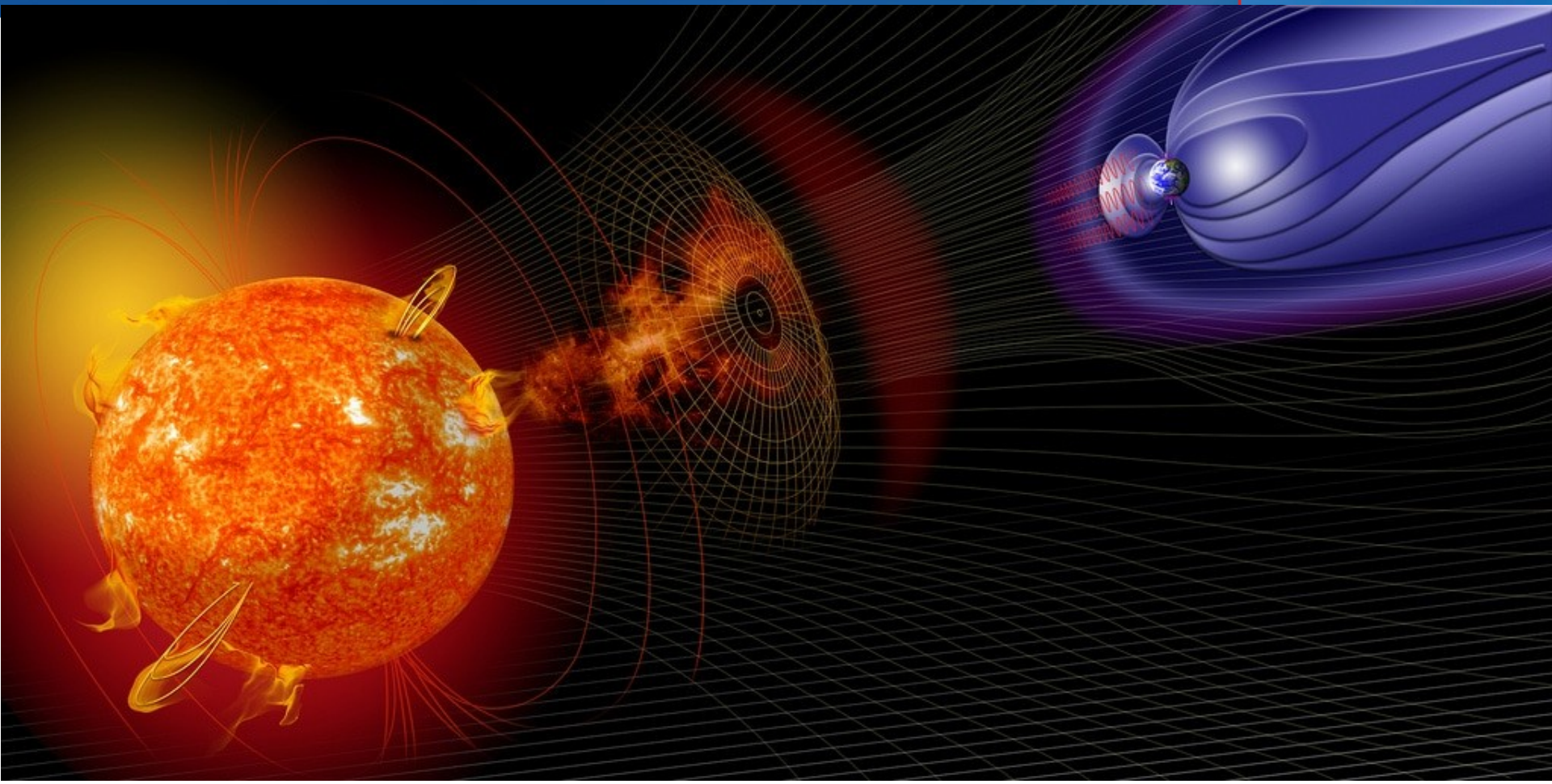
# IMBOT



„An automatic data checker for 1-second submissions to INTERMAGNET“
Roman Leonhardt, Conrad Observatorium, ZAMG

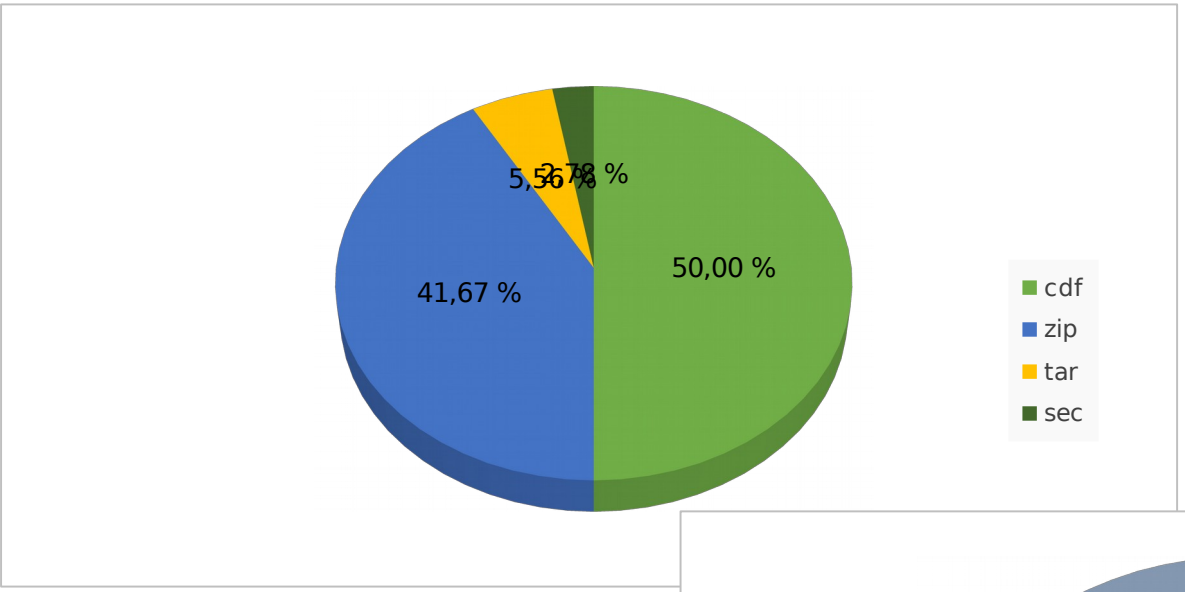
ZAMG
Zentralanstalt für Meteorologie und Geodynamik

# Background

- Since 2014 INTERMAGNET welcomes data submission in one second resolution
- For archiving and providing such data, the IMAGCDF format was introduced, based on NASA CDF
- Currently 36 observatories are submitting definitive one second data
- So far, one second data products are not checked and/or provided to end users by INTERMAGNET

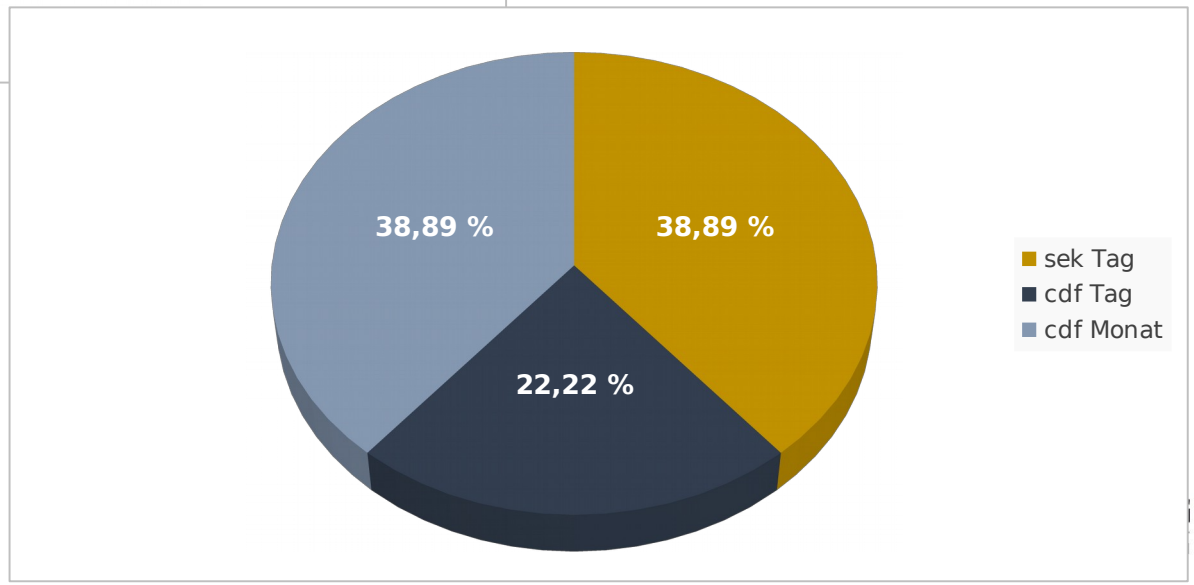
ZAMG
Zentralanstalt für Meteorologie und Geodynamik

# One second submissions for 2016




**36 data sets**

- different packing routines
- several formats (IAGA02, CDF1.0, 1.1)
- variable coverage



# One second submissions for 2016

**Typical issues encountered for 2016 submissions:**

- Wrong coverage:
  - files do not cover recommended time intervals
  - files contain data from last month/day, one second missing
- Wrong file names: at least according to naming convention
- Meta information incomplete
- Corrupted files (2)
- Wrong data content (1)
- Minor deviations from one-minute definitive data
- Noise level exceeds 100pT/sqrt(Hz) for about 10% of the data sets

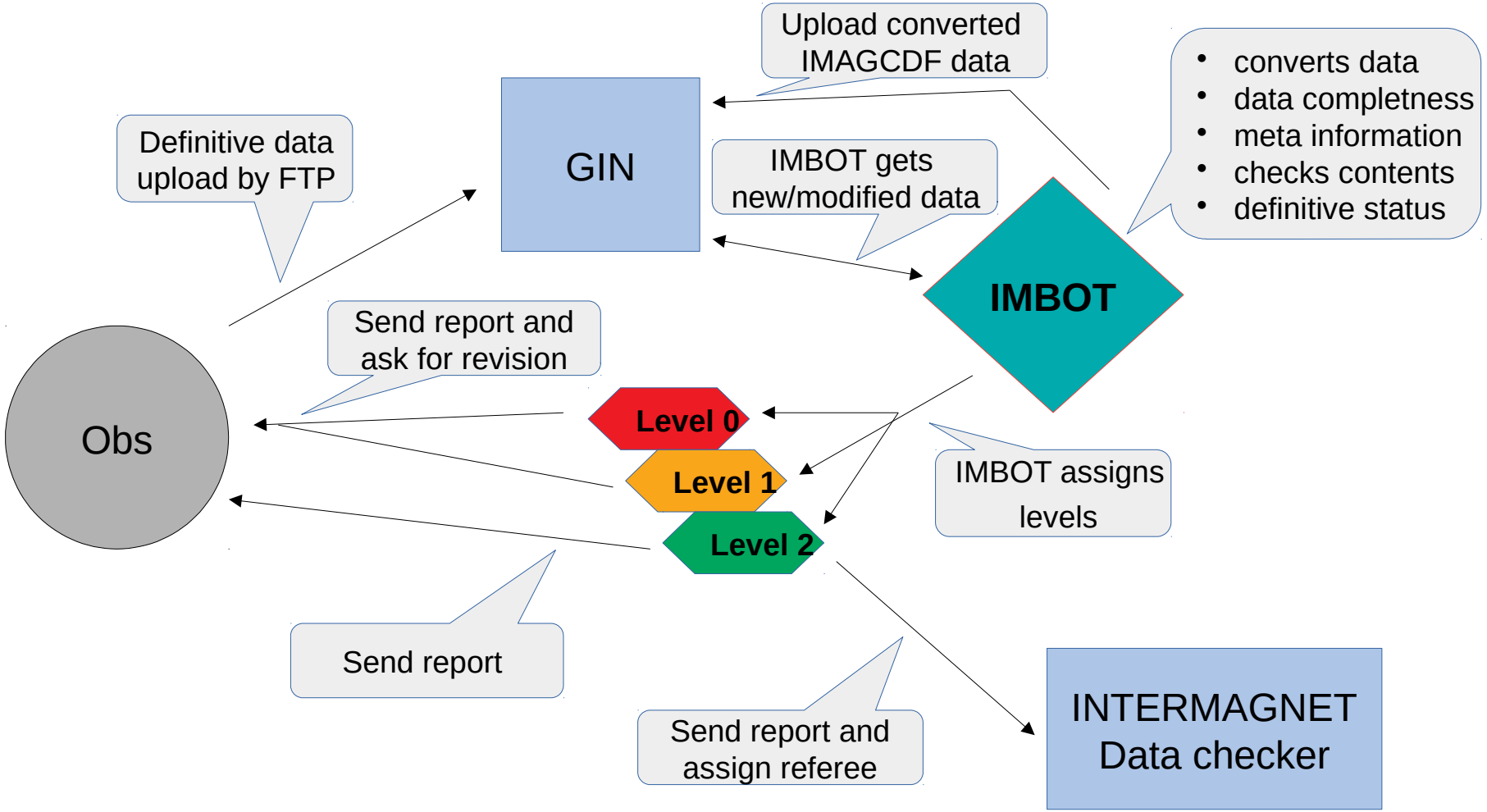
ZAMG
Zentralanstalt für Meteorologie und Geodynamik

# IMBOT – the principle

IMBOT stands for **I**NTER**M**AGNET ro**bot**

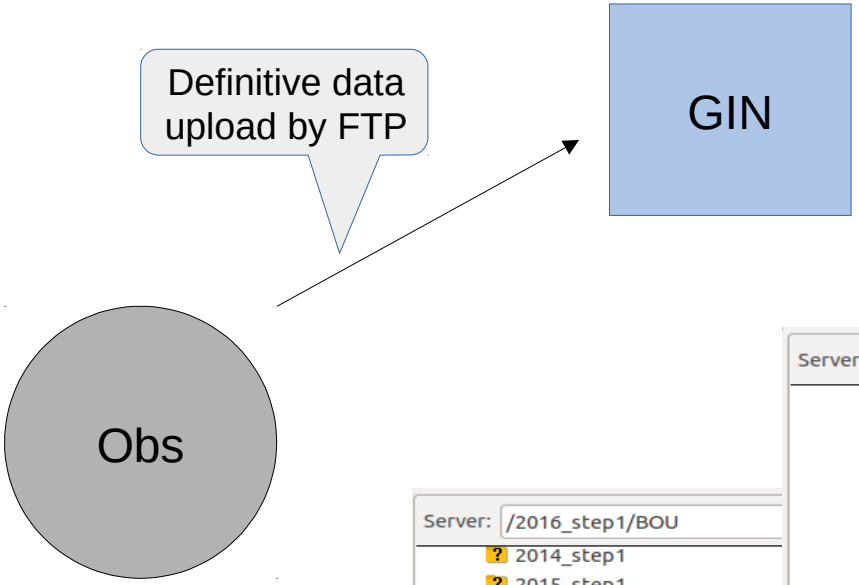
- IMBOT will analyse and re-format one-second data submissions into IM requested structures
- IMBOT works completely automatic
- IMBOT should minimize the workload for data providers, checkers, and users
- IMBOT needs to be based on a transparent evaluation process
- IMBOT should significantly reduce the time between submission and publication

# IMBOT – how does it work

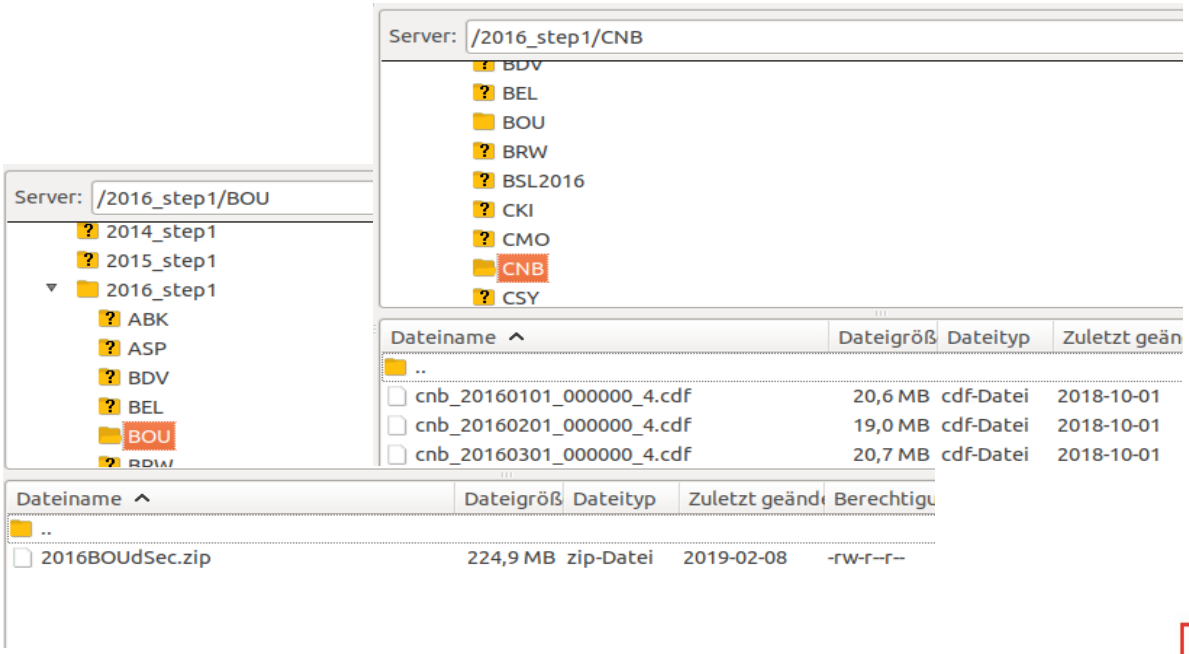



ZAMG
Zentralanstalt für Meteorologie und Geodynamik

# IMBOT – how does it work

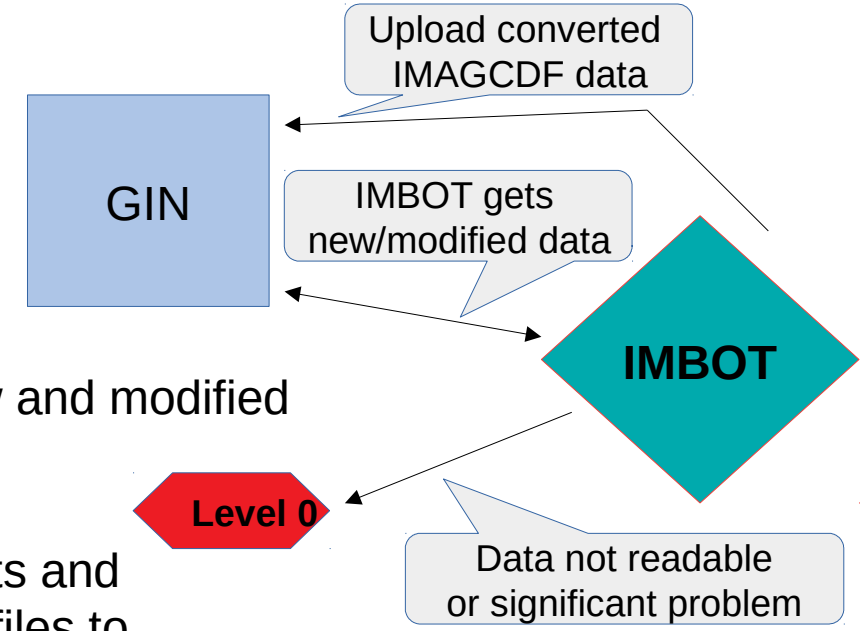



- Many different packing methods and data formats are used for uploading data




ZAMG
Zentralanstalt für Meteorologie und Geodynamik

# IMBOT – how does it work

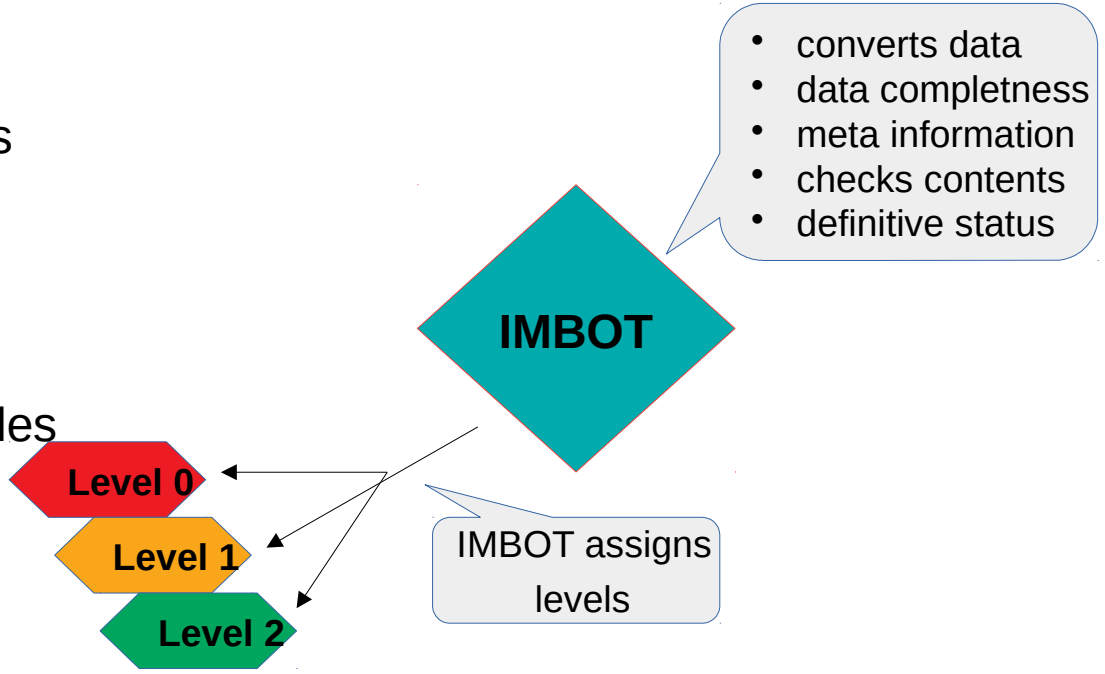



- IMBOT identifies all new and modified data sets from the GIN
- IMBOT extracts data sets and eventually converts the files to monthly IMAGCDF files as requested by IM
- Unreadable data or corrupted files lead to level 0. The submitter is asked to provide appropriate data
- Converted data is uploaded to the GIN

ZAMG
Zentralanstalt für Meteorologie und Geodynamik

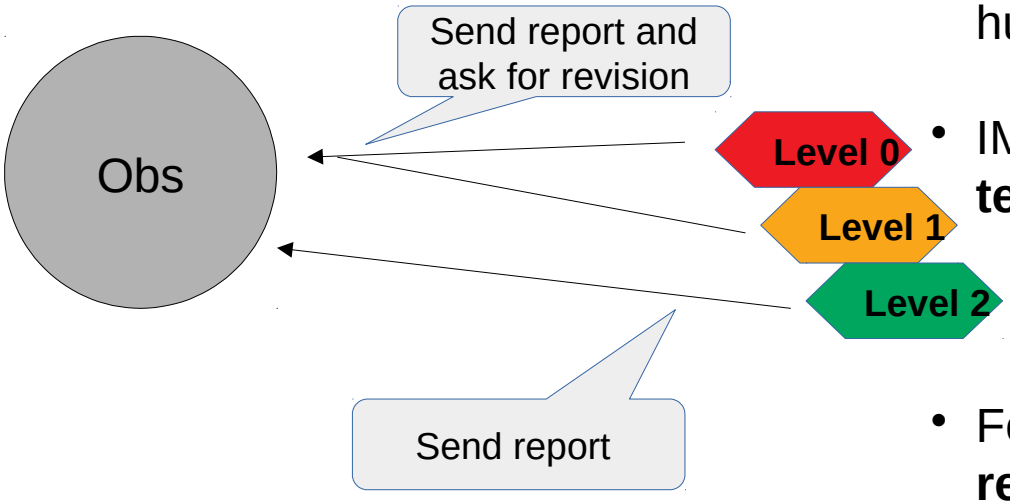
# IMBOT – data checking

- **Submitted files and formats**
  accepted are all readable formats (e.g. IAGA2002, IMAGCDF)
- **Meta information**
  meta information contained and consistent between all different files
- **Data content**
  Correct coverage and content
- **Data quality**
  not used as a criteria, but information provided in the report
- **Data consistency**
  data consistent with minute submission, definitive status




ZAMG
Zentralanstalt für Meteorologie und Geodynamik

# IMBOT – how does it work




- IMBOT creates an analysis **report**
- The **report** is uploaded along with the converted data and send to data submitters and, in case of level 2, to a human data checker
- IMBOT creates a meta data **review template** (for level 0 and level 1)
- For missing meta information only the **review template** must be filled out and uploaded
- **Report** and **review** template are also send out by e-mail

ZAMG
Zentralanstalt für Meteorologie und Geodynamik

# IMBOT – the e-mail

Dear data submitter,

you receive the following information as your e-mail address is connected to submissions of geomagnetic data products from unkown XXX observatory.
Your one-second data submission from 2018 has been automatically evaluated by IMBOT, an automatic data checker of INTERMAGNET.

The evaluation process resulted in

LEVEL 1

Your data has provisionally been accepted by INTERMAGNET. Congratulations!

In order to continue the evaluation process some issues need to be clarified. Please read the attached report and instructions.

If you have any questions regarding the evalutation process please check out the general instructions (github link) or contact the IMBOT manager.

Sincerely,
IMBOT

---

Instructions to update file and or meta information for re-evaluation of your data:
....

ZAMG
Zentralanstalt für Meteorologie und Geodynamik

# IMBOT – the report

https://github.com/geomagpy/IMBOT/blob/master/examples/level1_underreview.md

## OBSCODE - Level 1

## Analysis report for one second data from OBSCODE

### Issues to be clarified for level 2:

| Issue | Observed in months |
|---|---|
| header StandardLevel missing | 1,2,3,4,5,6,7,8,9,10,11,12 |
| StandardLevel full or partial - see TN8: 4.7 Relevant data standards | 1,2,3,4,5,6,7,8,9,10,11,12 |
| PartialStandDesc required for partial - see TN8: 4.7 Relevant data standards | 1,2,3,4,5,6,7,8,9,10,11,12 |

### Possible improvements (not obligatory):

| Improvements | Applicable for months |
|---|---|
| provide information on Terms | 1,2,3,4,5,6,7,8,9,10,11,12 |

### ImagCDF standard levels as provided by the submitter

| StandardLevel | Description | Validity |
|---|---|---|
| IMOS-01 | Time-stamp accuracy (centred on the UTC second): 0.01s | not provided |
| IMOS-02 | Phase response: Maximum group delay: ±0.01s | not provided |
| IMOS-03 | Maximum filter width: 25 seconds | not provided |
| IMOS-04 | Instrument amplitude range: ≥±4000nT High Lat., ≥±3000nT Mid/Equatorial Lat. | not provided |
| IMOS-05 | Data resolution: 1pT | not provided |
| IMOS-06 | Pass band: DC to 0.2Hz | not provided |
| IMOS-11 | Noise level: ≤100pT RMS | not provided - IMBOT indicates success |
| IMOS-12 | Maximum offset error (cumulative error between absolute observations): ±2. 5 nT | not provided |

ZAMG
Zentralanstalt für Meteorologie und Geodynamik

# IMBOT – the report

## Basic analysis information

- amount : 1
- type : .zip
- lastmodified : 1594113906.112519
- obscode : OBSCODE
- Readability test file : /media/leon/Images/DataCheck/tmp/OBSCODE/raw/obs20160210dsec.sec
- Readability : OK
- Data format : IAGA-2002
- Year : 2016
- MagPyVersion : 0.9.7
- Noiselevel : 10 pT
- NoiselevelStdDeviation : 1 pT

## Details on monthly evaluation

| Month 1 | Value |
|---|---|
| mean difference - x component | 0.00629 nT |
| mean difference - y component | 0.00528 nT |
| mean difference - z component | 0.0056 nT |
| stddev of difference - x component | 0.041 nT |
| stddev of difference - y component | 0.0409 nT |
| stddev of difference - z component | 0.0404 nT |
| amplitude of difference - x component | 0.229 nT |
| amplitude of difference - y component | 0.898 nT |

ZAMG
Zentralanstalt für Meteorologie und Geodynamik

# IMBOT – the report

| Month 1 | Value |
|---|---|
| mean difference - x component | 0.00629 nT |
| mean difference - y component | 0.00528 nT |
| mean difference - z component | 0.0056 nT |
| stddev of difference - x component | 0.041 nT |
| stddev of difference - y component | 0.0409 nT |
| stddev of difference - z component | 0.0404 nT |
| amplitude of difference - x component | 0.229 nT |
| amplitude of difference - y component | 0.898 nT |
| amplitude of difference - z component | 0.202 nT |
| Datalimits | [datetime.datetime(2016, 1, 1, 0, 0), datetime.datetime(2016, 1, 31, 23, 59, 59)] |
| N | 2678400 |
| Leap second update | None |
| Filled gaps | 0 |
| Difference to expected amount | 0.0 |
| Level | 1 |
| Samplingrate | 1.0 sec |
| delta F | mean delta F of 0.084 with a std of 0.199 |
| F | found independend f with sampling period: 1.0 sec |
| Definitive comparison | differences in peak amplitudes between definitive one-minute and one-second data products observed |
| Contact | ['observer@observatory.obs'] |

ZAMG
Zentralanstalt für Meteorologie und Geodynamik

# IMBOT – the review template

```
## Parameter sheet for additional/missing metainformation
## -----------------------------------------------------
## Please provide key - value pairs as shown below.
## The key need to correspond to the IMAGCDF key. Please
## check out the IMAGCDF format description at INTERMAGNET
## for details. Alternatively you can use MagPy header keys.
## Values must not contain special characters or colons.
## Enter "None" to indicate that a value is not available
## Comments need to start in new lines and every comment line.
## must start with a hash.
## Please note - you can also provide optional keys here.
##
## Example:
## Providing Partial standard value descriptions as requested:
# StandardLevel  :  partial
# PartialStandDesc  :  IMOS11,IMOS14,IMOS41


# Provide a valid standard level (full, partial), None is not accepted
StandardLevel  :  partial

# If Standard Level is partial, provide a list of standards met
PartialStandDesc  :  IMOS-01,IMOS-02,IMOS-03,IMOS-04,IMOS-05,IMOS-11,IMOS-14,IMOS-41

# If data is not available please confirm by MissingData  :  confirmed
MissingData  :  confirmed
```

ZAMG
Zentralanstalt für Meteorologie und Geodynamik

# IMBOT – the e-mail

Dear data submitter,

you receive the following information as your e-mail address is connected to submissions of geomagnetic data products from Mawson MAW observatory.
Your one-second data submission from 2018 has been automatically evaluated by IMBOT, an automatic data checker of INTERMAGNET.

The evaluation process resulted in

LEVEL 2

Your data has provisionally been accepted by INTERMAGNET. Congratulations!

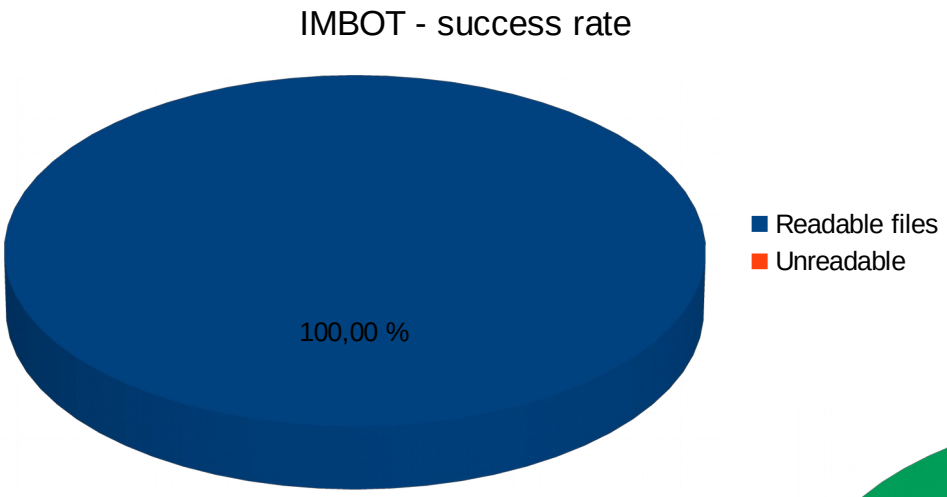
Your data fulfills all requirements for a final review. A level 2 data product is already an excellent source for high resolution magnetic information. Your data set has been assigned to an INTERMAGNET data checker for final evaluation regarding data quality.
Your data checker is Max Mustermann.
Please note that INTERMAGNET data checkers perform all check on voluntary basis beside their usual duties. So please be patient. The data checker will contact you if questions arise.

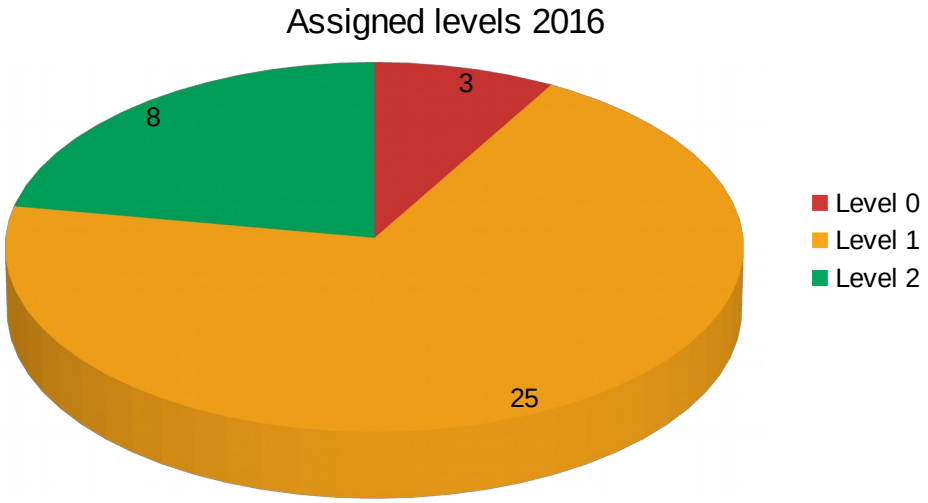
If you have any questions regarding the evalutation process please check out the general instructions (github link) or contact the IMBOT manager.

Sincerely,
IMBOT

ZAMG
Zentralanstalt für Meteorologie und Geodynamik

# IMBOT – 2016 analysis

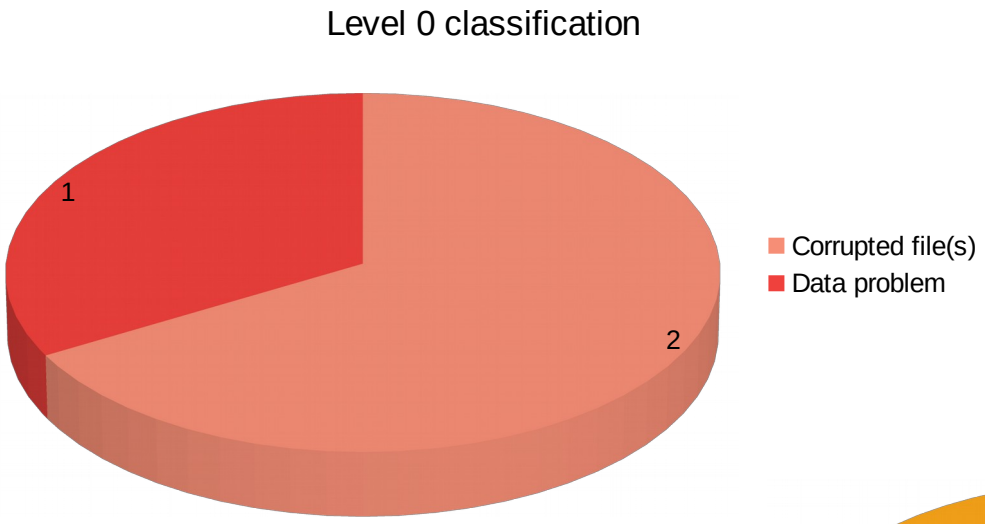


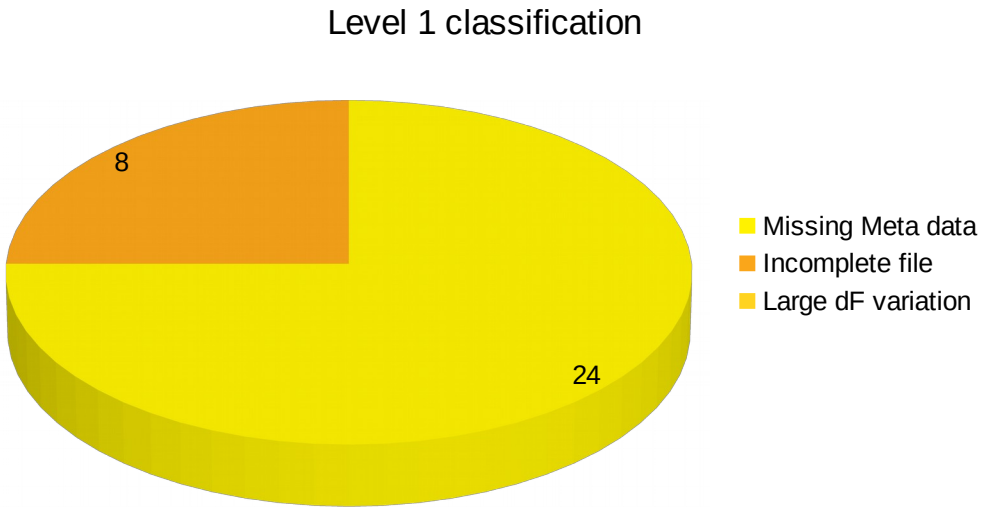






ZAMG
Zentralanstalt für Meteorologie und Geodynamik

# IMBOT – 2016 analysis






ZAMG
Zentralanstalt für Meteorologie und Geodynamik

# IMBOT – analysis summary

- IMBOT can test and convert **all** submitted data sets within hours
- Data checks indicate „valid“ data for the majority of all submissions
- By simply submitting a meta information update sheet, 75% of all submitted data sets will achieve level 2 classification
- IMBOT provides additional information e.g. on noise levels, individual deviations from submitted one-minute products, which is not used for automatic level classification, but might help the observatory to locate issues and improve data quality
- Data content of converted files is identical to original raw data (checked)
- *It is suggested that such reports are available also for end users*

# IMBOT – application

- A complete analysis needs approximately 30min to 1.5h for each observatory (dependend on IMBOT servers hardware).
- IMBOT is currently running, automatically analyzing all new submissions from 2018 and 2019 (uploading converted data to GIN is disabled)
- IMBOT is modular and can be easily extended and applied to other data submissions (e.g. variation data, one-minute submissions)
- Level assignement can be modified easily (requiring a decision of the IM data comitee). Further tresholds can be included
- *It is suggested that already level 1 data is provisionally accepted by IM as this data might already be useful for end users*
- *It is suggested that the current data level is fixed if no revision is submitted within three months*

ZAMG
Zentralanstalt für Meteorologie und Geodynamik

# IMBOT – application

- IMBOT fulfills all requirements as listed in the introduction
- It is possible to run IMBOT in test-mode with selected observatories and/or with a group of selected recievers of messages and reports
- If accepted by INTERMAGNET, IMBOT can be started any time provided…
  - that a list of human data checkers is available and their responsibilities are clarified
  - that GIN storage volume is capable of maintaining twice the amount of data for a short period (only in the beginning)