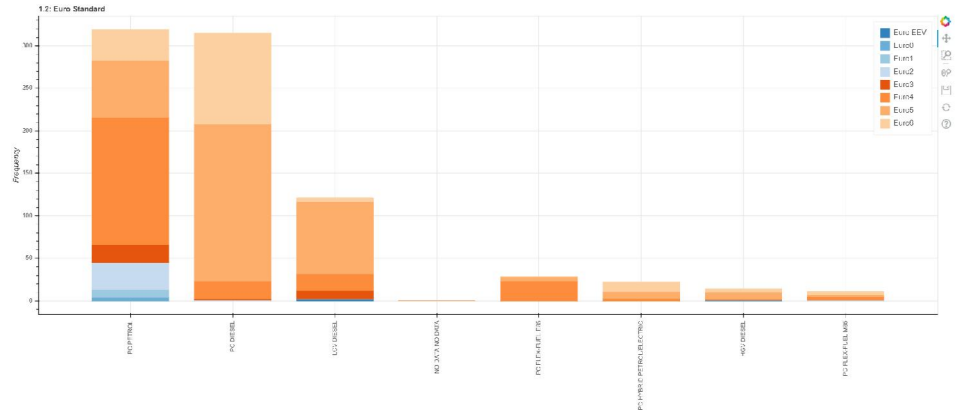
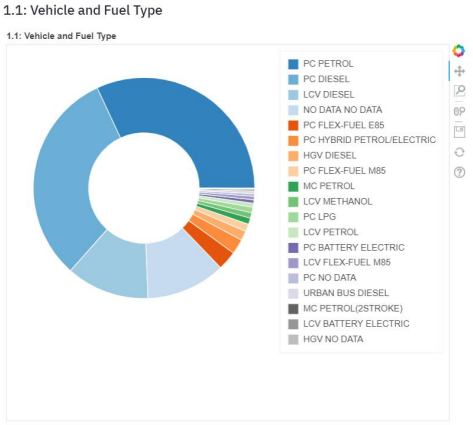
Progress Towards a Unified Storage Solution for Remote Emissions Sensing Data

James Tate and Chris Rushton

UNIVERSITY OF LEEDS
Institute for Transport Studies

CARES
CITY AIR REMOTE EMISSION SENSING





Current Practice

Conventional Vehicle Emission Remote Sensing

CARES
CITY AIR REMOTE EMISSION SENSING

Measurement Data (Flat):

As provided by commercial RS providers (fixed format?)

e.g. CSV file format.

Straight-forward data solution / sharing.

Relational Database:

Fixed schema forces conformity but lacks flexibility.

e.g. MySQL is a well understood enterprise level database solution.

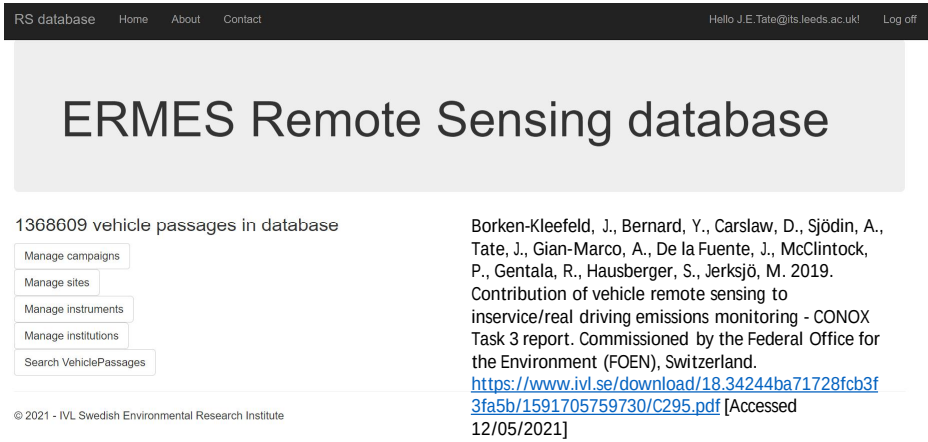
CONOX:

Comprehending

Combining

COmparing Real Driving Emissions

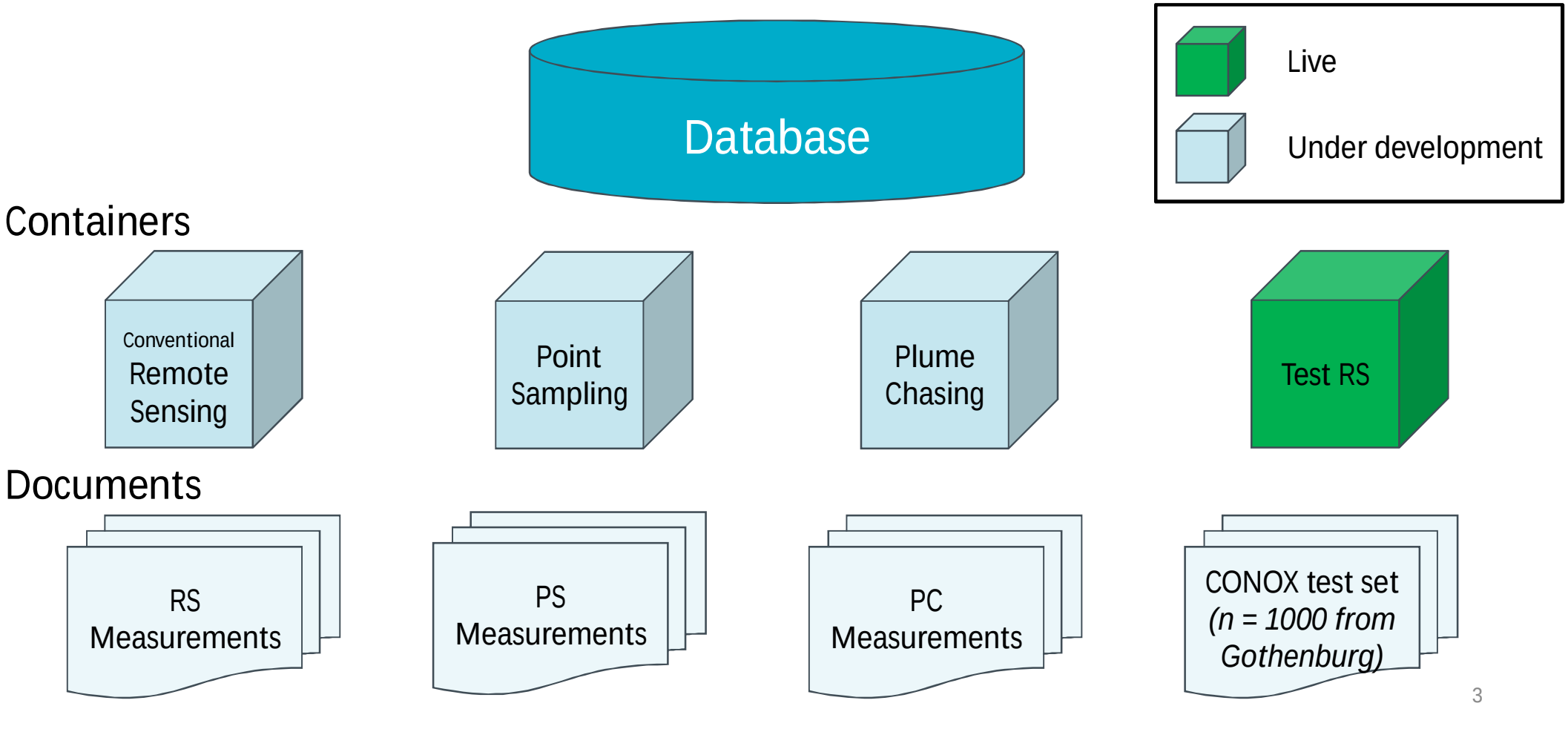
COllaborating on NOx RDE and Remote Sensing measurements.



Borken-Kleefeld, J., Bernard, Y., Carslaw, D., Sjödin, A., Tate, J., Gian-Marco, A., De la Fuente, J., McClintock, P., Gentala, R., Hausberger, S., Jerksjö, M. 2019. Contribution of vehicle remote sensing to inservice/real driving emissions monitoring - CONOX Task 3 report. Commissioned by the Federal Office for the Environment (FOEN), Switzerland. https://www.ivl.se/download/18.34244ba71728fcb3f3fa5b/1591705759730/C295.pdf [Accessed 12/05/2021]

CARES Project

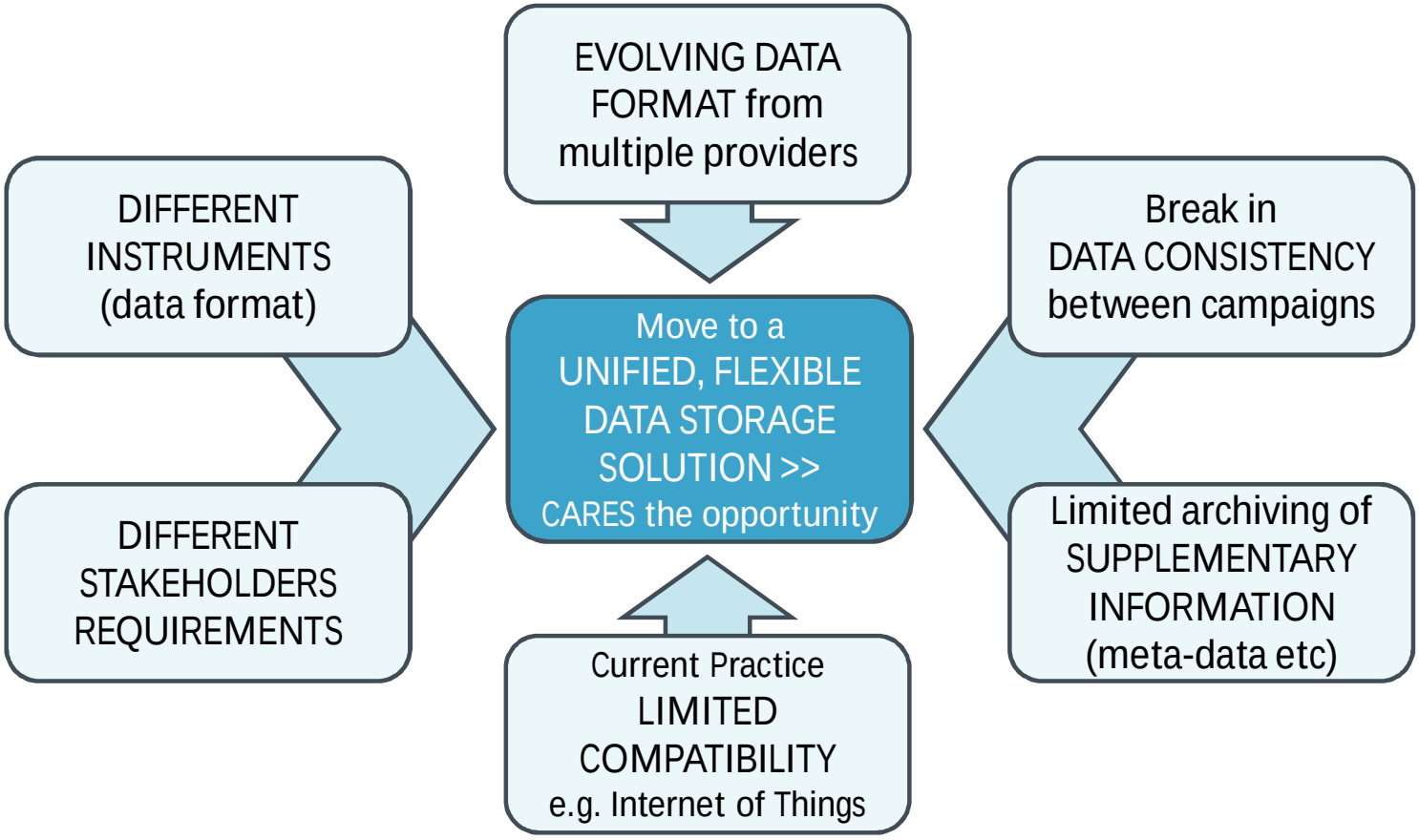
The data infrastructure challenge



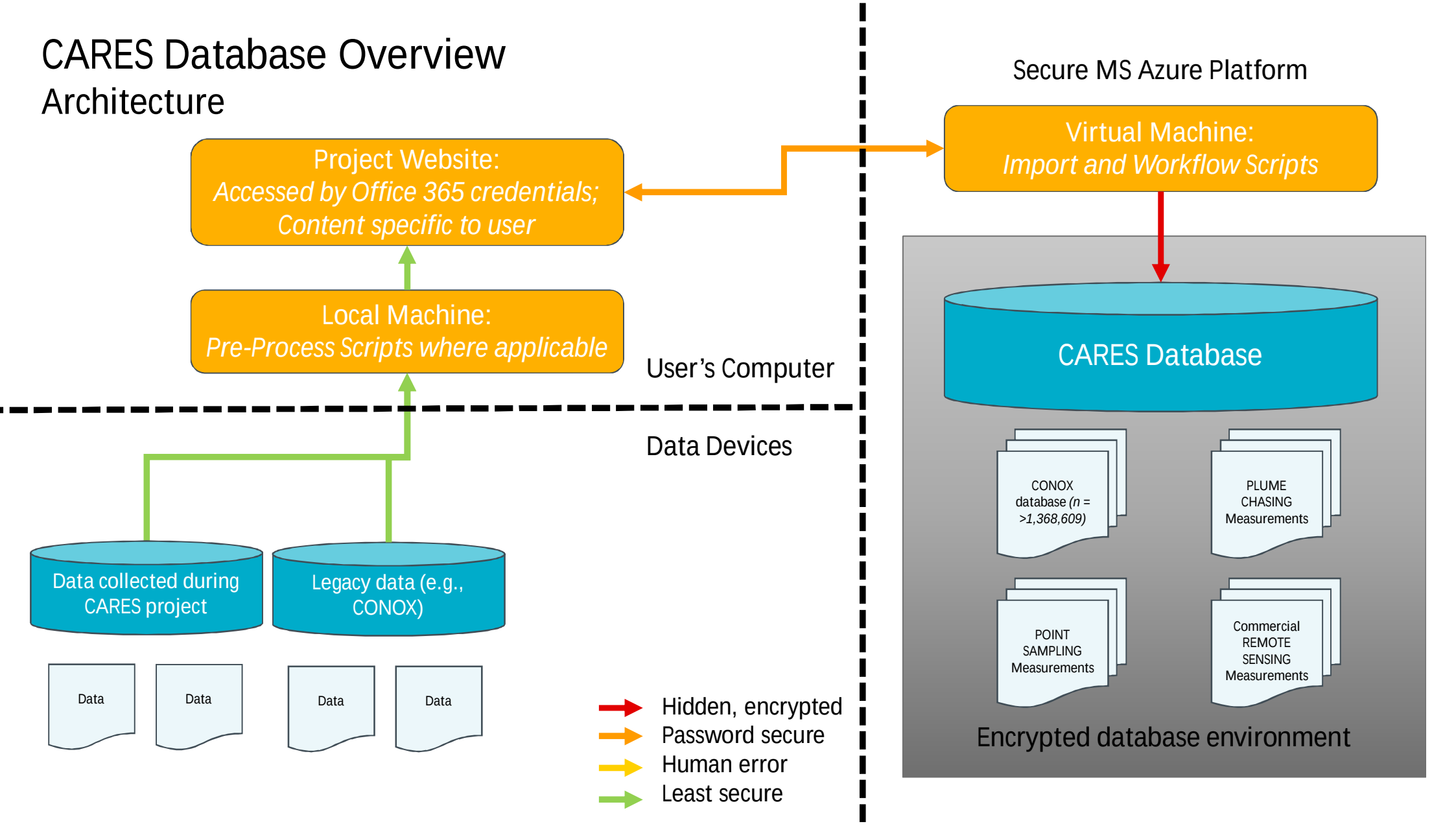
CARES Work Package 2

The data infrastructure challenge

The need for a unified, flexible solution



CARES Database Overview
Architecture



Data Journey

Traditional Remote Sensing

CARES
CITY AIR REMOTE EMISSION SENSING

Tasks performed before handover

Tasks performed between handover and retrieval

Post-retrieval analysis

Remote sensing deployment data collected

Initial data sense check

Data added onto database

Data quality control functions applied

Queried data set accessed

Data Collection → *Handover* → *Storage* → *Query* → *Analyse*

- Data collected using normal procedures
- Initial sense checks of data to be performed offline, including:
  - *Data exists in the file*
  - *Data is in the correct format*
- Some data checks available in GUI
  - *Prioritization of headers for checking*
  - *Headers updated to standard form*
  - *Ranges of data match CONOX or reference values already in database*
- Data set converted into database format
- Data is ready to be uploaded

- Data is received by the database
- Data is added to the database *in the form it was delivered*
- Data is stored on the database and remains encrypted
- Data retrieval request is received subject to permissions
- Quality control functions (*Deliverable 2.2 etc*) are applied to queried data
- Data is ready to be retrieved

- Data access controlled by Azure platform
- Security based on institution standards
- All actions auditable
- Anonymous *read-only* queries for some users
- Download for selected users
- User interface via *streamlit*
- Data is ready to be analyzed

Demonstration of the

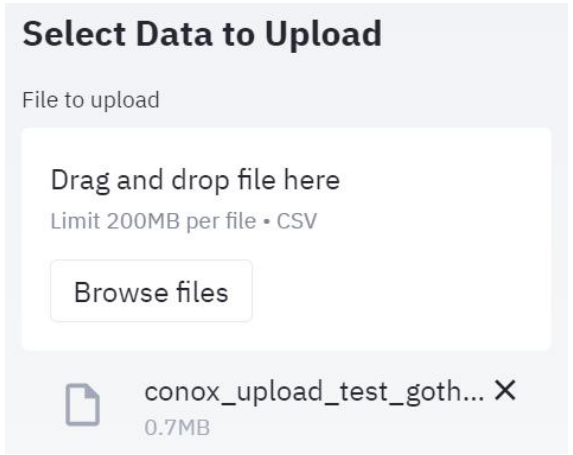
DEVELOPMENT CARES DATABASE INTERFACES

CARES Work Package 2

Data Interface

CARES
CITY AIR REMOTE EMISSION SENSING

- Data can be uploaded from a local file using familiar operating system file browsing systems
- Headings are updated to standard form using pre-configured templates
- Manual heading transformation is implemented but is very time consuming to use
- Basic statistics are applied to the data set comparing it to the data already in the database
  - Midpoint ratio checks that the data is broadly aligned
  - Kolomogorov - Smirnov (K-S) test for two samples applied
  - Thresholds are determined and applied



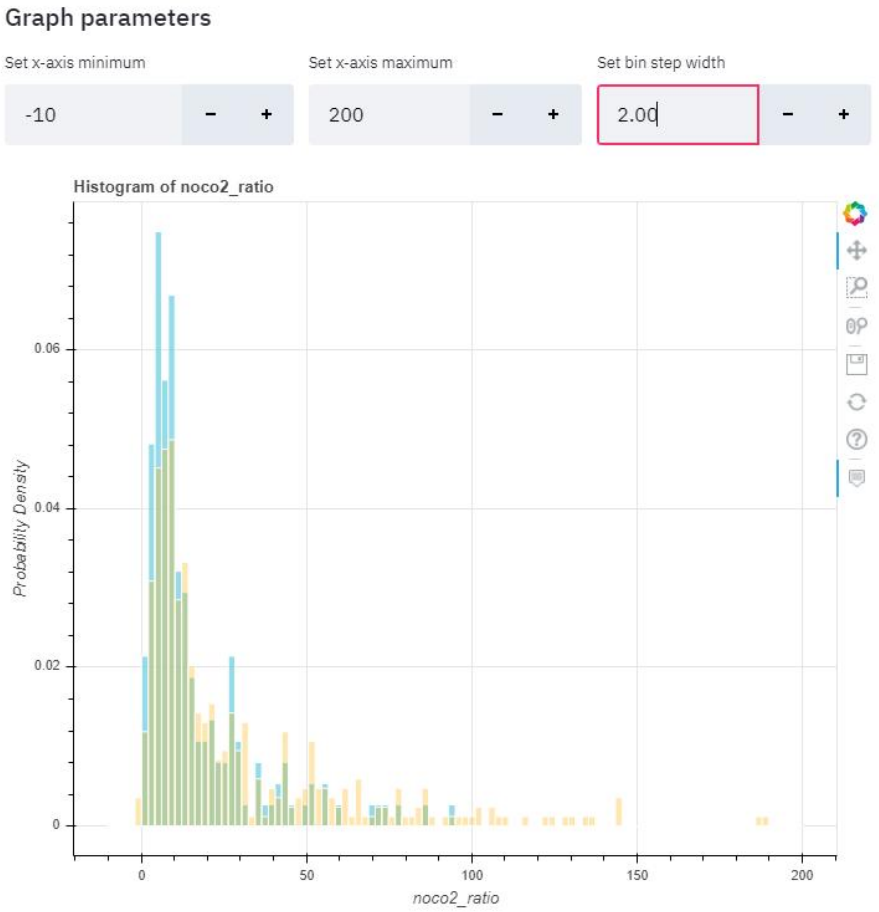
Variable	Midpoint Ratio	KS-Statistic	p-value	Overall RAG
Euro 3 Petrol PC				
NO:CO2	1	0.31	0.13	Green
HC:CO2	1	0.07	1	Green
CO:CO2	1	0.14	0.93	Green
Euro 5 Diesel PC				
NO:CO2	1	0.18	0	Amber
HC:CO2	1	0.05	0.85	Green
CO:CO2	1	0.07	0.52	Green
Fleet Dynamics				
Speed	0.97	0.08	0	Amber
Acceleration	-3.91	0.15	0	Red
VSP	1.58	0.07	0	Amber

CARES Work Package 2

Data Interface

CARES
CITY AIR REMOTE EMISSION SENSING

- The NO:CO2 ratio in the sample data set only met Amber conditions
- We can investigate this using a built in data viewer
- Test data (yellow) is different from reference data (blue) but an eyeball confirmation tells us that they are likely equivalent
- Data that is added to the database will still have amber status attached to it
- Future users can determine whether they accept amber or red rated data
- These tests can be performed for all pollutants
- Graphics can be easily saved directly from the interface



CARES Work Package 2

Data Interface

CARES
CITY AIR REMOTE EMISSION SENSING

- The acceleration parameter did not meet the amber or green criteria
- Data viewer shows us that there are more lower and negative acceleration values in the test data set
- User decision must be made as to whether this is caused by test location or by some error in the data
- If data is added to the database without modification it will be flagged as red.

Future outlook and closing remarks

CARES Work Package 2

Future Work – Remote Sensing

CARES
CITY AIR REMOTE EMISSION SENSING

- Adding functionality to current web platform:
  - *Improving scientific basis for upload decision*
  - *Improving the user interface for easier interaction*
  - *Include Office 365 authentication*
  - *Publish to online platform using Azure App Service*
- Increasing the content of the database
  - *Currently a small segment of CONOX*
  - *Plan to include all CONOX data*
  - *CARES characterisation and city demo data as available*
- Integrate data quality control checks
  - *Basic functions already developed as part of WP2*
  - *Investigate using natural language processing to improve quality of text based fields (e.g. standardised naming conventions)*
  - *Sense checking Euro standard*

Work Beyond CARES

Future Work – More New Methods

- Closer integration of database with instruments
  - *IoT connectivity*
  - *Increased stakeholder access*
  - *Larger instrument networks*
- Improved collaboration space
  - *Azure platform offers huge potential to work collaboratively*
  - *New tools can be developed to take advantage of new data handling approach*
  - *Secure environment to test new ideas*
  - *Bespoke dashboards for different stakeholders*