XML Schema
Activities at IRS
U.S. Department of Treasury
Internal Revenue Service

Dr. Glenda Hayes – MITRE
ghayes@mitre.org

Meeting IRS Business Needs with XML

- XML = Industry Standard for Data Exchange
  - Ease of tax compliance and lower cost for taxpayers
  - Ease of entry and lower cost for software companies
  - Subsidiary benefits such as shorter audit cycle time

- XML Schema + XSL = Encoded Business Rules
  - Fewer Errors
  - Reduced Costs for Validating Returns
  - Easier Transitions for Tax Changes
Benefits of Using XML for 1120

- 1120 represents ~85% data elements for IRS
- Significant overlap with Individual Tax (1040)
- Effort will reconcile duplicate data elements (e.g., string length for address)
- Meets external stakeholder request
- IS/ETA joint effort
- Scoped pilot
- Exploit larger base of developers
- Coordination with external groups
Three Approaches for Tax XML Schema Development

- **XBRL (eXtensible Business Reporting Language)**
  - Abstracted data exchange, `<item>` tag
  - 2-step validation process: XML parsers + XBRL process

- **EDI based (TIGERS)**
  - Tax Information Group for EC Requirements Standardization
  - Schema patterned after EDI transaction
  - Develop derived schemas that inherit from generic, Base Schema

- **Direct**
  - Schema patterned after Microsoft 1040 strawman and TIGERS Sales Tax Group

---

IRS XML Work Plan

- Draft XML Schemas for all 3 variants for 1120 kernel
  - XBRL
  - EDI-based (TIGERS)
  - Direct

- Develop Instance Documents for 1120 kernel
- Develop Stylesheets to Transform among Variants
- Coordinate with other Stakeholders re: Merits of each Variant
  - TIGERS
  - International Tax Agencies
  - Vendors
  - Corporations

- Develop all related 1120 forms and schedules in selected variant
One IRS Objective:
Validate numeric values

Value is too large for IRS systems (50 digits)

MITRE Approach for 1120 - Direct

IRS Data Types
IRS Common Structures

IRSXMLSchemaTypes.xsd
IRSXMLGeneralElements.xsd
IRS1120SchedASchema.xsd
IRS1120SchedCSchema.xsd
IRS1120SchedESchema.xsd
IRS1120SchedJSchema.xsd
IRS1120SchedKSchema.xsd
IRS1120SchedLSchema.xsd
IRS1120SchedM1Schema.xsd
IRS1120SchedM2Schema.xsd
IRS3115SchedASchema.xsd
IRS3115SchedBSchema.xsd
IRS3115SchedCSchema.xsd
IRS3115SchedDSchema.xsd
IRS3115SchedESchema.xsd
IRS4684Schema.xsd
IRS4797Schema.xsd
IRS851Schema.xsd
IRSBadDebsStatementSchema.xsd
IRSChargableContributionsStatementSchema.xsd
IRSControlledForeignPartnershipSchema.xsd
IRSDeminisExceptionSchema.xsd
IRSInstallmentSalesScheduleSchema.xsd
IRSInterestStatementSchema.xsd
IRSMarkToMarketElectionSchema.xsd
IRSSpecialNonaccualExperienceMethodScheduleSchema.xsd
IRSOtherDeductionsStatementSchema.xsd
IRSOtherIncomeStatementSchema.xsd
IRSStockOwnershipForeignCorpSchema.xsd
IRSTaxShelterStatementSchema.xsd

94 Schema Files – 1120

IRS1120Schema.xsd
Coordination Among Approaches

- **XBRL (eXtensible Business Reporting Language)**
  - XBRL 1.0 – Abstracted data exchange
  - XBRL 2.0 (in draft) – single-schema process, XBRL data types

- **TIGERS**
  - Tax Information Group for EC Requirements Standardization
  - Transition strategy – Develop XML vocabulary

- **Direct**
  - Domain-specific tags and business rules
  - Use of common data types and data structures

IRS Experiment
Current Status

- **Done:**
  - All forms and schedules for 1120 and 941 identified and included
    - Common Data Types: IRSXMLSchemaTypes.xsd
    - Common Data Structures: IRSXMLGeneralElements.xsd
    - Specific forms and schedules: 92

- **In Progress:**
  - Reconciling duplicate terms & conflicting data types
  - Developing instance documents
  - Updating IRS X12 info for TIGERS
  - Investigating viability of ebXML Transport as transmission specification

7 Nov 2001