

## Introduction to IRS e-file XML Schemas



1120 U.S. Corporation Income Tax Return

First calendar year 2001 or tax year beginning 2001 ending 2001

Check appropriate boxes:  Sole proprietor  Partnership  S corporation  Trust or estate  Beneficiary of a trust or estate  Estate of a decedent  Estate of a trust or estate  Other

Check applicable boxes:  International  Foreign asset  Non-foreign  International  Foreign asset  Non-foreign

March 18, 2002

## Schema Namespace

A single namespace is used for all e-file schemas

The standard SOAP schema (<http://schemas.xmlsoap.org/soap/envelope/>) is used without modification, except to work around a validation bug in XML Spy during development

2

## Naming Convention



An element name is made up of a meaningful phrase with each word capitalized and the spaces taken out  
When a name exceeds 30 characters, individual words are abbreviated, one at a time, until the length is reduced to 30 characters or less  
Attribute names start with lower case

## Element and Type Reuse



Elements are generally not reusable  
Types that are reusable across multiple schema files are defined in efileTypes.xsd. Reusable types are generally XML schema built-in types with additional length restrictions or character restrictions by regular expressions

## Amount Field Restrictions



Length restriction is typically set via facets such as `totalDigits` and/or `fractionDigits` on the field itself  
Range restriction facets such as `minInclusive` and `maxInclusive` may also apply

## ID Generation



An `IdType` is defined in `efileTypes.xsd` for the generation of Ids such as `TransmissionId`, `OriginId` (formerly `BatchId`), `ItemId`, `returnId`, `documentId`, etc.

A timestamp of `dateTime` type with up to 4-decimal fractional second may be used as an ID, but it must be unique within its parent element.

## Optional Field Attribute



A note attribute may have predefined phrase(s). It is equivalent to writing on the dotted line, when allowed.

7

## Fixed Document Attributes



Fixed attributes are for IRS internal use only. **DO NOT** include them in your XML data for transmission.

- ◆ `documentType` and `documentName` are used on a return document (a form, schedule, supporting info, election, etc.)
- ◆ `documentReference` is used on a field and serves as a pointer to another return document with a matching `documentName`

8





## Repeating Groups



A repeating group of elements (maxOccurs=“unbounded”) replaces paper attachments that are used merely to extend the limited space on the paper form

- ◆ e.g. a table showing only 4 rows but allows additional rows on paper attachments are replaced with a repeating group of the data elements.

## Return Documents



### Form and Schedules

- ◆ e.g. Form 1120, Form 4797, Schedule D, Itemized Other Income Schedule, ...

### Supporting Info

- ◆ e.g. Tax Shelter Disclosure, miscellaneous explanations, ...

### Election

- ◆ e.g. Securities or Commodities Held by a Trader - Mark-to-Market, ...

### Payment

## Transmission File



Based on SOAP 1.1  
with Attachments

SOAP envelope

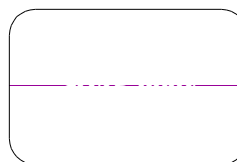
- ◆ SOAP Header
- ◆ SOAP Body

SOAP attachment

- ◆ Return document(s)

MIME boundaries

### SOAP Envelope



## Acknowledgement File

### Acknowledgement File



SOAP envelope with SOAP Body only (i.e. no SOAP header)

17

## Acknowledgement Envelope

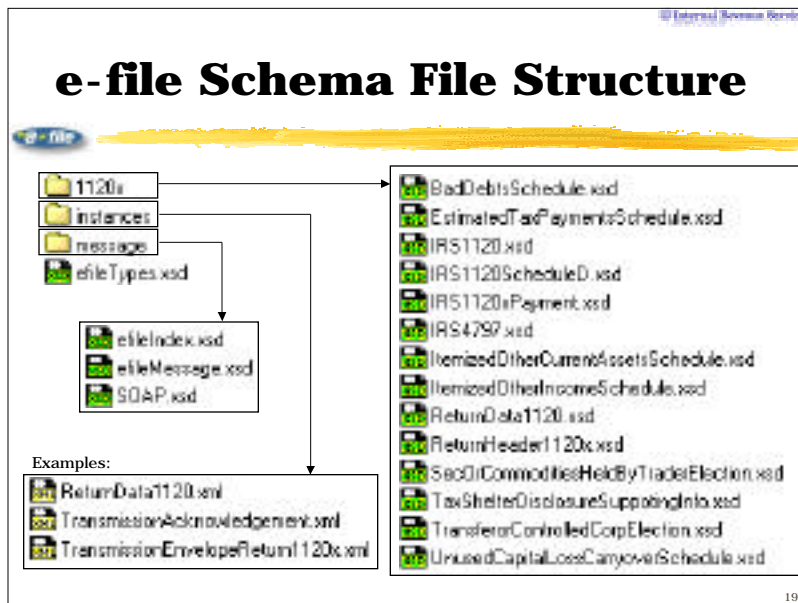
### SOAP Envelope



SOAP Body / AcknowledgementAction / Acknowledgement

- ◆ Transmission Acknowledgement
- Return Origin Acknowledgement(s)
  - Return (Item) Acknowledgement(s)

18



## efileTypes Schema

efileTypes.xsd defines data structures for:

- ◆ Reusable data types such as EIN, SSN, etc.
- ◆ Required attributes for each return and each return document

## **efileMessage Schema**



efileMessage.xsd defines data structures for:

- ◆ Transmission Header
- ◆ PIN Registration Origin (formerly as Batch Headers) Header
- ◆ Return Origin Header
- ◆ Base Return Header - return header elements common across families
- ◆ Acknowledgement

## **efileIndex Schema**



efileIndex.xsd includes schemas for efileTypes, efileMessage and the return headers

Used for validating XML data for the transmission envelope

## SOAP Schema

The SOAP is used as is but a slightly change was made to work around a validation bug in XML Spy during development

- ◆ all processContents attribute values were changed from “lax” to “strict”

23

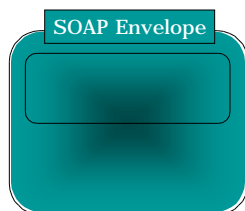
## 1120x Return Schemas

ReturnHeader1120x.xsd contains data elements that apply to the entire 1120x return

ReturnData1120.xsd specifies what and how many return documents can be included in the return  
Schemas for forms, schedules, supporting info, elections and payment

24

## Transmission File and Schema Files



## Acknowledgement File and Schema Files



All data elements for the acknowledgement file are defined in efileMessage.xsd

## **Sample XML Instances**



Sample XML data instances are included in the instances folder  
Use for reference and illustration of data structure only  
Real-world data are not used