



## IIT Madras BSc Degree

### Copyright and terms of use

**IIT Madras is the sole owner of the content available in this portal - [onlinedegree.iitm.ac.in](http://onlinedegree.iitm.ac.in) and the content is copyrighted to IIT Madras.**

- Learners may download copyrighted material for their use for the purpose of the online program only.
- Except as otherwise expressly permitted under copyright law, no use other than for the purpose of the online program is permitted.
- No copying, redistribution, retransmission, publication or exploitation, commercial or otherwise of material will be permitted without the express permission of IIT Madras.
- Learner acknowledges that he/she does not acquire any ownership rights by downloading copyrighted material.
- Learners may not modify, publish, transmit, participate in the transfer or sale, create derivative works, or in any way exploit, any of the content, in whole or in part.

# Beyond HTML

# HTML Evolution

# Markup languages

- Origins from late 60s
- Mostly used for typesetting and document management systems
- Problems?
  - Lack of standardization
  - Target audience: coders, publishers, academics?
  - Target output: print, other forms of media
  - Machine readability

# SGML

## Standard Generalized Markup Language

- Meant to be a base from which any markup language could be designed
- Basic postulates:
  - Declarative: Specify structure and attributes, not how to process
  - Rigorous: strict definition of structure, like databases
- DTD - Document Type Definition
  - Used to specify different families within this umbrella
  - Each could have its own tags, interpretation
- SGML *Applications*

# HTML

- Originally intended to be an *application* of SGML
- Very lenient with parsing - meant to be forgiving of errors
  - Not valid SGML
- HTML 2.0 - attempt to become SGML compliant
- Legacy support
  - Not truly SGML compliant
- HTML4 official definition - true SGML application
  - Limited usage
- HTML5 - **not** an SGML application - defines its own parsing rules

# XML

## eXtensible Markup Language

- Based on SGML
- Custom tags - multiple *applications* defined
- Focus on simplicity, generality and usability
- Both human-readable and machine-readable
- Well structured: can be used to represent complex data relationships, data structures etc.
- Examples:
  - MathML, RSS, Atom, SVG

# XML Example - RSS feeds

```
<?xml version="1.0" encoding="UTF-8" ?>
<rss version="2.0">
<channel>
  <title>RSS Title</title>
  <description>This is an example of an RSS feed</description>
  <link>http://www.example.com/main.htm</link>
  <copyright>2020 Example.com All rights reserved</copyright>
  <lastBuildDate>Mon, 06 Sep 2010 00:01:00 +000</lastBuildDate>
  <pubDate>Sun, 06 Sep 2009 16:20:00 +000</pubDate>
  <ttl>1800</ttl>

  <item>
    <title>Example entry</title>
    <description>Here is some text containing an interesting description</description>
    <link>http://www.example.com/blog/post/</link>
    <guid isPermaLink="false">7bd204c6-1655-4c27-aeee-53f933c5395</guid>
    <pubDate>Sun, 06 Sep 2009 16:20:00 +000</pubDate>
  </item>

</channel>
</rss>
```



# XML Example: SVG

```
<svg version="1.1"
      width="300" height="200"
      xmlns="http://www.w3.org/2000/svg">
  <rect width="100%" height="100%" fill="red" />
  <circle cx="150" cy="100" r="80" fill="green" />
  <text x="150" y="125" font-size="60"
        text-anchor="middle" fill="white">SVG</text>
</svg>
```



# XHTML

- Based on XML - not directly SGML
- Reformulation of HTML4 as applications of XML
- Main goal: clean up HTML specification
  - Modular and more extensible
- XML Namespaces: allow inter-operability with other XML applications

# HTML5

- Add support for latest features (multimedia support, canvases, ...)
- Remain easily readable and understandable to both human and machine
- Remain backward compatible
- Breaks away from SGML:
  - Not an SGML or XML application
  - Defines own parser

# HTML5

- Add support for latest features (multimedia support, canvases, ...)
- Remain easily readable and understandable to both human and machine
- Remain backward compatible
- Breaks away from SGML:
  - Not an SGML or XML application
  - Defines own parser

The **last** version of HTML!

- HTML Living Standard maintained by WHATWG (Web Hypertext Application Technology Working Group) - split away from W3C

## Extension?

- How to add new features? New tags?
- “Software defined”
  - Allow new tags to be added through JavaScript
  - Custom Elements - API supported by browsers
- Very powerful mechanism: arbitrary functionality possible
  - No new tags need to be brought into standard
- Potential problems:
  - Anyone can define a tag!?
  - Semantics (meaning) of tags may not be well thought out

Requirement? Javascript