

Written by Dr. Thapanapong Rukkanchanunt

02 Markup Languages 1

OUTLINE

- อะไรคือ Markup Languages
- Standard Generalized Markup Language (SGML)
- Hypertext Markup Language (HTML)
- Extensible Markup Language (XML)
- Extensible Hypertext Markup Language (XHTML)
- HTML 5.0, CSS and JavaScript

Markup Languages

- Markup Languages คือภาษาที่ใช้กำหนดรูปแบบของเอกสาร โดยจะถูกแยกส่วนออกมาจากเนื้อความอย่างชัดเจน
- โดยส่วนมากแล้วจะใช้ Tag เป็นตัวกำหนดรูปแบบ
- Application เช่น Microsoft Word ซ่อน Markup Languages ไว้เพื่อให้ง่ายต่อการใช้งานแก่บุคคลทั่วไป
- การสร้าง Web Application จำเป็นต้องใช้ความรู้ Markup Languages เนื่องจากเราไม่สามารถเปิดเว็บด้วย Application เพียงอย่างเดียว (เปิดเว็บใน Chrome, Firefox, Safari)

Standard Generalized Markup Language (SGML)

- SGML เป็นภาษาแม่ของ Markup Language ทุกชนิด (หรือเรียกว่าเป็น metalanguage)
- ถูกคิดค้นโดยบริษัท IBM และกลายเป็นภาษามาตรฐานในปี 1986
- SGML กำหนดให้ผู้ใช้สร้างและนิยามกฎที่จะใช้ในเอกสาร กฎเหล่านี้จะถูกเรียกว่า Document Type Definition (DTD) โดยจะเป็นไฟล์ที่แยกออกมาต่างหาก
- เราจำเป็นต้องอ้างอิงไฟล์ DTD ในเอกสารทุกครั้งที่เราต้องการใช้งาน

Hypertext Markup Language (HTML)

- Tim Berners-Lee จาก MIT เป็นคนคิดค้นภาษา HTML เนื่องจาก SGML มีความซับซ้อนมากและไม่สามารถใช้ได้กับอุปกรณ์ทุกชนิด
- ใน HTML เอกสารแต่ละชิ้นสามารถเชื่อมโยงหากันได้ด้วย Pointer ที่มีชื่อเรียกเฉพาะว่า Hyperlink
- Hyperlink เป็นคำสั่งที่จะเชื่อมโยงไปยังอีกจุดหนึ่งของไฟล์หรือไฟล์อื่น ๆ เมื่อมีการเข้าถึง link นั้น (ส่วนมากมาจากการคลิกเมาส์)
- World Wide Web ประกอบไปด้วย Documents ที่มี Hyperlink เชื่อมโยงซึ่งกันและกัน

HTML History

Year	Version
1991	Tim Berners-Lee invented HTML
1995	HTML Working Group defined HTML 2.0
1997	W3C Recommendation: HTML 3.2
1999	W3C Recommendation: HTML 4.01
2000	W3C Recommendation: XHTML 1.0
2014	W3C Recommendation: HTML5
2016	W3C Candidate Recommendation: HTML 5.1
2017	W3C Recommendation: HTML5.1 2nd Edition
2017	W3C Recommendation: HTML5.2

HTML vs. SGML

- HTML และ SGML เหมือนกันตรงที่ทั้งสองภาษาทำให้การแลกเปลี่ยนข้อมูลเป็นไปได้ง่ายเมื่อมีไฟล์ DTD ที่ใช้ควบคุมรูปแบบของเอกสาร
- ใน HTML เราไม่สามารถสร้าง DTD แถมยังมีข้อบังคับน้อยกว่า SGML อีกด้วย
- HTML จึงกลายมาเป็นมาตรฐานในการส่งข้อมูลผ่านอินเทอร์เน็ตของเอกสารบนเว็บ
- เวอร์ชันล่าสุดของ HTML คือเวอร์ชัน 5.0 ซึ่งแตกต่างจาก SGML โดยสิ้นเชิง การเปลี่ยนแปลงนี้ทำให้โครงสร้างของ HTML ง่ายขึ้นและไม่จำเป็นต้องใช้ DTD อีกต่อไป

DTD for HTML 4.0

- <https://www.w3.org/TR/WD-html40-970708/sgml/dtd.html>

HTML Example

```
3 <div *ngFor="let product of products">
4   <h3>
5     <a [title]="product.name + ' details'">
6       {{ product.name }}
7     </a>
8   </h3>
9
10  <p *ngIf="product.description">
11    Description: {{ product.description }}
12  </p>
13
14  <button (click)="share()">
15    Share
16  </button>
17 </div>
```

- HTML ประกอบไปด้วย Tag (สีน้ำเงิน) โดยจำนวน Tag ที่สามารถใช้ได้ถูกกำหนดไว้แล้ว โดยส่วนมาก Tag จะต้องกำหนดเป็นคู่คือ Tag ปิดและ Tag เปิด
- ภายใน Tag จะสามารถกำหนด Attribute (สีฟ้า) เพื่อเพิ่มคุณสมบัติของ Tag นั้น

ข้อดีของ HTML5

- มีการเพิ่ม <video> ทำให้ไม่ต้องลง Adobe Flash หรือ Microsoft Silverlight
- เพิ่ม <audio> สำหรับการใส่ไฟล์เสียงในเว็บอย่างง่าย
- เพิ่มความสามารถในการ drag and drop โดยไม่ต้องลง add-on เสริม
- เพิ่มความสามารถให้กับผู้เขียนโปรแกรม เช่น การตรวจสอบการกรอกแบบฟอร์ม
- เพิ่มความสามารถในการดึง Geolocation

Extensible Markup Language (XML)

- XML เป็นภาษาที่ใช้อธิบายข้อมูลเป็นลำดับชั้น
- มีลักษณะคล้ายกับ HTML คือมี Tag เปิดและปิด แต่ XML ไม่ได้กำหนดรูปแบบข้อมูล
- Tag ใน XML ไม่มี DTD มากำหนด เราจึงสามารถกำหนด Tag ได้เองตามความต้องการ
- เราสามารถแปลง XML ให้เป็นข้อมูลในรูปแบบต่าง ๆ ได้เช่น Web document, PDF document, comma-separated values (CSV) หรือ import ไปยัง Database Management System
- W3C เป็นองค์กรที่กำหนดการพัฒนาของ XML

XML – Sample Document

```
1 <?xml version="1.0" encoding="UTF-8"?>
2 <note>
3   <to>Tove</to>
4   <from>Jani</from>
5   <heading>Reminder</heading>
6   <body>Don't forget me this weekend!</body>
7 </note>
```

element

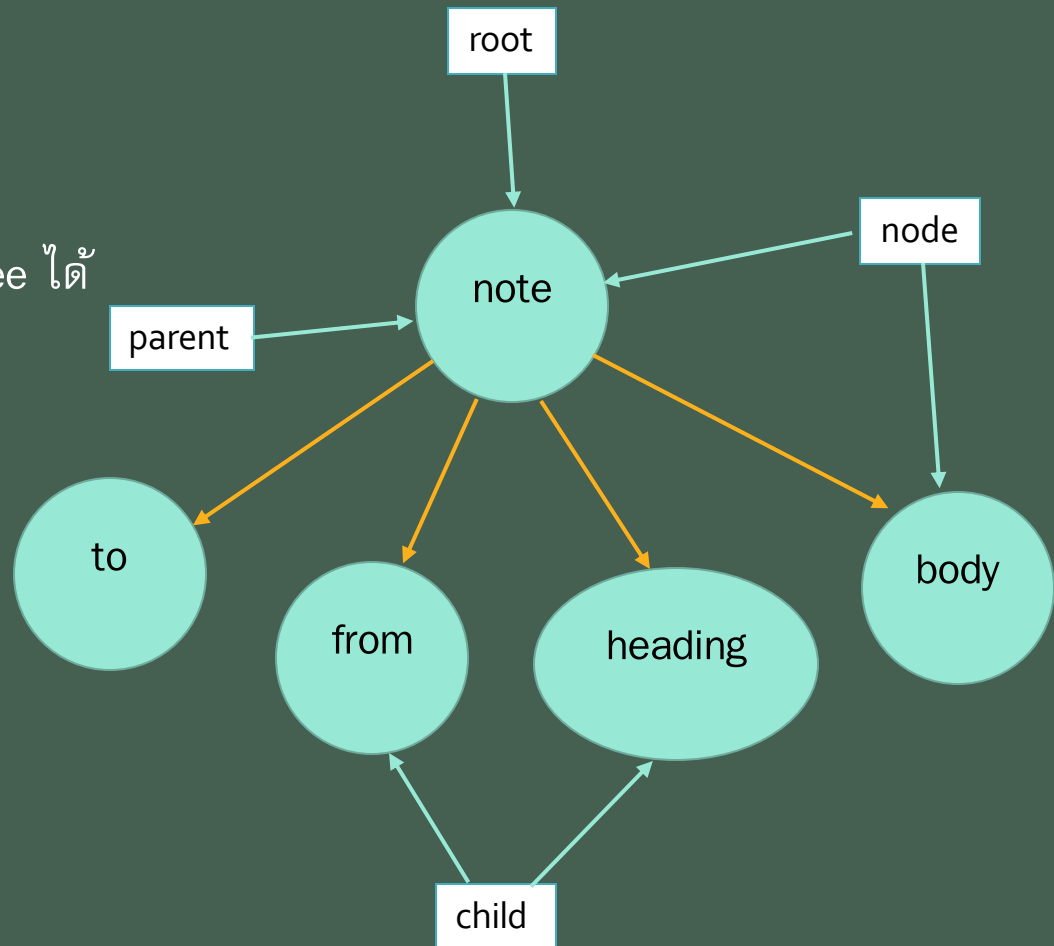


```
▼<note>
  <to>Tove</to>
  <from>Jani</from>
  <heading>Reminder</heading>
  <body>Don't forget me this weekend!</body>
</note>
```

XML Tree

- เราสามารถเขียน XML ให้อยู่รูป Tree ได้

```
▼<note>  
  <to>Tove</to>  
  <from>Jani</from>  
  <heading>Reminder</heading>  
  <body>Don't forget me this weekend!</body>  
</note>
```



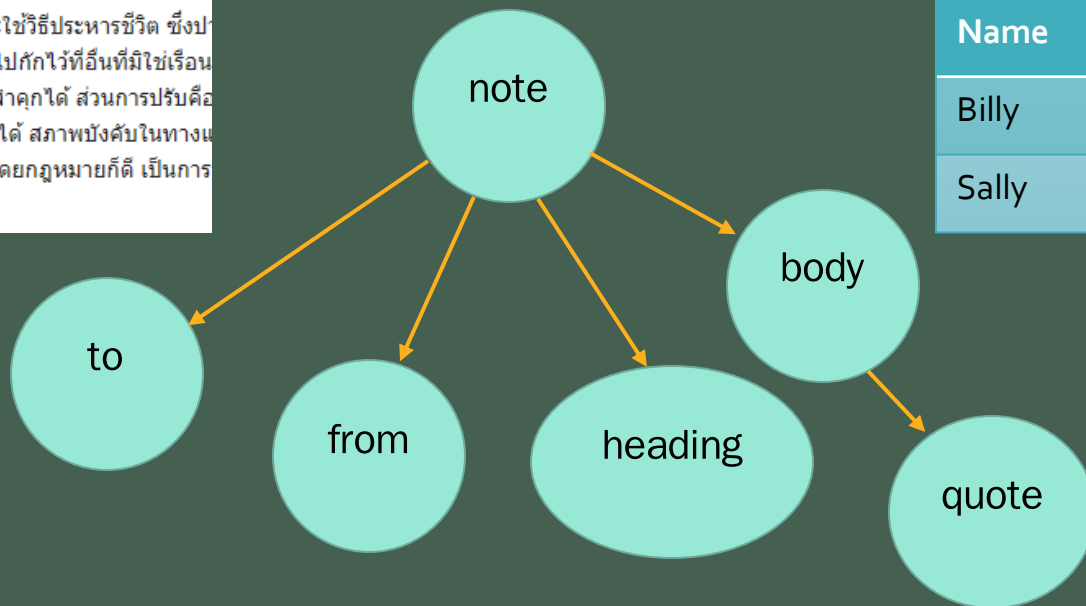
XML is Semi-Structured Data

Unstructured

Semi-structured

Structured

เพื่อให้กฎหมายเกิดความศักดิ์สิทธิ์ และประชาชนเคารพ
คล้ายคลึงกัน คือ หากเป็นโทษสูงสุดจะใช้วิธีประหารชีวิต ซึ่งป
เรือนจำ ซึ่งต่างกับกักขังเป็นการเอาตัวไปกักไว้ที่อื่นที่มีไขเรือน
สามเดือน ศาลจึงจะลงโทษกักขังแทนจำคุกได้ ส่วนการปรับคือ
ริบแล้วอาจสั่งทำลายทรัพย์สินนั้นเสียก็ได้ สภาพบังคับในทางแ
มีวัตถุประสงค์เป็นการต้องห้ามขัดแย้งโดยกฎหมายก็ดี เป็นการ
สิ้นใหม่ทดแทนแก่ผู้ถูกละเมิด เป็นต้น



Name	Age	Gender
Billy	24	M
Sally	30	F

XML Rules

- ทุกเอกสาร XML จะต้องมีการประกาศ declaration
 - `<?xml version="1.0" encoding="UTF-8" standalone="no" ?>`
- ทุก Tag เปิดจะต้องมี Tag ปิด
- Tag จะต้องไม่ทับซ้อนกัน
 - `<one><two></one></two>`
- เอกสาร XML มี element ที่เป็น root ได้เพียงอันเดียว
- ค่าของ Attribute จะต้องอยู่ภายในเครื่องหมายคำพูด หนึ่ง Attribute มีได้ค่าเดียว

XML Rules (cont.)

- ตัวอักษรที่เกี่ยวข้องกับ XML จะต้องมีการ encode เช่น
 - > > (greater than)
 - < < (less than)
 - & & (ampersand)
 - % %
 - ' ' (apostrophe)
 - " " (quotation)

XML Rules (cont.)

- ชื่อของ Tag หรือ element จะต้องเป็นไปตามข้อบังคับดังนี้
 - เริ่มต้นจากตัวอักษรหรือ underscore
 - ต่อจากนั้นสามารถประกอบไปด้วย ตัวอักษร, ตัวเลข, hyphen, period, underscore
 - ไม่มีช่องว่างในชื่อ
 - ตัวเลขใหญ่มีความสำคัญ (case-sensitive)

Extensible Hypertext Markup Language (XHTML)

- เราจะพบว่าข้อกำหนดใน HTML และ XML มีความแตกต่างกันพอสมควร ดังนั้นจึงการนำเอาข้อกำหนดทั้งสองมาผสมผสานกันจนเกิดมาเป็น XHTML
- แก้ไขความสับสนใน HTML เช่น Tag ปิด, การทับซ้อนของ Tag
- เพิ่มแน่นอนของ XML โดยการกำหนด DTD
- สามารถใช้ URL ด้านล่างในการตรวจสอบว่าหน้าเว็บของเราเป็นไปตามเงื่อนไขหรือไม่
 - <http://validator.w3.org/>

HTML to XHTML

- เพิ่ม XHTML declaration ไว้ที่บรรทัดแรกของเอกสาร (ทุก Tag เป็นการบังคับ)

```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN"
"http://www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd">

<html xmlns="http://www.w3.org/1999/xhtml">

<head>
  <title>Title of document</title>
</head>

<body>
  some content
</body>

</html>
```

HTML to XHTML (cont.)

- สลับลำดับ Tag ให้ถูกต้อง

```
<b><i>This text is bold and italic</b></i>
```

เปลี่ยนเป็น

```
<b><i>This text is bold and italic</i></b>
```

- เพิ่ม Tag ปิด

```
<p>This is a paragraph
```

เปลี่ยนเป็น

```
<p>This is a paragraph</p>
```

HTML to XHTML (cont.)

- Tag ว่างเปล่าจะต้องมีการปิด

A break: `
`

A horizontal rule: `<hr>`

An image: ``

เปลี่ยนเป็น

A break: `
`

A horizontal rule: `<hr />`

An image: ``

HTML to XHTML (cont.)

- ชื่อ Tag และ Attribute ต้องเป็นตัวพิมพ์เล็ก และค่าของ Attribute จะต้องอยู่ใน quote

```
<BODY>
```

```
<table width=100%>
```

```
</BODY>
```

เปลี่ยนเป็น

```
<body>
```

```
<table width="100%">
```

```
</body>
```

HTML to XHTML (cont.)

- Attribute จะต้องไม่มีการย่อ

```
<input type="checkbox" value="car" checked />
```

```
<input type="text" name="l" disabled />
```

เปลี่ยนเป็น

```
<input type="checkbox" value="car" checked="checked" />
```

```
<input type="text" name="l" disabled="disabled" />
```

HTML5, CSS and JavaScript

- การพัฒนาเว็บในปัจจุบันจะขึ้นอยู่กับเทคโนโลยีสามชิ้นคือ HTML5, Cascading Style Sheets (CSS), และ JavaScript โดยมักจะถูกรวมกันว่า HTML5 Family
- HTML5 ใช้กำหนดโครงสร้างของหน้าเว็บ
- CSS ใช้กำหนดรูปลักษณ์ของเว็บที่ถูกเขียนโดย Markup Language
- JavaScript เพื่อความสามารถให้เว็บตอบสนองผู้ใช้งาน
- ในคลาสนี้เราจะเรียนรู้การใช้งานสามเทคโนโลยีนี้