การศึกษาคุณลักษณะโครงสร้างของเอกสาร HTML จากเว็บไซต์ที่ให้ข้อมูลยา

สำนักหอสมุดกลาง

โดย
นางสาววรรณทิพย์ แสงทินทอง

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาเภสัชศาสตรมหาบัณฑิต
สาขาวิชาสารสนเทศศาสตร์ทางสุขภาพ
ภาควิชาสารสนเทศศาสตร์ทางสุขภาพ
บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร
ปีการศึกษา 2555
ลิขสิทธิ์ของบัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร

การศึกษาคุณลักษณะโครงสร้างของเอกสาร HTML จากเว็บไซต์ที่ให้ข้อมูลยา

สำนักหอสมุดกลาง

โดย
นางสาววรรณทิพย์ แสงทินทอง

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาเภสัชศาสตรมหาบัณฑิต
สาขาวิชาสารสนเทศศาสตร์ทางสุขภาพ
ภาควิชาสารสนเทศศาสตร์ทางสุขภาพ
บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร
ปีการศึกษา 2555
ลิขสิทธิ์ของบัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร

A STUDY ON STRUCTURE CHARACTERISTICS OF HTML DOCUMENTS FROM DRUG INFORMATIONS WEBSITES

สำนักหอสมุดกลาง

By
Miss Wantip Saengthinthong

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree
Master of Pharmacy Program in Health Informatics
Department of Health-Related Informatics
Graduate School, Silpakorn University
Academic Year 2012
Copyright of Graduate School, Silpakorn University

บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร อนุมัติให้วิทยานิพนธ์เรื่อง “การศึกษาคุณลักษณะโครงสร้างของเอกสาร HTML จากเว็บไซต์ที่ให้ข้อมูลยา” เสนอโดย นางสาววรรณทิพย์ แสงทินทอง เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาเภสัชศาสตรมหาบัณฑิต สาขาวิชาสารสนเทศศาสตร์ทางสุขภาพ

.....
(ผู้ช่วยศาสตราจารย์ ดร.ปานใจ ธารศนวงศ์)
คณบดีบัณฑิตวิทยาลัย
วันที่.....เดือน..... พ.ศ.....

สำนักหอสมุดกลาง

อาจารย์ที่ปรึกษาวิทยานิพนธ์
เภสัชกรหญิง อาจารย์ ดร. ลาวัลย์ ศรัทธาพุทธ
เภสัชกร รองศาสตราจารย์ ดร. วีรยุทธ์ เลิศนที

คณะกรรมการตรวจสอบวิทยานิพนธ์

..... ประธานกรรมการ
(เภสัชกร ผู้ช่วยศาสตราจารย์ ดร. สรวง รุ่งประกายพรรณ)
...../...../.....

..... กรรมการ
(เภสัชกรหญิง อาจารย์ ดร. ลาวัลย์ ศรัทธาพุทธ)
...../...../.....

..... กรรมการ
(เภสัชกร รองศาสตราจารย์ ดร. วีรยุทธ์ เลิศนที)
...../...../.....

..... กรรมการ
(ดร. อิทธิพันธ์ เมธเศรษฐ)
...../...../.....

52363202 : สาขาวิชาสารสนเทศศาสตร์ทางสุขภาพ
คำสำคัญ : เอกสารเอชทีเอ็มแอล, เว็บไซต์, ข้อมูลยา

วรรณทิพย์ แสงทินทอง : การศึกษาคุณลักษณะโครงสร้างของเอกสาร HTML จากเว็บไซต์ที่ให้ข้อมูลยา. อาจารย์ที่ปรึกษาวิทยานิพนธ์ : ภญ.อ.ดร.ลาวัลย์ ศรัทธาพุทธ และ ภก.รศ.ดร. วีรยุทธ์ เลิศนที. 117 หน้า

เนื่องจากข้อมูลด้านยาจากเว็บไซต์มีจำนวนเพิ่มมากขึ้น ทำให้โปรแกรมค้นหาเข้ามามีบทบาทสำคัญในการเป็นเครื่องมือค้นหา งานวิจัยนี้จึงมีวัตถุประสงค์เพื่อศึกษาคุณลักษณะโครงสร้างของเอกสาร HTML ที่เป็นปัจจัยต่อคุณภาพของเว็บไซต์ที่ให้ข้อมูลยาที่เป็นภาษาอังกฤษ และมีผลต่อการปรากฏในโปรแกรมค้นหา โดยทำการศึกษาเว็บไซต์ที่ให้ข้อมูลยาจำนวน 35 เว็บไซต์ ซึ่งได้จากการใช้โปรแกรมค้นหา 3 ชนิด คือ Google, Yahoo และ MSN คำสำคัญที่ใช้ในการสืบค้นแบ่งออกเป็น 2 กลุ่มคือ คำสืบค้นทั่วไปและคำสืบค้นที่เป็นชื่อเฉพาะทางยา จากนั้นศึกษาปัจจัยจำนวน 31 ตัวแปร ผลการวิเคราะห์ทางสถิติพบว่ามี 4 ตัวแปรที่ส่งผลต่ออันดับของเว็บไซต์และ 7 ตัวแปรที่ส่งผลต่อคุณภาพของเว็บไซต์ จึงสรุปได้ว่าคุณลักษณะเอกสาร HTML ที่ทำการศึกษามีความเกี่ยวข้องกับการจัดอันดับของเว็บไซต์และคุณภาพของเว็บไซต์ ดังนั้นในการออกแบบเว็บไซต์ที่ให้ข้อมูลยาจึงควรคำนึงถึงคุณลักษณะดังกล่าวเพื่อให้ได้เว็บไซต์ข้อมูลยาที่มีคุณภาพและปรากฏในอันดับต้นเมื่อค้นหาผ่านโปรแกรมค้นหา

สาขาวิชาสารสนเทศศาสตร์ทางสุขภาพ บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร
ลายมือชื่อนักศึกษา..... ปีการศึกษา 2555
ลายมือชื่ออาจารย์ที่ปรึกษาวิทยานิพนธ์ 1..... 2.....

52363202 : MAJOR : HEALTH INFORMATICS
KEY WORD : HTML DOCUMENTS / WEBSITES / DRUG INFORMATION
WANTIP SAENGTHINTHONG : A STUDY ON STRUCTURE CHARACTERISTICS OF HTML DOCUMENTS FROM DRUG INFORMATIONS WEBSITES. THESIS ADVISOR : LAWAN SRATTHAPHUT, Ph.D. AND ASSOC. PROF. VERAYUTH LERTNATTEE, Ph.D. 117 pp.

Due to the increasing of drug information from websites, search engines become important as a searching tool. The objectives of this research are to investigate the factors of HTML documents structure which affect drug information content quality and search engine result pages (SERP) rank. Thirty-five websites that provide the drug information in English query by using Google, Yahoo and MSN as search engines were studied. Keywords used in the search query are divided into two groups: the general keywords and the specific name of the drug. Then, 31 factors were investigated. Statistical analysis of the results obtained by the selected websites has been also carried out. The results illustrated that there were 4 factors which affect the rank of websites and there were 7 factors which affect the quality of websites. In conclusion, the investigated characteristics of HTML documents in drug information websites are related with ranking and quality and these should be concerned in the designing drug information websites to achieve drug information content quality and SERP rank.

Department of Health-Related Informatics Graduate School, Silpakorn University
Student's signature ..... Academic Year 2012
Thesis Advisor's signature 1..... 2.....

กิตติกรรมประกาศ

การทำวิทยานิพนธ์นี้ ข้าพเจ้าได้รับความช่วยเหลือ ความสนับสนุนและคำปรึกษาแนะนำอย่างดียิ่งจากผู้ทรงคุณวุฒิหลายท่าน ข้าพเจ้าขอขอบพระคุณ ภญ.อ.ดร.ลาวัลย์ ศรัทธาพุทธ ที่ให้ความรู้ คำปรึกษาแนะนำสั่งสอน และคอยติดตามการทำวิทยานิพนธ์ แก้ไขและปรับปรุงกระทั่งวิทยานิพนธ์ฉบับนี้เสร็จสมบูรณ์

ขอขอบพระคุณ ภก.รศ.ดร.วีรยุทธ เลิศนที ที่ให้คำปรึกษาและแนะนำ รวมทั้งช่วยตรวจทานและแก้ไขเล่มวิทยานิพนธ์เพื่อให้วิทยานิพนธ์นี้มีความถูกต้องและสมบูรณ์ยิ่งขึ้น

ขอขอบพระคุณนายณรงค์ฤทธิ์ สุขสาร ภญ.อ.ชนิสรา ลือวิพันธ์ คณาจารย์รวมทั้งพี่ๆ น้องๆ ภาควิชาสารสนเทศสาสตร์ทางสุขภาพ และพี่ๆ กลุ่มงานเภสัชกรรม รพ.บ้านโป่ง ทุกท่านที่คอยช่วยเหลือ ให้คำแนะนำในการทำวิทยานิพนธ์และเป็นกำลังใจในการทำวิทยานิพนธ์เป็นอย่างดี

ขอขอบพระคุณครอบครัวของข้าพเจ้า ที่คอยให้การสนับสนุนและเป็นกำลังใจในการทำวิทยานิพนธ์นี้จนลุล่วงได้ด้วยดี

สุดท้ายโครงการวิทยานิพนธ์นี้จะสำเร็จลงไม่ได้หากไม่ได้รับการสนับสนุนจาก 2 ทุนคือ ทุนวิจัยในโครงการวิจัยเรื่อง “การพัฒนาตัวสืบค้นข้อมูลยาผ่านเครือข่ายอินเทอร์เน็ตสำหรับบุคลากรสาธารณสุขและบุคคลทั่วไป” SURDI 52/01/23, SURDI 53/01/12 จากสถาบันวิจัยและพัฒนามหาวิทยาลัยศิลปากร และทุนอุดหนุนการทำวิจัยจากงบประมาณแผ่นดิน (หมวดเงินอุดหนุนทั่วไป) ประจำปีงบประมาณ 2555 ข้าพเจ้าขอขอบพระคุณผู้เกี่ยวข้องในหน่วยงานดังกล่าวทุกท่านที่ให้การสนับสนุนในทุกๆ ด้านจนกระทั่งการศึกษานี้สำเร็จลงด้วยดี

สารบัญ

		หน้า
บทคัดย่อภาษาไทย		ง
บทคัดย่อภาษาอังกฤษ		จ
กิตติกรรมประกาศ		ฉ
สารบัญตาราง		ฌ
สารบัญรูปภาพ		ฎ
บทที่		
1	บทนำ	1
	ความเป็นมาและความสำคัญของปัญหา	1
	ความมุ่งหมายและวัตถุประสงค์ของการศึกษา	4
	ประโยชน์ที่คาดว่าจะได้รับ	4
	ขอบเขตของการศึกษา	5
2	วรรณกรรมที่เกี่ยวข้อง	6
	ภาษา HTML และโครงสร้างของเอกสาร HTML	6
	คุณลักษณะของเว็บไซต์และองค์ประกอบของเว็บไซต์ที่ดี	16
	โปรแกรมค้นหา และหลักการทำ SEO	20
	ความน่าเชื่อถือของข้อมูลสุขภาพบนอินเทอร์เน็ตและการประเมินคุณภาพเว็บไซต์	31
3	วิธีดำเนินการวิจัย	33
	อุปกรณ์และเครื่องมือที่ใช้ในการทดลอง	33
	วิธีการศึกษา	34
	การคัดเลือกเว็บไซต์	34
	การประเมินคุณภาพของเว็บไซต์	35
	การจัดกลุ่มของเว็บไซต์	35
	การเก็บรวบรวมข้อมูลตัวแปรต่างๆ	35
	การวิเคราะห์คุณลักษณะของเว็บไซต์ด้วยสถิติเชิงพรรณนา	38
	การเปรียบเทียบค่าเฉลี่ยของตัวแปรด้วยวิธี Independent t-test	38
4	ผลการวิเคราะห์ข้อมูล	42
	การคัดเลือกเว็บไซต์	42
	การประเมินคุณภาพเว็บไซต์	43
	การจัดกลุ่มเว็บไซต์	45
	การเก็บรวบรวมข้อมูล	48
	การศึกษาด้วยสถิติเชิงพรรณนา	48

บทที่ หน้า
การศึกษาโครงสร้างของเอกสาร HTML และคุณลักษณะต่างๆ จากเว็บไซต์ที่อยู่อันดับต้น 48
การศึกษาโครงสร้างของเอกสาร HTML และคุณลักษณะต่างๆ จากเว็บไซต์ที่อยู่อันดับท้าย 63
การศึกษาโครงสร้างของเอกสาร HTML และคุณลักษณะต่างๆ จากเว็บไซต์ที่ที่มีคะแนนการประเมินตั้งแต่ 60 เปอร์เซ็นต์ขึ้นไป (เว็บไซต์กลุ่ม A) 72
การศึกษาโครงสร้างของเอกสาร HTML และคุณลักษณะต่างๆ จากเว็บไซต์ที่ที่มีคะแนนการประเมินต่ำกว่า 60 เปอร์เซ็นต์ (เว็บไซต์กลุ่ม B) 83
การเปรียบเทียบค่าเฉลี่ยของตัวแปรด้วยวิธี Independent t-test 94
การศึกษาตัวแปรที่มีผลต่อการจัดอันดับของเว็บไซต์บนหน้าแสดงผล 94
การศึกษาตัวแปรที่มีผลต่อคุณภาพของเว็บไซต์ 97
5 สรุป อภิปรายผลการศึกษาและข้อเสนอแนะ 100
สรุปและอภิปรายผลการศึกษา 100
ข้อเสนอแนะ 105
รายการอ้างอิง 106
ภาคผนวก 109
แบบประเมิน DISCERN 110
Critical values of the F-distribution 115
ประวัติผู้วิจัย 117

สารบัญตาราง

หน้า

ตารางที่		
1	ผลการคัดเลือกเว็บไซต์	42
2	ผลการประเมินคุณภาพเว็บไซต์ด้วย DISCERN	43
3	เว็บไซต์ที่มีคะแนนการประเมินคุณภาพตั้งแต่ 60 เปอร์เซ็นต์ขึ้นไป	46
4	เว็บไซต์ที่มีคะแนนการประเมินคุณภาพน้อยกว่า 60 เปอร์เซ็นต์	47
5	ปริมาณ Tag เฉลี่ยต่อหน้าเว็บเพจที่ปรากฏบนเอกสาร HTML ของเว็บไซต์อันดับต้น	50
6	คุณลักษณะของเอกสาร HTML ในส่วนของจุดเชื่อมโยงภายนอกและจุดเชื่อมโยงภายในของเว็บไซต์อันดับต้น	53
7	แสดงขนาดของหน้าเว็บเพจเฉลี่ยในแต่ละเว็บไซต์อันดับต้น, เวลาที่ใช้ในการดาวน์โหลด (Connection Rate 1.44 Mbps.), ขนาดภาพที่พบภายในเว็บเพจเฉลี่ย, ขนาดของ CSS ที่พบภายในเว็บเพจ, ภาพที่ไม่สามารถเรียกดูได้ในเว็บไซต์อันดับต้น (Dead Images) และจำนวนภาพรวมภายในเว็บไซต์อันดับต้น	55
8	ประเภทของเอกสาร HTML และการเข้ารหัสภาษา ของเว็บไซต์อันดับต้น	58
9	HTML error, HTML warning และ CSS error ของเว็บไซต์อันดับต้น	60
10	ค่า Page Rank เฉลี่ยของแต่ละเว็บไซต์อันดับต้น	62
11	ปริมาณ Tag เฉลี่ยต่อหน้าเว็บเพจที่ปรากฏบนเอกสาร HTML ของเว็บไซต์อันดับท้าย	64
12	คุณลักษณะของเอกสาร HTML ในส่วนของจุดเชื่อมโยงนอกและจุดเชื่อมโยงภายในของเว็บไซต์อันดับท้าย	66
13	แสดงขนาดของหน้าเว็บเพจเฉลี่ยในแต่ละเว็บไซต์อันดับต้น, เวลาที่ใช้ในการดาวน์โหลด (Connection Rate 1.44 Mbps.), ขนาดภาพที่พบภายในเว็บเพจเฉลี่ย, ขนาดของ CSS ที่พบภายในเว็บเพจ, ภาพที่ไม่สามารถเรียกดูได้ในเว็บไซต์อันดับต้น (Dead Images) และจำนวนภาพรวมภายในเว็บไซต์อันดับท้าย	67
14	ประเภทของเอกสาร HTML และการเข้ารหัสภาษา ของเว็บไซต์อันดับท้าย	69
15	HTML error, HTML warning และ CSS error ของเว็บไซต์อันดับท้าย	70
16	ค่า Page Rank เฉลี่ยของแต่ละเว็บไซต์อันดับท้าย	71
17	ปริมาณ Tag เฉลี่ยต่อหน้าเว็บเพจที่ปรากฏบนเอกสาร HTML ของเว็บไซต์กลุ่ม A	73
18	คุณลักษณะของเอกสาร HTML ในส่วนของจุดเชื่อมโยงภายนอกและจุดเชื่อมโยงภายในของเว็บไซต์กลุ่ม A	76

ตารางที่ หน้า

19 แสดงขนาดของหน้าเว็บเพจเฉลี่ยในแต่ละเว็บไซต์อันดับต้น, เวลาที่ใช้ในการดาวน์โหลด (Connection Rate 1.44 Mbps.), ขนาดภาพที่พบภายในเว็บเพจเฉลี่ย, ขนาดของ CSS ที่พบภายในเว็บเพจ, ภาพที่ไม่สามารถเรียกดูได้ในเว็บไซต์อันดับต้น (Dead Images) และจำนวนภาพรวมภายในเว็บไซต์กลุ่ม A..... 77

20 ประเภทของเอกสาร HTML และการเข้ารหัสภาษาของเว็บไซต์กลุ่ม A..... 80

21 HTML error, HTML warning และ CSS error ของเว็บไซต์กลุ่ม A..... 81

22 ค่า Page Rank เฉลี่ยของแต่ละเว็บไซต์กลุ่ม A..... 82

23 ปริมาณ Tag เฉลี่ยต่อหน้าเว็บเพจที่ปรากฏบนเอกสาร HTML ของเว็บไซต์กลุ่ม B..... 84

24 คุณลักษณะของเอกสาร HTML ในส่วนของจุดเชื่อมโยงภายนอกและจุดเชื่อมโยงภายในของเว็บไซต์กลุ่ม B..... 87

25 แสดงขนาดของหน้าเว็บเพจเฉลี่ยในแต่ละเว็บไซต์อันดับต้น, เวลาที่ใช้ในการดาวน์โหลด (Connection Rate 1.44 Mbps.), ขนาดภาพที่พบภายในเว็บเพจเฉลี่ย, ขนาดของ CSS ที่พบภายในเว็บเพจ, ภาพที่ไม่สามารถเรียกดูได้ในเว็บไซต์อันดับต้น (Dead Images) และจำนวนภาพรวมภายในเว็บไซต์กลุ่ม B..... 88

26 ประเภทของเอกสาร HTML และการเข้ารหัสภาษาของเว็บไซต์กลุ่ม B..... 91

27 HTML error, HTML warning และ CSS error ของเว็บไซต์กลุ่ม B..... 92

28 ค่า Page Rank เฉลี่ยของแต่ละเว็บไซต์กลุ่ม B..... 93

29 ค่าสถิติ Independent T-test ระหว่างตัวแปรต่างๆ ที่ทำการศึกษากับการจัดอันดับของเว็บไซต์บนหน้าแสดงผล..... 95

30 ค่าสถิติ Independent T-test ของตัวแปรต่างๆ ต่อคุณภาพของเว็บไซต์..... 98

31 แบบประเมิน DISCERN..... 110

32 ค่า Critical values of the F-distribution..... 115

สารบัญรูป

รูปที่		หน้า
1	ตัวอย่างโครงสร้างเอกสาร HTML	8
2	Web Page Analyzer - 0.98	36
3	Markup Validation Service	36
4	CSS Validation Service	37
5	โปรแกรม SeoQuake SEO extension version 2.5.9	37
6	การหาค่า PR ด้วย Google toolbar	38

สำนักหอสมุดกลาง

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

สำนักหอสมุดกลาง

ปัจจุบัน จำนวนผู้ใช้อินเทอร์เน็ตทั่วโลกโดยประมาณ 2.095 พันล้านคน หรือร้อยละ 30.2 ของประชากรทั่วโลก โดยเมื่อเปรียบเทียบในทวีปต่างๆ พบว่าทวีปที่มีผู้ใช้อินเทอร์เน็ตมากที่สุด คือ เอเชีย โดยคิดเป็นร้อยละ 44.0 ของผู้ใช้อินเทอร์เน็ตทั้งหมด เช่นเดียวกันกับจำนวนผู้ใช้อินเทอร์เน็ตในประเทศไทยที่มีการเปลี่ยนแปลงเพิ่มขึ้นอย่างรวดเร็ว พบว่ามีผู้ใช้อินเทอร์เน็ต 18.3 ล้านคน คิดเป็นร้อยละ 27.4 ของประชากรในประเทศ (Internet World Stats, 2009) ทั้งนี้เนื่องมาจากอินเทอร์เน็ตเป็นแหล่งที่ให้ข้อมูลจำนวนมากและหลากหลายรูปแบบ ทำให้ผู้ใช้สามารถสืบค้นข้อมูลต่างๆ ได้อย่างสะดวก รวดเร็ว ไม่มีข้อจำกัดในเรื่องของเวลาและสถานที่สามารถทำที่ใดก็ได้ ข้อมูลมีความทันสมัยเนื่องจากสามารถปรับแก้ไขได้ง่ายและสามารถทำได้ตลอดเวลา อีกทั้งยังประหยัดค่าใช้จ่ายกว่าการค้นหาข้อมูลที่อยู่ในรูปแบบสิ่งพิมพ์ การประยุกต์ใช้อินเทอร์เน็ตในปัจจุบันทำได้หลากหลาย อาทิเช่น ไปรษณีย์อิเล็กทรอนิกส์ หรือ อีเมล์ (e-Mail), สนทนา (Chat), อ่านหรือแสดงความคิดเห็นในเว็บบอร์ด, การติดตามข่าวสาร, การสืบค้นข้อมูล / การค้นหาข้อมูล, การชมหรือซื้อสินค้าออนไลน์, การดาวโหลด เกม เพลง ไฟล์ข้อมูล ฯลฯ, การติดตามข้อมูล ภาพยนตร์ รายการบันเทิงต่างๆ ออนไลน์, การเล่นเกมคอมพิวเตอร์ออนไลน์, การเรียนรู้ออนไลน์ (e-Learning), การประชุมทางไกลผ่านอินเทอร์เน็ต (Video Conference), โทรศัพท์ผ่านอินเทอร์เน็ต (VoIP), การอับโหลดข้อมูล หรือ อื่นๆ ข้อมูลที่มีอยู่บนอินเทอร์เน็ตนั้นมีด้วยกันหลากหลายรูปแบบทั้ง ที่เป็นเสียง วีดีโอ ไฟล์เอกสาร นามสกุลต่างๆ เช่น .doc, .pdf แต่ข้อมูลส่วนใหญ่มักจะได้จะอยู่ในรูปแบบของเอกสาร HTML (Hyper Text Markup Language) ซึ่งเป็นภาษาคอมพิวเตอร์รูปแบบหนึ่ง ที่มีโครงสร้างการเขียนโดยอาศัยตัวกำกับ (Tag) ควบคุมการแสดงผลข้อความ, รูปภาพ หรือวัตถุอื่นๆ

ผ่านโปรแกรมค้นดูเว็บ(Web browser) ภาษา HTML มีแม่แบบมาจากภาษา SGML (Standard Generalized Markup Language) ที่ตัดความสามารถบางส่วนออกไปเพื่อให้สามารถทำความเข้าใจและเรียนรู้ได้ง่าย ซึ่งปัจจุบันมีการพัฒนาและกำหนดมาตรฐานโดยองค์กร World Wide Web Consortium (W3C) (W3Schools, 2011)

จากการศึกษาพบว่า 93 ล้านคนของผู้ค้นหาข้อมูลหรือประมาณร้อยละ 8 ของผู้ค้นหาข้อมูลที่เป็นผู้ใหญ่ค้นหาข้อมูลด้านสุขภาพ (Sutherland L.A and others, 2005: 300-305) เช่น ข้อมูลที่เกี่ยวข้องกับโรค ยาและวิธีการรักษา เนื่องจากปัจจุบันผู้คนส่วนมากได้ให้ความสนใจด้านสุขภาพของตนเองเพิ่มขึ้น ผู้ป่วยสนใจที่จะศึกษาข้อมูลของยาที่ตนเองจะได้รับมากขึ้นรวมทั้งปัญหาต่างๆทางด้านสุขภาพของตนเอง เพื่อใช้เป็นแนวทางในการปฏิบัติตน แสดงให้เห็นว่าการสืบค้นข้อมูลยาในอินเทอร์เน็ตมีความนิยมมากขึ้น ด้วยเหตุนี้ทำให้ข้อมูลด้านยาและสุขภาพบนอินเทอร์เน็ตเพิ่มขึ้นอย่างรวดเร็วเช่นเดียวกันกับข้อมูลอื่นๆ

เนื่องจากข้อมูลที่ปรากฏบนอินเทอร์เน็ต มีจำนวนมากและหลากหลายทำให้ผู้คนส่วนมากใช้อินเทอร์เน็ตเป็นแหล่งข้อมูลอันดับแรกๆในการสืบค้นข้อมูล โดยร้อยละ 80 ของผู้ใช้อินเทอร์เน็ตใช้โปรแกรมค้นหา (Search engine) ในการสืบค้นข้อมูล (Bernard and Amanda, 2006: 248–263) ถึงแม้ว่าการสืบค้นข้อมูลสารสนเทศทางอินเทอร์เน็ตจะมีข้อดีอย่างเห็นได้ชัดมากมายดังกล่าวมาข้างต้นแต่การค้นหาข้อมูลผ่านทางอินเทอร์เน็ตยังมีข้อจำกัดบางประการคือ ข้อมูลที่อยู่บนอินเทอร์เน็ตนั้นมีความครอบคลุมกว้างขวางมีความหลากหลายมาก ทำให้ผู้ใช้ที่ไม่มีทักษะในการสืบค้นอาจเกิดปัญหาเนื่องจากได้ข้อมูลไม่ตรงตามความต้องการเป็นจำนวนมากทำให้เสียเวลา นอกจากนั้นข้อมูลบนอินเทอร์เน็ตยังสามารถถูกปรับปรุงแก้ไขได้อย่างรวดเร็วทำให้การอ้างอิงเอกสารทำได้ลำบาก

การค้นหาข้อมูลที่มีเนื้อหาที่ตรงกับความสนใจจากเว็บเพจกว่า 100 ล้านเว็บเพจ ที่มีรวมกันอยู่ในอินเทอร์เน็ตจึงเป็นเรื่องยาก ดังนั้น โปรแกรมค้นหา จึงเป็นเครื่องมือสำคัญที่ใช้ในการค้นหาข้อมูลต่างๆ ซึ่งในปัจจุบันมีโปรแกรมค้นหาที่ใช้เป็นเครื่องมือในการสืบค้นข้อมูลบนอินเทอร์เน็ตหลายโปรแกรม ให้ผู้ใช้อินเทอร์เน็ตสามารถเลือกใช้ได้หลากหลาย อาทิเช่น AltaVista, Lycos, Yahoo, MSN, Google ฯลฯ อย่างไรก็ตามถึงแม้ว่าโปรแกรมค้นหาจะช่วยให้สามารถทำการ

สืบค้นข้อมูลได้อย่างว่องไวและได้ข้อมูลจำนวนมาก แต่การสืบค้นข้อมูลที่เป็นข้อมูลจำเพาะโดยเฉพาะอย่างยิ่งข้อมูลยาที่ดีและมีคุณภาพนั้นยังทำได้ยาก เนื่องจากจำนวนผลลัพธ์หรือเอกสาร HTML ที่โปรแกรมค้นหานำมาแสดงมีจำนวนมาก ซึ่งมีทั้งข้อมูลที่มีคุณภาพและไม่มีคุณภาพปะปนกัน อีกทั้งพบว่ามีเพียง 20 อันดับแรกเท่านั้นที่ผู้ใช้อินเทอร์เน็ตทำการเลือกอ่าน ทำให้ลำดับการปรากฏของเอกสาร HTML มีความสำคัญต่อการสืบค้นข้อมูลยาผ่านโปรแกรมค้นหาบนอินเทอร์เน็ตและเนื่องจากลำดับของเอกสาร HTML ที่ปรากฏบนโปรแกรมค้นหามีความสำคัญต่อการสืบค้นข้อมูล ผู้สร้างเว็บไซต์จึงพยายามปรับปรุงและพัฒนาเว็บไซต์เพื่อให้สามารถค้นหาเจอในอันดับต้นๆของโปรแกรมค้นหาต่างๆ การพัฒนาเว็บไซต์เพื่อให้ได้ผลลัพธ์ดังกล่าวสามารถทำได้หลายวิธีเช่น การเป็นผู้สนับสนุนโปรแกรมค้นหา, การปรับปรุงในส่วนของ Tag HTML, การทำ link popularity ฯลฯ กระบวนการดังกล่าวถูกเรียกว่าการทำ SEO (Search Engine Optimization) ซึ่งปัจจุบันถือเป็นกลยุทธ์ทางการตลาดที่มีประสิทธิภาพใช้ต้นทุนต่ำแต่ให้ผลกำไรสูง สามารถเพิ่มจำนวนผู้เข้าชม ทำให้สินค้าหรือบริการที่ตนเองทำนั้นสามารถขายได้มากขึ้นหรือมียอดการใช้บริการที่เพิ่มขึ้น

อย่างไรก็ดีการตรวจสอบความถูกต้องหรือความน่าเชื่อถือของข้อมูลบนอินเทอร์เน็ตกระทำได้ยาก เนื่องจากเครือข่ายอินเทอร์เน็ตเป็นระบบอิสระไม่มีผู้ควบคุม ข้อมูลส่วนใหญ่ขาดการประเมินคุณภาพจึงมีทั้งข้อมูลที่ถูกต้องน่าเชื่อถือ ไม่ถูกต้องหรืออาจเป็นข้อมูลที่ถูกต้องเพียงบางส่วน ซึ่งหากข้อมูลเหล่านั้นเป็นข้อมูลจำเพาะเช่นเรื่องโรคและยาอาจจะก่อให้เกิดอันตรายต่อผู้บริโภคได้อย่างร้ายแรง จากการศึกษาพบว่าข้อมูลสุขภาพที่นำเสนอทางเว็บไซต์จำนวนมากให้ข้อมูลที่ไม่ถูกต้องในการให้คำแนะนำการดูแลสุขภาพ การใช้ยา การแนะนำหรือการขายผลิตภัณฑ์ทางสุขภาพมีการโฆษณาสรรพคุณเกินจริง ก่อให้เกิดผลเสียต่อผู้บริโภค เช่น เสียทรัพย์โดยไม่จำเป็น ที่สำคัญคือการเสียโอกาสที่จะได้รับการรักษาอย่างถูกต้องทันเวลา (Theodosiou and Green, 2003: 387-396) จากการสำรวจเว็บไซต์ด้านการแพทย์ประมาณ 60 เว็บไซต์ ของผู้เชี่ยวชาญด้านกุมารเวชศาสตร์ของมหาวิทยาลัยโอไฮโอด้านสาธารณสุขและยา พบว่ามีเพียง 12 เว็บไซต์เท่านั้นที่ให้ข้อมูลถูกต้อง หรือประมาณร้อยละ 80 ของเว็บไซต์สุขภาพที่ทำการสำรวจในการศึกษาครั้งนั้นยังให้ข้อมูลไม่ถูกต้องครบถ้วน นอกจากนี้งานวิจัยของ Crocco และคณะพบว่าข้อมูลการใช้ยาที่สืบค้นได้จากอินเทอร์เน็ตบางเว็บไซต์ เป็นข้อมูลที่ไม่ถูกต้อง ทำให้ผู้นำข้อมูลไปใช้อาจเกิดอันตรายถึงชีวิตได้ แต่ยังมีการรายงานอันตรายที่เกิดจากการใช้ข้อมูลน้อย (Crocco and others, 2002: 2869 – 2871) ด้วย

ปัจจัยข้างต้นทำให้มีการพัฒนาเครื่องมือที่ใช้ในการประเมินข้อมูลทางด้านสุขภาพดังกล่าวเช่น HON (Health on the Net Foundation), CyberGuide, DISCERN โดยเครื่องมือดังกล่าวข้างต้นมีวัตถุประสงค์เพื่อใช้เป็นเครื่องมือที่ช่วยในการประเมินคุณภาพข้อมูล ที่เกี่ยวกับการรักษา โดยออกแบบให้ผู้ใช้เป็นผู้ประเมินความเหมาะสมของเนื้อหาคู่กับความต้องการของผู้ใช้เอง (Health on the net foundation, 2011) แต่อย่างไรก็ดี ยังไม่มีการศึกษาในส่วนของโครงสร้างของเว็บไซต์ที่ให้ข้อมูลด้านยา

ดังนั้นในการศึกษาครั้งนี้จึงทำการศึกษาโครงสร้างของเอกสารที่อยู่ในรูปของเว็บเพจบนเว็บไซต์ที่ให้ข้อมูลทางด้านยาที่เป็นภาษาอังกฤษ โดยทำการศึกษาปัจจัยที่มีผลต่อคุณภาพของเว็บไซต์ที่ให้ข้อมูลยาที่เป็นภาษาอังกฤษ และอันดับของการปรากฏในโปรแกรมค้นหาเพื่อเป็นแนวทางในการพัฒนาเว็บไซต์ให้มีโครงสร้างที่เหมาะสมและเป็นไปตามความต้องการของโปรแกรมค้นหา

1.2 ความมุ่งหมายและวัตถุประสงค์ของการศึกษา

เพื่อศึกษาคุณลักษณะโครงสร้างของเอกสาร HTML ที่เป็นปัจจัยต่อคุณภาพของเว็บไซต์ที่ให้ข้อมูลยาที่เป็นภาษาอังกฤษ และมีผลต่อการปรากฏในโปรแกรมค้นหา

1.3 ประโยชน์ที่คาดว่าจะได้รับ

ได้แนวทางในการพัฒนาเว็บไซต์ที่ให้ข้อมูลยา ให้มีคุณลักษณะของโครงสร้างของเอกสาร HTMLเหมาะสม เป็นไปตามความต้องการของโปรแกรมค้นหา และเป็นประโยชน์สำหรับผู้บริโภคด้วย

1.4 ขอบเขตของการศึกษา

คัดเลือกเอกสาร HTML โดยการสุ่มจากเว็บไซต์ที่ให้ข้อมูลยาที่เป็นภาษาอังกฤษ ใช้โปรแกรมค้นหา 3 ชนิด คือ Google, MSN, Yahoo และใช้คำสำคัญจำนวน 8 คำสำคัญ แบ่งออกเป็น 2 กลุ่มคือ คำสืบค้นทั่วๆไปได้แก่ “Drug Information” และ “Medical Information” คำสืบค้นที่เป็นชื่อเฉพาะทางยาได้แก่ "acetaminophen", "amoxicillin”, ”celecoxib“, "hydrochlorothiazide“, "lipitor" และ "spironolactone” เลือกศึกษาคุณลักษณะโครงสร้างของเอกสาร HTML เฉพาะที่เป็นปัจจัยต่อคุณภาพของเว็บไซต์ที่ให้ข้อมูลด้านยา และการแสดงผลของโปรแกรมค้นหา โดยช่วงเวลาที่ทำการเก็บข้อมูลคือตั้งแต่พฤษภาคม - ตุลาคม 2553

ลิขสิทธิ์มหาวิทยาลัย

บทที่ 2

วรรณกรรมที่เกี่ยวข้อง

2.1 ภาษา HTML และโครงสร้างของเอกสาร HTML

สำนักหอสมุดกลาง

HTML ย่อมาจาก Hyper Text Markup Language จัดเป็นภาษาทางคอมพิวเตอร์รูปแบบหนึ่งที่ใช้สำหรับแสดงข้อมูลต่างๆ ไม่ว่าจะเป็น ข้อความ รูปภาพ เสียง ฯลฯ ผ่านทางเว็บเบราว์เซอร์ ภาษานี้ได้รับการพัฒนาขึ้นในปี ค.ศ. 1989 เพื่อตอบสนองความต้องการในการสื่อสารแลกเปลี่ยนข้อมูลกันของนักวิทยาศาสตร์ระหว่างสถาบันและมหาวิทยาลัยต่างๆ ทั่วโลก โดย Tim Berners-Lee นักพัฒนาของ CERN ได้พัฒนาภาษาที่มีรากฐานมาจาก SGML ซึ่งเป็นภาษาที่ซับซ้อนและยากต่อการเรียนรู้ จนมาเป็นภาษาที่ใช้ได้ง่ายและสะดวก ซึ่งในปัจจุบันได้รับการดูแลมาตรฐานโดย W3C (World Wide Web Consortium) (Tim Berners-Lee, 1990) ซึ่งภาษา HTML จะประกอบไปด้วย 2 ส่วนคือ ส่วนที่เป็นข้อความทั่วไป และ ส่วนที่เป็นคำสั่งที่ใช้ในกำหนดรูปแบบของข้อมูลในการแสดงผลซึ่ง เรียกว่าแท็ก (Tag) ซึ่งเป็นอักษรหรือคำที่อยู่ใน square brackets คือประกอบด้วยเครื่องหมาย น้อยกว่า (<) และมากกว่า (>) คร่อมตัวอักษรหรือคำ แท็ก 1 ชุดประกอบด้วย แท็กเปิด กับ แท็กปิด โดยที่ แท็กปิด จะมีเครื่องหมาย slash (/) อยู่หน้าอักษรหรือคำในแท็ก สำหรับข้อความที่ต้องการนำเสนอจะใส่ไว้ระหว่าง แท็กเปิด กับ แท็กปิด โดยแท็ก HTML ดังกล่าวแบ่งออกเป็น 2 แบบคือ (HTML Introduction)

แท็ก เดี่ยว มีรูปแบบดังนี้

<ชื่อแท็ก /> เช่น

แท็ก เปิด-ปิด มีรูปแบบดังนี้

<ชื่อแท็ก>...</ชื่อแท็ก> เช่น ข้อความที่ต้องการแสดง

นอกจากแท็ก ที่เป็นตัวกำหนดรูปแบบของข้อมูลในการแสดงผลแล้ว ยังอาจประกอบด้วยส่วนขยายความสามารถของแท็ก หรือที่เราเรียกกันว่า Attribute ซึ่งจะต้องใส่ภายในเครื่องหมายน้อยกว่า (<) และมากกว่า (>) ในส่วนแท็กเปิดเท่านั้น โดยในแต่ละคำสั่ง HTML จะมี Attribute แตกต่างกันไป และมีจำนวนไม่เท่ากัน การระบุ Attribute จำนวนมากกว่า 1 Attribute จะใช้ช่องว่างเป็นตัวคั่น

โดยทั่วไปภาษา HTML จะมีการจัดวางคำสั่งหลักที่เป็นมาตรฐานเหมือนกันทั่วโลก คือจะประกอบด้วยคำสั่งหลัก ๆ อยู่ 4 คำสั่งด้วยกันดังนี้

สำนักหอสมุดกลาง

<html>.... </html> เป็นคำสั่งหลักที่ทำหน้าที่บอกจุดเริ่มต้นและจุดสิ้นสุดของเอกสาร HTML

<head>.... </head> ทำหน้าที่กำหนดส่วนหัวของเอกสาร HTML เป็นส่วนที่ใช้อธิบายเกี่ยวกับข้อมูลเฉพาะของหน้าเว็บนั้น ๆ

<title>.... </title> เป็นคำสั่งที่ใช้กำหนดข้อความที่ต้องการนำมาแสดงผลบนแถบ Title bar คำสั่งนี้จะอยู่ภายในคำสั่ง <head>.... </head>

<body>.... </body> ใช้กำหนดข้อความและรูปแบบของคำสั่งต่าง ๆ ที่ใช้สำหรับปรับแต่งเอกสารที่จะนำเสนอทางจอภาพ จะอยู่หลังคำสั่ง <head>.... </head>

รูปแบบโครงสร้างของเอกสาร HTML จะประกอบด้วยกัน 3 ส่วนคือ ส่วนประกาศ Version ของ HTML ส่วนหัวเรื่อง (Head Section) และส่วนเนื้อหา (Body Section) ดังรูปที่ 1

```
<!doctype>
<html>
<head>
        <title> title ของเว็บเพจ </title>
        <meta name="keywords" content="คำสำคัญ" >
        <meta name="description" content="รายละเอียดอย่างย่อ" >
</head>
<body>
        <h1> หัวเรื่อง </h1>
        <p> เนื้อเรื่อง, การจัดวางลักษณะของหน้าเอกสาร </p>
</body>
</html>
```

รูปที่ 1 แสดงตัวอย่างโครงสร้างเอกสาร HTML

2.1.1 ส่วนประกาศ Version ของ HTML เป็นส่วนที่กำหนดวิธีการแปลข้อมูลภายใน HTML ทั้งหมด โดยจะมีคำสั่งคือ <!DOCTYPE> โดยควรจะใส่ในเอกสารทุกๆ หน้า โดยวางไว้บรรทัดแรกเสมอ เพื่อบอกให้เว็บเบราว์เซอร์ ทราบถึงรุ่นของคำสั่ง HTML และชนิดของเอกสาร (Document Type Definition : DTD) ที่ใช้ ซึ่งจะช่วยให้เว็บเบราว์เซอร์แปลเอกสารได้อย่างถูกต้อง (Raggett 2005)

2.1.2 ส่วนหัวเรื่อง (Head Section) เป็นส่วนที่ใช้อธิบายเกี่ยวกับข้อมูลเฉพาะของหน้าเว็บเพจนั้นๆ เช่น ชื่อเรื่องของหน้าเว็บ (Title) ชื่อผู้จัดทำเว็บ (Author) คำสำคัญสำหรับการค้นหา ฯลฯ โดยมีจะอยู่ภายในส่วนของ <head>.... </head> ความสำคัญของ ส่วน headers คือไว้สำหรับเพื่อให้โปรแกรมค้นหาสามารถเข้าถึงและเก็บข้อมูลต่างๆ ภายในเว็บไซต์นั้นๆ ได้ง่ายขึ้น โดยส่วนนี้อาจประกอบไปด้วยแท็กต่างๆ ดังนี้

2.1.2.1 แท็ก Title <title> ใช้กำหนดชื่อในแถบด้านบนของเว็บเบราว์เซอร์ (Title bar) ใช้เป็นชื่อเรื่องสำหรับเพจนั้นๆ เมื่อมีการเพิ่มในรายการโปรด และใช้แสดงชื่อสำหรับหน้าเว็บเพจนั้นๆ ในหน้าผลลัพธ์ที่ได้จากการค้นหาด้วยโปรแกรมค้นหา

2.1.2.2 Metadata Elements หรือ แท็ก Meta <meta> เป็นแท็กที่ให้รายละเอียดเกี่ยวกับเอกสาร HTML หน้านั้นๆ โดยเฉพาะ ทั้งในส่วนของคำสำคัญ ชื่อผู้เขียนเอกสาร วันที่แก้ไข

ล่าสุด และข้อมูลอื่น ๆ ซึ่งข้อมูลในส่วนนี้จะไม่ปรากฏบนหน้าเว็บเพจ แต่จะถูกอ่านและวิเคราะห์โดยโปรแกรมค้นหาและเว็บเบราว์เซอร์ แท็กนี้มีด้วยกัน 2 Attribute คือ name และ HTTP-EQUIV

2.1.2.2.1 Attribute name ใช้สำหรับกำหนดคำค้นหา รายละเอียดเกี่ยวกับเว็บไซต์เพื่อให้โปรแกรมค้นหา นำข้อมูลที่ระบุไว้ไปใช้ในการประมวลผลและจัดเก็บเว็บไซต์โดยรูปแบบคำสั่งและ Attribute value ของ name ที่สำคัญมีดังนี้คือ (Chris Sherman, 2005)

Keywords เป็นการบรรยายเว็บไซต์ด้วยคำสำคัญคือพวก คำ, วลี, กลุ่มคำที่สื่อความหมายเกี่ยวกับเว็บไซต์ ที่คนนิยมใช้ในการค้นหาข้อมูล สามารถใส่ได้หลายคำ และแบ่งคำโดยใช้เครื่องหมายจุลภาค (,) เช่น <meta name= “keywords” content= “drug information, HTML Documents” />

Description เป็นการบรรยายเว็บไซต์ด้วยประโยคสรุป แสดงรายละเอียดสั้น ๆ ของหน้าเว็บเพจที่กำลังแสดงผลอยู่ ไม่ควรเขียนให้สั้น หรือ ยาวจนเกินไป ข้อความที่เขียนควรสัมพันธ์กับเนื้อหาของหน้านั้นๆ เมื่อมีการค้นหา โปรแกรมค้นหาจะนำข้อความใน Description ไปแสดงในหน้าผลการค้นหา เช่น <meta name= “description” content= “การศึกษาคุณลักษณะโครงสร้างของเอกสาร HTML จากเว็บไซต์ที่ให้ข้อมูลยา” />

Robot เป็นคำสั่งที่ใช้สำหรับบอก Robot ของโปรแกรมค้นหา ให้ทำอะไรกับหน้าเว็บเพจได้บ้าง ซึ่งโดยปกติเมื่อ Robot ของโปรแกรมค้นหา มาเจอเว็บเพจจะเข้ามาเก็บข้อมูล หรือ index ข้อมูลในหน้าเว็บเพจนั้นๆ กลับไปฐานข้อมูลของตัวโปรแกรมค้นหาเพื่อนำไปประมวลผลต่อไป

นอกเหนือจากที่กล่าวมายังมี Attribute value อื่นๆ เช่น author copyright generator ใช้สำหรับบอก ผู้เขียนหน้านี้ ผู้เป็นเจ้าของลิขสิทธิ์ โปรแกรมที่ใช้สร้างเว็บเพจนี้ ฯลฯ

2.1.2.2.2 Attribute http-equiv ใช้สำหรับกำหนดชุดของตัวอักษรที่ใช้หรือหรือสำหรับสั่ง Refresh หน้าเว็บเพจหรือให้ไปที่เว็บไซต์ ไฟล์อื่นๆ ตามเวลาที่กำหนด (delay เป็นวินาที) Attribute value ของ http-equiv ที่สำคัญมีดังนี้คือ

Content-Type เป็นคำสั่งสำหรับกำหนดชุดตัวอักษรที่ใช้ในหน้าเว็บเพจ ว่าสามารถแสดงผลได้ถูกต้องด้วยชุดตัวอักษรแบบใด และเป็นเอกสารประเภทอะไร โดยปัจจุบันนิยมกำหนดแบบอักษรเป็น utf-8 และสำหรับการแสดงผลในภาษาไทยจะใช้ windows-874 หรือ tis-620 เช่น <meta http-equiv= “content-type” content= “text/html; charset=UTF-8” />

Content-language ใช้สำหรับระบุว่า หน้าเว็บไซต์นั้น ๆ มีเนื้อหาเป็นภาษาอะไร เช่น <meta http-equiv= “content-language” content= “th” />

Refresh เป็นคำสั่งสำหรับสั่งให้ Re-direct คือการให้เปลี่ยนหน้าเว็บเพจที่แสดงผลอยู่ ไปแสดงผลอีกหน้าที่เราตั้งไว้โดยอัตโนมัติหน้าเว็บเพจอื่นๆ หรือเพื่อ Refresh หน้าเว็บเพจปัจจุบัน เช่น <meta http-equiv ="Refresh" content="10; URL= http://www.pharm.su.ac.th"> โดยตัวเลขจะบอกระยะเวลาที่จะแสดงหน้าแรกก่อนจะเปลี่ยนไปยังหน้าที่กำหนด (แสดงผลเป็นวินาที)

จากการศึกษาผลกระทบของลักษณะเนื้อหาของหน้าเว็บในการแสดงผลหน้าเว็บบนโปรแกรมค้นหาและ ผลกระทบของการใช้แท็ก Meta ในการแสดงผลหน้าเว็บบนโปรแกรมค้นหา พบว่าแท็ก Title มีผลต่อการแสดงผลหน้าเว็บบนโปรแกรมค้นหาโดยจะสามารถทำให้อันดับที่ได้จากการค้นหาดีขึ้นถ้าหากเพิ่มความถี่ของคำสำคัญ ใน Title หรือทั้งใน Title และใน Full-text นอกจากนี้ยังพบว่าแท็ก Meta เป็นเครื่องมือที่สำคัญในการปรับปรุงการแสดงผลหน้าเว็บและ คำสำคัญใน Mata Tag ที่มาจากใน Title และใน Full-text จะให้ผลของอันดับที่ดีขึ้น (J. Zhang and A. Dimitroff, 2005: 665–715) นอกจากนี้การศึกษารูปแบบการใช้คำอธิบายใน Meta Tag ในหน้าเว็บเพจพบว่าเว็บเพจที่อยู่ใน 20 อันดับแรกเมื่อใช้ Google เป็นเครื่องมือค้นหา โดยใช้ภาษาต่างๆ กันเมื่อทำการวิเคราะห์แท็ก Meta ที่เป็นคำอธิบายและความยาวของคำอธิบาย พบว่ามีสัดส่วนที่แตกต่างกันโดยหน้าเว็บเพจที่ใช้ภาษายุโรปตะวันตกมีคำอธิบายในสัดส่วนที่สูง ในขณะที่หน้าเว็บเพจที่คำอธิบายเป็นภาษาจีนมีสัดส่วนที่ต่ำ (T. C. Craven, 2004: 479–493)

2.1.2.3 แท็ก Style หรือ Style sheet <style> ใช้กำหนดคุณลักษณะต่างๆ เช่น สี ขนาด แบบตัวอักษร ตำแหน่งการจัดวาง ให้กับ HTML element ต่างๆ โดยจะกำหนดในหน้าเว็บเพจนั้นๆ หรือสร้างเป็นไฟล์ style sheet แยกเป็นอีกไฟล์ แล้วเรียกใช้อีกทีเพื่อจะสามารถนำ style sheet เดียวกันมาใช้กับไฟล์ HTML ได้หลายไฟล์ได้ง่ายโดยที่เขียนเพียงแค่ครั้งเดียว และจำเป็นต้องกำหนด Attribute type ให้กับแท็กนี้ด้วย ส่วน Attribute value ที่เป็นไปได้คือ "text / css" เท่านั้น นอกจากนี้ ใน Tag style สามารถกำหนดชนิดของอุปกรณ์ที่ใช้ในการแสดงผลได้อีกด้วยซึ่งหากมีมากกว่า 1 ชนิด จะคั่นด้วยเครื่องหมาย (,) เช่น <style type="text/css" media="screen">.....</style>

2.1.2.4 แท็ก Link <link> ใช้กำหนดความสัมพันธ์ระหว่างเว็บเพจกับข้อมูลภายนอกเพื่อนำมาใช้งานในเว็บเพจนิยมใช้ในการเชื่อมโยงไปยัง Style sheets โดยมี Attribute ที่สำคัญคือ

href ไฟล์ที่ต้องการลิงก์ไป, URL ของเอกสารที่เกี่ยวข้อง

src ไฟล์ที่ต้องการแทรก

media ชนิดอุปกรณ์ที่เจาะจงนำไปใช้แสดงผล

rel อธิบายถึงความสัมพันธ์ไปยังหน้าเว็บอื่นหรือทรัพยากรอินเทอร์เน็ตอื่น ๆ เช่น <link href="index.html" rel=HOME>

rev อธิบายการย้อนกลับความสัมพันธ์แสดงวิธีการที่ทรัพยากรอื่นๆ มองของเอกสารปัจจุบัน เช่น <link href="index.html" rel=HOME rev=CHILD>

type MIME Type (Multipurpose Internet Mail Extension) เป็นการกำหนดชนิดข้อมูลมาตรฐานสำหรับอินเทอร์เนตในแบบ Content-type เป็นตัวบ่งชี้ว่าชนิดข้อมูลนั้นๆ คืออะไร เว็บเบราเซอร์จะนำข้อมูลนั้นมาโดยวิธีใด และจะจัดการกับข้อมูลเหล่านั้นต่อไปอย่างไร อาจแสดงเป็นชนิดข้อมูลหลัก type เปล่าๆ หรือแสดงชนิดข้อมูลรองด้วย type/subtype เช่น Text, Text/html

2.1.2.5 แท็ก Script <script> ใช้ในการกำหนดสคริปต์ฝั่งไคลเอ็นต์เช่น JavaScript สามารถเขียนเป็น function เพื่อใช้งาน หรือจะเขียนฟังก์ชั่นเก็บเป็นไฟล์ .js ก่อนแล้วค่อยเรียกใช้งานก็ได้ Attribute ที่จำเป็นต้องระบุให้กับ แท็กนี้คือ type (MIME Type) เป็นการกำหนดชนิดข้อมูลมาตรฐานสำหรับอินเทอร์เนตในแบบ Content-type เป็นตัวบ่งชี้ว่าชนิดข้อมูลนั้นๆ คืออะไร เว็บเบราว์เซอร์จะนำข้อมูลนั้นมาโดยวิธีใด และจะจัดการกับข้อมูลเหล่านั้นต่อไปอย่างไร อาจแสดงเป็นชนิดข้อมูลหลัก type เปล่าๆ หรือแสดงชนิดข้อมูลรองด้วย type/subtype

2.1.3 ส่วนเนื้อหา (Body Section)

เป็นส่วนที่แสดงเนื้อหาทั้งหมดที่จะแสดงทางหน้าจอ มีส่วนประกอบ ได้แก่ ข้อความ ตาราง ลิสต์ รูป ภาพ ลิงค์ เป็นต้น เนื่องจากในส่วนนี้จะเป็นส่วนการทำงานหลักของหน้าเว็บเพจ ดังนั้นในส่วนนี้จึงประกอบไปด้วย แท็กต่างๆ เป็นจำนวนมากตามลักษณะของข้อมูล ที่ต้องการนำเสนอโดยทั้งหมดจะอยู่ภายใต้แท็ก <body>...</body> (Amit Singhal, 2012)

2.1.3.1 แท็ก Body <body> เป็นแท็กเริ่มต้นและสิ้นสุดของเนื้อหาในเอกสารที่จะแสดงบนหน้าจอ โดยเราสามารถกำหนดรูปแบบ, ตัวอักษรหรือ คุณสมบัติต่างๆ ให้กับ Body Tag <body> โดยผ่าน Attribute ต่างๆ เช่น alink value = “colorname” ระบุสีของลิงก์ขณะเรียกใช้งานเอกสาร นอกจากนี้การกำหนดลักษณะพื้นหลังและแบบหรือลักษณะของตัวอักษรที่ใช้ในหน้าเว็บ

เพจ สามารถใช้งาน CSS Style Sheet แทนการกำหนดค่าผ่าน Attribute ของ <body> ได้เช่นเดียวกัน

2.1.3.2 แท็ก Heading <hn> เป็นแท็กที่ใช้สำหรับกำหนดลักษณะหัวเรื่องให้มีขนาดต่างๆ ตามความสำคัญ มีขนาดตัวอักษรใหญ่กว่าข้อความที่เป็นเนื้อหา ซึ่งจะถูกกำกับด้วยแท็ก <hn> โดย n คือตัวเลขแสดงขนาดของตัวอักษร ค่าของ n นั้นจะมีค่าอยู่ในช่วง 1 – 6 ซึ่ง n = 1 จะมีขนาดตัวอักษรหัวเรื่องใหญ่สุด ใช้สำหรับหัวข้อสำคัญที่สุดและ n = 6 จะมีขนาดตัวอักษรหัวเรื่องเล็กสุด สำหรับหัวข้อที่สำคัญน้อยสุด โดยโปรแกรมค้นหาจะใช้หัวเรื่องที่เรากำหนดผ่านแท็ก <h1-h6> ในการทำดรรชนีเนื้อหาในเว็บเพจของว่าเป็นเรื่องเกี่ยวกับอะไร เนื่องจาก <h1> เป็นหัวเรื่องสำคัญจึงควรมีได้เพียงครั้งเดียวในหน้าเว็บเพจนั้นๆ ส่วนหัวเรื่องย่อยๆ สามารถมีได้หลายหัวข้อย่อยได้แต่ไม่ควรมีเกิน <h6>

2.1.3.3 แท็ก Paragraph เป็นการแท็กที่ใช้ในการขึ้นย่อหน้าใหม่ และกำหนดตำแหน่งของข้อความที่จะแสดงผล โดยเว็บเบราว์เซอร์จะทำการเพิ่มแถวว่างให้ก่อนและหลังข้อความที่อยู่ในแท็ก โดยอัตโนมัติ ซึ่งเราจะเห็นในลักษณะของการเว้นบรรทัด หรือเห็นระยะห่างระหว่างข้อความชัดเจนขึ้น และแท็กนี้สามารถกำหนดตำแหน่งของข้อความภายในย่อหน้านั้นๆ ผ่าน Attribute align ตัวอย่างการเขียน ข้อความที่ต้องการแสดง โดยผลที่ได้คำว่า “ข้อความที่ต้องการแสดง” จะถูกจัดให้อยู่ชิดด้านขวาของส่วนแสดงผล

2.1.3.4 แท็ก Hyperlink <a> หรือแท็กลิงก์ เป็นการเชื่อมโยงหน้าเว็บเพจ โดยการกำหนดส่วนของข้อความที่ต้องการ หรือรูปภาพที่ได้เลือกไว้เพื่อเป็นจุดเชื่อมโยงไปยังเนื้อหาจุดอื่นๆ ซึ่งเนื้อหาที่เชื่อมโยงไปนั้นอาจจะอยู่ภายในหน้าเดียวกัน อยู่คนละหน้า หรืออาจอยู่คนละเว็บไซต์ก็ได้ ในกรณีที่เป็นเว็บไซต์เดียวกันหน้าเว็บเพจ รูปภาพหรือไฟล์ข้อมูลต่างๆ นั้นควรจะถูกจัดเก็บไว้ในที่เดียวกัน (Folder เดียวกัน) เพื่อความสะดวกในการเรียกใช้งาน เมื่อเลื่อนเมาส์ไปยังข้อความที่เป็นลิงก์ลูกศร จะเปลี่ยนรูปเป็นรูปมือเล็กๆ ทั้งนี้การเชื่อมโยงหน้าเว็บเพจจะประกอบด้วยส่วนสำคัญ 2 ส่วน คือ จุดที่ใช้สำหรับเชื่อมโยง (Link) เป็นส่วนที่ใช้เป็นจุดเชื่อมโยงไปยังเป้าหมาปลายทางอาจเป็นรูปภาพหรือข้อความก็ได้ เป้าหมาย (Target) เป็นส่วนปลายทางหรือจุดที่ต้องการให้เชื่อมโยงมาถึงจะอยู่ส่วนใดของเอกสารก็ได้แล้วแต่ผู้เขียนกำหนด การสร้างการเชื่อมโยงสามารถทำได้หลายวิธีดังนี้

2.1.3.4.1 การเชื่อมโยงไปยังหน้าเว็บเพจอื่นๆ ที่อยู่ในเว็บไซต์เดียวกันคือการเชื่อมโยงไปยังจุดเชื่อมโยงที่อยู่คนละหน้าเว็บเพจกัน โดยมีรูปแบบการใช้งานดังนี้ ข้อความที่ปรากฏบนหน้าเว็บเพจ เมื่อกำหนดหน้าเว็บเพจที่ต้องการเชื่อมโยงแล้ว สามารถระบุสถานที่ที่จะเปิดใช้หน้าเว็บเพจนั้นโดยอาศัย Attribute target

ตัวอย่าง บทที่ 2 วรรณกรรมที่เกี่ยวข้อง ซึ่งในตัวอย่างนี้เมื่อคลิกไปที่ข้อความจะเชื่อมโยงไปยัง "chapter2.html"

2.1.3.4.2 การเชื่อมโยงไปยังเว็บไซต์อื่นๆ สามารถทำได้โดยมีรูปแบบดังต่อไปนี้ ข้อความลิงก์

2.1.3.4.3 การเชื่อมโยงภายในหน้าเว็บเพจเดียวกันจะใช้ในกรณีที่หน้าเว็บเพจนั้นมีข้อมูลยาวมาก เราสามารถทำหัวข้อกำหนดไว้ให้เลือกอ่านได้ง่าย เมื่อผู้ชมทำการเลือกหัวข้อจะมีการเลื่อนตำแหน่งข้อมูลจากตำแหน่งเดิมไปยังเนื้อหาของหัวข้อที่ถูกเลือกนั้นๆ

ตัวอย่างการเชื่อมโยงภายในเว็บเพจเดียวกัน

```
<html>
<head>
        <title>การสร้างการเชื่อมโยง (Link) </title>
</head>
<body>
        <a name="top"> การสร้างลิงก์</a>
        เนื้อหาหรือข้อความที่ยาวๆ
        <a href="#top"> กลับไปหัวข้อการสร้างลิงก์ </a>
</body>
</html>
```

2.1.3.4.4 การเชื่อมโยงไปยังแฟ้มข้อมูลประเภทต่างๆ เพื่อให้ทำการดาวน์โหลดข้อมูลหรือเอกสารต่างๆ เช่น ไฟล์รูปภาพ, ไฟล์โปรแกรม, ไฟล์ซิป (.zip) และไฟล์เสียง เป็นต้น การสร้างจุดเชื่อมโยงหรือลิงก์มีรูปแบบการใช้งานคล้ายกับการเชื่อมโยงแบบอื่นๆ ที่กล่าวมาข้างต้น ต่างกันที่นามสกุลของไฟล์ที่ใช้ หากไฟล์จุดหมายปลายทางเป็นไฟล์ประเภทใด เบราว์เซอร์ก็จะทำงานให้สอดคล้องกับไฟล์ประเภทนั้น เช่น หากไฟล์จุดหมายปลายทางที่สร้างลิงค์ไว้เป็นรูปภาพ เว็บเบราว์เซอร์ก็จะเปิดรูปภาพนั้นแสดง แต่ถ้าเป็นไฟล์ประเภทอื่นก็จะพบหน้าจอของการดาวน์โหลด (Download) โดยรูปแบบคำสั่งที่ใช้ เป็นดังนี้คือ ข้อความลิงก์

2.1.3.4.5 การเชื่อมโยงเพื่อส่งจดหมายอิเล็กทรอนิกส์ (e-mail) เพื่อเป็นการให้ผู้เข้าชมเว็บไซต์สามารถติดต่อสื่อสารกับผู้จัดทำเว็บไซต์นั้นๆ โดยอาจเป็นการส่งคำติชมการสร้างเว็บ การแจ้งข่าวสารให้ผู้สร้างเว็บไซต์ทราบ เพื่อเป็นการสื่อสาร แบบ 2 ทาง (Two-way Communication) โดยรูปแบบคำสั่งที่ใช้ เป็นดังนี้คือ ข้อความลิงค์

2.1.3.5 แท็ก Image <img /> หรือแท็กรูปภาพ เป็นการใส่รูปภาพลงในเว็บเพจ เพื่อเพิ่มความน่าสนใจให้กับเว็บไซต์โดยนำรูปภาพมาใส่ในเว็บเพจร่วมกับข้อความ แต่การใส่รูปภาพในเว็บเพจนั้น ผู้สร้างเว็บไซต์ควรจัดเก็บรูปภาพนั้นไว้ในโฟลเดอร์เดียวกับเว็บเพจก่อนเพื่อให้ง่ายต่อการจัดการและการเคลื่อนย้ายส่วนภาพที่นำมาใส่ในเอกสารควรมีขนาดของไฟล์ภาพไม่เกิน 40 Kb เพราะภาพที่มีขนาดใหญ่เกินไปจะทำให้การโหลดข้อมูลช้าผู้เข้าชมต้องเสียเวลารอนาน รูปภาพที่นิยมใช้ในหน้าเว็บเพจมีอยู่ 2 ชนิด คือ รูปภาพชนิด Graphics Interchange Format (GIF) เป็นไฟล์รูปภาพที่มีการบีบข้อมูลทำให้ประหยัดเนื้อที่ มีนามสกุล .gif สามารถแสดงความละเอียดของภาพได้สูงสุด 256 สี ส่วนมากจะเป็นรูปภาพการ์ตูนต่างๆ ตัวอักษร ภาพเคลื่อนไหวที่ไม่ต้องการความละเอียดของมาก สามารถทำพื้นโปร่งใสได้ (transparent) รูปภาพชนิด Joint Photographic Experts Group (JPEG หรือ JPG) เป็นไฟล์ที่มีการบีบข้อมูลเช่นกันทำให้ไฟล์มีขนาดเล็ก มีนามสกุลเป็น .jpg แต่ไฟล์ รูปแบบ JPG มีความละเอียดของสีถึง 16.7 ล้านสี ซึ่งทำให้ภาพมีความคมชัดมากกว่า จึงนิยมใช้กับภาพถ่ายต่าง ๆ แต่ไม่สามารถทำภาพโปร่งใส หรือทำเป็นภาพเคลื่อนไหวได้

โดยในการแสดงรูปภาพนั้นจะใช้แท็ก <img /> ซึ่งเป็นแท็กเดี่ยว และจำเป็นต้องตามด้วย Attribute src (src มาจากคำว่า source ที่แปลว่าแหล่งข้อมูล) ส่วน value ของ src คือ URL เป็นตัวบอกตำแหน่งของภาพที่ต้องการแสดงที่ถูกจัดเก็บไว้ และ Attribute alt เป็นข้อความที่ใช้เพื่อระบุความจำเพาะของรูปภาพนั้นๆ โดยค่าของ alt นั้นจะเป็นข้อความที่ผู้สร้างเว็บไซต์เป็นผู้กำหนด โดย

ปกติจะแสดงเมื่อใช้เมาส์วางอยู่เหนือรูปภาพนั้นๆ Attribute alt จัดเป็นการข้อมูลทางเลือกสำหรับรูปถ้าผู้ใช้ ในกรณีที่ไม่สามารถเรียกดูภาพนั้นๆ ได้ซึ่งอาจเป็นเพราะข้อผิดพลาดในการเชื่อมต่อของระบบอินเทอร์เน็ตที่ช้า, ความผิดพลาดจาก Attribute src, หรือผู้ใช้เรียกผ่านโปรแกรมอ่านหน้าจอ (screen reader) มีรูปแบบคำสั่งดังนี้คือ นอกจากนี้เพื่อให้การแสดงรูปภาพที่อยู่บนเว็บเพจมีความสมบูรณ์มากยิ่งขึ้น สามารถเพิ่มเติมด้วย Attribute อื่นๆ ได้

2.1.3.6 แท็ก Form <form> เป็นการสร้างแบบฟอร์มที่พบเห็นตามเว็บไซต์ต่างๆ เช่น แบบสอบถาม, ใบสั่งซื้อของ, การลงทะเบียนเป็นสมาชิก, สมุดเยี่ยม (Guestbook), แสดงความคิดเห็น (Comment) เป็นต้น เพื่อใช้ในการส่งผ่านข้อมูลต่างๆ ไปยัง server โดยปกติแล้วการบันทึกค่าข้อมูลในแบบฟอร์มจะใช้ Server Script เช่น PHP, ASP เข้ามาช่วยจัดการ จะประกอบด้วยช่อง Element หลัก 3 ประเภท คือ ส่วนรับข้อมูล ส่วนที่ใช้ในการรับข้อมูลและ ส่วนที่ใช้รับข้อมูลในลักษณะของข้อความหลายบรรทัด เช่น ส่วนรับที่อยู่ ส่วนแสดงความคิดเห็น

2.1.3.7 แท็ก Frame <frame> เป็นแท็กที่ทำให้สามารถแสดงเอกสาร HTML ได้หลายหน้า (หลายไฟล์) พร้อมกันในหน้าต่างเบราว์เซอร์เดียวกัน เอกสาร HTML แต่ละหน้าก็คือ Frame ซึ่งแต่ละ Frame จะเป็นอิสระจาก Frame อื่นๆ แต่การใช้ Frame มีข้อเสีย คือ ยุ่งยากเวลาจะสั่งพิมพ์หน้าเว็บเพจทั้งหมด นอกจากการใช้ Frame ที่ได้กล่าวไปข้างต้นแล้ว ยังมีการใช้ frame อีกรูปแบบหนึ่ง ที่เรียกว่า Inline Frame (IFrame) เป็นการกำหนดให้สามารถแสดงหน้าเว็บเพจอื่นๆ ลงในหน้าเว็บเพจของเราได้ มักใช้ในส่วนที่เป็นโฆษณา ซึ่งเมื่อมีการแก้ไข code จะไม่ส่งผลกระทบกับหน้าเว็บเพจหลัก

2.1.3.8 แท็ก List <ul>, <li>, <ol>, <li>, <dl>, <dt>, <dd> เป็นแท็กที่ใช้ในการสร้างลิสต์รายการต่างๆ สามารถแบ่งออกเป็น 3 ประเภทได้แก่ ลิสต์รายการ แบบมีลำดับ (Ordered List) ลิสต์รายการแบบไม่มีลำดับ โดยใช้สัญลักษณ์กำกับ (Unordered List หรือ Bullet List) ในการทำลิสต์รายการ และลิสต์รายการแบบให้คำนิยาม (Definition List)

2.1.3.9 แท็ก Table <table> แท็กตารางเป็นคำสั่งที่ใช้ในการสร้างตารางลงในเอกสาร HTML โดยแท็กที่พบมากที่สุดได้แก่<table>, <tr>, <td>

2.2 คุณลักษณะของเว็บไซต์และองค์ประกอบของเว็บไซต์ที่ดี

2.2.1 สามารถจำแนกกลุ่มผู้ใช้งานเว็บไซต์ ตามลักษณะการเข้าถึงสามารถแบ่งได้เป็น 2 กลุ่ม ดังนี้คือ (กรรณิการ์ สวรรค์โพธิพันธุ์, 2550: 21-28)

2.2.1.1 การเข้าถึงเว็บไซต์โดยใช้ URL หรือชื่อของเว็บไซต์โดยตรง ซึ่งผู้ใช้งานกลุ่มนี้จะมีข้อมูลหรือเป้าหมายที่แน่นอน ซึ่งมักจะสอดคล้องกับความต้องการของตนเองอยู่แล้ว

2.2.1.2 การเข้าถึงเว็บไซต์โดยผ่านโปรแกรมค้นหา ผู้ใช้งานกลุ่มนี้จะทราบเพียงเป้าหมายที่ตนเองต้องการ แต่ไม่ทราบว่าจะหาข้อมูลจากเว็บไซต์ใดได้บ้าง จึงจำเป็นต้องใช้เครื่องมือที่ช่วยในการค้นหา

จากการแบ่งกลุ่มผู้ใช้เว็บไซต์ออกเป็น 2 กลุ่มข้างต้น จะสัมพันธ์กับปัจจัยที่ส่งผลกระทบต่อการออกแบบเว็บไซต์ ซึ่งสอดคล้องกับพฤติกรรมการใช้งานของผู้ใช้งาน

2.2.2 ปัจจัยที่ส่งผลกระทบต่อการออกแบบเว็บไซต์ คือ สิ่งที่นักพัฒนาเว็บไซต์ต้องคำนึงถึงเมื่อต้องการออกแบบเว็บไซต์ให้มีประสิทธิภาพ ในที่นี้จะจำแนกปัจจัยที่มีผลกระทบต่อการออกแบบเว็บไซต์ ออกเป็น 2 ปัจจัยคือ

2.2.2.1 ผู้ใช้งานเว็บไซต์ (User) เนื่องจากวัตถุประสงค์หลักของการออกแบบเว็บคือสามารถใช้งานได้อย่างมีประสิทธิภาพ และสร้างความพึงพอใจให้แก่ผู้ใช้งาน ดังนั้น กลุ่มผู้ใช้งานจึงถือเป็นตัวแปรสำคัญในการออกแบบเว็บไซต์ ผู้ใช้งานเว็บไซต์อาจหมายถึงกลุ่มผู้ใช้งานเว็บไซต์ทั่วไปทั้งที่เคยเข้ามา เยี่ยมชมหรือ ไม่เคยเข้ามาเยี่ยมชมก็ได้ โดยปัจจัยด้านนี้ได้แก่ พฤติกรรมของผู้ใช้งาน ซึ่งเป็นการแสดงออกถึงวิธีการใช้งานเว็บไซต์ เช่น วิธีการเข้าถึงเว็บไซต์ เวลาที่ใช้ในการชมข้อมูลในแต่ละหน้าเว็บเพจ เครื่องมือที่ใช้งานบ่อยและความพึงพอใจของผู้ใช้ เป็นต้น และจากการสำรวจนั้นพบว่าผู้ใช้ส่วนใหญ่ จะมีลักษณะการใช้งานดังนี้คือ ผู้ใช้จะนิยมเข้าถึงข้อมูลในลักษณะ “Deep Link” คือผ่านโปรแกรมค้นหา มากกว่าการเข้าถึงข้อมูลโดยใช้ URL หรือชื่อของเว็บไซต์นั้นโดยตรงในกรณีที่เข้าถึงเว็บไซต์โดยใช้โปรแกรมค้นหา ผู้ชมเว็บไซต์จะเลือกชมเว็บไซต์ที่ปรากฏในหน้าแสดงผล (SERP) แรก หรืออันดับต้นๆ เท่านั้น เมื่อเข้าถึงเว็บไซต์ ผู้ชมเว็บไซต์จะใช้เวลาในการชมเว็บเพจย่อย ซึ่งจะมีเนื้อหาที่ตรงกับความต้องการของผู้ชม มากกว่าการใช้เวลาในการชมหน้าโฮมเพจ และจะใช้เวลาเฉลี่ยเพียงเล็กน้อย เพื่อพิจารณาว่าเว็บเพจดังกล่าวมีข้อมูลที่เกี่ยวข้องกับที่ตนต้องการ หรือมีข้อมูลใดที่น่าสนใจ หากไม่พบก็จะยกเลิกการใช้งานเว็บนั้นทันที โดยส่วนที่ผู้ใช้จะเลือกชมข้อมูลมากที่สุดคือ

ส่วนเนื้อหา จากพฤติกรรมดังกล่าวจึงทำให้ผู้พัฒนาเว็บไซต์ พยายามทำให้เว็บไซต์ของตนสามารถค้นเจอเมื่อใช้โปรแกรมค้นหา และอยู่ในอันดับต้นๆ ของหน้าแสดงผล

2.2.2.2 สภาพแวดล้อม (Environment) สภาพแวดล้อมจัดเป็นปัจจัยเสริมที่ช่วยให้สามารถใช้งานเว็บไซต์ได้มีประสิทธิภาพมากขึ้น ซึ่งสิ่งที่มีผลมากที่สุดคือ การเปลี่ยนแปลงทางเทคโนโลยี โดยสามารถแบ่ง ได้เป็นปัจจัยย่อยๆ ดังนี้

เวลาที่ใช้ในการดาวน์โหลดข้อมูล ซึ่งการที่เว็บไซต์ มีจำนวนข้อมูลที่มาก หรือเป็นเว็บไซต์ที่มีการใช้งานในส่วนของกราฟฟิกมาก จะทำให้ใช้เวลาในการโหลดข้อมูลต่างๆ นานจนทำให้ผู้ชมเว็บไซต์เกิดความเบื่อหน่าย ดังนั้นผู้พัฒนาเว็บไซต์จึงต้องจำกัดในเรื่องของกราฟฟิกต่างๆ ซึ่งในปัจจุบันปัญหาดังกล่าวลดลงเนื่องจากมีการนำเทคโนโลยีอินเทอร์เน็ตความเร็วสูง (DSL หรือ ADSL) ซึ่งประมวลผลด้วยสัญญาณดิจิตอลมาใช้ ทำให้การนำเสนอข้อมูลบนเว็บเพจทำได้อย่างรวดเร็วขึ้นการจัดโครงร่างเว็บเพจ (Layout) ส่วนนี้มีความสัมพันธ์กับความละเอียดของจอภาพที่ใช้แสดงผล ซึ่งปัจจุบันนิยมออกแบบเว็บเพจบนความละเอียด 1024 x 768 pixels อย่างไรก็ตาม ในอนาคตต้องมีการพัฒนาจอภาพให้มีขนาดใหญ่ขึ้น ซึ่งอาจส่งผลกระทบต่อการแสดงผลของหน้าเว็บเพจเดิมได้

การทำงานข้ามแพลตฟอร์ม (Cross-Platform) เว็บเบราว์เซอร์ที่ใช้งานอยู่กันในปัจจุบันมีความหลากหลายขึ้น ไม่ได้จำกัดเพียง Internet Explorer เท่านั้นแต่ยังมีเว็บเบราว์เซอร์อื่นๆ อีกเช่น Safari Mozilla ดังนั้นการออกแบบควรมีการทดสอบการใช้งานกับเว็บเบราว์เซอร์ต่างๆ เพื่อป้องกันความผิดพลาดของการแสดงผลข้อมูล

2.2.3 คุณลักษณะของเว็บไซต์ที่ดีนอกจากการมีโครงสร้างของเอกสาร HTML ที่ถูกต้องและเหมาะสม, รูปแบบสวยงาม, เนื้อหาที่มีคุณภาพแล้ว เว็บไซต์ที่ดียังต้องมีองค์ประกอบอื่น ๆ ดังนี้ (จิตเกษม พัฒนศิริ, 2539)

2.2.3.1 มีรายการสารบัญแสดงรายละเอียดของเว็บเพจ ผู้สร้างควรแสดงรายการทั้งหมดที่เว็บเพจนั้นมีอยู่ให้ผู้ใช้ทราบ โดยอาจจัดทำในรูปของสารบัญ หรือลิงค์ การสร้างสารบัญนี้จะช่วยให้ผู้ใช้สามารถค้นหาข้อมูลภายในเว็บเพจได้อย่างรวดเร็ว แนวทางที่จะป้องกันไม่ให้ผู้ใช้หลงทางคือ การจัดสร้างแผนที่การเดินทาง (Site map) ขั้นพื้นฐานที่เว็บเพจนั้นก่อน ซึ่งได้แก่ การสร้างสารบัญให้ผู้ใช้ได้เลือกที่จะเดินทางไปยังส่วนใดของเว็บเพจได้จากจุดเริ่มของเว็บเพจที่สร้างขึ้น

2.2.3.2 เชื่อมโยงข้อมูลไปยังเป้าหมายได้ตรงกับความต้องการมากที่สุด ถ้าผู้สร้างเว็บทราบแหล่งข้อมูลอื่นว่าสามารถให้ความกระจ่างแก่ผู้ใช้ได้ ควรที่จะนำเอาแหล่งข้อมูลนั้นมาเขียนเป็นลิงค์ (Link) ซึ่งอาจสร้างอยู่ในรูปของตัวอักษร หรือรูปภาพก็ได้ แต่ควรที่จะแสดงจุดเชื่อมโยงกลับมายังหน้าแรก (Homepage) ของเว็บไซต์ที่กำลังใช้งานอยู่ด้วย เพื่อให้ผู้ใช้มีหนทางกลับสู่จุดเริ่มต้นได้

2.2.3.3 เนื้อหา กระชับ สั้นและทันสมัย เนื้อหาควรเป็นเรื่องที่กำลังมีความสำคัญ อยู่ในความสนใจของผู้คนหรือเป็นเรื่องที่ต้องการให้ผู้ใช้ทราบ การปรับปรุงเนื้อหา มีบริการใหม่เพิ่มขึ้น การปรับเปลี่ยนเว็บไซต์จึงเป็นเรื่องที่ต้องทำอยู่เสมอ เนื้อหาไม่ควรที่จะยาวเกินไป ขนาดที่ดูดีที่สุด คือ กำหนดให้ในแต่เว็บเพจแสดงผลได้เพียงหน้าเดียวเท่านั้น หรือให้แสดงได้ในจำนวนหน้าน้อยที่สุดเท่าที่จะทำได้

2.2.3.4 สามารถตอบโต้กับผู้ใช้ได้อย่างทันท่วงที ควรกำหนดจุดที่ผู้ใช้สามารถแสดงความคิดเห็นหรือให้คำแนะนำกับผู้สร้างได้ เช่น ใส่อีเมล์ (E-mail address) ลงในเว็บเพจ ตำแหน่งที่ควรเขียนเป็นส่วนบนสุดหรือล่างสุดของเว็บเพจนั้น ๆ ไม่ควรเขียนแทรกไว้ในตำแหน่งใด ๆ ของจอภาพ เพราะผู้ใช้อาจหาไม่พบ

2.2.3.5 มีรูปภาพประกอบการนำเสนอที่ดี แต่ไม่ควรมีรูปภาพมากเกินไป ควรเป็นรูปที่สื่อความหมายและตรงกับเนื้อหา ที่นิยมใช้กันมากที่สุด ได้แก่ การนำรูปภาพมาประกอบพื้นฉากหลัง ภาพที่นำมาประกอบนั้นไม่ควรจะมีสีสันที่ฉูดฉาดมากนัก เนื่องจากจะไปลดความเด่นของเนื้อหาลง ควรใช้ภาพที่มีสีอ่อน ๆ ไม่สว่างมาก เป็นรูปเรียบ ๆ ตัวอักษรที่นำมาแสดงบนจอภาพก็เช่นกัน ควรเลือกขนาดที่อ่านง่าย ไม่มีสีสันและลวดลายมากเกินความจำเป็น ไม่ควรที่จะนำรูปภาพขนาดใหญ่หรือมีจำนวนรูปภาพมากเกินไป บางเว็บจะมีการออกแบบเว็บเพจที่เน้นความสวยงามเป็นหลักจนละเลยความสำคัญของเนื้อหาไป

2.2.3.6 เข้าสู่กลุ่มเป้าหมายได้อย่างถูกต้อง เว็บเพจจำนวนมากที่สร้างขึ้นมาแล้วกลับไม่ได้รับความสนใจจากผู้ใช้ มีหลายสาเหตุที่ทำให้ผู้ใช้ไม่เข้ามาใช้บริการ แต่การกำหนดกลุ่มเป้าหมายที่ชัดเจนย่อมทำให้ผู้สร้างสามารถกำหนดเนื้อหา และเรื่องราวเพื่อให้ตรงกับความต้องการของผู้ใช้ได้มากกว่า

2.2.3.7 ใช้งานง่าย ประเด็นนี้เป็นประเด็นที่ผู้สร้างต้องให้ความสำคัญเป็นรายการแรกสุด เพราะไม่ว่าอะไรก็ตามถ้ามีความง่ายในการใช้งาน โอกาสที่จะประสบความสำเร็จย่อมสูงขึ้นเป็นลำดับ ทำอย่างไรจึงจะสร้างเว็บเพจให้ใช้งานได้ง่าย สิ่งนี้ย่อมขึ้นอยู่กับเทคนิคและประสบการณ์ของผู้สร้างแต่ละคน

2.2.3.8 เป็นมาตรฐานเดียวกัน กำหนดข้อมูลให้เป็นมาตรฐานเดียวกัน โดยอาจจะต้องมีการแบ่งเนื้อหาของข้อมูลออกเป็นส่วนๆ ไปจัดชุดข้อมูลให้เป็นกลุ่มเป็นหมวดหมู่จะทำให้ข้อมูลทุกอย่างดูเป็นระเบียบน่าอ่าน

2.3 โปรแกรมค้นหา (Search Engines) และหลักการทำ SEO (Search Engine Optimization)

เนื่องจากทุกวันนี้มีข้อมูลต่างๆ เกิดขึ้นเป็นจำนวนมากในโลกอินเทอร์เน็ตไม่ว่าจะเป็นข้อมูลภาพ เสียง หรือเอกสาร HTML ต่างๆ จึงจำเป็นต้องมีระบบค้นหาเกิดขึ้นเพื่ออำนวยความสะดวก โดยในปี 1993 นาย Matthew Gray จากสถาบัน MIT ได้พัฒนาโปรแกรมค้นหาที่เรียกว่า World Wide Web Wanderer ซึ่งจัดเป็นโปรแกรมค้นหาโปรแกรมแรก และหลังจากนั้นก็ได้มีการพัฒนาโปรแกรมค้นหาใหม่ๆ ขึ้นมาอย่างต่อเนื่อง โปรแกรมค้นหา คือโปรแกรมที่ช่วยในการสืบค้นหาข้อมูล โดยเฉพาะข้อมูลบนอินเทอร์เน็ต โดยครอบคลุมทั้งข้อความ รูปภาพ ภาพเคลื่อนไหว เพลง ซอฟต์แวร์ แผนที่ ข้อมูลบุคคล กลุ่มข่าว และอื่น ๆ ซึ่งแตกต่างกันไปแล้วแต่โปรแกรมหรือผู้ให้บริการแต่ละราย โปรแกรมค้นหาส่วนใหญ่จะค้นหาข้อมูลจาก Key word (คำสำคัญ) ที่ผู้ใช้ป้อนเข้าไป จากนั้นจะแสดงรายการผลลัพธ์ที่คิดว่าผู้ใช้น่าจะต้องการขึ้นมา

โปรแกรมค้นหาสามารถแบ่งออกเป็นหลายประเภทซึ่งเมื่อแบ่งตามรูปแบบโครงสร้างของระบบจะสามารถแบ่งโปรแกรมค้นหาออกเป็น 3 ประเภทคือ

ประเภทที่ 1 Crawler Based Search Enginesเป็นโปรแกรมค้นหาหาบนอินเตอร์เน็ตแบบอาศัยการบันทึกข้อมูล และ จัดเก็บข้อมูลเป็นหลัก ให้ผลการค้นหาแม่นยำที่สุด และการประมวลผลการค้นหาสามารถทำได้อย่างรวดเร็ว โดยการทำงานจะใช้ Crawler หรือ Spider หรืออีกชื่อคือ Robots ทำหน้าที่ไปตามอินเทอร์เน็ตเพื่อเก็บข้อมูลต่างๆ ของเว็บไซต์ นำมาประมวลผลและจัดเก็บลงในฐานข้อมูลของตนเอง เช่น Google, Yahoo, MSN เป็นต้น

ประเภทที่ 2 Web Directory หรือ Blog Directory เป็นกลุ่มแรกๆ ของโปรแกรมค้นหาที่มีโครงสร้างอยู่ในแบบฐานข้อมูล เนื่องจากมีระบบที่ไม่ยุ่งยากมาก อาศัยคนเป็นผู้เพิ่มข้อมูลไปในกลุ่มต่างๆ ทำให้สามารถค้นหาข้อมูลต่างๆ ที่เกี่ยวข้องกัน ในปริมาณมาก ๆ ทำงานคล้ายกับสมุดหน้าเหลือง โดยจะมีการสร้าง ดรรชนี ระบุหมวดหมู่ อย่างชัดเจน ซึ่งจะช่วยให้การค้นหา ข้อมูลต่าง ๆ ตามหมวดหมู่นั้น ๆ ได้รับการเปรียบเทียบอ้างอิง ในขณะที่เราค้นหาข้อมูลจะมีเว็บไซต์มากมาย หรือ Blog มากมายที่มีเนื้อหาคล้าย ๆ กันในหมวดหมู่เดียวกัน ให้เราเลือกที่จะหาข้อมูลได้ อย่างตรงประเด็นที่สุด ซึ่งจะช่วยลดระยะเวลาในการค้นหา ตัวอย่างเครื่องมือค้นหาประเภทนี้ได้แก่ Open Directory Project (ODP) หรือ DMOZ.org SANOOK

ประเภทที่ 3 Meta Search Engines เป็นโปรแกรมค้นหาที่จะไปดึงข้อมูลจากฐานข้อมูลของโปรแกรมค้นหาอื่นๆ มาประมวลผลในระบบของตนเอง แล้วจึงนำผลลัพธ์ที่ได้มาแสดงให้กับผู้ที่ใช้งานอีกทีหนึ่ง โดยการทำงานจะใช้การค้นหาโดยอาศัย Meta Tag ในภาษา HTML ซึ่งมีการประกาศชุดคำสั่งต่าง ๆ เป็นรูปแบบของ Tex Editor ด้วยภาษา HTML เนื่องจากบางครั้งผู้ให้บริการหรือผู้ออกแบบเว็บสามารถใส่อะไรเข้าไปใน Meta Tag เพื่อให้เกิดการค้นหาและพบเว็บ หรือ บล็อกของตนเอง จึงทำให้ผลการค้นหาข้อมูลผ่านโปรแกรมค้นหาประเภทนี้ให้ผลที่ไม่เที่ยงตรงเท่าที่ควร ตัวอย่างโปรแกรมค้นหาได้แก่ Dogpile, SurfWax, Metacrawler เป็นต้น

ซึ่งถ้าหากเรานำโปรแกรมค้นหาต่างๆ ดังที่กล่าวมาข้างต้นมาทำการจัดกลุ่มใหม่ตามความนิยมในการใช้งานจะสามารถแบ่งได้เป็นกลุ่มย่อยๆ ดังนี้

ประเภทที่ 1 Major Search Engines คือกลุ่มของโปรแกรมค้นหาที่มีผู้ใช้มากที่สุด โดยจะประกอบด้วยโปรแกรมค้นหาหลักๆ คือ Google Yahoo และ Live หรือ MSN

ประเภทที่ 2 Miner Search Engines คือกลุ่มที่ให้บริการสืบค้นข้อมูลและ ขายหรือแจกจ่ายฐานข้อมูลให้กับโปรแกรมค้นหาอื่นๆ เช่น ODP (DMOZ.com)

ประเภทที่ 3 Niche Search Engines หรือโปรแกรมค้นหาเฉพาะทาง ซึ่งจะใช้ในการค้นหาข้อมูลที่มีความจำเพาะเจาะจง

การทำงานของโปรแกรมค้นหาจะประกอบด้วยขั้นตอนหลักๆ อยู่ด้วยกัน 3 ขั้นตอนคือ

ค้นหา (Crawling) โดยในโปรแกรมค้นหาจะมี Robots, Spider หรือ Crawler เป็นของตัวเอง เพื่อค้าหาหน้าเว็บไซต์ หรือข้อมูลใหม่ๆ ที่เกิดขึ้นบนอินเทอร์เน็ต และทำการเก็บข้อมูล

จัดทำดรรชนี (Indexing) หลังจากที่ได้ข้อมูลมาแล้วจะมีการแยกเนื้อหาออกเป็นส่วนต่างๆ โดยจะแยกขั้นตอนออกเป็นขั้นตอนย่อยๆ คือ คัดแยก ซึ่งจะตัดเนื้อหาหรือหน้าเว็บที่เก็บมาแยกออกเป็นส่วนๆ ประกอบไปด้วย ตัวหนังสือ, ลิงค์, รูปภาพ และเนื้อหาอื่นๆ เช่น ไฟล์เอกสาร แฟ้มบีบอัด ไฟล์ภาพเคลื่อนไหว ต่างๆ และส่งต่อไปยังส่วนอื่นๆ คือ ลิงค์ จะถูกส่งไปให้ Robots เข้าไปเก็บข้อมูลอีกครั้ง ภาพ จะถูกส่งไปให้ Robots ตัวอื่นทำการจัดทำดรรชนีใหม่ ซึ่งจะใช้อัลกอริทึมคนละตัวกับการทำดรรชนีปกติ ไฟล์ประเภทอื่นๆ จะถูกส่งไปอ่าน เพื่อให้เข้าใจเนื้อหาของไฟล์นั้นๆ และส่งไปเข้าทำดรรชนี เช่นกัน

แสดงผล (Searching) หลังจากที่เนื้อหาของหน้าเว็บไซต์ถูกจัดเรียงลำดับ จะถูกนำมาใช้งานเมื่อผู้ใช้ทำการสืบค้นผ่านโปรแกรมค้นหานั้นๆ

เนื่องจากข้อมูลที่อยู่บนอินเทอร์เน็ตนั้นแบ่งออกเป็นหลายประเภทด้วยกันดังนั้นผู้พัฒนาโปรแกรมค้นหาจึงได้เขียนโปรแกรมเพื่อพัฒนา Robots ให้สามารถเก็บข้อมูลที่แตกต่างออกไปได้ง่ายขึ้น โดยจะแบ่ง Robots ออกเป็นประเภทต่างๆ คือ

General Bots จะใช้ในการเก็บหน้าเว็บไซต์ปกติ ซึ่งแต่ละค่ายโปรแกรมค้นหาก็จะมีชื่อเรียกของ Robots ที่แตกต่างกันออกไป

Bots for Business จะใช้งานในแง่ของธุรกิจเป็นหลัก เช่นใช้เพื่อตรวจสอบข้อมูลบางอย่างของ Pay-Per-Click ใช้สำหรับเว็บที่หารายได้จากโฆษณาของทั้ง Google, Yahoo เป็นต้น

Niche Bots จะเป็น Robots ที่เน้นเฉพาะทางด้านใดด้านหนึ่งเท่านั้น เช่น สำหรับการเก็บข้อมูลรูปภาพ วีดีโอ เพลง, สำหรับเก็บข่าว, สำหรับเก็บบล็อกภาษาญี่ปุ่น, สำหรับตรวจสอบเว็บไซต์ที่โกงจากการทำ SEO เป็นต้น

ปัจจุบันผู้ใช้งานอินเทอร์เน็ตส่วนใหญ่ใช้โปรแกรมค้นหาในการค้นหาข้อมูลแทนการพิมพ์ URL (Uniform Resource Locator) โดยใช้คำสำคัญ ป้อนลงใน Search Engine Box (ช่องสำหรับใส่คำสืบค้น) ต่างๆ ทำให้ค้นหาสิ่งที่ต้องการง่าย มีข้อมูลปรากฏให้เลือกเปรียบเทียบจากหลายแหล่งข้อมูลโดยแสดงผลออกมาหลายเว็บไซต์ เว็บไซต์ที่ปรากฏเป็นอันดับต้นๆ บนหน้าแสดงผล จะถูกเลือกเข้าไปดูข้อมูลมากที่สุด จึงทำให้เว็บไซต์ต้องการขึ้นอันดับ 1 ของคำค้นนั้นๆ เพื่อผลประโยชน์หลายๆ ด้าน (สุธี จันทร์แต่งพล, 2551)

Search Engine Optimization หรือ SEO เป็นการจัดทำ, ปรับปรุง หรือปรับแต่งหน้าเว็บไซต์ ด้วยกระบวนการต่างๆ ให้เหมาะสมสำหรับโปรแกรมค้นหาต่างๆ เพื่อพัฒนาให้เว็บไซต์ติดอันดับต้นๆ หรือปรากฏในหน้าแรกของโปรแกรมค้นหา ในการทำ SEO ควรจะประกอบไปด้วยบุคคลอย่างน้อย 3 กลุ่มคือ

นักการตลาด (Marketer) เพื่อทำหน้าที่ในการวางแผนเก็บข้อมูล กำหนดเป้าหมายของการทำ SEO

ผู้ดูแลเว็บไซต์ (Webmaster) เพื่อจัดทำและเขียนเนื้อหาภายในเว็บไซต์ ให้เหมาะกับโปรแกรมค้นหา

โปรแกรมเมอร์ (Programmer) เป็นกลุ่มที่จะต้องเข้าใจหลักการทำ SEO อย่างมาก เพื่อเขียนโปรแกรมให้สนับสนุนโปรแกรมค้นหาเข้าใจเนื้อหาของเว็บไซต์ได้อย่างถูกต้อง

นักออกแบบ (Web designer) เพื่อออกแบบให้เว็บไซต์ สวยงาม ใช้งานได้ง่าย และถูกต้องตามหลักการพื้นฐานของ SEO

โดยการทำ SEO สามารถแบ่งออกได้เป็น 2 ส่วนหลักคือ (Amit Singhal and Matt Cutts, 2011)

On-Page Factors (การปรับแต่งภายในเว็บ) โดยแบ่งออกเป็นได้ 3 ส่วนหลักๆ คือ

1. โครงสร้างของเว็บไซต์ (Website Structure) โดยทั่วไปเว็บไซต์จะพัฒนาด้วยภาษา HTML PHP ASP JSP ฯลฯ ซึ่งภาษาเหล่านี้ตัวโปรแกรมค้นหาไม่ได้ให้ความสำคัญ หรือนำไปใช้ในการจัดอันดับ แต่สิ่งที่สำคัญคือ รูปแบบการจัดโครงสร้างของหน้าเว็บไซต์โดยการใช้ HTML Tag ในการบ่งบอกความสำคัญของแต่ละส่วน

แท็ก Title– <title></title> ในส่วนของแท็ก Title ควรใส่คำสำคัญที่จะใช้ในการสืบค้นไว้โดยควรจะเป็นคำที่เกี่ยวข้องกับเนื้อหาของเอกสารนั้นๆ โดยการสร้างแท็กชื่อควรหลีกเลี่ยงการใช้ชื่อเดียวกันในแต่ละหน้าของเว็บเพจที่แตกต่างกัน เพื่อให้โปรแกรมค้นหาทราบความแตกต่างของหน้าเว็บเพจนั้นๆ กับหน้าเว็บเพจอื่นๆ นอกจากนี้ควรใช้คำที่สั้นแต่ได้ใจความ โดยปกติคำที่อยู่ในแท็กชื่อไม่ควรเกิน 200-250 ตัวอักษร หากชื่อยาวเกินไปการแสดงผลจะแสดงได้เพียงบางส่วนในผลการค้นหา และหลีกเลี่ยงการใส่คำที่ไม่จำเป็นลงในแท็กชื่อ (Alireza 2007: 501-506)

แท็ก Head– <h1></h1>, <h2></h2>, <h3></h3> แท็ก Head ถูกนำไปใช้ในการเน้นหัวข้อต่างๆ เพื่อให้ผู้ใช้งานสามารถเข้าใจได้ว่าเนื้อหาส่วนนี้เป็นเนื้อหาเกี่ยวกับอะไร ส่วนของโปรแกรมค้นหาก็เช่นกัน ได้นำข้อความส่วนของแท็กหัวเรื่องมาใช้ตรวจสอบโครงสร้าง และเนื้อหาของเว็บเพจว่ามีความเกี่ยวข้องกันอย่างไร และหากภายในเว็บไซต์มีเนื้อหาที่ตรงกับคำค้นหาที่ผู้ใช้งานค้นมาจากโปรแกรมค้นหา ตัวโปรแกรมก็จะจัดอันดับเว็บไซต์นั้นๆ ให้อยู่สูงกว่าเว็บไซต์อื่นๆ โดยทั่วไปหัวเรื่องที่เป็น h1 นั้นควรจะเป็นเรื่องหลักและใช้หรือมีคำสำคัญหลักปนอยู่ด้วย ในขณะที่หัวเรื่องรองก็ควรมีคำสำคัญรองอยู่ด้วยเช่นกัน นอกจากนี้ควรมีการกำหนดหัวเรื่องบนหน้าเว็บเพจให้เหมาะสม หลีกเลี่ยงการใช้แท็กหัวเรื่องที่มากเกินไป รวมถึงการใส่ข้อความทั้งหมดลงในแท็กหัวเรื่อง (Matt Cutts, 2009a)

แท็ก Meta Keyword– <meta name=”keyword” content=”Optimizing On-Page Factors for Search Engine”> เป็นส่วนที่ใช้บอกโปรแกรมค้นหาว่าเว็บไซต์นั้นๆ มีเนื้อหาเกี่ยวกับอะไร โดยการใส่คำสำคัญ ลงใน แท็ก Meta Keyword โดยเลือกใช้คำที่มีความเกี่ยวข้องกับเนื้อหาลงไปให้พอเหมาะ ไม่มากจนเกินไป ทางโปรแกรมค้นหาจะนำส่วนนี้มาใช้พิจารณาความสัมพันธ์ของส่วนต่างๆ เพื่อใช้ในการจัดอันดับ แต่ปัจจุบัน แท็ก Meta Keyword ถูกนำไปใช้ในด้านมืด โดยผู้สร้างเว็บไซต์บางรายนำคีย์เวิร์ดที่ไม่เกี่ยวข้องจำนวนมากใส่ไว้ที่นี่ ทำให้โปรแกรมค้นหาต่างๆ ตัดทอนความสำคัญหรือยกเลิกการใช้แท็กนี้ในการจัดอันดับ แต่อย่างไรก็ตามการใส่คำสำคัญลงในแท็ก Meta Keyword เป็นการช่วยยืนยันความถูกต้องของเนื้อหาในหน้าเว็บเพจนั้นๆ ว่าเกี่ยวข้องกับเรื่องอะไร และเป็นส่วนเสริมให้แท็กชื่อ มีความน่าเชื่อถือเพิ่มขึ้นด้วย และเช่นเดียวกันหากใส่คำค้นต่างๆ ลงไปมากเกินความจำเป็นก็จะทำให้โปรแกรมค้นหา จัดอันดับเว็บไซต์ไว้อันดับท้ายๆ ได้ ซึ่งโดยปกติแล้วตามมาตรฐานจะกำหนดไว้ที่ 250 ตัวอักษร โดยระหว่างคำสำคัญให้คั่นด้วยเครื่องหมาย (,)

แท็ก Meta Description– <meta name=”description” content=”Optimizing On-Page Factors for Search Engine”> ส่วนของ Meta Description จะเป็นส่วนที่ใช้แสดงรายละเอียดที่อยู่ใต้หัวข้อที่หน้าผลการค้นหาของโปรแกรมค้นหา โดยควรจะบอกรายละเอียดโดยสรุป และมีความเกี่ยวข้องกับเนื้อหา หรือคำสำคัญ ให้มากที่สุด สำหรับแท็กนี้จะไม่แสดงผลในหน้าเว็บบราวเซอร์เช่นเดียวกันกับ แท็ก Meta Keyword โดยความยาวของเนื้อหา หรือข้อความที่อยู่ในแท็กไม่ควรเกิน 250 ตัวอักษร แต่สามารถยืดหยุ่นได้ตามเนื้อหาของเว็บเพจ

แท็ก Meta Robots <meta name=”robots” content=”noindex, nofollow”> เป็นส่วนที่ใช้บอก Robots ให้กระทำตามรูปแบบการใช้งานที่กำหนดไว้โดยจะใส่ไว้ในส่วนหัวระหว่างแท็ก <head>..</head> เช่นเดียวกันกับแท็ก Meta tag อื่นๆ โดยคำสั่งจะมีด้วยกัน 4 รูปแบบคือ

noindex คือ ไม่ต้องเก็บเป็น index แต่ให้วิ่งตามหน้านี้

nofollow คือ ให้เก็บเป็น index แต่ไม่ต้องวิ่งตามลิงก์ทั้งหมดในหน้าเว็บนี้

noarchive คือ ไม่ต้องเก็บ Cache Page ในหน้านี้แต่จะเก็บ index ตามปกติ

nosnippet คือ ไม่ต้องแสดงข้อความสั้นๆ ใต้ชื่อเว็บไซต์ในหน้าผลการค้นหา

โดยหากทำการใส่ nofollow ไว้ที่หน้าใดก็ตามลิงก์ในหน้านั้นทั้งหมดจะไม่ถูกนำไปคิดในการจัดอันดับเลย

แท็ก Image <img> หรือแท็กที่เกี่ยวข้องกับรูปภาพ โดยภาพจัดเป็นส่วนประกอบของเว็บไซต์ที่สามารถสื่อสารได้อย่างตรงไปตรงมา และสามารถเพิ่มประสิทธิภาพการใช้งานให้เกิดประโยชน์สูงสุดได้ เพราะโปรแกรมค้นหามองความครบถ้วนสมบูรณ์ของเนื้อหา การมีภาพสามารถช่วยทำให้เนื้อหากระจ่างมากขึ้น ภาพทั้งหมดสามารถมีชื่อไฟล์ที่แตกต่างกัน รวมทั้งแอตทริบิวต์ “alt” โดยตัวแอตทริบิวต์นี้ช่วยให้สามารถเขียนข้อความสำรองที่จะปรากฏขึ้นเมื่อมีเหตุขัดข้องที่ทำให้ไม่สามารถแสดงรูปภาพได้ หรือหากผู้ใช้งานชมเว็บไซต์ผ่านเบราว์เซอร์ที่ไม่สนับสนุนภาพ ส่วนเนื้อหาของแอตทริบิวต์ alt จะเป็นตัวบอกข้อมูลเกี่ยวกับภาพนั้นๆ ให้ผู้ใช้ได้รับทราบ ดังนั้นควรที่จะตั้งชื่อไฟล์ภาพและข้อความ alt ให้กระชับแต่ได้ใจความ หลีกเลี่ยงการตั้งชื่อไฟล์ภาพด้วยคำทั่วๆ ไปเช่น "image1.jpg", "pic.gif", "1.jpg" หลีกเลี่ยงการยัดคำหลักลงในข้อความ alt หรือคัดลอกและวางประโยคทั้งหมดลงไป เนื่องจากการเขียนข้อความ alt ที่ยืดยาวโปรแกรมค้นหาอาจสงสัยว่าข้อความนั้นเป็นสแปม หากจำเป็นต้องมีการใช้ภาพเป็นลิงก์ ควรมีการสร้างข้อความ alt ทุกครั้งเพื่อช่วยให้โปรแกรมค้นหา ทำความเข้าใจเกี่ยวกับหน้าเว็บไซต์ที่ปลายลิงก์ได้ดียิ่งขึ้น นอกจากนี้ในการจัดเก็บภาพควรจัดเก็บภาพไว้ในไดเรกทอรี่เดียวกันซึ่งจะทำให้พาธที่ไปยังภาพไม่ซับซ้อน และใช้ไฟล์ที่ได้รับการสนับสนุนโดยทั่วไป ได้แก่ JEPG, GIF, PNG และ BMP

แท็ก P, B และ strong สำหรับแท็กทั้งสามนี้ ควรนำมาใช้งานควบคู่กันไป โดยแท็ก p (Paragraph) นั้นไม่มีผลมากนักต่อการจัดอันดับของโปรแกรมค้นหา แต่มักถูกใช้เพื่อเสริมให้ Robots สนใจในเนื้อหาส่วนนี้มากยิ่งขึ้น เพราะแท็กนี้ หมายถึงย่อหน้าที่มีเนื้อหาต่างๆ บรรจุอยู่ภายใน สำหรับแท็ก b (Bold) และ strong ให้ผลลัพธ์เหมือนกันคือ ทำให้ตัวหนังสือหนาขึ้น ต่างกันที่ b จะเป็นมาตรฐานที่เก่ากว่า แต่ควรหลีกเลี่ยงการทำตัวหนาทั้งย่อหน้า หรือทั้งหน้าเว็บเพจเพราะจะทำให้ Robots ที่เก็บข้อมูลไม่ทราบว่าเว็บไซต์นี้ต้องการเน้นคำใด

แท็ก ul, li เป็นแท็กที่ใช้ในการแสดงส่วนเนื้อหา สามารถนำมาใช้เพื่อแสดงรายการของลิสต์ต่างๆ ที่เชื่อมโยงไปยังหน้าเว็บเพจหรือเว็บไซต์อื่นๆ ซึ่งช่วยให้ Robots สามารถตามเข้าไปเก็บเนื้อหาได้ง่ายขึ้น

แท็ก a แท็กลิงก์ ถูกนำมาใช้เพื่อเป็นลิงก์ไปยังส่วนต่างๆ ในหน้าเว็บไซต์ โดยมีด้วยกันหลายรูปแบบ ทั้งลิงก์ทั่วไป, ลิงก์ที่เป็นจาวาสคริปต์ เป็นต้น ซึ่งรูปแบบของลิงก์ที่ Robots ชอบมากที่สุดคือ SEO เนื่องจากเป็นรูปแบบที่อธิบายชัดเจนว่าต้องการจะไปที่ไหน และแสดงเรื่องอะไร ซึ่งหากเว็บไซต์นั้นๆ มีรูปแบบของลิงก์ที่ผิดพลาดก็จะทำให้ Robots ไม่สามารถเก็บข้อมูลได้และมองว่านั่นไม่ใช่ลิงก์ โดยรูปแบบลิงก์ที่มักมีข้อผิดพลาดได้แก่ลิงก์

ที่เป็นจาวาสคริปต์, ลิงก์ที่เป็นดร็อปดาวน์เมนู เนื่องจากตัวบอทไม่สามารถเลือกหัวข้อต่างๆ และไปยังหน้าเว็บเพจเหล่านั้นได้เลย ผลคือทำให้หน้านั้นไม่ได้รับการเก็บข้อมูล ซึ่งการแก้ไขมีสองทางคือ เลิกใช้เมนูนี้ หรือเปลี่ยนเมนูนี้ไปใช้เมนูที่เป็น CSS Style และจาวาสคริปต์แทน, ลิงก์ที่อยู่ในลักษณะของแฟลชแอนิเมชั่น เนื่องจากการเข้าถึงลิงก์ดังกล่าวของบอททำได้ยากหรืออาจทำไม่ได้, ลิงก์ที่เป็นภาพหากการระบุชื่อของไฟล์ภาพนั้นไม่ชัดเจนการเก็บข้อมูลของโปรแกรมค้นหาก็จะผิดพลาดตามไปด้วย เช่นหากตั้งชื่อไฟล์ภาพว่า “menu1.jepg” โปรแกรมค้นหาจะไม่ทราบว่าลิงก์ดังกล่าวต้องการสื่ออะไร รู้แค่ว่าเป็นลิงก์ไปยังอีกหน้าหนึ่งเท่านั้น อีกประการหนึ่งคือการใส่ nofollow ลงในลิงก์ ตัวอย่างเช่น SEO จะนิยมใส่ในส่วนของเว็บไซต์ที่อนุญาตให้ผู้ใช้สามารถกรอกแสดงความคิดเห็น เพื่อเป็นการรักษาระดับลิงก์ออกไปยังเว็บไซต์อื่นๆ (Hodgson, 2001: 20-25)

นอกจากนี้ยังพบว่าส่วนของคุณลักษณะต่างๆ ของเว็บไซต์ที่ส่งผลต่อการปรากฏบนหน้าแสดงผลแสดงผลด้วยเช่น ค่า Page Rank, ปริมาณลิงก์ที่พบในหน้าเว็บเพจ จากการศึกษาพบว่าหากมีปริมาณลิงก์มากกว่า 100 ลิงก์ต่อหน้าเว็บเพจจะทำให้โปรแกรมค้นหาไม่ทำการเก็บข้อมูลเว็บเพจดังกล่าวซึ่งจะทำให้อันดับของเว็บเพจนั้นตกลง, ค่าการดาวน์โหลดหน้าเว็บเพจ จากการศึกษาของพบว่าหากเว็บเพจมีระยะเวลาการดาวน์โหลดที่มากกว่า 25-30 วินาทีจะทำให้ปริมาณของผู้เข้าชมลดลง โดยจากการศึกษานี้ระยะเวลาการดาวน์โหลดหน้าเว็บเพจเฉลี่ยอยู่ที่ 6 วินาที (Matt McGee, 2010, Barry Schwartz, 2009)

2. ชื่อและเนื้อหาของเว็บไซต์ (Website Content)

ตัวชี้แหล่งในอินเทอร์เน็ต หรือ โปรแกรมชี้แหล่งทรัพยากรสากล (Uniform Resource Locator, Universal Resource Locator) ที่เรียกโดยย่อว่า ยูอาร์แอล (URL) คือตัวระบุแหล่งทรัพยากรสากล (URI) ประเภทหนึ่ง ซึ่งใช้สำหรับระบุแหล่งที่อยู่ของทรัพยากรที่ต้องการ และมีกลไกบางอย่างสำหรับดึงข้อมูลทรัพยากรนั้น หรืออีกนัยหนึ่งคือที่อยู่บนอินเทอร์เน็ตของเว็บไซต์หรือข้อมูล ที่เราจะใช้ในการเรียกผ่านเว็บเบราว์เซอร์เพื่อเข้าถึงข้อมูลหรือเว็บไซต์นั้นๆ การสร้าง URL สำหรับเว็บไซต์ให้มีโครงสร้างที่เป็นระเบียบนอกจากจะช่วยให้โปรแกรมค้นหาทำการรวบรวมข้อมูลบนเว็บไซต์ได้ง่ายขึ้น ยังสามารถทำให้ผู้ใช้จดจำและเข้าถึงเว็บไซต์ได้ง่ายขึ้นด้วย

ชื่อของโดเมนก็เช่นเดียวกันมีส่วนสำคัญในการให้อันดับของโปรแกรมค้นหา เนื่องจากโดเมนเป็นเสมือนชื่อที่บ่งบอกว่าเว็บไซต์นั้นๆ เกี่ยวข้องกับอะไร ผู้ใช้งานอินเทอร์เน็ตทั่วไปรวมถึง

โปรแกรมค้นหาจะสามารถเข้าใจได้ว่าเว็บไซต์นั้นมีเนื้อหาอย่างไร ตรงกับความต้องการในการสืบค้นหรือไม่ ผู้ใช้บางรายอาจสร้างลิงค์ไปยังเว็บไซต์โดยการใช้ URL ของหน้าเว็บไซต์นั้นในการเชื่อมโยง หากพิมพ์ URL ผิดไป 1 ตัวก็ไม่สามารถที่จะเชื่อมต่อไปยังลิงค์นั้นได้เลย ดังนั้นหากชื่อโดเมนที่อยู่ใน URL ประกอบด้วยคำที่เกี่ยวข้องกับหน้าเว็บไซต์ ย่อมเป็นประโยชน์มากกว่าการตั้งชื่อโดเมนด้วยคำประหลาด หรือด้วยข้อความที่จำยาก ดังนั้นโครงสร้าง URL หรือชื่อโดเมนที่เหมาะสมควรจะมีลักษณะดังนี้คือ

ใช้คำที่ได้ใจความใน URL โดยการใช้คำที่มาจากเนื้อหาหรือโครงสร้างของเว็บไซต์ใน การสร้าง URL, หลีกเลี่ยงการใช้ URL ที่ยืดยาว มีพารามิเตอร์หรือรหัสเซสชั่นที่ไม่จำเป็น เช่น “www.searchengineoptimization-seothailand.com/a1x41/102125/” , หลีกเลี่ยงการใช้ชื่อหน้าเว็บทั่วไป เช่น "page1.html" และ หลีกเลี่ยงการใช้คำหลักที่ซ้ำซ้อน เช่น “seothailand-seo-seothailand-seo.html”

สร้างโครงสร้างไดเรกทอรี่ที่เรียบง่าย จัดเนื้อหาให้เป็นระเบียบ หลีกเลี่ยงการใช้ไดเรกทอรีย่อยที่ซับซ้อน เช่น "…/dir1/dir2/dir3/dir4/dir5/dir6/page.html" และหลีกเลี่ยงการใช้ชื่อไดเรกทอรีที่ไม่เกี่ยวข้องกับเนื้อหาใน URL

กำหนดรูปแบบ URL เดียวในการเข้าถึงเอกสาร เพื่อทำให้คะแนนความนิยมของหน้าเว็บนั้นกระจายไปตาม URL อื่นๆ หลีกเลี่ยงการผสม URL รูปแบบ www เข้ากับ URL รูปแบบที่ ไม่มี www ในโครงสร้างการเชื่อมโยงภายใน, หลีกเลี่ยงการให้หน้าเว็บต่างๆ จากโดเมนย่อยและไดเรกทอรี่รากเข้าถึงเนื้อหาเดียวกัน เช่น "domain.com/page.htm" และ "sub.domain.com/page.htm" และหลีกเลี่ยงการแทรกตัวพิมพ์ใหญ่ใน URL เพื่อผู้ใช้จะสามารถจดจำได้ง่ายกว่า

นอกจาก URL และชื่อโดเมนของเว็บไซต์ที่มีส่วนสำคัญแล้ว เนื้อหาหรือข้อความต่างๆ ที่เว็บไซต์ต้องการที่จะจะสื่อไปยังผู้อ่าน รวมถึงคำสำคัญ ที่ใช้ในการสืบค้นผ่านโปรแกรมค้นหา ซึ่งทั้ง 2 ล้วนส่งผลต่ออันดับในการค้นหาทั้งสิ้น (Jin Zhang and Alexandra Dimitroff, 2005a)

คำสำคัญที่ใช้ในการสืบค้นจะถูกแบ่งออกเป็น 3 กลุ่มใหญ่ๆ ด้วยกันคือ คำสำคัญหลัก (Major Keyword) โดยคำสำคัญนี้ควรจะเป็นคำที่ไม่เกิน 2-3 คำต่อกัน และควรเป็นคำแรกที่นึกถึงหรือใช้ในการค้นหา หรือบอกถึงเว็บไซต์, คำสำคัญรอง (Minor Keyword) เป็นคำอื่นๆ ที่อธิบายหรือสื่อถึงเว็บไซต์ที่เราทำได้ ในส่วนนี้ไม่ควรมีเกิน 10-20 คำ และคำสำคัญตามฤดูกาล (Seasonal

Keywords) คำสำคัญประเภทนี้มักพบตามเว็บไซต์ขายสินค้าหรือบริการที่เกี่ยวข้องกับฤดูกาล สามารถพบเห็นได้เป็นระยะ เหมาะสำหรับนำมาทำเป็นคำสำคัญในช่วงเวลาสั้นๆ ซึ่งเราควรแทรกคำสำคัญให้อยู่ในเนื้อหาของเว็บไซต์อย่างสม่ำเสมอ การที่เราจงใจปรับแต่งเว็บไซต์มากจนเกินไป อาจจะทำให้ผู้เยี่ยมชมรู้สึกถึงความไม่เป็นธรรมชาติของเนื้อหาหรืออาจทำให้เข้าใจในเนื้อหาได้ยากมากขึ้น ดังนั้นทางโปรแกรมค้นหาจึงมีการตรวจจับความหนาแน่นของคำสำคัญ ว่าเว็บไซต์ไหนมีการปรับแต่งจนดูว่าเป็นการจงใจเพื่อ ให้ได้อันดับดี เว็บไซต์เหล่านั้นจะถูกดันลงไปอยู่ในหลุม สำหรับความหนาแน่นของคำสำคัญ (Keywords Density) ที่ดียังไม่มีการระบุแน่ชัดโดยความหนาแน่นของคำสำคัญคิดได้จาก

Keywords Density = (จำนวนคำค้นหนึ่งๆ /จำนวนคำทั้งหมด) x 100

โดยทั้งจำนวนของคำค้นและจำนวนของคำที่นำมาใช้คิดนั้นขึ้นอยู่กับว่าใช้ในส่วนไหน แต่ที่ยากที่สุด คือ ส่วนเนื้อหา เนื่องจากมีคำเยอะมากปริมาณของคำค้นที่แนะนำให้ใช้อยู่จะประมาน 6-20 เปอร์เซ็นต์ โดยแบ่งเป็น 2 กรณีคือ กรณีเนื้อหาในหน้าเว็บไซต์มากกว่า 600 คำขึ้นไปควรใช้งานไม่เกิน 10 เปอร์เซ็นต์ และ กรณีเนื้อหามีน้อยกว่า 600 คำลงไปก็ไม่ควรเกิน 20 เปอร์เซ็นต์ สาเหตุที่ต้องมีสองกรณีเนื่องจาก ในเนื้อหาสั้นๆ ทำให้มีตัวหารน้อย ตัวเลขของความหนาแน่นก็จะมากขึ้นด้วย จะมีคำบางกลุ่มคำ หรือบางประโยคที่โปรแกรมค้นหาไม่นำมาใช้คิดความหนาแน่น ได้แก่ stop words และ stop phrases เพื่อเป็นการคัดกรองคำที่ไม่สำคัญหรือ ไม่เกี่ยวข้องกับโปรแกรมค้นหาออกไป ตัวอย่างของ stop words ได้แก่ a, the, is, am, are, so, that, this เป็นต้น ส่วน stop phrases เช่น ในประโยค “ This is a seo book” คำว่า “This is” จะถือเป็น stop phrases ส่วนในภาษาไทยนั้นยังไม่มีการระบุที่ชัดเจนเนื่องจากการเขียนประโยคในภาษาไทยจะเขียนติดกันไปทั้งหมดทำให้โปรแกรมค้นหาไม่สามารถที่จะแยกหรือตัดคำต่างๆ ในประโยคออกไปได้

เนื้อหา Content การสร้างเนื้อหาที่ดึงดูดความสนใจและเป็นประโยชน์ถือเป็นปัจจัยสำคัญในการสร้างเว็บไซต์ของคุณเหนือปัจจัยอื่นๆ ซึ่งแนวทางในการสร้างเว็บไซต์ให้เหมาะสมมีด้วยกันดังนี้คือ เขียนข้อความที่อ่านง่าย, หลีกเลี่ยงการเขียนข้อความอย่างสะเพร่า สะกดคำผิด ใช้หลักไวยากรณ์ไม่ถูกต้อง, หลีกเลี่ยงการฝังข้อความลงในรูปภาพเป็นเนื้อหาแบบข้อความ (เนื่องจากผู้ใช้อาจต้องการ คัดลอกและวางข้อความนั้น ซึ่งเครื่องมือค้นหาไม่สามารถอ่านได้) ครอบคลุมประเด็นเนื้อหา มีการจัดระเบียบเนื้อหาทำให้สามารถรับรู้ถึงจุดเริ่มต้นของประเด็นเนื้อหาและจุดสิ้นสุดได้ ซึ่งการแบ่งเนื้อหาออกเป็นส่วนๆ จะช่วยให้ผู้ใช้พบเนื้อหาที่ต้องการได้รวดเร็วขึ้น,

หลีกเลี่ยงการอัดเนื้อหาของประเด็นต่างๆ ไว้ในหน้าเว็บเดียวกันโดยไม่ได้จัดย่อหน้า แบ่งหัวข้อย่อย หรือการแบ่งเค้าโครงหน้า

เลือกใช้คำหรือภาษาที่สอดคล้องกับเนื้อหา สร้างเนื้อหาแปลกใหม่ไม่ซ้ำใคร Unique Content หลีกเลี่ยงการปรับปรุงข้อมูลเดิม หรือคัดลอกข้อความต่างๆ หลีกเลี่ยงการใส่เนื้อหาที่ทำซ้ำ หรือเนื้อหาที่ใกล้เคียงกับเนื้อหาเดิมลงทั่วทั้งเว็บไซต์ เนื่องจากโลกปัจจุบันทำให้ทุกอย่างดูจะสะดวกไปหมดการคัดลอกเนื้อหาเป็นเรื่องธรรมดาทำให้มีข้อมูลบนอินเทอร์เน็ตที่ซ้ำๆ กันเป็นจำนวนมาก ทางผู้พัฒนาโปรแกรมค้นหาจึงได้มีนโยบายในการลดคะแนนที่ให้กับเว็บไซต์ที่มีเนื้อหาซ้ำๆ รวมถึงอาจไม่เก็บข้อมูลของเว็บไซต์ที่ซ้ำเหล่านั้น ทำให้เว็บไซต์เหล่านั้นถูกมองเป็นเว็บขยะ ทำให้เว็บไซต์เหล่านี้ไม่สามารถทำอันดับได้ในโปรแกรมค้นหา

เสนอเนื้อหาหรือบริการพิเศษ สร้างความแตกต่างให้กับเว็บไซต์ เช่น นำเสนอผลงานวิจัยที่ได้คิดขึ้นเอง หรือนำเสนอข่าวสารต่างๆ ที่น่าสนใจ สร้างเนื้อหาโดยยึดผู้ใช้ของคุณเป็นสำคัญ ไม่ใช่เครื่องมือค้นหา โดยการออกแบบเว็บไซต์ให้สอดรับกับความต้องการของผู้เข้าชม และการจัดการเว็บไซต์ให้โปรแกรมค้นหาเข้าถึงได้ง่าย, หลีกเลี่ยงการสอดแทรกคำหลักที่ไม่จำเป็นโดยคำนึงถึงเครื่องมือค้นหาเป็นหลัก เนื่องจากเป็นข้อมูลที่มากเกิน หรือไม่ตรงกับความต้องการของผู้ค้นหา และหลีกเลี่ยงการซ่อนข้อความจากผู้ใช้เพื่อหลอกลวง แต่แสดง ข้อความนั้นแก่เครื่องมือค้นหา

3. ลิงก์ออกไปยังเว็บอื่น (Outbound-Links)

การทำลิงก์ออกไปยังเว็บอื่น ส่งผลต่อการทำ SEO เช่นกันเพราะโปรแกรมค้นหามองว่าการทำลิงก์ออกเป็นการให้แต้มที่ตัวเองเก็บสะสมไว้ ส่งต่อไปยังเว็บปลายทาง ดังนั้นหากเว็บเราทำลิงก์ออกไปยังเว็บไซต์อื่นจำนวนมากหมายความว่าเราได้มอบแต้มที่เราเก็บสะสมไว้ ให้กับคนอื่น โดยเฉพาะอย่างยิ่งหากเว็บปลายทางนั้นเป็นเว็บไซต์ที่ทางโปรแกรมค้นหา มองว่าเป็นเว็บไซต์ที่ไม่เหมาะสม หรือเข้าข่าย SPAM ทางโปรแกรมค้นหาจะลดเครดิตของเว็บไซต์เราทันที ดังนั้นในการสร้างลิงก์เข้าและลิงก์ออกควรจะควบคุมปริมาณลิงก์ให้สมดุลกัน หลีกเลี่ยงการสร้างลิงก์ออกที่มากเกินไป

Off-Page Factors การปรับแต่งเว็บจากภายนอก โดยจะมีองค์ประกอบอยู่ 3 ส่วนด้วยกันคือ

1. ลิงค์ขาเข้า (Link Building)

โปรแกรมค้นหาส่วนใหญ่พยายามหาระบบการจัดอันดับที่มีความยุติธรรม และส่งผลประโยชน์ต่อผู้ใช้ให้มากที่สุด ซึ่งก็เป็นเรื่องสมเหตุสมผลที่เว็บที่มีประโยชน์ หรือเป็นที่ชื่นชอบจะได้รับการทำลิงค์เข้าหาจำนวนมาก เมื่อทางโปรแกรมค้นหาพบว่ามีการทำลิงค์ให้กับเว็บไซต์ใดเว็บไซต์หนึ่งจำนวนมาก ก็จะตั้งสมมุติฐานขึ้นมาทันทีว่าเว็บนั้นจะต้องมีดีอะไร ผลก็คือ เว็บไซต์นั้นจะถูกโปรแกรมค้นหาจัดให้อยู่ในอันดับที่สูงกว่าเว็บไซต์อื่นๆ ในกลุ่มเดียวกันการทำลิงค์ (Link Building) ถือว่าเป็นเรื่องสำคัญสำหรับการทำ SEO ในยุคปัจจุบัน Google และ Yahoo ต่างนำเอาปัจจัยนี้มาเป็นปัจจัยหลักในการจัดอันดับ ดังนั้นผู้พัฒนาเว็บไซต์ควรหาลิงค์เข้าเว็บไซต์อยู่เสมอ ไม่ว่าจะเป็นการ submit directory หรือการใช้ Social Bookmark ก็สามารถเพิ่มลิงค์ให้กับเว็บไซต์ได้โดยไม่ต้องเสียค่าใช้จ่ายใดๆ หรือยอมที่จะเสียเงินซื้อลิงค์จากเว็บมีชื่อเสียง เพื่อช่วยดันอันดับเว็บไซต์ให้สูงขึ้น แต่แน่นอนว่าลิงก์ ดังกล่าว โปรแกรมค้นหาย่อมไม่ชอบ และอาจจะไม่นำลิงก์นั้นมาพิจารณาในการให้คะแนนการจัดอันดับ (Matt Cutts, 2009c; Rand Fishkin, 2009)

2. อายุของโดเมน (Age of Domain)

นอกจากชื่อของโดเมนแล้ว อายุของโดเมนก็มีส่วนในการให้อันดับของโปรแกรมค้นหาเช่นกัน จะเห็นได้ว่าอย่าง Google เอาเรื่องของอายุโดเมนมาใช้ในการจัดอันดับด้วย สังเกตได้จากโดเมนที่จดไว้ตั้งแต่ปี 2001 กับที่เพิ่งจดใหม่เมื่อลองทำการเขียนบทความ และนำออกแสดงพร้อมกันจะพบว่า Google จะมาทำการ Index บทความของเว็บไซต์ที่มีโดเมนเก่ากว่าก่อน แต่ทั้งนี้ขึ้นกับปัจจัยอื่นๆ ประกอบด้วย

3. การใช้งานของผู้ใช้ (Behavior of Visitors)

โปรแกรมค้นหาได้พยายามเข้ามามีบทบาทกับการใช้งานของผู้ใช้งานมากขึ้น มีการเรียนรู้อยู่ตลอดเวลา โดยศึกษาพฤติกรรมการใช้งานของผู้ใช้โปรแกรมค้นหา ว่ามีแนวโน้มไปอย่างไร หากสามารถทำเว็บไซต์ให้เป็นที่นิยมของผู้ใช้งานแล้วก็มีโอกาสที่โปรแกรมค้นหาจะจัดอันดับเว็บไซต์นั้นๆ ให้อยู่ในตำแหน่งที่สูงกว่าเว็บไซต์อื่นๆ

ซึ่งจากคุณลักษณะของเว็บไซต์และองค์ประกอบของเว็บไซต์ที่ดีและหลักการทำ SEO จึงเป็นที่มาของ Tag HTML ที่เลือกศึกษารวมถึงคุณลักษณะต่างๆ ของเอกสาร HTML ในครั้งนี้

2.4 ความน่าเชื่อถือของข้อมูลสุขภาพบนอินเทอร์เน็ตและการประเมินคุณภาพเว็บไซต์

จากการศึกษาพบว่าข้อมูลทางสุขภาพที่นำเสนอทางเว็บไซต์จำนวนมาก มีข้อมูลไม่ถูกต้องในการให้คำแนะนำการดูแลสุขภาพ การใช้ยา หรือมีการโฆษณาสรรพคุณของอาหารหรือยาที่เกินจริงก่อให้เกิดผลเสียต่อผู้บริโภค เช่น เสียทรัพย์โดยไม่จำเป็น ที่สำคัญคือการเสียโอกาสที่จะได้รับการรักษาอย่างถูกต้อง จึงมีหลายหน่วยงานได้จัดทำเครื่องมือ หรือกำหนดมาตรฐานที่ใช้ในการประเมินคุณภาพของเว็บไซต์ด้านสุขภาพ เช่น HON (Health on the Net Foundation) เป็นองค์กรที่มีวัตถุประสงค์เพื่อประเมินมาตรฐานของเว็บไซต์ด้านสุขภาพ, WWW.CYBERGUIDE เป็นเครื่องมือหนึ่งที่ใช้เป็นแนวทางในการประเมินคุณภาพของเว็บไซต์ และDISCERN เป็นเครื่องมือที่ใช้ในการประเมินคุณภาพข้อมูลที่เกี่ยวกับการรักษา Vincent Morel และคณะได้ทำการศึกษาความน่าเชื่อถือของเว็บไซต์ที่ให้ข้อมูลเรื่องโรคอารมณ์แปรปรวนโดยใช้ HON และ DISCERN จากเว็บไซต์ 80 เว็บไซต์พบว่ามีเพียง 34 เว็บไซต์ที่ผ่านเกณฑ์ โดยพบว่าเว็บไซต์ที่มีสัญลักษณ์ HON นำเสนอข้อมูลที่มีคุณภาพไม่แตกต่างจากเว็บไซต์ที่ไม่ผ่านการประเมินโดย HON ส่วนผลการประเมินโดยใช้ DISCERN พบว่าคุณภาพของข้อมูลอยู่ในเกณฑ์ดีแต่ควรปรับปรุงเนื้อหาให้อ่านง่ายขึ้น (Morel and others, 2008: 265-269)

งานวิจัยนี้จึงเลือกใช้ DISCERN เป็นเครื่องมือในการประเมินคุณภาพของเว็บไซต์ เนื่องจากเป็นเครื่องมือที่ประเมินทั้งในส่วนของความน่าเชื่อถือของเว็บไซต์และแหล่งข้อมูล รวมถึงความถูกต้องของเนื้อหา ซึ่ง DISCERN เป็นเครื่องมือที่มองในมุมมองของผู้ป่วย ซึ่งมหาวิทยาลัยออกซ์ฟอร์ดเป็นผู้พัฒนาขึ้น โดยออกแบบให้ผู้ใช้เป็นผู้ประเมินความเหมาะสมของเนื้อหาควบคู่กับความต้องการของผู้ใช้ โดย DISCERN แบ่งเป็น 3 ส่วน มีด้วยกันทั้งหมด 16 คำถามคือ (DISCERN, 2004)

คำถามที่ 1-8 เป็นการประเมินความน่าเชื่อถือของเว็บไซต์ ความน่าเชื่อถือของแหล่งข้อมูล

คำถามที่ 9-15 เป็นการประเมินความน่าเชื่อถือและความถูกต้องเกี่ยวกับเนื้อหาข้อมูลทางด้านของการรักษา

คำถามที่ 16 เป็นการประเมินคุณภาพโดยรวมของเว็บไซต์

โดยในการให้คะแนนจะจัดเป็นระดับคะแนน 5 ระดับ ดังนี้

ระดับ 5 คือมีข้อมูลนั้นสมบูรณ์ ครบถ้วน สามารถตอบคำถามข้อนั้นๆ ได้ชัดเจน

ระดับ 2-4 คือถ้าหากเว็บไซต์นั้นมีสามารถตอบคำถามข้อนั้นๆ ได้ โดยทั้งนี้ขึ้นอยู่กับวิจารณญาณของผู้ประเมิน โดยอาจจะหักคะแนนออกไปตามข้อบกพร่องต่างๆ

ระดับ 1 คือเว็บไซต์นั้นๆ ไม่มีข้อมูลที่สามารถตอบคำถามในข้อนั้นๆ ได้

ซึ่งเมื่อรวมทั้งหมด 16 คำถามแต่ละเว็บไซต์จะมีคะแนนเต็มคือ 80 คะแนน

บทที่ 3

วิธีดำเนินการวิจัย

การศึกษาคุณลักษณะโครงสร้างของเอกสาร HTML จากเว็บไซต์ที่ให้ข้อมูลยา เป็นงานวิจัยเชิงทดลองที่ต้องการทราบถึงคุณลักษณะโครงสร้างของเอกสาร HTML ที่เป็นปัจจัยต่อคุณภาพของเว็บไซต์ที่ให้ข้อมูลยาที่เป็นภาษาอังกฤษ และมีผลต่อการปรากฏในโปรแกรมค้นหา

3.1 อุปกรณ์และเครื่องมือที่ใช้ในการทดลอง

PC: Intel® Core™2 Duo CPU 2.00 GHz RAM 2.00 Gb

ระบบปฏิบัติการ Microsoft Windows XP พร้อมการเชื่อมต่อเครือข่ายอินเทอร์เน็ต

Web Browser (โปรแกรมค้นดูเว็บ): Mozilla Firefox, Internet explorer, Google Chrome

Search Engine (โปรแกรมค้นหา): Google, Yahoo และ MSN (Bing)

Microsoft Office: Excel, Word

Web Page Analyzer - 0.98

Markup Validation Service

CSS Validation Service

โปรแกรม PSPP

3.2 วิธีการศึกษา (Methodology)

ในส่วนของวิธีการศึกษาสามารถแบ่งออกเป็นขั้นตอนย่อย ดังนี้คือ

3.2.1 การคัดเลือกเว็บไซต์

คัดเลือกเว็บไซต์ที่จะทำการศึกษาโครงสร้างตามเกณฑ์ที่กำหนด และทำการสุ่มเอกสาร HTML โดยในการศึกษานี้ได้ทำการเลือกเว็บไซต์ที่ให้ข้อมูลยาที่เป็นภาษาอังกฤษ โดยสืบค้นเว็บไซต์ข้อมูลยาด้วยโปรแกรมค้นหาทั้งหมด 3 ชนิดคือ Google Yahoo และ MSN (Bing) และเลือกเว็บไซต์ที่อยู่อันดับต้นๆ ของหน้าแสดงผล (SERP) โดยเลือก 40 รายการแรกที่ปรากฏ และเว็บไซต์ที่อยู่อันดับท้ายๆ ของหน้าแสดงผล โดยเลือกอับรายการ ที่ 301 ถึงอันดับรายการที่ 340 ในแต่ละโปรแกรมค้นหาทั้งอันดับต้นและอันดับท้าย เมื่อทำการสืบค้นแล้วจะได้จำนวนรายการเว็บเพจ 80 รายการต่อ 1 เครื่องมือค้นหา และ 1 คำสำคัญ โดยคำสำคัญที่ใช้ในการสืบค้นแบ่งออกเป็น 2 กลุ่มคือ คำสืบค้นทั่วๆไป ได้แก่ “drug information” และ “medical information” คำสืบค้นที่เป็นชื่อเฉพาะทางยา ได้แก่ "acetaminophen" "amoxicillin" "celecoxib" "hydrochlorothiazide" "lipitor" และ "spironolactone" เมื่อทำการสืบค้นเสร็จสิ้น จะได้รายการเว็บเพจทั้งสิ้นจำนวน 1,920 หน้าเว็บเพจโดยช่วงเวลาที่ทำการเก็บข้อมูลคือตั้งแต่พฤษภาคม - ตุลาคม 2553 จากนั้นนำมาทำการคัดเลือกตามคุณลักษณะดังต่อไปนี้ คือ

เว็บไซต์ที่ซ้ำกัน เลือกมาเพียง 1 เว็บไซต์
เว็บไซต์ที่เป็นภาษาอังกฤษ
เว็บไซต์ที่ให้ข้อมูลทางยา (ไม่ใช่เว็บไซต์ขายสินค้าหรือบทคัดย่อของวารสาร)
เว็บไซต์ที่ไม่ใช่ลักษณะเป็นเว็บบอร์ด
เว็บไซต์ที่ไม่ใช่เว็บที่ให้ข้อมูลยาเสพติด
เว็บไซต์ที่สามารถเข้าถึงในช่วงเวลาที่ทำการประเมิน
เว็บไซต์ที่ไม่ต้องเสียค่าบริการ หรือต้องสมัครสมาชิก

เมื่อได้เว็บไซต์ที่ต้องการจะทำการสุ่มเลือกเว็บเพจจำนวน 20 เว็บเพจ จากเว็บไซต์ที่ทำการศึกษาโดยจะกำหนดให้เป็นหน้าโฮมเพจ 1 หน้า และหน้าอื่นๆ จำนวน 19 หน้าเพื่อเป็นตัวแทนของเว็บไซต์

3.2.2 การประเมินคุณภาพของเว็บไซต์

ประเมินคุณภาพของเนื้อหาเว็บไซต์ตามหลักของเครื่องมือประเมินคุณภาพ ได้แก่ DISCERN

3.2.3 การจัดกลุ่มของเว็บไซต์

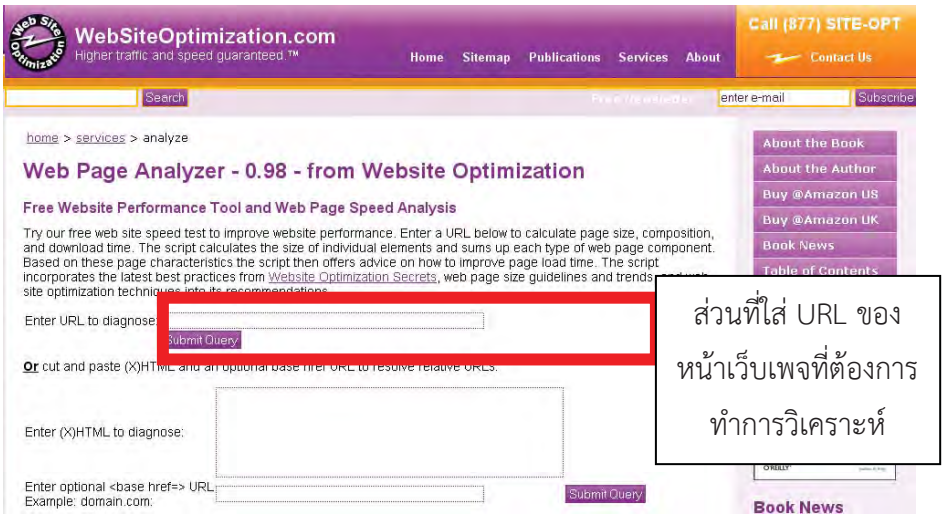
ทำการจัดกลุ่มเว็บไซต์ตามการประเมินคุณภาพเว็บไซต์โดยเพื่อให้ง่ายต่อการจำแนก แบ่งออกเป็น 2 ระดับคือ เว็บไซต์ที่มีคะแนนตั้งแต่ 60 เปอร์เซ็นต์ ขึ้นไปจัดเป็นเว็บไซต์กลุ่ม A และเว็บไซต์ที่มีคะแนนน้อยกว่า 60 เปอร์เซ็นต์ จัดเป็นเว็บไซต์กลุ่ม B

3.2.4 การเก็บรวบรวมข้อมูลตัวแปรต่างๆ ดังนี้

ทำการสุ่มเอกสาร HTML ที่ให้ข้อมูลด้านยาของแต่ละหน้าเว็บไซต์จากเว็บไซต์ทั้งหมดที่ทำการศึกษา จำนวนเว็บไซต์ละ 20 หน้า เพื่อใช้เป็นตัวแทนของเว็บไซต์ ในส่วนของตัวแปรต่างๆ ที่นำมาใช้ในการวิเคราะห์ จะคัดเลือกจากการมีความสัมพันธ์กับการให้อันดับของโปรแกรมค้นหา (Search Engine Optimization) และเกี่ยวเนื่องกับคุณภาพของเว็บไซต์ โดยคัดเลือกตัวแปรที่เป็น Tag HTML จาก HTML Reference

3.2.4.1 ปริมาณ Tag ที่ทำการศึกษาที่ปรากฏบนเอกสาร HTML ในส่วนของ Tag ที่ปรากฏบนเอกสาร HTML ทั้งหมด จะใช้วิธีการคัดแยกด้วยมือ โดยการคัดแยกทีละหน้าเว็บเพจจากเว็บเพจที่ทำศึกษาทั้งหมด จากนั้นใช้ Microsoft Excel ในการจำแนกกลุ่ม รวบรวมและนับจำนวน ของ Tag ต่างๆ

3.2.4.2 ขนาดของหน้าเอกสาร HTML ขนาดของภาพที่พบภายในเว็บเพจ และเวลาที่ใช้ในการดาวน์โหลด ในส่วนนี้จะทำการเก็บรวบรวมโดยใช้ Web Page Analyzer - 0.98 จากเว็บไซต์ www.websiteoptimization.com ซึ่งเป็นเครื่องมือช่วยในการจำแนกขนาดและความเร็วของเว็บไซต์ ผ่านอินเทอร์เน็ต โดยในส่วนของขนาดของหน้าเอกสาร HTML จะเลือกเก็บในส่วนของ Total size , Total Image size และในส่วนของเวลาที่ใช้ในการดาวน์โหลด จะทำการเลือกเก็บในส่วนของค่าที่ Connection Rate 1.44Mb.



รูปที่ 2 แสดง Web Page Analyzer - 0.98 [Online] http://www.websiteoptimization.com/services/analyze/

3.2.4.3 Doctype (ประเภทของเอกสาร HTMLที่ใช้), Encoding (การเข้ารหัสภาษา), HTML error และ HTML warning ในการศึกษาส่วนนี้จะใช้ Markup Validation Service ของ W3C เป็นเครื่องมือในการจำแนกประเภทของเอกสาร HTML ที่ใช้, การเข้ารหัสภาษา, ข้อผิดพลาดต่างๆ ในส่วนของ HTML ทั้ง error และ warning

รูปที่ 3 แสดง Markup Validation Service [Online] http://validator.w3.org/

3.2.4.4 CSS error ในการศึกษาคุณลักษณะเรื่อง CSS error จะใช้ CSS Validation Service ของ W3C เป็นเครื่องมือในการจำแนกข้อผิดพลาดต่างๆ ในส่วนของ CSS โดยจะเก็บเฉพาะในส่วนของ error

รูปที่ 4 แสดง CSS Validation Service [Online] http://jigsaw.w3.org/css-validator/

3.2.4.5 Internal- external link (จุดเชื่อมโยงภายใน-ภายนอก) ในการศึกษาจุดเชื่อมโยงภายใน-ภายนอก จะใช้โปรแกรม SeoQuake SEO extension version 2.5.9 ซึ่งเป็น Add-ons for Firefox เป็นเครื่องมือในการนับและรวบรวม จุดเชื่อมโยงภายในและภายนอก

รูปที่ 5 แสดงโปรแกรม SeoQuake SEO extension version 2.5.9 [Online] http://www.seoquake.com/

3.2.4.6 Google Page Rank หรือค่าลำดับคะแนนที่ Google ประเมินให้กับคุณภาพของเนื้อหา ในหน้าเว็บเพจแต่ละหน้า ที่อยู่ในเว็บไซต์โดยการศึกษานี้จะใช้ google toolbar ซึ่งติดตั้งบนเว็บเบราเซอร์ Google Chrome ในการหาค่าดังกล่าว

รูปที่ 6 แสดงการหาค่า PR ด้วย Google toolbar [Online]
http://www.google.com/toolbar/ie/index.html

3.2.5 การวิเคราะห์คุณลักษณะของเว็บไซต์ด้วยสถิติเชิงพรรณนา

ศึกษาปริมาณของ Tag HTML และคุณลักษณะต่างๆของเอกสาร HTML เฉลี่ย, จุดเชื่อมโยงภายนอกและจุดเชื่อมโยงภายใน, ศึกษาขนาดของหน้าเว็บเพจเฉลี่ยในแต่ละเว็บไซต์ และเวลาที่ใช้ในการดาวน์โหลด, ศึกษาประเภทของเอกสาร HTML, การเข้ารหัสภาษา, ศึกษาปริมาณของค่าความผิดพลาดเฉลี่ยของเว็บไซต์ที่ให้ข้อมูลยาที่เป็นภาษาอังกฤษ โดยศึกษาทั้งในส่วนของ HTML error, HTML warning และ CSS error, ศึกษา Page Rank, ศึกษาขนาดภาพที่พบ, ขนาดของ CSS, หลังจากทำการเก็บรวบรวมข้อมูล และจำแนกข้อมูลตัวแปรออกเป็นกลุ่มย่อยต่างๆ แล้วทำการแยกศึกษาเป็นหัวข้อได้แก่ เว็บไซต์อันดับต้น เว็บไซต์อันดับท้าย เว็บไซต์กลุ่ม A และเว็บไซต์กลุ่ม B

3.2.6 การเปรียบเทียบค่าเฉลี่ยของตัวแปรด้วยวิธี Independent t-test

ศึกษาเปรียบเทียบค่าเฉลี่ยของตัวแปรระหว่างเว็บไซต์อันดับต้นกับเว็บไซต์อันดับท้าย และเปรียบเทียบค่าเฉลี่ยของตัวแปรระหว่างเว็บไซต์กลุ่ม A และเว็บไซต์กลุ่ม B ด้วยสถิติแบบ Independent t-test ที่ระดับความเชื่อมั่น 95 เปอร์เซ็นต์ โดยใช้ PSPP เป็นเครื่องมือในการคำนวณทางสถิติ

วิธีการคำนวณ Independent t-test

สูตรการคำนวณค่า T-Test

1. Separate variance t-test ในกรณีที่ assume ว่า $\sigma^2_1 \neq \sigma^2_2$

ใช้สูตร $$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

$$df = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{\left(\frac{S_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{S_2^2}{n_2}\right)^2}{n_2 - 1}}$$

เมื่อ $\bar{X}_1$, $\bar{X}_2$ แทนค่าเฉลี่ยของกลุ่มตัวอย่างกลุ่มที่ 1, 2

S_1^2 , S_2^2 แทนความแปรปรวนของกลุ่มตัวอย่างกลุ่มที่ 1, 2

n_1 , n_2 แทนขนาดของกลุ่มตัวอย่างกลุ่มที่ 1, 2

df แทนชั้นแห่งความเป็นอิสระ (degree of freedom)

ข้อตกลงเบื้องต้นของการใช้สูตรนี้

1. กลุ่มตัวอย่างทั้งสองกลุ่มต้องเป็นอิสระจากกัน และต้องได้มาจากการสุ่ม จากกลุ่มประชากรที่มีการแจกแจงเป็นโค้งปกติ
2. คุณลักษณะที่ต้องการศึกษาภายในกลุ่มต้องเป็นอิสระจากกัน
3. มีเหตุผลพอที่จะเชื่อได้ว่าความแปรปรวนของกลุ่มประชากรทั้งสองกลุ่ม ไม่เท่ากัน ($\sigma^2_1 \neq \sigma^2_2$)
4. ขนาดของกลุ่มตัวอย่างไม่เท่ากัน ($n_1 \neq n_2$)

2. Pooled t-test ในกรณีที่ assume ว่า $\sigma^2_1 = \sigma^2_2$

ใช้สูตร $$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{S_p^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$$

$$df = n_1 + n_2 - 2$$

เมื่อ $\bar{x}_1$, $\bar{x}_2$ แทนค่าเฉลี่ยของกลุ่มตัวอย่างกลุ่มที่ 1, 2

S_p^2 แทนความแปรปรวนร่วม (Pooled variance)

$$S_p^2 = \frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}$$

n_1, n_2 แทนขนาดของกลุ่มตัวอย่างกลุ่มที่ 1, 2

df แทนชั้นแห่งความเป็นอิสระ (degree of freedom)

ข้อตกลงเบื้องต้นของการใช้สูตรนี้

1. กลุ่มตัวอย่างทั้งสองกลุ่มต้องเป็นอิสระจากกัน และต้องได้มาจากการสุ่ม จากกลุ่มประชากรที่มีการแจกแจงเป็นโค้งปกติ
2. คุณลักษณะที่ต้องการศึกษาภายในกลุ่มต้องเป็นอิสระจากกัน
3. ความแปรปรวนของกลุ่มประชากรทั้งสองกลุ่มเท่ากัน ($\sigma^2 1 = \sigma^2 2$) แต่ไม่ทราบค่า

หมายเหตุ สูตรนี้เขียนได้อีกลักษณะหนึ่งดังนี้

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$$

$$df = n_1 + n_2 - 2$$

ในกรณีที่ไม่สามารถตัดสินใจได้ว่า $\sigma^2_1 = \sigma^2_2$ หรือ $\sigma^2_1 \neq \sigma^2_2$ ควรต้องมีการทดสอบความแปรปรวนก่อน ด้วยการทดสอบค่าเอฟ หรือ F-test

ในการทดสอบความแปรปรวน ว่า $\sigma^2_1 = \sigma^2_2$ หรือไม่ ตั้งสมมุติฐานดังนี้

$$H_0: \sigma_1^2 = \sigma_2^2$$

$$H_1: \sigma_1^2 \neq \sigma_2^2$$

สูตรการทดสอบค่าเอฟ F-test คือ

$$F = \frac{S_1^2}{S_2^2} \quad \text{เมื่อ } S_1^2 > S_2^2 \quad \text{โดยมี } df_1 = n_1\text{-}1,\ df_2 = n_2\text{-}1$$

$$\text{หรือ } F = \frac{S_2^2}{S_1^2} \quad \text{เมื่อ } S_2^2 > S_1^2 \quad \text{โดยมี } df_1 = n_2\text{-}1,\ df_2 = n_1\text{-}1$$

นั่นคือ เอาความแปรปรวนที่มีค่ามากกว่าเป็นตัวเศษ และให้ df ของตัวเศษเป็น df_1 df ของตัวส่วน เป็น df_2 เสมอ นำค่า F ที่คำนวณได้จากสูตรไปเปรียบเทียบกับ ค่าวิกฤต F ที่ได้จากตาราง Critical values of F (ตารางที่ 32 ภาคผนวก)

ถ้า F ที่ได้จากการคำนวณ ≥ ค่าวิกฤต F ที่ได้จากตาราง แสดงว่าความแปรปรวนของประชากรสองกลุ่มแตกต่างกัน ($\sigma_1^2 \neq \sigma_2^2$) จะต้องใช้สูตรการทดสอบ t-test ในข้อ 1. Non pooled t-test หรือ Separate variance t-test

ถ้า F ที่ได้จากการคำนวณ < ค่าวิกฤต F ที่ได้จากตาราง แสดงว่าความแปรปรวนของประชากรสองกลุ่มเท่ากัน ($\sigma_1^2 = \sigma_2^2$) จะต้องใช้สูตรการทดสอบ t-test ในข้อ 2. Pooled variance t-test หรือ Pooled t-test

บทที่ 4

ผลการวิเคราะห์ข้อมูล

จากการรวบรวมข้อมูลและศึกษาเว็บไซต์ที่ให้ข้อมูลด้านยาที่เป็นภาษาอังกฤษ ได้ผลการศึกษาและวิเคราะห์แบ่งออกเป็นหัวข้อต่างๆ ดังนี้คือ

4.1 การคัดเลือกเว็บไซต์

จากการศึกษาเมื่อทำการคัดเลือกเว็บไซต์ตามเกณฑ์ที่กำหนดไว้ในหัวข้อที่ 3.2.1 แล้วจะได้ผลการคัดเลือกดังตารางที่ 1

ตารางที่ 1 แสดงผลการคัดเลือกเว็บไซต์

รายชื่อเว็บไซต์	
เว็บไซต์อันดับต้น	เว็บไซต์อันดับท้าย
www.bccancer.bc.ca	drugs.medsort.com
www.cancer.gov	www.allina.com
www.drugdigest.org	www.aol.com
www.druginformationonline.org	www.canada.com
www.drugs.com	www.drug3k.com
www.empr.com	www.everyonehealthy.com
www.familydoctor.org	www.grubbscare.com
www.health.yahoo.com	www.mdonweb.org
www.healthline.com	www.medindia.net
www.healthsquare.com	www.ralphs.com
www.mayoclinic.com	
www.medical-clinic.org	
www.medicinenet.com	
www.medinfo.co.uk	

ตารางที่ 1 แสดงผลการคัดเลือกเว็บไซต์ (ต่อ)

รายชื่อเว็บไซต์	
เว็บไซต์อันดับต้น	เว็บไซต์อันดับท้าย
www.muschealth.com	
www.myoptumhealth.com	
www.nhsdirect.nhs.uk	
www.nlm.nih.gov	
www.parkinsons.org.uk	
www.patient.co.uk	
www.pdrhealth.com	
www.riteaid.com	
www.rxlist.com	
www.walgreens.com	
www.webmd.com	

จากตารางมีเว็บไซต์ที่ได้รับการคัดเลือกจำนวนทั้งสิ้น 35 เว็บไซต์ โดยจัดเป็นเว็บไซต์ที่อยู่อันดับต้นๆ จำนวน 25 เว็บไซต์ และเว็บไซต์ที่อยู่อันดับท้ายๆ จำนวน 10 เว็บไซต์

4.2 การประเมินคุณภาพเว็บไซต์

เมื่อทำการคัดเลือกเว็บไซต์ที่ให้ข้อมูลด้านยาที่เป็นภาษาอังกฤษตามเกณฑ์ที่กำหนดแล้ว ในงานวิจัยนี้ได้กำหนดให้ประเมินคุณภาพเว็บไซต์ที่ได้ด้วย DISCERN ซึ่งเป็นเครื่องมือที่ใช้ในการประเมินคุณภาพเว็บไซต์ที่ให้ข้อมูลทางสุขภาพที่ได้รับความน่าเชื่อถือและเป็นที่นิยม โดยใช้ผู้ทำการประเมินคุณภาพ 2 คนจากนั้นนำผลที่ได้มาทำการคิดค่าเฉลี่ยและคำนวณเป็นเปอร์เซ็นต์ ได้ผลการประเมินดังตารางที่ 2

ตารางที่ 2 แสดงผลการประเมินคุณภาพเว็บไซต์ด้วย DISCERN

รายชื่อเว็บไซต์	คะแนน		คะแนนเฉลี่ย	คิดเป็นเปอร์เซ็นต์
	ประเมิน 1	ประเมิน 2	เต็ม 80	
www.bccancer.bc.ca	65	72	68.5	86
www.cancer.gov	46	49	47.5	59
www.drugdigest.org	50	52	51	64
www.druginformationonline.org	28	21	24.5	31

ตารางที่ 2 แสดงผลการประเมินคุณภาพเว็บไซต์ด้วย DISCERN (ต่อ)

รายชื่อเว็บไซต์	คะแนน		คะแนนเฉลี่ย	คิดเป็นเปอร์เซ็นต์
	ประเมิน 1	ประเมิน 2	เต็ม 80	
www.drugs.com	66	65	65.5	82
www.empr.com	57	61	59	74
www.familydoctor.org	50	47	48.5	61
www.health.yahoo.com	50	53	51.5	64
www.healthline.com	60	63	61.5	77
www.healthsquare.com	63	66	64.5	81
www.mayoclinic.com	68	70	69	86
www.medical-clinic.org	45	50	47.5	59
www.medicinenet.com	58	62	60	75
www.medinfo.co.uk	51	54	52.5	66
www.muschealth.com	43	48	45.5	57
www.myoptumhealth.com	45	49	47	59
www.nhsdirect.nhs.uk	60	62	61	76
www.nlm.nih.gov	55	59	57	71
www.parkinsons.org.uk	35	41	38	48
www.patient.co.uk	40	45	42.5	53
www.pdrhealth.com	42	45	43.5	54
www.riteaid.com	42	47	44.5	56
www.rxlist.com	67	69	68	85
www.walgreens.com	31	37	34	43
www.webmd.com	51	60	55.5	69
drugs.medsort.com	55	64	59.5	74
www.allina.com	51	56	53.5	67
www.aol.com	46	50	48	60
www.canada.com	33	48	40.5	51
www.drug3k.com	29	37	33	41
www.everyonehealthy.com	48	52	50	63
www.grubbscare.com	33	47	40	50
www.mdonweb.org	30	39	34.5	43
www.medindia.net	32	38	35	44
www.ralphs.com	45	59	52	65

จากการประเมินคุณภาพเว็บไซต์ที่ให้ข้อมูลด้านยาที่เป็นภาษาอังกฤษ จำนวน 35 เว็บไซต์พบว่า มีเว็บไซต์ที่มีคะแนนมากกว่า 80 เปอร์เซ็นต์ จำนวน 5 เว็บไซต์ คิดเป็น 14.29 เปอร์เซ็นต์ เว็บไซต์ที่มีคะแนนระหว่าง 60-79 เปอร์เซ็นต์ จำนวน 15 เว็บไซต์ คิดเป็น 42.86 เปอร์เซ็นต์ เว็บไซต์ที่มีคะแนนระหว่าง 50-59 เปอร์เซ็นต์ จำนวน 9 เว็บไซต์ คิดเป็น 25.71 เปอร์เซ็นต์ และเว็บไซต์ที่มีคะแนนน้อยกว่า 50 เปอร์เซ็นต์ จำนวน 6 เว็บไซต์ คิดเป็น 17.14 เปอร์เซ็นต์

ทั้งนี้เว็บที่ได้คะแนนการประเมินสูงสุด ได้แก่ เว็บไซต์ www.bccancer.bc.ca และ www.mayoclinic.com โดยมีคะแนนการประเมินอยู่ที่ 86 เปอร์เซ็นต์ ส่วนเว็บไซต์ที่มีคะแนนการประเมินต่ำที่สุดได้แก่เว็บไซต์ www.drug3k.com ซึ่งมีคะแนนการประเมินอยู่ที่ 41เปอร์เซ็นต์ เมื่อพิจารณาจากข้อมูลที่ปรากฏอยู่ในเว็บไซต์เหล่านั้นพบว่า เว็บที่มีคะแนนการประเมินสูง จะเป็นเว็บไซต์ที่มีเนื้อหาครอบคลุมทั้งส่วนของโรค แนวทางในการรักษา และยาที่ใช้ในการรักษา รวมไปถึงมีส่วนการอ้างอิงข้อมูลจากแหล่งข้อมูลที่มีความน่าเชื่อถือ และมีการระบุวันเดือนปีที่ให้ข้อมูลได้ชัดเจน มีการกำหนดเป้าหมายและวัตถุประสงค์ของเว็บไซต์ไว้อย่างชัดเจน รวมทั้งระบุกลุ่มเป้าหมายของเว็บไซต์ด้วยว่าเป็นเนื้อหาที่เหมาะกับบุคคลกลุ่มใด ซึ่งจะแตกต่างจากเว็บไซต์ที่ได้คะแนนประเมินต่ำส่วนใหญ่จะให้ข้อมูลได้ไม่ครบถ้วน โดยอาจจะขาดส่วนใดส่วนหนึ่ง หรือมากกว่าหนึ่งส่วนไป เช่นไม่มีการระบุแหล่งที่มาของข้อมูล หรือไม่มีการอ้างอิงแหล่งข้อมูลที่มีความน่าเชื่อถือ จากการวิจัยยังพบว่าเว็บอันดับท้ายๆ สามารถพบเว็บไซต์ที่มีคุณภาพของข้อมูลที่มากและน้อยได้เช่นเดียวกันกับเว็บไซต์อันดับต้น นอกจากนี้พบว่าเว็บไซต์ที่ให้ข้อมูลด้านสุขภาพส่วนมากจะไม่มีกระดานสำหรับให้ผู้เข้าชมได้แลกเปลี่ยนความคิดเห็นจะมีเพียงแต่ส่วนที่ใช้สำหรับกรอกข้อสักถามแล้วรอทางเว็บไซต์ตอบ หรือให้ข้อมูลกลับเท่านั้น

4.3 การจัดกลุ่มเว็บไซต์

เมื่อทำการประเมินคุณภาพของเว็บไซต์ที่ให้ข้อมูลด้านยาที่เป็นภาษาอังกฤษแล้ว ทำการจัดกลุ่มของเว็บไซต์ออกตามผลการประเมินคุณภาพจะสามารถแบ่งเว็บไซต์ได้เป็น 2 ระดับเพื่อให้ง่ายต่อการวิเคราะห์ผลดังนี้คือ เว็บไซต์ที่มีคะแนนการประเมินคุณภาพตั้งแต่ 60 เปอร์เซ็นต์ขึ้นไป (แสดงในตารางที่ 3) และเว็บไซต์ที่มีคะแนนการประเมินคุณภาพน้อย 60 เปอร์เซ็นต์ (แสดงในตารางที่ 4)

ตารางที่ 3 แสดงเว็บไซต์ที่มีคะแนนการประเมินคุณภาพตั้งแต่ 60 เปอร์เซ็นต์ขึ้นไป

รายชื่อเว็บไซต์	คะแนนเฉลี่ย (เปอร์เซ็นต์)
www.bccancer.bc.ca	86
www.mayoclinic.com	86
www.rxlist.com	85
www.drugs.com	82
www.healthsquare.com	81
www.healthline.com	77
www.nhsdirect.nhs.uk	76
www.medicinenet.com	75
www.empr.com	74
www.nlm.nih.gov	71
www.webmd.com	69
www.medinfo.co.uk	66
www.drugdigest.org	64
www.health.yahoo.com	64
www.familydoctor.org	61
drugs.medsort.com	74
www.allina.com	67
www.ralphs.com	65
www.everyonehealthy.com	63
www.aol.com	60

จากตารางที่ 3 เมื่อทำการจัดกลุ่มตามการประเมินคุณภาพเว็บไซต์ พบว่ามีเว็บไซต์ที่มีคะแนนการประเมินคุณภาพตั้งแต่ 60 เปอร์เซ็นต์ขึ้นไปจำนวน 20 เว็บไซต์จากเว็บไซต์ที่ทำการศึกษาทั้งหมด 35 เว็บไซต์ คิดเป็น 57.14 เปอร์เซ็นต์ โดยแบ่งเป็นเว็บไซต์ที่มาจากเว็บอันดับต้นๆ จำนวน 15 เว็บไซต์คิดเป็น 42.86 เปอร์เซ็นต์ของเว็บไซต์ทั้งหมด และเว็บไซต์ที่มาจากเว็บอันดับท้ายๆ จำนวน 5 เว็บไซต์คิดเป็น 14.29 เปอร์เซ็นต์ของเว็บไซต์ทั้งหมด ซึ่งผลดังกล่าวแสดงให้เห็นว่าเว็บไซต์ที่อยู่ในอันดับท้ายของหน้าแสดงผล สามารถเป็นเว็บไซต์ที่มีคะแนนการประเมินคุณภาพมากกว่า 60 เปอร์เซ็นต์ได้เช่นเดียวกัน

ตารางที่ 4 แสดงเว็บไซต์ที่มีคะแนนการประเมินคุณภาพน้อยกว่า 60 เปอร์เซ็นต์

รายชื่อเว็บไซต์	คะแนนเฉลี่ย (เปอร์เซ็นต์)
www.cancer.gov	59
www.medical-clinic.org	59
www.myoptumhealth.com	59
www.muschealth.com	57
www.riteaid.com	56
www.pdrhealth.com	54
www.patient.co.uk	53
www.parkinsons.org.uk	48
www.walgreens.com	43
www.druginformationonline.org	31
www.canada.com	51
www.grubbscare.com	50
www.medindia.net	44
www.mdonweb.org	43
www.drug3k.com	41

จากตารางที่ 4 พบว่ามีเว็บไซต์ที่มีคะแนนการประเมินคุณภาพน้อยกว่า 60 เปอร์เซ็นต์ จำนวน 15 เว็บไซต์ คิดเป็น 42.86 เปอร์เซ็นต์ของเว็บไซต์ทั้งหมด โดยแบ่งเป็นเว็บที่มาจากเว็บไซต์อันดับต้น จำนวน 10 เว็บไซต์คิดเป็น 28.57 เปอร์เซ็นต์ของเว็บไซต์ทั้งหมด และเว็บไซต์ที่มาจากเว็บอันดับท้ายจำนวน 5 เว็บไซต์คิดเป็น 14.29 เปอร์เซ็นต์ของเว็บไซต์ทั้งหมด ซึ่งพบว่าเว็บไซต์ที่อยู่ในอันดับต้น เมื่อทำการค้นหาผ่านโปรแกรมค้นหาจะมีอันตราส่วนของเว็บไซต์ที่มีคะแนนการประเมินคุณภาพน้อยกว่า 60 เปอร์เซ็นต์ อยู่ถึง 29 เปอร์เซ็นต์ อาจเป็นผลให้ผู้ค้นหาได้ข้อมูลที่ไม่ครบถ้วนสมบูรณ์จากเว็บไซต์ดังกล่าว จึงสามารถสรุปได้ว่าคุณภาพของเนื้อหาภายในเว็บไซต์เป็นองค์ประกอบหนึ่งที่จะทำให้เว็บไซต์อยู่ในอันดับต้น หรืออันดับท้ายเมื่อทำการค้นหาผ่านโปรแกรมค้นหา แต่อย่างไรก็ดีไม่จำเป็นว่าเว็บไซต์ที่มีคุณภาพของเนื้อหาดีจะได้อยู่ในหน้าแรกๆ ของหน้าแสดงผลการค้นหาเสมอไปเนื่องจากยังมีองค์ประกอบอื่นๆ อีกที่ส่งผลต่อการจัดอันดับของเว็บไซต์ของโปรแกรมค้นหาเช่นกัน

4.4 การเก็บรวบรวมข้อมูล

ในการสุ่มเก็บรวบรวมข้อมูลจากเอกสาร HTML ที่ให้ข้อมูลด้านยาที่เป็นภาษาอังกฤษจากเว็บไซต์ทั้งหมด 35 เว็บไซต์ตั้งแต่เดือนพฤษภาคม - ตุลาคม 2553 โดยสุ่มเอกสาร HTML จากแต่ละเว็บไซต์จำนวนเว็บไซต์ละ 20 หน้าซึ่งกำหนดให้มีหน้าโฮมเพจของแต่ละเว็บไซต์ 1 หน้าและหน้าอื่นๆ อีก 19 หน้ารวมได้ทั้งสิ้นจำนวน 700 หน้าและทำการคัดเลือกตัวแปรที่ใช้ในการศึกษา โดยในการศึกษานี้ได้ตัวแปรแบ่งเป็น 2 ส่วนดังนี้คือ

4.4.1 Tag ที่ปรากฏบนเอกสาร HTML ทั้งหมด 12 Tag คือ a, h1, h2, h3, h4, h5, h6, img, link, meta, p และ title

4.4.2 คุณลักษณะต่างๆ ได้แก่ ขนาดของหน้าเอกสาร HTML (Total Page Size), เวลาที่ใช้ในการดาวน์โหลด (Download time), ประเภทของเอกสาร HTML (Doctype), การเข้ารหัสภาษา (Encoding), จุดเชื่อมโยงภายใน-ภายนอก (Internal-External link), ข้อผิดพลาดต่างๆ ทั้งในส่วนของ HTML และ CSS

4.5 การศึกษาด้วยสถิติเชิงพรรณนา

4.5.1 การศึกษาโครงสร้างของเอกสาร HTML และคุณลักษณะต่างๆ จากเว็บไซต์ที่อยู่อันดับต้น ดังนี้

1. การศึกษาปริมาณ Tag HTML เฉลี่ยที่ปรากฏบนเอกสาร HTML เมื่อทำการศึกษาในส่วนของปริมาณ Tag ที่ปรากฏบนเอกสาร HTML ทั้งหมด โดยใช้วิธีการคัดแยกด้วยมือทีละหน้าเว็บเพจ จากเว็บเพจที่ทำศึกษาทั้งหมด จากนั้นใช้โปรแกรม Microsoft Excel ในการจำแนกกลุ่มรวบรวม, นับจำนวน ของ Tag ต่างๆ และทำการหาค่าเฉลี่ย รวมทั้งหาจำนวนเว็บเพจที่พบ Tag HTML เทียบกับจำนวนเว็บเพจทั้งหมดที่ศึกษา

2. การศึกษาปริมาณจุดเชื่อมโยงภายนอกและจุดเชื่อมโยงภายในโดยใช้โปรแกรม SeoQuake SEO extension version 2.5.9 ซึ่งเป็นโปรแกรมเสริมที่ติดตั้งบนเว็บเบราเซอร์ Firefox เป็นตัวเก็บรวบรวมจากนั้นนำค่าที่ได้มาทำการหาค่าเฉลี่ยต่อหน้าเว็บเพจ

3. การศึกษาขนาดของหน้าเว็บเพจเฉลี่ยในแต่ละเว็บไซต์, เวลาที่ใช้ในการดาวน์โหลด (Connection Rate 1.44 Mbps.), ขนาดภาพที่พบภายในเว็บเพจ, ภาพที่ไม่สามารถเรียกดูได้ใน

เว็บไซต์ (Dead Images) และขนาดของ CSS ที่พบภายในเว็บเพจโดยใช้ Web Page Analyzer-0.98 จากเว็บไซต์ ww.websiteoptimization.com ซึ่งเป็นเครื่องมือที่ช่วยในการจำแนกขนาดและความเร็วของเว็บไซต์

4. การศึกษาประเภทของเอกสาร HTML และการเข้ารหัสภาษาโดยการใช้ Markup Validation Service ของ W3C เป็นเครื่องมือที่ในการจำแนกประเภท

5. การศึกษาปริมาณของค่าความผิดพลาดเฉลี่ยของเว็บไซต์ที่ให้ข้อมูลยาที่เป็นภาษาอังกฤษ โดยศึกษาทั้งในส่วนของ HTML error, HTML warning และ CSS error โดยใช้ Markup Validation Service เป็นเครื่องมือในการศึกษาส่วนของ HTML error และ HTML warning และใช้ CSS Validation Service เป็นเครื่องมือในการจำแนกข้อผิดพลาดต่างๆ ในส่วนของ CSS error รวมถึงศึกษาจำนวนการเชื่อมโยงทั้งหมดที่พบ, การเชื่อมโยงที่ไม่สามารถเข้าถึงได้ในเว็บไซต์ (Dead Links) โดยใช้ W3C Link Checker เป็นเครื่องมือในการเก็บข้อมูล

6. การศึกษา Google Page Rank ซึ่งเป็นค่าลำดับคะแนนที่ Google ประเมินให้กับหน้าเว็บเพจแต่ละหน้า โดยในการศึกษานี้ได้ใช้ Google toolbar ซึ่งติดตั้งบนเว็บเบราเซอร์ในการหาค่าดังกล่าว จากนั้นนำค่าที่ได้ของแต่ละหน้าที่ทำการศึกษามาหาค่าเฉลี่ยของเว็บไซต์

ตารางที่ 5 แสดงปริมาณ Tag เฉลี่ยต่อหน้าเว็บเพจที่ปรากฏบนเอกสาร HTML ของเว็บไซต์อันดับต้น

เว็บไซต์อันดับต้น	Tag HTML ที่ศึกษา											
	a	h1	h2	h3	h4	h5	h6	img	link	meta	p	title
www.bccancer.bc.ca	154.9	0.1	2.65	12.55	1.75	0.05	0.2	367.1	2	16.15	279.05	1
www.cancer.gov	72.75	0	0.05	0	1	0	0	84.6	5	5	14.25	1
www.drugdigest.org	66.25	1.4	0	0	0	0	0	70.7	9	0.1	5	2.05
www.druginformationonline.org	53.65	0.05	6.2	0.25	0.2	0	0	9	3.05	3	9.75	1
www.drugs.com	208.45	1	13.15	0	0	0	0	15.65	6.15	6	80.3	1
www.empr.com	227	0.95	14.95	28.15	41.8	0	0	20.35	5	5	47.6	1
www.familydoctor.org	56.8	2	0	8.4	0.05	0	0	24.35	6	5.95	29.1	1
www.health.yahoo.com	174.8	2.15	0	8.15	0.95	0	0	16.6	1.95	1.05	7.7	1
www.healthline.com	241.5	1.9	0.25	2.85	0	0	0	55.75	7.85	2.1	9.85	1
www.healthsquare.com	81.25	2.4	6.6	1.6	1.4	0	0	10.35	3	3	10.1	1
www.mayoclinic.com	80.7	4.2	5.4	7.65	3.15	0	0	12.05	4	8	24.5	1
www.medical-clinic.org	69.45	0.95	1.9	0.95	0.95	0.95	0	56.8	1	3.05	10.1	1
www.medicinenet.com	193.7	1.2	1	12.6	2.45	3.1	0	27.2	12.2	7	41.05	0
www.medinfo.co.uk	30.45	1	1	6.25	0.75	0	0	2.1	1	1.6	17.9	1

ตารางที่ 5 แสดงปริมาณ Tag เฉลี่ยต่อหน้าเว็บเพจที่ปรากฏบนเอกสาร HTML ของเว็บไซต์อันดับต้น (ต่อ)

เว็บไซต์อันดับต้น	Tag HTML ที่ศึกษา											
	a	h1	h2	h3	h4	h5	h6	img	link	meta	p	title
www.muschealth.com	54.7	0	0	0	0	0	0	12.05	4.8	4.3	12.6	1
www.myoptumhealth.com	185.15	2.1	9.2	1.25	0	0	0	12.25	4	1	7.3	0
www.nhsdirect.nhs.uk	130.2	2	3.95	1.05	1.3	0	0	4.4	12.9	14.4	17.35	1
www.nlm.nih.gov	30.45	1	11.1	3.5	0	0	0	4.9	7.9	4.9	22.4	1
www.parkinsons.org.uk	60.75	1	6.1	0.3	0	0	0	6.45	7	4	21.45	1
www.patient.co.uk	95.45	0.85	7.15	0.05	0	0.05	0	24	6.9	4.05	13.55	1
www.pdrhealth.com	32.3	2	1.2	1.1	0	0	0	10.85	4.05	8.9	12.95	1.6
www.riteaid.com	5.4	0	0	1.5	0	0	0	3.15	3.05	3.35	2.05	1
www.rxlist.com	296.75	1.3	4	25.55	22.35	22.3	0	31.45	9.2	5	225.45	1
www.walgreens.com	180.2	0.95	6.85	0.1	0	0	0	14.25	10.95	5	19.95	1
www.webmd.com	322	1.9	1.1	7.15	4.05	1	7.65	32.9	13.8	4.95	51.3	1
ค่าเฉลี่ยของแต่ละ Tag HTML	124.20	1.30	4.15	5.24	3.29	1.10	0.31	37.17	6.07	5.07	39.70	0.99
จำนวนเว็บเพจที่พบ Tag HTML เทียบกับจำนวนเว็บเพจทั้งหมดที่ศึกษา (เปอร์เซ็นต์)	96.20	78.60	70.40	69.00	44.80	16.20	4.20	100.00	100.00	96.20	99.80	92.00

จากตารางที่ 5 การศึกษาปริมาณ Tag ต่างๆ ที่พบในเว็บไซต์ที่ให้ข้อมูลด้านยา พบว่าปริมาณTag HTML เฉลี่ยต่อหน้าเว็บเพจในเว็บไซต์อันดับต้น มีปริมาณเฉลี่ยของ Tag a มากที่สุด รองมาได้แก่ Tag p และ Tag img ตามลำดับ ซึ่งพบว่าเว็บไซต์ที่มีปริมาณ Tag a มากที่สุด ได้แก่ www.webmd.com โดยค่าเฉลี่ยของ Tag a อยู่ที่ 124.20 ในส่วนของ Tag h1 พบว่าเว็บไซต์ ที่พบมากที่สุดได้แก่ www.mayoclinic.com ซึ่งค่าเฉลี่ยของ Tag h1 เท่ากับ 1.30 Tag h2, h3 และ h4 พบมากในเว็บไซต์ www.empr.com ซึ่งค่าเฉลี่ยของ Tag ในเว็บไซต์อันดับต้นคือ 4.15 5.24 และ 3.29 Tag h5 พบปริมาณ Tag น้อยมากโดยพบมากในเว็บ www.rxlist.com และ Tag h6 พบว่ามีปริมาณ Tag น้อยที่สุดเมื่อเทียบกับ Tag อื่นๆ โดยพบเพียง 2 เว็บไซต์ ได้แก่ www.webmd.com และ www.bccancer.bc.ca ในส่วนของ Tag img ซึ่งพบปริมาณของ Tag มากเป็นอันดับต้นๆ พบว่ามีมากในเว็บไซต์ www.bccancer.bc.ca โดยค่าเฉลี่ยของ Tag ในเว็บอันดับต้นเท่ากับ 37.17 Tag link พบมากในwww.webmd.com Tag meta พบมากในเว็บไซต์ www.bccancer.bc.ca ในส่วนของ Tag p ซึ่งเป็นอีกตัวแปรที่มีปริมาณ Tag เฉลี่ยมาก พบว่าเว็บไซต์ที่มีปริมาณ Tag p มาก ได้แก่ www.bccancer.bc.ca และ www.rxlist.com ตามลำดับ โดยค่าเฉลี่ยปริมาณ Tag ในเว็บอันดับต้นเท่ากับ 39.70 ในส่วนของ Tag title พบว่าส่วนใหญ่ จะมีปริมาณ Tag เฉลี่ยอยู่ที่ 1 มีเพียงบางเว็บไซต์ ที่มีมากกว่า และพบว่าไม่พบ Tag ดังกล่าวในบางเว็บไซต์ จากการศึกษาอาจกล่าวได้ว่าเว็บไซต์อันดับต้นส่วนใหญ่จะประกอบไปด้วย Tag HTML ที่ใช้มากอยู่ด้วยกัน 3 Tag คือ Tag a ซึ่งเป็น Tag ที่ใช้ในการสร้างจุดเชื่อมโยงไปยังเอกสารอื่นๆ โดยการใช้ แอททริบิว href หรือใช้ในการสร้างจุดเชื่อมโยงภายในเอกสารนั้นๆ เอง เรียกได้ว่าเป็นการเชื่อมโยงหลายมิติ, Tag img ซึ่งเป็น Tag ที่ใช้ในการกำหนดภาพภายในเอกสาร HTML เมื่อศึกษาเปอร์เซ็นต์ของเว็บเพจที่พบแท็กนั้นๆ เทียบกับจำนวนเว็บเพจทั้งหมดที่ศึกษา พบว่าเว็บไซต์อันดับต้น จะพบ Tag img และ Tag link ในทุกหน้าเว็บเพจที่ทำการศึกษา ส่วน Tag ที่พบน้อยที่สุดคือ Tag h6 พบการใช้เพียง 4.20 เปอร์เซ็นต์ จากเว็บเพจของเว็บไซต์อันดับต้นที่ทำการศึกษา จากการศึกษาก่อนหน้าพบว่าการกำหนด Tag h1 และ meta tag ดีจะส่งผลให้สามารถค้นเจอได้ง่าย

ตารางที่ 6 แสดงคุณลักษณะของเอกสาร HTML ในส่วนของจุดเชื่อมโยงภายนอกและจุดเชื่อมโยงภายในของเว็บไซต์อันดับต้น

เว็บไซต์อันดับต้น	คุณลักษณะ	
	จุดเชื่อมโยงภายนอก	จุดเชื่อมโยงภายใน
www.bccancer.bc.ca	15.05	138.85
www.cancer.gov	15.35	57.4
www.drugdigest.org	7	49.65
www.druginformationonline.org	52.35	0
www.drugs.com	41.15	155.9
www.empr.com	192.05	33
www.familydoctor.org	16	38.8
www.health.yahoo.com	56.1	116.75
www.healthline.com	3.1	231.5
www.healthsquare.com	58.3	20.1
www.mayoclinic.com	24	50.45
www.medical-clinic.org	5.05	0
www.medicinenet.com	130.35	57.95
www.medinfo.co.uk	10.3	17.7
www.muschealth.com	49	0
www.myoptumhealth.com	154.25	0
www.nhsdirect.nhs.uk	117.8	0
www.nlm.nih.gov	13.35	17.1
www.parkinsons.org.uk	4.45	54.55
www.patient.co.uk	24.2	67.75
www.pdrhealth.com	2.15	14.6
www.riteaid.com	2.4	2.9
www.rxlist.com	55.7	196.2

ตารางที่ 6 แสดงคุณลักษณะของเอกสาร HTML ในส่วนของจุดเชื่อมโยงภายนอกและจุดเชื่อมโยงภายในของเว็บไซต์อันดับต้น (ต่อ)

เว็บไซต์อันดับต้น	คุณลักษณะ	
	จุดเชื่อมโยงภายนอก	จุดเชื่อมโยงภายใน
www.walgreens.com	20.1	157.55
www.webmd.com	283.7	14.15
ค่าเฉลี่ยของจำนวนการเชื่อมโยงแต่ละเว็บเพจ	54.13	59.71
จำนวนเว็บเพจที่พบการเชื่อมโยง เทียบกับจำนวนเว็บเพจทั้งหมดที่ศึกษา (เปอร์เซ็นต์)	96.20	76.00

จากตารางที่ 6 จุดเชื่อมโยงภายนอกและจุดเชื่อมโยงภายในนั้นพบว่ามีบางเว็บไซต์ที่ไม่มีการเชื่อมโยงภายในเว็บไซต์ซึ่งทำให้การเข้าถึงข้อมูลต่างๆ ภายในเว็บไซต์ทำได้ยาก ส่วนการเชื่อมโยงไปยังเว็บไซต์ภายนอกนั้นพบว่าหากเชื่อมโยงไปยังเว็บไซต์ที่ให้ข้อมูลที่มีความน่าเชื่อถือมาก และมีเนื้อหาที่เกี่ยวข้องสัมพันธ์กันจะส่งผลให้เว็บไซต์ได้รับความนิยมจากโปรแกรมค้นหามากยิ่งขึ้น

ตารางที่ 7 แสดงขนาดของหน้าเว็บเพจเฉลี่ยในแต่ละเว็บไซต์อันดับต้น, เวลาที่ใช้ในการดาวน์โหลด (Connection Rate 1.44 Mbps.), ขนาดภาพที่พบภายในเว็บเพจเฉลี่ย, ขนาดของ CSS ที่พบภายในเว็บเพจ, ภาพที่ไม่สามารถเรียกดูได้ในเว็บไซต์อันดับต้น (Dead Images) และจำนวนภาพรวมภายในเว็บไซต์อันดับต้น

เว็บไซต์อันดับต้น	ขนาดของหน้าเอกสารHTML (กิโลไบต์)	เวลาที่ใช้ในการดาวน์โหลด (วินาที)	ขนาดของภาพ (กิโลไบต์)	ขนาดของ CSS (กิโลไบต์)	รูปภาพที่ไม่สามารถเข้าถึงได้	จำนวนรูปภาพรวมเฉลี่ยต่อหน้าเว็บเพจ
www.bccancer.bc.ca	35.77	2.83	18.20	0	0	126.05
www.cancer.gov	161.66	3.35	19.38	13.90	0	19.1
www.drugdigest.org	290.09	7.61	37.15	2.12	1	17.75
www.druginformationonline.org	29.15	2.98	0.76	0.18	0	0.15
www.drugs.com	117.56	3.3	47.87	2.70	0	14.1
www.empr.com	365.58	4.63	87.77	44.28	0	17.45
www.familydoctor.org	126.08	2.24	20.95	24.41	0	10.55
www.health.yahoo.com	214.23	1.28	83.92	32.82	0	7.9
www.healthline.com	190.10	0.89	91.53	12.88	0	27.3
www.healthsquare.com	76.53	2.15	4.87	8.49	0	2.7
www.mayoclinic.com	211.09	3.11	49.80	61.86	0	4.8
www.medical-clinic.org	50.14	2.6	15.77	0	0	2.6
www.medicinenet.com	241.64	1.75	88.49	56.45	0	15.8
www.medinfo.co.uk	14.80	2.3	4.80	0.02	0	1.95

ตารางที่ 7 แสดงขนาดของหน้าเว็บเพจเฉลี่ยในแต่ละเว็บไซต์อันดับต้น, เวลาที่ใช้ในการดาวน์โหลด (Connection Rate 1.44 Mbps.), ขนาดภาพที่พบภายในเว็บเพจเฉลี่ย, ขนาดของ CSS ที่พบภายในเว็บเพจ, ภาพที่ไม่สามารถเรียกดูได้ในเว็บไซต์อันดับต้น (Dead Images) และจำนวนภาพรวมภายในเว็บไซต์อันดับต้น (ต่อ)

เว็บไซต์อันดับต้น	ขนาดของหน้าเอกสารHTML (กิโลไบต์)	เวลาที่ใช้ในการดาวน์โหลด (วินาที)	ขนาดของภาพ (กิโลไบต์)	ขนาดของ CSS (กิโลไบต์)	รูปภาพที่ไม่สามารถเข้าถึงได้	จำนวนรูปภาพรวมเฉลี่ยต่อหน้าเว็บเพจ
www.muschealth.com	161.51	2.42	40.59	2.18	0	7.25
www.myoptumhealth.com	102.18	3.15	30.52	0	0	5.05
www.nhsdirect.nhs.uk	708.16	11.31	82.34	165.18	0	4
www.nlm.nih.gov	108.53	1.5	6.99	38.43	0	1.95
www.parkinsons.org.uk	299.45	2.8	53.88	6.35	0	5.15
www.patient.co.uk	42.32	3.81	21.46	6.56	0.05	24.85
www.pdrhealth.com	388.05	3.1	29.06	20.24	0	7.95
www.riteaid.com	23.98	2.02	6.48	1.70	0	2.3
www.rxlist.com	285.00	3.55	138.71	44.99	0	27
www.walgreens.com	98.11	2.7	52.55	18.27	0.05	9.25
www.webmd.com	176.35	2.16	59.09	1.86	0	23.25
ค่าเฉลี่ยต่อหน้าเว็บเพจ	180.72	3.18	43.72	22.64	0.04	15.45
รวมทั้งหมด	4518.05	79.54	1092.94	565.88	1.1	386.2

จากตารางที่ 7 การศึกษาขนาดของหน้าเว็บเพจเฉลี่ยในแต่ละเว็บไซต์และเวลาที่ใช้ในการดาวน์โหลด โดยใช้ Web Page Analyzer-0.98 จากเว็บไซต์ ww.websiteoptimization.com ซึ่งเป็นเครื่องมือที่ช่วยในการจำแนกขนาดและความเร็วของเว็บไซต์พบว่าเว็บไซต์ www.nhsdirect.nhs.uk มีขนาดหน้าเว็บเพจเฉลี่ยขนาดใหญ่ที่สุดเมื่อเทียบจากเว็บไซต์ที่ทำการศึกษาทั้งหมดโดยมีขนาดหน้าเว็บเพจเฉลี่ยอยู่ที่ 708.16 กิโลไบต์และเว็บไซต์ www.medinfo.co.uk เป็นเว็บไซต์ที่มีขนาดหน้าเว็บเพจเฉลี่ยน้อยที่สุด โดยมีขนาดหน้าเว็บเพจเฉลี่ยเพียง 14.80 กิโลไบต์เมื่อทำการศึกษาในเรื่องของเวลาที่ใช้ในการดาวน์โหลดหน้าเว็บเพจพบว่าเว็บไซต์ที่ใช้เวลาในการดาวน์โหลดหน้าเว็บเพจมากที่สุดได้แก่เว็บไซต์ www.nhsdirect.nhs.uk โดยใช้เวลาในการเข้าถึง 11.31 วินาที และเว็บไซต์ที่ใช้เวลาในการเข้าถึงน้อยที่สุดได้แก่ www.health.yahoo.com โดยใช้เวลา 1.28 วินาที จากการศึกษานี้อาจสรุปได้ว่าเว็บไซต์ที่มีขนาดของหน้าเว็บเพจขนาดใหญ่ จะใช้ระยะเวลาในการเข้าถึงหน้าเว็บเพจ นานกว่าเว็บเพจที่มีหน้าเว็บเพจขนาดเล็ก แต่ทั้งนี้ขึ้นอยู่กับปัจจัยอื่นๆ ที่เป็นองค์ประกอบด้วย เนื่องจากผลการสำรวจพบว่ามีบางเว็บไซต์ ที่ให้ผลไม่เป็นไปตามลักษณะข้างต้น จากงานวิจัยอื่นๆ พบว่าผู้เข้าชมเว็บไซต์จะยกเลิกการเข้าชมหน้าเว็บไซต์นั้นๆ หากใช้เวลาในการดาว์นโหลดหน้าเว็บเพจนานกว่า 15 วินาที นอกจากนี้เว็บไซต์อันดับต้นจะมีขนาดของหน้าเว็บเพจเฉลี่ยที่เล็กและมีค่าเวลาในการดาวน์โหลดหน้าเว็บของเว็บต้นเฉลี่ยอยู่ที่ 3.18 วินาทีจึงอาจกล่าวได้ว่าระยะเวลาที่ใช้ในการดาว์นโหลดหน้าเว็บเพจและขนาดของเว็บเพจส่งผลต่ออันดับการค้นหาเมื่อทำการค้นหาผ่านโปรแกรมค้นหา โดยพบว่าเว็บไซต์ที่ใช้เวลาในการดาว์นโหลดหน้าเว็บเพจน้อย จะมีอันดับที่ดีกว่าเว็บไซต์ที่สามารถเข้าถึงได้ยากกว่า หรือใช้เวลาในการดาว์นโหลดนานกว่า

ในส่วนของขนาดภาพที่พบภายในเว็บเพจรวมเฉลี่ย พบว่าเว็บไซต์ www.rxlist.com มีขนาดภาพรวมมากที่สุด คือ 138.71 กิโลไบต์ส่วนเว็บไซต์ www.druginformationonline.org เป็นเว็บไซต์ที่มีขนาดของขนาดภาพรวมเฉลี่ยน้อยที่สุด โดยมีขนาดเท่ากับ 0.76 กิโลไบต์ซึ่งส่วนของขนาดภาพนั้นส่งผลต่อ Data transfer rate โดยหากไฟล์รูปภาพที่ใช้มีขนาดเล็ก ก็จะช่วยให้พื้นที่ในการใช้งานทำให้สามารถรองรับการทำงานอื่นๆ ได้เพิ่มขึ้น โดยขนาดของภาพที่พบภายในเว็บเพจรวมเฉลี่ยของเว็บไซต์อันดับต้นมีค่าเท่ากับ 43.72 กิโลไบต์

การศึกษาขนาดของ CSS ที่พบภายในเว็บเพจเฉลี่ยโดยรวม พบว่ามีเว็บไซต์ที่ไม่มีการใช้ CSS ในเว็บไซต์อันดับต้นบางเว็บไซต์ได้แก่ www.bccancer.bc.ca, www.medical-clinic.org, www.myoptumhealth.com นอกจากนี้ยังพบว่าเว็บไซต์ www.nhsdirect.nhs.uk เป็นเว็บที่มีขนาด CSS ที่พบภายในเว็บเพจใหญ่ที่สุดคือ 165.18 กิโลไบต์โดยในส่วนของ CSS นั้นจะช่วยทำ

หน้าที่ในการแสดงผลข้อมูล (การตกแต่ง) ทำให้เนื้อหาและข้อมูลถูกแยกออกจากกัน ซึ่งการใช้ CSS จะทำให้โปรแกรมค้นหาสามารถหาข้อมูลได้แม่นยำขึ้น ไม่ต้องเสียเวลาในการคัดแยกโค้ดที่ไม่เกี่ยวข้องออกก่อน โดยค่าเฉลี่ยของขนาด CSS ที่พบภายในเว็บเพจเท่ากับ 22.64 กิโลไบต์

จากการศึกษาภาพที่ไม่สามารถเรียกดูได้ภายในเว็บไซต์ พบว่ามีทั้งหมด 3 เว็บไซต์ที่พบภาพที่ไม่สามารถเรียกดูได้ในเว็บไซต์ คือ www.drugdigest.org, www.patient.co.uk และ www.walgreens.com คิดเป็น 20 เปอร์เซ็นต์จากเว็บไซต์ที่ทำการศึกษาทั้งหมด นอกจากนี้พบว่าแต่ละเว็บไซต์จะมีจำนวนรูปภาพต่อหน้าเว็บเพจเฉลี่ยที่แตกต่างกันออกไป จากการศึกษานั้นพบว่าการที่มีรูปภาพบนเว็บไซต์จะช่วยให้เว็บไซต์ดูน่าสนใจมากยิ่งขึ้นแต่หากมีจำนวนรูปที่มากเกินไปจะส่งผลต่ออันดับของการค้นหาผ่านการค้นหา เนื่องจากบางครั้งผู้สร้างเว็บไซต์ไม่ได้กำหนดค่าให้กับรูปภาพนั้นๆ ทำให้ Robot ที่เข้ามาเก็บข้อมูลบนเว็บไซต์ ไม่สามารถเข้าใจได้ว่ารูปนั้นคือรูปอะไร มีความสำคัญอย่างไรกับเว็บไซต์นั้นๆ ทำให้คะแนนของเว็บไซต์นั้นลดลง

ตารางที่ 8 แสดงประเภทของเอกสาร HTML และการเข้ารหัสภาษา ของเว็บไซต์อันดับต้น

เว็บไซต์อันดับต้น	Encoding			ประเภทของเอกสาร HTML	
	utf-8	iso-8859-1	windows-1252	html	xhtml
www.bccancer.bc.ca	20	0	0	20	0
www.cancer.gov	0	20	0	20	0
www.drugdigest.org	20	0	0	20	0
www.druginformationonline.org	0	20	0	0	20
www.drugs.com	20	0	0	0	20
www.empr.com	20	0	0	0	20
www.familydoctor.org	20	0	0	0	20
www.health.yahoo.com	20	0	0	20	0
www.healthline.com	20	0	0	20	0
www.healthsquare.com	20	0	0	0	20
www.mayoclinic.com	20	0	0	0	20
www.medical-clinic.org	0	20	0	0	20
www.medicinenet.com	20	0	0	0	20
www.medinfo.co.uk	0	20	0	20	0
www.muschealth.com	20	0	0	1	19
www.myoptumhealth.com	20	0	0	20	0

ตารางที่ 8 แสดงประเภทของเอกสาร HTML และการเข้ารหัสภาษา ของเว็บไซต์อันดับต้น (ต่อ)

เว็บไซต์อันดับต้น	Encoding			ประเภทของเอกสาร HTML	
	utf-8	iso-8859-1	windows-1252	html	xhtml
www.nhsdirect.nhs.uk	20	0	0	0	20
www.nlm.nih.gov	20	0	0	0	20
www.parkinsons.org.uk	20	0	0	0	20
www.patient.co.uk	0	20	0	0	20
www.pdrhealth.com	0	20	0	0	20
www.riteaid.com	20	0	0	0	20
www.rxlist.com	0	20	0	0	20
www.walgreens.com	20	0	0	0	20
www.webmd.com	20	0	0	0	20
รวมทั้งหมด	360	140	0	141	359
ค่าเฉลี่ยต่อหน้าเว็บเพจ	14.40	5.60	0.00	5.64	14.36

จากการศึกษาประเภทของเอกสาร HTML และการเข้ารหัสภาษา โดยการใช้ Markup Validation Service ของ W3C เป็นเครื่องมือที่ในการจำแนกประเภท พบว่าในเว็บไซต์อันดับต้นมีรูปแบบของการเข้ารหัสภาษาด้วยกัน 2 รูปแบบหลักๆ คือ utf-8 และ iso-8859-1 โดยรูปแบบที่นิยมใช้มากที่สุดได้แก่ utf-8 ซึ่งเป็นรูปแบบที่เป็นมาตรฐานสามารถรองรับการใช้งานภาษาได้หลายภาษา รองลงมาได้แก่รูปแบบ iso-8859-1 ซึ่งเป็นชุดอักขระมาตรฐานของภาษายุโรปทางตะวันตก ส่วนรูปแบบ windows-1252 ไม่พบในเว็บไซต์อันดับต้นเลย ในส่วนของประเภทของเอกสาร HTML ที่ใช้ในการศึกษานี้พบรูปแบบการใช้อยู่ด้วยกัน 2 รูปแบบหลักๆ คือ html และ xhtml โดยแต่ละรูปแบบสามารถลงย่อยได้อีกแต่ทั้งนี้การศึกษานี้ขอแยกออกเป็น 2 กลุ่มหลักๆ เพื่อง่ายต่อการเข้าใจ ซึ่งรูปแบบที่พบการใช้มากที่สุดได้แก่ xhtml โดย xhtml มีข้อดีคือจะมีข้อมูลที่เป็นเท็กซ์มากกว่าเว็บทั่วๆ ไป ทำให้โปรแกรมค้นหาสามารถหาข้อมูลได้ง่ายกว่า, เป็นภาษาแบบ Semantic ทำให้ โปรแกรมค้นหาเข้าใจโครงสร้างเอกสารได้ง่ายขึ้นและรู้ว่าข้อความใดมีความสำคัญมากน้อยจากการใช้ Tag ของเราในการ Mark up

ตารางที่ 9 แสดง HTML error, HTML warning และ CSS error ของเว็บไซต์อันดับต้น

เว็บไซต์อันดับต้น	HTML error	HTML warning	CSS error	Dead Links	Total Links
www.bccancer.bc.ca	497.55	17.15	2.95	0	536.2
www.cancer.gov	80.75	32.45	9	0	140
www.drugdigest.org	185	34.8	31	1.05	132.1
www.druginformationonline.org	45.35	2	0.95	0	66.55
www.drugs.com	0	0	73	0	237.65
www.empr.com	115	54.9	134	0	259.05
www.familydoctor.org	62.25	5.5	6	0	68.75
www.health.yahoo.com	50.75	43.05	505.7	0	188.45
www.healthline.com	168.1	134.1	230.25	0.95	333.55
www.healthsquare.com	64.95	9	22	0	90.5
www.mayoclinic.com	110.2	86.85	23.8	0.95	114.95
www.medical-clinic.org	123.25	0	0	0	131.1
www.medicinenet.com	110.1	17.7	111.55	0	231.35
www.medinfo.co.uk	0	2	0.05	0	30.6
www.muschealth.com	59.35	3.6	13.7	0	65.05
www.myoptumhealth.com	147.65	16.55	232	0	177.7
www.nhsdirect.nhs.uk	22.15	0	4	0	174
www.nlm.nih.gov	2.5	0.3	5.7	0	41.55
www.parkinsons.org.uk	6.05	0	6	0	71.7
www.patient.co.uk	104.85	44.85	10	0.05	125.6
www.pdrhealth.com	286.45	106.15	14.05	1	46.3
www.riteaid.com	3.05	0.4	0.95	0	10.2
www.rxlist.com	186.75	17.65	83.95	0.95	301.5
www.walgreens.com	84.9	72.6	26	0.15	181.9
www.webmd.com	230.8	102.65	88.55	0	352.3
รวมทั้งหมด	2747.75	804.25	1635.15	5.1	4108.6
ค่าเฉลี่ยต่อหน้าเว็บเพจ	109.91	32.17	65.406	0.204	164.344
จำนวนเว็บเพจที่พบ Tag HTML เทียบกับจำนวนเว็บเพจทั้งหมดที่ศึกษา (เปอร์เซ็นต์)	92.00	76.60	87.60	-	-

จากการศึกษาปริมาณของค่าความผิดพลาดเฉลี่ยของเว็บไซต์ทั้งในส่วนของ HTML error, HTML warning และ CSS error โดยใช้ Markup Validation Service เป็นเครื่องมือในการศึกษาส่วนของ HTML error และ HTML warning และใช้ CSS Validation Service เป็นเครื่องมือในการจำแนกข้อผิดพลาดต่างๆ ในส่วนของ CSS error พบว่าเว็บไซต์เกือบทั้งหมดมีความผิดพลาดทาง HTML ทั้งที่เป็นความผิดพลาดแบบ error และ warning มีเพียงเว็บไซต์เดียวที่ไม่พบค่าความผิดพลาดทั้งแบบ error และ warning คือ www.drugs.com ส่วนเว็บไซต์ www.medinfo.co.uk ไม่พบ HTML error พบเพียงส่วนของ HTML warning และเว็บไซต์ www.parkinsons.org.uk, www.nhsdirect.nhs.uk และ www.medical-clinic.org พบเพียงส่วนของ HTML error นอกจากในส่วนของ HTML แล้ว ในส่วนของ CSS พบว่ามีเพียงเว็บไซต์เดียวที่ไม่มีค่าความผิดพลาดในส่วนนี้ คือ www.medical-clinic.org ส่วนเว็บไซต์ที่พบค่าความผิดพลาดของ HTML มากที่สุด ทั้งแบบ error และ warning คือ www.pdrhealth.com และพบว่าเว็บไซต์ www.health.yahoo.com มีค่าความผิดพลาดของ CSS มากที่สุด ซึ่งจากค่าความผิดพลาดต่างๆ อาจส่งผลต่อการแสดงผลบนหน้าเว็บไซต์ เช่นไม่สามารถเข้าถึงลิงค์ต่างๆ ตามที่ผู้พัฒนาเว็บไซต์ต้องการได้ หรือ การแสดงผลรูปภาพไม่แสดงเนื่องจากเนื่องจากการส่งค่า บนเว็บไซต์เกิดความผิดพลาดจาก HTML ที่เขียน และความผิดพลาดเหล่านี้อาจส่งผลให้ Robot ที่มีหน้าที่ในการเก็บข้อมูลของโปรแกรมค้นหาไม่สามารถเข้าถึงหรือไม่เข้าใจเกี่ยวกับ HTML เหล่านั้น

ในส่วนของการศึกษาการเชื่อมโยงที่ไม่สามารถเข้าถึงได้ในเว็บไซต์ (Dead Link) โดยใช้ W3C Link Checker เป็นเครื่องมือช่วยในการหาการเชื่อมโยงที่ไม่สามารถเข้าถึงได้ พบว่ามีเว็บไซต์อันดับต้น จำนวน 18 เว็บไซต์ที่ไม่พบการเชื่อมโยงที่ไม่สามารถเข้าถึงได้ในเว็บไซต์ ซึ่งจำนวนของการเชื่อมโยงที่ไม่สามารถเข้าถึงได้ภายในเว็บไซต์จะส่งผลกระทบต่ออันดับของหน้าเว็บไซต์ที่ได้จากการค้นหาผ่านโปรแกรมค้นหา โดยจะทำให้อันดับของเว็บไซต์นั้นๆ ลดลงรวมถึงคะแนนความน่าเชื่อถือของเว็บไซต์จะลดลงด้วย ดังนั้นในการทำเว็บไซต์ควรที่จะทำการตรวจสอบการเชื่อมโยงภายในเว็บไซต์ด้วย

ตารางที่ 10 แสดงค่า Page Rank เฉลี่ยของแต่ละเว็บไซต์อันดับต้น

เว็บไซต์อันดับต้น	Page Rank
www.bccancer.bc.ca	2.75
www.cancer.gov	4.2
www.drugdigest.org	0.35
www.druginformationonline.org	1.4
www.drugs.com	4.65
www.empr.com	2.65
www.familydoctor.org	4.5
www.health.yahoo.com	0.7
www.healthline.com	3.7
www.healthsquare.com	2
www.mayoclinic.com	3.95
www.medical-clinic.org	0.3
www.medicinenet.com	4.85
www.medinfo.co.uk	3.7
www.muschealth.com	0.3
www.myoptumhealth.com	0.4
www.nhsdirect.nhs.uk	0.35
www.nlm.nih.gov	5.1
www.parkinsons.org.uk	2.85
www.patient.co.uk	2.2
www.pdrhealth.com	0.3
www.riteaid.com	0.35
www.rxlist.com	4.75
www.walgreens.com	1.15
www.webmd.com	3.35
รวมทั้งหมด	60.8
ค่าเฉลี่ยต่อหน้าเว็บเพจ	2.43

จากการศึกษา Google Page Rank ซึ่งเป็นค่าลำดับคะแนนที่ Google ประเมินให้กับหน้าเว็บเพจแต่ละหน้า โดยในการศึกษานี้ได้ใช้ Google toolbar ซึ่งติดตั้งบนเว็บเบราเซอร์ในการหาค่าดังกล่าว จากนั้นนำค่าที่ได้ของแต่ละหน้าที่ทำการศึกษามาหาค่าเฉลี่ยของเว็บไซต์ พบว่าเว็บไซต์ที่มีค่า Page Rank สูงสุดได้แก่ www.nlm.nih.gov โดยมีค่า Page Rank เท่ากับ 5.1 รองมาได้แก่

www.medicinenet.com และ www.rxlist.com โดยมีค่า Page Rank เท่ากับ 4.85 และ 4.75 ตามลำดับ ส่วนเว็บไซต์ที่มีค่า Page Rank น้อยที่สุดมีค่า Page Rank เพียง 0.3 เท่านั้น ซึ่งโดยปกติแล้วค่า Page Rank จะถูกกำหนดค่าไว้ตั้งแต่ 0-10 ยิ่งค่า Page Rank มีค่ามากยิ่งทำให้ได้รับการจัดลำดับผลลัพธ์จากการค้นหาในอันดับที่ดีจาก Google แต่พบว่าค่า Page Rank นั้นจะพิจารณาต่อหน้าเว็บเพจเท่านั้นไม่ได้หมายรวมถึงทั้งเว็บไซต์ จากการศึกษาพบว่าในแต่ละเว็บไซต์หน้าที่เป็นหน้าโฮมเพจจะมีค่า Page Rank สูงสุด ส่วนหน้าเว็บเพจอื่นๆ จะมีค่าที่แตกต่างกันออกไปขึ้นกับปัจจัยหลายๆ ด้าน พบว่าเว็บไซต์ที่อยู่อันดับต้นๆ ไม่ได้มีค่า Page Rank สูงเหมือนกันทุกเว็บไซต์ ส่วนเว็บไซต์ที่อยู่อันดับท้ายๆ อาจมีค่า Page Rank ที่สูงได้เช่นเดียวกัน

4.5.2 การศึกษาโครงสร้างของเอกสาร HTML และคุณลักษณะต่างๆ จากเว็บไซต์ที่อยู่อันดับท้ายศึกษา 5 ข้อหลักเช่นเดียวกับการศึกษาในหัวข้อ 4.5.1

1. การศึกษาปริมาณ Tag HTML เฉลี่ยที่ปรากฏบนเอกสาร HTML และจำนวนเว็บเพจที่พบ Tag HTML เทียบกับจำนวนเว็บเพจทั้งหมดที่ศึกษา

2. ปริมาณจุดเชื่อมโยงนอกและจุดเชื่อมโยงภายใน

3. การศึกษาขนาดของหน้าเว็บเพจเฉลี่ยในแต่ละเว็บไซต์, เวลาที่ใช้ในการดาวน์โหลด (Connection Rate 1.44 Mbps.), ขนาดภาพที่พบภายในเว็บเพจ, ภาพที่ไม่สามารถเรียกดูได้ในเว็บไซต์และขนาดของ CSS ที่พบภายในเว็บเพจ

4. การศึกษาประเภทของเอกสาร HTML และการเข้ารหัสภาษา

5. การศึกษาปริมาณของค่าความผิดพลาดเฉลี่ยของเว็บไซต์อันดับท้าย ได้แก่ HTML error, HTML warning และ CSS error, จำนวนการเชื่อมโยงทั้งหมดที่พบ, การเชื่อมโยงที่ไม่สามารถเข้าถึงได้ในเว็บไซต์

6. การศึกษา Google Page Rank เฉลี่ยของเว็บไซต์อันดับท้าย

ตารางที่ 11 แสดงปริมาณ Tag เฉลี่ยต่อหน้าเว็บเพจที่ปรากฏบนเอกสาร HTML ของเว็บไซต์อันดับท้าย

เว็บไซต์อันดับท้าย	Tag HTML ที่ศึกษา											
	a	h1	h2	h3	h4	h5	h6	img	link	meta	p	title
drugs.medsort.com	55.25	15.8	32.35	23.35	8.35	0.65	0	39	2.95	0.95	289.75	1.95
www.allina.com	30.7	0	0.95	6.7	7.4	0	0	12.75	6.95	0.35	3.15	1
www.aol.com	182.25	1	2	13.65	7.6	0	0.05	28.45	2.05	3.95	23.15	1
www.canada.com	173.25	5	4.55	12.35	1.2	0	0	16.65	5.55	4.9	23.85	1
www.drug3k.com	98.9	0	0	0	0	0	0	112	4	5	3.1	0.95
www.everyonehealthy.com	170.3	1.9	6.65	3.45	0.5	0	0	7.9	1.95	0	17.6	1
www.grubbscare.com	6.4	0.1	0	0	0	0	0	0.15	0.05	0.1	12.9	1
www.mdonweb.org	14	0	0	0	0	0	0	8.4	1.95	5.85	9.6	1
www.medindia.net	228.55	1.25	3.05	0.05	0.25	0.05	0	30.1	3.9	4	53	1
www.ralphs.com	74	4.45	29.95	1.05	5.3	0.95	0	6.35	2.35	4.2	176.6	1
ค่าเฉลี่ยของแต่ละ Tag HTML	103.36	2.95	7.95	6.06	3.06	0.17	0.01	26.18	3.17	2.93	61.27	1.09
จำนวนเว็บเพจที่พบ Tag HTML เทียบกับจำนวนเว็บเพจทั้งหมดที่ศึกษา (เปอร์เซ็นต์)	100.00	59.50	68.00	60.00	48.00	11.00	0.50	90.50	90.50	70.50	100.00	99.50

จากตารางที่ 11 พบว่าเว็บไซต์อันดับท้ายมีปริมาณเฉลี่ยของ Tag a มากที่สุด รองมาได้แก่ Tag p และ Tag img ตามลำดับ เว็บไซต์ที่มีปริมาณ Tag a มากที่สุด ได้แก่ www.medindia.net พบว่าค่าเฉลี่ยของ Tag a ของเว็บไซต์อันดับท้ายเท่ากับ 103.36 ในส่วนของ และเมื่อดูจากจำนวนเพจทั้งหมดที่พบเทียบกับจำนวนเพจที่ศึกษา จะทราบว่ามีการใช้ในทุกหน้าเว็บเพจที่ทำการศึกษาTag h1 พบว่าเว็บไซต์ที่พบ Tag นี้มากที่สุดได้แก่ drugs.medsort.com ซึ่งค่าเฉลี่ยของ Tag h1 เท่ากับ 2.95 ซึ่งมี 3 เว็บไซต์ที่ไม่มีการใช้ Tag ดังกล่าว Tag h6 พบว่ามีปริมาณ Tag น้อยที่สุดเมื่อเทียบกับ Tag อื่นๆ โดยพบเพียง 1 เว็บไซต์ ได้แก่ www.aol.com ในส่วนของ Tag img ซึ่งพบปริมาณของ Tag มากเป็นอันดับต้นๆ โดยค่าเฉลี่ยของ Tag นี้ในเว็บอันดับท้ายเท่ากับ 26.18, ในส่วนของ Tag p ซึ่งเป็นอีกตัวแปรที่มีปริมาณ Tag เฉลี่ยมาก พบว่ามีการใช้ในทุกเว็บเพจของเว็บไซต์อันดับท้ายเมื่อดูจากจำนวนเพจทั้งหมดที่พบเทียบกับจำนวนเพจที่ศึกษา พบค่าเฉลี่ยปริมาณ Tag ในเว็บอันดับท้ายๆ เท่ากับ 61.27 จากการศึกษาอาจกล่าวได้ว่าเว็บไซต์อันดับท้าย ส่วนใหญ่จะประกอบไปด้วย Tag HTML ที่ใช้มากอยู่ด้วยกัน 3 Tag เช่นเดียวกันกับเว็บไซต์อันดับต้นคือ Tag a, Tag p และ Tag img

ตารางที่ 12 แสดงคุณลักษณะของเอกสาร HTML ในส่วนของจุดเชื่อมโยงนอกและจุดเชื่อมโยงภายในของเว็บไซต์อันดับท้าย

เว็บไซต์อันดับท้าย	คุณลักษณะ	
	จุดเชื่อมโยงนอก	จุดเชื่อมโยงภายใน
drugs.medsort.com	5.25	27.9
www.allina.com	24.35	0.65
www.aol.com	129.05	50.35
www.canada.com	39.1	11.3
www.drug3k.com	39.45	52.5
www.everyonehealthy.com	162.15	8.1
www.grubbscare.com	2.25	0.95
www.mdonweb.org	1.3	0
www.medindia.net	196.05	2.75
www.ralphs.com	9.55	58.65
ค่าเฉลี่ยของจำนวนการเชื่อมโยงแต่ละเว็บเพจ	60.85	21.32
จำนวนเว็บเพจที่พบการเชื่อมโยง เทียบกับจำนวนเว็บเพจทั้งหมดที่ศึกษา (เปอร์เซ็นต์)	90.50	60.00

จากตารางที่ 12 การศึกษาจุดเชื่อมโยงภายนอกและจุดเชื่อมโยงภายในพบว่ามีเว็บไซต์ที่ไม่มีการเชื่อมโยงภายในเว็บไซต์ได้แก่ www.mdonweb.org ซึ่งทำให้การเข้าถึงข้อมูลต่างๆ ภายในเว็บไซต์ทำได้ยาก และจากเว็บไซต์อันดับท้ายที่ทำการศึกษาพบว่ามีจำนวนของการเชื่อมโยงภายในเฉลี่ยน้อยกว่าการเชื่อมโยงภายนอกมาก เมื่อคิดเป็นจำนวนเพจทั้งหมดที่พบเทียบกับจำนวนเพจที่ศึกษาให้ผลเช่นเดียวกัน

ตารางที่ 13 แสดงขนาดของหน้าเว็บเพจเฉลี่ยในแต่ละเว็บไซต์อันดับท้าย, เวลาที่ใช้ในการดาวน์โหลด (Connection Rate 1.44 Mbps.), ขนาดภาพที่พบภายในเว็บเพจเฉลี่ย, ขนาดของ CSS ที่พบภายในเว็บเพจ, ภาพที่ไม่สามารถเรียกดูได้ในเว็บไซต์อันดับท้าย (Dead Images) และจำนวนภาพรวมภายในเว็บไซต์อันดับท้าย

เว็บไซต์อันดับท้าย	ขนาดของหน้าเอกสารHTML (กิโลไบต์)	เวลาที่ใช้ในการดาวน์โหลด(วินาที)	ขนาดของภาพ (กิโลไบต์)	ขนาดของ CSS (กิโลไบต์)	รูปภาพที่ไม่สามารถเข้าถึงได้	จำนวนรูปภาพรวมเฉลี่ยต่อหน้าเว็บเพจ
drugs.medsort.com	726.94	10.53	18.16	0	0	13.05
www.allina.com	104.37	7.81	8.44	61.30	0.95	6.3
www.aol.com	130.85	1.94	46.30	21.68	0	9.35
www.canada.com	239.06	3.63	48.29	6.30	0	12.95
www.drug3k.com	164.80	5.44	53.01	37.37	0.1	12.75
www.everyonehealthy.com	157.38	5.2	117.77	0.76	0.1	6.75
www.grubbscare.com	10.75	1.88	2.01	0	0	0.1
www.mdonweb.org	29.58	2.32	1.95	3.52	0	0.4
www.medindia.net	93.16	4.4	18.81	23.13	0.95	14.45
www.ralphs.com	222.06	5.08	18.33	68.70	0	5.3
รวมทั้งหมด	1878.96	48.23	333.07	222.75	2.1	81.4
ค่าเฉลี่ยต่อหน้าเว็บเพจ	187.90	4.823	33.31	22.27	0.21	8.14

จากตารางที่ 13 การศึกษาขนาดของหน้าเว็บเพจเฉลี่ยในแต่ละเว็บไซต์และเวลาที่ใช้ในการดาวน์โหลด โดยใช้ Web Page Analyzer-0.98 จากเว็บไซต์ ww.websiteoptimization.com ซึ่งเป็นเครื่องมือที่ช่วยในการจำแนกขนาดและความเร็วของเว็บไซต์พบว่า drugs.medsort.com มีขนาดหน้าเว็บเพจเฉลี่ยขนาดใหญ่ที่สุดเมื่อเทียบจากเว็บไซต์อันดับท้ายทั้งหมดที่ทำการศึกษาและ www.grubbscare.com เป็นเว็บไซต์ที่มีขนาดหน้าเว็บเพจเฉลี่ยน้อยที่สุด โดยขนาดหน้าเว็บเพจเฉลี่ยของเว็บไซต์อันดับท้ายมีค่าเท่ากับ 187.9 กิโลไบต์เมื่อทำการศึกษาในเรื่องของเวลาที่ใช้ในการดาวน์โหลดหน้าเว็บเพจพบว่า เว็บไซต์ที่ใช้เวลาในการดาวน์โหลดหน้าเว็บเพจมากที่สุดได้แก่เว็บไซต์ drugs.medsort.com โดยใช้เวลาในการเข้าถึง 10.53 วินาที และเว็บไซต์ที่ใช้เวลาในการเข้าถึงน้อยที่สุดได้แก่ www.grubbscare.com โดยใช้เวลา 1.88 วินาที จากการศึกษานี้สรุปได้ว่าเว็บไซต์ที่มีขนาดของหน้าเว็บเพจขนาดใหญ่ จะใช้ระยะเวลาในการเข้าถึงหน้าเว็บเพจ นานกว่าเว็บเพจที่มีหน้าเว็บเพจขนาดเล็ก นอกจากนี้เว็บไซต์อันดับท้ายมีค่าเวลาในการดาวน์โหลดหน้าเว็บของเว็บต้นเฉลี่ยอยู่ที่ 4.823 วินาที

ในส่วนของขนาดภาพที่พบภายในเว็บเพจรวมเฉลี่ย พบว่าเว็บไซต์ www.everyonehealthy.com มีขนาดภาพรวมมากที่สุด คือ 117.77 กิโลไบต์ส่วนเว็บไซต์ www.mdonweb.org เป็นเว็บไซต์ที่มีขนาดของขนาดภาพรวมเฉลี่ยน้อยที่สุด โดยมีขนาดเท่ากับ 8.44 กิโลไบต์โดยขนาดของภาพที่พบภายในเว็บเพจรวมเฉลี่ยของเว็บไซต์อันดับท้าย มีค่าเท่ากับ 33.31 กิโลไบต์

การศึกษาขนาดของ CSS ที่พบภายในเว็บเพจเฉลี่ยโดยรวม พบว่ามี 2 เว็บไซต์ที่ไม่มีการใช้ CSS ได้แก่ drugs.medsort.com และ www.grubbscare.com นอกจากนี้พบว่าค่าเฉลี่ยของขนาด CSS ที่พบภายในเว็บเพจเท่ากับ 22.27 กิโลไบต์

จากการศึกษาภาพที่ไม่สามารถเรียกดูได้ภายในเว็บไซต์ พบว่ามีทั้งหมด 4 เว็บไซต์ที่พบภาพที่ไม่สามารถเรียกดูได้ในเว็บไซต์ คิดเป็น 40 เปอร์เซ็นต์ของเว็บไซต์อันดับท้าย นอกจากนี้พบว่าแต่ละเว็บไซต์จะมีจำนวนรูปภาพต่อหน้าเว็บเพจเฉลี่ยที่แตกต่างกันออกไป เช่นเดียวกันกับเว็บไซต์อันดับต้น

ตารางที่ 14 แสดงประเภทของเอกสาร HTML และการเข้ารหัสภาษาของเว็บไซต์อันดับท้าย

เว็บไซต์อันดับท้าย	การเข้ารหัสภาษา			ประเภทของเอกสาร HTML	
	utf-8	iso-8859-1	windows-1252	html	xhtml
drugs.medsort.com	20	0	0	0	20
www.allina.com	19	1	0	1	19
www.aol.com	20	0	0	0	20
www.canada.com	1	0	19	19	1
www.drug3k.com	20	0	0	0	20
www.everyonehealthy.com	20	0	0	19	1
www.grubbscare.com	20	0	0	19	1
www.mdonweb.org	0	1	19	20	0
www.medindia.net	0	0	20	20	0
www.ralphs.com	20	0	0	20	0
รวมทั้งหมด	140	2	58	118	82
ค่าเฉลี่ยต่อหน้าเว็บเพจ	14.00	0.20	5.80	11.80	8.20

จากตารางพบว่า เว็บไซต์อันดับท้าย จะมีการใช้ windows-1252 ถึง 28.5 เปอร์เซ็นต์ และมีการใช้ iso-8859-1 เพียงแค่ 1 เปอร์เซ็นต์เท่านั้น แต่มีการใช้ utf-8 มากเช่นเดียวกันกับเว็บไซต์อันดับต้น คือ 70.5 เปอร์เซ็นต์ ซึ่งจากการศึกษาอาจกล่าวได้ว่าการเข้ารหัสภาษาในเว็บไซต์ส่วนใหญ่นั้นจะใช้รูปแบบ utf-8 ซึ่ง Unicode (utf-8) คือ character ที่สามารถรองรับการทำงานระบบหลายภาษา (Multi language) ระบบ utf-8 จะเก็บข้อมูล 1 - 4 bytes ต่อ 1 ตัวอักษรจึงเป็นที่นิยมใช้กันอย่างกว้างขวาง ส่วน windows-1252 เป็นชุดอักขระมาตรฐานของ ภาษา ยุโรป ทางตะวันตก Western European languages ของ windows และ iso-8859-1 เป็นชุดอักขระมาตรฐานของภาษา ยุโรป ทางตะวันตก Western European languages เมื่อเปรียบเทียบในส่วนของประเภทของภาษา (คำสั่งภาษาอังกฤษทำหน้าที่บอกให้เว็บเบราเซอร์ทราบว่า เว็บเพจของเราใช้ภาษา HTML หรือ XHTML และคำสั่ง HTML/XHTML ที่ใช้เป็นรุ่นใด และบอกชนิดของเอกสาร (Document Type Definition : DTD) ที่ใช้) ซึ่งจะช่วยให้เว็บเบราเซอร์แปลเอกสารได้อย่างถูกต้องพบว่าเว็บไซต์อันดับท้ายมีการใช้ประเภทของเอกสาร HTML แบบ xhtml 1.0 transitional มากที่สุดคือ 40.50 เปอร์เซ็นต์ เนื่องจาก เป็นรูปแบบที่สามารถใช้ HTML 4.0 ร่วมด้วยได้ และสามารถเรียกใช้ทั้ง CSS กับ XHTML ซึ่งมีข้อดีคือจะช่วยให้บราวเซอร์รุ่นเก่าๆ โดยเฉพาะ IE6 นั้นสามารถแปลเอกสารได้

ใกล้เคียงกัน นอกจากนี้ในกรณีที่ XHTML บางแท็กไม่สามารถใช้กับ บราวเซอร์รุ่นเก่าๆ ก็สามารถที่จะใช้ html คู่ไปได้ และแบบที่มีการใช้น้อยที่สุดคือ html5 ซึ่งไม่พบการใช้เลยในเว็บไซต์อันดับท้าย อาจเนื่องมาจาก html5 เป็นรูปแบบของเอกสาร HTML ที่อยู่ในระหว่างการพัฒนา โดยมีลักษณะเหมือนมาตรฐานตัวก่อนหน้าทั้ง HTML 4.01 และ XHTML 1.1 ที่ใช้ในการจัดโครงสร้างและการแสดงผลของเนื้อหาสำหรับเว็บเพจโดยมีคุณลักษณะเด่นที่สำคัญได้แก่ การใช้งานวิดีโอ, การแสดงตำแหน่งทางภูมิศาสตร์, การเก็บไฟล์ในลักษณะออฟไลน์, การแสดงกราฟิกส์, input types แบบใหม่ เช่น search, number, range, color, tel, url, email, date, month, week, time, datetime, datetime-local โดยคุณสมบัติเด่นหลายอย่างไม่จำเป็นต้องติดตั้งซอฟต์แวร์เพิ่ม เช่น flash เหมือนที่ผ่านมาในการใช้งาน HTML 4 จึงทำให้การใช้งานยังไม่เป็นที่นิยมนัก นอกจากนี้พบว่ารูปแบบที่มีการใช้ในเว็บอันดับท้ายๆ มากกว่าได้แก่รูปแบบ html 4.01 และ html 4.01 transitional และรูปแบบ xhtml 1.0 strict (ในการศึกษานี้จัดกลุ่มรูปแบบการใช้เป็น 2 รูปแบบหลักๆ คือ html และ xhtml)

ตารางที่ 15 แสดง HTML error, HTML warning และ CSS error ของเว็บไซต์อันดับท้าย

เว็บไซต์อันดับท้าย	HTML error	HTML warning	CSS error	Dead Links	Total Links
drugs.medsort.com	132.15	4.15	9.75	0	92.7
www.allina.com	62.8	3.85	53.15	0.95	47.35
www.aol.com	78.65	63.2	228.6	1	208.05
www.canada.com	107.95	22.25	63.6	0	187.35
www.drug3k.com	1301.65	105.15	8.95	0.1	208.6
www.everyonehealthy.com	13.25	23.05	6.95	0.1	179.2
www.grubbscare.com	26.4	4.2	0.2	0	6.5
www.mdonweb.org	77	6.3	10.55	0.05	11.4
www.medindia.net	101.65	39.95	13.35	0.95	287.3
www.ralphs.com	43.95	23.35	7.55	0.05	82.45
รวมทั้งหมด	1945.45	295.45	402.65	3.2	1310.9
ค่าเฉลี่ยต่อหน้าเว็บเพจ	194.545	29.545	40.265	0.32	131.09
จำนวนเว็บเพจที่พบ Tag HTML เทียบกับจำนวนเว็บเพจทั้งหมดที่ศึกษา (เปอร์เซ็นต์)	100.00	100.00	90.50	-	-

จากตารางแสดงผลการศึกษาค่าความผิดพลาดของเอกสาร HTML ที่พบในเว็บไซต์ที่ให้ข้อมูลยาที่เป็นภาษาอังกฤษ ทั้งในส่วนของ HTML error, HTML warning และ CSS error พบว่าเว็บไซต์ที่อยู่อันดับท้ายมีความผิดพลาดของเอกสาร HTML ทั้งที่เป็น แบบ error และ warning คิดเป็น 100 เปอร์เซ็นต์ นอกจากนี้ยังพบว่าเว็บไซต์อันดับท้าย พบค่า warning เฉลี่ยเท่ากับ 29.545 ส่วนค่าความผิดพลาดที่เป็น CSS error เฉลี่ยเท่ากับ 40.265 ซึ่งในส่วนของ error และ warning ต่างๆ หากมีการแก้ไขหรือปรับปรุงเว็บเพจให้มีค่าความผิดพลาดน้อยลงหรือไม่มีความผิดพลาดเลย อาจส่งผลให้เว็บไซต์มีอันดับดีขึ้นกว่าเดิมได้

ในส่วนของการศึกษาการเชื่อมโยงที่ไม่สามารถเข้าถึงได้ในเว็บไซต์ พบว่ามีเว็บไซต์เพียง 3 เว็บไซต์ที่ไม่พบการเชื่อมโยงที่ไม่สามารถเข้าถึงได้ในเว็บไซต์ คิดเป็น 30 เปอร์เซ็นต์ของเว็บไซต์อันดับท้าย

ตารางที่ 16 แสดงค่า Page Rank เฉลี่ยของแต่ละเว็บไซต์อันดับท้าย

เว็บไซต์อันดับต้น	Page Rank
drugs.medsort.com	0.05
www.allina.com	0.25
www.aol.com	4.4
www.canada.com	1.7
www.drug3k.com	1.75
www.everyonehealthy.com	0.15
www.grubbscare.com	0.15
www.mdonweb.org	0.15
www.medindia.net	0.9
www.ralphs.com	0.4
รวมทั้งหมด	9.9
ค่าเฉลี่ยต่อหน้าเว็บเพจ	0.99

จากตารางพบว่าเว็บไซต์ที่อยู่ในอันดับท้ายๆ มีค่าเฉลี่ย Page Rank เพียง 0.99 ซึ่งถือว่าน้อยมาก (Google Page Rank จัดอันดับคะแนนไว้ตั้งแต่ 0-10) อาจกล่าวได้ว่าเว็บไซต์อันดับท้ายจะมีค่า Page Rank น้อย

4.5.3 การศึกษาโครงสร้างของเอกสาร HTML และคุณลักษณะต่างๆ จากเว็บไซต์ที่ที่มีคะแนนการประเมินตั้งแต่ 60 เปอร์เซ็นต์ขึ้นไป (เว็บไซต์กลุ่ม A) โดยศึกษา 6 ข้อหลักเช่นเดียวกับการศึกษาในหัวข้อ 4.5.1

1. การศึกษาปริมาณ Tag HTML เฉลี่ยที่ปรากฏบนเอกสาร HTML และจำนวนเว็บเพจที่พบ Tag HTML เทียบกับจำนวนเว็บเพจทั้งหมดที่ศึกษา

2. ปริมาณจุดเชื่อมโยงภายนอกและจุดเชื่อมโยงภายใน

3. การศึกษาขนาดของหน้าเว็บเพจเฉลี่ยในแต่ละเว็บไซต์, เวลาที่ใช้ในการดาวน์โหลด (Connection Rate 1.44 Mbps.), ขนาดภาพที่พบภายในเว็บเพจ, ภาพที่ไม่สามารถเรียกดูได้ในเว็บไซต์ (Dead Images) และขนาดของ CSS ที่พบภายในเว็บเพจ

4. การศึกษาประเภทของเอกสาร HTML และการเข้ารหัสภาษา

5. การศึกษาปริมาณของค่าความผิดพลาดเฉลี่ยของเว็บไซต์ ได้แก่ HTML error, HTML warning และ CSS error, จำนวนการเชื่อมโยงทั้งหมดที่พบ, การเชื่อมโยงที่ไม่สามารถเข้าถึงได้ในเว็บไซต์

6. การศึกษา Google Page Rank เฉลี่ยของเว็บไซต์

ตารางที่ 17 แสดงปริมาณ Tag เฉลี่ยต่อหน้าเว็บเพจที่ปรากฏบนเอกสาร HTML ของเว็บไซต์กลุ่ม A

เว็บไซต์กลุ่ม A	Tag HTML ที่ศึกษา											
	a	h1	h2	h3	h4	h5	h6	img	link	meta	p	title
www.bccancer.bc.ca	154.9	0.1	2.65	12.55	1.75	0.05	0.2	367.1	2	16.15	279.05	1
www.drugdigest.org	66.25	1.4	0	0	0	0	0	70.7	9	0.1	5	2.05
www.drugs.com	208.45	1	13.15	0	0	0	0	15.65	6.15	6	80.3	1
www.empr.com	227	0.95	14.95	28.15	41.8	0	0	20.35	5	5	47.6	1
www.familydoctor.org	56.8	2	0	8.4	0.05	0	0	24.35	6	5.95	29.1	1
www.health.yahoo.com	174.8	2.15	0	8.15	0.95	0	0	16.6	1.95	1.05	7.7	1
www.healthline.com	241.5	1.9	0.25	2.85	0	0	0	55.75	7.85	2.1	9.85	1
www.healthsquare.com	81.25	2.4	6.6	1.6	1.4	0	0	10.35	3	3	10.1	1
www.mayoclinic.com	80.7	4.2	5.4	7.65	3.15	0	0	12.05	4	8	24.5	1
www.medicinenet.com	193.7	1.2	1	12.6	2.45	3.1	0	27.2	12.2	7	41.05	0
www.medinfo.co.uk	30.45	1	1	6.25	0.75	0	0	2.1	1	1.6	17.9	1
www.nhsdirect.nhs.uk	130.2	2	3.95	1.05	1.3	0	0	4.4	12.9	14.4	17.35	1

ตารางที่ 17 แสดงปริมาณ Tag เฉลี่ยต่อหน้าเว็บเพจที่ปรากฏบนเอกสาร HTML ของเว็บไซต์กลุ่ม A

เว็บไซต์กลุ่ม A	Tag HTML ที่ศึกษา											
	a	h1	h2	h3	h4	h5	h6	img	link	meta	p	title
www.nlm.nih.gov	30.45	1	11.1	3.5	0	0	0	4.9	7.9	4.9	22.4	1
www.rxlist.com	296.75	1.3	4	25.55	22.35	22.3	0	31.45	9.2	5	225.45	1
www.webmd.com	322	1.9	1.1	7.15	4.05	1	7.65	32.9	13.8	4.95	51.3	1
drugs.medsort.com	55.25	15.8	32.35	23.35	8.35	0.65	0	39	2.95	0.95	289.75	1.95
www.allina.com	30.7	0	0.95	6.7	7.4	0	0	12.75	6.95	0.35	3.15	1
www.aol.com	182.25	1	2	13.65	7.6	0	0.05	28.45	2.05	3.95	23.15	1
www.everyonehealthy.com	170.3	1.9	6.65	3.45	0.5	0	0	7.9	1.95	0	17.6	1
www.ralphs.com	74	4.45	29.95	1.05	5.3	0.95	0	6.35	2.35	4.2	176.6	1
จำนวนเว็บเพจที่พบ Tag HTML เทียบกับจำนวนเว็บเพจทั้งหมดที่ศึกษา (เปอร์เซ็นต์)	100	88.75	78.25	88.75	64.00	20.50	5.50	100	100	85.25	99.75	95.00

จากตารางที่ 17 พบว่าเว็บไซต์กลุ่ม A มีปริมาณเฉลี่ยของ Tag a มากที่สุด รองมาได้แก่ Tag p และ Tag img ตามลำดับ เว็บไซต์ที่มีปริมาณ Tag a มากที่สุด ได้แก่ www.rxlist.com พบว่าค่าเฉลี่ยของ Tag a เท่ากับ 296.75 และเมื่อดูจากจำนวนเพจทั้งหมดที่พบเทียบกับจำนวนเพจที่ศึกษาจะทราบว่ามีการใช้ Tag a Tag img และ Tag link ในทุกหน้าเว็บเพจที่ทำการศึกษา พบว่ามี 1 เว็บไซต์ที่ไม่มีการใช้ Tag h1 และ Tag h6 พบว่ามีปริมาณ Tag น้อยที่สุดเมื่อเทียบกับ Tag อื่นๆ โดยพบเพียง 3 เว็บไซต์ ในส่วนของ Tag img ซึ่งพบปริมาณของ Tag มากเป็นอันดับต้นๆ โดยค่าเฉลี่ยของ Tag นี้ในเว็บอันดับท้ายเท่ากับ 39.52, ในส่วนของ Tag p ซึ่งเป็นอีกตัวแปรที่มีปริมาณ Tag เฉลี่ยมาก มีค่าเฉลี่ยปริมาณ Tag ในเว็บไซต์กลุ่ม A เท่ากับ 68.95 จากการศึกษาอาจกล่าวได้ว่าเว็บไซต์กลุ่ม A ส่วนใหญ่จะประกอบไปด้วย Tag HTML ที่ใช้มากอยู่ด้วยกัน 3 Tag เช่นเดียวกันกับเว็บไซต์อันดับต้นคือ Tag a, Tag p และ Tag img

นอกจากนี้จากตารางที่ 18พบว่าในส่วนของจุดเชื่อมโยงภายนอกและจุดเชื่อมโยงภายในนั้นมีการใช้ จุดเชื่อมโยงภายนอกในทุกเว็บเพจของเว็บไซต์กลุ่ม A ที่ทำการสำรวจคิดเป็น 100 เปอร์เซ็นต์ ซึ่งเว็บเพจส่วนใหญ่มีการเชื่อมโยงไปยังเว็บไซต์ที่มีความน่าเชื่อและมีเนื้อหาเกี่ยวข้องกับข้อมูลที่นำเสนอในเว็บเพจนั้นๆ ในส่วนของจุดเชื่อมโยงภายในพบว่ามีเพียง www.nhsdirect.nhs.uk ที่ไม่มีการเชื่อมโยงภายในซึ่งอาจทำให้ผู้เยี่ยมชมเว็บไซต์ลำบากในการติดตามอ่านในหน้าเว็บเพจอื่นๆ ของเว็บไซต์

ตารางที่ 4.18 แสดงคุณลักษณะของเอกสาร HTML ในส่วนของจุดเชื่อมโยงภายนอกและจุดเชื่อมโยงภายในของเว็บไซต์กลุ่ม A

เว็บไซต์กลุ่ม A	คุณลักษณะ	
	จุดเชื่อมโยงภายนอก	จุดเชื่อมโยงภายใน
www.bccancer.bc.ca	15.05	138.85
www.drugdigest.org	7	49.65
www.drugs.com	41.15	155.9
www.empr.com	192.05	33
www.familydoctor.org	16	38.8
www.health.yahoo.com	56.1	116.75
www.healthline.com	3.1	231.5
www.healthsquare.com	58.3	20.1
www.mayoclinic.com	24	50.45
www.medicinenet.com	130.35	57.95
www.medinfo.co.uk	10.3	17.7
www.nhsdirect.nhs.uk	117.8	0
www.nlm.nih.gov	13.35	17.1
www.rxlist.com	55.7	196.2
www.webmd.com	283.7	14.15
drugs.medsort.com	5.25	27.9
www.allina.com	24.35	0.65
www.aol.com	129.05	50.35
www.everyonehealthy.com	162.15	8.1
www.ralphs.com	9.55	58.65
จำนวนเว็บเพจที่พบการเชื่อมโยงเทียบกับจำนวนเว็บเพจทั้งหมดที่ศึกษา (เปอร์เซ็นต์)	100.00	89.50

ตารางที่ 19 แสดงขนาดของหน้าเว็บเพจเฉลี่ยในแต่ละเว็บไซต์กลุ่ม A, เวลาที่ใช้ในการดาวน์โหลด (Connection Rate 1.44 Mbps.), ขนาดภาพที่พบภายในเว็บเพจเฉลี่ย, ขนาดของ CSS ที่พบภายในเว็บเพจ, ภาพที่ไม่สามารถเรียกดูได้ในเว็บไซต์กลุ่ม A (Dead Images) และจำนวนภาพรวมภายในเว็บไซต์กลุ่ม A

เว็บไซต์กลุ่ม A	ขนาดของหน้าเอกสารHTML (กิโลไบต์)	เวลาที่ใช้ในการดาวน์โหลด (วินาที)	ขนาดของภาพ (กิโลไบต์)	ขนาดของ CSS (กิโลไบต์)	รูปภาพที่ไม่สามารถเข้าถึงได้	จำนวนรูปภาพรวมเฉลี่ยต่อหน้าเว็บเพจ
www.bccancer.bc.ca	35.77	2.83	18.20	0	0	126.05
www.drugdigest.org	290.09	7.61	37.15	2.12	1	17.75
www.drugs.com	117.56	3.3	47.87	2.70	0	14.1
www.empr.com	365.58	4.63	87.77	44.28	0	17.45
www.familydoctor.org	126.08	2.24	20.95	24.41	0	10.55
www.health.yahoo.com	214.23	1.28	83.92	32.82	0	7.9
www.healthline.com	190.10	0.89	91.53	12.88	0	27.3
www.healthsquare.com	76.53	2.15	4.87	8.49	0	2.7
www.mayoclinic.com	211.09	3.11	49.80	61.86	0	4.8
www.medicinenet.com	241.64	1.75	88.49	56.45	0	15.8
www.medinfo.co.uk	14.80	2.3	4.80	0.02	0	1.95
www.nhsdirect.nhs.uk	708.16	11.31	82.34	165.18	0	4
www.nlm.nih.gov	108.53	1.5	6.99	38.43	0	1.95
www.rxlist.com	285.00	3.55	138.71	44.99	0	27

ตารางที่ 19 แสดงขนาดของหน้าเว็บเพจเฉลี่ยในแต่ละเว็บไซต์กลุ่ม A, เวลาที่ใช้ในการดาวน์โหลด (Connection Rate 1.44 Mbps.), ขนาดภาพที่พบภายในเว็บเพจเฉลี่ย, ขนาดของ CSS ที่พบภายในเว็บเพจ, ภาพที่ไม่สามารถเรียกดูได้ในเว็บไซต์กลุ่ม A (Dead Images) และจำนวนภาพรวมภายในเว็บไซต์กลุ่ม A (ต่อ)

เว็บไซต์กลุ่ม A	ขนาดของหน้าเอกสารHTML (กิโลไบต์)	เวลาที่ใช้ในการดาวน์โหลด (วินาที)	ขนาดของภาพ (กิโลไบต์)	ขนาดของ CSS (กิโลไบต์)	รูปภาพที่ไม่สามารถเข้าถึงได้	จำนวนรูปภาพรวมเฉลี่ยต่อหน้าเว็บเพจ
www.webmd.com	176.35	2.16	59.09	1.86	0	23.25
drugs.medsort.com	726.94	10.53	18.16	0.00	0	13.05
www.allina.com	104.37	7.81	8.44	61.30	0.95	6.3
www.aol.com	130.85	1.94	46.30	21.68	0	9.35
www.everyonehealthy.com	157.38	5.2	117.77	0.76	0.1	6.75
www.ralphs.com	222.06	5.08	18.33	68.70	0	5.3
ค่าเฉลี่ยต่อหน้าเว็บเพจ	225.16	4.06	51.57	32.45	0.10	17.17
รวมทั้งหมด	4503.11	81.17	1031.48	648.94	2.05	343.3

จากตารางที่ 19 พบว่าเว็บไซต์กลุ่ม A จะมีขนาดของหน้าเว็บเพจเฉลี่ยและเวลาที่ใช้ในการดาวน์โหลดหน้าเว็บเพจเฉลี่ย เท่ากับ 225.16 กิโลไบต์และ 4.06 วินาที ตามลำดับ ทั้งนี้อาจเนื่องมาจากเนื้อหาของข้อมูลด้านยาที่พบ, multimedia ต่างๆ พบในเว็บไซต์กลุ่ม A มีปริมาณมาก แต่เวลาที่ใช้ในการดาวน์โหลดหน้าเว็บเพจพบว่าไม่นานมาก ทั้งนี้อาจเนื่องมาจากหลายปัจจัยเช่น โดเมนที่เป็นที่ตั้งของเว็บไซต์นั้นๆ มีความเสถียรสามารถเข้าถึงได้ง่าย ขนาดรูปภาพมีความเหมาะสม ไม่ใหญ่จนเกินไปมีการปรับขนาดพอดี ดังนั้นหากต้องการให้เว็บไซต์มีคุณภาพดี ควรจะต้องมีเนื้อหาข้อมูลที่ถูกต้องและมากเพียงพอ และต้องมีการปรับปรุงโครงสร้างของเอกสาร HTML ให้สามารถเข้าถึงได้โดยใช้เวลาในการเข้าถึงสั้นด้วย

ในส่วนของขนาดภาพที่พบภายในเว็บเพจรวมเฉลี่ย พบว่าขนาดของภาพที่พบภายในเว็บเพจรวมเฉลี่ยของเว็บไซต์กลุ่ม A มีค่าเท่ากับ 51.57 กิโลไบต์พบว่า www.rxlist.com มีขนาดของภาพที่พบเฉลี่ยสูงสุดคือ 138.71 กิโลไบต์

การศึกษาขนาดของ CSS ที่พบภายในเว็บเพจเฉลี่ยโดยรวม พบว่ามี 2 เว็บไซต์ที่ไม่มีการใช้ CSS ได้แก่ www.bccancer.bc.ca และ drugs.medsort.com นอกจากนี้พบว่าค่าเฉลี่ยของขนาด CSS ที่พบภายในเว็บเพจเท่ากับ 51.57 กิโลไบต์

จากการศึกษาภาพที่ไม่สามารถเรียกดูได้ภายในเว็บไซต์กลุ่ม A พบว่ามีทั้งหมด 3 เว็บไซต์ที่พบภาพที่ไม่สามารถเรียกดูได้ในเว็บไซต์ นอกจากนี้พบว่าแต่ละเว็บไซต์จะมีจำนวนรูปภาพต่อหน้าเว็บเพจเฉลี่ยที่แตกต่างกันออกไป เช่นเดียวกันกับเว็บไซต์กลุ่มอื่นๆ

การศึกษาประเภทของเอกสาร HTML และการเข้ารหัสภาษา พบว่า เว็บไซต์กลุ่ม A ไม่พบการใช้ encoding รูปแบบ windows-1252เลย แต่มีการใช้ utf-8 เป็นส่วนมากส่วน รูปแบบ iso-8859-1 พบมีการใช้เพียง 3 เว็บไซต์จาก 20 เว็บไซต์ ซึ่งจากการศึกษาอาจกล่าวได้ว่าการเข้ารหัสภาษาในเว็บไซต์ส่วนใหญ่นั้นจะใช้รูปแบบ utf-8 ส่วน windows-1252 เป็นชุดอักขระที่ได้รับความนิยมน้อยสุดจากการศึกษานี้ ส่วน iso-8859-1 เป็นชุดอักขระที่มีการใช้บ้างในบางเว็บไซต์กลุ่ม A เมื่อเปรียบเทียบในส่วนของประเภทของภาษา พบว่าเว็บไซต์กลุ่ม A มีการใช้ประเภทของเอกสาร HTML แบบ xhtml มากที่สุดคือ แต่ในบางเว็บไซต์ก็มีการใช้รูปแบบปนกันทั้ง 2 รูปแบบไม่ได้เจาะจงให้เป็นรูปแบบเดียวกันหมดทั้งเว็บไซต์ ผลดังตารางที่ 20

ตารางที่ 20 แสดงประเภทของเอกสาร HTML และการเข้ารหัสภาษาของเว็บไซต์กลุ่ม A

เว็บไซต์กลุ่ม A	การเข้ารหัสภาษา			ประเภทของเอกสาร HTML	
	utf-8	iso-8859-1	windows -1252	html	xhtml
www.bccancer.bc.ca	20	0	0	20	0
www.drugdigest.org	20	0	0	20	0
www.drugs.com	20	0	0	0	20
www.empr.com	20	0	0	0	20
www.familydoctor.org	20	0	0	0	20
www.health.yahoo.com	20	0	0	20	0
www.healthline.com	20	0	0	20	0
www.healthsquare.com	20	0	0	0	20
www.mayoclinic.com	20	0	0	0	20
www.medicinenet.com	20	0	0	0	20
www.medinfo.co.uk	0	20	0	20	0
www.nhsdirect.nhs.uk	20	0	0	0	20
www.nlm.nih.gov	20	0	0	0	20
www.rxlist.com	0	20	0	0	20
www.webmd.com	20	0	0	0	20
drugs.medsort.com	20	0	0	0	20
www.allina.com	19	1	0	1	19
www.aol.com	20	0	0	0	20
www.everyonehealthy.com	20	0	0	19	1
www.ralphs.com	20	0	0	20	0
ค่าเฉลี่ยต่อหน้าเว็บเพจ	17.95	2.05	0.00	7.00	13.00
รวมทั้งหมด	359	41	0	140	260

ตารางที่ 21 แสดง HTML error, HTML warning และ CSS error ของเว็บไซต์กลุ่ม A

เว็บไซต์กลุ่ม A	HTML error	HTML warning	CSS error	Dead Links	Total Links
www.bccancer.bc.ca	497.55	17.15	2.95	0	536.2
www.drugdigest.org	185	34.8	31	1.05	132.1
www.drugs.com	0	0	73	0	237.65
www.empr.com	115	54.9	134	0	259.05
www.familydoctor.org	62.25	5.5	6	0	68.75
www.health.yahoo.com	50.75	43.05	505.7	0	188.45
www.healthline.com	168.1	134.1	230.25	0.95	333.55
www.healthsquare.com	64.95	9	22	0	90.5
www.mayoclinic.com	110.2	86.85	23.8	0.95	114.95
www.medicinenet.com	110.1	17.7	111.55	0	231.35
www.medinfo.co.uk	0	2	0.05	0	30.6
www.nhsdirect.nhs.uk	22.15	0	4	0	174
www.nlm.nih.gov	2.5	0.3	5.7	0	41.55
www.rxlist.com	186.75	17.65	83.95	0.95	301.5
www.webmd.com	230.8	102.65	88.55	0	352.3
drugs.medsort.com	132.15	4.15	9.75	0	92.7
www.allina.com	62.8	3.85	53.15	0.95	47.35
www.aol.com	78.65	63.2	228.6	1	208.05
www.everyonehealthy.com	13.25	23.05	6.95	0.1	179.2
www.ralphs.com	43.95	23.35	7.55	0.05	82.45
รวมทั้งหมด	2136.9	643.25	1628.5	6	3702.25
ค่าเฉลี่ยต่อหน้าเว็บเพจ	106.85	32.16	81.43	0.30	185.11
จำนวนเว็บเพจที่พบ Tag HTML เทียบกับจำนวนเว็บเพจทั้งหมดที่ศึกษา (เปอร์เซ็นต์)	90.00	85.25	95.00	-	-

จากตารางแสดงผลการศึกษาค่าความผิดพลาดของเอกสาร HTML ที่พบในเว็บไซต์ที่ให้ข้อมูลยาที่เป็นภาษาอังกฤษ ทั้งในส่วนของ HTML error, HTML warning และ CSS error พบว่าเว็บไซต์กลุ่ม A มีความผิดพลาดของเอกสาร HTML ทั้งที่เป็น แบบ error และ warning รวมทั้งความผิดพลาดที่เป็น CSS error ด้วย โดยพบ CSS error มากสุด คือ 95.00 เปอร์เซ็นต์ รองลงมาได้แก่

HTML error พบ 90.00 เปอร์เซ็นต์ และ HTML warning พบ 85.25 เปอร์เซ็นต์ ในส่วนของการศึกษาการเชื่อมโยงที่ไม่สามารถเข้าถึงได้ในเว็บไซต์ พบว่ามีเว็บไซต์ 8 เว็บไซต์จาก 20 เว็บไซต์ที่ไม่พบการเชื่อมโยงที่ไม่สามารถเข้าถึง คิดเป็น 40 เปอร์เซ็นต์ของเว็บไซต์กลุ่ม A ซึ่งจัดว่าเป็นจำนวนมาก

ตารางที่ 22 แสดงค่า Page Rank เฉลี่ยของแต่ละเว็บไซต์กลุ่ม A

เว็บไซต์กลุ่ม A	Page Rank
www.bccancer.bc.ca	2.75
www.drugdigest.org	0.35
www.drugs.com	4.65
www.empr.com	2.65
www.familydoctor.org	4.5
www.health.yahoo.com	0.7
www.healthline.com	3.7
www.healthsquare.com	2
www.mayoclinic.com	3.95
www.medicinenet.com	4.85
www.medinfo.co.uk	3.7
www.nhsdirect.nhs.uk	0.35
www.nlm.nih.gov	5.1
www.rxlist.com	4.75
www.webmd.com	3.35
drugs.medsort.com	0.05
www.allina.com	0.25
www.aol.com	4.4
www.everyonehealthy.com	0.15
www.ralphs.com	0.4
ค่าเฉลี่ยต่อหน้าเว็บเพจ	2.63
รวมทั้งหมด	52.6

จากตารางพบว่าเว็บไซต์ที่อยู่ในกลุ่ม A มีค่าเฉลี่ย Page Rank 2.63 โดยพบว่า www.nlm.nih.gov มีค่า Page Rank สูงสุดคือ 5.1 และค่าที่ต่ำที่สุดคือ 0.05 ได้แก่ drugs.medsort.com

4.5.4 การศึกษาโครงสร้างของเอกสาร HTML และคุณลักษณะต่างๆ จากเว็บไซต์ที่ที่มีคะแนนการประเมินต่ำกว่า 60 เปอร์เซ็นต์ (เว็บไซต์กลุ่ม B) โดยศึกษา 6 ข้อหลักเช่นเดียวกับการศึกษาในหัวข้อ 4.5.1

1. การศึกษาปริมาณ Tag HTML เฉลี่ยที่ปรากฏบนเอกสาร HTML, และจำนวนเว็บเพจที่พบ Tag HTML เทียบกับจำนวนเว็บเพจทั้งหมดที่ศึกษา

2. การศึกษาปริมาณจุดเชื่อมโยงภายนอกและจุดเชื่อมโยงภายใน

3. การศึกษาขนาดของหน้าเว็บเพจเฉลี่ยในแต่ละเว็บไซต์, เวลาที่ใช้ในการดาวน์โหลด (Connection Rate 1.44 Mbps.), ขนาดภาพที่พบภายในเว็บเพจ, ภาพที่ไม่สามารถเรียกดูได้ในเว็บไซต์ (Dead Images) และขนาดของ CSS ที่พบภายในเว็บเพจ

4. ประเภทของเอกสาร HTML และการเข้ารหัสภาษา

5. การศึกษาปริมาณของค่าความผิดพลาดเฉลี่ยของเว็บไซต์ ได้แก่ HTML error, HTML warning และ CSS error, จำนวนการเชื่อมโยงทั้งหมดที่พบ, การเชื่อมโยงที่ไม่สามารถเข้าถึงได้ในเว็บไซต์

6. การศึกษา Google Page Rank เฉลี่ยของเว็บไซต์กลุ่ม B

ตารางที่ 23 แสดงปริมาณ Tag เฉลี่ยต่อหน้าเว็บเพจที่ปรากฏบนเอกสาร HTML ของเว็บไซต์กลุ่ม B

เว็บไซต์กลุ่ม B	Tag HTML ที่ศึกษา											
	a	h1	h2	h3	h4	h5	h6	img	link	meta	p	title
www.cancer.gov	72.75	0	0.05	0	1	0	0	84.6	5	5	14.25	1
www.druginformationonline.org	53.65	0.05	6.2	0.25	0.2	0	0	9	3.05	3	9.75	1
www.medical-clinic.org	69.45	0.95	1.9	0.95	0.95	0.95	0	56.8	1	3.05	10.1	1
www.muschealth.com	54.7	0	0	0	0	0	0	12.05	4.8	4.3	12.6	1
www.myoptumhealth.com	185.15	2.1	9.2	1.25	0	0	0	12.25	4	1	7.3	0
www.parkinsons.org.uk	60.75	1	6.1	0.3	0	0	0	6.45	7	4	21.45	1
www.patient.co.uk	95.45	0.85	7.15	0.05	0	0.05	0	24	6.9	4.05	13.55	1
www.pdrhealth.com	32.3	2	1.2	1.1	0	0	0	10.85	4.05	8.9	12.95	1.6
www.riteaid.com	5.4	0	0	1.5	0	0	0	3.15	3.05	3.35	2.05	1
www.walgreens.com	180.2	0.95	6.85	0.1	0	0	0	14.25	10.95	5	19.95	1
www.canada.com	173.25	5	4.55	12.35	1.2	0	0	16.65	5.55	4.9	23.85	1
www.drug3k.com	98.9	0	0	0	0	0	0	112	4	5	3.1	0.95

ตารางที่ 23 แสดงปริมาณ Tag เฉลี่ยต่อหน้าเว็บเพจที่ปรากฏบนเอกสาร HTML ของเว็บไซต์กลุ่ม B (ต่อ)

เว็บไซต์กลุ่ม B	Tag HTML ที่ศึกษา											
	a	h1	h2	h3	h4	h5	h6	img	link	meta	p	title
www.grubbscare.com	6.4	0.1	0	0	0	0	0	0.15	0.05	0.1	12.9	1
www.mdonweb.org	14	0	0	0	0	0	0	8.4	1.95	5.85	9.6	1
www.medindia.net	228.55	1.25	3.05	0.05	0.25	0.05	0	30.1	3.9	4	53	1
ค่าเฉลี่ยต่อหน้าเว็บเพจ	88.73	0.95	3.08	1.19	0.24	0.07	0.00	26.71	4.35	4.10	15.09	0.97
จำนวนเว็บเพจที่พบ Tag HTML เทียบกับจำนวนเว็บเพจทั้งหมดที่ศึกษา (เปอร์เซ็นต์)	93.67	52.33	58.33	36.67	21.33	7.00	0.00	93.67	93.67	93.67	100.00	93.00

จากตารางที่ 23 พบว่าเว็บไซต์อันดับท้ายมีปริมาณเฉลี่ยของ Tag a มากที่สุด รองมาได้แก่ Tag img และ Tag p ตามลำดับ เว็บไซต์ที่มีปริมาณ Tag a มากที่สุด ได้แก่ www.medindia.net พบว่าค่าเฉลี่ยของ Tag a เท่ากับ 228.55 และเมื่อดูจากจำนวนเพจทั้งหมดที่พบเทียบกับจำนวนเพจที่ศึกษาจะทราบว่ามีการใช้ Tag p ในทุกหน้าเว็บเพจที่ทำการศึกษา Tag h1 พบว่ามี 5 เว็บไซต์ที่ไม่มีการใช้ Tag ดังกล่าว และไม่พบการใช้ Tag h6 เลยในเว็บไซต์กลุ่ม B ในส่วนของ Tag img มีค่าเฉลี่ยของ Tag นี้ในเว็บไซต์กลุ่ม B เท่ากับ 26.71, ในส่วนของ Tag p ซึ่งเป็นอีกตัวแปรที่มีปริมาณ Tag เฉลี่ยมาก มีค่าเฉลี่ยปริมาณ Tag ในเว็บไซต์กลุ่ม B เท่ากับ 15.09 จากการศึกษาอาจกล่าวได้ว่าเว็บไซต์กลุ่ม B ส่วนใหญ่จะประกอบไปด้วย Tag HTML ที่ใช้มากอยู่ด้วยกัน 3 Tag เช่นเดียวกันกับเว็บไซต์อันดับต้นคือ Tag a, Tag img และ Tag p และไม่มีการใช้ Tag h6 เลย

ตารางที่ 24 แสดงปริมาณจุดเชื่อมโยงภายนอกและจุดเชื่อมโยงภายในของเว็บไซต์กลุ่ม B

เว็บไซต์กลุ่ม B	คุณลักษณะ	
	จุดเชื่อมโยงภายนอก	จุดเชื่อมโยงภายใน
www.cancer.gov	15.35	57.4
www.druginformationonline.org	52.35	0
www.medical-clinic.org	5.05	0
www.muschealth.com	49	0
www.myoptumhealth.com	154.25	0
www.parkinsons.org.uk	4.45	54.55
www.patient.co.uk	24.2	67.75
www.pdrhealth.com	2.15	14.6
www.riteaid.com	2.4	2.9
www.walgreens.com	20.1	157.55
www.canada.com	39.1	11.3
www.drug3k.com	39.45	52.5
www.grubbscare.com	2.25	0.95
www.mdonweb.org	1.3	0
www.medindia.net	196.05	2.75
ค่าเฉลี่ยต่อหน้าเว็บเพจ	40.50	28.15
จำนวนเว็บเพจที่พบการเชื่อมโยงเทียบกับจำนวนเว็บเพจทั้งหมดที่ศึกษา (เปอร์เซ็นต์)	87.33	47.33

จากตารางที่ 24 พบว่าส่วนของจุดเชื่อมโยงภายนอกและจุดเชื่อมโยงภายใน นั้นมีการใช้การเชื่อมโยงภายนอกมากกว่าการเชื่อมโยงภายใน นอกจากนี้ยังพบว่าในเว็บไซต์กลุ่ม B มี 5 เว็บไซต์ที่ไม่มีการใช้การเชื่อมโยงภายในเลยซึ่งอาจทำให้การเข้าถึงข้อมูลต่างๆ ภายในเว็บไซต์นั้นทำได้ยาก

ตารางที่ 25 แสดงขนาดของหน้าเว็บเพจเฉลี่ยในแต่ละเว็บไซต์กลุ่ม B, เวลาที่ใช้ในการดาวน์โหลด (Connection Rate 1.44 Mbps.), ขนาดภาพที่พบภายในเว็บเพจเฉลี่ย, ขนาดของ CSS ที่พบภายในเว็บเพจ, ภาพที่ไม่สามารถเรียกดูได้ในเว็บไซต์กลุ่ม B (Dead Images) และจำนวนภาพรวมภายในเว็บไซต์กลุ่ม B

เว็บไซต์กลุ่ม B	ขนาดของหน้าเอกสารHTML (กิโลไบต์)	เวลาที่ใช้ในการดาวน์โหลด (วินาที)	ขนาดของภาพ (กิโลไบต์)	ขนาดของ CSS (กิโลไบต์)	รูปภาพที่ไม่สามารถเข้าถึงได้	จำนวนรูปภาพรวมเฉลี่ยต่อหน้าเว็บเพจ
www.cancer.gov	161.66	3.35	19.38	13.90	0	19.1
www.druginformationonline.org	29.15	2.98	0.76	0.18	0	0.15
www.medical-clinic.org	50.14	2.6	15.77	0	0	2.6
www.muschealth.com	161.51	2.42	40.59	2.18	0	7.25
www.myoptumhealth.com	102.18	3.15	30.52	0	0	5.05
www.parkinsons.org.uk	299.45	2.8	53.88	6.35	0	5.15
www.patient.co.uk	42.32	3.81	21.46	6.56	0.05	24.85
www.pdrhealth.com	388.05	3.1	29.06	20.24	0	7.95
www.riteaid.com	23.98	2.02	6.48	1.70	0	2.3
www.walgreens.com	98.11	2.7	52.55	18.27	0.05	9.25
www.canada.com	239.06	3.63	48.29	6.30	0	12.95
www.drug3k.com	164.80	5.44	53.01	37.37	0.1	12.75
www.grubbscare.com	10.75	1.88	2.01	0	0	0.1

ตารางที่ 25 แสดงขนาดของหน้าเว็บเพจเฉลี่ยในแต่ละเว็บไซต์กลุ่ม B, เวลาที่ใช้ในการดาวน์โหลด (Connection Rate 1.44 Mbps.), ขนาดภาพที่พบภายในเว็บเพจเฉลี่ย, ขนาดของ CSS ที่พบภายในเว็บเพจ, ภาพที่ไม่สามารถเรียกดูได้ในเว็บไซต์กลุ่ม B (Dead Images) และจำนวนภาพรวมภายในเว็บไซต์กลุ่ม B (ต่อ)

เว็บไซต์กลุ่ม B	ขนาดของหน้าเอกสารHTML (กิโลไบต์)	เวลาที่ใช้ในการดาวน์โหลด (วินาที)	ขนาดของภาพ (กิโลไบต์)	ขนาดของ CSS (กิโลไบต์)	รูปภาพที่ไม่สามารถเข้าถึงได้	จำนวนรูปภาพรวมเฉลี่ยต่อหน้าเว็บเพจ
www.mdonweb.org	29.58	2.32	1.95	3.52	0	0.4
www.medindia.net	93.16	4.4	18.81	23.13	0.95	14.45
ค่าเฉลี่ยต่อหน้าเว็บเพจ	126.26	3.11	26.30	9.31	0.08	8.29
รวมทั้งหมด	1893.90	46.6	394.53	139.69	1.15	124.3

จากตารางที่ 25 พบว่าเว็บไซต์กลุ่ม B จะมีขนาดของหน้าเว็บเพจเฉลี่ยเท่ากับ 126.26 กิโลไบต์และเวลาที่ใช้ในการดาวน์โหลดหน้าเว็บเพจเฉลี่ย เท่ากับ 3.11 วินาที

ในส่วนของขนาดภาพที่พบภายในเว็บเพจรวมเฉลี่ย พบว่าขนาดของภาพที่พบภายในเว็บเพจรวมเฉลี่ยของเว็บไซต์กลุ่ม B มีค่าเท่ากับ 26.3 กิโลไบต์พบว่า www.parkinsons.org.uk มีขนาดของภาพที่พบเฉลี่ยสูงสุดคือ 53.89 กิโลไบต์

การศึกษาขนาดของ CSS ที่พบภายในเว็บเพจเฉลี่ยโดยรวมของเว็บไซต์กลุ่ม B พบว่ามี 3 เว็บไซต์ที่ไม่มีการใช้ CSS ได้แก่ www.medical-clinic.org, www.myoptumhealth.com และ www.grubbscare.com นอกจากนี้พบว่าค่าเฉลี่ยของขนาด CSS ที่พบภายในเว็บเพจกลุ่ม B เท่ากับ 139.7 กิโลไบต์

จากการศึกษาภาพที่ไม่สามารถเรียกดูได้ภายในเว็บไซต์กลุ่ม B พบว่ามีทั้งหมด 4 เว็บไซต์ที่พบภาพที่ไม่สามารถเรียกดูได้ในเว็บไซต์จาก 15 เว็บไซต์ นอกจากนี้พบว่าแต่ละเว็บไซต์จะมีจำนวนรูปภาพต่อหน้าเว็บเพจเฉลี่ยที่แตกต่างกันออกไป เช่นเดียวกันกับเว็บไซต์กลุ่มอื่นๆ แต่จำนวนของรูปภาพจะน้อยกว่ากลุ่มอื่นๆมีค่าเฉลี่ยอยู่ที่ 8.29 รูปต่อหน้าเว็บเพจ

ตารางที่ 26 แสดงประเภทของเอกสาร HTML และการเข้ารหัสภาษาของเว็บไซต์กลุ่ม B

เว็บไซต์กลุ่ม B	การเข้ารหัสภาษา			ประเภทของเอกสาร HTML	
	utf-8	iso-8859-1	windows-1252	html	xhtml
www.cancer.gov	0	20	0	20	0
www.druginformationonline.org	0	20	0	0	20
www.medical-clinic.org	0	20	0	0	20
www.muschealth.com	20	0	0	1	19
www.myoptumhealth.com	20	0	0	20	0
www.parkinsons.org.uk	20	0	0	0	20
www.patient.co.uk	0	20	0	0	20
www.pdrhealth.com	0	20	0	0	20
www.riteaid.com	20	0	0	0	20
www.walgreens.com	20	0	0	0	20
www.canada.com	1	0	19	19	1
www.drug3k.com	20	0	0	0	20
www.grubbscare.com	20	0	0	19	1
www.mdonweb.org	0	1	19	20	0
www.medindia.net	0	0	20	20	0
ค่าเฉลี่ยต่อหน้าเว็บเพจ	9.40	6.73	3.87	7.93	12.07
รวมทั้งหมด	141	101	58	119	181

จากตารางพบว่า เว็บไซต์กลุ่ม B มีการใช้ encoding กระจายทั้ง 3 รูปแบบ โดยพบการใช้รูปแบบ utf-8 มากสุด และรูปแบบ windows-1252 น้อยสุด ในส่วนของประเภทของภาษา พบว่าเว็บไซต์กลุ่ม B มีการใช้ประเภทของเอกสาร HTML แบบ xhtml มากที่สุดเช่นเดียวกันกับเว็บไซต์กลุ่มอื่น

ตารางที่ 27 แสดง HTML error, HTML warning และ CSS error ของเว็บไซต์กลุ่ม B

URL เว็บไซต์กลุ่ม B	HTML error	HTML warning	CSS error	Dead Links	Total Links
www.cancer.gov	80.75	32.45	9	0	140
www.druginformationonline.org	45.35	2	0.95	0	66.55
www.medical-clinic.org	123.25	0	0	0	131.1
www.muschealth.com	59.35	3.6	13.7	0	65.05
www.myoptumhealth.com	147.65	16.55	232	0	177.7
www.parkinsons.org.uk	6.05	0	6	0	71.7
www.patient.co.uk	104.85	44.85	10	0.05	125.6
www.pdrhealth.com	286.45	106.15	14.05	1	46.3
www.riteaid.com	3.05	0.4	0.95	0	10.2
www.walgreens.com	84.9	72.6	26	0.15	181.9
www.canada.com	107.95	22.25	63.6	0	187.35
www.drug3k.com	1301.65	105.15	8.95	0.1	208.6
www.grubbscare.com	26.4	4.2	0.2	0	6.5
www.mdonweb.org	77	6.3	10.55	0.05	11.4
www.medindia.net	101.65	39.95	13.35	0.95	287.3
ค่าเฉลี่ยต่อหน้าเว็บเพจ	170.42	30.43	27.29	0.15	114.48
รวมทั้งหมด	2556.3	456.45	409.3	2.3	1717.25
จำนวนเว็บเพจที่พบ Tag HTML เทียบกับจำนวนเว็บเพจทั้งหมดที่ศึกษา (เปอร์เซ็นต์)	100.00	80.67	86.33	-	-

จากตารางแสดงผลการศึกษาค่าความผิดพลาดของเอกสาร HTML ที่พบในเว็บไซต์ที่ให้ข้อมูลยาที่เป็นภาษาอังกฤษ ทั้งในส่วนของ HTML error, HTML warning และ CSS error พบว่าเว็บไซต์กลุ่ม B มีความผิดพลาดของเอกสาร HTML แบบ error ในทุกหน้าเว็บเพจที่ทำการศึกษา ส่วน HTML warning มี 2 เว็บไซต์ที่ไม่พบความผิดพลาด ส่วนความผิดพลาดที่เป็น CSS error พบว่ามี 1 เว็บไซต์ที่ไม่พบความผิดพลาดดังกล่าว

ในส่วนของการศึกษาการเชื่อมโยงที่ไม่สามารถเข้าถึงได้ในเว็บไซต์ พบว่ามีเว็บไซต์ 6 เว็บไซต์จาก 15 เว็บไซต์ที่ไม่พบการเชื่อมโยงที่ไม่สามารถเข้าถึง คิดเป็น 40 เปอร์เซ็นต์ของเว็บไซต์กลุ่ม B

ตารางที่ 28 แสดงค่า Page Rank เฉลี่ยของแต่ละเว็บไซต์กลุ่ม B

URL เว็บไซต์กลุ่ม B	Page Rank
www.cancer.gov	4.2
www.druginformationonline.org	1.4
www.medical-clinic.org	0.3
www.muschealth.com	0.3
www.myoptumhealth.com	0.4
www.parkinsons.org.uk	2.85
www.patient.co.uk	2.2
www.pdrhealth.com	0.3
www.riteaid.com	0.35
www.walgreens.com	1.15
www.canada.com	1.7
www.drug3k.com	1.75
www.grubbscare.com	0.15
www.mdonweb.org	0.15
www.medindia.net	0.9
ค่าเฉลี่ยต่อหน้าเว็บเพจ	1.21
รวมทั้งหมด	18.1

จากตารางพบว่าเว็บไซต์ที่อยู่ในกลุ่ม B มีค่าเฉลี่ย Page Rank 1.21 โดยพบว่า www.cancer.gov มีค่า Page Rank สูงสุดคือ 4.2 และค่าที่ต่ำที่สุดคือ 0.15 พบด้วยกัน 2 เว็บไซต์ ได้แก่ www.grubbscare.com และ www.mdonweb.org

4.6 การเปรียบเทียบค่าเฉลี่ยของตัวแปรด้วยวิธี Independent t-test เพื่อดูความสัมพันธ์ของตัวแปรต่างๆ กับคุณภาพของเว็บไซต์และการจัดอันดับของเว็บไซต์บนหน้าแสดงผล โดยมีตัวแปรที่ทำการศึกษาทั้งหมด 31 ตัวแปรได้แก่

Tag HTML จำนวน 12 Tag ได้แก่ a, h1, h2, h3, h4, h5, h6, img, link, meta, p, title

คุณลักษณะต่างๆ จำนวน 11 คุณลักษณะ ได้แก่ จุดเชื่อมโยงภายนอก, จุดเชื่อมโยงภายใน, ขนาดของเอกสาร HTML (ไบต์), เวลาที่ใช้ในการดาวน์โหลด (วินาที), Page Rank, ขนาดของรูปภาพ (ไบต์), ขนาดของ CSS (ไบต์), รูปภาพที่ไม่สามารถเข้าถึงได้, จำนวนรูปภาพรวมเฉลี่ยต่อหน้าเว็บเพจ, จำนวนรูปภาพที่ไม่สามารถเข้าถึงได้, จำนวนการเชื่อมต่อ

ค่าความผิดพลาดของเอกสาร HTML 3 ชนิด ได้แก่ HTML error, HTML warning, CSS error

ประเภทของเอกสาร HTML 2 รูปแบบ ได้แก่ html, xhtml และการเข้ารหัสภาษา 3 รูปแบบได้แก่ utf-8, iso-8859-1, window-1252

4.6.1 การศึกษาความสัมพันธ์ระหว่างตัวแปรต่างๆ ที่ทำการศึกษากับการจัดอันดับของเว็บไซต์บนหน้าแสดงผล โดยศึกษาตัวแปรต่างๆ ทั้ง 31 ตัวแปรดังกล่าวไว้ในหัวข้อ 4.6 ที่มีผลต่อการจัดอันดับของเว็บไซต์ที่ให้ข้อมูลยาที่เป็นภาษาอังกฤษบนหน้าหน้าแสดงผล ได้แบ่งกลุ่มตัวอย่างที่ทำการศึกษาออกเป็น 2 กลุ่มคือกลุ่มเว็บไซต์อันดับต้น จำนวน25 เว็บไซต์ และกลุ่มเว็บไซต์อันดับท้าย จำนวน 10 เว็บไซต์ โดยอ้างอิงกลุ่มของเว็บไซต์ตามหัวข้อ 3.2.1 ได้ผลการศึกษาดังตารางที่ 29 ดังนี้

จากตารางแสดงค่าสถิติ t, องศาแห่งความเป็นอิสระ (degree of Freedom: df), ค่าระดับนัยสำคัญที่ขอบเขตวิกฤติแบบสองทาง Sig. (2-tailed) และ ค่าต่ำสุด-สูงสุดที่ระดับความเชื่อมั่น 95 เปอร์เซ็นต์ (95% Confidence Interval of the Difference) ของตัวแปรต่างๆ ต่อการจัดอันดับของเว็บไซต์บนหน้าแสดงผล พบว่ามีตัวแปรที่ส่งผลให้เว็บไซต์ทั้งกลุ่ม 2 กลุ่มแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ p-values < 0.05 ได้แก่ HTML link Tag, จุดเชื่อมโยงภายในและ Page Rank ที่อยู่ในส่วนของคุณลักษณะต่างๆ รวมถึงการเข้ารหัสภาษาแบบ iso-8859-1 ที่ส่งผลให้กลุ่มเว็บไซต์มีอันดับการปรากฏบนหน้าแสดงผลที่แตกต่างกัน พบว่าการเข้ารหัสภาษาในเว็บไซต์อันดับต้นจะพบการเข้ารหัสภาษาแบบ iso-8859-1 มากกว่าเว็บไซต์อันดับท้าย และเว็บไซต์ที่อยู่ในอันดับต้น จะมีค่า

Page Rank ที่มากกว่าเว็บไซต์ที่อยู่อันดับท้าย นอกจากนี้ในส่วนของ HTML Tag พบว่าเว็บไซต์อันดับต้น จะมีการใช้ HTML link Tag มากกว่าด้วยเช่นกัน

ตารางที่ 29 แสดงค่าสถิติ Independent T-test ระหว่างตัวแปรต่างๆ ที่ทำการศึกษากับการจัดอันดับของเว็บไซต์บนหน้าแสดงผล

t-test for Equality of Means					
ตัวแปร	T	df	Sig.(2-tailed)	95% Confidence Interval of the Difference	
				Lower	Upper
a	0.65	33	0.52	-41.78	83.46
h1	-1.07	9.28	0.31	-5.15	1.84
h2	-0.94	9.92	0.37	-12.78	5.19
h3	-0.28	33	0.78	-6.82	5.18
h4	0.08	33	0.94	-4.18	4.63
h5	0.65	33	0.52	-0.9	2.76
h6	0.63	33	0.53	-0.31	0.93
img	0.46	33	0.65	-25.05	47.04
link	2.35	33	0.02*	0.93	4.87
meta	1.68	33	0.1	0.02	4.27
p	-0.76	33	0.45	-88.75	45.62
title	-0.77	33	0.45	-0.35	14
จุดเชื่อมโยงภายนอก	-0.25	33	0.8	-61.75	48.31
จุดเชื่อมโยงภายใน	2.49	32.84	0.02*	7.02	69.78
ขนาดของหน้าเอกสาร HTML(ไบต์)	-0.11	33	0.91	-15280	137731
เวลาที่ใช้ในการดาวน์โหลด (วินาที)	-1.9	33	0.07	-3.6	0.32
HTML error	-0.67	9.58	0.52	-365.99	196.72
HTML warning	0.19	33	0.85	-23.78	29.03
CSS error	0.65	33	0.52	-39.17	89.45
Page Rank	2.32	33	0.03*	0.31	2.57
ขนาดของรูปภาพ	0.79	33	0.44	-16417.4	37237.6
หมายเหตุ : * p-values <0.05					

ตารางที่ 29 แสดงค่าสถิติ Independent T-test ระหว่างตัวแปรต่างๆ ที่ทำการศึกษากับการจัดอันดับของเว็บไซต์บนหน้าแสดงผล (ต่อ)

t-test for Equality of Means					
ตัวแปร	T	df	Sig.(2-tailed)	95% Confidence Interval of the Difference	
				Lower	Upper
ขนาดของ CSS (ไบต์)	0.03	33	0.98	-21567.7	222.88.83
จำนวนรูปภาพที่ไม่สามารถเข้าถึงได้	-1.27	10.92	0.23	-0.45	0.12
จำนวนรูปภาพรวมเฉลี่ยต่อหน้าเว็บเพจ	0.92	33	0.36	-3.24	17.86
จำนวนการเชื่อมต่อที่ไม่สามารถเข้าถึงได้	-0.75	33	0.46	-0.45	0.21
จำนวนการเชื่อมต่อ	0.77	33	0.45	-46.03	112.53
utf-8	0.12	33	0.91	-6.73	7.53
iso-8859-1	2.94	24.25	0.01*	1.61	9.19
window-1252	-1.96	9	0.08	-12.48	0.88
html	-1.76	33	0.09	-13.57	1.25
xhtml	1.76	33	0.09	-1.25	13.57
หมายเหตุ : * p-values <0.05					

4.6.2 การศึกษาตัวแปรที่มีผลต่อคุณภาพของเว็บไซต์ ซึ่งจะศึกษาผลของตัวแปรต่างๆ ดังหัวข้อ 4.6 ทั้ง 31 ตัวแปร โดยแบ่งกลุ่มตัวอย่างที่ทำการศึกษาออกเป็นเว็บไซต์กลุ่ม A (เว็บไซต์ที่มีคะแนนการประเมินคุณภาพเฉลี่ยตั้งแต่ 60 เปอร์เซ็นต์ขึ้นไป) จำนวน 20 เว็บไซต์ และเว็บไซต์กลุ่ม B (เว็บไซต์ที่มีคะแนนการประเมินคุณภาพเฉลี่ยต่ำกว่า 60 เปอร์เซ็นต์) จำนวน 15 เว็บไซต์ผลการศึกษาปรากฏดังนี้

จากตารางที่ 30 แสดงค่าสถิติ t, องศาแห่งความเป็นอิสระ (degree of Freedom: df), ค่าระดับนัยสำคัญที่ขอบเขตวิกฤติแบบสองทาง Sig.(2-tailed) และ ค่าต่ำสุด-สูงสุดที่ระดับความเชื่อมั่น 95 เปอร์เซ็นต์ (95% Confidence Interval of the Difference) ของตัวแปรต่างๆ ต่อคุณภาพของเว็บไซต์ พบว่ามีตัวแปรที่ส่งผลให้เว็บไซต์ที่ให้ข้อมูลด้านยาที่เป็นภาษาอังกฤษทั้งกลุ่ม 2 กลุ่มแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ p-values < 0.05 ได้แก่ Tag HTML จำนวน 3 Tag คือ Tag h3, h4 และ p ในส่วนของคุณลักษณะต่างๆ พบว่า Page Rank, ขนาดภาพที่พบ และ ขนาด CSS รวมถึงตัวแปร utf-8 ที่ส่งผลให้เว็บไซต์กลุ่ม A และเว็บไซต์กลุ่ม B มีคุณภาพแตกต่างกันอย่างมีนัยสำคัญทางสถิติ ซึ่งพบว่าปริมาณของ Tag p ซึ่งเป็น Tag ที่ใช้ในการจัดพารากราฟของข้อความในเว็บไซต์กลุ่ม A มีปริมาณมากกว่าเว็บไซต์กลุ่ม B มาก เช่นเดียวกับ HTML h3 และ h4 Tag ที่เป็นส่วนของหัวเรื่องซึ่งให้ผลสอดคล้องกับการศึกษาจำนวนเพจทั้งหมดที่พบเทียบกับจำนวนเพจที่ศึกษา โดยพบว่า Tag ทั้ง 2 มี ค่าเฉลี่ยในเว็บไซต์กลุ่ม A มากกว่าเว็บไซต์กลุ่ม B ถึง 50.08เปอร์เซ็นต์ และ 42.67 เปอร์เซ็นต์ ตามลำดับ ในส่วนของขนาดรูปภาพที่บอกขนาดของภาพเฉลี่ยที่พบในเว็บไซต์ที่ให้ข้อมูลด้านยาและขนาด CSS ซึ่งเป็นตัวที่กำหนดคุณลักษณะของตัวอักษรหรือการจัดการหน้าเว็บเพจนั้นพบว่าในเว็บไซต์กลุ่ม A จะมีขนาดของข้อมูลมากกว่าเว็บไซต์กลุ่ม B และส่วนของตัวแปร utf-8 ซึ่งเป็นส่วนของการเข้ารหัสภาษานั้นพบว่าเว็บไซต์กลุ่ม A มีปริมาณการเข้ารหัสภาษาแบบ utf-8 มากกว่ารูปแบบ iso-8859-1และไม่พบการใช้รูปแบบ window-1252 เลย

ตารางที่ 30 แสดงค่าสถิติ Independent T-test ของตัวแปรต่างๆ ต่อคุณภาพของเว็บไซต์

t-test for Equality of Means					
ตัวแปร	t	df	Sig.(2-tailed)	95% Confidence Interval of the Difference	
				Lower	Upper
a	1.83	33	0.08	-3.79	107.11
h1	1.56	33	0.13	-0.24	3.11
h2	1.66	24.76	0.11	-0.91	8.45
h3	3.65	25.45	0*	3.27	11.71
h4	2.33	19.09	0.03*	0.53	9.9
h5	1.03	33	0.31	-0.93	3.6
h6	0.89	33	0.38	-0.38	1.17
img	0.59	33	0.56	-27.03	52.63
link	1.31	33	0.2	-0.72	3.84
meta	0.52	33	0.6	-1.6	2.87
p	2.55	19.85	0.02*	9.81	97.89
title	0.65	33	0.52	-0.16	0.32
จุดเชื่อมโยงภายนอก	1.14	33	0.26	-19.25	73.69
จุดเชื่อมโยงภายใน	1.8	33	0.08	-2.25	74.33
ขนาดของหน้าเอกสาร HTML(ไบต์)	1.8	33	0.08	-5230.62	203021
เวลาที่ใช้ในการดาวน์โหลด (วินาที)	1.32	23.61	0.2	-0.54	2.44
HTML error	-0.82	33	0.42	-239.7	112.55
HTML warning	0.13	33	0.89	-24.33	27.79
CSS error	1.58	33	0.12	-9.49	117.76
Page Rank	2.72	32.03	0.01*	0.36	2.49
ขนาดของรูปภาพ	2.44	28.88	0.02*	4100.69	46444.03
หมายเหตุ : * p-values <0.05					

ตารางที่ 30 แสดงค่าสถิติ Independent T-test ของตัวแปรต่างๆ ต่อคุณภาพของเว็บไซต์ (ต่อ)

t-test for Equality of Means					
ตัวแปร	t	df	Sig.(2-tailed)	95% Confidence Interval of the Difference	
				Lower	Upper
ขนาดของ CSS (ไบต์)	2.49	22.84	0.02*	3944.67	42323.04
จำนวนรูปภาพที่ไม่สามารถเข้าถึงได้	0.27	33	0.79	-0.16	0.21
จำนวนรูปภาพรวมเฉลี่ยต่อหน้าเว็บเพจ	1.24	33	0.22	-3.92	21.68
จำนวนการเชื่อมต่อที่ไม่สามารถเข้าถึงได้	1.1	33	0.28	-0.13	0.42
จำนวนการเชื่อมต่อ	1.86	33	0.07	-2.28	143.54
utf-8	2.86	21.39	0.01*	2.35	14.75
iso-8859-1	-1.64	22.19	0.12	-10.61	1.24
window-1252	-1.87	14	0.08	-8.3	0.57
html	-0.28	33	0.78	-7.76	5.89
xhtml	0.28	33	0.78	-5.89	7.76
หมายเหตุ : * p-values <0.05					

บทที่ 5

สรุป อภิปรายผลการศึกษาและข้อเสนอแนะ

จากงานวิจัยศึกษาคุณลักษณะโครงสร้างของเอกสาร HTML ที่เป็นปัจจัยต่อคุณภาพของเว็บไซต์ที่ให้ข้อมูลยาที่เป็นภาษาอังกฤษ และมีผลต่อการปรากฏในโปรแกรมค้นหา ได้ผลโดยสรุปดังนี้

5.1 สรุปและอภิปรายผลการศึกษา

5.1.1 สามารถสรุปลักษณะโครงสร้างของเว็บไซต์แต่ละประเภทได้ดังต่อไปนี้คือ

เว็บไซต์อันดับต้น จะประกอบไปด้วยแท็ก a p และ img มากสุดตามลำดับ ส่วนแท็กที่พบน้อยสุดได้แก่แท็ก h6 พบว่ามีแท็ก img และ link ในทุกๆ เพจของเว็บไซต์อันดับต้นที่ทำการศึกษามีการใช้ external link มากกว่า internal link และพบว่าค่าเวลาเฉลี่ยในการดาวน์โหลดหน้าเว็บไซต์อยู่ที่ 3.2 วินาที รูปแบบของรหัสการเข้าภาษาพบรูปแบบ utf-8 มากที่สุด ไม่พบการใช้รหัสภาษาแบบ windows-1252 พบประเภทเอกสารเป็น xhtml เป็นส่วนมาก และพบว่าค่า Page Rank เฉลี่ยอยู่ที่ 2.4

เว็บไซต์อันดับท้าย จะประกอบไปด้วยแท็ก a p และ img เฉลี่ยมากสุดตามลำดับ ส่วนแท็ก h6 มีปริมาณเฉลี่ยน้อยสุด นอกจากนี้พบแท็ก h1 ในทุกเพจ และพบว่ามีการใช้ external link มากกว่า internal link และยังพบเว็บไซต์ที่ไม่มีการใช้ internal link เลย ค่าเวลาเฉลี่ยในการดาวน์โหลดอยู่ที่ 4.8 วินาที พบรูปแบบการเข้ารหัสภาษาดังนี้คือพบ utf-8 มากกว่า windows-1252 และพบรูปแบบ iso-8859-1 น้อยที่สุด พบว่าประเภทเอกสารเป็น html เป็นส่วนมาก พบ HTML error-warning พบในทุกเว็บเพจที่ทำการศึกษา และพบว่าค่า Page Rank เฉลี่ยอยู่ที่ 0.99

เว็บไซต์กลุ่ม A จะพบแท็ก a p และ img มีปริมาณเฉลี่ยมากสุดตามลำดับ และพบว่าแท็ก h6 มีปริมาณเฉลี่ยน้อยสุด นอกจากนี้ยังพบแท็ก a, img และ link ในทุกเพจ พบว่ามีการใช้ external link ในทุกเว็บเพจ และในส่วนของการเข้ารหัสภาษาและประเภทของเอกสารพบว่าไม่พบการใช้รหัส

ภาษาแบบ windows-1252 ส่วนประเภทเอกสารเป็น xhtml เป็นส่วนมาก เวลาที่ใช้ในการดาวน์โหลดหน้าเพจเฉลี่ยอยู่ที่ 4 วินาที และพบว่าค่า Page Rank เฉลี่ยอยู่ที่ 2.6

เว็บไซต์กลุ่ม B จะพบแท็ก a img และ p มีปริมาณเฉลี่ยมากสุดตามลำดับ และพบว่าไม่มีการใช้แท็ก h6 เลยในทุกเว็บเพจที่ทำการศึกษาแต่พบว่ามีการใช้แท็ก h1 ในทุกเพจ นอกจากนี้ยังพบว่ามีการใช้ external link มากกว่า internal link และพบว่าค่าเวลาเฉลี่ยในการดาวน์โหลดอยู่ที่ 3.1 วินาที พบว่าการเข้ารหัสภาษาจะพบการใช้ utf-8 iso-8859-1 และ windows-1252 มากไปน้อยตามลำดับ และพบประเภทเอกสารเป็น xhtml เป็นส่วนมาก นอกจากการศึกษายังพบว่าพบHTML error พบในทุกเว็บเพจ และค่า Page Rank เฉลี่ยอยู่ที่ 1.21

ดังนั้นจึงกล่าวได้ว่าโครงสร้างของเอกสาร HTML มีผลต่อคุณภาพและการปรากฏในโปรแกรมค้นหา โดยจากการศึกษาปริมาณแท็ก HTML และคุณลักษณะต่างๆของเอกสาร HTML ระหว่างเว็บไซต์ที่อยู่ในอันดับต้น และเว็บไซต์ที่อยู่ในอันดับท้าย ของโปรแกรมการค้นหา และความสัมพันธ์ระหว่างเว็บไซต์ที่มีระดับคุณภาพของเว็บไซต์ต่างๆ กันพบว่า ปัจจัยที่มีผลทำให้อันดับของเว็บไซต์เมื่อทำการค้นหาผ่านโปรแกรมค้นหาต่างกันอย่างมีนัยสำคัญได้แก่แท็ก link, การเชื่อมโยงภายใน, Page Rank และการเข้ารหัสภาษาแบบ iso-8859-1 ส่วนปัจจัยที่ทำให้คุณภาพของเว็บไซต์แตกต่างกันอย่างมีนัยสำคัญได้แก่แท็ก h3, h4, p, Page Rank, ขนาดของรูปภาพ, ขนาดของ CSS และการเข้ารหัสภาษาแบบ utf -8

5.1.2 การพัฒนาคุณภาพและโครงสร้างของเว็บไซต์

จากการศึกษาพบว่าเว็บไซต์ที่มีคะแนนการประเมินคุณภาพตั้งแต่ 60 เปอร์เซ็นต์ขึ้นไป สามารถพบได้ทั้งจากเว็บไซต์ที่อยู่ในอันดับต้นๆ และเว็บไซต์ที่อยู่ในอันดับท้ายๆ เมื่อทำการค้นหาผ่านโปรแกรมค้นหา จึงทำให้ทราบว่าอันดับของเว็บไซต์ที่ปรากฏบนหน้าแสดงผล ไม่ได้เป็นตัวบอกคุณภาพของเว็บไซต์ ดังนั้นในการพัฒนาเว็บไซต์นอกจากเนื้อหาของข้อมูลต้องมีคุณภาพที่ดีเว็บไซต์ที่ดีควรที่จะอยู่ในอันดับต้นๆ เมื่อทำการค้นหาผ่านโปรแกรมค้นหาด้วย

เว็บไซต์ที่ดีและมีคุณภาพควรประกอบไปด้วย

เนื้อหาของเว็บไซต์ที่ถูกต้องชัดเจน มีประโยชน์สอดคล้องกับคำค้นภายในหน้าเว็บเพจนั้นๆ และควรมีองค์ประกอบอื่นๆ ด้วยเช่น รูปภาพ วีดีโอ มีการอ้างอิงที่ถูกต้อง ระบุวันเวลาการอับ

เดตข้อมูลที่ชัดเจน มีความน่าเชื่อถือทั้งในด้านเนื้อหาและที่มาของเนื้อหา ควรเป็นเว็บไซต์ที่มีเนื้อหาใหม่เสมอ ไม่เป็นข้อความที่คัดลอกมาจากเว็บไซต์อื่นๆ (Amit Singhal and Matt Cutts, 2011) โดยหากเป็นเว็บไซต์ที่ให้ข้อมูลยาควรจะประกอบไปด้วยหัวข้อต่างๆ ดังนี้คือ เว็บไซต์นั้นต้องมีการแสดงชื่อและรายละเอียดในการติดต่อของผู้จัดทำเว็บไซต์ แสดงข้อมูลของผู้สนับสนุน ควรให้ผู้เข้าใช้ทราบว่าข้อมูลนั้นเหมาะกับผู้ใช้ระดับใดผู้ใช้ทั่วไปหรือผู้เชี่ยวชาญ เว็บไซต์ต้องแสดงวันและเวลาที่มีการปรับปรุงข้อมูล มีการแสดงชื่อยาและสาระสำคัญในการออกฤทธิ์ แสดงข้อบ่งใช้ วิธีใช้ยา วิธีการเก็บรักษา ผลข้างเคียง ข้อห้ามใช้และข้อควรระวังต่างๆ นอกจากนี้ควรแสดงรายละเอียดของบริษัทผู้ผลิตหรือผู้จัดจำหน่ายและข้อมูลที่อยู่ที่สามารถติดต่อได้ด้วย ซึ่งเป็นไปตามข้อแนะนำในการสืบค้นข้อมูลยาบนอินเทอร์เน็ตขององค์การอนามัยโลก (World Health Organization, 2009)

โครงสร้างของเว็บไซต์จะต้องประกอบไปด้วยแท็ก HTML ที่ครบถ้วนสมบูรณ์ เช่น alt ซึ่งเป็นส่วนที่อธิบายเกี่ยวกับรูปภาพ บอกให้โปรแกรมค้นหาทราบว่ารูปภาพดังกล่าวคือภาพอะไร นอกจากนี้ยังทำให้ผู้ชมเว็บไซต์ได้ทราบเกี่ยวกับรูปภาพนั้นๆ ด้วยหากเกิดความผิดพลาดในการแสดงผลของภาพนั้นๆ เช่นเดียวกับการตั้งชื่อไฟล์ควรตั้งให้ชัดเจนเช่นกัน เว็บไซต์ที่ดีควรมีการใช้ Title ที่เหมาะสมไม่ยาวเกินไปเนื่องจาก ข้อความที่ยาวเกินจะไม่ถูกแสดงบนหน้าแสดงผลซึ่งจะทำให้ผู้อ่านอ่านไม่รู้เรื่องได้ ในส่วนของการใช้ Heading แท็กนั้นพบว่าควรมีการกำหนด h1-h6 ที่เหมาะสมซึ่ง h1 ควรที่จะใส่ในทุกหน้าเว็บเพจ โดยแต่ละเว็บเพจควรจะมีเพียงแค่ 1 อันเท่านั้น (Matt Cutts, 2009) ส่วน h2-h4 ควรที่จะมีการกำหนดเช่นกันซึ่งควรมีปริมาณของแต่ละแท็กไม่มากเกินไป และจากการศึกษานี้พบว่า h3 และ h4 เป็นปัจจัยหนึ่งที่มีผลต่อคุณภาพของเว็บไซต์ อาจเนื่องมาจากการกำหนดหัวเรื่องสำหรับหน้าเว็บเพจนั้นๆ จะทำให้ผู้ใช้สามารถเข้าใจในเนื้อหาในหน้านั้นได้ง่ายและมีกรอบที่ชัดเจนกว่าการอ่านแบบไม่มีหัวเรื่อง

ในส่วนของ meta แท็กนั้นจากการศึกษาของ Jin Zhang และ Alexandra Dimitroff ในปี 2005 นั้นพบว่าเป็นแท็กที่มีผลต่อการจัดอันดับของโปรแกรมค้นหา แต่ในการศึกษาครั้งนี้แท็ก meta ไม่ได้เป็นตัวแปรที่ทำให้อันดับของเว็บไซต์แตกต่างกันอย่างมีนัยสำคัญซึ่งสอดคล้องกับการให้ความสำคัญของการจัดอันดับของ Google ในปัจจุบันที่ไม่ให้ค่าน้ำหนักในส่วนของ meta keyword ซึ่งเป็นผลมาจากผู้สร้างเว็บไซต์บางกลุ่มได้ทำการใส่คำสำคัญที่ไม่ตรงกับเนื้อหาของข้อมูลหรือมีการใช้คำสำคัญที่มากเกินไปลงไปในส่วนของ meta keyword เป็นการหลอกโปรแกรมค้นหาเพื่อให้เว็บไซต์ของตนปรากฏในอันดับต้นทั้งที่เนื้อหาไม่สอดคล้องกับคำค้นที่ใส่ ทำให้นักพัฒนาโปรแกรมค้นหานำส่วนนี้ออกจากการจัดอันดับ (Matt Cutts, 2009) อย่างไรก็ดีในส่วนของ meta description นั้นจะ

ถูกนำมาแสดงเป็นคำอธิบายเวลาที่ผู้ค้นหาค้นด้วยดังนั้นในส่วนของ meta แท็กก็ยังควรที่ต้องกรอกข้อมูลให้ครบถ้วนเช่นกันเพื่อให้ผู้ค้นหาเข้าใจได้ว่าเว็บเพจนั้นๆ แสดงข้อมูลอะไรที่น่าสนใจบ้าง

จากการศึกษานี้พบว่าแท็ก link ซึ่งเป็นการเชื่อมโยงไปยังหน้าเว็บเพจอื่น หรือเอกสารอื่นๆ นั้นเป็นปัจจัยที่ทำให้อันดับของการแสดงผลบนหน้าแสดงผลแตกต่างกันอย่างมีนัยสำคัญนั่นอาจเป็นผลมาจากหลายปัจจัย เช่น ความน่าเชื่อถือหรือคุณภาพของ link ปลายทาง ยิ่งหากเป็นเว็บไซต์ที่มีความน่าเชื่อถือมากก็จะส่งผลให้เว็บเพจนั้นๆ ได้รับความน่าเชื่อถือไปด้วย แต่หากใส่ปริมาณ link มากเกินไปอาจทำให้อันดับลดลงได้เช่นเดียวกัน นอกจากนี้ในส่วนของ Internal link หรือจุดเชื่อมโยงภายในก็มีผลต่ออันดับของการแสดงผลเช่นเดียวกัน เนื่องจากการเชื่อมโยงภายในเว็บไซต์ทำให้ผู้ใช้สะดวกในการเข้าถึงข้อมูล แต่ควรจะมีในปริมาณที่เหมาะสมไม่มากหรือน้อยเกินไป โดยพบว่าไม่ควรเกิน 100 การเชื่อมต่อต่อหน้าเว็บเพจ และยังพบว่าการมีการเชื่อมต่อหลายๆ อันในหน้าเว็บเพจจะทำให้เราเสียคะแนนการจัดอันดับของโปรแกรมค้นหาไป ดังนั้นเราควรจะแบ่งหมวดหมู่ และจัดโครงสร้างของเว็บไซต์ให้ง่ายต่อการเข้าถึงด้วยการจัดโครงสร้างแบบสถาปัตยกรรมเว็บแนวราบหรือ Flat site Architecture ซึ่งเป็นการออกแบบโครงสร้าง โดยมีวัตถุประสงค์ที่จะให้ผู้ใช้คลิกน้อยครั้งที่สุด ไม่ว่าจะไปเว็บเพจใดก็ตาม และตัดการเชื่อมโยงที่ไม่จำเป็นออก รวมถึงตรวจสอบ คำสั่งการเชื่อมโยงเพื่อไม่ให้มีการเชื่อมโยงที่เข้าถึงไม่ได้ด้วย (Rand Fishkin, 2009)

ในการศึกษานี้พบว่าขนาดภาพที่พบภายในเว็บเพจและอัตราส่วนระหว่างขนาดภาพที่พบในเว็บเพจกับขนาดของเว็บเพจทั้งหมดมีผลต่อคุณภาพของเว็บไซต์ โดยในแต่ละเว็บเพจควรมีจำนวนของรูปภาพที่ไม่มากหรือน้อยเกินไปเพราะหากเป็นเว็บไซต์ที่มีแต่เนื้อหาอย่างเดียวก็จะทำให้เป็นเว็บไซต์ที่ไม่น่าสนใจ และมีการกำหนดขนาดของรูปภาพที่เหมาะสมไม่ใหญ่เกินไปเพราะอาจเป็นผลทำให้เสียเวลาในการดาว์นโหลดหน้าเว็บเพจนานขึ้น หรือหากรูปภาพมีขนาดที่เล็กเกินไปจะทำให้เห็นภาพได้ไม่ชัดเจน จากการศึกษาพบว่า CSS เป็นอีกปัจจัยหนึ่งที่ส่งผลให้คุณภาพของเว็บไซต์แตกต่างกันอาจเป็นผลมาจาก CSS ภายในเว็บเพจจะช่วยทำหน้าที่ในการแสดงผลข้อมูล (การตกแต่ง) ทำให้เนื้อหาและข้อมูลถูกแยกออกจากกัน ช่วยลดระยะเวลาในการดาวน์โหลดหน้าเว็บเพจนั้นๆลงได้ และลดความผิดพลาดที่จะเกิดจากการเขียน code HTML ซึ่งจะทำให้ Robot ในโปรแกรมค้นหาสามารถกรองข้อมูลภายในเว็บเพจได้แม่นยำขึ้น ไม่ต้องเสียเวลาในการคัดแยกโค้ดที่ไม่เกี่ยวข้องออกก่อนซึ่งส่งผลให้อันดับบนหน้าแสดงผลดีขึ้น นอกจากนี้การใช้ CSS ในการปรับปรุงเว็บเพจจะช่วยให้สามารถแก้ไขเว็บเพจได้ง่ายขึ้น และทำให้รูปแบบของเว็บเพจไปในทิศทางเดียวกันดูเป็นระเบียบเรียบร้อยง่ายต่อการอ่าน

รูปแบบของการเข้ารหัสภาษาที่นิยมใช้มากที่สุดได้แก่ utf-8 ซึ่งเป็นรูปแบบที่เป็นมาตรฐานสามารถรองรับการใช้งานภาษาได้หลายภาษา ส่วนของประเภทของเอกสาร HTML พบว่ารูปแบบที่ใช้มากที่สุดได้แก่ xhtml เนื่องจาก เป็นรูปแบบที่ยืดหยุ่นช่วยให้บราวเซอร์รุ่นเก่าๆ สามารถแปลเอกสารได้ใกล้เคียงกัน

ค่า Page Rank เป็นปัจจัยตัวหนึ่งที่ Google ใช้ในการจัดอันดับของการแสดงผลค่าที่ยิ่งสูงจะยิ่งทำให้ได้รับการจัดลำดับผลลัพธ์จากการค้นหาในอันดับที่ดีจาก Google แต่ก็ไม่ได้เป็นเช่นนั้นเสมอไปเพราะยังมีปัจจัยอื่นๆ ด้วย และค่า Page Rank นั้นจะพิจารณาต่อหน้าเว็บเพจเท่านั้นไม่ได้หมายรวมถึงทั้งเว็บไซต์ จากการศึกษาพบว่าในแต่ละเว็บไซต์หน้าที่เป็นหน้าโฮมเพจจะมีค่า Page Rank สูงสุด ส่วนหน้าเว็บเพจอื่นๆ จะมีค่าที่แตกต่างกันออกไป ซึ่งพบว่าเว็บไซต์ที่มีคุณภาพดีจะมีค่า Page Rank เฉลี่ยที่สูงด้วย (Susan, 2011)

5.2 ข้อเสนอแนะ

5.2.1 เนื่องจากในการทดลองนี้ได้ทำการเก็บข้อมูลโดยใช้การคัดแยกด้วยมือเป็นส่วนใหญ่ ดังนั้นอาจมีข้อบกพร่องจากการเก็บข้อมูล รวมถึงทำให้สามารถเก็บตัวอย่างได้ไม่มากดังนั้นในการศึกษาต่อไปคือการพัฒนาในส่วนของเครื่องมือที่ใช้ช่วยในการเก็บข้อมูลเพื่อเพิ่มความสะดวกในการเก็บข้อมูลและลดความผิดพลาดต่างๆ และอาจมีการทดลองซ้ำเพื่อให้ได้ค่าที่ถูกต้องและแม่นยำเพิ่มขึ้น รวมถึงทำการศึกษากับจำนวนกลุ่มทดลองมากขึ้นเพื่อจะได้เห็นความแตกต่างของข้อมูลที่ชัดเจน

5.2.2 ในงานวิจัยนี้เป็นการศึกษาในเว็บไซต์ที่ให้ข้อมูลด้านยา แต่เครื่องมือที่ใช้ในการประเมินคุณภาพของเว็บไซต์เป็นเครื่องมือที่ใช้ประเมินคุณภาพของเว็บไซต์ด้านสุขภาพทั่วไป ในการศึกษาครั้งต่อไปจึงควรเลือกใช้เครื่องมือที่จำเพาะต่อเนื้อหามากยิ่งขึ้น

5.2.3 ข้อมูลบนอินเทอร์เน็ตมีการเปลี่ยนแปลงรวดเร็วและตลอดเวลาดังนั้นการเข้าถึงข้อมูลของเว็บไซต์ในช่วงที่ทำการเก็บข้อมูลกับช่วงเวลาอื่นอาจแตกต่างกัน โครงสร้างและคุณลักษณะต่างๆของเว็บไซต์อาจเปลี่ยนแปลงตามข้อกำหนดใหม่ๆได้ ดังนั้นการศึกษานี้จึงเป็นเพียงแนวทางในการพัฒนาโครงสร้างและคุณลักษณะของเว็บไซต์ที่ให้ข้อมูลด้านยาเบื้องต้นเพื่อเท่านั้น

รายการอ้างอิง

กัลยา วานิชย์บัญชา. (2553). **การวิเคราะห์เชิงปริมาณ.** พิมพ์ครั้งที่ 1. กรุงเทพฯ: ห้างหุ้นส่วนจำกัดสามลดา

ลาวัลย์ ศรัทธาพุทธ. (2555). **คู่มือการใช้ซอฟต์แวร์เสรีทางสถิติ PSPP สำหรับผู้เริ่มต้น.**หนังสือชุดของโครงการสร้างคุณธรรมจริยธรรมเรื่อง “การส่งเสริมการใช้ซอฟต์แวร์โอเพนซอร์ส”. นครปฐม :โรงพิมพ์มหาวิทยาลัยศิลปากร.

สุธี จันทร์แต่งพล. (2551). **SEO ปรับเว็บให้แรง แต่งให้ติด**อันดับ. กรุงเทพฯ: โปรวิชั่น.

เอกสารงานวิจัย เรื่อง “การพัฒนาตัวสืบค้นข้อมูลยาผ่านเครือข่ายอินเทอร์เน็ตสำหรับบุคลากรสาธารณสุขและบุคคลทั่วไป.” SURDI 52/01/23 สถาบันวิจัยและพัฒนา มหาวิทยาลัยศิลปากร

Allison Woodruff , Paul M. Aoki, Eric Brewer, and others. (1996). “An investigation of documents from the World Wide Web.” **Computer Networks and ISDN Systems**. 963-980.

Amit Singhal. (2012). “Search quality highlights: 40 changes for February”. Access 17 May 2013. Available from http://insidesearch.blogspot.com/2012/02/search-quality-highlights-40-changes.html

Amit Singhal and Matt Cutts. (2011). “Finding more high-quality sites in search” Access 17 May 2013. Available from http://googleblog.blogspot.com/2011/02/finding-more-high-quality-sites-in.html

Aysu Betin Can and Nazife Baykal (2007). “MedicoPort: A medical search engine for all journal homepage.” **Computer Methods and Programs in Biomedicine**. 86, 1 (April 2007): 73–86.

Barry Schwartz. (2009). “Site Speed, Google’s Next Ranking Factor” Access 17 January. Available from http://searchengineland.com/site-speed-googles-next-ranking-factor-29793

Bernard. J. Jansen and Amanda Spink. (2006). “How are we searching the World Wide Web? A comparison of nine search engine transaction logs.” Information Processing and Management 42: 248–263.

Charnock, D. and Shepperd, S**.** (2004). “Learning to DISCERN online: applying an appraisal tool to health websites in a workshop setting”. **Health Education Research** 2004 19: 440-446.

Chris Sherman. (2005). “ Metacrawlers and Metasearch Engines.” **Search Engine Watch**. Mar 23. Access 10 December 2009. Available from http://searchenginewatch.com/2156241

CyberGuide. Accessed 1 January 2010 .Available from http://www.cyberbee.com

Crocco, Anthony G., M. Villasis-Keever and others. (2002). “Analysis of cases of harm associated with use of health information on the internet.” **JAMA.** 287,21 (2002) : 2869-71

Cyrus Shepard. (2013). “How to Rank: 25 Step SEO Master Blueprint”. Access 17 May 2013. Available from http://www.seomoz.org/blog/how-to-rank

DISCERN. Accessed 1 January 2010. Available from http://www.discern.org.uk

Deshpande A. and Jadad A.R. (2006). “Healia, a search engine for finding high quality and personalized health information.” The Journal of Men's Health & Gender. 3,4 (December 2006) : 418–419.

Health on the net foundation. Accessed 1 January 2010. Available from http://www.hon.ch

HTML Introduction. Accessed 7 May 2011. Available from http://www.w3schools.com/html/html_intro.asp

Internet Usage and World Population Statistics are for December 31, 2009. [Online]. Accessed 7 May 2010 .Available from http://www.internetworldstats.com/stats.htm

Jin Zhang and Alexandra Dimitroff. (2005a). “The impact of webpage content characteristics on webpage visibility in search engine results (Part I).” **Information Processing and Management.** 41,3 (May 2005) : 665–690

Jin Zhang and Alexandra Dimitroff. (2005b). “The impact of metadata implementation on webpage visibility in search engine results (Part II).” **Information Processing and Management.** 41,3 (May 2005) : 691–715.

Jonathan Hodgson. (2001). “Do HTML Tags Flag Semantic Content?.” **IEEE**: 20-25.

Matt Cutts. (2009a). “More than one H1 on a page: good or bad?”. GoogleWebmasterHelp. Access 17 May 2013. Available from http://www.youtube.com/watch?v=GIn5qJKU8VM

Matt Cutts. (2009b). “Google doesn’t use the keywords meta tag in web search”. Access 17 May 2013. Available from http://googlewebmastercentral.blogspot.com/2009/09/google-does-not-use-keywords-meta-tag.html

Matt Cutts. (2009c). “How many links per page?”. Access 17 May 2013. Available from http://www.mattcutts.com/blog/how-many-links-per-page/

Matt McGee. (2011). “It’s Official: Google Now Counts Site Speed As A Ranking Factor” Access 13 January 2013. Available from

http://searchengineland.com/google-now-counts-site-speed-as-ranking-factor-39708

Mei Kobayashi and Koichi Takeda. (2000). “Information retrieval on the Web.” **ACM Computing Surveys (CSUR)**. 32,2 (June 2000) : 144–173.

Morel, Vincent. And others. (2008). “Quality of web-base information on bipolar disorder.” Journal of Affective Disorder. 110 (October 2008): 265-269

Rand Fishkin. (2009). “Whiteboard Friday-Using Trust Rank to Guide Your Link Building”. Access 17 May 2013. Available from

Ron Jones. (2008a). “SEO Site Structure 101, Part 1.” **Search Engine Watch**. [Online]. Access 10 December 2009. Available from http://searchenginewatch.com/3632183

Ron Jones. (2008b). “SEO Site Structure 101, Part 2.” Search Engine Watch. [Online]. Access 10 December 2009. Available from http://searchenginewatch.com/3632212

Sukanya Wiboonrangsan and Suang Roongprakaypan. (2010). “Quality of internet websites providing drug information in English.” **SURF**. January 2010.

Sutherland L. A., Wildemuth B., Campbell M. K., and others. (2005). “Unraveling the Web: An Evaluation of the Content Quality, Usability, and Readability of Nutrition Web Sites.” **Journal of Nutrition Education and Behavior**. 36,6 (November-December 2005) : 300-305.

Susan Moskwa. (2011). “Beyond PageRank: Graduating to actionable metrics”. Access 17 May 2013. Available from http://googlewebmastercentral.blogspot.com/2011/06/beyond-pagerank-graduating-to.html

Theodosiou, Louise and Jonathan Green. (2003). “Emerging challenges in using health information from the internet.” **Advances in Psychiatric Treatment**. 9 (2003) : 378-396.

Timothy C. Craven. (2004). “Variations in use of meta tag descriptions by Web pages in different languages.” **Journal of Information Science**. 30,3 (June 2004) : 268-279.

Webmaster Tools. (2013). “Webmaster Guidelines”. Access 17 May 2013. Available from http://support.google.com/webmasters/bin/answer.py?hl=en&answer=35769

ภาคผนวก

ตารางที่ 31 แบบประเมินข้อมูล DISCERN

Item of DISCERN	Score				
	1	2	3	4	5
Section 1 : IS THE PUBLICATION RELIABLE?					
1.Are the aims clear					
2.Does it achieve its aims					
3.Is it relevant					
4.Is it clear what sources of information were use to compile the publication					
5.Is it clear when the information used or reported in the publication was produce					
6.Is it balance and unbiased					
7.Does it provide details of additional sources of support and information					
8.Does it refer to areas of uncertainty					
Section 2 : HOW GOOD IS THE QUALITY OF INFORMATION ON TREATMENT CHOICES?					
9. Does it describe how each treatment works?					
10. Does it describe the benefits of each treatment?					
11. Does it describe the risks of each treatment?					
12. Does it describe what would happen if no treatment is used?					
13. Does it describe how the treatment choices affect overall quality of life?					
14. Is it clear that there may be more than one possible treatment choice?					
15. 15 Does it provide support for shared decision-making?					

ตารางที่ 31 แบบประเมินข้อมูล DISCERN (ต่อ)

Item of DISCERN	Score				
	1	2	3	4	5
Section 3 : OVERALL RATING OF THE PUBLICATION					
16. Based on the answers to all of the above questions, rate the overall quality of the publication as a source of information about treatment choices					

รายละเอียดของแต่ละหัวข้อของแบบประเมิน

Section 1 IS THE PUBLICATION RELIABLE?

1 Are the aims clear?

HINT Look for a clear indication at the beginning of the publication of:

what it is about

what it is meant to cover (and what topics are meant to be excluded)

who might find it useful.

If the answer to Question 1 is ‘No’, go directly to Question 3

2 Does it achieve its aims?

HINT Consider whether the publication provides the information it aimed to as outlined in Question 1.

3 Is it relevant?

HINT Consider whether:

the publication addresses the questions that readers might ask

recommendations and suggestions concerning treatment choices are realistic or appropriate.

4 Is it clear what sources of information were used to compile the publication (other than the author or producer)?

HINT Check whether the main claims or statements made about treatment choices are accompanied by a reference to the sources used as evidence, e.g. a research study or expert opinion.

Look for a means of checking the sources used such as a bibliography/ reference list or the addresses of the experts or organizations quoted.

Rating note: In order to score a full ‘5’ the publication should fulfil both

hints. Lists of *additional* sources of support and information (Q7) are not necessarily sources of *evidence* for the current publication.

5 Is it clear when the information used or reported in the publication was produced?

HINT Look for:

dates of the main sources of information used to compile the publication

date of any revisions of the publication (but not dates of reprinting)

date of publication (copyright date).

Rating note: The hints are placed in order of importance – in order to score a full '5' the dates relating to the first hint should be found.

6 Is it balanced and unbiased?

HINT Look for:

a clear indication of whether the publication is written from a personal or objective point of view

evidence that a *range* of sources of information was used to compile the publication, e.g. more than one research study or expert

evidence of an external assessment of the publication.

Be wary if:

the publication focuses on the advantages or disadvantages of one particular treatment choice without reference to other possible choices

the publication relies primarily on evidence from single cases (which may not be typical of people with this condition or of responses to a particular treatment)

the information is presented in a sensational, emotive or alarmist way.

7 Does it provide details of additional sources of support and information?

HINT Look for suggestions for further reading or for details of other organisations providing advice and information about the condition and treatment choices.

8 Does it refer to areas of uncertainty?

HINT

Look for discussion of the gaps in knowledge or differences in expert opinion concerning treatment choices.

Be wary if the publication implies that a treatment choice affects

everyone in the same way, e.g. 100% success rate with a particular treatment.

Section 2 HOW GOOD IS THE QUALITY OF INFORMATION ON TREATMENT CHOICES?
N.B. The questions apply to the treatment (or treatments) described *in the publication.* Self-care is considered a form of treatment throughout this section.

9 Does it describe how each treatment works?
HINT Look for a description of how a treatment acts on the body to achieve its effect.
10 Does it describe the benefits of each treatment?
HINT Benefits can include controlling or getting rid of symptoms, preventing recurrence of the condition and eliminating the condition, both short-term and long-term.
11 Does it describe the risks of each treatment?
HINT Risks can include side-effects, complications and adverse reactions to treatment, both short-term and long-term.
12 Does it describe what would happen if no treatment is used?
HINT Look for a description of the risks and benefits of postponing treatment, of watchful waiting (i.e. monitoring how the condition progresses without treatment) or of permanently forgoing treatment.
13 Does it describe how the treatment choices affect overall quality of life?
HINT Look for:

description of the effects of the treatment choices on day-to-day activity
description of the effects of the treatment choices on relationships with family, friends and carers.

14 Is it clear that there may be more than onepossible treatment choice?
HINT Look for:

a description of who is most likely to benefit from each treatment choice mentioned, and under what circumstances
suggestions of alternatives to consider or investigate further (including choices not fully described in the publication) before deciding whether to select or reject a particular treatment choice.

15 Does it provide support for shared decision-making?
HINT Look for suggestions of things to discuss with family, friends, doctors or

other health professionals concerning treatment choices.

Section 3 OVERALL RATING OF THE PUBLICATION

16 Based on the answers to all of the above questions, rate the overall quality of the publication as a sourceof information about treatment choices

ตารางที่ 32 แสดง Critical values of the F-distribution for the signicance level α = 0:05

df2	df 1																		
	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54	241.88	243.91	245.95	248.01	249.05	250.10	251.14	252.20	253.25	254.31
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41	19.43	19.45	19.45	19.46	19.47	19.48	19.49	19.50
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.37
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.21
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88

ตารางที่ 32 แสดง Critical values of the F-distribution for the signicance level α = 0:05 (ต่อ)

df2	df 1																		
	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.20	2.13	2.05	2.01	1.96	1.91	1.86	1.81	1.76
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.15	2.07	1.99	1.95	1.90	1.85	1.80	1.75	1.69
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20	2.13	2.06	1.97	1.93	1.88	1.84	1.79	1.73	1.67
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.77	1.71	1.65
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18	2.10	2.03	1.94	1.90	1.85	1.81	1.75	1.70	1.64
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	1.62
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47	1.39
120	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96	1.91	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.35	1.25
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	1.00

ประวัติผู้วิจัย

ชื่อ-สกุล	นางสาววรรณทิพย์ แสงทินทอง
ที่อยู่	19/50 หมู่ 4 โรงแรมไทยนำริเวอร์ไซต์ ถนนแสงชูโต ตำบลปากแรต อำเภอบ้านโป่ง จังหวัดราชบุรี
ที่ทำงาน	โรงพยาบาลบ้านโป่ง อำเภอบ้านโป่ง จังหวัดราชบุรี
ประวัติการศึกษา	
พ.ศ.2551	สำเร็จการศึกษาเภสัชศาสตรบัณฑิต มหาวิทยาลัยศิลปากร
พ.ศ.2552	ศึกษาต่อระดับเภสัชศาสตรมหาบัณฑิต สาขาวิชาสารสนเทศศาสตร์ทางสุขภาพ มหาวิทยาลัยศิลปากร อำเภอเมืองนครปฐม จังหวัดนครปฐม
ประวัติการทำงาน	
พ.ศ.2552-2553	เภสัชกรนอกเวลา ประจำร้านยาอำเภอบ้านโป่ง จังหวัดราชบุรี
พ.ศ.2553-2554	เภสัชกร โรงพยาบาลซานคามิลโล จังหวัดราชบุรี
พ.ศ.2555-ปัจจุบัน	เภสัชกร โรงพยาบาลบ้านโป่ง จังหวัดราชบุรี