

บทที่ 2

ทฤษฎีและระบบงานที่เกี่ยวข้อง

2.1 ทฤษฎีที่เกี่ยวข้อง

2.1.1 ทฤษฎีที่เกี่ยวข้องกับการพัฒนาเว็บ

1. Responsive Web

การทำ Responsive Web Design [4] คือ การออกแบบเว็บไซต์ด้วยแนวคิดที่จะทำให้เว็บไซต์สามารถแสดงผลได้อย่างเหมาะสมบนอุปกรณ์ที่แตกต่างกันไปโดยใช้โค้ดร่วมกัน เพื่อแก้ปัญหาการพัฒนาเว็บไซต์หลายเวอร์ชัน ซึ่งทำให้มีต้นทุนเพิ่มขึ้นทั้งในด้านเวลาและค่าใช้จ่ายในการพัฒนา



ภาพประกอบที่ 2.1 ภาพรวมของ Online Education System

ที่มา: <https://falconsearch.com/responsive-website-design-in-uae/>

การพัฒนา Responsive Web Design จะใช้เทคนิคในการพัฒนาได้ 3 แบบ คือ Fluid Grid, Flexible Images และ CSS3 Media Queries อธิบายได้ดังนี้

- Fluid Grid คือ การออกแบบ Grid ให้เป็นแบบ Relative ให้มีความยืดหยุ่นซึ่งก็คือ การที่ไม่ได้กำหนดขนาดของ Grid แบบตายตัวแต่กำหนดให้สัมพันธ์กับสิ่งอื่นๆ เช่น กำหนดความกว้างเป็น % หรือการใช้ font-size ที่มีหน่วยเป็น em เป็นต้น

- Flexible Images คือ การกำหนดขนาดของรูปต่างๆ ให้มีความสัมพันธ์กับ ขนาดของหน้าจอแสดงผล เช่น รูปที่มีขนาดใหญ่มาก เวลาแสดงในโทรศัพท์มือถือที่มีขนาดหน้าจอเล็ก อาจจะเต็มหน้าจอเกินไป จึงควรปรับขนาดลงเพื่อให้แสดงผลได้อย่างเหมาะสม

- CSS3 Media Queries คือ การออกแบบ style sheets ให้กับขนาดของอุปกรณ์ต่างๆ โดย ออกแบบหน้าจอที่มีขนาดเล็กที่สุด แล้วจึงออกแบบหน้าจอขนาดใหญ่ตามลำดับ ซึ่งการเขียน style sheets จะช่วยลดความซ้ำซ้อนและสามารถแก้ไขโค้ดภายหลังได้สะดวก

ประโยชน์ของ Responsive Web Design

- สะดวกและลดความยุ่งยากในการพัฒนารวมทั้งช่วยลดค่าใช้จ่ายในการดูแลปรับปรุงเว็บไซต์
- เว็บไซต์มีความยืดหยุ่น ผู้ใช้งานสามารถใช้งานเว็บไซต์ได้สะดวกเมื่อเปิดเว็บไซต์ด้วยอุปกรณ์ที่มีขนาดหน้าจอใด เว็บไซต์สามารถแสดงข้อมูลได้เหมาะสมกับขนาดของหน้านั้น
- ช่วยให้เซิร์ฟเวอร์ไม่ทำงานหนักเนื่องจากเว็บไซต์ไม่ต้องเปลี่ยนเส้นทาง (redirect) ไปยังหน้า Mobile ก่อนที่จะแสดงผล

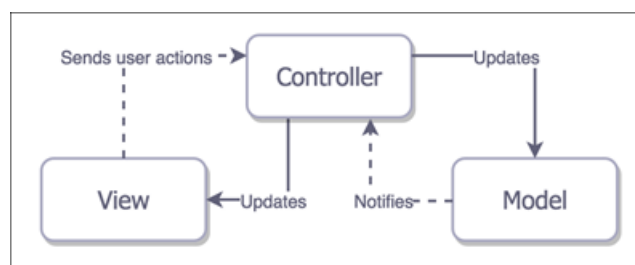
2. สถาปัตยกรรมในการพัฒนา

1) Model-View-Controller (MVC)

Model-View-Controller (MVC) [4] คือ สถาปัตยกรรมซอฟต์แวร์ประเภทหนึ่งที่เป็นที่นิยมนำมาใช้งาน ซึ่งเป็นรูปแบบการเขียนโปรแกรมตามแผนสถาปัตยกรรม (Architectural pattern) ในด้านวิศวกรรมซอฟต์แวร์ โดยจะแยกส่วนของซอฟต์แวร์ออกเป็น ส่วนๆ เช่น ส่วนของการทำงานของระบบ ส่วนของการแสดงผล ส่วนของการรับค่าและควบคุม เป็นต้น ซึ่งการแยกส่วนจะทำให้เข้าใจและเขียนซอฟต์แวร์ได้ง่าย สามารถทำงานร่วมกันเป็นทีมได้

โดยสถาปัตยกรรมแบบ Model-View-Controller (MVC) ได้แบ่งการพัฒนาออกเป็น 3 ส่วน อธิบายได้ดังนี้

- Model : M คือ ส่วนที่ไว้สำหรับจัดเตรียมข้อมูลเพื่อส่งให้กับส่วนคอนโทรลเลอร์ ซึ่งผู้ใช้งานไม่สามารถเข้าถึงข้อมูลได้โดยตรงจะต้องร้องขอข้อมูลผ่านคอนโทรลเลอร์ เพื่อร้องขอข้อมูลตามที่ผู้ใช้ต้องการ
- Views : V คือ ส่วนแสดงผลเพื่อแสดงผลข้อมูลให้กับผู้ใช้ ซึ่งมีหลากหลายรูปแบบ ขึ้นอยู่กับวัตถุประสงค์ในการแสดงผลข้อมูล หรือเรียกว่าส่วนประสานงานกับผู้ใช้ (User interface) นั้นเอง
- Controller : C คือ ส่วนที่รับคำร้องขอมาจากผู้ใช้งาน และเลือกรูปแบบการสนองตอบกลับไปยังผู้ใช้ตามการกำหนดค่าของซอฟต์แวร์ ภายในเว็บแอปพลิเคชันหน้าที่ของคอนโทรลเลอร์คือรับค่า Get หรือ Post และเลือกติดต่อกับโมเดลเพื่อรับข้อมูล ในส่วนการแสดงผลจะออกแบบฟอร์มไว้รองรับข้อมูลที่คอนโทรลเลอร์ส่งให้ซึ่งอาจจะเป็นเอกสารเอชทีเอ็มแอลเพื่อใช้แสดงผลที่หน้าเว็บเบราว์เซอร์



ภาพประกอบที่ 2.2 การทำงานของสถาปัตยกรรม MVC

ที่มา: https://toutiao.io/posts/147792/app_preview

3) Bootstrap

Bootstrap [6] คือ Front-end Framework ที่ได้รวม HTML, CSS และ JS เข้าด้วยกัน สำหรับพัฒนา Web ที่รองรับทุก Smart Device หรือเรียกว่า Responsive Web หรือ Mobile First Bootstrap ถูกพัฒนาขึ้นโดยทีมงานจาก Twitter หรือ Twitter.com ที่มากความสามารถ

Bootstrap ไม่ใช่แค่ CSS Framework สำหรับ HTML Web ธรรมดาเท่านั้น แต่เป็น Framework สำหรับ CMS ด้วย และ Bootstrap มีความสามารถมากมายทำให้ Bootstrap ไม่ใช่แค่ CSS Framework สำหรับ HTML Web ธรรมดาเท่านั้น แต่เป็น Framework สำหรับ CMS ด้วย

Bootstrap [6] มีจุดเด่นที่สามารถอธิบายได้ดังนี้

1. Theme จะเห็นได้ว่า bootstrap สามารถปรับเปลี่ยนรูปแบบได้มากมาย จากไฟล์ CSS โดยสามารถแก้ไขได้ด้วยตนเอง หรือดาวน์โหลดจากอินเทอร์เน็ตได้ฟรี ไม่มีลิขสิทธิ์
2. Responsive Design ตัว bootstrap เองมีการพัฒนาโดยคำนึงถึงหลักการ mobile-first approach หรือการพัฒนาเว็บแอปพลิเคชัน โดยโฟกัสไปที่หน้าจอขนาดเล็ก ก่อน ซึ่งวิธีนี้ ถือเป็น best practice ของการทำ responsive web ในปัจจุบัน นอกจากนี้ CSS ของ bootstrap จะรวม responsive features มาให้ในตัว รองรับการทำงานกับทุกอุปกรณ์
3. Grid System ใน bootstrap มีการออกแบบการใช้คอลัมน์คือ mobile-first responsive fluid grid system โดยมีจำนวน 12 คอลัมน์ ผู้ออกแบบสามารถกำหนดขนาดของคอลัมน์ให้กับ devices ในแต่ละแบบได้
4. Components ใน bootstrap มี Components style sheets สำหรับ สิ่งที่เราต้องใช้อย่าง ไม่ว่าจะเป็น navigation, breadcrumbs รวมไปถึง pagination JavaScript jQuery plugins ต่างๆ ไม่ว่าจะเป็น modal, carousel หรือ tooltip เป็นต้น
5. Browser support ตัว bootstrap รองรับการทำงานกับ Internet Explorer 8 ขึ้นไป ส่วน Browser อื่นนั้นยังคงรองรับอยู่ แต่เราต้องใส่ respond.js เพิ่มเข้าไปด้วย เพื่อให้ใช้ media queries ได้

3) ภาษาที่ใช้ในการพัฒนา

1) HTML

การสร้างเว็บไซต์ให้ตอบสนองการทำงานตามความต้องการของผู้ใช้นั้น ผู้พัฒนาเว็บไซต์ต้องใช้เทคโนโลยีในการสร้างเว็บไซต์ที่หลากหลาย รวมทั้งภาษาที่ใช้พัฒนาก็ต้องมีความหลากหลายเพื่อให้สามารถทำงานตามที่ต้องการได้

ภาษา HTML เป็นภาษาประเภทมาร์คอัพ (Markup) คือ มีการแสดงข้อมูลและรูปแบบการแสดงผลเข้าด้วยกันสามารถสร้างและพัฒนาได้ง่าย โดยผู้ใช้ไม่จำเป็นต้องมีความรู้ทางการเขียนโปรแกรมมาก่อน หากเขียน TAG HTML การแสดงผลอาจจะไม่เป็นไปตามที่ต้องการ ปัจจุบันนี้ภาษา HTML ถูกพัฒนาให้สามารถรองรับการทำงานได้หลากหลายอุปกรณ์โดยเฉพาะอุปกรณ์เคลื่อนที่อย่างสมาร์ทโฟน แท็บเล็ต เป็นต้น

HTML [7] คือ ภาษา HTML ย่อมาจากคำว่า Hyper Text Markup Language เป็นภาษาที่ใช้ในการกำหนดการแสดงผลข้อมูลภายในหน้าเว็บเพจเพื่อแสดงผลบนโปรแกรมบราวเซอร์ ประกอบด้วย คำสั่งที่อยู่ภายในเครื่องหมายวงเล็บ <> เรียกว่า TAG เป็นส่วนควบคุมให้โปรแกรมบราวเซอร์ แสดงผลข้อมูลตามชื่อ TAG การเปิดการควบคุมคำสั่ง HTML ด้วย และปิดการ ควบคุมคำสั่งใน TAG จะประกอบด้วย Attribute ซึ่งเป็นส่วนขยายของ TAG เป็นการเพิ่มคุณสมบัติของการแสดงผล TAG HTML ให้เป็นไปตามที่กำหนด

การสร้างเว็บเพจด้วยภาษา HTML ประกอบด้วยโครงสร้าง 3 ส่วน สามารถอธิบายได้ดังนี้

1.1 <HTML>...</HTML> กำหนดจุดเริ่มต้นและจุดสิ้นสุดเว็บเพจภาษา HTML

1.2 <HEAD>...</HEAD> ส่วนหัวของเว็บเพจ ประกอบด้วย TAG ดังนี้

1.2.1 <META>กำหนดข้อมูลเกี่ยวกับเว็บเพจเช่น คำค้นหา (KEYWORD) คำอธิบายเว็บเพจ (Description) ชื่อผู้จัดทำเว็บไซต์ (Author) กำหนดอักขระภาษาของเว็บเพจและกำหนดเวลารีเฟรชหน้าเว็บเพจ (Refresh page) เป็นต้น และที่สำคัญ TAG META เป็น TAG เดี่ยวไม่มี TAG ปิด

1.2.2 <TITLE>...</TITLE> กำหนดชื่อของเว็บเพจ

1.2.3 <STYLE>...</STYLE> กำหนดรูปแบบ CSS การแสดงผลของเว็บเพจ
<LINK>...</LINK> เชื่อมโยงไฟล์ CSSจากภายนอก

1.2.4 <SCRIPT>...</SCRIPT> ควบคุมเว็บเพจด้วยภาษา JavaScript

1.3 <BODY>....</BODY>แสดงข้อมูลต่างๆ บนเว็บเพจอยู่ใน TAG HTML

ตัวอย่างของโครงสร้างภาษา HTML แสดงได้ดังภาพประกอบที่ 2.3

```
<HTML>
<HEAD>
<TITLE>การสร้างเว็บเพจด้วย HTML</TITLE>
<META NAME="KEYWORDS" CONTENT="HTML, เว็บเพจ">
<META NAME="DESCRIPTION" CONTENT="มหาวิทยาลัยราชภัฏเชียงราย">
<META NAME="AUTHOR" CONTENT="มหาวิทยาลัยราชภัฏเชียงราย">
<META HTTP-EQUIV="CONTENT-TYPE" CONTENT="TEXT/HTML; CHARSET=
</HEAD>
<BODY>
.....TAG NAME และข้อมูลต่างๆ ที่แสดงผลบนบราวเซอร์.....
</BODY>
</HTML>
```

ภาพประกอบที่ 2.3 โครงสร้างภาษา HTML

ที่มา: <http://srinuan.crru.ac.th/CIT3206/pdf/04HTML1.pdf>

2) JavaScript

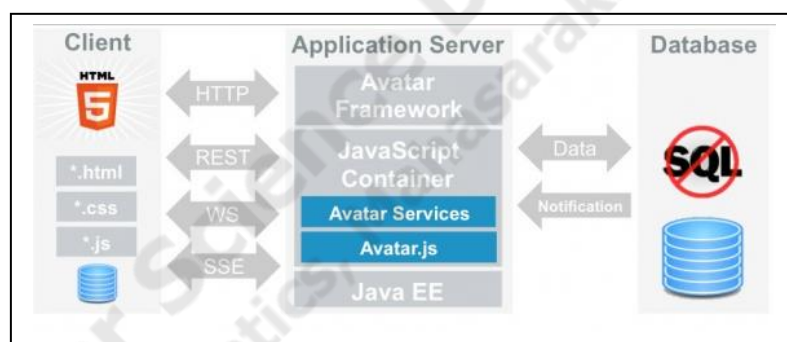
JavaScript [8] เป็นภาษาคอมพิวเตอร์สำหรับการเขียนโปรแกรมบนระบบอินเทอร์เน็ต ซึ่งเป็นภาษาสคริปต์เชิงวัตถุ สามารถใช้ร่วมกับภาษา HTML เพื่อการสร้างและพัฒนาเว็บไซต์ ทำให้เว็บไซต์มีการเคลื่อนไหว สามารถตอบสนองผู้ใช้งานได้มากขึ้น หรือเรียกว่า (Object Oriented Programming) ที่มีเป้าหมายในการออกแบบและพัฒนาโปรแกรมในระบบอินเทอร์เน็ต สำหรับผู้เขียนภาษา HTML สามารถทำงานข้ามแพลตฟอร์มได้โดยทำงานร่วมกับภาษา HTML และภาษาจาวา (Java) ได้ทั้งฝั่งไคลเอนต์ (Client) และฝั่งเซิร์ฟเวอร์ (Server) ซึ่งมีลักษณะการเขียนแบบ โปรโตไทป์ (Prototyped-based Programming) ส่วนมากใช้ในหน้าเว็บเพื่อประมวลผลข้อมูลที่ฝั่งของผู้ใช้งาน แต่ก็ยังมีใช้เพื่อเพิ่มเติมความสามารถในการเขียนสคริปต์โดยฝังอยู่ในโปรแกรมอื่นๆ

จาวาสคริปต์ (JavaScript) ถูกพัฒนาขึ้นโดย Netscape Communications Corporation โดยใช้ชื่อว่า Live Script ออกมาพร้อมกับ Netscape Navigator 2.0 เพื่อใช้สร้างเว็บเพจโดยติดต่อกับเซิร์ฟเวอร์แบบ Live Wire

ต่อมาเน็ตสเคปจึงได้ร่วมมือกับบริษัทซันไมโครซิสเต็มส์ปรับปรุงระบบของบราวเซอร์ เพื่อให้สามารถติดต่อกับภาษาจาวาได้ และได้ปรับปรุง Live Script ใหม่ เมื่อปี 2538 แล้วตั้งชื่อใหม่ว่า JavaScript ซึ่งสามารถทำให้การสร้างเว็บเพจ มีลูกเล่นต่างๆ มากมาย และยังสามารถติดต่อกับผู้ใช้ได้อย่างทันที เช่น การใช้เมาส์คลิก หรือ การกรอกข้อความในฟอร์ม เป็นต้น

ประโยชน์ของ JavaScript สามารถอธิบายได้ดังนี้

- เขียนโปรแกรมแบบง่ายๆได้ โดยไม่ต้องพึ่งภาษาอื่น
- มีคำสั่งที่ตอบสนองกับผู้ใช้งาน เช่น เมื่อผู้ใช้คลิกที่ปุ่ม หรือ Checkbox ก็สามารถสั่งให้เปิดหน้าต่างใหม่ได้ ทำให้เว็บไซต์ของเรามีปฏิสัมพันธ์กับผู้ใช้งานมากขึ้น ด้วยข้อดีของ JavaScript ทำให้เว็บไซต์ต่างๆนำคำสั่งจาวาสคริปต์ไปใช้
 - สามารถเขียนหรือเปลี่ยนแปลง HTML Element ได้ นั่นคือสามารถเปลี่ยนแปลงรูปแบบการแสดงผลของเว็บไซต์ได้ หรือหน้าแสดงเนื้อหาสามารถซ่อนหรือแสดงเนื้อหาได้แบบง่ายๆ
 - สามารถใช้ตรวจสอบข้อมูลได้ สังเกตว่าเมื่อเรากรอกข้อมูลบางเว็บไซต์ เช่น Email เมื่อเรากรอกข้อมูลผิดจะมีหน้าต่างฟ้องขึ้นมาว่าเรากรอกผิด หรือลืมกรอกอะไรบางอย่าง เป็นต้น
 - สามารถใช้ในการตรวจสอบผู้ใช้ได้ เช่น ตรวจสอบว่าผู้ใช้ใช้เว็บเบราว์เซอร์อะไร
 - สร้าง Cookies (เก็บข้อมูลของผู้ใช้ในคอมพิวเตอร์ของผู้ใช้เอง) ได้



ภาพประกอบที่ 2.4 คุณสมบัติของ JavaScript

ที่มา: <https://www.itgenius.co.th/article/คุณสมบัติของ%20JavaScript%20ทำอะไรได้บ้าง.html>

ข้อจำกัดของ JavaScript การทำงานของ JavaScript เกิดขึ้นบนเบราว์เซอร์ (เรียก client-side script) ดังนั้นไม่ว่าจะใช้เซิร์ฟเวอร์อะไร หรือที่ไหน ก็ยังคงสามารถใช้ JavaScript ในเว็บเพจได้ ต่างกับภาษาสคริปต์อื่น เช่น Perl, PHP หรือ ASP ซึ่งต้องแปลความและทำงานที่ตัวเครื่องเซิร์ฟเวอร์ (เรียกว่า server-side script) ดังนั้นจึงต้องใช้บนเซิร์ฟเวอร์ที่สนับสนุนภาษาเหล่านี้เท่านั้น

อย่างไรก็ตามจากลักษณะดังกล่าวก็ทำให้ JavaScript มีข้อจำกัด คือไม่สามารถรับและส่งข้อมูลต่างๆ กับเซิร์ฟเวอร์โดยตรง เช่น การอ่านไฟล์จากเซิร์ฟเวอร์ เพื่อนำมาแสดงบนเว็บเพจ หรือรับข้อมูลจากผู้ชม เพื่อนำไปเก็บบนเซิร์ฟเวอร์ เป็นต้น ดังนั้นงานลักษณะนี้ จึงยังคงต้องอาศัยภาษา server-side script อยู่

3) PHP

PHP [9] คือ ภาษาคอมพิวเตอร์ในลักษณะเซิร์ฟเวอร์-ไซด์ สคริปต์และเป็นภาษาที่อยู่ในลิขสิทธิ์ของโอเพนซอร์ส (Open source) ปัจจุบันนิยมนำภาษาพีเอชพีมาใช้ในการพัฒนาเว็บไซต์ นอกจากจะง่ายต่อการพัฒนาและสามารถโต้ตอบได้รวดเร็ว อีกทั้งยังสามารถทำงานร่วมกับระบบปฏิบัติการ ได้เกือบทุก Platform เช่น Windows, Unix (Linux), Mac เป็นต้น และจะทำการแสดงผลออกมาในรูปแบบ HTML เพื่อให้ผู้ใช้งานสามารถเรียกแสดงผลผ่านได้เว็บเบราว์เซอร์ ในภาษา PHP นั้นจะมีรากฐานโครงสร้างคำสั่งมาจากภาษาซี ภาษาเพิร์ล และภาษาจาวา ซึ่งง่ายต่อการเรียนรู้

PHP [10] เป็นภาษาจำพวก Scripting Language คำสั่งต่าง ๆ จะเก็บอยู่ในไฟล์ที่เรียกว่าสคริปต์ (Script) และเวลาใช้งานต้องอาศัยตัวแปลชุดคำสั่ง ตัวอย่างของภาษาสคริปต์เช่น JavaScript, Perl เป็นต้น

ลักษณะของ PHP ที่แตกต่างจากภาษาสคริปต์แบบอื่น ๆ คือ PHP ได้รับการพัฒนาและออกแบบมาเพื่อใช้งานในการสร้างเอกสารแบบ HTML โดยสามารถสอดแทรกหรือแก้ไขเนื้อหาได้โดยอัตโนมัติ ดังนั้นจึงกล่าวได้ว่า PHP เป็นภาษาที่เรียกว่า Server-Side หรือ HTML-Embedded Scripting Language เป็นเครื่องมือที่สำคัญชนิดหนึ่ง ที่ช่วยให้เราสามารถสร้างเอกสารแบบ Dynamic HTML ได้อย่างมีประสิทธิภาพและมีลูกเล่นมากขึ้น อาจจะสามารถได้ว่า PHP ได้รับการพัฒนาขึ้นมา เพื่อแทนที่ Server Side Include (SSI) รูปแบบเดิม ๆ โดยให้มีความสามารถ และมีส่วนเชื่อมต่อกับเครื่องมือชนิดอื่นมากขึ้น เช่น ติดต่อกับคลังข้อมูลหรือ Database

ข้อดีของภาษา PHP

- สร้างฟอร์มโต้ตอบ หรือรับส่งข้อมูลกับผู้ใช้ได้
- แทรกโค้ด PHP เข้าไประหว่างโค้ดภาษา HTML ได้ทันที
- มีฟังก์ชันสนับสนุนการทำงานที่หลากหลาย
- สามารถติดต่อกับฐานข้อมูลได้หลากหลาย เช่น Access MySQL Oracle SQLServer
- สนับสนุนการติดต่อกับโปรโตคอลได้หลากหลาย เช่น HTTP POP3
- ใช้งานได้ฟรี เพราะเป็น Opensource
- ใช้งานได้บนทุกระบบปฏิบัติการ ไม่ว่าจะเป็น UNIX, Linux หรือ Windows
- ภาษาสคริปต์ใช้คำสั่งที่เข้าใจง่าย
- ใช้ร่วมกับคำสั่ง XML ได้ทันที

4) CSS

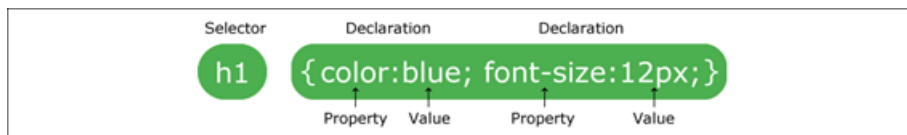
CSS ย่อมาจาก Cascading Style Sheets [11] เป็นอีกหนึ่งภาษาสไตลชีตที่มีรูปแบบการเขียน Syntax ที่เฉพาะ เหมือนกันกับ HTML และ XHTML ซึ่งเป็นภาษาที่ใช้สำหรับการจัดรูปแบบและตกแต่งเอกสาร HTML หรือ XHTML ไม่ว่าจะป็นหน้าตา สี สัน ตัวอักษร เส้นขอบ พื้นหลัง ระยะห่างเพื่อให้เว็บเพจมีความยืดหยุ่น มีความน่าสนใจ รวมไปถึงการควบคุมลักษณะของข้อมูลในการนำเสนอซึ่งผู้นำเว็บสามารถที่จะทำการแก้ไข ได้โดยง่าย และสะดวกรวดเร็ว ซึ่ง CSS นี้จะถูกกำหนดมาตรฐานโดย W3C (World Wide Web Consortium) ในการกำหนดค่าต่างๆ ให้กับเอกสาร HTML หรือ XHTML นั้น เราสามารถดึงเอา Element ของ HTML เช่น <body>, <p>, <h1> มากำหนดคุณสมบัติต่างๆ ได้ตามที่เราต้องการ ซึ่งทำให้เว็บไซต์ของเรามีความเป็นมาตรฐานมากยิ่งขึ้น

ประโยชน์ของ CSS

ในการใช้งาน CSS ซึ่งไม่ใช่เพียงแต่นำมาใช้กับการตกแต่งเอกสาร HTML หรือเว็บเพจเท่านั้น แต่ CSS ยังสามารถที่จะใช้ในการตกแต่งข้อมูลที่ได้มาจาก XML ได้อีกด้วย สำหรับประโยชน์ของ CSS นั้นมีอยู่มากมาย ขึ้นอยู่กับว่าเราต้องการที่จะนำเอาไปประยุกต์ใช้ในด้านไหน ในที่นี้ขอกกล่าวถึงประโยชน์ของ CSS ในการนำเอามาใช้ในการพัฒนาเว็บเพจในปัจจุบัน ซึ่งประโยชน์ของ CSS มีดังนี้

- ในการใช้ CSS จัดรูปแบบการแสดงผล จะช่วยให้เว็บเพจของเราลดการใช้ภาษา HTML ทำให้ซอร์จโค้ด ภายในเอกสาร HTML ลดน้อยลง ซึ่งจะเหลือเพียงส่วนที่เป็นเนื้อหา ทำให้เข้าใจง่ายยิ่งขึ้น และสามารถทำการแก้ไขเอกสารทำได้ง่ายและรวดเร็ว
- เนื่องจากซอร์จโค้ดมีจำนวนที่ลดลง ทำให้ขนาดของไฟล์ที่เขียนขึ้นมีขนาดเล็กลงด้วย ซึ่งจะทำให้การโหลดหน้าเว็บไซต์มีความเร็วเพิ่มขึ้น
- การเขียน CSS ไว้ชุดคำสั่งเดียว มีผลต่อเอกสาร HTML ทั้งหมด ซึ่งสามารถที่จะกำหนดรูปแบบในการแสดงผลได้ทั้งหน้าเดียวหรือหลายหน้าก็ได้ทำให้ง่ายต่อการปรับปรุงแก้ไข และไม่ต้องตามหาที่แก้ไขทั้งหน้าเอกสารนั่นเอง
- การนำ CSS มาสร้างเลเอาท์ในเว็บเพจ จะมีผลให้เว็บเพจของเราโหลดได้เร็วมากยิ่งขึ้นและซอร์จโค้ดมีความเป็นระเบียบมากยิ่งขึ้น
- CSS ทำให้เราสามารถที่จะเปลี่ยนแปลงการแสดงผลในหน้าเว็บเพจได้อย่างง่ายดาย โดยที่เราไม่ต้องไปยุ่งเกี่ยวกับข้อมูลในเอกสาร HTML
- Web Browser หลายตัวที่รองรับการแสดงผลของ CSS ซึ่งทำให้เราสามารถควบคุมการแสดงผลให้เหมือนกัน หรือใกล้เคียงกัน
- CSS สามารถกำหนดการแสดงผลในรูปแบบที่เหมาะสมกับสื่อหลากหลายชนิด เช่น PDA, โทรศัพท์มือถือ, สื่อสิ่งพิมพ์หรือทางหน้าจอ เป็นต้น

- ในการนำเอา CSS มาใช้งานนั้น ทำให้เว็บไซต์ของเรามีมาตรฐาน เนื่องจาก W3C แนะนำให้ใช้ CSS แทนการตกแต่งเว็บเพจโดยใช้ HTML



ภาพประกอบที่ 2.5 โครงสร้างคำสั่ง CSS

ที่มา: <https://www.w3schools.com/css/>

จากภาพประกอบที่ 2.5 จะแสดงให้เห็นโครงสร้างคำสั่งอย่างง่ายของ CSS [12] ซึ่งประกอบไปด้วย selectors, declarations, property, value สามารถอธิบายได้ดังนี้

1. Selector คือ ตำแหน่งหรือจุดที่ HTML element อยู่แล้วเราต้องการกำหนดสไตล์ หรือ CSS เช่นข้อความส่วนหัว h1,h2,h3...n <h1>head title</h1> พารากราฟ(paragraph) <p>paragraph</p> และนี่คือ selector ตัวอย่างที่เราต้องการใส่ CSS เข้าไป

2. Declaration คือ ตำแหน่งที่เราต้องการกำหนดลักษณะของสไตล์นั่นเอง ประกอบไปด้วย property และ value ซึ่งถ้ามีหลายๆ property ก็จำเป็นต้องคั่นด้วย เซมิโคลอน (;) คั่นไปเรื่อยๆ

- property เช่น กำหนดฟอนต์ font-size จัดข้อความให้อยู่ตรงกลาง text-align เป็นต้น
- value คือค่าของ property เช่น font-size : 16px ; text-align : center เป็นต้น

```
// Example of good basic formatting practices
.styleguide-format {
  color: #000;
  background-color: rgba(0, 0, 0, .5);
  border: 1px solid #0f0;
}

// Example of individual selectors getting their own lines (for error reporting)
.multiple,
.classes,
.get-new-lines {
  display: block;
}

// Avoid unnecessary shorthand declarations
.not-so-good {
  margin: 0 0 20px;
}
.good {
  margin-bottom: 20px;
}
```

ภาพประกอบที่ 2.6 ตัวอย่างคำสั่ง CSS

ที่มา : <https://css-tricks.com/enforcing-css-syntax-style/>

4) เครื่องมือ (Tools)

1. XAMPP

XAMPP [13] คือเป็นหนึ่งในโปรแกรมสำหรับจำลองเครื่องคอมพิวเตอร์ส่วนบุคคลให้ทำงานในลักษณะของเว็บเซิร์ฟเวอร์ นั่นคือการจำลองเครื่องคอมพิวเตอร์ให้เป็นทั้งเครื่องแม่และเครื่องลูกในเครื่องเดียวกัน เพื่อเป็นการทดสอบสคริปหรือเว็บไซต์ในเครื่อง โปรแกรมนี้เป็นโปรแกรมฟรีและไม่มีค่าใช้จ่ายใด ๆ ภายในโปรแกรม XAMPP ประกอบด้วย Apache, PHP, MySQL, phpMyAdmin, Perl ซึ่งเป็นโปรแกรมพื้นฐานที่รองรับการทำงาน CMS ซึ่งเป็นชุดโปรแกรมสำหรับออกแบบเว็บไซต์ที่ได้รับความนิยม

ในปัจจุบันไฟล์สำหรับติดตั้งโปรแกรม XAMPP นั้นอาจมีขนาดใหญ่ เนื่องจากมีชุดควบคุมการทำงานที่ช่วยให้การปรับแต่งส่วนต่าง ๆ ง่ายขึ้น นอกจากนี้โปรแกรม XAMPP นั้นยังรองรับระบบปฏิบัติการหลายตัว เช่น Windows, Linux, Apple และสามารถทำงานได้ทั้งบนระบบปฏิบัติการแบบ 32 bit และ 64 bit สิ่งที่น่าสนใจกว่าโปรแกรมอื่นคือมีตัวช่วยติดตั้ง CMS ที่เรียกว่า BitNami ซึ่งช่วยให้คุณติดตั้ง CMS รุ่นใหม่ ๆ ที่ได้รับความนิยมอีกด้วย

โปรแกรม XAMPP สามารถใช้งานได้ 4 OS [14] ได้แก่

- Windows สามารถใช้งานได้กับ windows รุ่น 2000, 2003, xp, vista, windows 7
- Linux สำหรับ SuSE, RedHat, Mandrake, Debian และ Ubuntu
- Mac OS X
- Solaris สำหรับ Solaris 8 และ Solaris 9

ในการใช้งานเบื้องต้นให้ดับเบิลคลิก Xampp Control Panel Application แล้วทำการคลิกปุ่ม start จากนั้นสามารถใช้งานได้ โดยเปิด Browser ขึ้นมาพิมพ์ localhost หรือ 127.0.0.1

ข้อจำกัดด้านเทคนิค

- เครื่องคอมพิวเตอร์ควรมี RAM ไม่ต่ำกว่า 128 MB
- Hard disk มีพื้นที่มากกว่า 320 MB
- CPU ไม่กำหนดขั้นต่ำ

2. Live Stream [15]

ในปัจจุบันนี้ บริการสตรีมมิ่งบนอินเทอร์เน็ตกำลังเป็นที่นิยมมากและจะเข้ามาแทนที่ช่องโทรทัศน์ในไม่กี่ปีข้างหน้า เพราะทุกวันนี้ การใช้ชีวิตประจำวันของบางคนก็แทบไม่ได้รับรู้ข่าวสารจากโทรทัศน์เลย ด้วยสาเหตุต่างๆไม่ว่าจะเป็นทำงานกลับดึก เรียนพิเศษ หรือจากการเดินทางที่ใช้เวลานานๆ เลยเลือกที่จะรับข่าวสารต่างๆจากช่องทางใกล้ตัวมากกว่านั้นคือ มือถือ โดยแค่มี 3G 4G

หรือ Wi-Fi ปริมาณอินเทอร์เน็ตที่มากพอก็สามารถดูได้แล้ว ใช้โปรเน็ตรายเดือนที่ปัจจุบันมีการลดราคา และโปรโมชั่นแข่งขันกัน ก็สามารถดูสตรีมมิ่งรายการต่างๆ สดผ่านโทรศัพท์มือถือได้ทันที แถมน้ำปลาสด ช่วงเวลานั้นไปก็สามารถดูย้อนหลังได้ด้วย

คำว่า "สตรีมมิ่ง" [16] จะถูกนำไปใช้ในกรณีที่คุณสามารถเล่นไฟล์มัลติมีเดียบนเครื่องคอมพิวเตอร์ได้โดยไม่ต้องมีการดาวน์โหลดจากอินเทอร์เน็ตจนครบไฟล์ เนื่องจากการดาวน์โหลดไฟล์มัลติมีเดียทั้งไฟล์จะใช้เวลาดาวน์โหลดมาก ดังนั้นการเล่นไฟล์มัลติมีเดียจากอินเทอร์เน็ตด้วยการสตรีมมิ่ง จะทำให้สามารถแสดงผลข้อมูลได้ก่อนที่ไฟล์ทั้งหมดจะถูกส่งผ่านเข้ามายังเครื่องคอมพิวเตอร์นั่นเอง

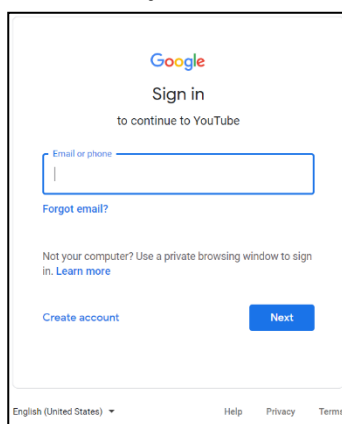
สำหรับการทำให้การสตรีมมิ่งสามารถเล่นไฟล์มัลติมีเดียได้อย่างสมบูรณ์นั้น คอมพิวเตอร์ที่ใช้จะต้องประมวลผลประมวลผลได้เร็วพอด้วย เนื่องจากข้อมูลที่ถูกส่งเข้ามายังเครื่องนอก จะต้องได้รับการจัดเก็บเข้าไปในหน่วยความจำบัฟเฟอร์แล้ว มันยังต้องมีการแปลงข้อมูลเหล่านั้น เพื่อนำไปแสดงผลในรูปของเสียง หรือวิดีโอ อีกด้วย ซึ่งถ้าขั้นตอนใดขั้นตอนหนึ่งล่าช้า คุณก็จะสังเกตเห็นได้ทันทีว่าวิดีโอหรือเสียงมีการกระตุก หรือแฉ่งเป็นระยะๆ (การกำหนดขนาดของหน่วยความจำบัฟเฟอร์ ความเร็วของเครื่องคอมพิวเตอร์ และอินเทอร์เน็ต เป็นตัวแปรสำคัญในการปรับแต่งให้การเล่นสตรีมมิ่งบนเครื่องคอมพิวเตอร์ราบรื่น)

Streaming สามารถนำมาประยุกต์ใช้ในรูปแบบต่างๆ ได้หลากหลาย โดยสามารถนำมาใช้ให้บริการได้ดังต่อไปนี้

- บริการ Video On demand เป็นการบันทึก Video Clip ไว้บน Server ให้ผู้ใช้สามารถเลือกดูคลิปวิดีโอผ่านระบบอินเทอร์เน็ต
- บริการ Video Broadcasting เป็นการถ่ายทอดสดวิดีโอผ่านระบบอินเทอร์เน็ตแบบ Real time

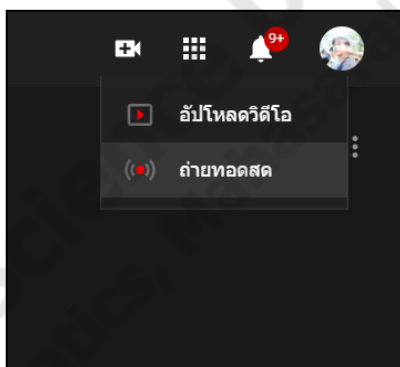
เนื่องจากระบบ Streaming เป็นการทำงานผ่านระบบอินเทอร์เน็ต ในลักษณะของ Web Application จึงทำให้สามารถพัฒนาซอฟต์แวร์ขึ้นมาใช้งานได้หลากหลายตามความต้องการ ซึ่งการทำไลฟ์สตรีมในโครงการปริญญาโทนี้มีวิธีการทำที่สามารถอธิบายได้ดังนี้

1. เปิดหน้า YouTube แล้วทำการ Login เข้าสู่ระบบ



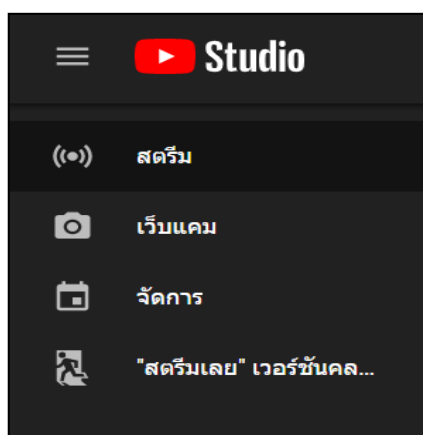
ภาพประกอบที่ 2.7 แสดงหน้าจอ login

2. คลิกเลือก ถ่ายทอดสด



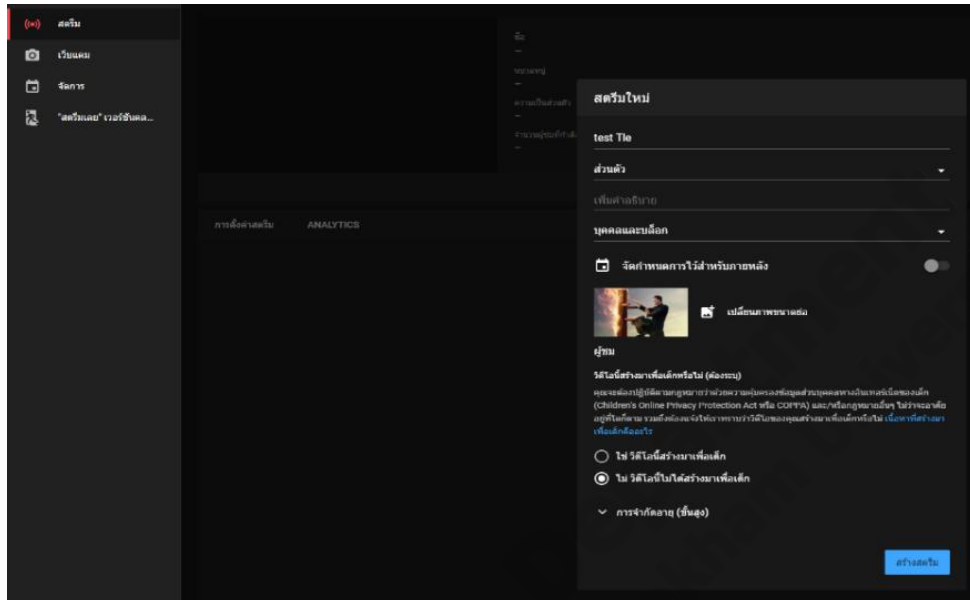
ภาพประกอบที่ 2.8 แสดงหน้าจอตัวเลือกการทอดสด

3. เมื่อเลือก ถ่ายทอดสด จะพบหน้าจอดังนี้ เลือกที่ สตรีม



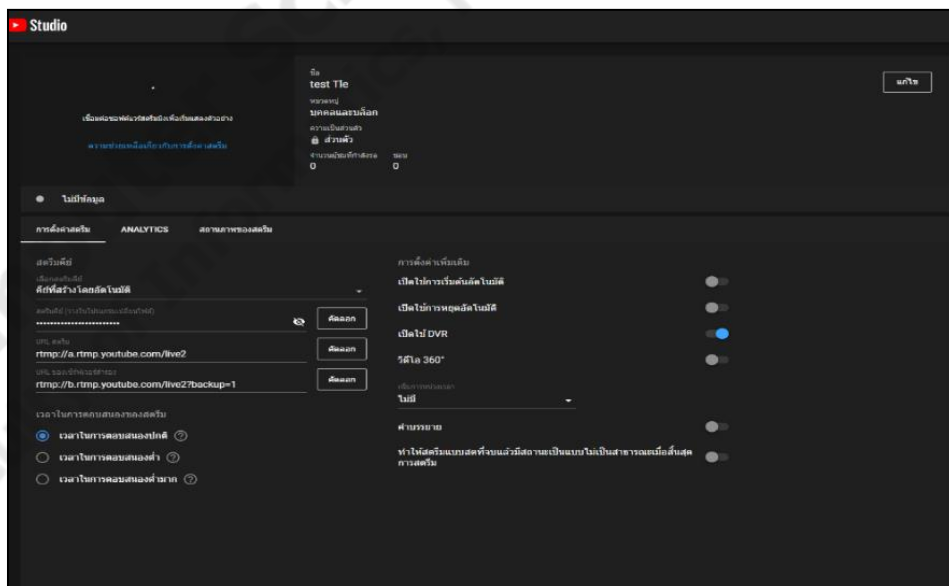
ภาพประกอบที่ 2.9 หน้าจอแสดงรูปแบบการถ่ายทอดสด

4. เมื่อเลือก สตรีม จะพบหน้าจอดังนี้ ใส่รายละเอียดงานที่ต้องการจะถ่ายทอดสดแล้วกด สร้างสตรีม



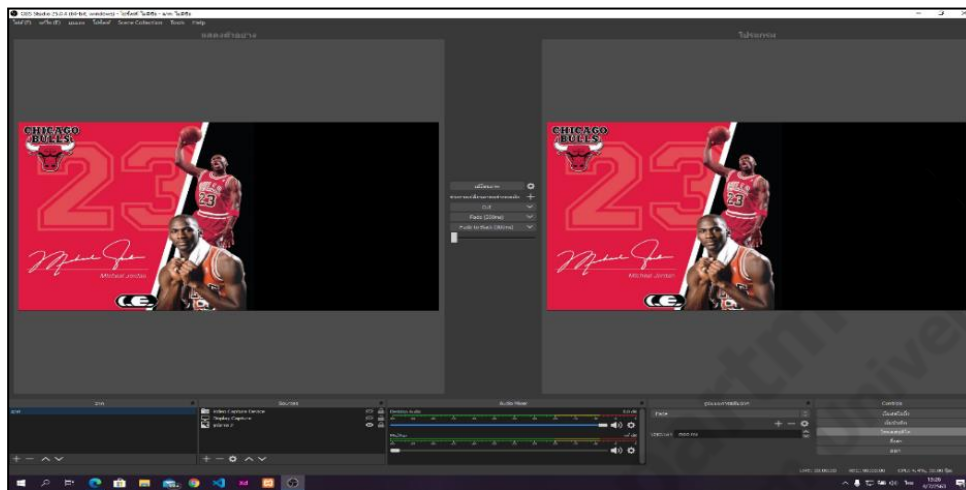
ภาพประกอบที่ 2.10 แสดงหน้าจอการใส่รายละเอียดการสตรีม

5. จะพบหน้าจอตั้งนี้ เลือกที่คัดลอก สตรีมคีย์



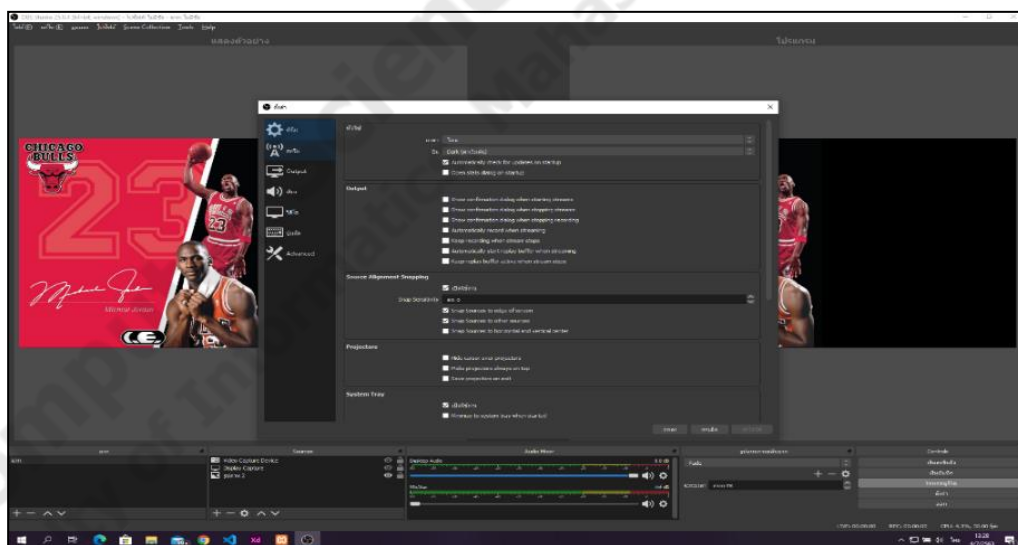
ภาพประกอบที่ 2.11 แสดงหน้าจอรายละเอียดการตั้งค่า

6. เลือกโปรแกรมสำหรับ encode (ในที่นี้ขอตกลงโปรแกรม OBS Studio)



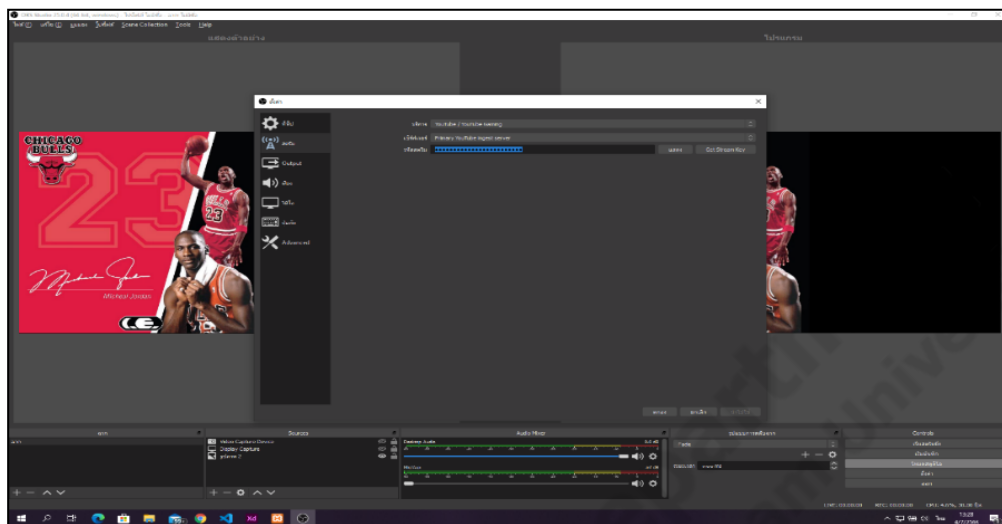
ภาพประกอบที่ 2.12 หน้าจอแสดงโปรแกรมที่ใช้สตรีม

7. เลือกที่ ตั้งค่า แล้วจะขึ้นหน้าต่างดังนี้ เลือกที่สตรีม



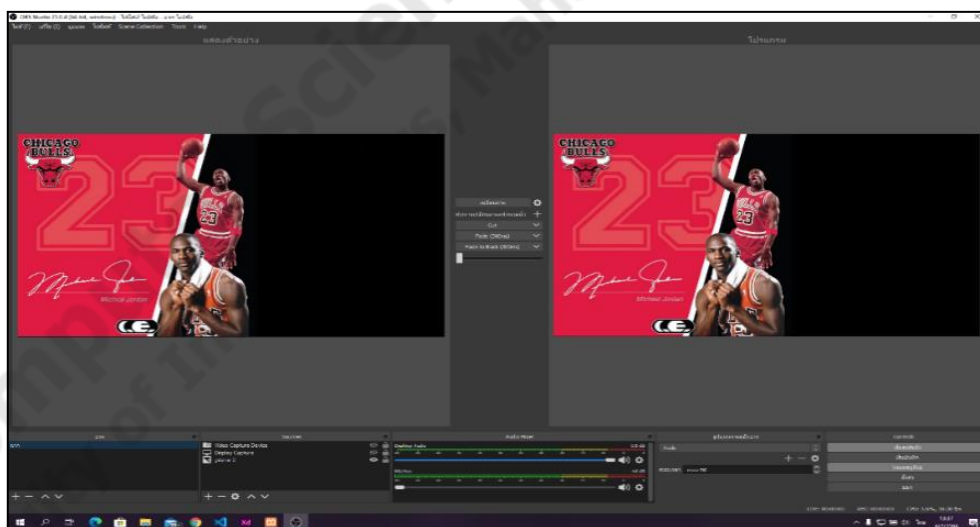
ภาพประกอบที่ 2.13 หน้าจอแสดงการตั้งค่าโปรแกรม

8. จะพบหน้าต่างนี้ นำสตรีมคีย์ที่คัดลอกมาวางไว้ที่ รหัสสตรีม คลิกที่ นำไปใช้ จากนั้น คลิกตกลง



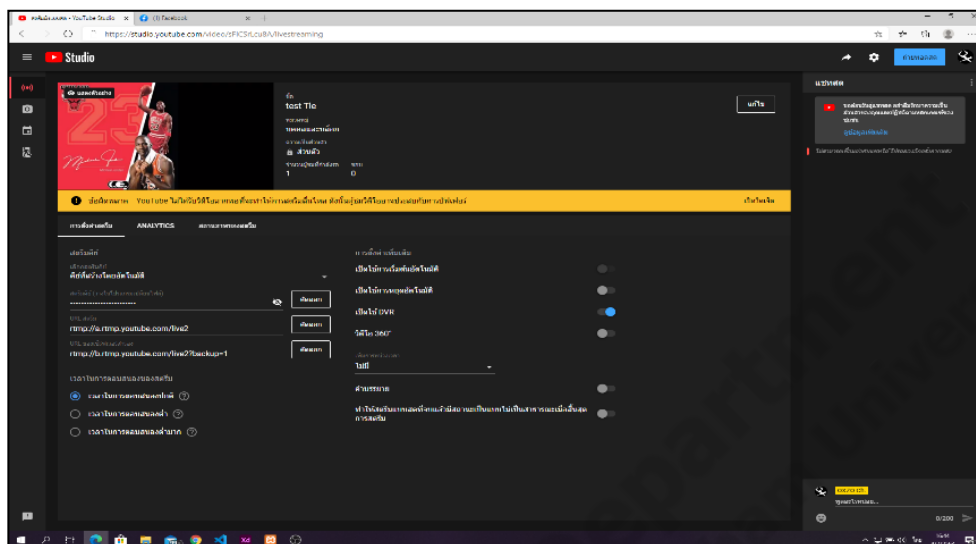
ภาพประกอบที่ 2.14 หน้าจอแสดงการตั้งค่าโปรแกรม

9. คลิกเริ่มสตรีม



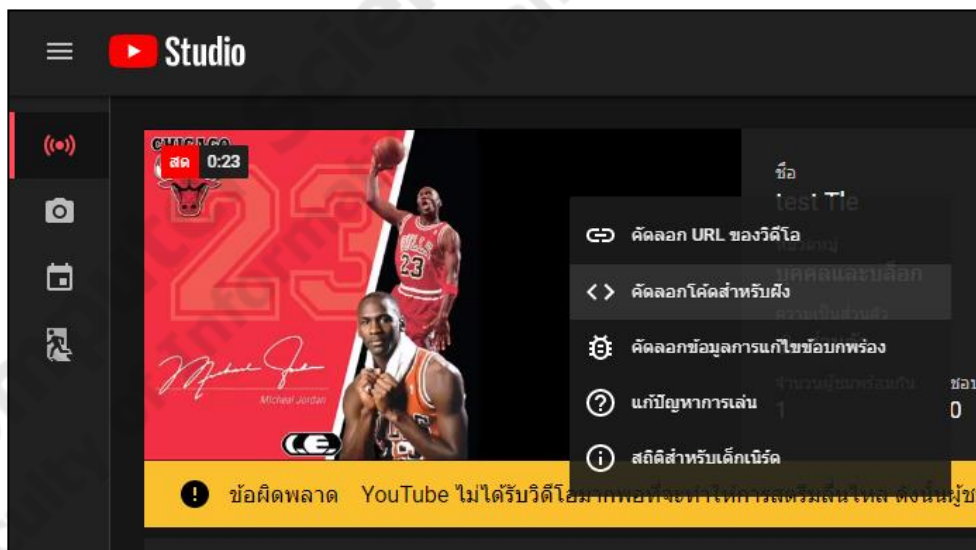
ภาพประกอบที่ 2.15 หน้าจอแสดงการสตรีม

10. เข้ามาที่หน้า YouTube คลิกที่ ถ่ายทอดสด



ภาพประกอบที่ 2.16 หน้าจอแสดงการสตรีมบน Youtube

11. คลิกขวาที่วิดีโอที่ถ่ายทอดสดดังนี้ แล้วคลิก คัดลอกโค้ดสำหรับฝัง



ภาพประกอบที่ 2.17 หน้าจอแสดงตัวเลือกการสตรีม

12. นำโค้ดที่คัดลอกมาวาง

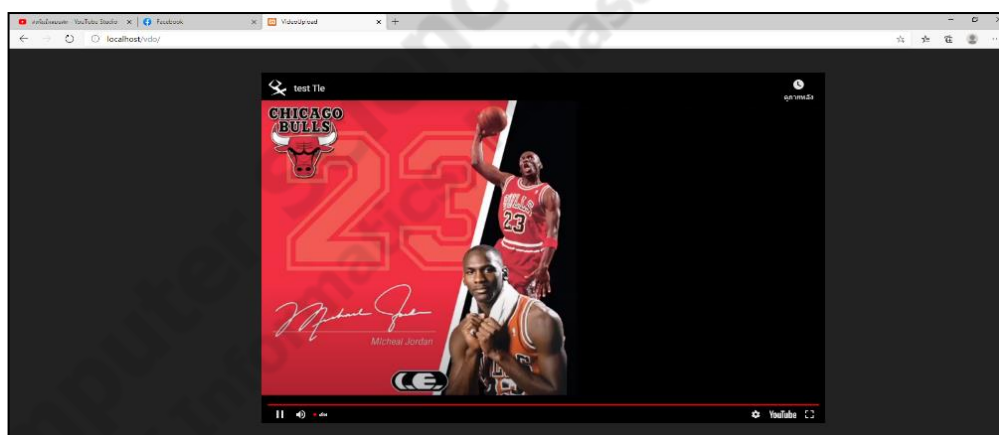
```

1 <html>
2 <head>
3     <meta charset="UTF-8">
4     <title>VideoUpload</title>
5     <link rel="stylesheet" href="styler.css">
6 </head>
7 <body>
8     <div id="box">
9         <form action="upload.php" enctype="multipart/form-data" method="post">
10
11             <iframe width="1880" height="720" src="https://www.youtube.com/embed/sFICsrLcu8A" frameborder="0"
12                 allow="accelerometer; autoplay; encrypted-media; gyroscope; picture-in-picture" allowfullscreen></iframe>
13         </form>
14     </div>
15 </body>
16 </html>

```

ภาพประกอบที่ 2.18 หน้าจอแสดงโค้ดสตรีม

13. เริ่มรับชมได้



ภาพประกอบที่ 2.19 หน้าจอแสดงวิดีโอสตรีม

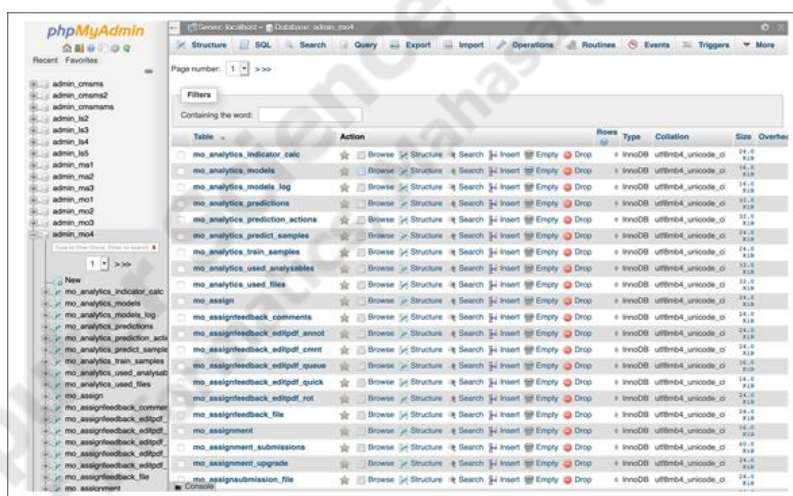
2.1.2 ฐานข้อมูล

พีเอชพีมายแอดมิน (phpMyAdmin) [18] คือ โปรแกรมที่ถูกพัฒนาโดยใช้ภาษาพีเอชพี (PHP) เพื่อใช้ในการบริหารจัดการฐานข้อมูลมายเอชคิวแอล (MySQL) แทนการคีย์คำสั่ง เนื่องจากถ้าใช้ฐานข้อมูลที่เป็นมายเอชคิวแอล บางครั้งจะมีความยุ่งยากในการทำงาน ดังนั้นจึงมีเครื่องมือในการจัดการฐานข้อมูลขึ้นมาเพื่อให้สามารถจัดการตัว ดีบีเอ็มเอช (DBMS) ที่เป็นมายเอชคิวแอล ได้ง่ายและสะดวกยิ่งขึ้น โดยพีเอชพีมายแอดมินก็ถือเป็นเครื่องมือชนิดหนึ่งในการจัดการ

พีเอชพีมายแอดมิน เป็นส่วนต่อประสานที่สร้างโดยภาษาพีเอชพี ซึ่งใช้จัดการฐานข้อมูลมายเอชคิวแอล ผ่านเว็บเบราว์เซอร์โดยสามารถที่จะทำการสร้างฐานข้อมูลใหม่ หรือทำการสร้างเทเบิลใหม่ๆ และยังมีฟังก์ชันที่ใช้สำหรับการทดสอบการควิรี่ข้อมูลด้วยภาษาเอสคิวแอล พร้อมกันนั้นยังสามารถทำการเพิ่ม ลบ แก้ไข หรือแม้กระทั่งใช้คำสั่งต่างๆ เหมือนกับการใช้ภาษาเอสคิวแอลในการสร้างตารางข้อมูลโดยตรง พีเอชพีมายแอดมิน ตัวนี้จะทำงานบนเว็บเซิร์ฟเวอร์เป็น พีเอชพีแอปพลิเคชัน ที่ใช้ควบคุมจัดการ มายเอชคิวแอลเซิร์ฟเวอร์

ความสามารถของ พีเอชพีมายแอดมิน (phpMyAdmin) อธิบายได้ดังนี้

1. สร้างและลบดาต้าเบส (Database)
2. สร้างและจัดการ Table เช่น แทรกเรคคอร์ด (record), ลบเรคคอร์ด (record), แก้ไขเรคคอร์ด (record), ลบเทเบิล (Table), แก้ไขฟิว (field)
3. โหลดเท็กซ์ไฟล์เข้าไปเก็บเป็นข้อมูลในตารางได้
4. หาผลสรุป (Query) ด้วยคำสั่ง SQL



ภาพประกอบที่ 2.20 ตัวอย่างหน้าเว็บพีเอชพีมายแอดมิน (phpMyAdmin)

ที่มา : <https://www.phpmyadmin.net>

ตารางที่ 2.1 ตารางแสดงการเปรียบเทียบระบบงานที่เกี่ยวข้อง

ฟังก์ชันการทำงาน	Wbi	Google Classroom	Online Education System
เอกสารประกอบการสอน	/	/	/
Video	-	/	/
Live	-	/	/
การจัดการข้อสอบแบบ Multiple Choice และแบบการเติมคำ	-	/	/
การส่งการบ้าน	/	/	/
การส่งงาน	/	/	/
แจ้งข่าวสาร	/	/	/
แจ้งคะแนน	/	/	/
ปฏิทินกำหนดการ	-	/	/
การเช็คชื่อเข้าเรียน	/	/	/
สืบค้นข้อมูล	-	-	/
การจัดทำรายงาน	-	-	/

2.2 การประมวลผลภาษาธรรมชาติ (Natural Language Processing : NLP)

การประมวลผลภาษาธรรมชาติ หรือ NLP [19] เป็นสาขาย่อยของภาษาศาสตร์วิทยาการคอมพิวเตอร์และปัญญาประดิษฐ์ที่เกี่ยวข้องกับการโต้ตอบระหว่างคอมพิวเตอร์และภาษามนุษย์ (ธรรมชาติ) โดยการใช้โปรแกรมคอมพิวเตอร์ในการประมวลผลและวิเคราะห์ข้อมูลภาษาธรรมชาติจำนวนมาก เช่น การรู้จำเสียงพูด เทคนิคที่ใช้ในการประมวลผลภาษาธรรมชาติที่ผู้วิจัยได้นำมาใช้ในงานวิจัยมีดังนี้

2.2.1 การตัดคำ (Words segmentation)

การตัดคำ [20] เป็นการแบ่งส่วนคำหรือการโทเค็น (Tokenization) คำถือเป็นงานพื้นฐานในการประมวลผลภาษาธรรมชาติ (NLP) สำหรับการตัดคำภาษาไทยนั้นจะมีความยุ่งยากมากกว่า เช่นเดียวกับภาษาจีน ญี่ปุ่น และเกาหลี เป็นภาษาที่ไม่มีกรแบ่งกลุ่ม มักเขียนเป็นคำเรียงและต่อเนื่อง

หากถ้าเป็นการตัดคำในภาษาอังกฤษ ฝรั่งเศส หรือสเปน ซึ่งเป็นภาษาที่มีรากฐานมาจากภาษาละตินมักจะไม่มีปัญหาและทำการตัดคำได้ง่ายกว่าภาษาที่ไม่ได้มีรากฐานมาจากภาษาละติน เพราะ

ภาษาเหล่านี้มักมีเครื่องหมายคั่นคำ (Delimiting) เช่น ช่องว่าง (Spaces) เซมิโคลอน (Semi-Colon) จุลภาค (Comma) เครื่องหมายคำพูด (Quote) และจุด (Period)

สำหรับภาษาที่ไม่แบ่งส่วนมีการศึกษาเทคนิคเพื่อแก้ไขปัญหาการแบ่งคำ สามารถแบ่งออกเป็น 2 วิธี คือ ตามพจนานุกรม (Dictionary-Based: DCB) และการเรียนรู้ด้วยเครื่อง (Machine Learning Based: MLB)

จากการศึกษาพบว่าในการตัดคำภาษาไทยที่มีประสิทธิภาพและเป็นที่น่าพอใจ มักจะเป็นการตัดคำที่ได้จากวิธีการตัดคำตามพจนานุกรม โดยวิธีการตัดคำตามพจนานุกรม จะเป็นการใช้ชุดคำศัพท์จากพจนานุกรมในการแยกและการแบ่งข้อความ ขั้นตอนการแยกจะค้นหาชุดของอักขระตามพจนานุกรมเพื่อค้นหาคำที่ตรงกัน ประสิทธิภาพการตัดคำตามพจนานุกรมจะขึ้นอยู่กับคุณภาพและขนาดของคำที่ต้องการตัด พจนานุกรมที่ใช้มีคำที่ค่อนข้างง่ายและตรงไปตรงมาซึ่งมักจะมีปัญหา เช่น ปัญหาคำที่ไม่รู้จัก เป็นคำที่ไม่พบในพจนานุกรม หรือปัญหาความกำกวมของคำที่ทำการตัด ปัญหาเหล่านี้สามารถแก้ไขด้วยเทคนิคต่างๆ เช่น

2.2.1.1 เทคนิคการตัดคำภาษาไทยด้วยพจนานุกรมแบบเปรียบเทียบคำที่ยาวที่สุด (Longest Matching) [21]

ในภาษาไทยมีการเขียนตัวอักษรโดยไม่มีขอบเขตของคำที่ชัดเจน โดยคำที่เขียนมักขึ้นอยู่กับบริบทซึ่งมีหลากหลายวิธีในการแบ่งเป็นคำ ยกตัวอย่างเช่น

คำว่า “อาจอง” สามารถแบ่งออกเป็น

“อา - จอง” หรือ “อาจ - อง”

หรือคำว่า “นั่งตากลม” สามารถแบ่งออกเป็น

“นั่ง - ตา - ลม” หรือ “นั่ง - ตาก - ลม”

จากตัวอย่างดังกล่าวจะเห็นได้ว่ามีความซับซ้อนในการระบุคำจึงได้มีการนำเทคนิคการเปรียบเทียบคำที่ยาวที่สุดมาใช้ในการแบ่งคำ การทำงานจะอ่านข้อความจากซ้ายไปขวาแล้วนำคำที่ได้ไปเปรียบเทียบกับคำในพจนานุกรมและเลือกคำที่ยาวที่สุด อย่างไรก็ตามคำที่ยาวที่สุดที่ได้อาจไม่สอดคล้องกับความเป็นจริง

2.2.1.2 เทคนิคการตัดคำภาษาไทยด้วยพจนานุกรมแบบที่สอดคล้องมากที่สุด (Maximal Matching) [21]

เทคนิคการตัดคำภาษาไทยด้วยพจนานุกรมแบบที่สอดคล้องมากที่สุด [12] เป็นอีกเทคนิคที่ใช้ในการแก้ไขข้อบกพร่องของเทคนิคการตัดคำภาษาไทยด้วยพจนานุกรมแบบเปรียบเทียบคำที่ยาวที่สุด โดยเป็นวิธีการตัดคำที่ทำให้คำที่ได้มีจำนวนคำที่น้อยที่สุด และเลือกตัดคำที่เป็นไปได้ทั้งหมดของประโยคนั้นๆ ก่อน จากนั้นจึงจะทำการเลือกรูปแบบที่เหมาะสมที่สุดโดยการพิจารณาจาก จำนวน

คำที่ตัดได้ และรูปแบบของประโยคที่มีจำนวนคำน้อยที่สุดจะถูกคัดเลือกให้เป็นรูปแบบที่สอดคล้องที่สุด

2.2.2 การสร้างตัวแทนข้อความ (Text Representation) [22]

เนื่องจากปัจจุบันคอมพิวเตอร์ไม่สามารถจำแนกหมวดหมู่ของข้อความที่เป็นภาษาธรรมชาติได้โดยตรง จึงต้องมีการจำลองข้อความให้อยู่ในรูปแบบที่คอมพิวเตอร์สามารถอ่านเข้าใจ และสามารถเรียนรู้ได้ โดยการสร้างตัวแทนข้อความที่นิยมใช้ก็คือการจำลองเอกสารให้อยู่ในแบบจำลองปริภูมิเวกเตอร์ (Vector Space Model: VSM)

การสร้างตัวแทนข้อความในรูปแบบของเวกเตอร์ เป็นหนึ่งในวิธีการแทนเอกสารที่ไม่มีโครงสร้าง (Unstructured Text Document) โดยการแบ่งข้อความให้อยู่ในรูปของถุคำ (Bag-of-Words: BOW) ซึ่งจะเก็บอยู่ในรูปแบบของเวกเตอร์ โดยกำหนดให้เอกสารแต่ละฉบับเปรียบเสมือนเวกเตอร์ของคำ ขนาดของเวกเตอร์ขึ้นอยู่กับจำนวนของคำที่ปรากฏอยู่ในเอกสาร โดยใช้วิธีการหาค่าความถี่ของคำ หรือเรียกว่าการหาค่าน้ำหนักของคำ (Term Weighting) มักแทนค่าด้วยเลขฐานสองคือ จะมีค่าตั้งแต่ 0 ถึง 1 หากค่าเป็น 0 หมายความว่าไม่มีคำนั้นอยู่ในเอกสาร และถ้าหากค่าเป็น 1 ก็หมายความว่าพบคำนั้นในเอกสาร ซึ่งจะได้รูปแบบที่มีลักษณะของการแทนความสัมพันธ์ระหว่างคำ (Words: W) และเอกสาร (Documents: D) ด้วยเวกเตอร์ 2 มิติ

	w_1	w_2	w_3	w_4	...	w_i
D_1	$w_{1,1}$	$w_{1,2}$	$w_{1,3}$	$w_{1,4}$...	$w_{1,i}$
D_2	$w_{2,1}$	$w_{2,2}$	$w_{2,3}$	$w_{2,4}$...	$w_{2,i}$
...
D_j	$w_{j,1}$	$w_{j,2}$	$w_{j,3}$	$w_{j,4}$...	$w_{j,i}$

ภาพประกอบที่ 2.21 ตัวอย่างของเวกเตอร์สเปซในรูปแบบเมทริกซ์

2.2.3 การวิเคราะห์ความคล้ายคลึง (Similarity Analysis) [23]

การวิเคราะห์ความคล้ายคลึงเป็นวิธีการในการวิเคราะห์ว่าวัตถุ (Object) 2 ชิ้นมีความคล้ายคลึงหรือความสอดคล้องกันหรือไม่ เทคนิคที่ใช้ในการวิเคราะห์ความคล้ายคลึงมีหลายเทคนิคที่เป็นที่นิยม เช่น ความคล้ายคลึงกันของโคไซน์ (Cosine Similarity) เป็นต้น

1. ความคล้ายคลึงกันของโคไซน์ (Cosine Similarity)

ความคล้ายคลึงกันของโคไซน์ [21] เป็นการวัดความคล้ายคลึงกันที่ใช้กันทั่วไปสำหรับเวกเตอร์ที่มีค่าความจริงซึ่งใช้ในการดึงข้อมูล เพื่อให้คะแนนความคล้ายคลึงกันของเอกสารในแบบจำลองปริภูมิเวกเตอร์ ซึ่งสามารถคำนวณดังสมการที่ (2.6)

$$\text{similarity} = \cos(\theta) = \frac{A \times B}{\|A\| \|B\|} = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}} \quad (2.1)$$

Computer Science Department
Faculty of Informatics, Maharakham University