

Cleaner Data for a Better National Emission Inventory

Sally Dombrowski,
Douglas Solomon, Rene Carrier
U. S. Environmental Protection Agency

National Emission Inventory 1996 - 2002

- 1996 – Anything goes
- 1999
 - Standardized format – National Input Format
 - QC Tool Developed
 - Central Data Exchange Submittals
- 2002
 - Agencies Accustom to NIF format
 - QC Tool Widely Used
 - Still too many errors to correct

Current Process Is Flawed

- Problems for both EPA and submitting agencies
 - Simple format errors not being corrected
 - EPA not finding errors until after data is received.
 - Frequent re-engagement with S/L/Ts to correct data

A New Approach

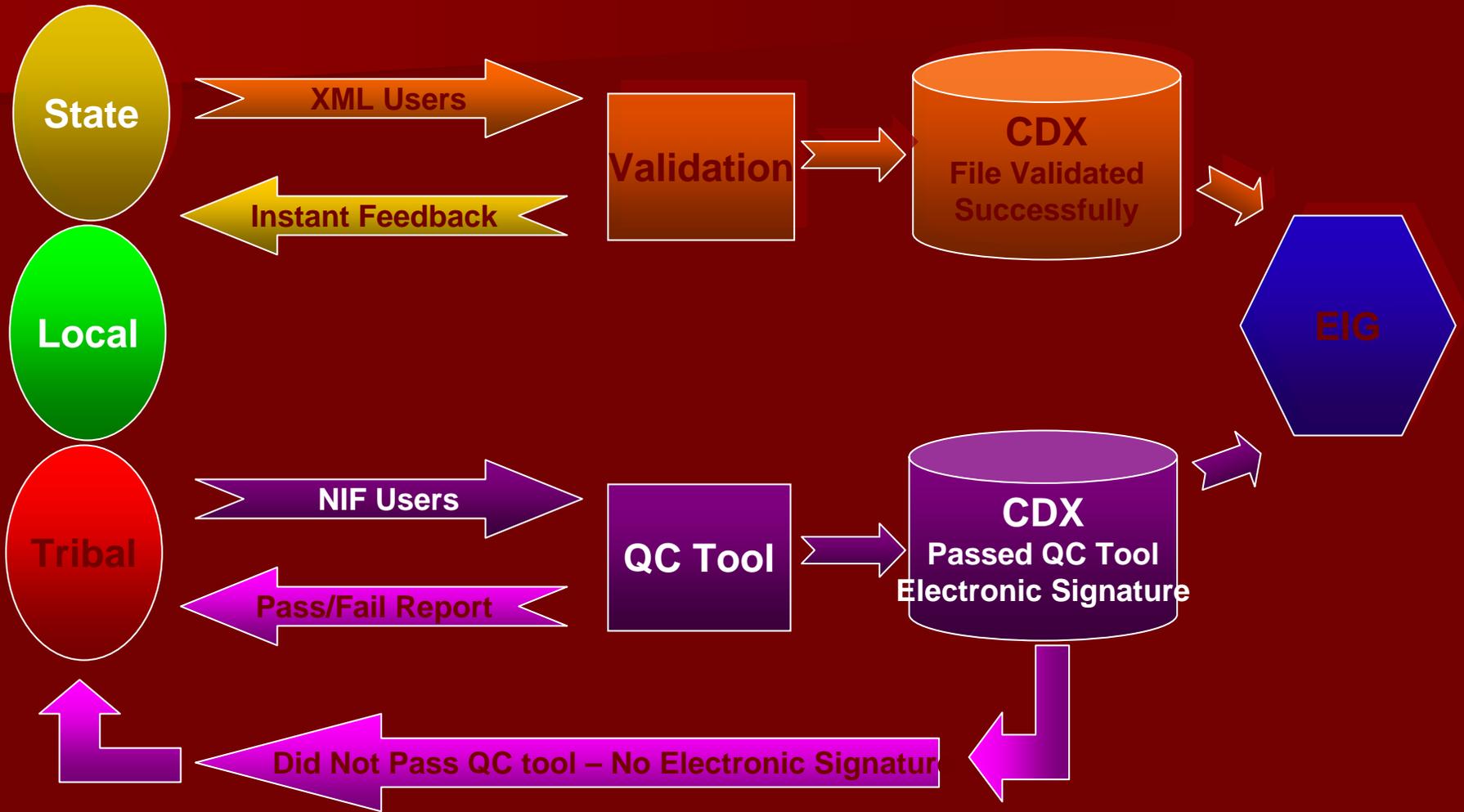
- Only accept data files which have passed specific QC guidelines
- Standardized QC practices will clarify EPA's data expectations
 - Help with automation and facilitate the practice.
- Data checks to be done at the Central Data Exchange (CDX)
 - File returned to submitter if it fails

Pass/Fail System

- Little difference for XML users
- New Basic/Format Content Checker
 - Two report system
 - Auto default
 - Electronic signature or license number



Central Data Exchange (CDX) Quality Control



What Will Be Checked?

- Correct Codes
- Correct Units
- Referential Integrity
- Absolute Ranges (7 days in a week)
- Stack Parameters?

Benefits

- S/L/T will save resources by not having to re-engage with data repeatedly.
- EPA would receive data free of format errors.
 - Resource spent on data analysis instead of data correction.

Questions?