Development of Functional Requirements for an EIA Corporate Internet Data Collection System

FEDCASIC 2007

www.eia.doe.gov

Introduction

- Background and objective of the project
- Three part focus of the work
- Research we did before writing requirements
- Detailed functional requirements' areas

Background & Purpose

- Project Purpose -- Develop the functional requirements for an EIA wide Internet Data Collection System (IDC)
- Interoffice team was formed to get an agency-wide perspective and buy-in

Three Part Focus

- The primary work focused on three parts:

  1) The development of IDC "principles"
  2) The research of other Federal Statistical agencies' approaches and literature search of Internet data collection to determine the best practices and approaches relevant to these EIA principles
  3) The detailed key functional requirements for the development of a successful application at EIA

IDC Principles

- IDC Team's work first centered on the development of IDC "principles"
- These principles are generalizations which were accepted and used as the basis for reasoning in developing the functional requirements
- These principles represented three dimensions—the respondent point of view, the survey program point of view, and the EIA corporate point of view
- For each of these viewpoints, the IDC Team agreed upon which characteristics would be important

Respondent Point of View

- Ease of access and use
- Ready and useful help and prompt support
- Ability to report large amounts of data without keying or using direct data transfer
- Ability to interrupt and complete surveys later, save/print/forward, and transmit information in both directions
- Confidentiality protection
- 24-7 access
- Reduced burden compared to other modes of reporting
- Other incentives for respondents ("What's in it for me?"/first results)

Survey Program Point of View

- Reduced survey processing cycle time
- Reduced data entry costs, mailing costs, and costs associated with data editing and follow-up with respondents
- Instrument design and flexibility (individual survey needs differ)
- Flexibility and ease of changing/modifying survey definitions, instructions, forms, respondent sets or other metadata directly by program offices at minimal cost

Survey Program Point of View (cont)

- Reduced reporting errors by editing at data capture.
  --Edits (within cells, across cells, and across reports or periods); Ease and flexibility of adding or modifying edits and monitoring edit performance within IDC; Ease of resolving edit failures within IDC (including bypasses, comments, or change data
- Respondent case management and survey management capability
  --Electronic nonresponse notification
- Ability to process resubmission/revisions, ability to disallow revisions, or restrict resubmissions
- Ability to measure process performance in real time and historically (reliability, robustness, and flexibility)
- Ease of integration, interface, and synchronization with responses from other modes of reporting and survey processing system
- Data integrity and back-up

EIA Corporate Point of View

- A centralized platform
- Shared architecture and infrastructure for IDC (reduction in cost of systems' development and maintenance) and user management
- One entry point
- Common look and feel
- Security (for EIA)
- Ability to support multiple surveys
- 508 compliant
- Life cycle records' management

The Research: Visits to Statistical Agencies

- 11 broad questions sent in advance of visit
- Visits: BLS, Census (2), NASS, Westat/Blaise (2), EIA (4)
- Main finding: three approaches:
  1) fully centralized, limited flexibility

  2) centralized core structure w/survey flexibility and/or control

  3) fully de-centralized (i.e. not common)
- Other findings: Little evidence of cost savings; evidence of higher quality data; implementations are few thus far; often started with paper image but quickly moved on; take-up is low (other than in EIA); help support underestimated; organizational culture and buy-in is critical (What's in it for me?)

Research: Literature Search

- Information from published/unpublished papers, workshops and conference were reviewed for best practices.
- Highlighted best practice was usability testing to determine how the respondent interacts and behaves within the IDC. The perspective of the respondents to the particular survey is the most important factor.
- Overview and summary of the research on usability at the U.S. Census Bureau reviewed in a paper *Designing Edits for Electronic Economic Surveys and Censuses: Issues and Guidelines,* by Anderson, Murphy, Nichols, Sigman, and Willimack that identifies best practices.
- Conferences such as the annual FedCASIC, Work Session of the UNECE Task Force on Electronic Data Reporting for Primary Data Collection, *FCSM-GSS Workshop on Web-Based Data Collection* have provided a forum to address many of the issues in Internet data collection.

Recommendations: Overview

- Recommend: Follow 4 basic guiding rules supported by detailed requirements. The IDC:

  1) **easy** to implement, modify, or migrate a survey to the IDC
  2) tools enabling the survey group ownership
  3) **value added from the respondent's view**
  4) promotes **high quality data** through editing features, user notifications, clear navigation, etc

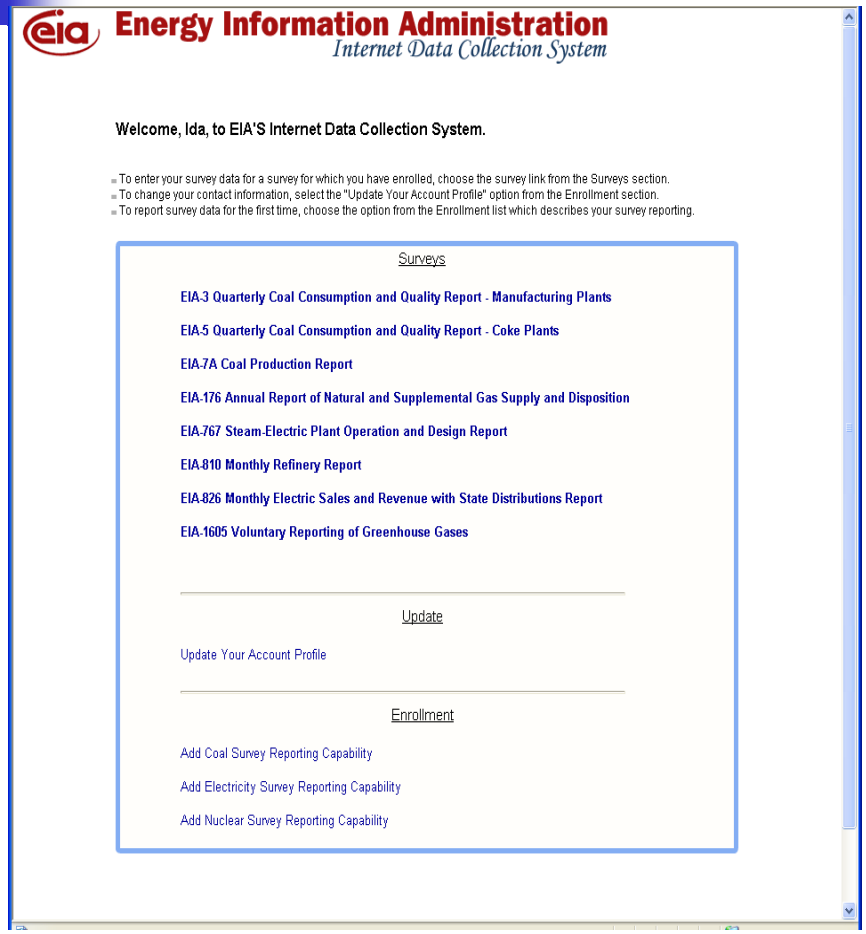
- Recommend: Form a **partnership** among the support offices and the program offices

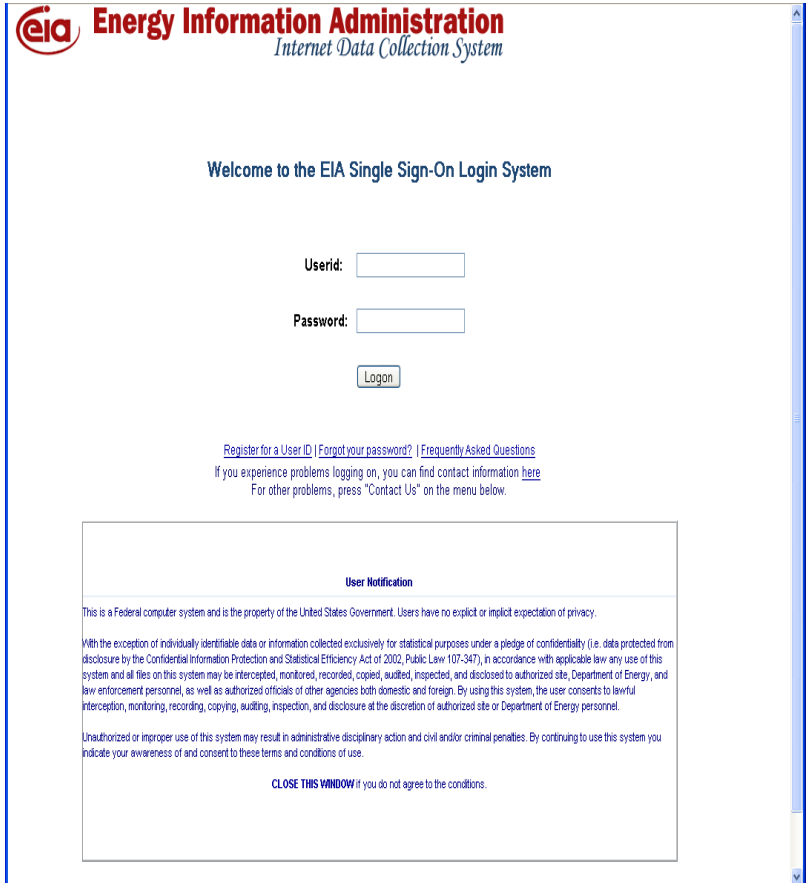
Detailed Key Functional Requirements

- The resulting functional requirements for EIA's ***next generation*** of internet data collection
- Intended to be a framework to be used by developers
- Focus on ***what*** is required, rather than ***how*** to satisfy those requirements.
- Requirements were categorized by: Graphical User Interface, Front-end editing, User notification, Interoperability, Performance measures, Retention and disposition, Essential support services, Security, Administrative module, and Implementation plan

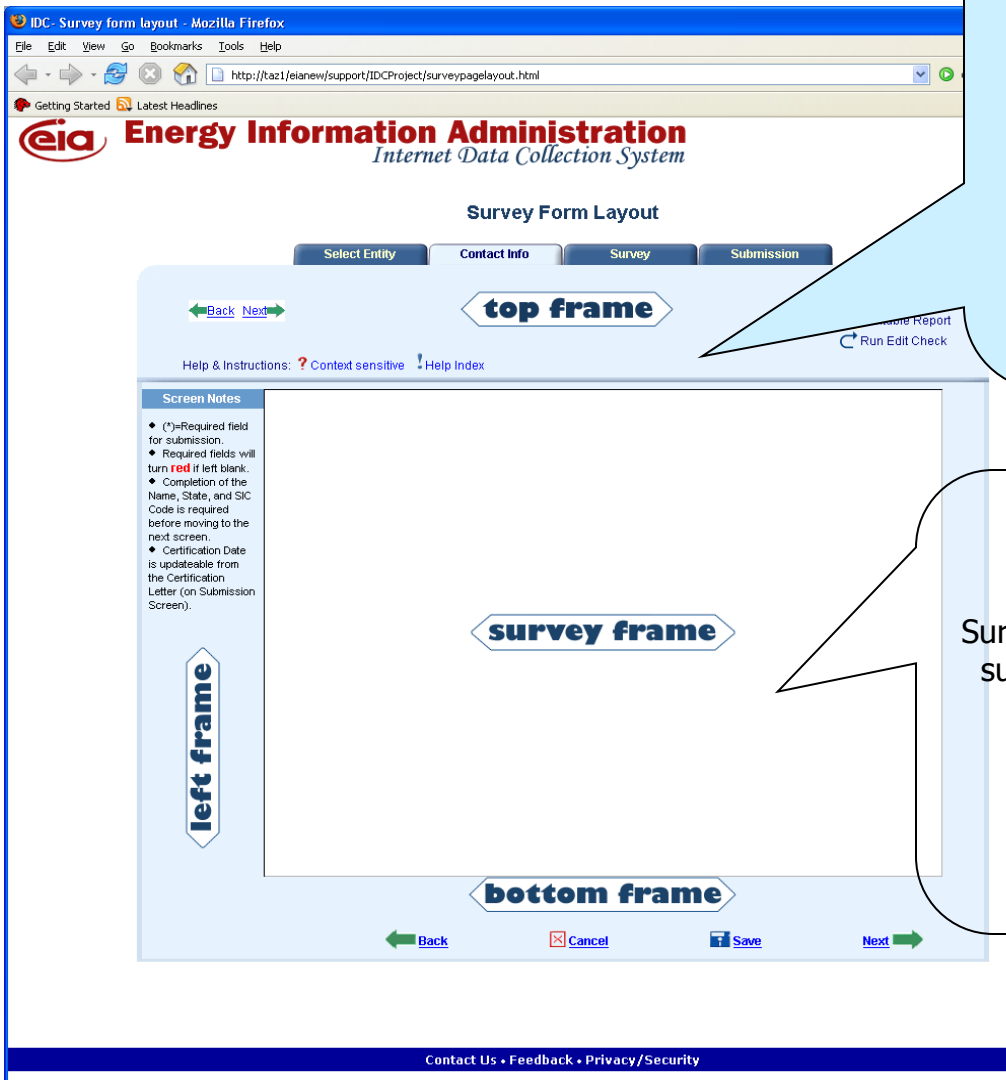
Recommendations:

Data Collection (Graphical User Interface - **Single Sign On**)





GUI-Common Look and Feel



Common Aspects of IDC

Infrastructure
Single Sign On
Registration and Enrollment
Look, Feel and Functionality of "Blue" area of GUI including Help Function
Database Schema and Application Access to Database
Data Import Function
Administrative Module
Repository for Source Code and Documents

Survey-Specific Aspect of IDC

Survey groups may construct their surveys as they wish subject to the requirements in the IDC Requirements Document

www.eia.doe.gov

Recommendations:

Data Collection **("Front End" Editing)**

- **Prevent errors** through drop downs
- Provide capability for "**hard**" edits—**must correct**
- Provide capability for "**soft**" edits **—can correct, comment or bypass**
- Provide ability to **implement various types of edit rules** such as consistency or range edits
- Provide data needed for the **various types of edits**
- Provide capability for survey managers to modify

Recommendations:
Data Collection **(User Notification)**

- Provide **dual notification** that official submission received
- Provide **non-response notification** and **targeted emails** (respondent or non-respondent) as defined by the survey
- Provide **notification of data not saved/submitted** if browser closed prior to that action.

Recommendations: Interoperability

- **IDC/Respondent Communication**

  1) Respondent imports or performs data entry

  2) Respondent can attach supplementary materials

  3) IDC displays contact information and historical submissions in browser

- **IDC/Processing System Communication**

  1) File can be exported from IDC to Processing system

  2) Processing group can choose to use IDC database directly

  3) Processing system passes respondent list and (possibly) revised respondent data back to IDC

Recommendations:

Performance Measures

- **Measures about the Survey**

  --Response Measures, Edit Failure & Correction/Bypass Measures, Corrections/Resubmissions Measures
- **Measures about the Collection Mode**

  --Access counts, Respondent Evaluation/Usability
- **Measures on Contribution of IDC Mode to Overall Survey Performance**

  --Cost per respondent, "back-end" edit failure/call back/corrections comparisons to other modes, change in meeting/exceeding dissemination deadlines

Recommendations:

Retention and Disposition

The IDC will have the capability of:

- Scheduling data retention for IDC data
- Placing an identifiable metadata tag on records
- Providing access to IDC data records using search criteria (for authorized users only)

Recommendations:

Essential Support Services

- **Two help desks**: (1) maintained by USC to assist with systems problems/questions/concerns, (2) maintained by the program office to assist respondents with survey specific questions.
- **Usability Testing**: all surveys must undergo usability testing with (1) a sample of actual survey respondents, (2) internal EIA users.
- **Documentation**: 1) online survey instructions and user guides for the respondents. 2) systems documentation and other internally used documentation maintained in a central location for all of EIA.

Recommendations:

Security

- Comply with the EIA Information Systems Security Policy
- Data transmitted between the client and the server protected from interception and be readable only by the intended recipient
- A Security and Privacy statement on the Logon page of SSO
- Passwords hashed and client and the server authenticated to one another.
- Individual work sessions are closed to unauthorized users from other workstations.

Summary

- The Corporate IDC
  --based on **agreed "principles"**
  --translated to **high level requirements**
  --drove **detailed requirements**
    --through researching **best practices of other agencies'** approaches
    --**literature search** on internet data collection to determine the best practices and approaches **relevant to these EIA principles**