Big Data in Meteorology

• • •

Operational Data Distribution, Processing and Archival at the German Meteorological Service

Outline

The German Meteorological Service

Your single source for weather and climate

Government organization

Basic and applied research

Operational support for specialized customers

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

DWD

DWD law

Provision of meteorological services
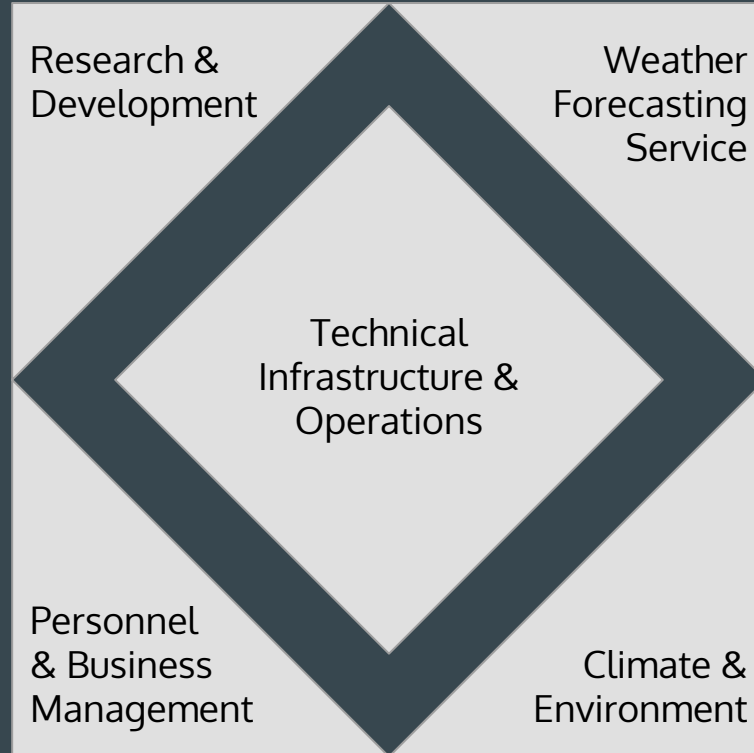
Safeguarding of aviation and shipping

Warnings about dangerous weather phenomena

Recording, monitoring, evaluation, and forecasting of the atmosphere and its interactions with other systems

Monitoring for radioactive trace elements and the forecasting of their dispersion
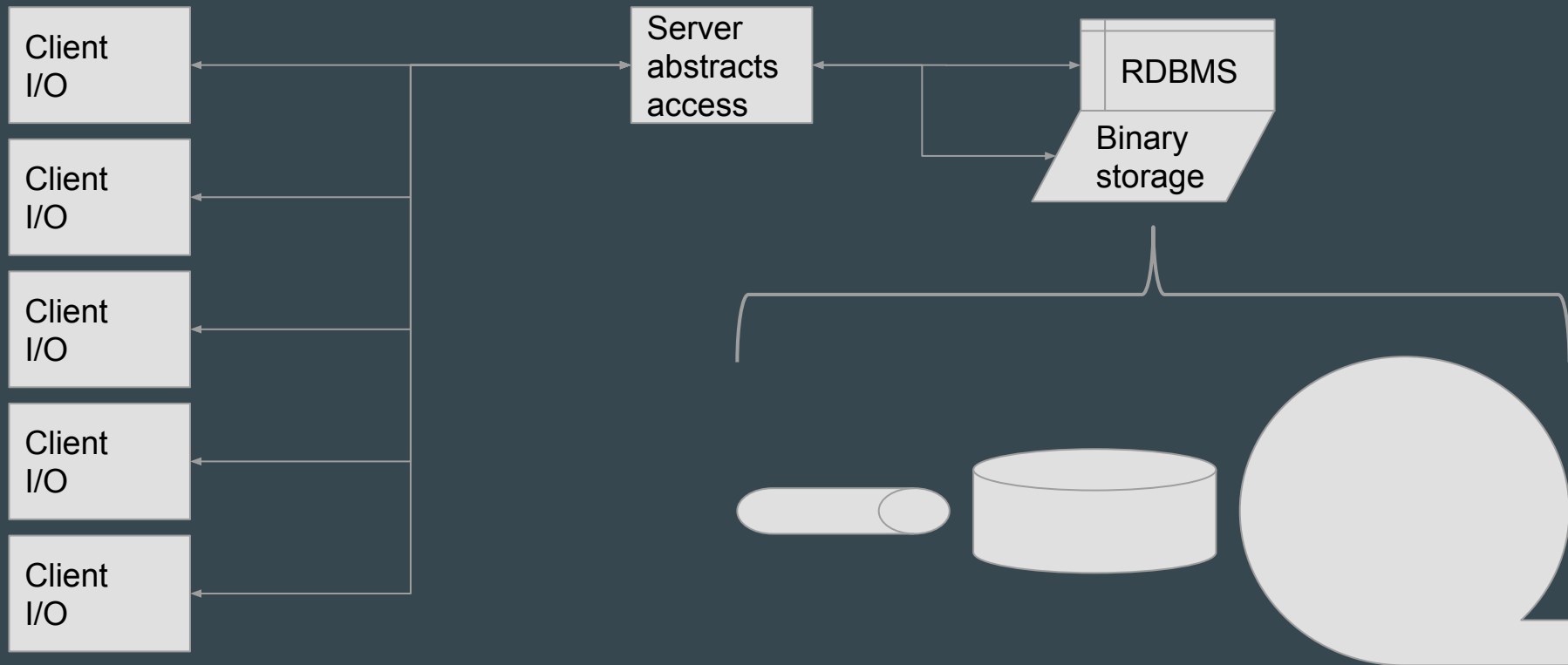
Business areas

Research &
Development

Weather
Forecasting
Service

Technical
Infrastructure &
Operations

Personnel
& Business
Management

Climate &
Environment

Volume

Yesterday's weather isn't old news

Binary data warehouse approach

Multiple storage media

Work is done on local copies

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

DWD

Proprietary data access with SKY

Client I/O

Client I/O

Client I/O

Client I/O

Client I/O

Server abstracts access

RDBMS

Binary storage

Data volume per database
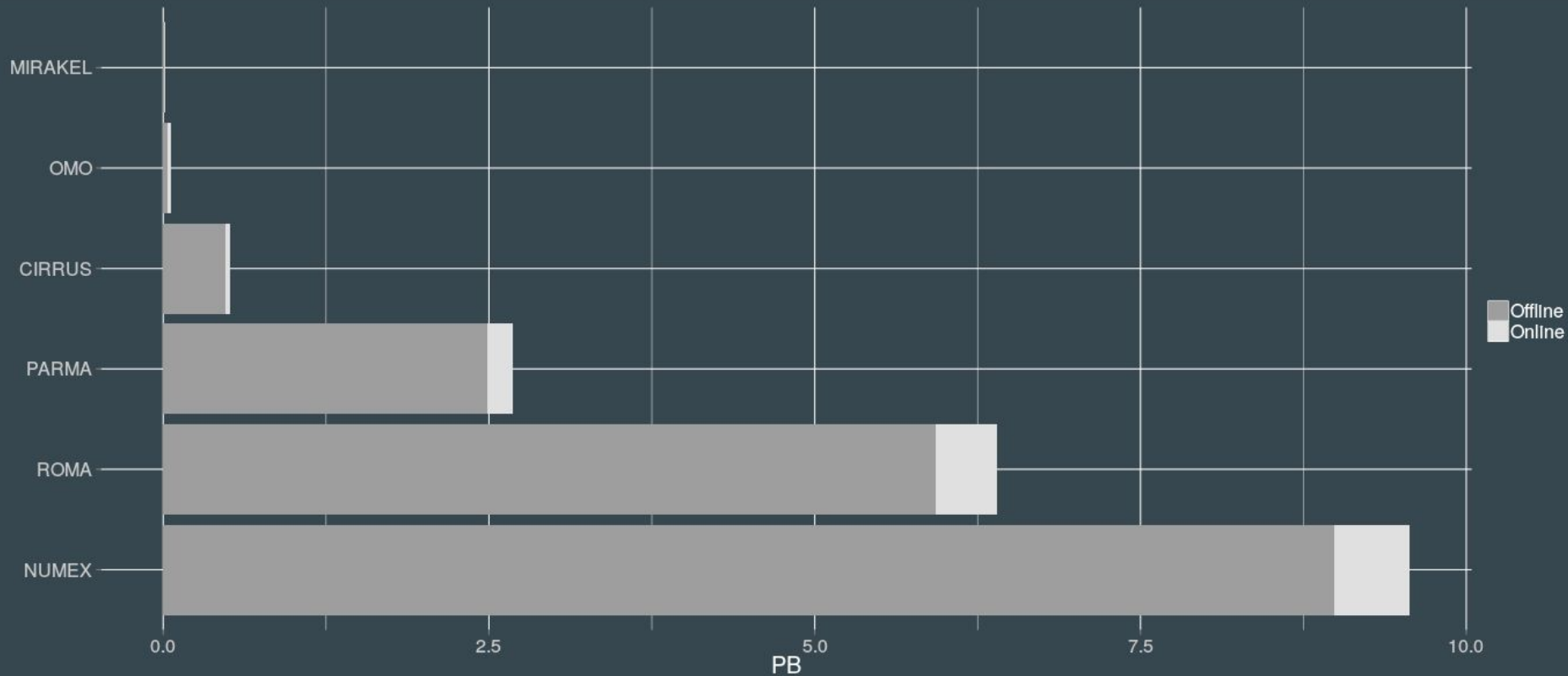
Velocity

The future's most interesting now

Users on a live update cycle

- Forecasters
- External users
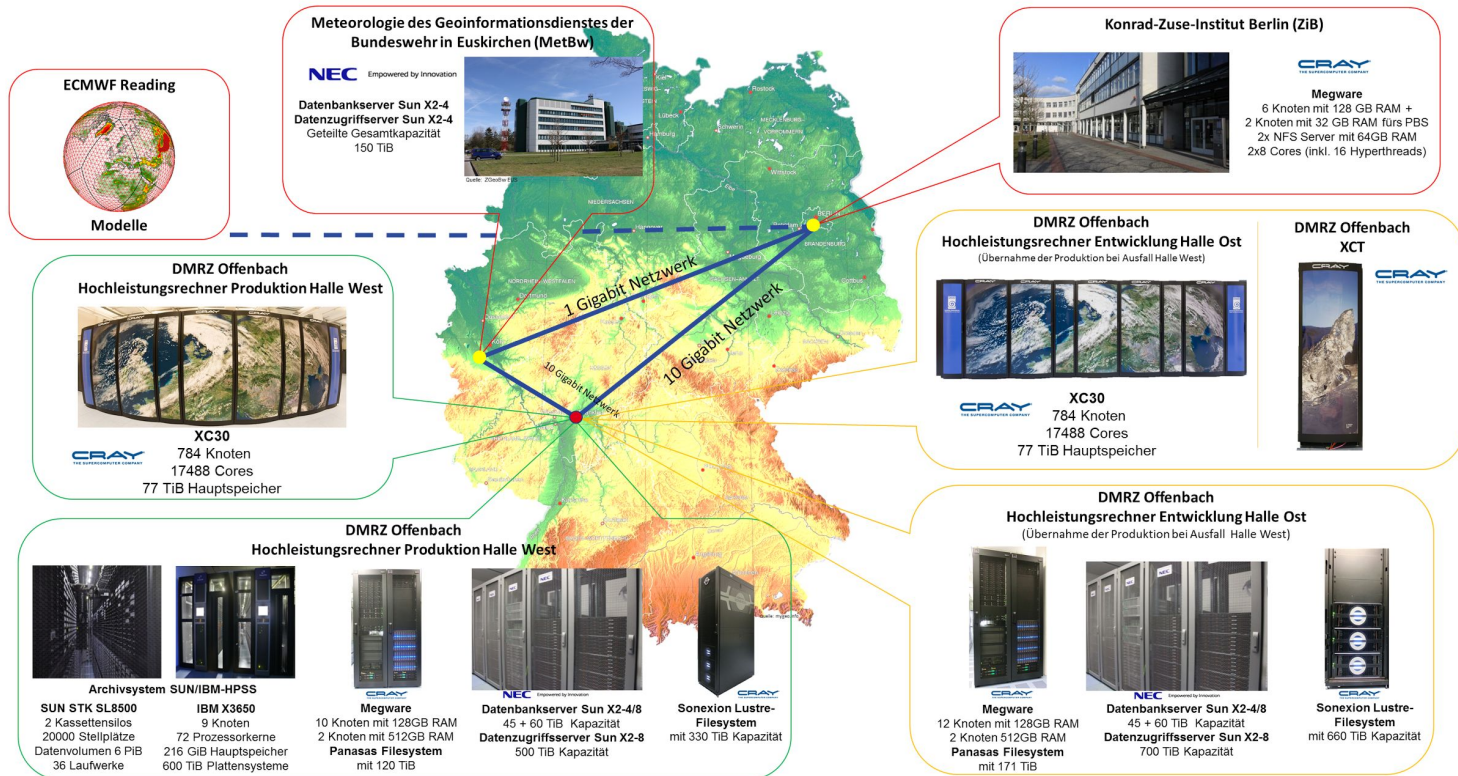- Georedundant backup

Streaming formats are used

Deutscher Wetterdienst
Wetter und Klima aus einer Hand
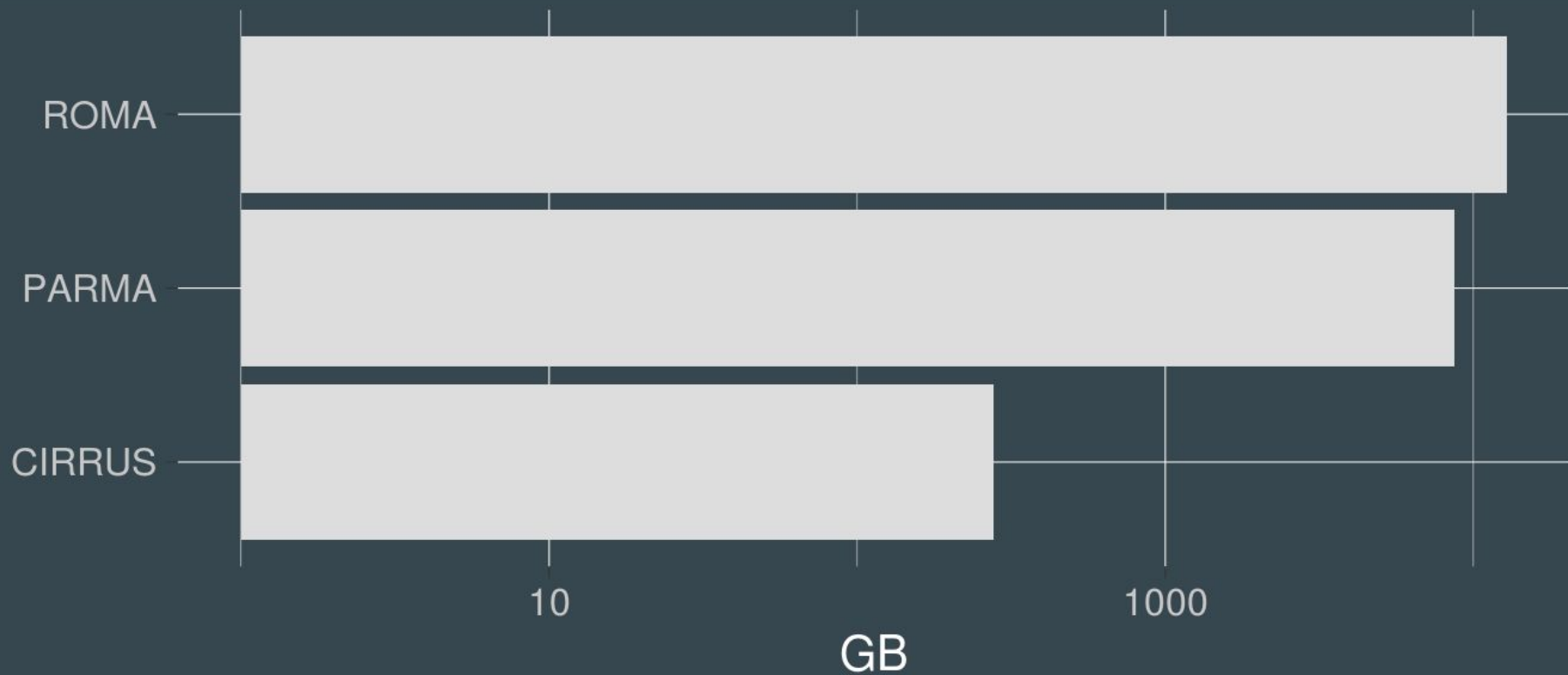
DWD

Daily model runs

- 3 nested models
- 8 runs each
- 20 member ensembles

Computing infrastructure

Daily data growth for binary DBs

Transaction log generation for MIRAKEL
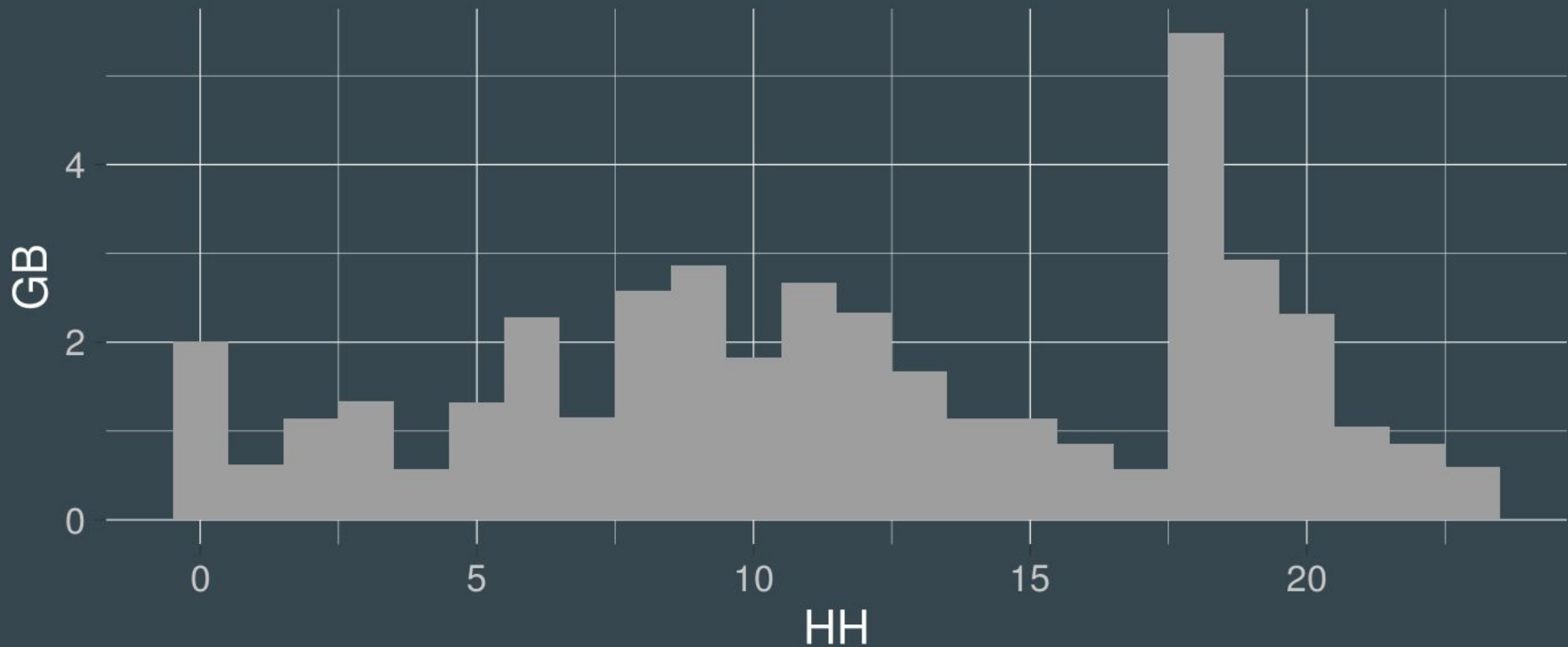
Veracity

The search for truth

Observation errors

Encoding errors

Local vocabularies

Global Observing System

POLAR ORBITING SATELLITE

GEOSTATIONARY SATELLITE

SATELLITE IMAGES

AIRCRAFT

SATELLITE GROUND STATION

WEATHER RADAR

OCEAN DATA BUOY

SATELLITE SOUNDINGS

UPPER-AIR STATION

WEATHER SHIP

SURFACE STATION

AUTOMATIC STATION

NMS

Something's bound to go wrong

QualiMET
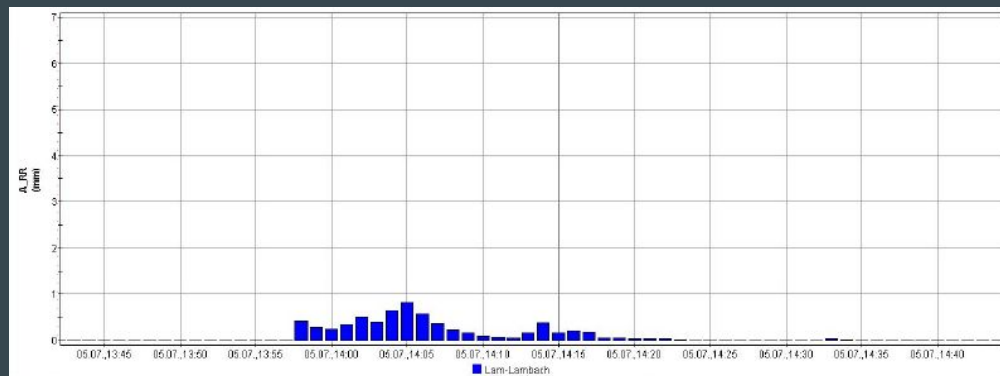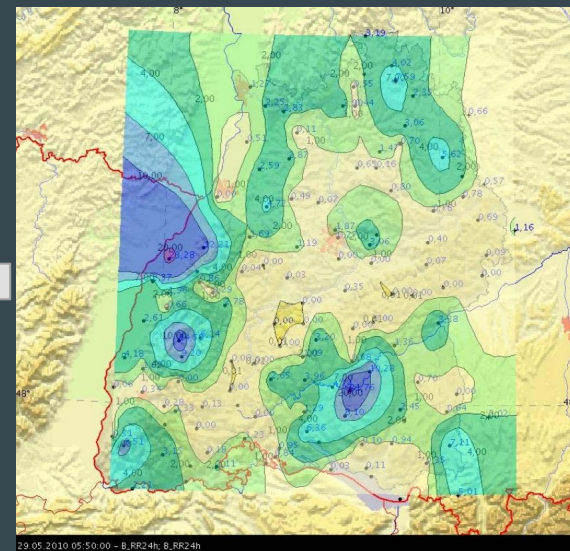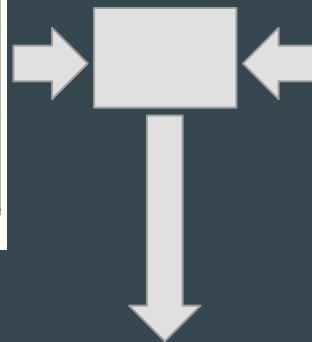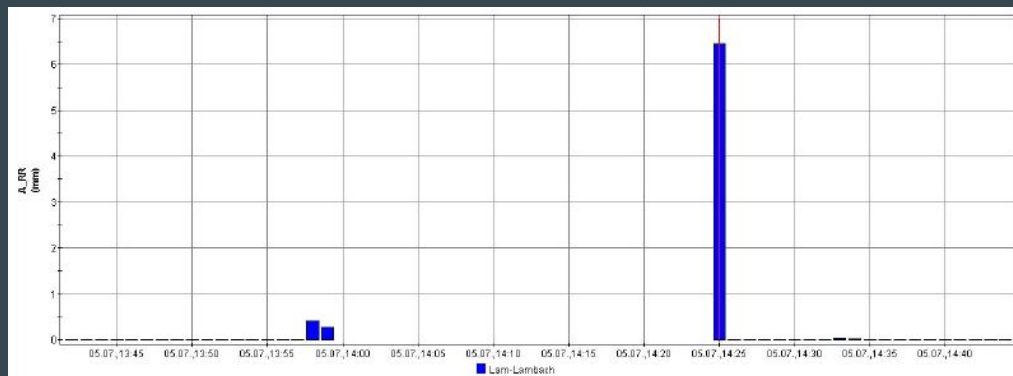
Transmission monitoring

Encoding monitoring
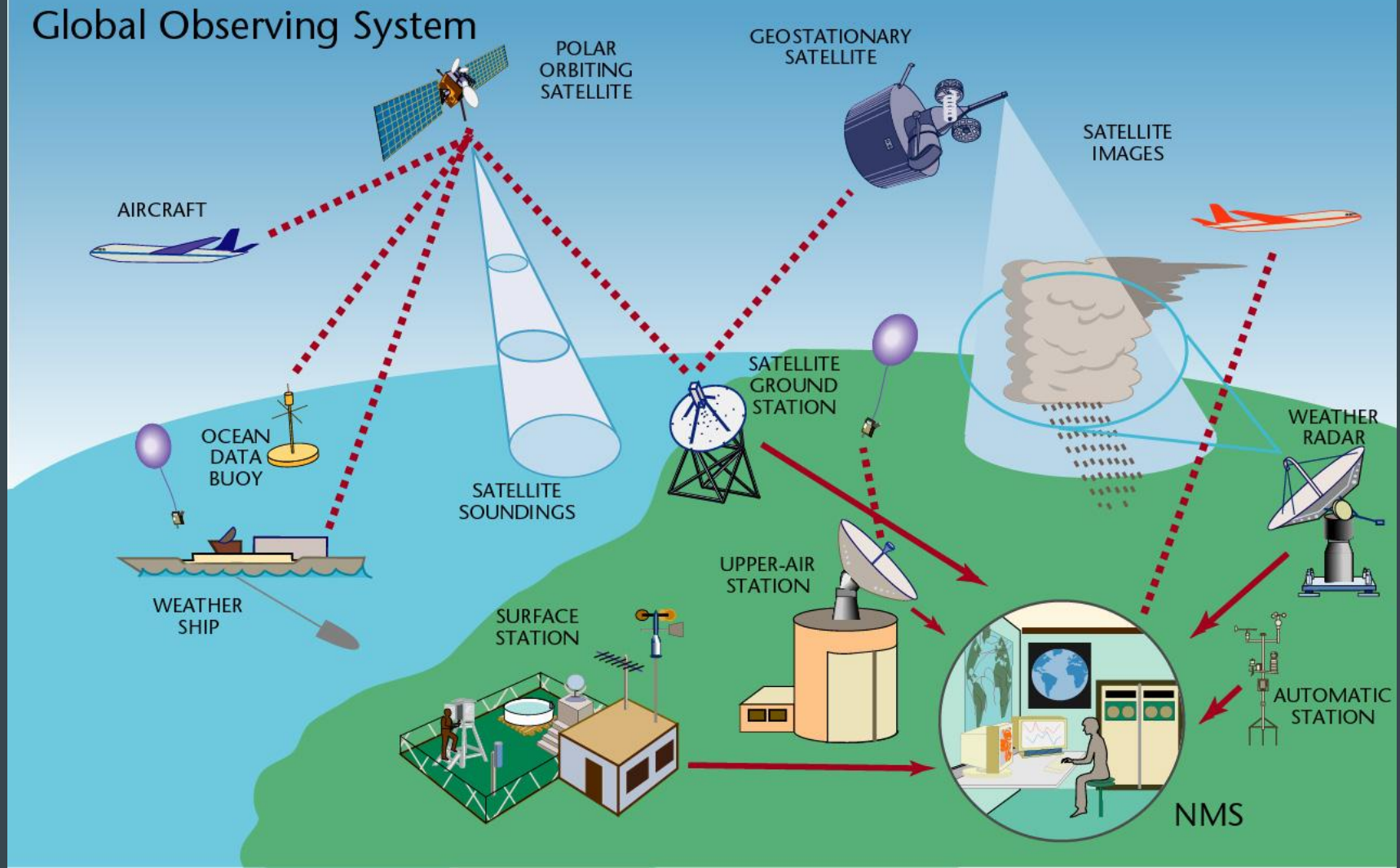
40 pair(s) from 88 Message(s) checked and 14 comparison(s) with problems

FM12 = 40 Message(s) ⟵⟶ FM94 = 48 Message(s)

SV	IIiii	Datum	UTC	iX	h	VV	N	DDD	FF	TT	TD	RH	QFE	QFF	geopH	a	pp	ww	W1	W2	Nh	Cl	Cm	Ch
FM12	40153	22.02.2016	00:00	1	0	0	9	220	12.3	7.1	6.9		894.3		1352	7	-2.3	62	9	6				
FM94	40153	22.02.2016	00:00	1	30	50		220	12.4	7.1	6.9	99	894.3		1352	7	-2.3	62	9	6				
FM12	40153	22.02.2016	06:00	1	0	800	9	270	5.1	6.9	6.8		897.0		1377	2	2.2	50	6	5				
FM94	40153	22.02.2016	06:00	1	30	800		270	5.1	6.9	6.8	99	897.0		1377	2	2.2	50	6	5				
FM12	40153	22.02.2016	09:00	1	200	3000	8	320	4.1	7.2	6.8		899.0		1395	2	2.0	60	9	6	8	4	-1	-1
FM94	40153	22.02.2016	09:00	1	250	3000	8	320	4.1	7.2	6.8	97	899.0		1395	2	2.0	60	9	6	8	4		
FM12	40153	22.02.2016	12:00	1	0	800	9	330	6.7	7.0	6.6		899.4		1399	2	0.5	60	9	5				
FM94	40153	22.02.2016	12:00	1	30	800		330	6.7	7.0	6.6	97	899.4		1399	2	0.5	60	9	5				
FM12	40153	22.02.2016	15:00	1	0	600	9	290	4.1	7.3	7.0		901.5		1418	2	2.1	60	6	5				
FM94	40153	22.02.2016	15:00	1	30	600		290	4.1	7.3	7.0	98	901.5		1418	2	2.1	60	6	5				
FM12	40153	22.02.2016	18:00	1	300	4000	7	310	3.6	6.7	6.4		903.8		1438	2	2.3	61	6	5	5	4	6	-1
FM94	40153	22.02.2016	18:00	1	450	4000	7	310	3.6	6.7	6.4	98	903.8		1438	2	2.3	61	6	5	5	4	6	
FM12	40153	22.02.2016	21:00	1	2500	10000	7	340	6.7	6.6	5.9		905.1		1450	2	1.3	1	6	2	7	0	3	-1
FM94	40153	22.02.2016	21:00	1	2550	10000	7	340	6.7	6.6	5.9	96	905.1		1450	2	1.3	1	6	2	7	0	3	
FM12	40179	22.02.2016	00:00	1	600	2000	8	260	6.2	13.6	11.6		1000.5	1004.2		7	-2.0	95	9	7	8	3	-1	-1
FM94	40179	22.02.2016	00:00	1	800	2000	8	260	6.2	13.6	11.6	88	1000.5	1004.2		7	-2.0	95	9	7	8	3		
FM12	40179	22.02.2016	09:00	1	600	15000	7	230	2.1	13.1	11.4		1004.4	1008.1		2	1.8	2	6	2	7	8	-1	-1
FM94	40179	22.02.2016	09:00	1	800	15000	7	230	2.1	13.1	11.4	90	1004.4	1008.1		2	1.8	2	6	2	7	8		

Variety

#iwokeuplikethis

Heterogeneous data received

Formatting more difficult with new partners

Diverse user needs

DWD

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

Standardized ingress

Now: 1 format / data source, homogenization on arrival

Global Observing System

POLAR ORBITING SATELLITE

GEOSTATIONARY SATELLITE

SATELLITE IMAGES

AIRCRAFT

SATELLITE GROUND STATION

WEATHER RADAR

OCEAN DATA BUOY

SATELLITE SOUNDINGS

WEATHER SHIP

UPPER-AIR STATION

SURFACE STATION

AUTOMATIC STATION

NMS
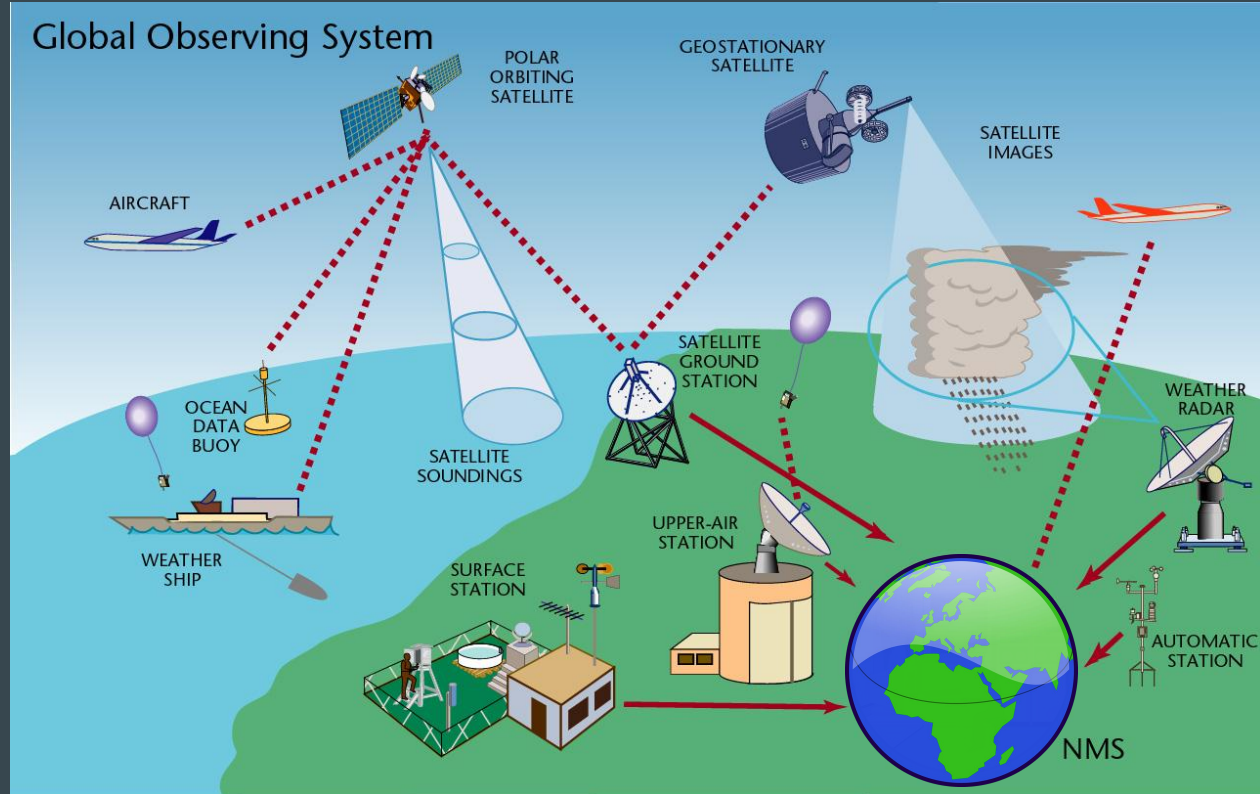
Standardized ingress

Now: 1 format / data source, homogenization on arrival

Future: Homogenization before transmission
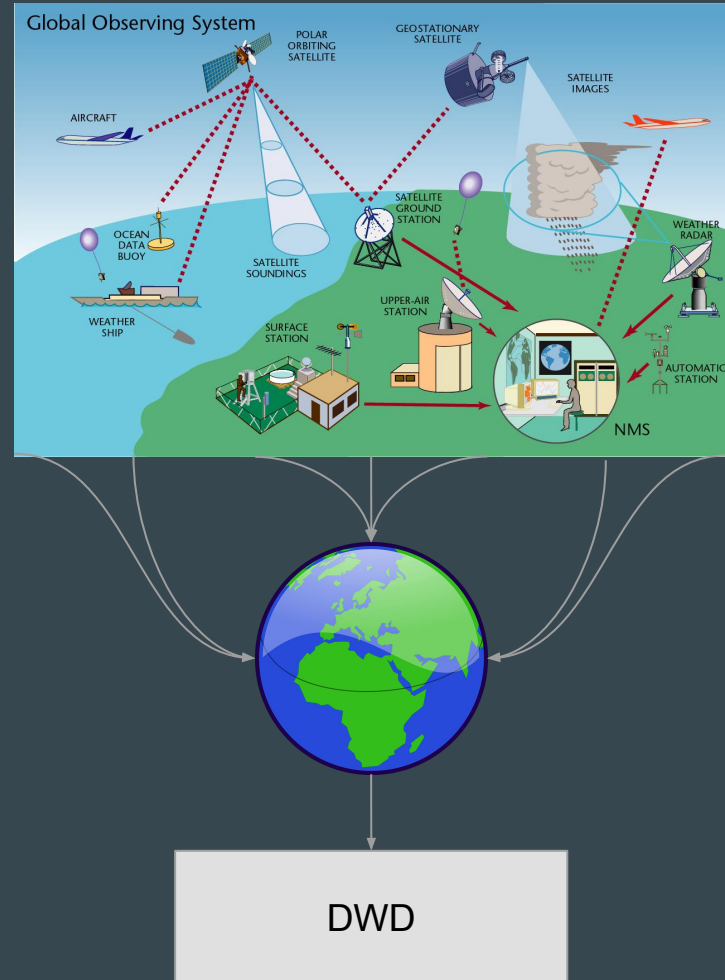
Standardized egress

Current: standardized outputs + some
specialized formats (fax, phone, etc.)

BUFR

DWD

Custom
formats

GRIB

Standardized egress

Current: standardized outputs + some specialized formats (fax, phone, etc.)

Trend: more users, more formats

BUFR

DWD

Custom formats

GRIB
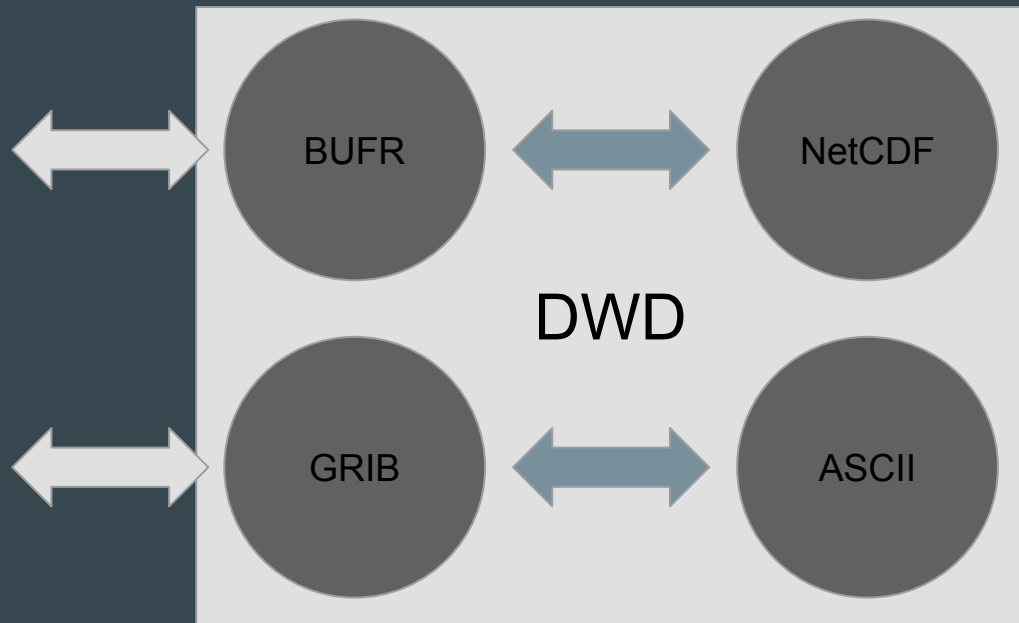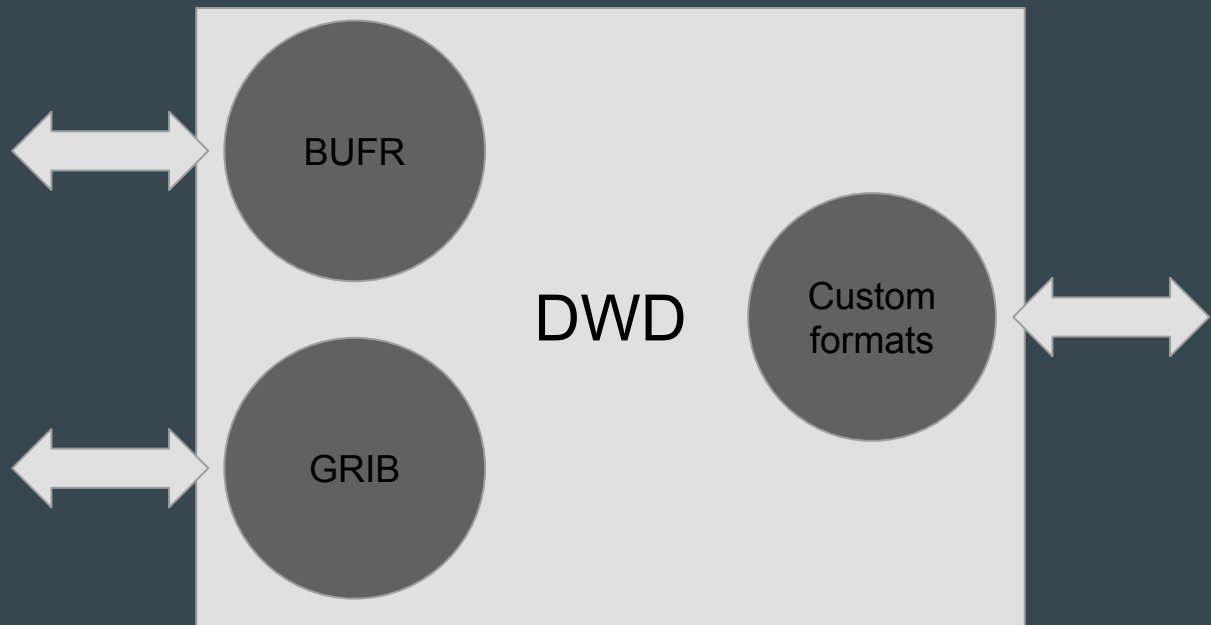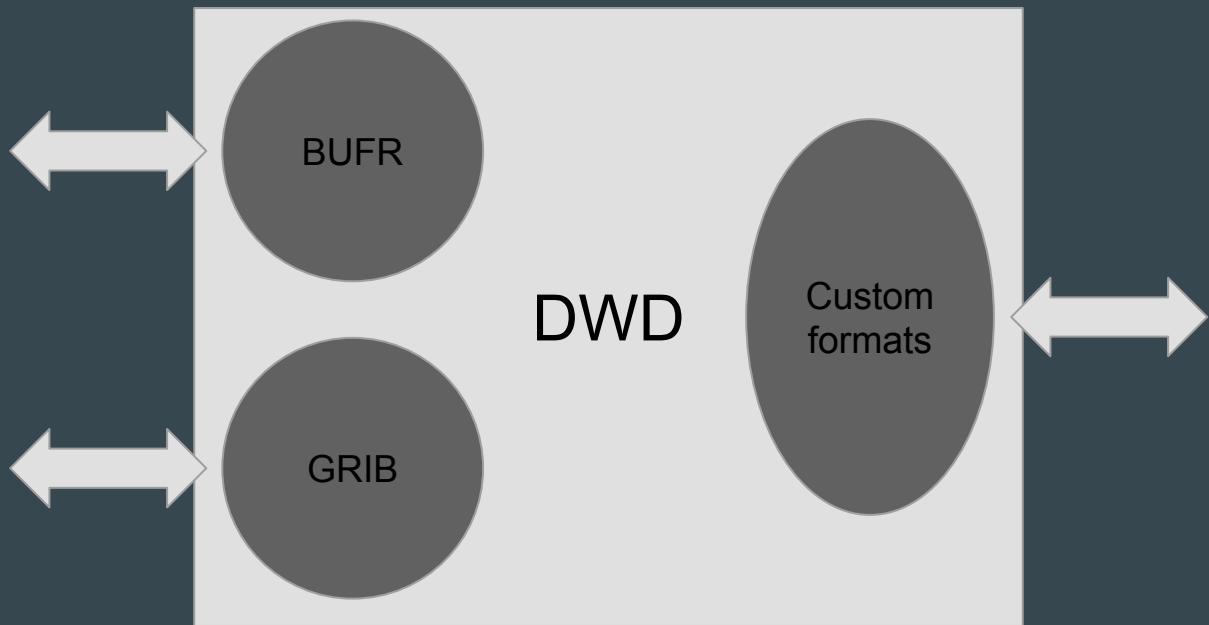
Standardized egress

Current: standardized outputs + some specialized formats (fax, phone, etc.)

Trend: more users, more formats

Future: Perhaps more standardized output formats (web services, etc.?)
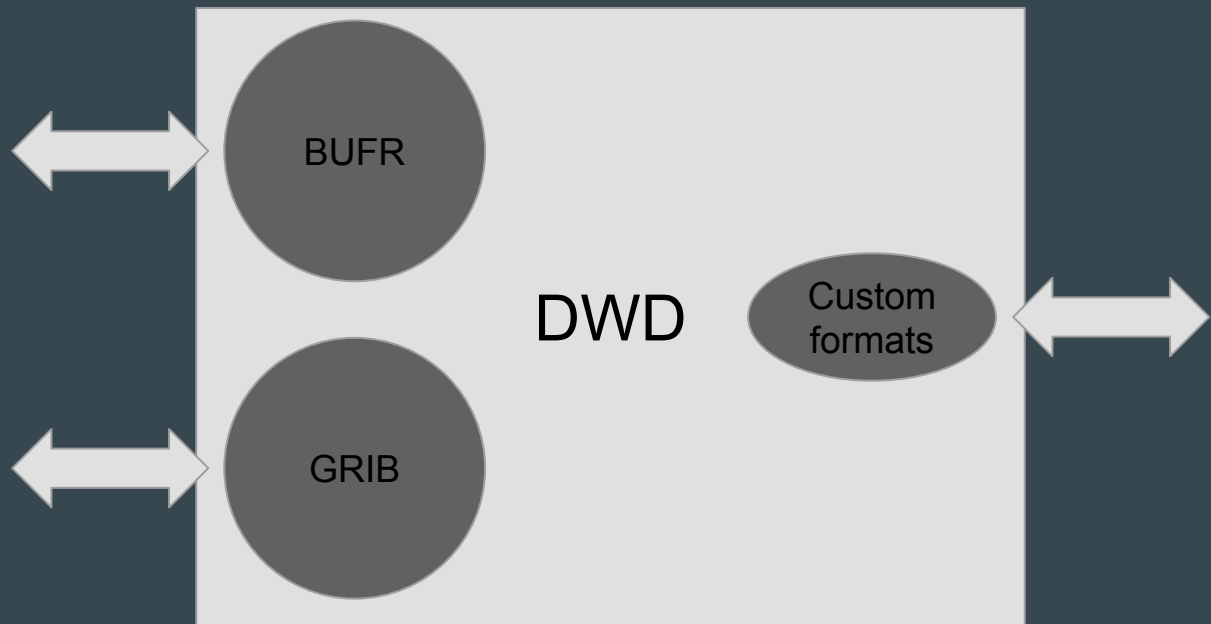
Value

All data are equal...
But some are more equal than others

Raw observations

Time series

Model outputs

...Integrity checks

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

DWD

Log am Bord der hamburger Brig Berg-Henrich auf der Reise von

Bemerkungen

Chungking

Januar 91 · Februar · März

	Januar 9ᵘ	9ᵘ	Max	Min	Regen	Februar 9ᵘ	9ᵘ	Max	Min	Rgn.	März 9ᵘ	9ᵘ	Max	Min	Rgn.
1	46	46	49	45		49	52	50	41		58	59	63	56	
2	47	52	57	45		48	45	47	44		59	56	62	56	
3	57	47	58	44		44	44	45	43		53	53	56	52	
4	43	49	60	43		44	45	47	43		53	52	60	52	
5	50	50	56	49		44	42	44	42		54	53	58	52	
6	50	46	56	46		42	43	45	40		52	54	60	52	
7	48	50	58	47		43	43	45	41		53	55	62	52	
8	50	50	55	50		44	43	45	41		51	52	60	51	
9	50	49	54	48	Regen Schauer	43	39	44	37		53	54	57	51	
10	47	47	54	47		43	41	43	36		53	51	54	50	
11	54	51	58	52		40	38	44	36		51	51	54	49	
12	50	50	54	50		36	42	47	34		54	52	65	49	
13	50	48	54	48		55	45	55	43		50	56	60	47	
14	47	49	58	47		52	49	55	38		56	54	57	51	
15	48	48	54	48		49	49	54	45		53	55	60	50	
16	50	50	58	50		47	47	50	45		57	55	65	51	
17	47	52	58	48		47	48	54	45		54	56	59	60	
18	51	51	54	51		47	47	52	46		52	55	60	48	
19	48	47	52	47		54	50	55	48		55	54	55	53	
20	46	45	52	45		50	51	56	47		52	50	52	50	
21	45	45	54	45		53	53	59	49		50	49	50	48	
22	46	47	59	46		54	58	64	51		48	49	50	47	
23	52	52	64	51		60	59	68	48		49	48	49	48	
24	51	49	56	49		57	58	61	50		49	51	52	47	
25	45	48	54	45		59	60	66	56		52	54	58	52	Regen Schauer
26	49	49	52	48		63	59	74	56		54	58	65	51	
27	48	48	58	48		63	62	70	58		56	60	70	52	
28	46	50	56	45		67	62	69	52		63	66	70	56	
29	50	49	51	45							60	60	66	59	
30	46	44	51	43							58	64	76	54	
31	42	44	46	40							60	67	77	54	
Mittel	48.4	48.5	55.2	46.9		49.9	49.1	53.9	44.8		54.0	54.9	59.7	51.3	

Outlook

More is more

Increasing demands

- Data sources
- Forecast horizons
- Data volume
- Tailored user solutions

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

DWD

Risks and opportunities

Increasing data volume

User demand for direct access

More data sources

Faster updates

Better HSM

Separation of concerns: ETL, REST

More standardization

Increased scale-out capabilities

Daniel Lee

Deutscher Wetterdienst (DWD)
Data ingress and decoding

Daniel.Lee@dwd.de

Tel: +49 (69) 8062-2706
Fax: +49 (69) 8062-3829
http://www.dwd.de