Electronic Greenhouse Gas Data Collection and Management: Some Key Insights

U.S. GHG Reporting Program

Kong Chiu
U.S. EPA
World Bank Partnership for Market Readiness
Regional GHG MRV Training Workshop
17 – 19 September, 2014
Izmir, Turkey

Presentation Outline

- U.S. GHG Data System Overview
- Key Business Challenges for the Data System
- Insights for System Design and Development
  - Architecture
  - Data collection
  - High quality data = high quality submissions
  - Coordination of software and regulation development
  - Electronic signature
  - Managing entity relationships

Overview of
EPA Electronic Greenhouse Gas Data System

GHG Data System Overview

- Greenhouse Gas Data System Supports
  - Collection, verification and publication of GHG Data collected under the Mandatory Reporting of Greenhouse Gases Rule (40 CFR Part 98)
- Key Components
  - Electronic Greenhouse Gas Reporting Tool (e-GGRT)
  - Integrated Verification Process (iVP)
  - Publication Portal (FLIGHT)
  - Business Intelligence (Spago)
- Related
  - EPA EnviroFacts
  - EPA Central Data Exchange (CDX), Facility Registry Service (FRS)

Electronic Reporting Data Flow

EPA GHG Data System

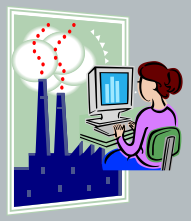
e-GGRT Electronic Greenhouse Gas Reporting Tool

EPA EnviroFacts:
Serviceable, searchable and separately hosted copy of non-CBI dataset.

State-Specific Service Oriented data flow using EnviroFacts API

Downloadable XLS, XML & HTML Data Files

Data Verification (EPA)



Reporters
汇报方
Facilities and Suppliers

e-GGRT Electronic Greenhouse Gas Reporting Tool

e-GGRT, interactive web-based, data reporting tool

EPA

Data Collection

e-GGRT Datamart

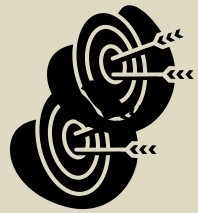
GHGRP Data Publication Website (FLIGHT)
(ghgdata.epa.gov)

Data Publication

Key Development Challenges

The System Development Business Challenge

- **2007, December:** FY2008 Appropriations Act
- **2009, April:** GHG Reporting Rule Proposed
- **2009, October:** GHG Reporting Rule Finalized
  - Electronic reporting only (1st for EPA)
  - Collect GHG data across most sectors of economy
  - Over two dozen unique industry source categories
  - Multiple GHG measurement methodologies
  - Multiple tiers
  - Dozens of fuels and fuel types
  - Sensitive or confidential data
  - Thousands of expected reporters
  - Thousands of users, all new entrants with steep learning curve
  - Has to support EPA verification of the data
- **2010, October & December:** Technical corrections and clarifications
- **2010, December:** Proposal to defer collection of inputs to equations
- **2011, March:** (initial) First reporting deadline
- **2011, September:** (final) First reporting deadline



Data Collected under US GHG Reporting Program

- **Facility or supplier information includes:**
  - Name, address, latitude/longitude (in some cases)
  - North American Industrial Classification Codes (NAICS)
  - Parent company information
- **GHG emissions include:**
  - CO2 equivalent emissions (metric tons)
    - across all applicable sources
    - by each applicable subpart (source category) at the facility
    - excluding biogenic emissions
    - biogenic emissions
  - CO2 equivalent quantity from supplier categories (metric tons)
  - CH4, N20 emissions for the facility by subpart
  - Emissions of each fluorinated GHG (F-Gas)

Data Collected, cont'd

- **Special Data Includes:**
  - Explanation of calculation methodology changes during the reporting year
  - Description of Best Available Monitoring Methods (BAMM) used during the reporting year
  - Requests for extension of the use of Best Available Monitoring Methods
  - Supporting documentation (document upload) for BAMM extension requests
  - Identification of each data element for which a missing data procedure was used
  - Total number of hours in the year that a missing data procedure was used for each element
  - Geologic sequestration monitoring, reporting and verification plan (document upload under Subpart RR)

Data Collected (3)

- **Source Specific Data Includes*:** (Example from Subpart Q, Iron and Steel)
  - Unit identification
  - Unit type (e.g. taconite indurating furnace, Electric Arc Furnace etc…)
  - Annual CO2 emissions for each unit
  - Annual quantity taconite pellets, coke, sinter, iron and raw steel for each unit
  - Method used (i.e. carbon mass balance or site-specific emission factor) for each unit
  - Annual mass of each process inputs and outputs to determine CO2 emissions
  - Annual volume of each type of gaseous and liquid fuel

**in some cases collection of specific data elements may have been deferred*

Special Challenge:
Confidential Business Information (CBI)

Approach to Confidentiality Determinations (1)

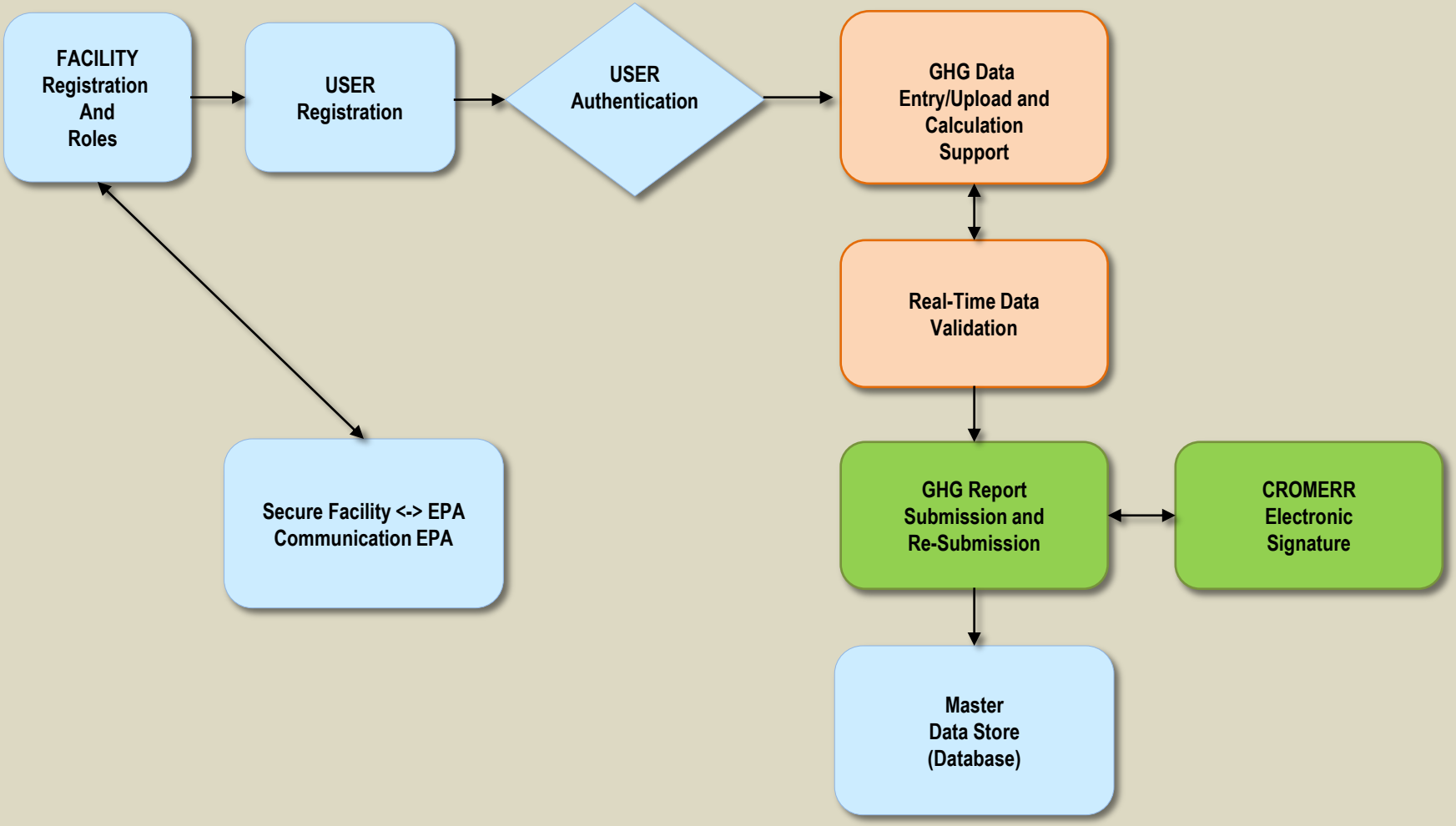
- Case-by-case confidentiality determinations not practical, given the number or reporters (~8,000) and data elements (~2,000)
  - Would not result in timely release of data
  - Burden on reporters and the agency
- From 2010-2012, the EPA proposed confidentiality determinations for the data elements required to be reported using a notice and comment process.
- The confidentiality of each reported data element was determined using a two-step approach:
  1. Grouping data elements into 11 data categories (e.g., inputs to emission equations, emissions, and unit/process operating characteristics that are not inputs to emission equations for direct emitter source categories) and
  2. Making confidentiality determinations either categorically or on the basis of individual data elements.
- The EPA has now largely finalized confidentiality determinations for data elements except those in the "inputs to emission equations" category. (EPA proposed a rule addressing inputs in September 2013)

Examples

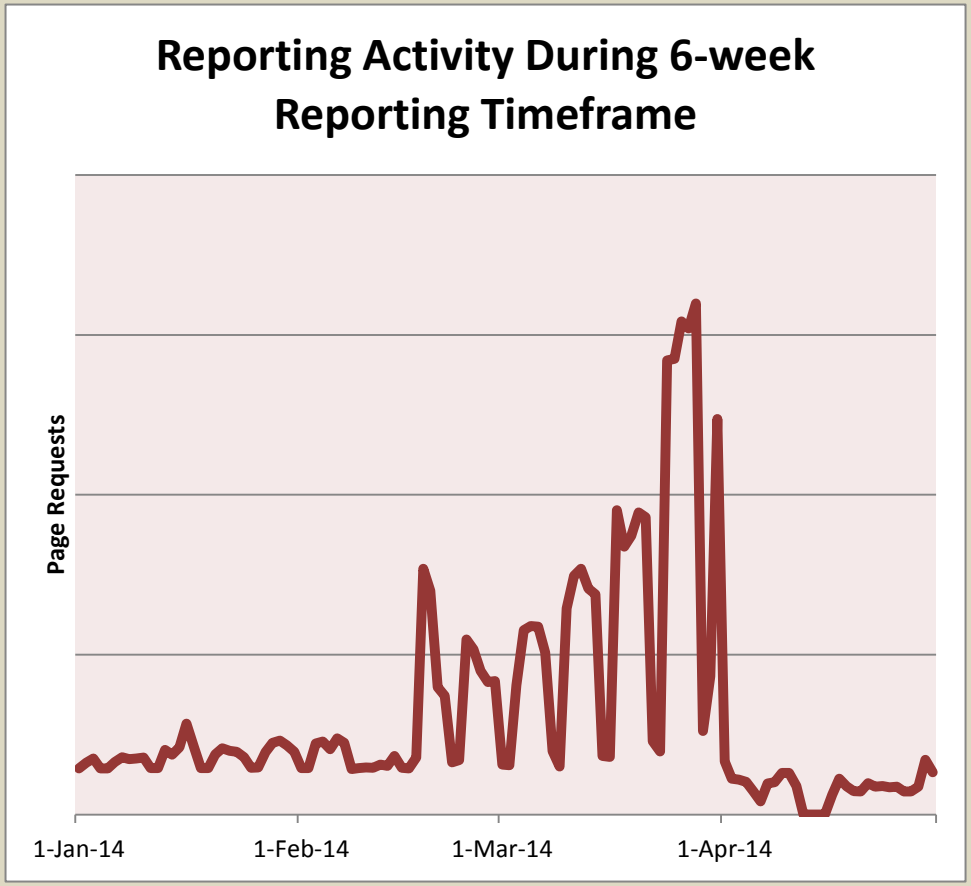
Source Category	Description of Data	Status (confidentiality determination as of 12/2013)
Q: Iron and Steel	Unit Identification Number	Emission Data
	Unit type	Emission Data
	Annual CO2 emissions for each unit	Emission Data
	Annual quantity taconite pellets, coke, sinter, iron and raw steel for each unit	CBI
	Method used	Emission Data
	Carbon content of each process input used to determine CO2 emissions	Input to Emission Equation Deferred Until 2015
C: Combustion	Unit ID number	Emission Data
	Maximum rated heat input capacity	Emission Data
	Types of fuel combusted during the report year	Emission Data
	Methodology (i.e. Tier) used to calculate emissions	Emission Data
	Annual CO2 mass emissions for each type of fuel combusted during the reporting year	Emission Data

Design and Development Insights

Include Key Electronic Reporting Functions



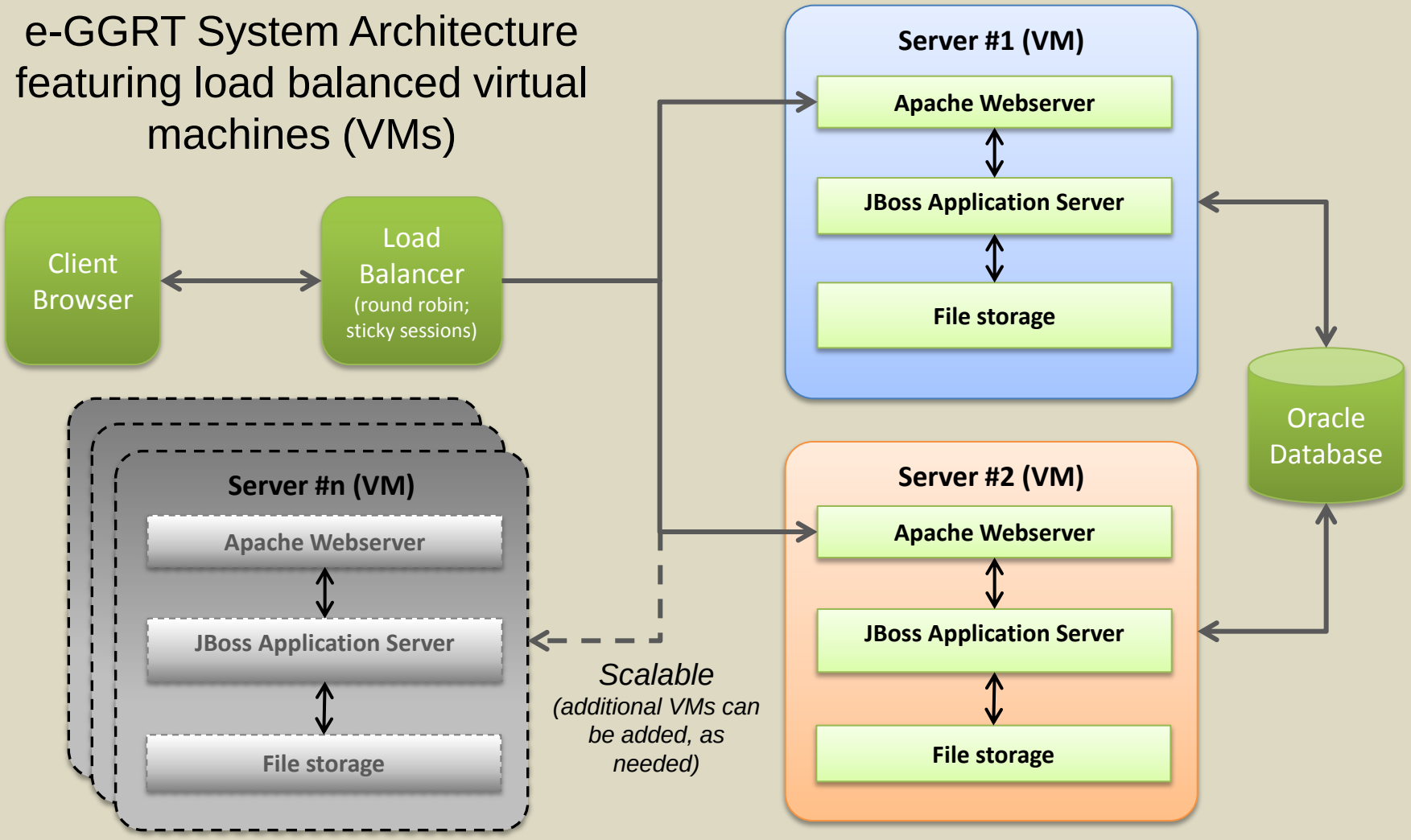
Manage Load through Architecture



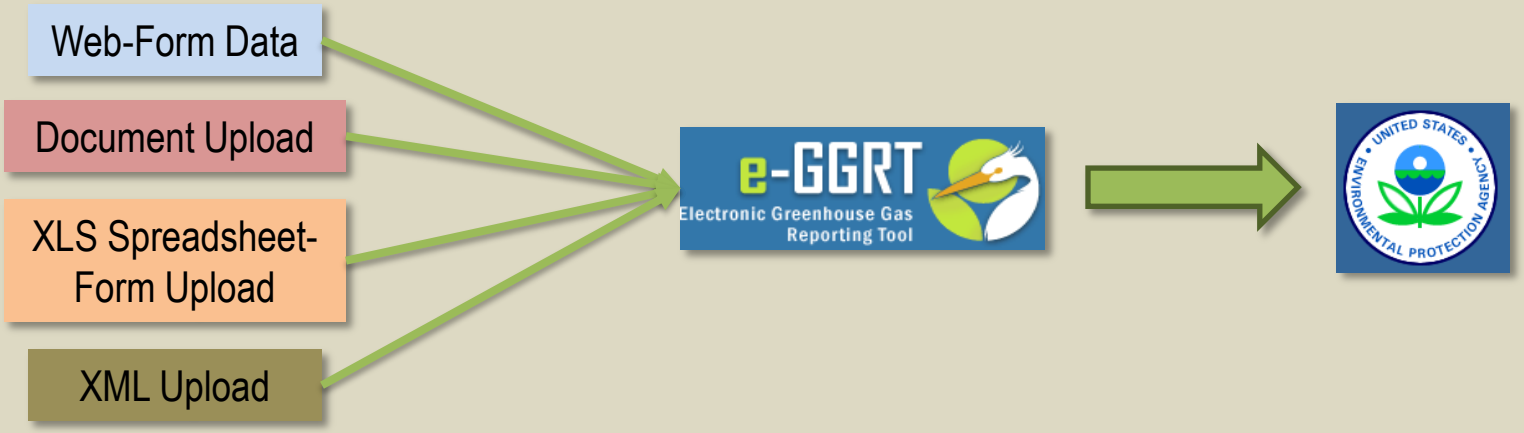
- Scalability
- Annual reporting deadline results in spikes in usage
- Architecture permits addition/removal of servers without disruption

Courtesy of Science Applications International Corporation

Scalable Architecture



Support Different Electronic Data Formats

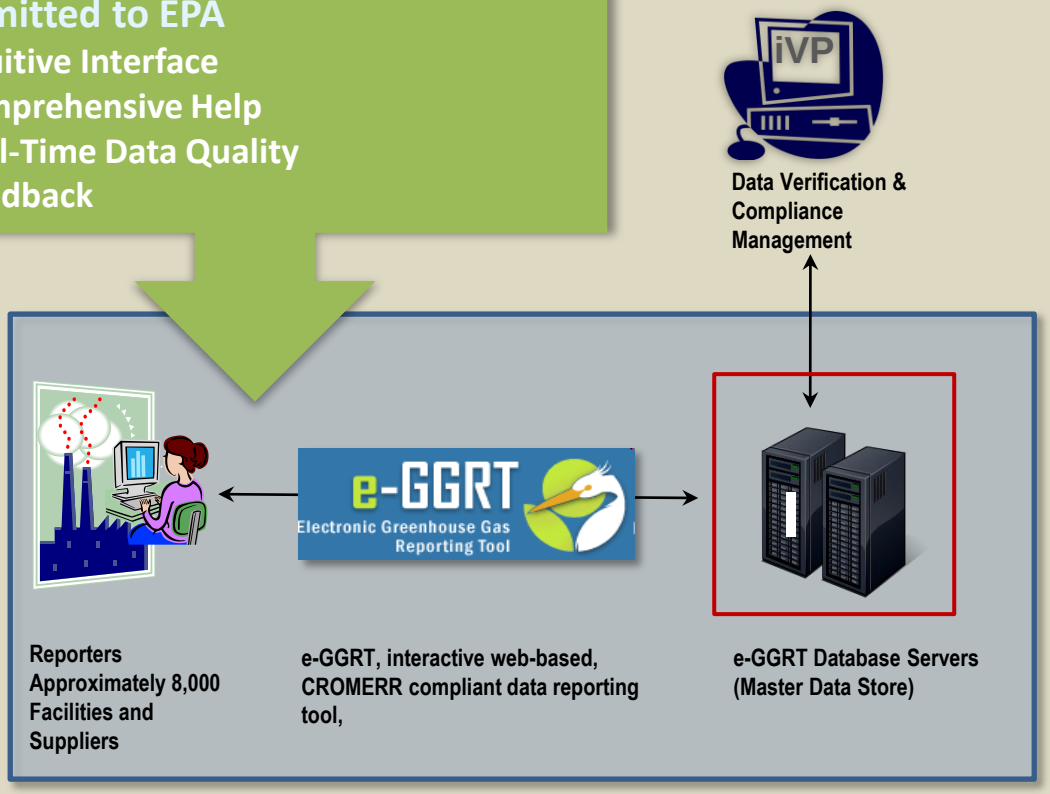


Data Format	Used For
Numerical data, via web-form	Emissions and activity data, entered directly into e-GGRT web-form fields
Free text, via web-form	Explanations of exceptions and special circumstances, entered directly into e-GGRT web-form
Documents (PDF, XLS, Doc, ZIP etc)	MRV plans under subpart RR, Supporting information under subpart W BAMMs
Numerical/text via XLS smart form	Used in certain subpart modules to ingest facility's GHG data, compatible with verification module
Numerical/text via XML file	Used in lieu of web-forms to upload GHG annual report into e-GGRT

High Quality Data begins with High Quality Submissions

Design software and system to improve data quality before user submits it to EPA

Improve data quality *before* it is submitted to EPA
- Intuitive Interface
- Comprehensive Help
- Real-Time Data Quality Feedback

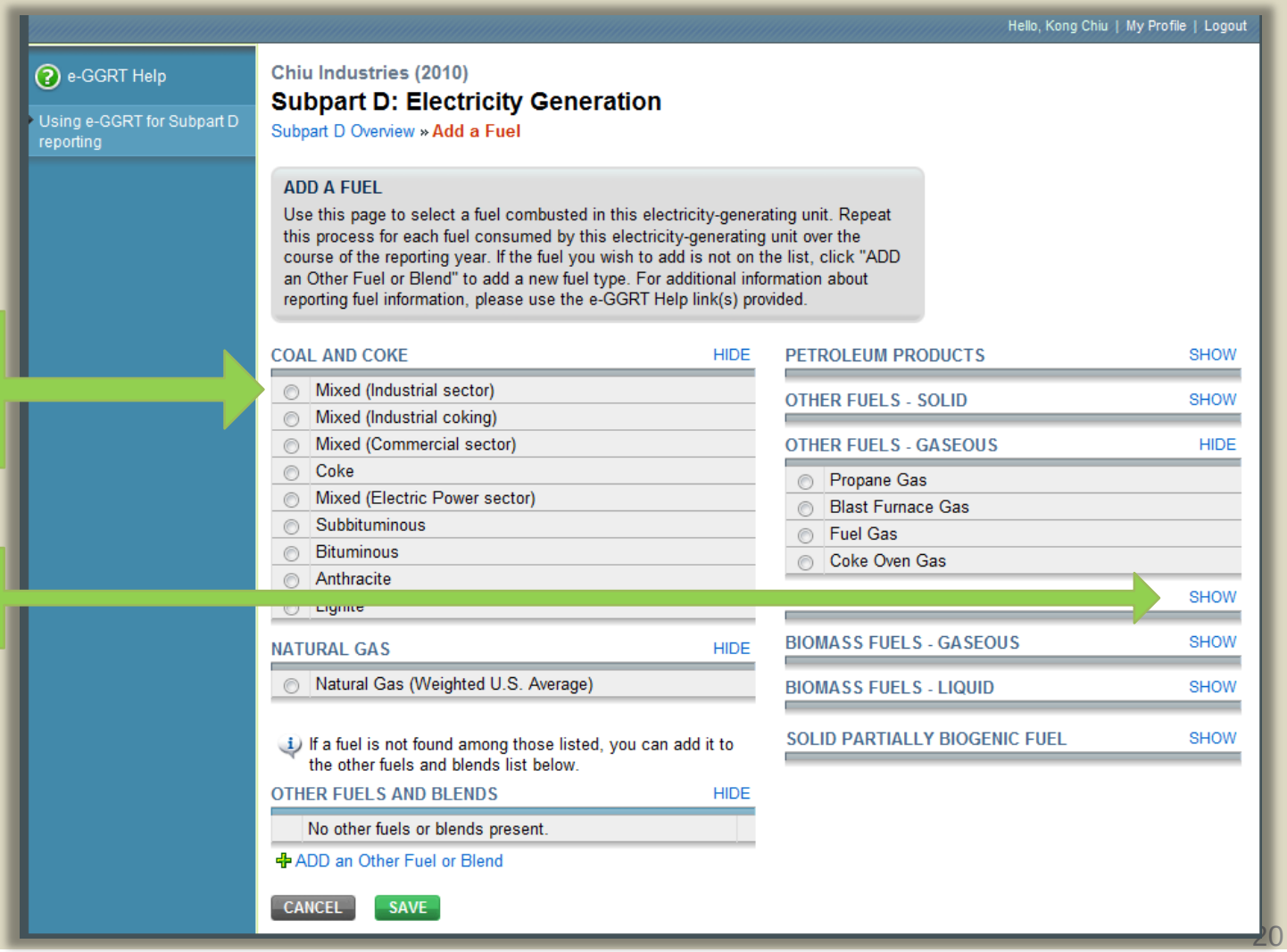


High Quality Submissions (cont'd)

- Intuitive User Interface (Table C-1)

Easy to find and select multiple Fuels/types

Lists Expand and Contract

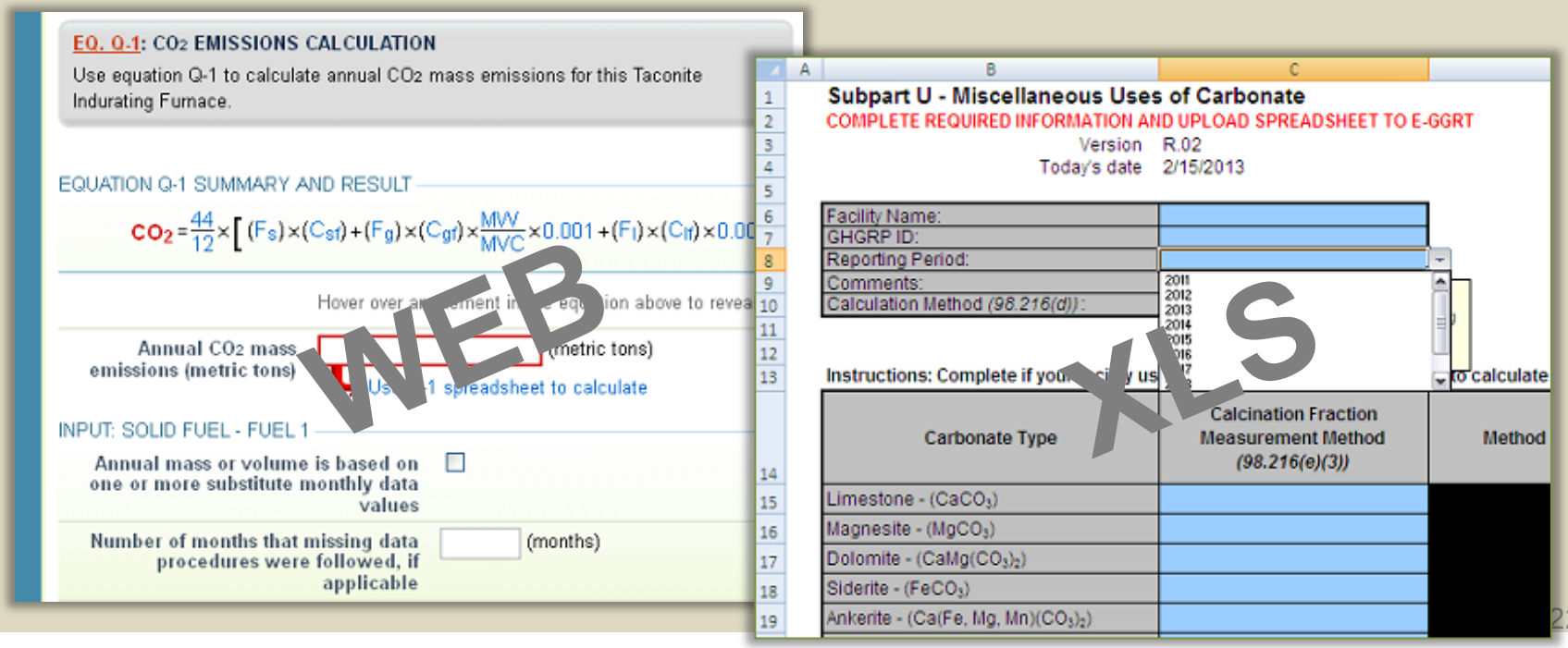


High Quality Submissions (cont'd)

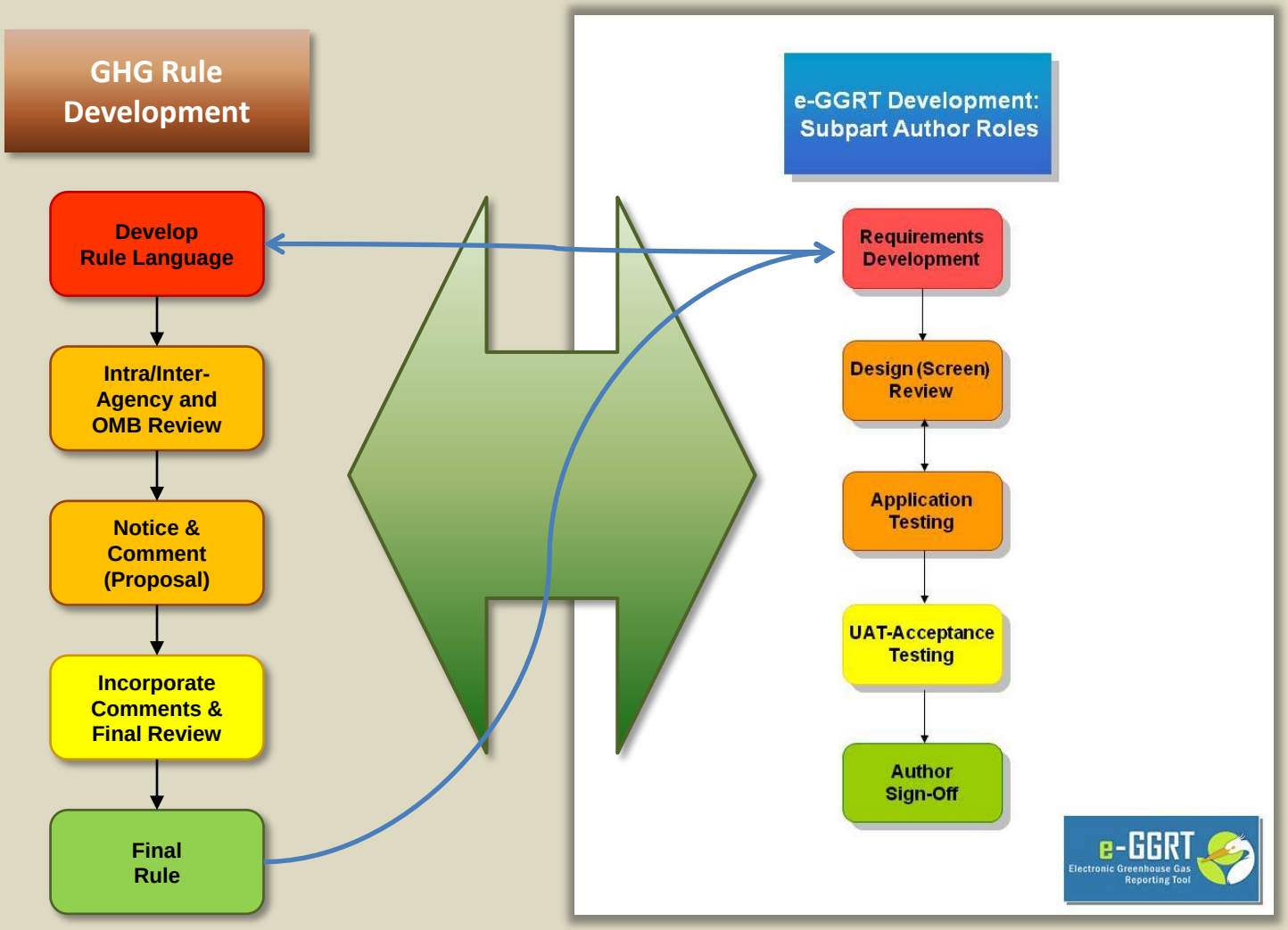
- Real-Time Data Quality Feedback

Hybridize Data Entry

- Challenge: Short timeline, changing business requirements
- Web form data entry
  - User friendly
  - Significant development and testing effort
  - Direct parsing of entered data
- Spreadsheet Reporting form (Microsoft XLS)
  - Faster development and testing
  - Harder to parse data



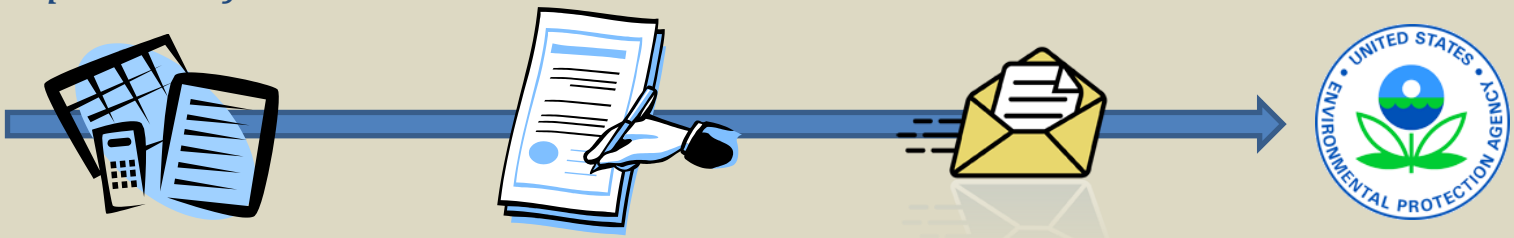
Coordinate Regulation and Data System Development



Electronic Signatures

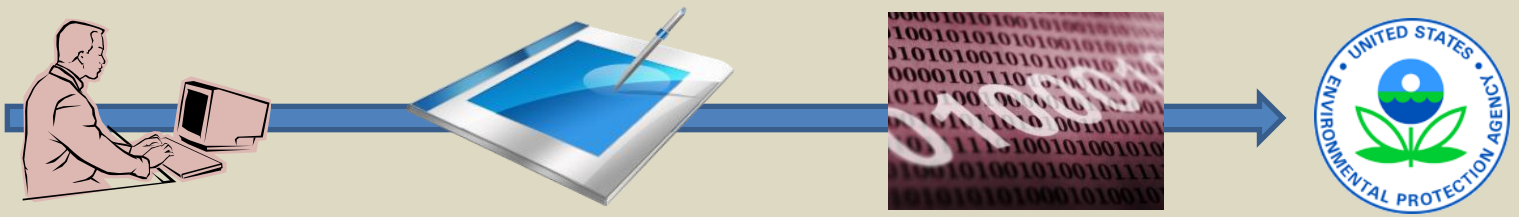
Paper vs Electronic Reporting

Paper workflow



Complete Forms | Sign | Mail to EPA

Electronic workflow

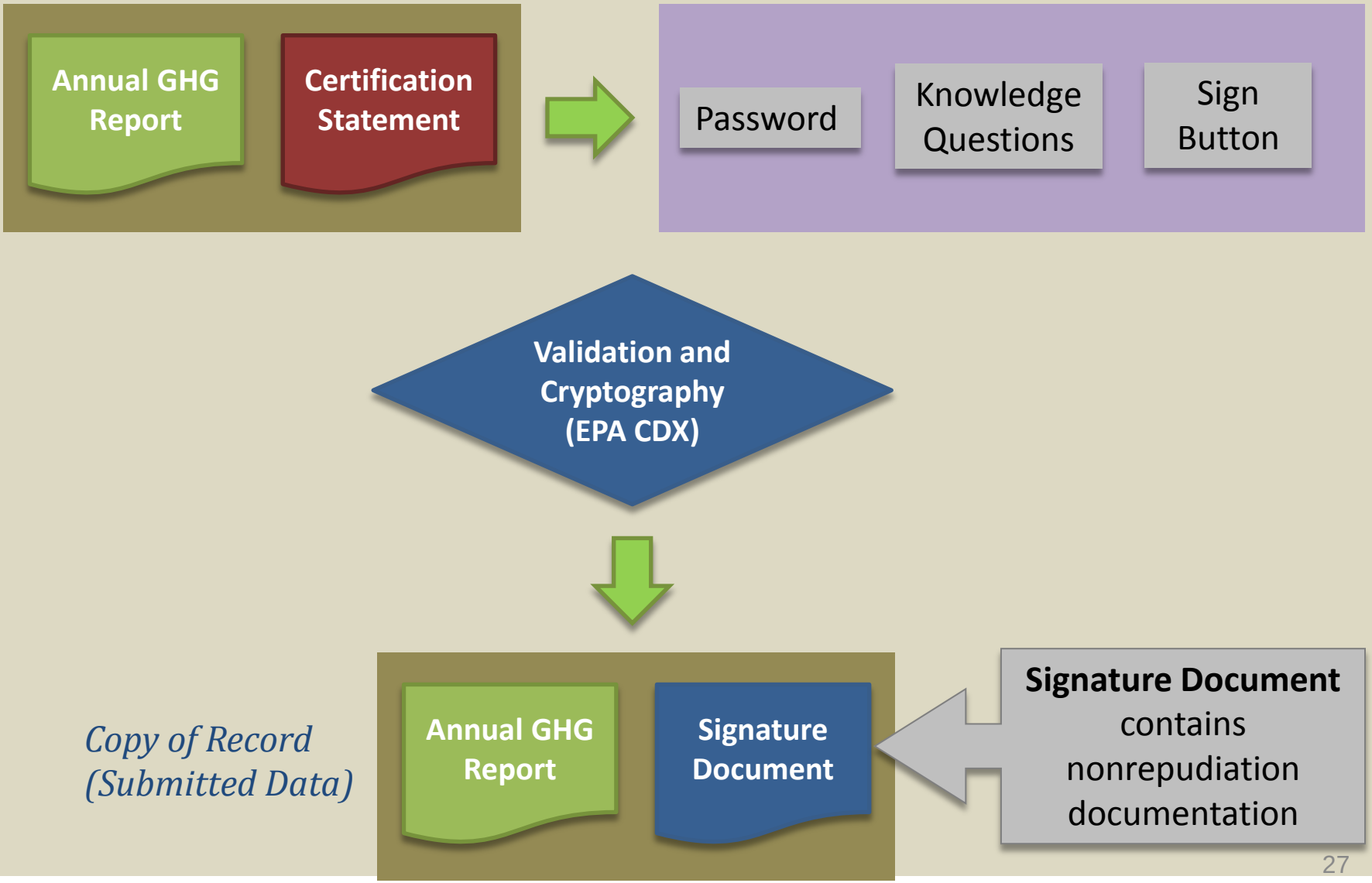


Enter Data | Electronic Signature | Electronic Submission to EPA

CROMERR, US Code, Title 40, Part 3

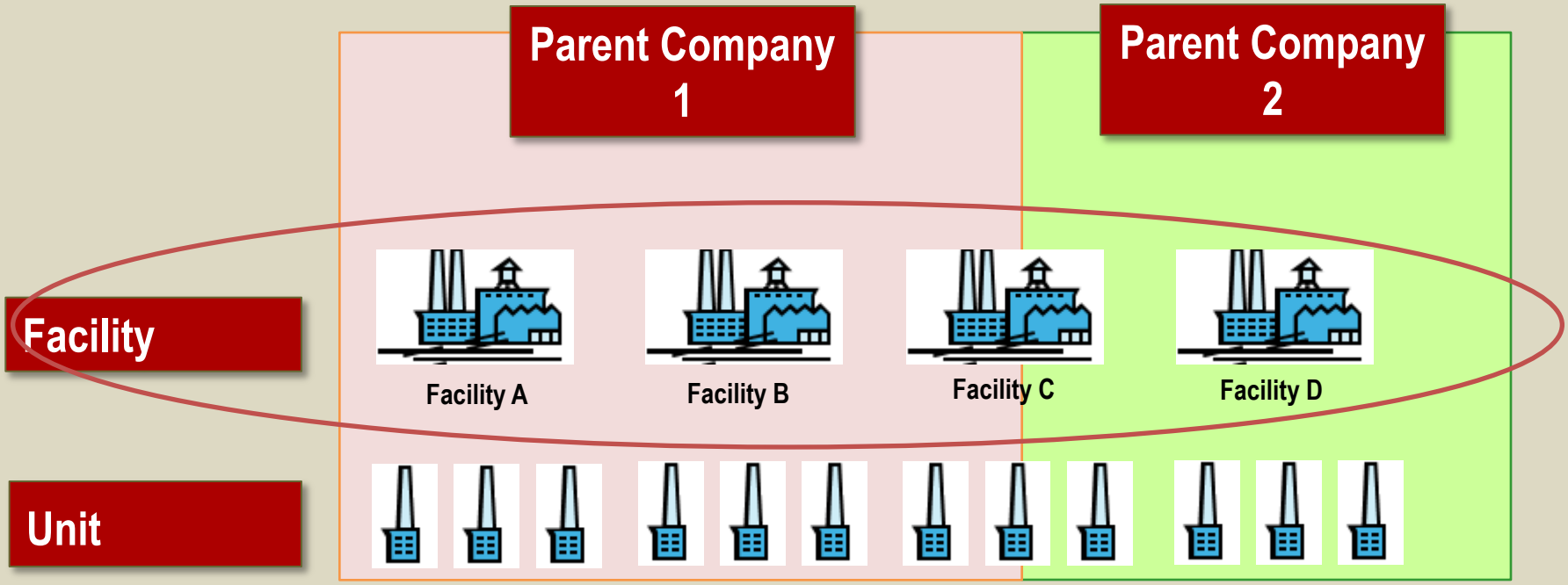
- Cross Media Electronic Reporting Rule (CROMERR)
- Electronic documents must have valid electronic signature if paper version requires signature
- The electronic document receiving system must be able to prove, legally that:
  - The electronic document was not altered without detection during transmission or after receipt
  - The electronic document was submitted knowingly
  - Any identified submitter or signatory had the opportunity to review the copy of record in a human-readable format
  - Each electronic signature was valid at the time of signing
  - Each signatory has signed an **electronic signature agreement**
  - The Identity of every individual using an electronic signature device has been determined with legal certainty

e-GGRT CROMERR Workflow



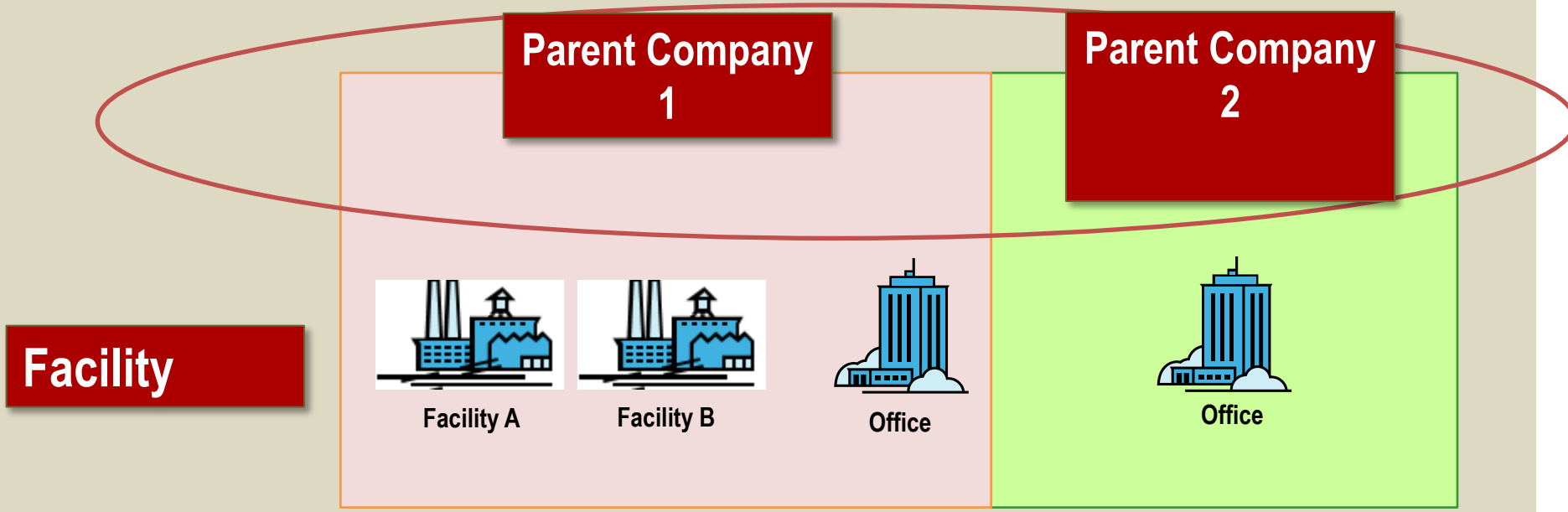
Managing Facilities and Users Electronically

Reporting Level: Direct Emissions



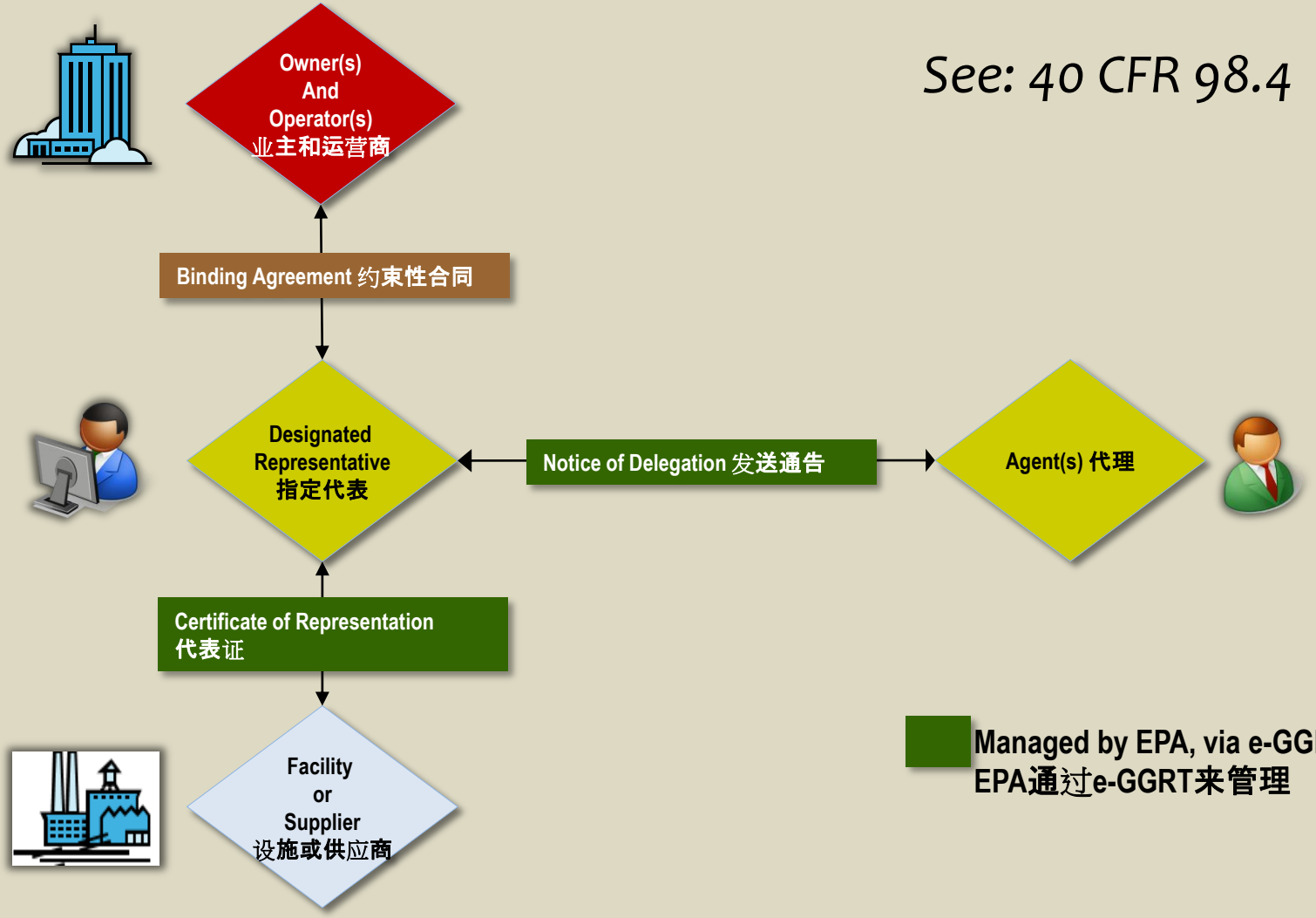
- **Examples:** Iron and Steel, Power Generation, Cement Production

Reporting Level: Suppliers
汇报级别:



- **Examples:** Suppliers of Petroleum Products, Suppliers of Industrial GHGs, Natural Gas Supply

Designated Representative for Facility or Supplier



See: 40 CFR 98.4

Managed by EPA, via e-GGRT
EPA通过e-GGRT来管理

Thank You! For More Information:

- Part 98 (U.S. GHG Reporting Rule) Info:
  - www.epa.gov/ghgreporting
- Published Data:
  - ghgdata.epa.gov

Kong Chiu
US EPA
1200 Pennsylvania Ave NW
Washington, DC 20460 USA
1+202-343-9309
chiu.kong@epa.gov