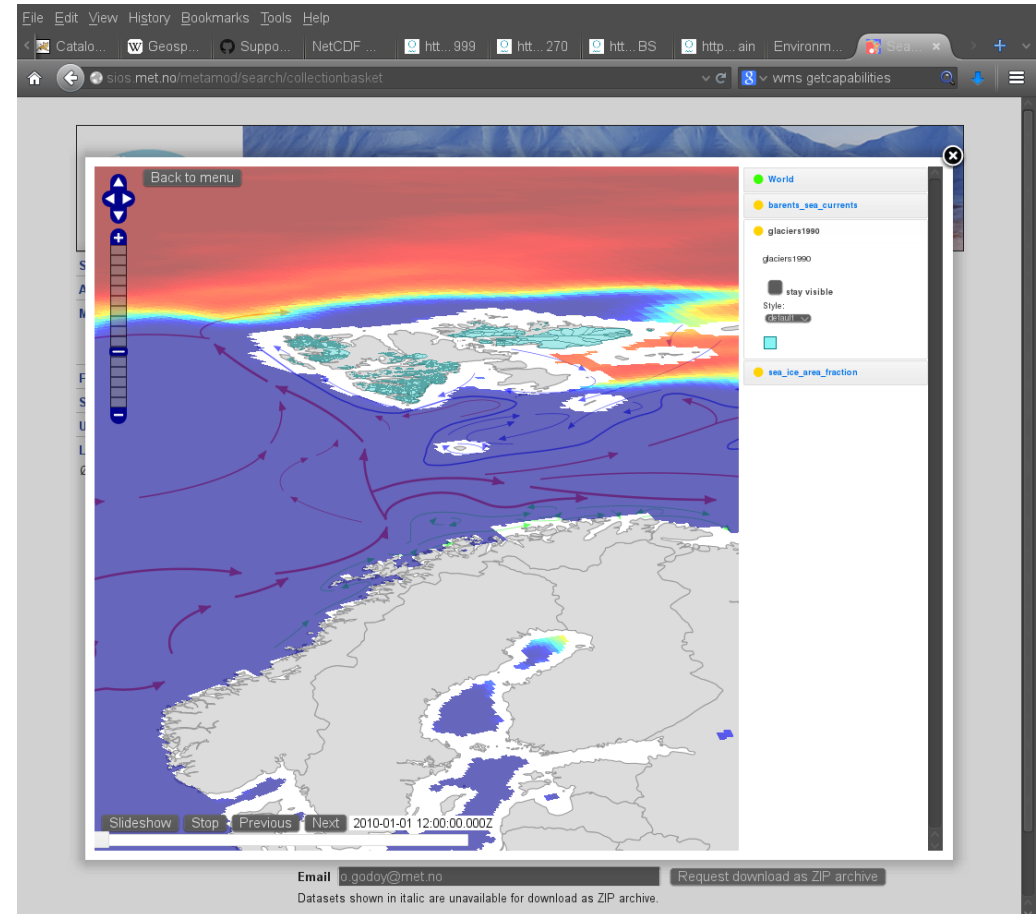
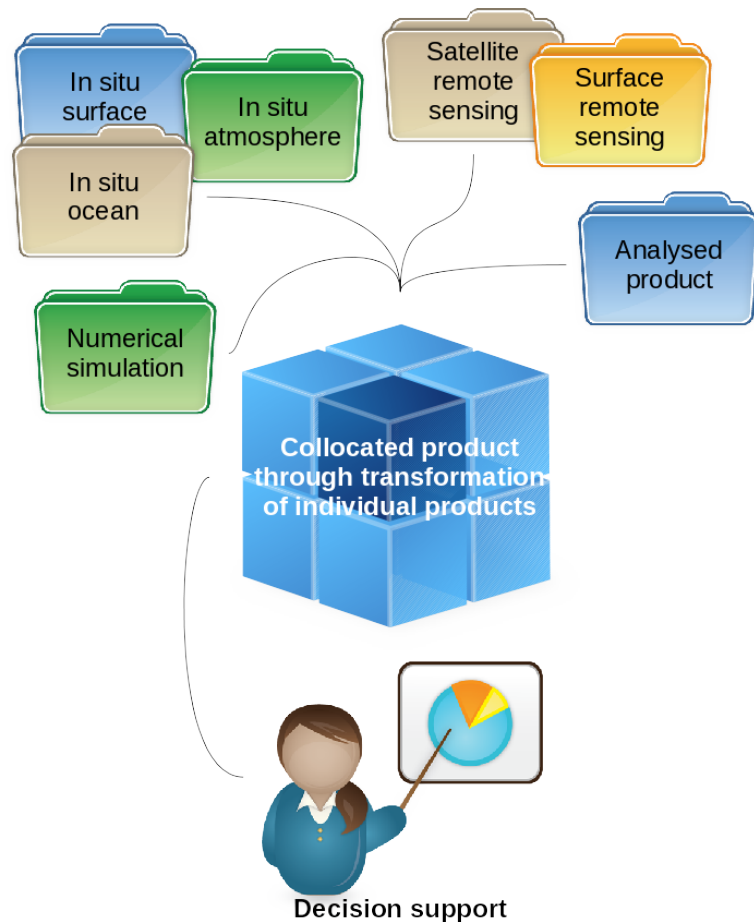
https://www.nordatanet.no

Long term data archives



DataSpan

Combining data







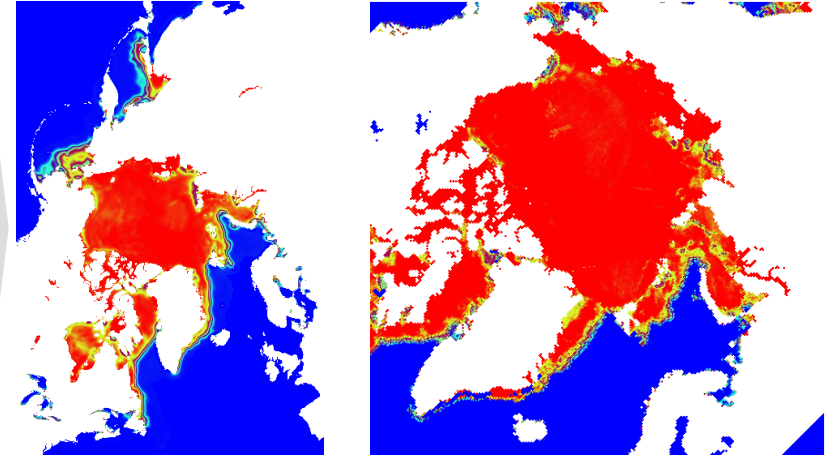
Various slides

Øystein Godøy

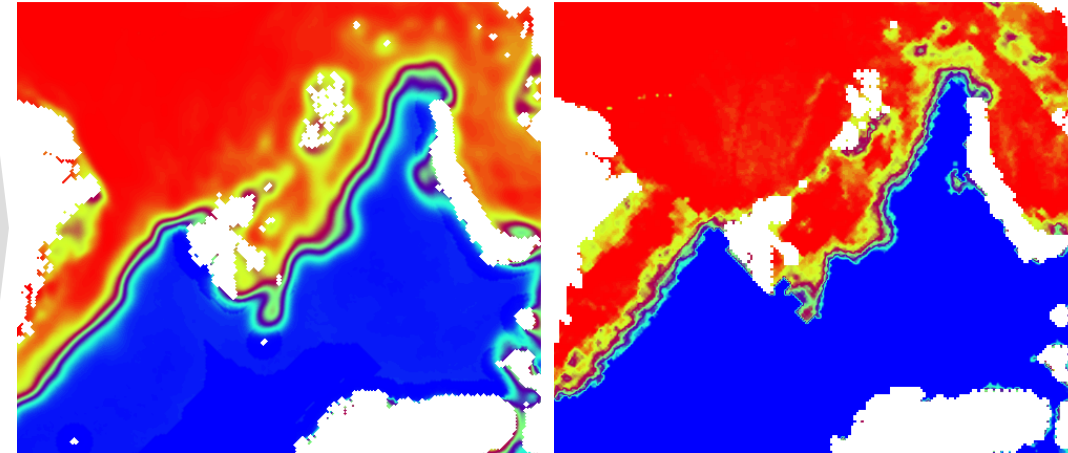
Transformations

Transformations allow users to do comparisons of products and to extract tailored products for their specific need

Search results



Transformation request



NetCDF

- NetCDF is a container you can put almost anything into
- Standardised formulations exist for
 - Gridded data
 - Timeseries
 - Trajectories
 - Including time series of trajectories
 - Profiles
 - Including time series of profiles
 - Geometries (next release)
- Can easily be served as data streams using OPeNDAP
- Can be integrated directly in tools like
 - R
 - Matlab
 - Ferret
 - Python
 - Excel
 - Check e.g. NETCDF4Excel on GitHub
- If you access data through OPeNDAP you do not have to download data

Best Practices

data.xls

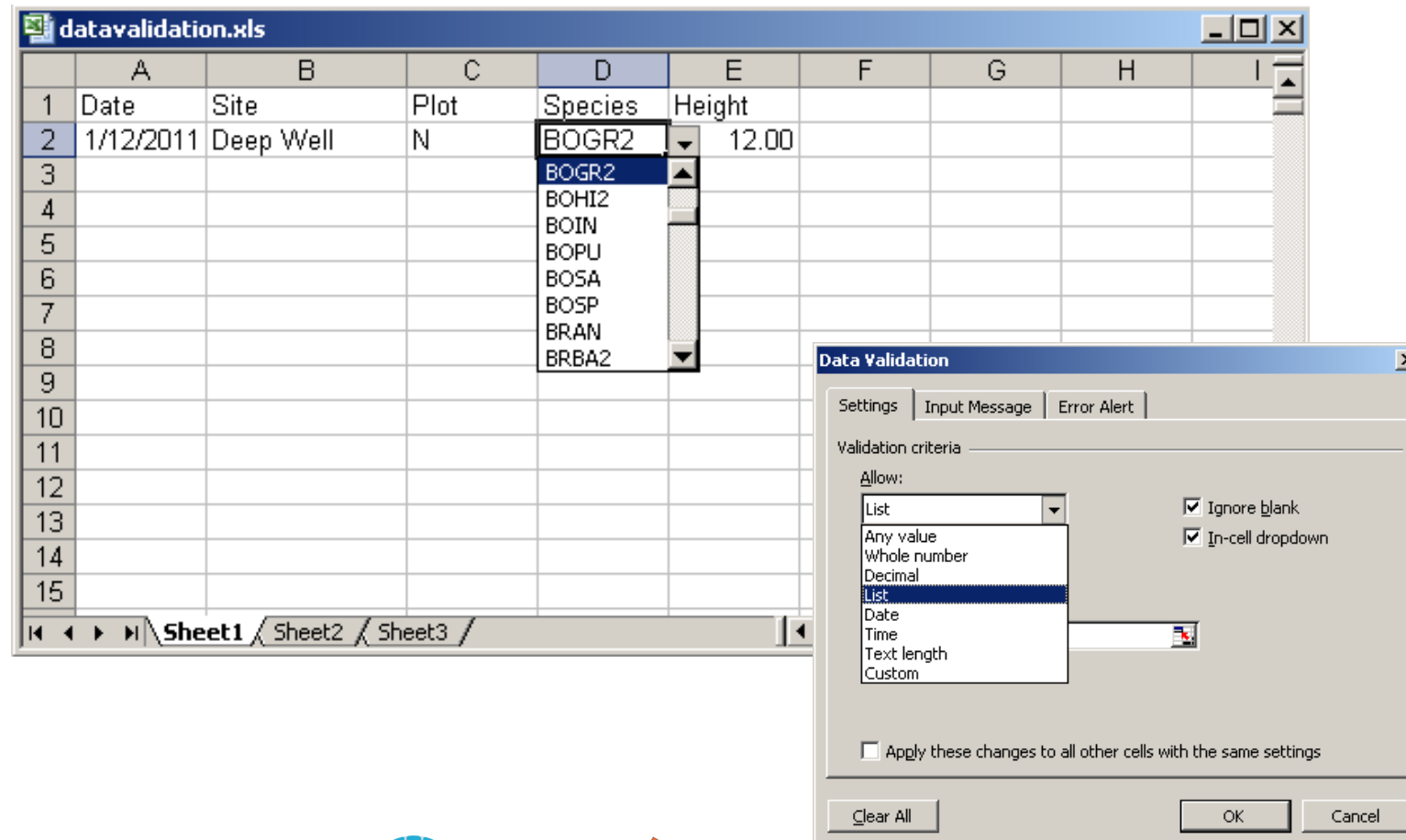
	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Site	Date	Plot	Species	Weight	Adult		Rodent Trapping 3/15/2010						
2	DeepWell	2/13/2010		1 DIPO	12.1	j		Site	Plot	Adult	RodentSp	Weight		
3	Deep Well	Feb-10		2 Pero	13.22	j		DW		1 y	Pero	12		
4	rioSalado	2/13/2010	1a	pero	16	N		RS		2 j	PERO	escaped <15		
5	riuSladu	"	1*	CleGap	18.92	gut away		RS		3 ri	Clegap	91		
6				Mean1	15.06									
7														
8														
9														
10														
11														
12	Rodent Trapping	MJK & ALN	10-Apr-10											
13	Site	Plot	Adult	Species	grams	Ccmments								
14	deep well		1 y	woodrat	13									
15	riosalado		2 y	PERO	24.5									
16	riosalado		3 y	Clegap	91									
17														
18														
19														
20														

SEV_SmallMammalData_v.5.25.2010.xls

	A	B	C	D	E	F	G	H
1	Date	Site	Plot	Species	Weight	Adult	Comments	
2	2/5/2010	Deep Well		1 DIPO	13.2	y		
3	2/4/2010	Deep Well		1 CLEGAP	11.6	j		
4	2/5/2010	Rio Salado		1 DIPO	14.2	y		
5	2/5/2010	Rio Salado		2 PERO	10.1	y		
6	3/15/2010	Deep Well		1 DIPO	15.2	y	plot burned	
7	3/15/2010	Deep Well		2 DIPO	21.7	y	pregnant	
8	3/15/2010	Rio Salado		1 CLEGAP	16.2	j		
9								
10								
11								
12								
13								

- Columns of data are consistent: only numbers, dates, or text
- Consistent Names, Codes, Formats (date) used in each column
- Data are all in one table, which is much easier for a statistical program to work with than multiple small tables which each require human intervention

If you use Excel or similar





D5																
1	Glacier	Holtedahlfonna	Weather Conditions	5 m s-1 wind, clear sky, sampling for C2S3 project												
2	Area	Ablation, Stake 2	Coordinates/GP	North			78.931367									
3	Date	17.04.16		East			13.304033									
4	Time	afternoon		Altitude (m a.s.l.)												
5	Observer(s)	TM, JCG, AS, EB														
6	General comments	110 cm pit														
7		Sampling for Iso, Ions, Bio and BC														
8																
9	Temperatures:															
10	Air temp 1m [°C]		-14.5													
11	Air/snow temp [°C]															
12	Snow/ice temp [°C]		-8.2													
13	Comments		stratigraphy by JCG and AS													
14																
15	Layer, from top to bottom	Top [cm]	Bottom [cm]	Depth center	Type 1	Type 2	Diameter min [mm]	Diameter max [mm]	Hardness	Hardness code	Depth Center [cm]	Density [g/cm³]	Depth [cm]	Temp [deg C]		
16	1	0	1	0.5	wind packed	and sun crust	0.2	0.2	1f	3	4	136	5	-11.6		
17	2	1	8	4.5	wind broken		0.1	0.3	Fist	1	14	309	8	-12.1		
18	3	8	20	14	faceted	clusters	0.1	0.1	1f	3	23	334	10	-12.1		
19	4	20	26	23	rounded and faceted		0.1	0.3	pencil	4	32	381	20	-11.3		
20	5	26	27	26.5	melt refreeze	clusters			knife	5	39	482	26	-11.7		
21	6	27	35	31	faceted		0.3	0.3	pencil	4	47	475	35	-11.1		
22	7	35	37	36	depth hoar	clusters	1	1	4 f	2	58	306	37	-11.1		
23	8	37	42	39.5	faceted		0.4	0.4	pencil	4	62	394	42	-10.8		
24	9	42	57	49.5	faceted		0.5	0.5	1f	3	66	361	50	-10.3		
25	10	57	60	58.5	depth hoar		1	1	4 f	2	75	325	57	-10		
26	11	60	63	61.5	faceted		0.4	0.4	1f	3	82	265	60	-9.9		
27	12	63	71	67	melt refreeze				pencil	4	100	207	71	-9.5		
28	13	71	73	72	faceted		1	1	4 f	2			79	-9.3		
29	14	73	79	76	melt refreeze				knife	5			85	-9.1		
30	15	79	80	79.5	horizontal ice layer				ice	6			91	-8.7		
31	16	80	85	82.5	faceted		0.5	0.5	1f	3			100	-8.6		
32	17	85	88	86.5	depth hoar		1.5	1.5	4 f	2			108	-8.5		
33	18	88	91	89.5	melt refreeze				pencil	4			110	-8.2		
34	19	91	108	99.5	depth hoar		2	2	Fist	1						
35	20	108	110	109	horizontal ice layer				ice	6						
36																
37																
38																
39																
40																
41																
42																
43																
44																



C13 stratigraphy by AS

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	Glacier	Kongsvegen	Weather Condition	Little wind, 13 minus, almost fully clear, nice day in general, sampling for C2S3 project													
2	Area	Ablation, Stake 2	Coordinates/GPS	North	8751951.45	78.83018											
3	Date	13.04.16		East	451534.753	12.75838											
4	Time	afternoon		Altitude (m a.s.l.)	230.728												
5	Observer(s)	TM, JCG, AS, EB, HH															
6	General comments	Little wind, partly cloudy, 20 cm snow pit, not much melt refrozen stuff, most likely all is superimposed ice															
7		Sampling for Iso, Ions, Bio and BC															
8																	
9	Temperatures:																
10	Air temp 1m [°C]		-13.9														
11	Air/snow temp [°C]																
12	Snow/ice temp [°C]		-11.6														
13	Comments	stratigraphy by AS															
14																	
15	Layer, from top to bottom	Top [cm]	Bottom [cm]	Depth center	Type 1	Type 2	Diameter min [mm]	Diameter max [mm]	Hardness	Hardness code	Depth Center [cm]	Density [g/cm ³]	Depth [cm]	Temp [deg C]			
16	1	0	1	0.5	wind packed		0.1	0.1	pencil	4	2	147	1	-12.6			
17	2	1	5	3	wind broken		0.2	0.3	1 finger	3	10	318	5	-12.4			
18	3	5	16	10.5	clustered rounded		0.3	0.4	pencil	4	18	224	16	-11.6			
19	4	16	20	18	faceted				1 finger	4			20	-11.6			
20																	
21																	
22																	
23																	
24																	
25																	
26																	
27																	
28																	
29																	
30																	
31																	
32																	
33																	
34																	
35																	
36																	
37																	
38																	
39																	
40																	
41																	
42																	
43																	
44																	

Preferably do not split content over columns and cells

Does this belong under general comments?

Use full names

What is type? Explain vocabularies.

How to interpret, which standard?

File

Edit

View

History

Bookmarks

Tools

Help

<

Metel

Unive

cryo.r

Data

Clima

Confi

Test aggr

Dash

Kalen

timeS

cf-cor

NetCDF C

Dw X

Dataset h

Dw Data

>

+

▼

←

→

↶

🏠

🔒

https://data.gbif.no/dwcexcel/

...

📁

☆

🔍 Search

⚙️ Most Visited

🔗 Getting Started

🌐 NorDataNet | Arctic D...

📄 XPath and XSLT with l...

🌐 Parsing XML and HTM...

🌐 xslt.md.html

DwC Excel Template Generator

English ▼

Set language

Check the boxes next to the terms you want to include in your template and click the **Create template** button.

[Darwin Core Terms: A quick reference guide](#)

RECORDLEVEL

☐ modified

☐ institutionID

☐ collectionID

☐ datasetID

☐ institutionCode

☐ collectionCode

☐ datasetName

☐ ownerInstitutionCode

☒ basisOfRecord

☐ informationWithheld

☐ dataGeneralizations

☐ dynamicProperties

OCCURRENCE

☒ occurrenceID

☐ catalogNumber

☐ recordNumber

☐ recordedBy

☒ individualCount

☒ organismQuantity

☒ organismQuantityType

☐ sex

☐ lifeStage

☐ reproductiveCondition

☐ behavior

☐ establishmentMeans

institutionCode [Institution Code]

The name (or acronym) in use by the institution having custody of the object(s) or information referred to in the record.

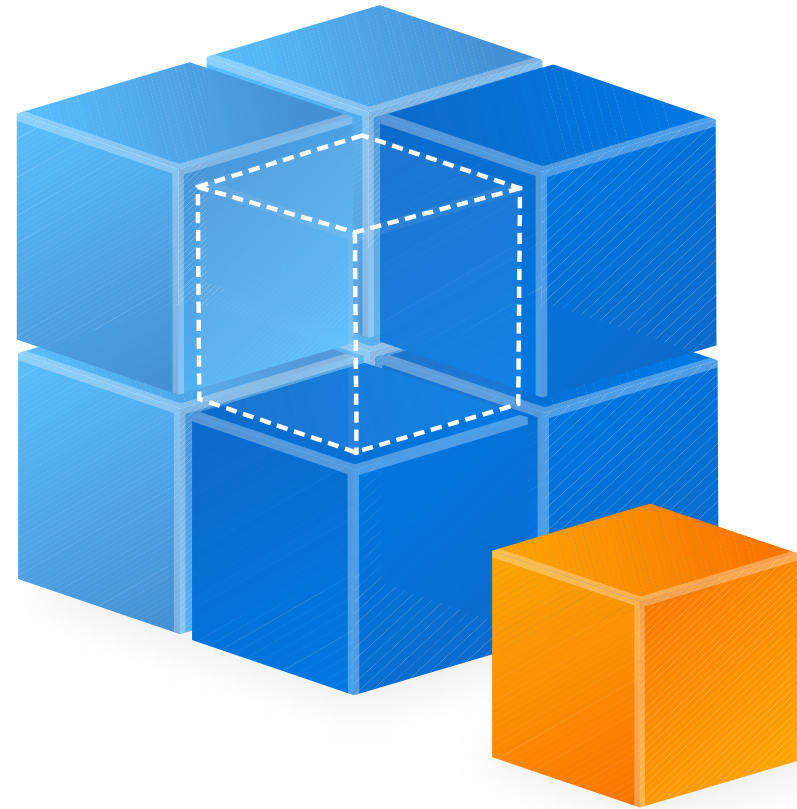
Examples: "MVZ", "FMNH", "AKN-CLO", "University of California Museum of Paleontology (UCMP)"

NorDataNet

Norwegian Scientific Data Network

Specific Arctic challenges

- Monitoring by both scientific and operational communities
 - Collaboration needed
- Scientific communities good at archiving
- Operational communities good at real time exchange and combining efforts
- Bridging needed



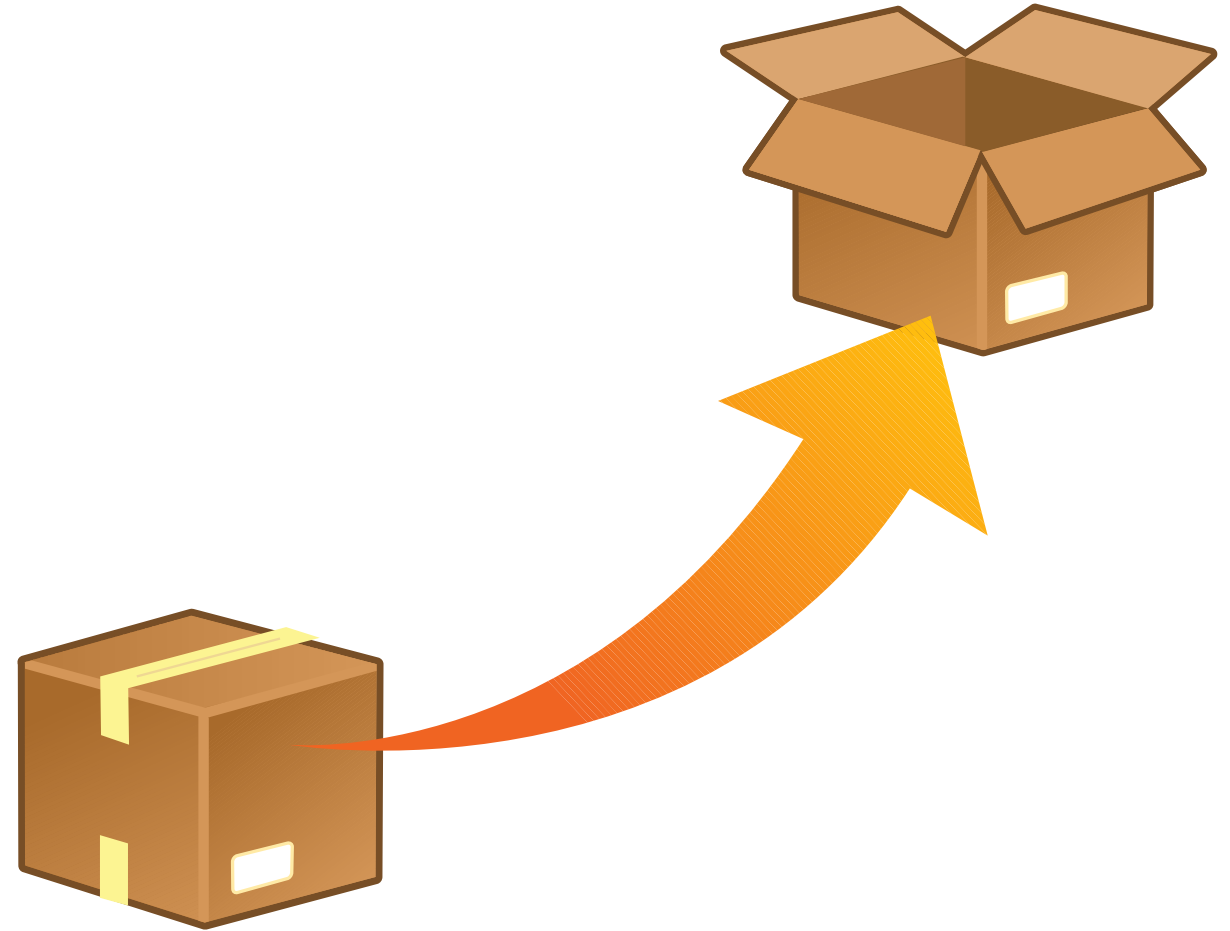
Summary

- Document and share data
- Use standards
 - And contribute
- Track usage
 - Across communities



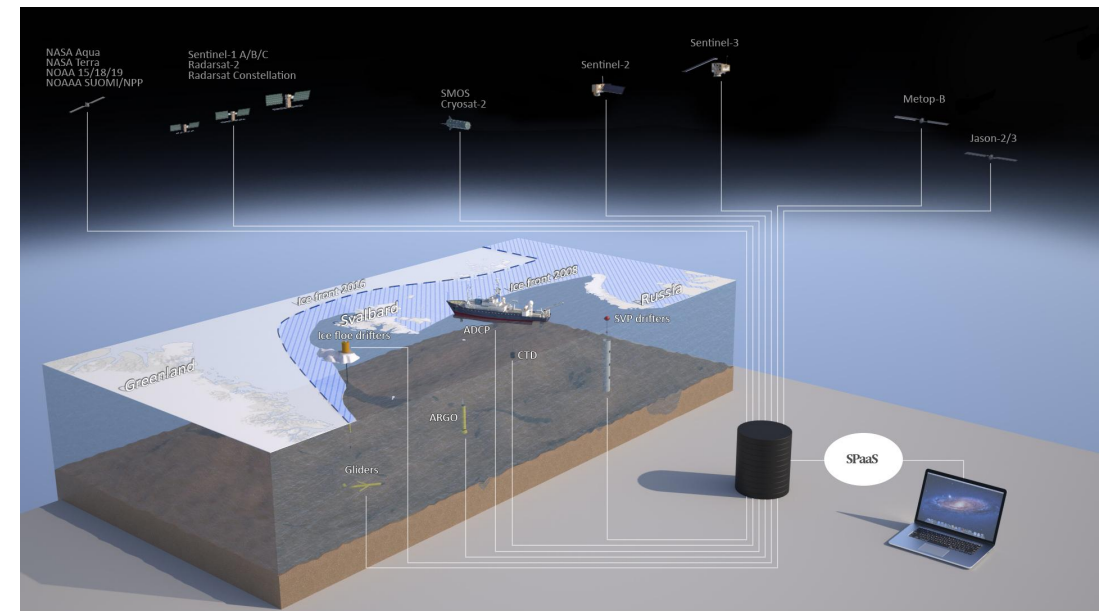
Issues for discussion

- Sharing data
 - Bottlenecks sharing data in a proper form?
 - Gaps in education?
 - Mitigation?
- Experience using data generated by others?
 - Bottlenecks and mitigation
- Standardisation
 - Necessary?
 - How to navigate?



Moving towards

- Data management required by funding agencies
- Integration of data centres
- Work flow management
- Funding agency requirements
 - Projects must have a data plan
 - Data underlying scientific publications have to be open
 - Data plan (DCC)
 - Data summary
 - FAIR data
 - Making data findable, including provisions for metadata
 - Making data openly accessible
 - Making data interoperable
 - Increase data re-use
 - Allocation of resources
 - Data security
 - Ethical aspects
- Scientific Platforms
 - European Open Science Cloud



Courtesy of Morten W. Hansen, NERSC

Challenges

