Additional automated checks on CDR

Automatic checks
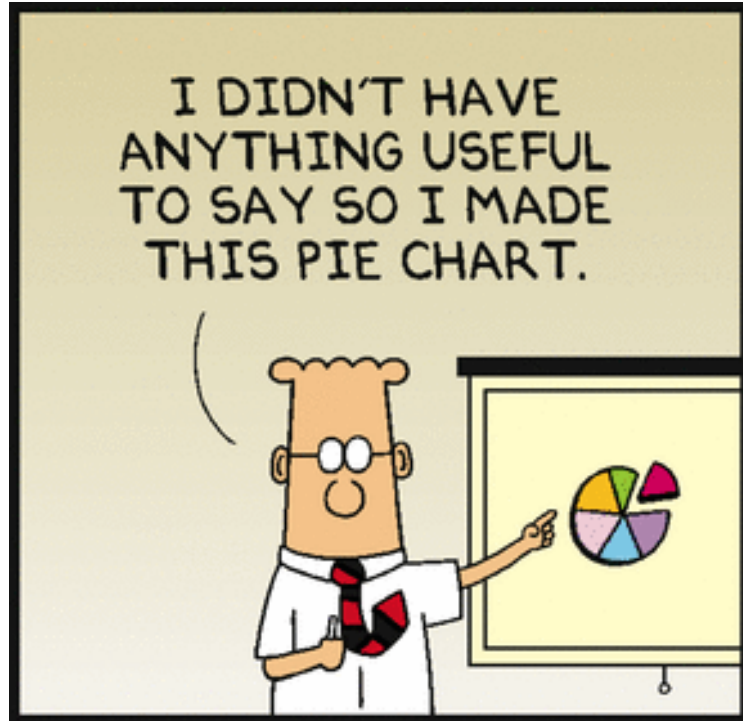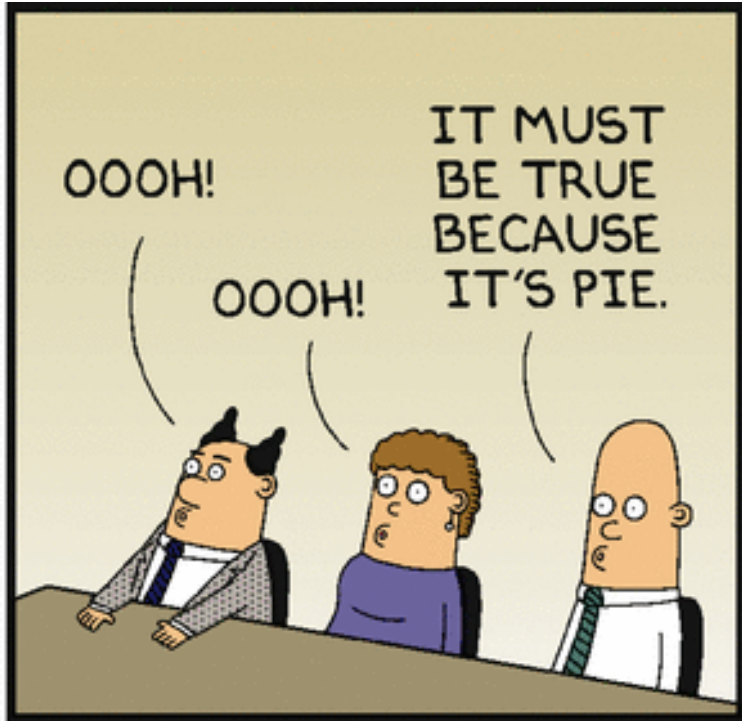
- Still crucial for

  – Country

  – EEA (both IT & content)

  – Commission

- Categories:

  – BLOCKER

  – ERROR

  – WARNING

  – INFO

  – OK

  – SKIPPED

- **BLOCKER**:

  – **A critical error.** The envelope can not be released. Normally, a blocker is an error in the format of the file, or in the structure or content of the data. Such a critical error makes it impossible for the delivery to be harvested and integrated into the database. The envelope can only be released if every incorrect file is removed and replaced by corrected files.

Check categories   *reminder!!!*

- **ERROR**:

  - **An important error.** The envelope can be released, but part of its content may be excluded from the European database

  - Data Reporters are strongly advised to correct errors.

  - A clarification or a resubmission **WILL** be requested, when the data is processed and the final feedback is added to the envelope.

- **WARNING** :

  – **An issue that may be an error.** Data Reporters are advised to check the correctness of the records or values that raised the warning. The envelope can be released.

  – Data Reporters are strongly advised to check warnings.

  – A clarification or a resubmission **MAY** be requested, when the data is processed and the final feedback is added to the envelope.

- **INFO**:

  - **Other issues related to the quality of the data or content of the submission.**

  - The envelope can be released.

- **OK**:

  - **The automatic QC did not detect quality issues.**

  - The envelope can be released.

- **SKIPPED**:

  – Data check has been executed, but there was "Nothing found to check", typically because of missing optional data

On-going improvements on QA
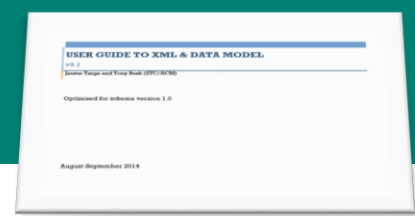
- Continue to ensure QA works 100%

- Improvements on small bugs on cross-checks

Improve on dataset D   *if possible!*

- Improvements on dataflow D to ensure full reporting (compared to previous deliveries)

- Cross-check between element (ef:supercedes) and samplingPoint id

Improve on E1a

- Automatic QA attempts to catch errors/bugs before processing data (on-going improvements)

Checks	Original
E27	Check that all values (between @@) include as many fields as declared under swe:DataRecord
E28	The data array should not end with "@@". Please note that @@ is a block separator.
E29	Check for unexpected spaces around all values (between comma separator) under swe:values
E30	Check for possible outliers reported. List of valid values found outside the range according to http://dd.eionet.europa.eu/vocabulary/aq/primaryObservationRange/
E31	Check for date overlaps between consecutive data blocks within swe:values
E32	Check that all data submitted via CDR has been fully verified. The verification flag must be 1 for all data.

on Air Pollution and Climate Change Mitigation

4sfera

European Environment Agency

Improve on E1a (pending improvements on QA/QC)

- E30 to take into account country specific ranges as previous checks in DEM

- Investigating if some statistical checks could be done on CDR
  - Negative/Zero annual averages, unexpected range...

- Possible new checks

    – When reporting G, ensure that full E1a and E1b have been reported

Any Questions?

EEA-ETC/ACM contact:
Jaume Targa (4sfera Innova)
Principal consultant

AQD IPR e-Reporting Helpdesk:
aqipr.helpdesk@eionet.europa.eu

m. +34 679 38 01 01
e. jaume.targa@4sfera.com
skype jaume.targa.4sfera
twitter @4sfera
web www.4sfera.com

You can follow us on twitter @4sfera

European Topic Centre
on Air Pollution and
Climate Change Mitigation

INNOVA
4sfera

European Environment Agency