

LOFAR: Supporting HDF5

ADASS XXV, Sydney, 2015-10-26

Dr. Jan David Mol

HDF5 in LOFAR



- Output since 2011
 - Initially processed using format converters
 - Post-processing pipelines slow to adopt
- Why HDF5?
 - Describes ALL meta data
 - Self contained
 - Reduces programmer effort & errors
 - Aids automated processing

Format Conversion?



- Existing tooling:
 - Support old formats
 - Hard to adjust
- Converters seem sufficient, but:
 - Loses meta data
 - III-described data → custom extensions, errors
 - Hard to process automatically
 - Has cost
 - Read+write all data
 - Must stream data

Old vs New: Options



- Upgrade tooling
 - Allows a cleaner tool chain → easier to extend
 - Allows cleaner formats → fewer errors
 - Can be hard (no resources, no internals knowledge)
- Extend old formats → "MS Word" approach
 - Within format
 - Wrap format (not backwards compatible)
 - Add extra file
- Do nothing
 - Everyone writes custom converters → more effort & errors
 - Tools can't use new meta data → less science

Risk of not Coordinating



HOW STANDARDS PROLIFERATE: (SEE: A/C CHARGERS, CHARACTER ENCODINGS, INSTANT MESSAGING, ETC.)

SITUATION: THERE ARE 14 COMPETING STANDARDS.



500N:

SITUATION: THERE ARE 15 COMPETING STANDARDS.

http://xkcd.com/927