Summary of proposed approach for mapping of historical E-PRTR and LCP data into the new 'EU Registry' and 'E-PRTR & LCP Integrated Reporting' databases

Background

The new 'EU Registry'[1] and 'E-PRTR and LCP Integrated Reporting'[2] data flows provide a streamlined reporting mechanism which covers a number of specific legislative reporting requirements, namely:

- Reporting of European Pollutant Release and Transfer (E-PRTR) data;
- Reporting of data on Large Combustion Plants (LCPs); and
- Reporting of other data required under the Industrial Emissions Directive (IED);

Each of the two data flows has a new dedicated database for the storage of collected data.

Historical E-PRTR and LCP data are also stored separately in existing databases. It is proposed to import, where possible, the historical data from these databases into the new database structures in order to provide a continuous time series of data collected in these datasets. This document sets out the proposed approach for mapping these data into the new databases and, importantly, sets out the input and support required from reporting countries in order to be able to successfully apply the methodology.

1: Mapping of Historical E-PRTR Data

1.1 Mapping E-PRTR facilities which are <u>also</u> reported to the EU Registry in 2017

The 2017 EU Registry reporting (reporting deadline of June 30th 2019) requires reporting of all currently functioning E-PRTR facilities (even those with releases and transfers below the reporting threshold). These data will include the specification of an InspireID for all of these facilities (which is the key identifier for geographical entities in the data flow) and also the definition of a parent site and associated details. Most of these facilities will have been reported in previous reporting years and will therefore have previously had a NationalID assigned under the E-PRTR reporting (the NationalID was the key identifier under the E-PRTR dataflow). This scenario is represented by Scenario B in Figure 1 below. The identification of the relationship between the InspireID and the NationalID is therefore necessary in order to create a continuous time-series of data.

Reporting countries will therefore be asked to provide the NationalID for each of the facility InspireIDs in the 2017 EU Registry report. This will be requested in a simple Microsoft Excel format and to further simplify the process a series of country envelopes will be created on CDR to collect this information from reporting countries (countries can then upload the file to the envelope). Collecting the data via CDR will also allow the data to be readily available in the future.

Countries will be asked to provide the <u>most recently</u> reported National ID for the facility from the 2017 E-PRTR data submission (as in some cases a facility may have had more than 1 National ID over its lifetime).

---

[1] For further information see: https://cdr.eionet.europa.eu/help/euregistry
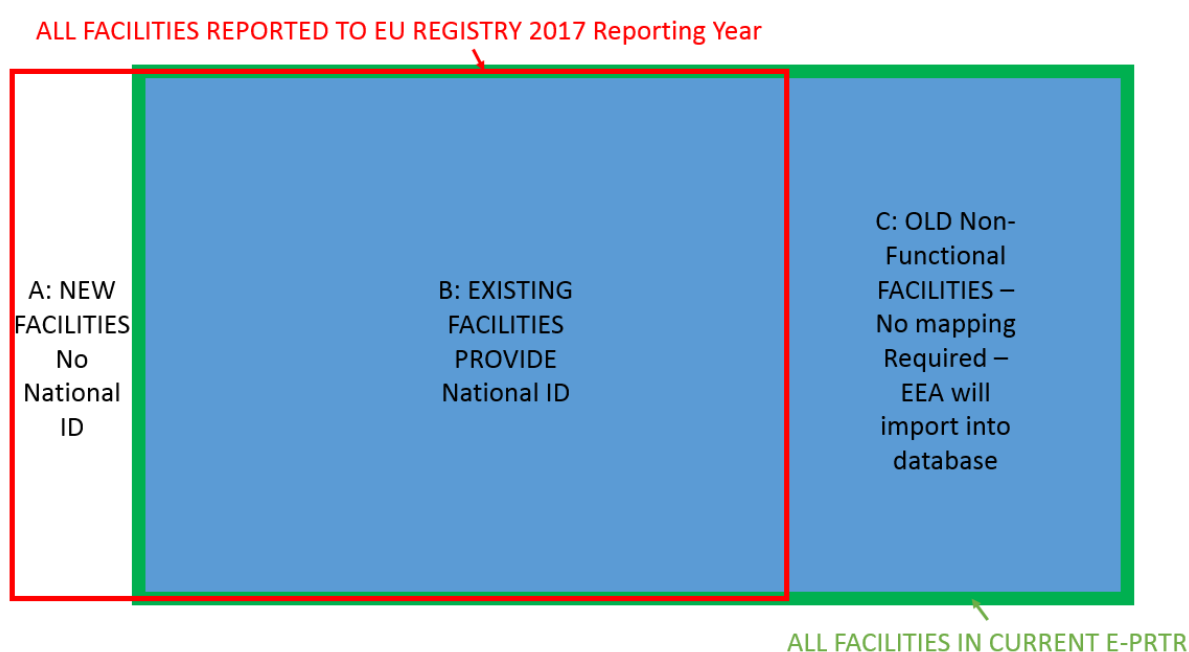[2] For further information see: https://cdrtest.eionet.europa.eu/help/eprtr_lcp

In a small number of cases the facility may not have been reported before in E-PRTR reporting (e.g. where emissions have always been below threshold), and therefore no National ID will be required. In these cases there will be no historical data to map into the new databases. This represents Scenario 'A' in Figure 1 below.

The existing E-PRTR database will remain the repository for the definitive E-PRTR dataset from 2007 to 2017, as it is proposed that only relevant fields will be imported into the EU Registry. It is not proposed to make any adjustments to the data, the data will be imported into the new database exactly as it was submitted by the reporting country. There are some new fields in the EU Registry and E-PRTR and LCP Integrated Reporting database which did not feature in the E-PRTR dataset (e.g. 'status' or 'facility type'). Where necessary these fields will be completed as part of the mapping (e.g. facility type will be set to 'EPRTR' for all historical E-PRTR facility data).

Figure 1: Schematic of proposed mapping scenarios

ALL FACILITIES REPORTED TO EU REGISTRY 2017 Reporting Year

A: NEW FACILITIES No National ID

B: EXISTING FACILITIES PROVIDE National ID

C: OLD Non-Functional FACILITIES – No mapping Required – EEA will import into database

ALL FACILITIES IN CURRENT E-PRTR

1.2 Mapping of old E-PRTR facilities not reported in the 2017 EU Registry submission

Facilities which are no longer operational (Scenario C in Figure 1) will not be reported in the EU Registry 2017 submission, but their historical data is available in the E-PRTR database. This data needs to be harvested into the new databases in order to generate a complete dataset from 2007 to date (to be used in the new website which will be launched in 2020).

To simplify the reporting requirements, reporting countries will not be asked to assign Inspire IDs to these old sites. Instead, the EEA will generate dummy InspireIDs and import the data for these facilities into the new database. This will also require the creation a dummy Production Site, which will have the same name and location and also a dummy Inspire ID for the site level. Otherwise the import process will be the same as outlined in Section 1.1 above.

2. Mapping of Historical LCP Data

For LCP, it is proposed to only initially import LCP data for the reporting year 2016 as the data model for LCP reporting changed for the 2016 reporting year. Similar to the approach for E-PRTR facilities, reporters will be asked to provide a Plant ID (associated with LCP reporting) for each LCP installation part reported in the 2017 EU Registry submission. This will allow all 2017 LCP thematic data to be imported into the new database.

EEA will also use this mapping to harvest data from the 2016 LCP submission. Some LCPs (a small number) reported in 2016 may no longer be operational and thus were not reported to the EU Registry in 2017 and therefore will not have an Inspire ID. For this small number of plants we will consult with each reporting country to identify how best to address them. Countries may wish to assign an InspireID and define a relationship with a parent site and installation. Otherwise the EEA can create dummy data for the purposes of populating the database in cases where these plants are decommissioned.

3. Proposed Timeframe for Mapping Activities and QA Checks

Reporting countries will be asked in Q4 2019 to provide an EPRTR NationalID for all facility InspireIDs reported in the 2017 EU Registry dataset. A reporting template (Excel) will be provided.

The deadline of December 31 2019 corresponds with the final resubmission deadline for the existing E-PRTR dataflow. This resubmission will allow countries a final opportunity to update 2007 to 2017 E-PRTR data, and it is considered a practical option to align the deadline for the mapping and for the resubmission as the list of NationalIDs in the E-PRTR database will not change once countries have made their final resubmissions.

The EPRTR database will effectively be closed from 31 December 2019 as this is the last date for resubmissions. Once all data is harvested into the database the actual transfer of data in the new databases can begin in February/March 2020. An ETC task will be created to support this work and to apply QA checks on the imported data.

These QA checks will be designed to verify that the data transfer has been successful and will include checks such as:

- Comparison of the number of facilities per country in each year between the E-PRTR and new databases;
- Comparison of the total releases or transfers of pollutants;
- Frequency of reporting of individual pollutants.

Reporters will also be asked to submit PlantID data for each LCP installation part reported in the 2017 EU Registry dataset. The priority will be to import the E-PRTR dataset into the new databases, hence the import of LCP data may not happen until later in 2020 depending on available resources.

Should you have any questions in relation to this process then please contact Ian Marnane: ian.marnane@eea.europa.eu.