



# *XML, the SEMI XML Study Task Force, and Possible Applicability to eDiagnostics*

**Chris Brandson**  
**SEMI XML Study Task Force Co-Chair**





## XML Overview

**Why XML Will Be Successful**

**XML HTML and SGML**

**The Need for Schemas**

**XML Schema Datatypes and Structures**

**XML Namespaces**

**XML Study Task Force**

**eDiagnostics and XML**

**References and Further Investigation**



# Why XML Will Be Successful

## Pervasiveness

- Developed by W3 (WWW), IETF (brought us the internet)
- Backed by Microsoft, IBM, Oracle, etc.
- Supported by IE5
- Driven by eCommerce: BizTalk, etc.
- Freely Available Information and Utilities: SAX, DOM, etc.

## Flexibility

- By making tags (elements) MetaData can apply meaning to data
- XML is expressive enough to express: A Graphic Image, A Word Document, a web page, a Chemical Molecule
- Applications communicate using a shared “Vocabulary”
- Filters can adapt to a host of applications – apply new meaning by translating between vocabularies
- Easy to bridges to Proprietary, Legacy, Specific Applications
- XML opens the door to a host of middleware possibilities





# XML, HTML and SGML

## HTML

```
<tr valign=top>
<td align=right valign=top>
<font size="+3">
```

HTML Markup is about layout!

## XML

```
<?xml version = "1.0">
<!DOCTYPE SemiEquipMessage "http://10.12.207.12/E9999.dtd">
<!ELEMENT DINAME (#PCDATA)>
<SemiEquipMessage>
<MESSAGENAME> EVENT </MESSAGENAME>
<DATAREPORT>
  <DINAME>PROCESSTIME</DINAME><DIVALUE FORMAT=TIME>12:23</DIVALUE>
  <DINAME>THICKNESS</DINAME><DIVALUE FORMAT=A>123.4</DIVALUE>
</DATAREPORT>
</SemiEquipMessage>
```

XML Markup is about MetaData!

**XML (HTML also) is derived from SGML, but is vastly less complex. The XML spec is only 36 pages long, yet is very powerful.**





## The Need for Schemas

- Everything in XML is PCDATA or CDATA and it all gets converted to CDATA or NDATA
  - ◆ i.e. it is all just "strings" or external data
- Need to define types which are not so open
- A "schema" is a set of type constraining rules - DTDs are schemas!
- But there are generic types, naturally many groups have attempted to develop generic schemas:
 

◆ XML Data	XDR	DCD
◆ SOX	DDML	WDDI

  - Some were even W3 Working Drafts!
- W3C Decided to Develop "XML Schema" for general use
- Why use "XML Schemas" rather than others:
  - ◆ Supported by W3C, future XML will be based on it,
  - ◆ Open, not proprietary
  - ◆ Will be approved W3C Recommendation



# XML Schema Datatypes & Structures

- Schemas add data types:

## Primitive

string (CDATA)

boolean

number

dateTime (ISO8601)

binary

URI

## Generated

integer

decimal

real

date

time

timeperiod

## User Generated

```
<datatype name = "monthday">
```

```
<basetype name="integer"/>
```

```
<minInclusive> 1 </minInclusive>
```

```
<maxInclusive> 31 </maxInclusive>
```

```
</datatype>
```

- It is possible to build a rich set of datatypes from this basis

- Structures can be defined and a content model defined

- ◆ Elements/Attributes can have sequence, choice, mixed, any, all, empty

- These become archetypes which can be refined

- ◆ This is very powerful because extensions can be made to encompass new features and functionality without violating the content rule of the original schema.

- ◆ But care must be taken in the construction of the original schema and the extensions!



## XML Namespaces

- Somewhat problematic area, none the less appear to be here to stay
  - ◆ Official W3C Recommendation
  - ◆ Widely Supported (Microsoft Office 2000)
- Uses the ":", which XML 1.0 spec reserves but does not constrain

```
<music:music xmlns:music="http://www.music.org/music">  
<music:note type="half" freq="A440" instrument="trumpet" volume="loud"/>  
</music:music>
```
- XML does not "understand" the ":", they just get mapped.
- Reservation of namespace relies on:
  - ◆ "good faith" of URI /URL owner (www.music.org) controls "music"
  - ◆ the author's discipline, the URI is never checked
- Note: When Schemas are finalized this may well be addressed



## XML Overview

### XML Study Task Force

**SEMI, XML and XML Study Task Force History**

**Task Force Charter**

**Task Force Scope**

**Initial Timeline**

**Reference Implementation Model**

**Data Formats, Reference Data & Implementors**

**Proposed Architecture**

**eDiagnostics and XML**

**References and Further Investigation**



# SEMI, XML and XML Study Task Force History

## • SEMI XML Activities

- ◆ Ballot 3157 Wafer Map Data
- ◆ Semi Japan Task Force
- ◆ SEED E36
- ◆ XML Study Task Force

## • XML Study Task Force History

- ◆ Interest in XML remained after Equipment Data Collection Task Force was sunset
- ◆ Formed at Winter Standards Meetings La Jolla CA, Feb. 2000
  - Established Charter and Scope
- ◆ 1<sup>st</sup> Full Meeting at SEMICON/West
  - Established Timeline, Formats, Reference Data Sets and Reference Implementors
  - Plan XML Step at SEMICON/Southwest
  - Separate Events Focus Group Formed
  - Selected XML Schemas as methodology
  - Launched web site at [www.semiml.org](http://www.semiml.org) (SemiML)
- ◆ Event Focus Group
  - After much discussion decided to explore re-using ERS E53 as baseline
- ◆ Next General Telecon Scheduled for Friday, 29<sup>th</sup> September
  - Contact Chris Brandson ([chris.brandson@brooks.com](mailto:chris.brandson@brooks.com)) to confirm





## XML Study Task Force - Charter

Conduct a feasibility study of where and why XML could be used in the Semiconductor Industry and depending on the conclusions reached possibly:

- Prepare a guideline document and/or prototypes and examples
- Review the current XML usage in the Semiconductor Industry; usage in other industries; and, where possible and appropriate, make use of these existing bodies of work
- Prepare a roadmap for adoption and/or a migration path for the use of XML in the industry



## XML Study Task Force - Scope

### The Task Force WILL Examine:

- Issues of data transmission, storage and presentation (Machine to Machine, and User Interface)
- XML specification techniques
- What are the appropriate domains for XML application
- Preparing guidance for XML usage
- DTDs/Schemas specific to the semiconductor industry

Specifically excluded from the scope of the task force, initially is:

- Complete replacement or complete mapping of SEMI E5 to XML
- New XML DTDs/Schema Languages
- Data Security Issues
- Transport Layers
- Quality of Service Issues



## XML Study Task Force - Initial Timeline

### Phase I

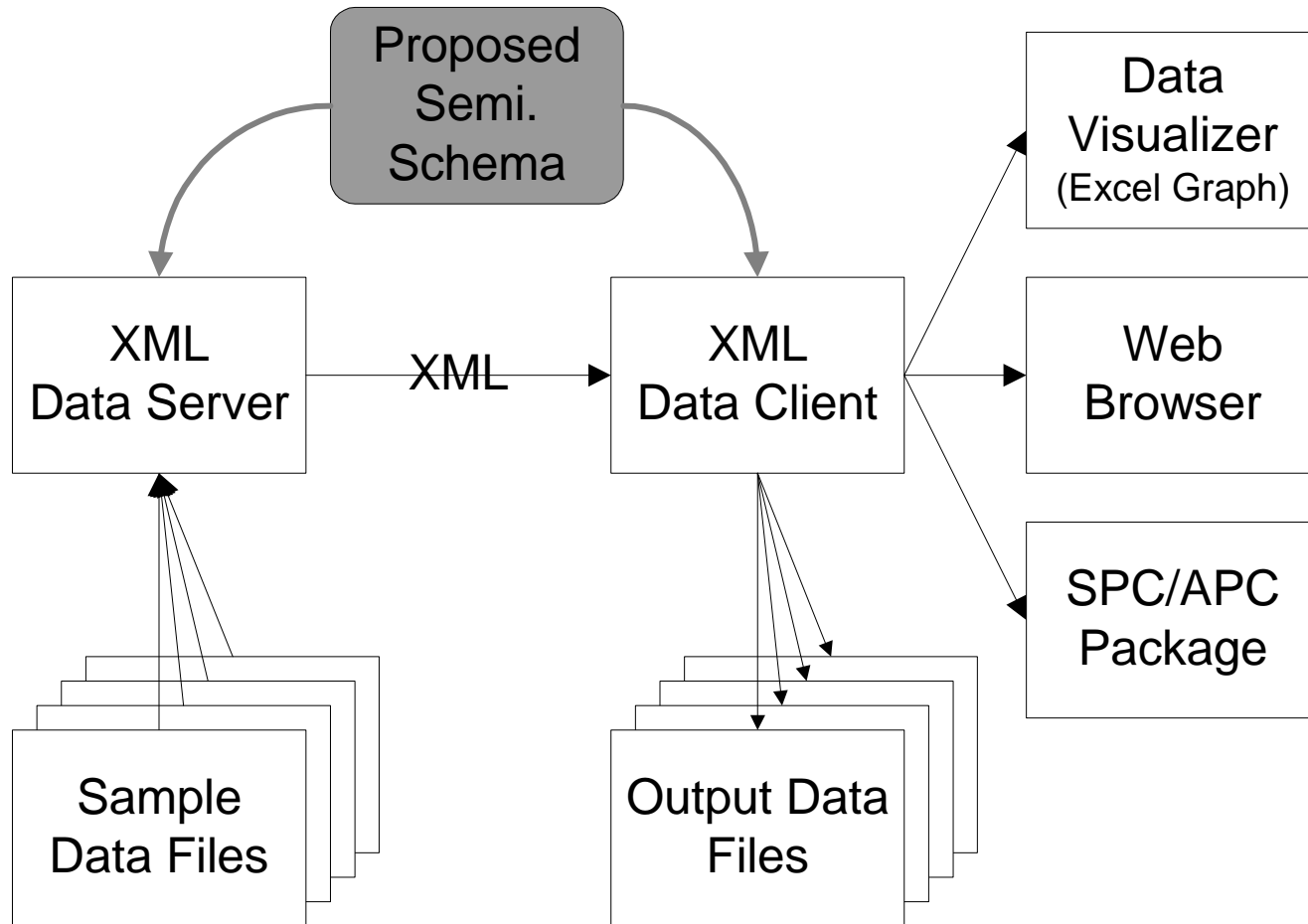
- Define Only Data Types
- Assume connection already established
- Work on Data Stream from Tool
- Do not include any control aspects
- Prepare Reference Implementation
- Prepare Ballot of Data Types

### Phase II

- Add Data Stream Control (Subscribe, Start Stop)
- Do not include any control aspects
- Prepare Ballot of On Stream Control
- Reference Implementation



# Reference Implementation Model





# Data Formats, Reference Data and Implementors

## Data Formats

- Timestamp
- Name, Value Pairs
- Spectral Data through Time (RGA, OES, etc) n-dim
- Commercial Format to be replaced with standard formats:
  - ◆ metrology data, yield data, netlists, lot-number-wafer-number lists
- Wafer Map Data (Task Force)
- Video Images
- Recipe (may have issues here)
- OEE Data

## Reference Data

- OES Data, Trace Data Files      Metrology/Stepper Data
- MFC Data

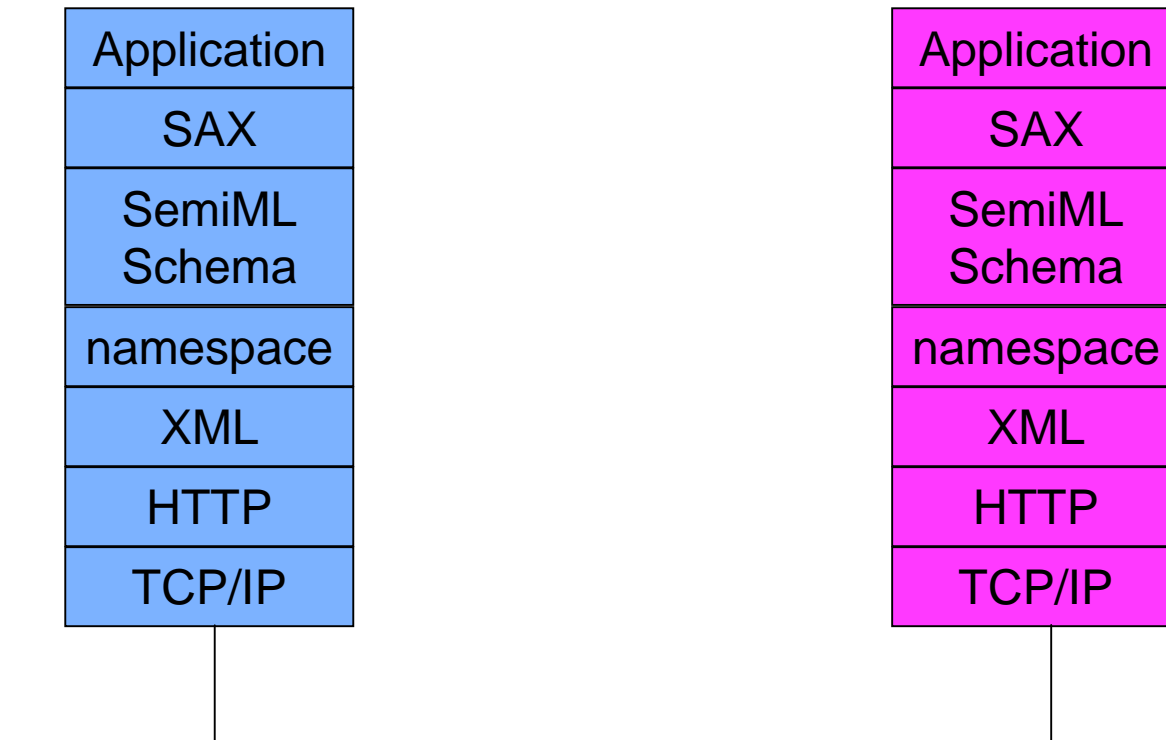
## Reference Implementors:

- Brooks Automation      Avantcom      IBM
- SDI      IPC





# Proposed Architecture





## XML Overview

## XML Study Task Force

## eDiagnostics and XML

eDiagnostics and XML Task Force Activities

eDiagnostics and XML Task Force Next Steps

References and Further Investigation



## eDiagnostics and XML Task Force Activities

- The general requirement of tools collecting (storing) and sending data is a natural application for XML
- The purposes to which the data is put is closely aligned with the initial area of focus of the XML Study Task Force
- XML is:
  - ◆Widely Available
  - ◆Cross Platform
  - ◆Robust Against Change (if designed carefully)
- Transformations and Filters are a basic feature of XML
- XML does not dictate internal implementation and design
  - ◆very usable in an "external only" configuration



## eDiagnostics and XML Study TF Next Steps

- May need to re-organize XML Study Task Force
- Remove "Study" from Task Force Name
- Officially Break into Focus Areas
- Need to examine in detail if there is a good match
  - ◆ the task force needs to review this
- At this point the potential for synergy looks excellent
- eDiagnostics may be the driver that says it is time for XML in SEMI Stds.
- And the Task Force can move on to developing the XML standards



## XML Overview

## XML Study Task Force

## eDiagnostics and XML

## References and Further Investigation



## References and Further Investigations

The Charles F. Goldfarb Series of XML Books by Prentice Hall [www.phptr.com](http://www.phptr.com)

A list which claims to list all XML books in print is at [www.xmlbooks.com](http://www.xmlbooks.com)

An excellent tutorial on Schemas by Roger L. Costello is freely available

And there is some good material at [www.xfront.com](http://www.xfront.com)

The standards themselves are located at [www.w3.org](http://www.w3.org)

Some specific ones are:

XML 1.0 (REC-xml-19980210)

XML Namespaces (REC-xml-names-19990114)

[www.w3.org/TR/xmlschema-0/](http://www.w3.org/TR/xmlschema-0/) (Primer)

[www.w3.org/TR/xmlschema-1/](http://www.w3.org/TR/xmlschema-1/) (Structures)

[www.w3.org/TR/xmlschema-2/](http://www.w3.org/TR/xmlschema-2/) (Datatypes)

The background document for Schemas (W3C Note) is very good.

The annotated XML spec (see Goldfarb) is easier reading than the spec.

Then of course there's the XML Study Task Force site at [www.semiml.org](http://www.semiml.org)

