SEMI E36-95
SPECIFICATION FOR ELECTRONIC DOCUMENTATION INTERCHANGE

NOTE: The SEED DTD V1.0 may be freely used by SEMI-members and other parties in the semiconductor industry without permission or payment of royalties or other charge. Use of the SEED DTD by any party outside of the semiconductor industry requires a separate license from Infrastructures for Information.

1 Purpose

The purpose of this standard is to establish how documents shall be exchanged electronically between two or more entities.

2 Scope

The scope of this standard is to define the process of exchanging documents electronically, the services required to support a document exchange session, and how to validate an exchange session.

This standard also addresses methodologies for creation of documents using the underlying data structures required to support the electronic document exchange process.

3 Definitions

3.1 *Database* — A database is a system that provides data management services to the user. Those services include:

- data storage;
- data retrieval;
- security;
- data integrity.

3.2 *Communications System* — A communications system is the methods and technologies used to permit and support a document exchange session. This standard assumes a full functioning communications system between entities capable of delivering application messages. In the OSI model, this standard is concerned only with the application layer.

3.3 *Conforming System* — A system that provides all the services defined in this standard. A conforming system may be made up of numerous dispersed application components. A conforming system must be able to exchange conforming documents with other conforming systems.

3.4 *Conforming Document* — A conforming document is one that meets the structural rules defined in a conforming system for the class of document. A conforming document created by one conforming system has to be able to be exchanged with another conforming system. If necessary, a conforming system has to be able to deliver to the other conforming system the structural rules for the document class of the document being exchanged.

3.5 *Conforming Information Component* — A conforming information component is one that meets the structural rules defined in the conforming system for the class of information component. A conforming information component created by one conforming system has to be able to be exchanged with another conforming system.

3.6 *DTD* — Document Type Definition. As specified in ISO 8879-"Rules, determined by an application, that apply SGML to the markup of documents of a particular type. A document type definition includes a formal specification, expressed in a document type declaration, of the element types, element relationships and attributes and references that can be represented by markup. It thereby defines the vocabulary of the markup for which SGML defines the syntax." ISO 8879 4.105.

3.7 *Document* — A document is a logical assembly of information components, that, if correctly ordered, is a means of transmitting an internally consistent piece of information. A document is made up of one or more information components. A document may be an information component.

3.8 *Document Class* — A document class is defined by a structural model that governs its existence in the conforming system.

3.9 *Document Exchange* — Document exchange means:

- the passing of a document or one or more of its components from one conforming system to one or more other systems;
- the optional return of that document or one or more of its components or annotation to that document from one or more receiving locations to the origination system;

3.10 *Document Exchange Session* — A document exchange session is the process that contains the document exchange.

SUPERSEDED

AVA Documents
Document Sharing Hub
Full document details: SEMI E36-0095

3.11 *Document Presentation* — Document presentation is the process and means by which the document shall be made visually or otherwise understandable to the end user. The presentation must be consistent with the intent of the originating system.

Document presentation must be able to enforce security and access privilege of the information components of the presented document.

Document presentation requires that:

- each title be identifiable by the application to the text of the title of a font or format different from the following text;
- each paragraph be identifiable by the application to the text of the paragraph of a font or format different from the following text;
- each section be identifiable by the assignment to that section of a sequential number or other identifier;
- each sub-section be identifiable by the assignment to that sub-section of a sequential number or other identifier that is preceded by the identifier of the section or sub-section that the sub-section is in;
- each list be formatted as the type of list specified in the list definition;
- each table be formatted, at a minimum, as a grid, with each cell in the grid clearly recognizable;
- each emphasis text object be formatted as specified in the emphasis definition;
- each graphic, table, formula, must be placed in the correct sequential position as defined in the text flow;
- each revision be identifiable by the application to the text of the revision of a font or format different from the preceding and following text. Each revision level must be separately identifiable;
- each soft-cross reference must automatically publish, at a minimum, the identifier to the sub-section level of the referenced point;

3.12 *Information Component* — An information component is:

- a uniquely addressable unit of information within the domain;
- an identifier to an addressable unit of information outside the domain.

For the purpose of this standard, the term document may be used to represent an information component.

3.13 *Non-Conforming Document* — A non-conforming document is one that does not meet the structural rules defined in the conforming system. A conforming system has to be able to process non-conforming documents.

3.14 *Non-Conforming System* — A system that does not provide all of the services defined in this standard. A non-conforming system may interact with a conforming system.

3.15 *Presentation Document* — A presentation document is a document that has gone through the document presentation process.

3.16 *SEED* — Semiconductor Electronic Equipment Documentation.

3.17 *SGML* — Standard Generalized Markup Language. The Standard Generalized Markup Language is an International Standards Organization "ISO" standard. It is ISO 8879. The ISO Reference Number is ISO 8879-1986-(E).

3.18 *Security Profile* — The security profile defines the access privileges of a participating system. Security profile implementation is specified by the participating systems.

3.19 *Structural Rules* — The structural rules for a document class is the SGML DTD (see ISO 8879) that defines the rules of encoding for a document.

3.20 *Version* — A version of a document in a state wherein the document has completed a formal approval process. The process must be auditable and available for review as part of the document history.

4 Document Exchange Methodology

A document exchange session may occur under two conditions. It may be a response to a request for a document from a conforming system or it may be a submission of document.

Part 1 - Request for a Document

- a session is initiated by a request to access the conforming system;

SUPERSEDED

Documents
Document Sharing Hub
Full document details: SEMI E36-0095

- the conforming system accepts or rejects the request to access based on the security profile of the requestor;
- if accepted the conforming system permits the user to request a document. A request may include one or more selection criteria joined by Boolean operators.
- if the conforming system finds one or more documents that match the selection criteria a profile of each document is returned to the requestor for selection;
- a request for one of the profiled documents is submitted;
- if the session is within a conforming system or between conforming systems the selected document may be delivered as a presentation document or as a conforming document;
- if the session is between a conforming system and a non-conforming system the selected document may be delivered from the conforming system as a conforming document or as a non-conforming document;

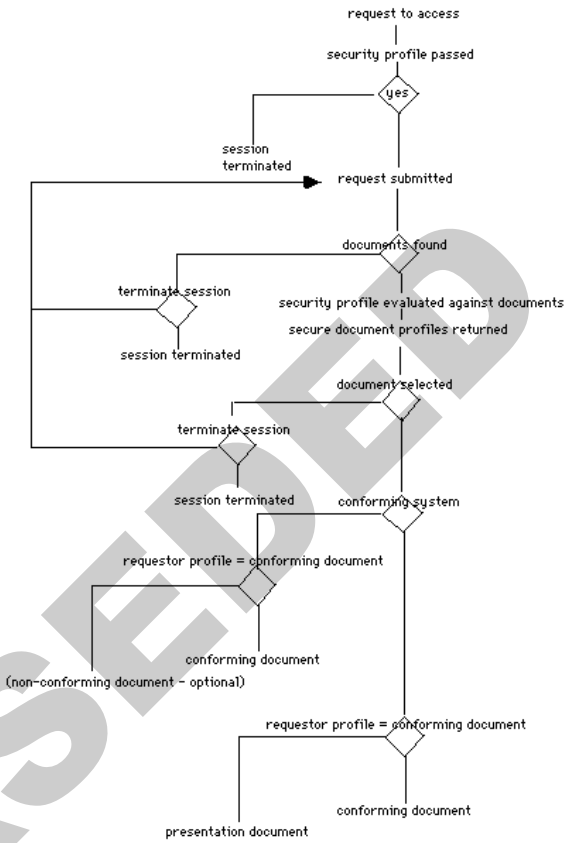

Figure 1

Document Exchange Session Part 1

Part 2 - Submission of a Document

- the requestor may submit conforming annotation components about information components of selected document to the delivering conforming systems. Information components about information components become part of the history of the target information component and must be able to travel with the component. Information components about information components can be submitted in any order at any time;
- optionally receipt information components about information components must be acknowledged;

SUPERSEDED

AVA Documents
Document Sharing Hub
Full document details: SEMI E36-0095

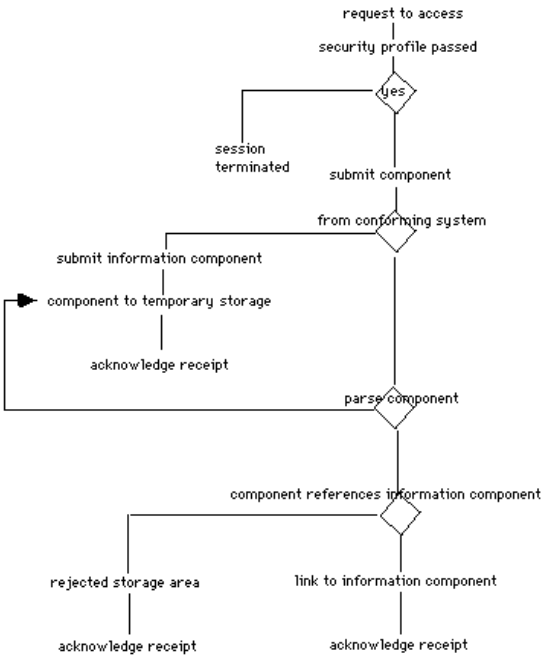

Figure 2

Document Exchange Session Part 2

5 Provided Services

A conforming system shall provide the following services:

5.1 *Conforming Documents and Non-Conforming Documents*

- the ability to create and manipulate conforming documents;
- the ability to create and manipulate one or more classes of conforming documents;
- the ability to import conforming documents from other systems into the conforming system;
- the ability to import non-conforming documents into the conforming system;
- the ability to turn conforming documents into non-conforming documents;

5.2 *Presentation Documents*

- the ability to convert conforming documents into presentation documents.

5.3 *Search and Retrieval*

- the ability to search for a conforming document by name, date of creation, creator, date last edited, version, document control number, document class, controlled keywords;
- the ability to join search criteria with Boolean operators;
- the ability to provide profiles of documents returned from the search to assist in making a final selection;
- the ability to work, within a session, with each document in the returned selection without having to resubmit the search;

5.4 *Information Components*

- the ability to uniquely identify and track information components;
- the ability to convert information components into presentation documents;
- the ability to use information components as identifiers of objects not within the conforming system;
- the ability to convert conforming information components to non-conforming information components in a manner that minimizes information loss. Structural information held in the conforming information component does not have to be retained with the exception of tabular material which must be converted to common (i.e., tab-delimited) tabular form that can be manipulated electronically;
- the ability to convert non-conforming components/documents into conforming components/documents. This may be done by loading the document into a local word processor and manually applying conforming SGML encoding to the document, or processing the document through a filter to apply conforming SGML encoding, or any combination of the two.
- components may include text, tables, mathematical equations, and graphics.

6 Conforming Structure for a SEED Document

SEED DTD V1.0

©copyright Infrastructures for Information Inc.

Version 1.0.1 December 1993

Version 1.0 November 1993

SUPERSEDED

Documents
Document Sharing Hub
Full document details: SEMI E36-0095

NOTE: The SEED DTD V1.0 may be freely used by SEMI-members and other parties in the semiconductor industry without permission or payment of royalties or other charge. Use of the SEED DTD by any party outside of the semiconductor industry requires a separate license from Infrastructures for Information.

This document describes the SGML DTD for the SEED Initiative. It is the data dictionary for the DTD known as <!DOCTYPE SEEDOC PUBLIC"-//SEEDOC//SEED//EN">.

For each element, the following is identified:

- element name
- element purpose - what data object or structural object does this element identify
- the content model for the elements
- the attributes of the element
- a discussion of where this element may appear or other information about the element that is required to understand it.

Conventions followed: all elements, attributes, and SGML key words are, where appropriate, bolded.

General notes on the DTD.

The SEED DTD is designed to provide the maximum flexibility possible within a single data model capable of supporting both composition and data managment. The objective of the DTD is to:

- allow the identification of structural objects such as paragraphs, lists, graphics, sections, and sub-sections as needed;
- allow the identification of objects as needed for data management purposes.

SUPERSEDED

Documents
Document Sharing Hub
Full document details: SEMI E36-0095

7 The SGML Elements

Tag Name:	AUTHOR
Tag Purpose:	To identify an author in a citation.
Tag Content Model:	pcdata.
Tag Attributes:	type - type of author, editor, personal, corporate, joint, etc.
	key - professonial orientation
Discussion:	This may be used in the citation.

Tag Name:	APPROVED
Tag Purpose:	To identify the version level of the instance. Everything within the content model has been approved.
Tag Content Model:	designat followed by docno followed by repeatable sections
Tag Attributes:	source -an authorship or other identifier
	security - options of UC (unclassified) or C (classified)
	restrict - numeric tokens to indicate the restriction level - takes precedence of security
	release - numeric tokens to indicate the release level of the instance
	level - numeric tokens to indicate the classification level of the instance
	custom - if a control object who it is controlled for
	start - the effective start date of this content model
	end - the effective end date of this content model - these values may be overridden by start and end dates within the content model that fall within the range established by these dates
	item - if this content model has been replaced the id of the item model that replaced it.
Discussion:	Instances go through versions ie: legacy, draft, approved. Each version is part of the permanent history of the instance. If a SEEDDOC is approved then everything within it is approved.
	If the document is not approved ie: legacy or draft individual content models within the SEEDOC can be approved.

Tag Name:	CITATION
Tag Purpose:	To identify a citation.
Tag Content Model:	an optional no, repeatable author or title, optional pub, vol, ed, date, repeatable page.
Tag Attributes:	none
Discussion:	Used for bibliographic citations.

Tag Name:	CONTROL
Tag Purpose:	To place an object that does not have security control under security control
Tag Content Model:	ANY
Tag Attributes:	id - so it can be found
	security - options of UC (unclassified) or C (classified)
	restrict - numeric tokens to indicate the restriction level - takes precedence of security
	release - numeric tokens to indicate the release level of the instance
	level - numeric tokens to indicate the classification level of the instance
	custom - if a control object who it is controlled for
Discussion:	This is used to place any object under security control. As most objects have security control, it is most appropriate for in-line objects such as special paragraphs, titles, keywords.

SUPERSEDED

Documents
Document Sharing Hub
Full document details: SEMI E36-0095

Tag Name:	DESIGNAT
Tag Purpose:	To permit the assignment of a designation to the SEEDOC.
Tag Content Model:	pcdata
Tag Attributes:	none
Discussion:	This is used to permit the assignment of a descriptive designation. This might be a structured character sequence identifying the source of the document, the date, or any other important identifying information. Note: This should not be a cont

Tag Name:	DRAFT
Tag Purpose:	To identify the version level of the instance.
Tag Content Model:	designat followed by docno followed by repeatable sections
Tag Attributes:	source -an authorship or other identifier
	security - options of UC (unclassified) or C (classified)
	restrict - numeric tokens to indicate the restriction level - takes precedence of security
	release - numeric tokens to indicate the release level of the instance
	level - numeric tokens to indicate the classification level of the instance
	custom - if a control object who it is controlled for
	start - the effective start date of this content model
	end - the effective end date of this content model - these values may be overridden by start and end dates within the content model that fall within the range established by these dates
	item - if this content model has been replaced the id of the item model that replaced it.
Discussion:	Instances go through versions ie: legacy, draft, approved. Each version is part of the permanent history of the instance.
	The content model is in draft form. Individual content models within it may be approved.

Tag Name:	ED
Tag Purpose:	To identify an edition in a citation.
Tag Content Model:	pcdata.
Tag Attributes:	none
Discussion:	To identify the edition of a source noted in a citation.

Tag Name:	EMPH
Tag Purpose:	To identify emphasis.
Tag Content Model:	ANY
Tag Attributes:	Derived from the CALS 38784B DTD. EMPH is a global include.

Tag Name:	FOOTNOTE
Tag Purpose:	To identify a footnote.
Tag Content Model:	Repeatable Para or Table or List or a Citation
Tag Attributes:	id to uniquely identify it
Discussion:	A footnote should have a xftnote which identifies, using an IDREF, where in the text the footnote refers to. This is because the structure of this DTD requires that footnotes be at the end of the section content model.

SUPERSEDED

AVA Documents
Document Sharing Hub
Full document details: SEMI E36-0095

Tag Name:	FORMULA
Tag Purpose:	To identify a formula.
Tag Content Model:	optional title, optional rc
Tag Attributes:	id to uniquely identify it
	name - this is an entity type attribute thus the name of the formula is a file entity
	data - the notation - see notation listings
Discussion:	FORMULA is a global include to use where the formula is done as notation.

Tag Name:	GRAPHIC
Tag Purpose:	To identify that a graphic that is anchored at the location of the tag.
Tag Content Model:	optional title, optional rc
Tag Attributes:	id to uniquely identify it
	name - this is an entity type attribute thus the name of the graphic is a file entity
	type - what type of graphic - chemistry, table, math, chart, picture, graphic, other.
	width - numeric for how wide the graphic is to be
	depth - numeric for how deep the graphic is to be
	hscale -numeric for horizontal scaling
	vscale - numeric for vertical scaling
	text - misc. text - usually the caption
	rotation - numeric for rotation of the graphic
	fig - misc. text - usually the figure number
	status - the status of this graphic within the version - approved, draft or review.
	security - options of UC (unclassified) or C (classified)
	level - numeric tokens to indicate the classification level of the instance
	restrict - numeric tokens to indicate the restriction level - takes precedence of security
	release - numeric tokens to indicate the release level of the instance
	custom - if a control object who it is controlled for
	start - the effective start date of this content model
	end - the effective end date of this content model - this values may be overridden by start and end dates within the content model that fall within the range established by these dates
	item - if this content model has been replaced the id of the item model that replaced it.
Discussion:	Derived from the CALS 38784B DTD. GRAPHIC is a global include. If the status attribute is set to approved, the section has been approved,
	if draft then this is a draft section, if review, then section is been polled for approval and should be treated as "read-only".

Tag Name:	ITEM
Tag Purpose:	To identify a list item.
Tag Content Model:	optional title followed by repeatable para or table or pcdata or list or graphic.
Tag Attributes:	idref - so it can point to something.
Discussion:	The tag can only appear in a list.

SUPERSEDED

Documents
Document Sharing Hub
Full document details: SEMI E36-0095

Tag Name:	KWORD
Tag Purpose:	To identify a keyword.
Tag Content Model:	pcdata
Tag Attributes:	id to uniquely identify it.
	type to identify the "type" of keyword; ie: place name, product name etc.
Discussion:	To be used to identify keywords for searching and for creation of vocabulary control and specialized indexes. The id allows the building of internal links using the xref idref attribute. KWORD is a global include.

Tag Name:	LEGACY
Tag Purpose:	To identify material that has been converted.
Tag Content Model:	designat followed by docno followed by repeatable sections
Tag Attributes:	source -an authorship or other identifier
	security - options of UC (unclassified) or C (classified)
	level - numeric tokens to indicate the classification level of the instance
	restrict - numeric tokens to indicate the restriction level - takes precedence of security
	release - numeric tokens to indicate the release level of the instance
	custom - if a control object who it is controlled for
	start - the effective start date of this content model
	end - the effective end date of this content model - these values may be overridden by start and end dates within the content model that fall within the range established by these dates
	item - if this content model has been replaced the id of the item model that replaced it.
Discussion:	Instances go through versions ie: legacy, draft, approved. Each version is part of the permanent history of the instance. Material within this content model is legacy material -
	material that was created in another form and converted to this form for development in this environment. This is different from third party material converted for use in this environment.

Tag Name:	LIST
Tag Purpose:	To identify a list.
Tag Content Model:	optional title followed by optional para or table or list and required repeatable item.
Tag Attributes:	id so if can be found
	type for the type of list - bullet, number, square, dash, indent, other, alphauc, alphalc, mono, none
	name - to identify the type of list such as TOC, glossary, index etc.
	status - the status of this list within the version - approved, draft or review.
	security - options of UC (unclassified) or C (classified)
	level - numeric tokens to indicate the classification level of the instance
	restrict - numeric tokens to indicate the restriction level - takes precedence of security
	release - numeric tokens to indicate the release level of the instance
	custom - if a control object who it is controlled for
	start - the effective start date of this content model
	end - the effective end date of this content model - this values may be overridden by start and end dates within the content model that fall within the range established by these dates
	item - if this content model has been replaced the id of the item model that replaced it.
Discussion:	The outermost element in a list.

SUPERSEDED

AVA Documents
Document Sharing Hub
Full document details: SEMI E36-0095

Tag Name:	NO
Tag Purpose:	To identify a book number in a citation.
Tag Content Model:	pcdata.
Tag Attributes:	none
Discussion:	none

Tag Name:	PAGE
Tag Purpose:	To identify the page numbers cited in a citation.
Tag Content Model:	pcdata.
Tag Attributes:	none
Discussion:	none

Tag Name:	PARA
Tag Purpose:	To identify paragraphs.
Tag Content Model:	pcdata.
Tag Attributes:	none
Discussion:	Paragraphs are basic structural objects.

Tag Name:	PUB
Tag Purpose:	To identify a publication in a citation.
Tag Content Model:	pcdata.
Tag Attributes:	none
Discussion:	none

Tag Name:	RC
Tag Purpose:	To identify raw character data
Tag Content Model:	rcdata.
Tag Attributes:	none
Discussion:	Used in the graphics and formula tags if necessary.

Tag Name:	REVISION
Tag Purpose:	To identify a revision in the material
Tag Content Model:	ANY
Tag Attributes:	who - who made the revision, this value should be the system identifier of the user who made the revision date - the date of the revision, this value should be automatically entered by the system level - the level of the revision, this can be managed by the system or set by the user. note - optional text to explain the revision.
Discussion:	There are two basic strategies for dealing with revisions. The first requires that new material is held within revision tags and the level is set to new, old material is held within revision tags and the level set to old. There can be as many "olds" as needed. The date stamp indicates the order in which revision were done. This strategy allows the system and the user to identify each revision.

SUPERSEDED

Documents
Document Sharing Hub
Full document details: SEMI E36-0095

	The second requires only that old material be held within revisions. New material is not tagged. This strategy reduces some of the tagging effort but fails to clearly identify new material. Manual revision tagging is used if only material revisions need to be identified. Auto tagging can be done if all revisions, regardless of importance, need to be tracked.

Tag Name:	SECTIONS
Tag Purpose:	To identify a section in a document.
Tag Content Model:	optional title, followed by repeatable para or table or list or repeatable subsect, followed by optional repeatable footnote
Tag Attributes:	id - so it can be found name - the name of the section order - a specific order for the section to appear in the SEEDOC status - the status of this section within the version - approved, draft or review. security - options of UC (unclassified) or C (classified) level - numeric tokens to indicate the classification level of the instance restrict - numeric tokens to indicate the restriction level - takes precedence of security release - numeric tokens to indicate the release level of the instance custom - if a control object who it is controlled for start - the effective start date of this content model end - the effective end date of this content model - this values may be overridden by start and end dates within the content model that fall within the range established by these dates item - if this content model has been replaced the id of the item model that replaced it.
Discussion:	A basic building block of a document. Sections have the option of being "named," that is they can be assigned a name such as Reference Documents, or Parts, or Warnings. The naming convention can be used to cause the section to be managed or presented specially. A named Warning section might have different compostion than a named Parts section.

Tag Name:	SEEDOC
Tag Purpose:	To identify a complete document conforming to the SEED DTD.
Tag Content Model:	legacy or draft or approved
Tag Attributes:	id - so it can be found security - options of UC (unclassified) or C (classified) level - numeric tokens to indicate the classification level of the instance restrict - numeric tokens to indicate the restriction level - takes precedence of security release - numeric tokens to indicate the release level of the instance custom - if a control object who it is controlled for start - the effective start date of this content model end - the effective end date of this content model - this values may be overridden by start and end dates within the content model that fall within the range established by these dates item - if this content model has been replaced, the id of the item model that replaced it.
Discussion:	none

SUPERSEDED

Documents
Document Sharing Hub
Full document details: SEMI E36-0095

Tag Name:	SUBSECT
Tag Purpose:	To identify a sub-section in a section.
Tag Content Model:	optional title, followed by repeatable para or table or list or repeatable subsect, followed by optional repeatable footnote
Tag Attributes:	id - so it can be found
	name - the name of the sub-section
	order - a specific order for the sub-section to appear in the section
	status - the status of this sub-section within the version - approved, draft or review.
	security - options of UC (unclassified) or C (classified)
	level - numeric tokens to indicate the classification level of the instance
	restrict - numeric tokens to indicate the restriction level - takes precedence of security
	release - numeric tokens to indicate the release level of the instance
	custom - if a control object who it is controlled for
	start - the effective start date of this content model
	end - the effective end date of this content model - these values may be overridden by start and end dates within the content model that fall within the range established by these dates
	item - if this content model has been replaced the id of the item model that replaced it.
Discussion:	A basic building block of a document. Sub-sections have the option of being "named", that is they can be assigned a name such as Parts, or Warnings.
	The naming convention can be used to cause the sub-section to be managed or presented specially. A named Warning section might have different compostion than a named Parts section.
	Sub-sections are recursive. If the composition system autonumbers, then sub-sections are used to assign numbering.

Tag Name:	TABLE
Tag Purpose:	To identify a table.
Tag Content Model:	optional tblhead followed by a tblbody followed by an optional tblfoot.
Tag Attributes:	id for unique ID
	status - the status of this table within the version - approved, draft or review.
	security - options of UC (unclassified) or C (classified)
	level - numeric tokens to indicate the classification level of the instance
	restrict - numeric tokens to indicate the restriction level - takes precedence of security
	release - numeric tokens to indicate the release level of the instance
	custom - if a control object who it is controlled for
	start - the effective start date of this content model
	end - the effective end date of this content model - this values may be overridden by start and end dates within the content model that fall within the range established by these dates
	item - if this content model has been replaced the id of the item model that replaced it.
Discussion:	The table is a basic structural element that appears through the entry.

SUPERSEDED

Documents
Document Sharing Hub
Full document details: SEMI E36-0095

Tag Name:	TBLBODY
Tag Purpose:	To identify the body of a table.
Tag Content Model:	a tblcols followed by repeatable tblrow
Tag Attributes:	none
Discussion:	The body of the table maps to a spreadsheet model. There is one cell for every position in the grid whether it is empty or not or overridden by another cell. A table body must be able to be mapped directly into industry-accepted means of exchanging table information (i.e., tab, tab return, or comma, comma, return delimited).

Tag Name:	TBLCELL
Tag Purpose:	To identify a table cell.
Tag Content Model:	Repeatable para or graphic or list or a table
Tag Attributes:	halign horizontal alignment valign vertical alignment char character alignment tsep top separator bsep bottom lsep left separator rsep right separator colstart column where a cell starts for spanners colend column where a cell ends for spanners rowstart row where a cell starts for spanners rowend row where a cell ends for spanners
Discussion:	Table cells can include text or graphics or formulas.

Tag Name:	TBLCOL
Tag Purpose:	To identify a table column.
Tag Content Model:	empty
Tag Attributes:	valign vertical alignment |char character alignment tsep top separator bsep bottom lsep left separator rsep right separator width width of the column units units of measure used in width
Discussion:	There is one tblcol for each column in the table.

Tag Name:	TBLCOLS
Tag Purpose:	To identify table columns.
Tag Content Model:	Repeatable tblcol
Tag Attributes:	none
Discussion:	Used to identify the column definition for the table. The number of columns is defined by counting the number of tblcol entries in this content model.

SUPERSEDED

AVA Documents
Document Sharing Hub
Full document details: SEMI E36-0095