การพัฒนาวิธีการค้นหาข้อมูลบนอินเตอร์เน็ตกับข้อมูลทะเบียนพันธุ์ไม้ดอกไม้ประดับ

DEVELOPMENT OF INTERNET RETRIEVAL DATABASE FOR REGISTERED ORNAMENTAL PLANTS

พุธ นาทีสุวรรณ

PUTH NATEESUWAN

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

สาขาวิชาการศึกษาวิทยาศาสตร์ (คอมพิวเตอร์)

บัณฑิตวิทยาลัย

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

พ.ศ. 2544

ISBN 974-648-412-5

สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง

การพัฒนาวิธีการค้นหาข้อมูลบนอินเตอร์เน็ต กับข้อมูลทะเบียนพันธุ์ไม้ดอกไม้ประดับ

DEVELOPMENT OF INTERNET RETRIEVAL DATABASE FOR REGISTERED ORNAMENTAL PLANTS

ว ๔๖๘ ก
T040790

พุธ นาฑีสุวรรณ
PUTH NATEESUWAN

เลขหมู่.....
เลขทะเบียน 40790
วัน, เดือน, ปี 26 พ.ย. 2544

.b.....
.i.....

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาการศึกษาวิทยาศาสตร์ (คอมพิวเตอร์)
บัณฑิตวิทยาลัย
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
พ.ศ. 2544

ISBN 974-648-412-5

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

DEVELOPMENT OF INTERNET RETRIEVAL DATABASE FOR REGISTERED ORNAMENTAL PLANTS

PUTH NATEESUWAN

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENT FOR THE DEGREE OF
MASTER OF SCIENCE IN SCIENCE EDUCATION (COMPUTER)
SCHOOL OF GRADUATE STUDIES
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG
2001

ISBN 974-648-412-5

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

COPYRIGHT 2001
SCHOOL OF GRADUATE STUDIES
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บัณฑิตวิทยาลัย

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ใบรับรองวิทยานิพนธ์

หัวข้อวิทยานิพนธ์ การพัฒนาวิธีการค้นหาข้อมูลบนอินเตอร์เน็ต กับข้อมูลทะเบียนพันธุ์ไม้ดอกไม้ประดับ

DEVELOPMENT OF INTERNET RETRIEVAL DATABASE FOR REGISTERED ORNAMENTAL PLANTS

ชื่อนักศึกษา นายพุธ นาฑีสุวรรณ

รหัสประจำตัว 41064249

ปริญญา วิทยาศาสตรมหาบัณฑิต

สาขาวิชา การศึกษาวิทยาศาสตร์

อาจารย์ผู้ควบคุมวิทยานิพนธ์ ผศ.ดร.พรรณี ลีกิจวัฒนะ

อาจารย์ผู้ควบคุมวิทยานิพนธ์ร่วม ดร.สุรสิทธิ์ ราตรี

อาจารย์ศุภร เหมินทร์

คณะกรรมการสอบวิทยานิพนธ์		ลายมือชื่อ
ผศ.ดร.พรรณี	ลีกิจวัฒนะ	
ดร.สุรสิทธิ์	ราตรี	
อาจารย์ศุภร	เหมินทร์	
รศ.ดร.รวีวรรณ	ชินะตระกูล	
ผศ.ดร.เลิศลักษณ์	กลิ่นหอม	

วัน/เดือน/ปี ที่สอบ 16 สิงหาคม 2544 เวลา 8.30 น. เป็นต้นไป

สถานที่สอบ ณ ห้องสมาคมศิษย์เก่าบัณฑิตศึกษา คณะครุศาสตร์อุตสาหกรรม

บัณฑิตวิทยาลัยรับรองแล้ว

(รศ.ดร.บุญวัฒน์ อัตชู)

คณบดีบัณฑิตวิทยาลัย

วันที่ 10 **เดือน** สิงหาคม **พ.ศ.** 2544

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หัวข้อวิทยานิพนธ์	การพัฒนาวิธีการค้นหาข้อมูลบนอินเตอร์เน็ต กับข้อมูลทะเบียนพันธุ์ไม้ดอกไม้ประดับ
นักศึกษา	พุธ นาฑีสุวรรณ
รหัสประจำตัว	41064249
ปริญญา	วิทยาศาสตรมหาบัณฑิต
สาขาวิชา	การศึกษาวิทยาศาสตร์ (คอมพิวเตอร์)
พ.ศ.	2544
อาจารย์ผู้ควบคุมวิทยานิพนธ์	ผศ.ดร.พรรณี ลีกิจวัฒนะ
อาจารย์ผู้ควบคุมวิทยานิพนธ์ร่วม	ดร.สุรสิทธิ์ ราตรี อาจารย์ศุภร เหมินทร์

บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์ เพื่อพัฒนาวิธีการค้นหาข้อมูลบนอินเตอร์เน็ตกับข้อมูลทะเบียนพันธุ์ไม้ดอกไม้ประดับ และทำการเปรียบเทียบประสิทธิภาพของการค้นหาข้อมูลบนอินเตอร์เน็ต ด้วยวิธีการของผู้วิจัย กับวิธี Breath-First Search

กลุ่มตัวอย่างที่ใช้ในการวิจัยครั้งนี้ คือ รายชื่อพันธุ์ไม้ดอกไม้ประดับในเมืองไทย จากพจนานุกรมไม้ดอกไม้ประดับในเมืองไทย และสารานุกรมไม้ดอกไม้ประดับในประเทศไทย จำนวน 1,849 ชื่อ ผู้วิจัยทำการทดลองโดยค้นหาชื่อพันธุ์ไม้ดอกไม้ประดับ จำนวน 17 ครั้ง โดยใช้วิธีการของผู้วิจัย และวิธี Breath-First Search กับข้อมูลที่เพิ่มขึ้นครั้งละ 20 เปอร์เซ็นต์ของข้อมูลครั้งก่อน โดยเริ่มต้นครั้งแรกที่ 100 รายการ นำผลที่ได้มาวิเคราะห์และเปรียบเทียบประสิทธิภาพ โดยทำการศึกษาเปรียบเทียบใน 4 ด้านคือ ความครบถ้วนของข้อมูล ด้านความตรงต่อความต้องการ ด้านเวลาที่ใช้ในการค้นหาข้อมูล และด้านการใช้เนื้อที่ในการเก็บข้อมูลเพื่อการค้นหา ผลการวิเคราะห์ข้อมูลพบว่าวิธีการของผู้วิจัยมีประสิทธิภาพเท่ากันกับวิธี Breath-First Search 2 ด้าน คือ ด้านความครบถ้วนของข้อมูล และด้านความตรงต่อความต้องการ แต่มีประสิทธิภาพสูงกว่า 2 ด้านคือ ด้านเวลาที่ใช้ในการค้นหาข้อมูล และด้านการใช้เนื้อที่ในการเก็บข้อมูลเพื่อการค้นหา ซึ่งเป็นไปตามสมมุติฐานของการวิจัย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Thesis Title	Development of Internet Retrieval Database for Registered Ornamental Plants
Student	Mr.Puth Nateesuwan
Student ID	41064249
Degree	Master of Science
Programme	Science Education (Computer)
Year	2001
Thesis Advisor	Asst.Prof.Dr. Punnee Leekitchwatana
Thesis Co-advisor	Dr. Surasit Ratree
	Mr.Suporn Hemindra

ABSTRACT

This study is aimed to develop the searching information in the internet and in the species credenticals of garden plants and trees as well as to compare the efficiency of searching the information on the internet between the research-constructed method and the Breath-First Search.

The samples of this study are 1,849 names of the species of ornamental plants from the dictionary of Thai garden plants and trees and the encyclopedia of Thai garden plants and tree. The research has experimented searching the names for 17 times, and for each experiment both researcher-constructed and Breath-First Search method with increasing name 20 percentage each time, by start with 100 names. The data are then operated to compare the methods in 4 aspects namely data completeness, validity, searching duration and space utilization for storage and searching. The results reveal that both methods are equally efficient in completeness and validity; however, the researcher-constructed method is better than the Breath-First Search in searching duration and space utilization. The results are coincided with the hypothesis that the research-constructed method is more efficient than the Breath-First Search.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จได้ด้วยความกรุณาเป็นอย่างยิ่งจาก ผศ.ดร.พรรณี ลีกิจวัฒนะ อาจารย์ผู้ควบคุมวิทยานิพนธ์ และ ดร.สุรสิทธิ์ ราตรี อาจารย์ผู้ควบคุมวิทยานิพนธ์ร่วม ที่ได้ให้คำแนะนำชี้แนะแนวทางในการดำเนินงาน และแนวทางในการปรับปรุงแก้ไขข้อบกพร่องต่างๆ เพื่อให้วิทยานิพนธ์ฉบับนี้สมบูรณ์ยิ่งขึ้น ผู้วิจัยขอ กราบขอบพระคุณอย่างสูงมา ณ โอกาสนี้

ขอกราบขอบพระคุณ อาจารย์ศุภร เหมินทร์ อาจารย์ผู้ควบคุมวิทยานิพนธ์ร่วม ที่ได้กรุณาให้คำแนะนำในเรื่องของพรรณไม้ดอกไม้ประดับ อันเป็นข้อมูลที่สำคัญต่อการวิจัยในครั้งนี้เป็นอย่างมาก และคำชี้แนะต่างเพื่อให้วิทยานิพนธ์ฉบับนี้สมบูรณ์ยิ่งขึ้น

ขอกราบขอบพระคุณ รศ.ดร รวีวรรณ ชินะตระกูล ผศ.ดร.เลิศลักษณ์ กลิ่นหอม ซึ่งเป็นกรรมการสอบ ที่ให้การช่วยเหลือ แนะนำ ตรวจสอบ แก้ไข ข้อบกพร่องต่างๆ และให้แนวคิดในการทำวิจัยครั้งนี้

ขอกราบขอบพระคุณ ดร.วิทย์ เที่ยงบูรณธรรม และ คุณประสิทธิ์ คำภูแสน สำหรับข้อมูลที่นำมาใช้ในการวิจัยครั้งนี้

เนื่องจากทุนการวิจัยครั้งนี้บางส่วนได้รับการสนับสนุนจากทุนอุดหนุนการวิจัยของบัณฑิตวิทยาลัยสถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง ผู้วิจัยขอขอบคุณมา ณ โอกาสนี้

ขอขอบคุณ คุณณัฏฐา บุญอยู่ คุณนิธิวรรณ รุ่งรังษี และเพื่อน ๆ อีกหลายคนที่ไม่สามารถเอ่ยชื่อได้หมด ในความกรุณาช่วยเหลือมากมาย ให้การทำวิจัยครั้งนี้ประสบความสำเร็จด้วยดี

สุดท้ายขอน้อมระลึกถึงพระคุณ คุณพ่อ คุณแม่ ผู้ให้ทุกสิ่งทุกอย่างในชีวิต รวมทั้งกำลังใจที่ดีที่สุดตลอดมา

พุธ นาฑีสุวรรณ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

หน้า

บทคัดย่อภาษาไทย I

บทคัดย่อภาษาอังกฤษ II

กิตติกรรมประกาศ III

สารบัญ IV

สารบัญตาราง VI

สารบัญภาพ VII

บทที่ 1 บทนำ 1

1.1 ความสำคัญและความเป็นมาของปัญหา 1

1.2 วัตถุประสงค์ของการวิจัย 3

1.3 สมมติฐานการวิจัย 3

1.4 ขอบเขตของการวิจัย 4

1.5 กรอบแนวคิดที่ใช้ในการวิจัย 4

1.6 นิยามศัพท์เฉพาะที่ใช้ในการวิจัย 6

บทที่ 2 เอกสารและงานวิจัยที่เกี่ยวข้อง 8

2.1 การค้นหาข้อมูลสารสนเทศ 8

2.2 อัลกอริทึมเพื่อการค้นหาแบบโครงสร้างของข้อมูล 9

2.3 อัลกอริทึมเพื่อการค้นหาแบบโครงสร้างข้อมูล 15

2.4 อัลกอริทึมที่ใช้ในการค้นหาข้อมูลสารสนเทศ 17

2.5 ทฤษฎีที่ใช้เป็นพื้นฐานของแนวความคิดในวิธีการของผู้วิจัย 20

2.6 ความหมายและความสำคัญของไม้ดอกไม้ประดับ 22

2.7 โครงสร้างของทะเบียนพันธุ์ไม้ดอกไม้ประดับ 23

บทที่ 3 วิธีดำเนินการวิจัย 24

3.1 ประชากรและกลุ่มตัวอย่าง 24

3.2 เครื่องมือที่ใช้ในการวิจัย 24

3.3 การดำเนินการทดลอง 39

3.4 การวิเคราะห์ข้อมูล 41

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

	หน้า
บทที่ 4 ผลการวิเคราะห์ข้อมูล	42
4.1 เปรียบเทียบประสิทธิภาพของการค้นหาข้อมูลบนอินเตอร์เน็ตที่เป็นผลจากการทดลองด้วยวิธีการที่ผู้วิจัยสร้างขึ้น กับวิธีการแบบ breath-first search	42
4.2 สรุปผลการวิเคราะห์ข้อมูล	47
บทที่ 5 สรุปผลการวิจัย อภิปรายผลและข้อเสนอแนะ	48
5.1 สรุปผลการวิจัย	48
5.2 อภิปรายผลการวิจัย	49
5.3 ข้อเสนอแนะ	51
บรรณานุกรม	52
ภาคผนวก	54
ประวัติผู้เขียน	59

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง

ตารางที่		หน้า
3.1	แสดงคุณสมบัติของพยัญชนะกับข้อมูลทะเบียนพันธุ์ไม้ดอกไม้ประดับ.....	27
3.2	แสดงสระในภาษาไทย.....	28
3.3	แสดงความถี่การใช้สระในข้อมูลทะเบียนพันธุ์ไม้ดอกไม้ประดับ.....	29
3.4	แสดงวรรณยุกต์ และตัวอักษรอื่น ๆ ในภาษาไทย.....	30
3.5	แสดงความถี่การใช้วรรณยุกต์ในข้อมูลทะเบียนพันธุ์ไม้ดอกไม้ประดับ.....	30
3.6	แสดงประเภทของคำค้น.....	40
4.1	แสดงผลการทดลองหาประสิทธิภาพการค้นหาข้อมูลบนอินเตอร์เน็ตกับข้อมูลทะเบียนพันธุ์ไม้ดอกไม้ประดับ โดยวิธีการที่ผู้วิจัยสร้างขึ้น และวิธีการแบบ Breath-first search.....	43

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญภาพ

ภาพที่		หน้า
2.1	แสดงการค้นหาข้อมูลแบบ depth-first search.....	11
2.2	แสดงขั้นตอนการค้นหาข้อมูลแบบ breath-first search.....	12
2.3	แสดงไบนารี่ทรีของคำว่า "01100100010111..." หลังการบันทึกข้อมูลผ่านไป 8 ตัวอักษร.....	13
2.4	แสดงแพททรีเซียทรีของคำว่า "01100100010111..." หลังการบันทึกข้อมูลผ่านไป 8 ตัวอักษร.....	14
2.5	แสดงเพิ่มค่า 16 ลงในตารางแฮชชิ่งโดยการเข้าฟังก์ชันแฮชชิ่ง เพื่อหาช่องว่าง ในตาราง.....	15
2.6	แสดงขั้นตอนการสร้างดัชนี.....	17
2.7	โครงสร้างทะเบียนพันธ์ไม้ดอกไม้ประดับที่ใช้ในประเทศไทย.....	23
3.1	แสดงตัวพยัญชนะภาษาไทย.....	25
3.2	แสดงขั้นตอนการวิเคราะห์คุณสมบัติของข้อมูลทะเบียนพันธุ์ไม้ดอกไม้ประดับ.....	26
3.3	แสดงแนวความคิดในการมองชื่อไทยของพืชเป็นวัตถุ.....	31
3.4	แสดงแนวความคิดในการเชื่อมความสัมพันธ์ของวัตถุ.....	32
3.5	แสดงการเตรียมข้อมูลก่อนการค้นหา.....	33
3.6	แสดงการค้นหาข้อมูลจากวัตถุที่มีคุณสมบัติเดียวกัน.....	34
3.7	แสดงกลุ่มของข้อมูลเพื่อให้ผู้ใช้ได้เลือกข้อมูลที่ตนต้องการ.....	34
3.8	แสดงขั้นตอนในการแสดงข้อมูลที่ผู้ใช้ได้เลือกแล้ว.....	35
3.9	Data Flow Diagram (Context Diagram).....	35
3.10	Data Flow Diagram (Level 0).....	36
3.11	Data Flow Diagram (Level 1 of Pre-processing Module).....	36
3.12	Data Flow Diagram (Level 1 of Query Module).....	37
3.13	FlowChart แสดงโครงสร้างการค้นหาข้อมูล.....	38
3.14	แสดงขั้นตอนการทำการทดลอง.....	40
4.1	กราฟแสดงผลวิเคราะห์ในด้านความครบถ้วนของข้อมูล ทั้ง 2 วิธี.....	44
4.2	กราฟแสดงผลวิเคราะห์ในด้านความตรงต่อความต้องการ ทั้ง 2 วิธี.....	45
4.3	กราฟแสดงผลวิเคราะห์ในด้านเวลาที่ใช้ในการค้นหาข้อมูล ทั้ง 2 วิธี.....	46
4.4	กราฟแสดงผลวิเคราะห์ในด้านการใช้เนื้อที่ในการเก็บข้อมูลเพื่อการค้นหา ทั้ง 2 วิธี	46

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญภาพ (ต่อ)

ภาพที่		หน้า
4.5	กราฟแสดงผลวิเคราะห์ในด้านเวลาที่ใช้ในการค้นหาข้อมูล ทั้ง 2 วิธี.....	47
4.6	กราฟแสดงผลวิเคราะห์ในด้านการใช้เนื้อที่ในการเก็บข้อมูลเพื่อการค้นหา ทั้ง 2 วิธี	47
5.1	แสดงขั้นตอนการค้นหาข้อมูลระหว่างวิธีการของผู้วิจัย และวิธี Breath-Fir Search	50

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1
บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ในปัจจุบันการพัฒนาทางด้านเทคโนโลยีต่าง ๆ ได้เข้ามามีบทบาทสำคัญในการเปลี่ยนแปลงสังคมมนุษย์ให้เข้าสู่สังคมยุคสารสนเทศ (Information Society) การติดต่อสื่อสารข้อมูลสามารถติดต่อถึงกันได้ทั่วโลกเพียงชั่วพริบตา โดยมีเครื่องคอมพิวเตอร์เป็นเครื่องมือสำคัญในการเชื่อมต่อเครือข่ายการสื่อสาร ด้วยสมรรถนะของคอมพิวเตอร์ในการเก็บและรับส่งข้อมูลได้เป็นจำนวนมาก จึงได้มีการนำเข้ามาใช้ในทุกวงการ การสื่อสารทางเครือข่ายคอมพิวเตอร์เป็นที่รู้จักกันอย่างกว้างขวางและเป็นที่นิยมใช้สื่อสารกันระหว่างบุคคลและสถาบันต่าง ๆ ในปัจจุบัน การสื่อสารทางเครือข่ายอินเตอร์เน็ต (Internet) ซึ่งเป็นเครือข่ายคอมพิวเตอร์สากลขนาดใหญ่ที่เชื่อมโยงต่อกับเครือข่ายคอมพิวเตอร์อื่น ๆ มากมายทั่วโลก ทำให้มีการติดต่อสื่อสารเชื่อมโยงกับแหล่งข้อมูลได้อย่างกว้างขวางทั่วโลกทำให้ผู้ใช้สามารถเข้าถึงแหล่งข้อมูลต่าง ๆ ในเครือข่ายได้อย่างมากมาย ไม่ว่าจะเป็นข้อมูลทางด้านการศึกษา ด้านการค้าหรือด้านธุรกิจบันเทิง โดยทำการค้นหาข้อมูลได้จากอินเตอร์เน็ต

ในปัจจุบันการสื่อสารผ่านอินเตอร์เน็ต ได้มีการพัฒนาโปรแกรมต่าง ๆ ให้ผู้ใช้สามารถใช้งานได้อย่างสะดวกด้วยวิธีการที่ง่ายสะดวก และราคาไม่แพง ทั้งยังเป็นการสื่อสารที่สามารถส่งและรับข้อมูลได้ทั้งข้อความที่เป็นตัวอักษร ภาพ และเสียง ดังนั้นด้วยสมรรถนะที่สมบูรณ์และเป็นการสื่อสารที่กว้างไกล โดยไม่จำกัดระยะทางและเวลา จึงสามารถที่จะนำมาใช้ประโยชน์กับการศึกษาได้อย่างมีประสิทธิภาพ วิธีการค้นหาข้อมูลบนอินเตอร์เน็ต จึงเป็นเครื่องมือที่สำคัญสำหรับการค้นหาข้อมูลในปัจจุบัน เพราะผู้ใช้ทุกคนต้องการข้อมูลในสิ่งที่ตนคาดหวังและเวลาที่ใช้ในการค้นหาไม่มากจนเกินไป ซึ่งการพัฒนาวิธีการค้นหาข้อมูลบนอินเตอร์เน็ตให้สามารถค้นหาข้อมูลที่ผู้ใช้ต้องการ ได้ตรงตามความต้องการและใช้เวลาที่รวดเร็วในการค้นหา (ทิพย์เกสร บุญอำไพ. 2540 : 1)

วิธีการค้นหาบนอินเตอร์เน็ตนั้น ควรจะมีการใช้วิธีการทางด้านปัญญาประดิษฐ์ให้มากขึ้น โดยคำนึงถึงโครงสร้างของข้อมูลที่จะนำเสนอ พิจารณาจากข้อมูลว่าเราสามารถรู้เป้าหมาย (Goal) ของผู้ใช้ได้หรือไม่ ถ้าไม่สามารถรู้ได้เลยว่าผู้ใช้ต้องการอะไรให้ใช้วิธีการของปัญญาประดิษฐ์ทางในส่วนวิธีการ Uninformed แต่ถ้าสามารถรู้ได้ว่าผู้ใช้ต้องการอะไรให้ใช้วิธีการในส่วน Informed

วิธีการค้นหาแบบ Informed เป็นวิธีการค้นหาที่โปรแกรมสามารถรู้เป้าหมาย (Goal) ของผู้ใช้ได้ ได้แก่งานที่ประยุกต์ในการสร้างเกมส์ต่าง ๆ เช่น ทิคแทคโท (Tic-Tac-Toe) ที่มีช่องสำหรับเล่นอยู่ 9 ช่อง โปรแกรมสามารถสร้างเส้นทางได้ว่าก่อนเลยว่าควรจะเล่นอย่างไรจึงจะปลอดภัย และสามารถชนะได้เร็วที่สุด หรือหมากรุก เป็นต้น วิธีการค้นหานั้นโปรแกรมจะสร้างเส้นทางขึ้นมาตั้งแต่ต้นจนจบ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ทำให้โปรแกรมทราบว่าเส้นทางที่ใช้สามารถเอาชนะคู่ต่อสู้ได้ จากการศึกษางานวิจัยต่าง ๆ พอจะสรุปได้ว่า เนื่องจากในปัจจุบันยังไม่ค่อยมีผู้ให้ความสนใจกับการพัฒนาวิธีการค้นหาข้อมูล การค้นหาข้อมูลในปัจจุบันเราจะเห็นกันโดยทั่วไปแล้วว่าจะสร้างขึ้นมาโดยอาศัยเครื่องมือที่ใช้กันอยู่ทั่วไป เช่น ผ่านทาง ODBC DBE เป็นต้น การใช้ข้อมูลผ่านทางเครื่องมือเหล่านี้มีประสิทธิภาพสูงและเหมาะสมกับการทำงานในลักษณะขององค์กรมาก แต่เมื่อนำมาประยุกต์ใช้บนอินเตอร์เน็ตแล้ว จะมีข้อยุ่งยาก และน่ารำคาญอยู่มากเพราะจะต้องคำนึงถึงองค์ประกอบหลาย ๆ อย่างเช่น เนื้อที่โฮมเพจที่ได้รับมา หน่วยความจำของเครื่อง เซิร์ฟเวอร์ที่ให้บริการ เนื้อที่ที่ใช้เก็บข้อมูลของเซิร์ฟเวอร์ที่ให้บริการ ซึ่งการค้นหาข้อมูลผ่านทางไดร์เวอร์นี้ก็ยังคงวิธีการพื้นฐานอยู่ คือการเทียบแบบ ASCII นั่นเอง ฉะนั้นทิศทางของงานวิจัยประเภทนี้ จึงจำเป็นที่จะต้องอาศัยวิธีการทางด้านปัญญาประดิษฐ์ (Artificial Intelligent) เพื่อช่วยพัฒนาประสิทธิภาพในการค้นหา โดยเฉพาะกับข้อมูลเทคนิคที่มีลักษณะเฉพาะ เช่น ข้อมูลทางด้านเคมี เกษตร แพทย์ หรือแม้กระทั่งข้อมูลทางด้านคอมพิวเตอร์เองก็ตาม ซึ่งวิธีการทางด้านนี้จะเริ่มจากการใช้วิธีแบบ Depth-First Search แต่มีข้อจำกัดอยู่มากจึงมีการพัฒนาวิธีการแบบ Breath-First Search ขึ้นมาเพื่อช่วยลดปัญหาของวิธีเดิมแต่วิธีการแบบ Breath-First Search นี้มีข้อจำกัดคือต้องการการเข้าถึงหน่วยความจำเป็นจำนวนมาก

ข้อมูลที่กล่าวถึงเรื่องของไม้ดอกไม้ประดับของไทยนั้นมีอยู่น้อยในทะเบียนข้อมูล ทั้ง ๆ ที่ไม้ดอกไม้ประดับมีความสำคัญต่อชีวิตประจำวันของเราอยู่มาก ให้ความเจริญตาเจริญใจ ให้ร่มเงา ฯลฯ จะเห็นได้ว่าประเทศต่าง ๆ ก็ยอมรับว่าพืชเหล่านี้มีความสำคัญให้ด้านอาหารตา พักผ่อนหย่อนใจ ถึงกับยอมลงทุนด้วยจำนวนมหาศาลในการสร้างสวนสาธารณะปลูกไม้ดอกไม้ประดับนานาชนิด ให้คนชมได้พักผ่อนหย่อนใจ และให้คลายอารมณ์อันตึงเครียด บ้านเราดินฟ้าอากาศเหมาะแก่การปลูกไม้ดอกไม้ประดับหลายชนิด คนไทยนิยมปลูกต้นไม้อยู่แล้วหากแต่ยังขาดเครื่องมือที่จำเป็นคือ ความรู้ทางพืช ซึ่งการจะศึกษาหาความรู้ทางพืชนั้น จำเป็นจะต้องใช้ชื่อทางวิทยาศาสตร์ของพืชเป็นปัจจัยหลักในการศึกษา หรือค้นคว้าหาความรู้ แต่คนไทยจะชินกับชื่อไทยของพืชมากกว่า ทำให้การศึกษาหรือค้นคว้าหาความรู้นั้นถูกจำกัดลงไป คือไม่สามารถค้นจากหนังสือ หรือแหล่งข้อมูลต่างประเทศได้ เนื่องจากชื่อไทยของพืชนั้น ชาวต่างชาติไม่รู้จัก

ปัญหาเกี่ยวกับชื่อไทยของพืชอีกอย่างหนึ่งก็คือ การเรียกชื่อพันธุ์ไม้ดอกไม้ประดับต่าง ๆ ในปัจจุบันยังสับสนอยู่มาก มีทั้งพันธุ์ไม้ชนิดเดียวกันแต่มีหลายชื่อ และหลายชนิดในชื่อเดียวกัน ที่ไม่มีชื่อไทยก็มากต้องเรียกทับศัพท์เอาจากชื่อสามัญหรือชื่อวิทยาศาสตร์ ซึ่งก็มีถูกบ้างผิดบ้างเพี้ยนบ้างก็ตามถนัด หรือไม่ก็ตั้งชื่อขึ้นเองตามใจชอบของแต่ละคนซึ่งบางทีไปซ้ำกับชื่อเก่าที่เป็นไม้คนละชนิดกัน หรือชื่อดั้งเดิมมีอยู่แล้วแต่ไปตั้งขึ้นใหม่อีก หากยังคงเป็นเช่นนี้ต่อไปเรื่อย ๆ นับวันจะสับสนยิ่งขึ้นทุกที คงจะทำความยุ่งยากให้กับผู้ที่สนใจหรือศึกษาหาความรู้ไม่น้อยทีเดียว จะอาศัยหนังสือตำราหรือสิ่งพิมพ์ต่าง ๆ บางทีก็ไม่ค่อยจะตรงกันนักประกอบกับขณะนี้มีผู้สนใจค้นคว้าหาพันธุ์ไม้ดอกไม้ประดับ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ทางทุ่งนาป่าเขาในบ้านเรา และนำเข้ามาจากต่างประเทศมากยิ่งขึ้นเรื่อย ๆ (พันธุ์ทิพย์ บริพัตร และคณะ. 2520 : 3)

การจำแนกสังคมพืชและเรียกชื่อที่ผิดพลาด อาจก่อให้เกิดผลเสียอย่างมหันต์ในแง่การพัฒนาพื้นที่สังคมพืช ดังเช่นปัญหาที่เกิดขึ้นกับสังคมพืชป่าพรุในปัจจุบัน การพัฒนาพื้นที่ป่าพรุดั้งเดิมที่มีคุณค่าทางเศรษฐกิจอย่างมหาศาล ทั้งในด้านการป่าไม้และการอนุรักษ์พื้นที่ไว้เป็นแหล่งเก็บน้ำจืดและความชุ่มชื้นตามธรรมชาติ โดยทำทางระบายน้ำออกจากพรุสู่ทะเล จะเป็นการทำลายทรัพยากรธรรมชาติผืนสุดท้ายของประเทศที่มีคุณค่ายิ่งให้สูญสิ้นไปโดยความรู้เท่าไม่ถึงการณ์ (สืบ นาคะเสถียร. 2542 : 135)

ดังนั้นผู้วิจัย จึงสนใจพัฒนาวิธีการค้นหาข้อมูลบนอินเตอร์เน็ตในยุคเทคโนโลยีสารสนเทศกับข้อมูลเฉพาะกลุ่มในส่วนของทะเบียนพันธุ์ไม้ดอกไม้ประดับให้มีประสิทธิภาพมากยิ่งขึ้น เพื่อช่วยส่งเสริมการนำเสนอข้อมูล และการใช้ข้อมูลประเภทนี้บนอินเตอร์เน็ตมากยิ่งขึ้นต่อไปในอนาคต

1.2 วัตถุประสงค์ของการวิจัย

1. เพื่อพัฒนาวิธีการค้นหาข้อมูลบนอินเตอร์เน็ตกับข้อมูลทะเบียนพันธุ์ไม้ดอกไม้ประดับให้มีประสิทธิภาพ
2. เพื่อเปรียบเทียบประสิทธิภาพของการค้นหาข้อมูลบนอินเตอร์เน็ต ด้วยวิธีการของผู้วิจัยกับวิธี แบบ Breath-First Search

1.3 สมมติฐานการวิจัย

การค้นหาข้อมูลบนอินเตอร์เน็ตกับข้อมูลทะเบียนพันธุ์ไม้ดอกไม้ประดับ ด้วยวิธีการของผู้วิจัย มีประสิทธิภาพสูงกว่าวิธี Breath-First Search โดยเปรียบเทียบในด้านความครบถ้วนของข้อมูล ด้านความตรงต่อความต้องการ ด้านเวลาที่ใช้ในการค้นหาข้อมูล และด้านการใช้เนื้อที่ในการเก็บข้อมูลเพื่อการค้นหา

1.4 ขอบเขตของการวิจัย

1.4.1 ประชากรและกลุ่มตัวอย่าง

1.4.1.1 ประชากรที่ใช้ในการวิจัย

ประชากรที่ใช้ในการวิจัย คือ รายชื่อพันธุ์ไม้ดอกไม้ประดับในเมืองไทย จากพจนานุกรมไม้ดอกไม้ประดับในเมืองไทย (วิทย์ เที่ยงบูรณธรรม. 2536) และสารานุกรมไม้ดอกไม้ประดับในประเทศไทย (สวัสดิ์ หรั่งเจริญ. 2525) จำนวน 2,000 ชื่อ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.4.1.2 **กลุ่มตัวอย่างที่ใช้ในการวิจัย**

กลุ่มตัวอย่างที่ใช้ในการวิจัย คือ รายชื่อพันธุ์ไม้ดอกไม้ประดับในเมืองไทย จากพจนานุกรมไม้ดอกไม้ประดับในเมืองไทย (วิทย์ เที่ยงบูรณธรรม. 2536) และสารานุกรมไม้ดอกไม้ประดับในประเทศไทย (สวัสดิ์ หรั่งเจริญ. 2525) จำนวน 1,849 ชื่อ โดยเริ่มจาก 100 ชื่อ และทำการเพิ่มครั้งละ 20 เปอร์เซ็นต์ จำนวน 17 ครั้ง ได้ชื่อ 1849 ชื่อ

1.4.2 **ลักษณะของการพัฒนาวิธีการค้นหาข้อมูลทะเบียนพันธุ์ไม้ดอกไม้ประดับในประเทศไทย**

การพัฒนาวิธีการค้นหาข้อมูลทะเบียนพันธุ์ไม้ดอกไม้ เป็นลักษณะการค้นหาข้อมูลจากชื่อไทยของพืชเพื่อที่จะเทียบไปยังชื่อวิทยาศาสตร์ของพืช และเปรียบเทียบประสิทธิภาพของการค้นหา คือ วิธีการของผู้วิจัย กับวิธี Breath-First Search แล้วทำการเก็บข้อมูล นำมาสรุปและวิเคราะห์ใน 4 ด้าน ดังนี้

1. ความครบถ้วนของข้อมูล
2. ความตรงต่อความต้องการ
3. เวลาที่ใช้ในการค้นหาข้อมูล
4. การใช้เนื้อที่ในการเก็บข้อมูลเพื่อการค้นหา

1.4.3 **การดำเนินการทดลอง**

ในการทดลองเริ่มทำการทดลองและเก็บข้อมูลตั้งแต่ 1 พฤษภาคม 2543 ถึง วันที่ 1 พฤศจิกายน 2543

1.5 กรอบแนวคิดที่ใช้ในการวิจัย

การวิจัยครั้งนี้ ผู้วิจัยจะทำการสร้างวิธีการค้นหาข้อมูลบนอินเตอร์เน็ตโดยศึกษาจากงานวิจัยต่าง ๆ เช่น งานวิจัยของ Richard Entlich และ Jan Olsen ได้ทำการศึกษาและพัฒนาการสร้างโปรแกรมห้องสมุดดิจิตอล โดยการพัฒนาระบบทรี (Tree) ใช้โมเดลแบบ CORE (Chemical Online Retrieval Experiment) เข้ามาช่วย และพัฒนาบนระบบภายในองค์กรเป็น Unix-Based Client-Server เพื่อให้ผู้ใช้เข้ามาค้นหาข้อมูล ชื่อว่านิวตัน (Newton) งานวิจัยของ Mark Lager ที่ศึกษาถึงวิธีการค้นหาข้อมูลบนอินเตอร์เน็ต ทำให้ทราบได้ว่าเครื่องมือที่ใช้สำหรับค้นหาข้อมูลบนอินเตอร์เน็ตมี 2 ประเภทด้วยกันคือ Index และ Search Engines ซึ่งทั้ง 2 อย่างนี้ต้องทำงานไปด้วยกัน คือ Index จะสร้างดัชนีขึ้นมาเพื่อให้ Search Engines เข้ามาค้นหาข้อมูลที่เป็นดัชนี งานวิจัยของ Ellman, J และ Tail, J. ได้ศึกษาถึงการพิจารณาข้อมูลเพื่อใช้ในการค้นหาข้อมูลผ่านเว็ป พบว่าการใช้วิธีการธรรมชาติ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สำหรับมนุษย์ในการค้นหาข้อมูลนั้น จะช่วยให้ผู้ใช้สามารถกำหนดสิ่งที่ตนเองต้องการรู้ (Keyword) และค้นหาข้อมูลที่ตนเองต้องการได้เป็นที่น่าพอใจ ตรงกับความต้องการ งานวิจัยของ G. Michael Youngblood ได้ทำการศึกษาและนำเสนอข้อมูลเรื่อง Web Haunting : Design of a Simple Intelligent Web Search Agent โดยงานชิ้นนี้เขาได้นำเสนอแนวทางในการสร้างอัลกอริทึมในการค้นหาและต้องการให้มีการศึกษาวิจัยเพิ่มที่ส่วนของวิธีค้นหาโดยใช้วิธีการของปัญญาประดิษฐ์ (Artificial Intelligent)

จากการศึกษางานวิจัยต่าง ๆ พอจะสรุปได้ว่า เนื่องจากในปัจจุบันยังไม่ค่อยมีผู้ให้ความสนใจกับการพัฒนาวิธีการค้นหาข้อมูล การค้นหาข้อมูลในปัจจุบันเราจะเห็นกันโดยทั่วไปแล้วว่าจะสร้างขึ้นมาโดยอาศัยเครื่องมือที่ใช้กันอยู่ทั่วไป เช่น ผ่านทาง ODBC DBE เป็นต้น การใช้ข้อมูลผ่านทางเครื่องมือเหล่านี้มีประสิทธิภาพสูงและเหมาะสมกับการทำงานในลักษณะขององค์กรมาก แต่เมื่อนำมาประยุกต์ใช้บนอินเตอร์เน็ตแล้ว จะมีข้อยุ่งยาก และน่ารำคาญอยู่มากเพราะจะต้องคำนึงถึงองค์ประกอบหลาย ๆ อย่างเช่น เนื้อที่โฮมเพจที่ได้รับมา หน่วยความจำของเครื่อง เซิร์ฟเวอร์ที่ให้บริการ เนื้อที่ที่ใช้เก็บข้อมูลของเซิร์ฟเวอร์ที่ให้บริการ ซึ่งการค้นหาข้อมูลผ่านทางไดร์เวอร์นี้ก็ยังคงวิธีการพื้นฐานอยู่ คือการเทียบแบบ ASCII นั่นเอง ฉะนั้นทิศทางของงานวิจัยประเภทนี้ จึงจำเป็นที่จะต้องอาศัยวิธีการทางด้านปัญญาประดิษฐ์ (Artificial Intelligent) เพื่อช่วยพัฒนาประสิทธิภาพในการค้นหา โดยเฉพาะกับข้อมูลเทคนิคที่มีลักษณะเฉพาะ เช่น ข้อมูลทางด้านเคมี เกษตร แพทย์ หรือแม้กระทั่งข้อมูลทางด้านคอมพิวเตอร์เองก็ตาม ซึ่งวิธีการทางด้านนี้จะเริ่มจากการใช้วิธีแบบ Depth-First Search แต่มีข้อจำกัดอยู่มากจึงมีการพัฒนาวิธีการแบบ Breath-First Search ขึ้นมาเพื่อช่วยลดปัญหาของวิธีเดิม แต่วิธีการแบบ Breath-First Search นี้มีข้อจำกัดคือ ต้องการการเข้าถึงหน่วยความจำเป็นจำนวนมาก

วิธีการของผู้วิจัย ตามแนวความคิดที่สร้างขึ้นคือ การนำเอาคุณสมบัติของอักษรภาษาไทยมาพิจารณาสร้างรูปแบบวิธีการค้นหาเฉพาะตัวขึ้น โดยจะมองพยัญชนะ สระ วรรณยุกต์ และอักษรอื่น ๆ ให้เป็นวัตถุอิสระต่อกัน ซึ่งมีค่าความสำคัญในระดับต่าง ๆ และเมื่ออักษรถูกมองเป็นวัตถุ ให้แต่ละวัตถุสามารถมีความสัมพันธ์กันได้อย่างอิสระต่อกัน โดยการค้นหาจะใช้คุณสมบัติของความสัมพันธ์ของวัตถุมาพิจารณาใช้การค้นหาข้อมูลซึ่งจะช่วยลดจำนวนเวลา และจำนวนเนื้อที่ในการค้นหาข้อมูลได้มาก

การหาประสิทธิภาพของวิธีการค้นหาข้อมูลบนอินเตอร์เน็ตกับข้อมูลทะเบียนพันธุ์ไม้ดอกไม้ประดับ จะใช้ดัชนี 4 ดัชนีด้วยกัน คือ

1.ความครบถ้วนของข้อมูล คือ ความสามารถของวิธีการค้นหาข้อมูลที่สามารถค้นหา (Query) ข้อมูล และดึงข้อมูล (Retrieve) ข้อมูลทะเบียนพันธุ์ไม้ดอกไม้ประดับได้ครบถ้วนตามที่มีอยู่ในฐานข้อมูลจริง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. ความตรงต่อความต้องการ (Relevance) คือ ความสามารถของวิธีการค้นหาข้อมูลที่สามารถเรียกข้อมูลทะเบียนพันธุ์ไม้ดอกไม้ประดับค้นคืนมาได้ถูกต้องตามที่ผู้ใช้ต้องการ

3. เวลาที่ใช้ในการค้นหาข้อมูล (Execution Time) คือ จำนวนเวลาของวิธีการค้นหา ที่ใช้ไปในการค้นหาข้อมูลทะเบียนพันธุ์ไม้ดอกไม้ประดับ ซึ่งจะมีหน่วยเป็นวินาที

4. การใช้เนื้อที่ในการเก็บข้อมูลเพื่อการค้นหา (Memory Space) คือ จำนวนของเนื้อที่ที่ถูกใช้ไปเพื่อการค้นหาข้อมูลตามวิธีการค้นหาข้อมูล ซึ่งจะมีหน่วยเป็นไบต์ (Byte)

การเปรียบเทียบประสิทธิภาพระหว่างวิธีการค้นหาทั้ง 2 วิธี คือการนำเอาค่าผลที่ได้จากการหาประสิทธิภาพของดัชนีทั้ง 4 ด้านที่ได้กล่าวไปแล้วข้างต้น มาทำการเปรียบเทียบกันระหว่างวิธีการของผู้วิจัย กับวิธี Breath-First Search

1.6 นิยามศัพท์เฉพาะที่ใช้ในการวิจัย

เพื่อให้เข้าใจความหมายของศัพท์เฉพาะในการวิจัยตรงกัน ผู้วิจัยจึงกำหนดนิยามศัพท์ที่ใช้ในการวิจัย ดังนี้

1. การค้นหาข้อมูลบนอินเตอร์เน็ต หมายถึง การค้นหาข้อมูลที่อยู่บนอินเตอร์เน็ตโดยใช้วิธีการค้นหาข้อมูลใด ๆ ผ่านทางเครื่องคอมพิวเตอร์ที่ผู้ใช้ระบุความต้องการในการค้นหาข้อมูล เพื่อนำข้อมูลไปใช้ให้เกิดประโยชน์ในด้านต่าง ๆ เช่น เพื่อศึกษาหาความรู้ หรือความบันเทิง เป็นต้น

2. ข้อมูลทะเบียนพันธุ์ไม้ดอกไม้ประดับ หมายถึง ข้อมูลทะเบียนพันธุ์ไม้ดอกไม้ประดับเฉพาะในประเทศไทย จากพจนานุกรมไม้ดอกไม้ประดับในเมืองไทย และสารนุกรมไม้ดอกไม้ประดับในเมืองไทย

3. คำค้น หมายถึง กลุ่มของอักขระ หรือ วลี ที่ระบุขึ้น เพื่อที่จะใช้นำมาเป็นข้อมูลสำหรับการค้นหาข้อมูลจากฐานข้อมูล

4. ประสิทธิภาพของของวิธีการค้นหาข้อมูล หมายถึง การวัดความสามารถของวิธีการค้นหาในด้าน ความครบถ้วนของข้อมูล ความถูกต้องของข้อมูล เวลาที่ใช้ในการค้นหาข้อมูล และ การใช้เนื้อที่ในการเก็บข้อมูลเพื่อการค้นหา

4.1 ด้านความครบถ้วนของข้อมูล หมายถึง สามารถค้นหาข้อมูลคืนมาได้ครบถ้วนตามที่มีข้อมูลอยู่จริง

4.2 ด้านความตรงต่อความต้องการ หมายถึง สามารถค้นหาข้อมูลคืนมาได้ตรงตามที่ต้องการ

4.3 ด้านเวลาที่ใช้ในการค้นหาข้อมูลเพื่อการค้นหา หมายถึง สามารถค้นหาข้อมูลคืนมาได้รวดเร็ว หรือไม่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.4 ด้านการใช้เนื้อที่ในการเก็บข้อมูล หมายถึง สามารถใช้เนื้อที่ในการค้นหาข้อมูลคืนมา โดยใช้เนื้อที่ของการเก็บเท่าใด

5. การเปรียบเทียบประสิทธิภาพของวิธีการค้นหาข้อมูล หมายถึง การนำเอาผลที่ได้จากการหาประสิทธิภาพที่ได้จากการทดลองของวิธีการของผู้วิจัย และวิธี Breath-First Search มาเปรียบเทียบกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2
เอกสารและงานวิจัยที่เกี่ยวข้อง

ในการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องกับงานวิจัยครั้งนี้ ขอนำเสนอโดยแยกรายละเอียดออกเป็นหัวข้อต่าง ๆ ดังนี้

2.1 การค้นหาข้อมูลสารสนเทศ

2.2 อัลกอริทึมเพื่อการค้นหาแบบโครงสร้างข้อมูล

2.3 อัลกอริทึมพื้นฐานเพื่อการค้นหาข้อมูลสารสนเทศ

2.4 อัลกอริทึมที่ใช้ในการค้นหาข้อมูลสารสนเทศ

2.5 ทฤษฎีที่ใช้เป็นพื้นฐานของแนวความคิดในวิธีการของผู้วิจัย

2.6 ความหมายและความสำคัญของไม้ดอกไม้ประดับ

2.7 โครงสร้างของทะเบียนพันธุ์ไม้ดอกไม้ประดับ

2.1 การค้นหาข้อมูลสารสนเทศ

ระบบการค้นหาข้อมูลถูกพัฒนาขึ้นเป็นครั้งแรกในช่วงปี ค.ศ. 1940 โดยมีองค์กร และหน่วยงานต่าง ๆ อาทิ กลุ่มของมหาวิทยาลัย บริษัท และหอสมุดสาธารณะใหญ่ ๆ เป็นผู้ริเริ่มและใช้ระบบการค้นหากับข้อมูลประเภทหนังสือ วารสาร และเอกสารต่าง ๆ ต่อมาได้มีการนำระบบค้นหานี้เข้าไปสู่ระบบธุรกิจซึ่งจะให้บริการฐานข้อมูลที่เก็บรวบรวมข้อมูลแต่ละประเภทไว้ ดำเนินการให้บริการในการค้นหา ระบบค้นหาที่ได้รับความนิยม และเป็นที่รู้จักกันเป็นอย่างดี คือ พจนานุกรมและสารานุกรมประเภทต่าง ๆ (B. Frakes และ Baeza-Yates. 1992 : 1) ผู้ใช้สามารถเลือกข้อมูลที่ต้องการตามลำดับอักษร เช่น "abstract" จะเปิดพจนานุกรมจาก ดัชนี "a" จากนั้นจึงดูตามลำดับที่อักษรของคำที่ต้องการ จนพบคำดังกล่าว

ในช่วงต่อมาในยุคที่เทคโนโลยีสารสนเทศเข้ามามีบทบาทกับชีวิตประจำวันมากขึ้น จึงมีการสร้างแนวคิดของระบบค้นหาเดิมมาสร้างโดยใช้เทคโนโลยี Korfhage and DeLutis (1969 : 31) ได้อธิบายความหมายของระบบสารสนเทศ (Information System) ว่าระบบสารสนเทศจะประกอบด้วย 2 ส่วนใหญ่ ๆ คือ เอ็กโตซิสเต็ม (Ectosystem) และ เอนโดซิสเต็ม (Endosystem) ซึ่งส่วนที่ผู้ออกแบบระบบสามารถสัมผัสได้คือส่วนของของเอนโดซิสเต็ม (Endosystem) เท่านั้น ส่วนเอ็กโตซิสเต็ม (Ectosystem) นั้นจะขึ้นอยู่กับปัจจัยต่าง ๆ ของระบบซึ่งอยู่นอกเหนือความสามารถของผู้ออกแบบ ปัจจัยต่าง ๆ นั้นอาจมาจากสาเหตุหลายประการ เช่นประเภทของข้อมูล อุปกรณ์ หรือเทคโนโลยีที่เอื้อต่อระบบสารสนเทศ เป็นต้น ในมุมมองของผู้ออกแบบระบบนั้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สามารถควบคุมส่วนของเอนโดซิสเต็ม (Endosystem) เช่น การระบุอุปกรณ์ (Equipment) อัลกอริทึม (Algorithms) และวิธีการ (Procedures) กับระบบที่ออกแบบได้

Baker (1968 : 16) กล่าวไว้ว่า เอ็กโตซิสเต็ม (Ectosystem) สามารถแบ่งได้เป็น 3 ส่วน คือ ผู้ใช้ (User) จะเป็นผู้ที่บันทึกข้อมูลเข้าสู่ระบบ หรือค้นหาข้อมูลจากระบบ ผู้สนับสนุน (Funder) เป็นบุคคลหรือองค์กรที่สนับสนุนทางด้านเงินลงทุนในการพัฒนาระบบ และผู้ให้บริการ (Servers) เป็นผู้ที่จัดการทางด้านบริการให้กับผู้ใช้

R. Korfhage. (1997 : 5 - 6) กล่าวไว้ว่า เอนโดซิสเต็ม (Endosystem) จึงเป็นส่วนที่ผู้ออกแบบจำเป็นต้องควบคุมให้ได้มากที่สุด ประกอบด้วย 4 ส่วน คือ สื่อ (Media) ที่นำมาบันทึกข้อมูล อุปกรณ์ (Device) สำหรับการประมวลผล อัลกอริทึม (Algorithm) เพื่อทำให้อุปกรณ์สามารถทำงานได้ และโครงสร้างข้อมูล (Data Structure) เพื่อการจัดการข้อมูล

การระมัดระวังในการเลือก อัลกอริทึมที่ใช้สำหรับการบันทึกข้อมูล ค้นหาข้อมูล และส่วนอื่น ๆ ที่ทำให้ข้อมูลมีสามารถในระดับต่าง ๆ กันระหว่างประสิทธิภาพ และประสิทธิผลของระบบสารสนเทศ เป็นส่วนสำคัญที่ควรให้ความสนใจพัฒนาระบบสารสนเทศให้มีประสิทธิภาพมากกว่าการพัฒนาจากการเลือกสื่อ (Media) หรือ อุปกรณ์ (Device) และการเลือกโครงสร้างข้อมูล (Data Structure) ก็เช่นกัน ผู้ออกแบบระบบควรให้ความสำคัญต่อการเลือกโครงสร้างข้อมูล (Data Structure) ให้เหมาะสมกับอัลกอริทึมสำหรับการค้นหา

2.2 อัลกอริทึมเพื่อการค้นหาแบบโครงสร้างข้อมูล

ประเด็นที่มีการโต้เถียงกันมากภายในกลุ่มของนักวิจัยคือ ความสัมพันธ์ระหว่างคิวรี (Query) กับเอกสาร (Document) นักวิจัยบางกลุ่มเชื่อว่าการนำเอาตำแหน่งที่คิวรี (Query) เพื่อแบ่งความแตกต่างระหว่างเอกสารที่ค้นหานั้น ไม่มีความจำเป็นที่จะพิจารณาเอกสาร (Bollmann-Sdorra and Raghavan. 1993 :579) โดยอ้างว่า ตามปกติคิวรีจะต้องมีขนาดสั้นกว่าเอกสารที่ถูกค้นหา เพราะเหตุว่าจำนวนเทอมเหมือนกันจะถูกแสดงเพียงแค่ครั้งเดียวในคิวรีโดยที่ความถี่ของการเกิดขึ้นของเทอมในเอกสารจะไม่มีผลต่อการค้นหา คิวรีไม่จำเป็นจะต้องอยู่ในรูปแบบของภาษาอย่างเอกสาร แต่ก็มีกระแสของนักวิจัยอีกกลุ่มเชื่อว่าความถี่ของการเกิดขึ้นของเทอมในเอกสารมีผลต่อการค้นหาข้อมูล (Harman 1995) ส่วนมากคิวรีหมายถึง หัวข้อ (Topic) ส่วนเอกสารจะเป็นเนื้อหาที่บรรยายละเอียด อย่างไรก็ดีการทำงานของคิวรีและเอกสารจำเป็นที่จะต้องมีการกำหนดโครงสร้าง

B. Frakes and Baeza-Yates (1992 : 18 – 23) อธิบายถึงโครงสร้างข้อมูล (Data Structure) สำหรับใช้จัดการข้อมูล คือ เซิชทรี (Search Tree) ดิจิตอลทรี (Digital Tree) และแฮช

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

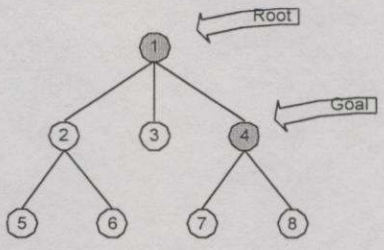
ชิ่ง (Hashing) ว่าไม่เป็นเพียงแค่การบันทึกข้อมูลลงในหน่วยความจำเสริม (Secondary Memory) เท่านั้น แต่เป็นส่วนประกอบของอัลกอริทึมในการค้นหาข้อมูล ไม่เพียงแต่สามารถบรรยายในรูปแบบของอาร์เรย์ (Array) เท่านั้น แต่สามารถบรรยายในรูปของโครงสร้างที่เป็นที่นิยม และรู้จักกันดี นำมาสร้างโดยตารางค้นหาแบบคงที่ (Static Search Table) เวกเตอร์แบบบิต (Bit Vector) ที่ใช้สำหรับการจัดการค้นหา หรือแม้กระทั่งซัฟฟิคอาร์เรย์ (Suffix Array) เป็นต้น

เซิชทรี (Search Tree) รูปแบบของเซิชทรีที่เป็นที่รู้จักกันเป็นอย่างดี คือ ไบนารีเซิชทรี (Binary Search Tree) ซึ่งแต่ละโหนด (Node) ถูกบรรจุด้วยคีย์ (Key) และที่ปลายด้านซ้ายของโหนดถูกบันทึกด้วยค่าคีย์ที่เล็กกว่าของทรีย่อย (Sub Tree) ค่าที่บันทึก คือคีย์ระดับก่อนหน้า (Parent Key) ส่วนอีกด้านจะบันทึกด้วยค่าคีย์ที่ใหญ่กว่าคีย์ระดับก่อนหน้า ข้อสังเกตคือ ไบนารีเซิชทรีต้องการเนื้อที่ของหน่วยความจำหลัก (Main Memory) ในการประมวลผล อย่างไรก็ตามสำหรับหน่วยความจำสำรอง (Secondary Memory) นั้น ยังมีวิธีค้นหาข้อมูลที่ดีกว่าเพราะว่าโหนดภายในมักจะใหญ่กว่า เช่น เดปเฟิร์สเซิช (Depth-First Search) เบรดเฟิร์สเซิช (Breath-First Search หรือ บีทรี (B-Tree) ซึ่งสามารถสร้างความสมดุลย์ภายในโครงสร้างได้ดีกว่าไบนารีเซิชทรี

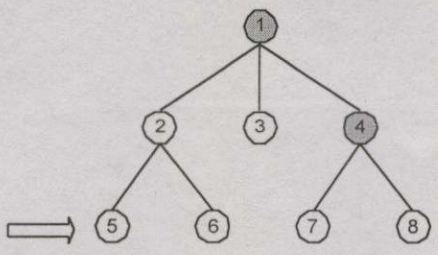
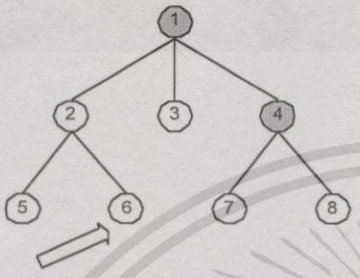
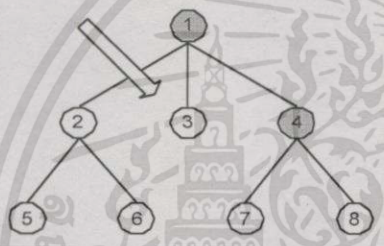
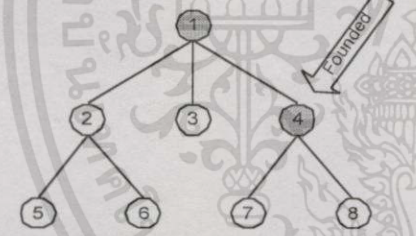
เดปเฟิร์สเซิช (Depth-First Search) เบรดเฟิร์สเซิช (Breath-First Search) ได้รับความสนใจมากสำหรับการนำมาใช้ในการเรียน และการศึกษาเพื่อนำมาใช้พัฒนางาน ไม่ว่าจะเป็นทางด้านโครงสร้างข้อมูล (Data Structure) โครงสร้างแฟ้ม (File Structure) หรือแม้กระทั่งทางด้านปัญญาประดิษฐ์ (Artificial Intelligent) ดังนั้นเนื้อหาส่วนนี้จึงมีอยู่มากสามารถสรุปได้ดังนี้ (Gilles Brassard and Paul Bratley. 1988. ; Mark J. Folk and Bill Zoellick. 1992 ; George F. Luger and William A. Stubblefield. 1993 ; Thomas A. Standish. 1994 ; Greg W. Scragg. 1997)

การค้นหาแบบ เดปเฟิร์สเซิช (Depth-First Search) เป็นวิธีหนึ่งในการค้นหาข้อมูลของกราฟ ซึ่งวิธีการที่ใช้นี้จะคล้ายคลึงกับวิธีการแบบพรีออเดอร์ (Preorder) คือจะค้นหาข้อมูลที่อยู่ตำแหน่งจุดยอดทางด้านลึกแล้วค่อย ๆ ย้อนกลับออกมาทีละระดับของแต่ละโหนด ตามทิศทางจากโหนดซ้ายสุดไปยังขวาสุด (Left-to-Right) ดังตัวอย่างที่จะแสดงต่อไปนี้

สมมติให้กราฟมีข้อมูล 1 2 3 4 5 6 7 และ 8 บรรจุอยู่ในแต่ละโหนด

	กำหนดให้ โหนดที่มีค่า 1 เป็น Root และโหนดที่มีค่า 4 เป็น Goal

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

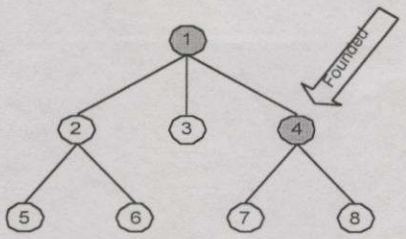
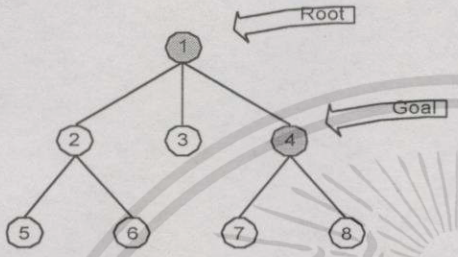
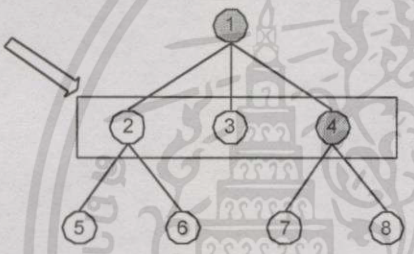
	การค้นหาจะผ่านโหนดที่มีค่า 1,2 และ 5 ซึ่งจะเห็นได้ว่าการค้นหาจะลงไปยังระดับสุดท้ายทางด้านซ้ายแล้ว
	เมื่อถึงระดับสุดท้ายแล้วก็จะพิจารณาในระดับเดียวกัน แล้วค่อย ๆ ย้อนกลับออกมา โดยในขั้นนี้จะมาพิจารณาที่โหนดที่มีค่า 6
	แล้วย้อนกลับออกมาในระดับก่อนหน้า มาพิจารณาในโหนดที่มีค่า 3
	เมื่อโหนดที่มีค่า 3 ไม่ใช่ค่าที่ต้องการก็ไปพิจารณาโหนดต่อไปในระดับเดียวกัน คือ โหนดที่มีค่า 4 ซึ่งเป็นข้อมูลที่ต้องการ การค้นหามีการผ่านทั้งหมด 6 โหนด

ภาพที่ 2.1 แสดงขั้นตอนการค้นหาข้อมูลแบบ Depth-First Search

การค้นหาแบบเบรดเฟิร์สเซิช (Breath-First Search) เป็นวิธีการที่พัฒนาขึ้นมาเพื่อช่วยลดข้อบกพร่องของวิธีการแบบเดปเฟิร์สเซิช (Depth-First Search) ซึ่งถูกนำมาพัฒนาจนเป็นที่รู้จักกันในนามอัลกอริทึมของ Dijkstra (Dijkstra's Algorithm) เพื่อช่วยลดเวลาในการค้นหาโดยค้นหาให้น้อยที่สุด (Shortest paths) คือจะเริ่มพิจารณาจากรากแรกก่อน โดยจะกวาดข้อมูลที่อยู่ในระดับเดียวกันก่อนถ้าในระดับไม่พบข้อมูลก็ค่อยก้าวลงไปทีละระดับ จนพบข้อมูล หรือจนถึงระดับสุดท้าย ซึ่งจะช่วยลดจำนวนเวลาในการค้นหาข้อมูลได้มากกว่าวิธีแบบเดปเฟิร์สเซิช (Depth-First Search) ดังตัวอย่างที่จะแสดงต่อไปนี้

สมมติให้กราฟมีข้อมูล 1 2 3 4 5 6 7 และ 8 บรรจุอยู่ในแต่ละโหนด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

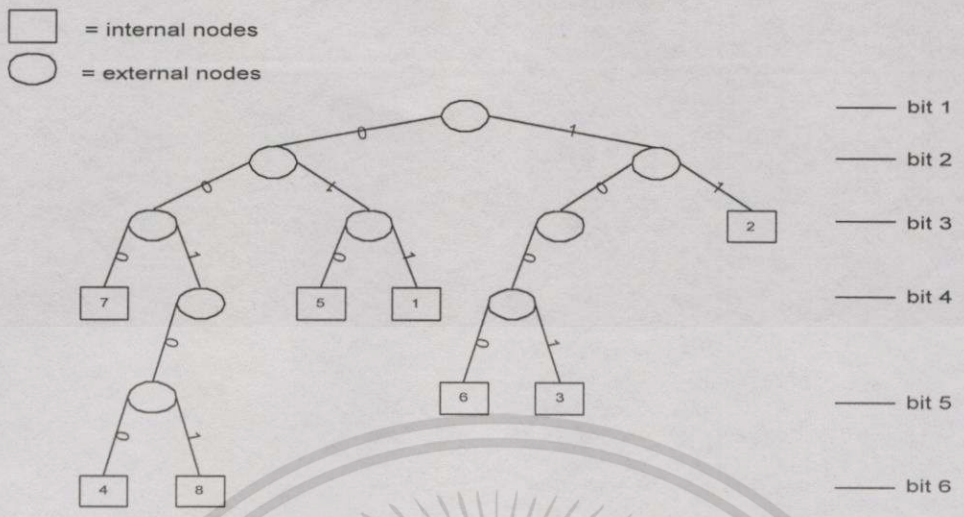
	กำหนดให้ โหนดที่มีค่า 1 เป็น Root และโหนดที่มีค่า 4 เป็น Goal
	พิจารณาในระดับแรก โหนดที่มีค่าเป็น 1 เป็นค่าที่ไม่ต้องการก็ลงลึกมาอีกระดับ พิจารณาในโหนดที่มีค่า 2,3 และ 4
	เมื่อโหนดที่มีค่า 2 และ 3 ไม่ใช่ค่าที่ต้องการก็ไปพิจารณาโหนดต่อไปในระดับเดียวกัน คือโหนดที่มีค่า 4 ซึ่งเป็นข้อมูลที่ต้องการ การค้นหามีการผ่านทั้งหมด 4 โหนด

ภาพที่ 2.2 แสดงขั้นตอนการค้นหาข้อมูลแบบ Breath-First Search

1. ดิจิตอลทรี (Digital Tree) การค้นหาข้อมูลให้มีประสิทธิภาพนั้น จำเป็นต้องใช้สิ่งที่เรียกว่าดัชนี (Indices) หนึ่งในดัชนีที่ดีที่สุดสำหรับการค้นหาคือ ไบนารีดิจิตอลทรี (Binary Digital Tree) สร้างจากกลุ่มของสตริงค์ย่อย (Sub-Strings) การใช้โครงสร้างข้อมูลวิธีนี้สามารถใช้อัลกอริทึมสำหรับค้นหาได้หลาย ๆ วิธี

ทไรซ์ (Tries) คือโครงสร้างทรีแบบย้อนกลับ (Recursive Tree Structure) ใช้ในการจำแนกคำที่ทำหน้าที่เสมือนกลุ่มของอักษรและใช้เพื่อการค้นหา ถูกประดิษฐ์ขึ้นโดย De La Briandais กล่าวถึงหลักการทำงานคือ ถ้าอักษรมีลำดับแล้วก็สามารถสร้างทรีในรูปแบบของพจนานุกรม (Lexicographically Ordered Tree) ที่มีลำดับได้ ค่าของราก (Root) ของทรีจะบรรจุอักษรแรกไว้ ส่วนลำดับถัดมาจากรากบรรจุอักษรในลำดับถัดไปเรื่อย ๆ นั่นหมายความว่าทรีแต่ละโหนดจะบรรจุเพียงอักษรเดียว ส่วนค่าของคำที่เหลือจะอยู่ในโหนดภายนอก ภาพที่ 2.3 แสดงไบนารี่ทรีของคำว่า "01100100010111..." หลังจากผ่านการบันทึกข้อมูลแล้ว เริ่มต้นจากตำแหน่งที่ 1 ถึงตำแหน่งที่ 8

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 2.3 แสดงไบนารี่ทรีของคำว่า "01100100010111..." หลังการบันทึกข้อมูลผ่านไป 8 ตัวอักษร

ความสูงของทรีจะเป็นค่าของโหนดในช่วงที่ยาวที่สุดนับจากรากถึงโหนดภายนอก (Eternal Node) โดยขนาดของแต่ละโหนดจากรากถึงโหนดภายนอกเชื่อมต่อกันเป็นความสูงของทรี เฉลี่ยความสูงของทรี เป็นลอการิทึมมิค (Logarithmic) ของ (Square-Integrable Probability Distribution) ใด ๆ สามารถกำหนดสมการเป็น

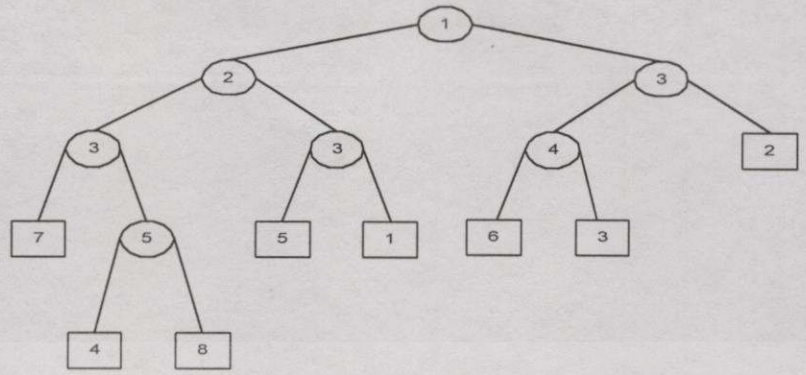
$$\chi_{(n)} = 2log_2(n) + O(log_2(n))$$

สำหรับไบนารีทรีที่บรรจุ n สตริงค์

การหาค่าเฉลี่ยของโหนดภายในที่ถูกใช้ไประหว่างการค้นหาในไบนารีทรีกับ n สตริงค์ เป็น $\log_2 n + O(1)$ จะมีค่าเป็น $(n/\ln 2) + O(n)$ (Knuth 1973)

แพททรีเซียทรี (Patricia tree) เป็นทรีที่มีความสามารถพิเศษในการกำจัดโหนดส่วนเกิน ย่อมาจาก "Practical Algorithm To Retrieve Information Code In Alphanumerical" มีหน้าที่ตรวจการเก็บค่าของแต่ละโหนดที่เชื่อมไปยังบิตต่อไป ดังภาพที่ 2.4 แสดงในรูปของแพททรีเซียทรี เพื่อเทียบกับไบนารีทรี

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 2.4 แสดงแพททรีเซียทรีของคำว่า "01100100010111..." หลังการบันทึกข้อมูลผ่านไป 8 ตัวอักษร

สำหรับ n สตริงค์ เช่นดัชนีที่มีโหนดภายนอกจำนวน n และโหนดภายในจำนวน n-1 นั้น แต่ละโหนดภายในเกิดขึ้นจากคู่ลำดับของพอยน์เตอร์กับตัวนับ ดังนั้น ความต้องการของเนื้อที่ในการค้นหาเป็น $O(n)$ ส่วนความน่าจะเป็นในการสร้างดัชนีใน $O(n\ \chi(n))$ ครั้ง ค่าของ $\chi(n)$ เป็นสิ่งที่สามารถบ่งบอกความสูงของทรีได้ แต่สำหรับแพททรีเซียทรีจะเป็นค่าของลอการิทึม เป็น $\log_2 n + O(\log_2 n)$

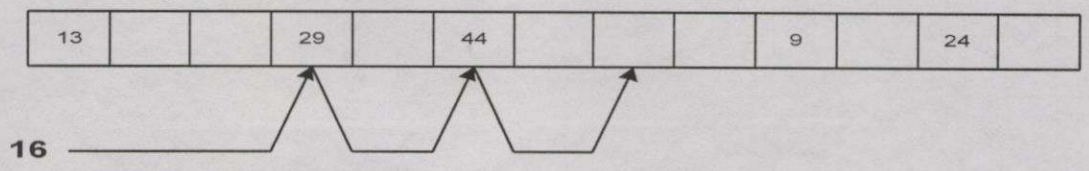
1. แฮชชิ่ง (Hashing)

ฟังก์ชันแฮชชิ่ง h(x) เป็นวิธีการเทียบค่า x เป็นค่าตัวเลขซึ่งบ่งบอกเป็นช่วง ถูกออกแบบขึ้นมาเพื่อสร้างค่าเฉพาะมีคุณสมบัติในการแบ่งออกเป็นช่วง ค่าเฉพาะดังกล่าว เรียกว่า ซิกเนเจอร์ (Signature)

หลักการทำงานของฟังก์ชันแฮชชิ่ง คือการเทียบค่าของกลุ่มคีย์ไปยังช่อง (Slot) ภายในตารางแฮชชิ่ง (Hashing Table) แต่ถ้าค่าที่ได้จากฟังก์ชันเกิดมีผลกับช่องเดียวกันโดยที่ค่าคีย์ต่างกันอาจเรียกว่าเกิดการชน (Collision) เหตุการณ์นี้สามารถเกิดขึ้นได้แต่แฮชชิ่งจะมีวิธีการที่สามารถจัดการกับปัญหานี้ 2 วิธี คือ การเปิดตำแหน่งเพิ่ม (Open Addressing) และการเลือกทับตำแหน่งเดิม (Overflow Addressing)

สำหรับการเปิดตำแหน่งเพิ่ม จะนำค่าคีย์ที่ชนคำนวณใหม่ โดยเข้าฟังก์ชันแฮชชิ่งอีกครั้ง (Rehashing) เพื่อหาค่าเฉพาะแล้วนำผลที่ได้มาเข้าช่องในตาราง ข้อจำกัดคือ จำเป็นต้องจัดโครงสร้างใหม่เมื่อตารางแฮชชิ่งเต็ม ตัวอย่างภาพที่ 2.5 กำหนดตารางแฮชชิ่งที่ขนาด 13 ช่อง และฟังก์ชันแฮชชิ่ง h(x) = x mod 13

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 2.5 แสดงเพิ่มค่า 16 ลงในตารางแฮชชิ่งโดยการเข้าฟังก์ชันแฮชชิ่งเพื่อหาช่องว่างในตาราง

ส่วนการเลือกทับตำแหน่งเดิม เมื่อเกิดการชนวิธีการนี้จะเลือกบันทึกคีย์ทับพื้นที่เดียวกัน เช่นค่าเฉพาะเกิดซ้ำกับค่าเดิมที่มีอยู่แล้ว ค่าใหม่จะเกิดการเชื่อมเข้าด้วยกันกับค่าเดิม ข้อจำกัดคือ เมื่อเกิดการเชื่อมต่อกันมากขึ้น การค้นหาข้อมูลอาจกลายไปสู่การค้นหาแบบเชิงเส้น (Linear Search) ซึ่งจะส่งผลให้ผลการค้นหามีอัตราที่ช้ากว่า

การปรับปรุงวิธีค้นหาจำเป็นต้องจัดหาวิธีอื่นเพื่อช่วยให้ความสามารถในการจัดการข้อมูลของแฮชชิ่งมีประสิทธิภาพ ซึ่งมีอยู่หลายวิธีเช่นการนำเอาทรีมาช่วยในการจัดการตารางแฮชชิ่งเพื่อป้องกันปัญหาดังกล่าวที่อาจเกิดขึ้นได้

2.3 อัลกอริทึมพื้นฐานเพื่อการค้นหาข้อมูลสารสนเทศ

B. Frakes and Baeza-Yates (1992 : 23 – 25) กล่าวถึงอัลกอริทึมพื้นฐานไว้ว่า เป็นเรื่องยากในการจำแนกอัลกอริทึมในการค้นหาข้อมูลสารสนเทศ ออกเป็นประเภทต่าง ๆ ตามประโยชน์ใช้สอย แต่สามารถแยกอย่างหยาบได้เป็น 3 ประเภท การแบ่งประเภทไม่ระบุถึงอัลกอริทึมที่ไม่มีผลต่อการประมวลผลในการค้นหา เช่น อัลกอริทึมที่เกี่ยวข้องกับผู้ใช้ (User Interface Algorithm) เพราะเหตุว่าอัลกอริทึมเหล่านี้จะยึดติดกับโปรแกรมคอมพิวเตอร์ที่ดำเนินการ

1. อัลกอริทึมค้นหา (Retrieval Algorithms)

ความสำคัญสำหรับอัลกอริทึมในการค้นหาข้อมูลสารสนเทศ คืออัลกอริทึมค้นหา ที่แบ่งข้อมูลจากฐานข้อมูลออกเพื่อการค้นหา สามารถแบ่งอัลกอริทึมนี้ออกเป็น 2 วิธี ตามความต้องการเนื้อที่ในการค้นหา คือ

การค้นหาโดยใช้ลำดับ (Sequential Scanning) ผลที่เกิดขึ้นในกรณีที่ร้ายแรงที่สุดของการค้นหาคือจะใช้เวลามากเท่ากับขนาดของคิวรี และขนาดของฐานข้อมูล แต่อีกกรณีทางด้านดี คือ สามารถค้นหาได้อย่างรวดเร็วโดยใช้เวลาลดลงตามสัดส่วนของอักษรที่ค้นหา

การค้นหาโดยใช้ดัชนี (indexed) สามารถเพิ่มความเร็วในการค้นหาข้อมูล ขนาดของดัชนีโดยปกติจะมีค่าเป็นสัดส่วนกับขนาดของฐานข้อมูล และเวลาในการค้นหาเป็นสัดส่วนย่อยระหว่างขนาดของอักษรที่ค้นหากับดัชนี

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตามปกติสามารถกำหนดสมการในการพิจารณาข้อผิดพลาดที่อาจเกิดจากการค้นหาข้อมูล กำหนดให้ สตริงค์ t, คิวรี q, และ ข้อมูลที่ถูกจัดให้อยู่ในรูปแบบที่สามารถค้นหาข้อมูลได้เป็น $t \in \Sigma^* q \Sigma^*$ และการพิจารณาค่าบางค่า หรือทั้งหมดจากสิ่งที่กำหนดให้ดังนี้

- ตำแหน่งซึ่งปรากฏขึ้นจากค่า q ที่ปรากฏ คือ ถ้า $t \in \Sigma^* q \Sigma^*$ แล้วค่าตำแหน่งที่ได้จะเป็น $m >= 0$ จาก $t \in \Sigma^* q \Sigma^*$
- จำนวนค่าที่ปรากฏขึ้นของรูปแบบสตริงค์ เป็นค่าผลบวกของ m ที่เกิดขึ้นจากลำดับก่อนหน้า
- ค่าตำแหน่งทั้งหมดเป็นค่ารูปแบบที่ปรากฏ คือกลุ่มของค่าผลบวกของ m

การแก้ไขข้อผิดพลาดต่าง ๆ โดยการเพิ่มประสิทธิภาพของการค้นหาเป็นสิ่งที่สำคัญมาก เพราะการหวังให้เกิดการค้นหาทันที (Online Query) โดยใช้เวลาน้อยที่สุดนั้น จำเป็นต้องได้รับการส่งเสริมอย่างจริงจังให้มีการสร้างอัลกอริทึมหลายวิธีขึ้น

2. อัลกอริทึมกรองข้อมูล (Filtering Algorithms)

การกรองข้อมูลจะเกิดขึ้นระหว่างที่เกิดการนำข้อมูลเข้าสู่ระบบ และการประมวลผลในการกรองสตริงค์จนเกิดผลลัพธ์ออกมาให้แก่ผู้ใช้ โดยใช้วิธีการแปลงค่า (Transformation) ข้อมูลของระบบสารสนเทศ เพื่อประโยชน์แก่การลดจำนวนขนาดที่บันทึก หรือเพื่อต้องการให้การค้นหาง่ายขึ้น การกรองข้อมูลขั้นพื้นฐานมีลักษณะดังต่อไปนี้

- กำจัดคำที่ไม่สมควรใช้ และไม่สื่อความหมายออกไป
- เปลี่ยนรูปแบบอักษรจากใหญ่ (Uppercase) ให้เป็นอักษรขนาดเล็ก (Lowercase)
- เปลี่ยนรูปแบบของตัวเลข และวันที่ให้อยู่ในรูปแบบเดียวกันทั้งหมด
- ตรวจสอบตัวสะกดให้ถูกต้อง
- สร้างดัชนี
- จัดลำดับของดัชนี

ข้อเสียของวิธีกรองข้อมูลคือ มักจะไม่เอื้อต่อการค้นหาข้อมูลที่มีความซับซ้อนสูง ๆ เช่น ถ้าต้องการระบุเงื่อนไขในการค้นหามาก ๆ โดยกำหนดให้ค้นหาข้อมูลเฉพาะตัวอักษรใหญ่ ค้นหาสัญลักษณ์พิเศษ หรือค้นหาคำเฉพาะที่ไม่มีความหมาย

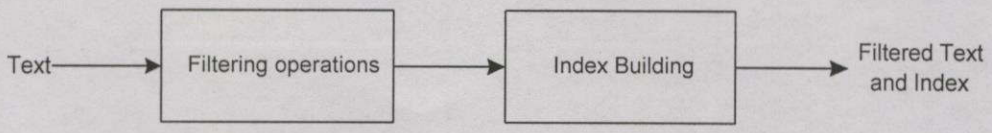
3. อัลกอริทึมดัชนี (Indexing Algorithms)

การใช้ประโยชน์จากความหมายของดัชนีคือ การสร้างโครงสร้างข้อมูลที่อนุญาตให้สามารถค้นหาข้อมูลได้ด้วยความรวดเร็ว ซึ่งมีด้วยกันหลายวิธีขึ้นอยู่กับความแตกต่างของผลลัพธ์ที่ต้องการ ซึ่งผู้ออกแบบระบบจะเป็นผู้ออกแบบตามความเหมาะสมเอง

โดยทั่วไปแล้วก่อนสร้างดัชนี ข้อมูลที่ได้จำเป็นต้องผ่านการกรองก่อน เพื่อให้เกิดการสร้างดัชนีที่มีประสิทธิภาพ ดังภาพที่ 2.6

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง



ภาพที่ 2.6 แสดงขั้นตอนการสร้างดัชนี

ถ้าต้องการคาดเดาเวลาที่ใช้สำหรับการค้นหาข้อมูล สามารถคำนวณได้จากเวลาที่ใช้ไปในการสร้างดัชนีหักลบด้วยความสามารถในการค้นหา เช่น ถ้าต้องการสร้างดัชนีโดยใช้เวลา $O(n \log n)$ แล้ว สามารถกำหนดเวลาที่ค้นหาน้อยที่สุดเป็น $O(\log n)$ เมื่อหักลบด้วยเวลาในการสร้างดัชนีเรียบร้อยแล้ว

2.4 อัลกอริทึมที่ใช้ในการค้นหาข้อมูลสารสนเทศ

อัลกอริทึมสำหรับการค้นหาข้อมูลมีอยู่มากมายหลายวิธี สำหรับงานวิจัยนี้จะนำเสนอ อัลกอริทึม 3 แบบด้วยกันดังนี้ คือ

1. เลกซิคอล อนาไลซิล (Lexical Analysis)

เลกซิคอล อนาไลซิล เป็นวิธีการประมวลผลด้วยการแปลงค่าอินพุทสตรีม (Input Stream) ของคำให้เป็นโทเคน (Token) วิธีการนี้มีส่วนประกอบ 2 ส่วนสำคัญ คือ ส่วนการสร้างดัชนี (Automatic Indexing) และส่วนการค้นหา (Query Processing) ซึ่งส่วนแรกเป็นการประมวลผลโดยอาศัยอัลกอริทึมที่เหมาะสมตามความต้องการของระบบโดยผู้ออกแบบระบบเป็นผู้กำหนดขึ้นเพื่อสร้างดัชนีขึ้นมา โทเคนอาจอยู่ในรูปแบบใด ๆ เพื่อความสะดวกต่อการค้นหาเช่น การใช้ตัวเลข (Digit) ยติภังค์ (Hyphen) วรรคตอนประเภทต่าง ๆ (Punctuation) หรือรูปแบบพิเศษที่ผู้ออกแบบกำหนดขึ้นใช้เองแทนค่าโทเคน เป็นต้น และส่วนต่อมาระบบจะแปลงคำค้นของผู้ใช้ให้อยู่ในรูปของโทเคนที่ตกลงกันไว้ แล้วจึงเปรียบเทียบโทเคนกับดัชนีเพื่อหาข้อมูลที่มีความสัมพันธ์กัน อาจใช้วิธีการเช่น บูลีน (Boolean) หรือฟังก์ชันน้ำหนัก (Weight Function) เป็นต้น

สำหรับวิธีการทำงานในการสร้างของอัลกอริทึมเลกซิคอล อนาไลซิล อัลกอริทึมประเภทนี้จะเริ่มต้นที่สถานะ 0 ซึ่งแต่ละอักษรที่ผ่านเข้ามาในอัลกอริทึมนี้จะปรับเปลี่ยนสถานะไปเรื่อย ๆ อัลกอริทึมจะอ่านค่าจนกระทั่งพบค่าพิเศษ ซึ่งค่าพิเศษที่กำหนดขึ้นจะเป็นเครื่องมือในการแบ่งแยกแต่ละโทเคนออกจากกัน หลังจากนั้นจะกลับไปอ่านค่าอักษรอื่นต่อไป สถานะสุดท้ายเป็นการเปลี่ยนรูปแบบโทเคนโดยใช้เครื่องหมายพิเศษที่กำหนดเพื่อบ่งบอกให้ทราบว่าจบคำแล้ว

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

40790

โดยสรุปอัลกอริทึมเลกซิคอล อนาไลซิล จำเป็นต้องใช้ดัชนีเพื่อการค้นหาข้อมูล การตัดสินใจต่อดัชนีของรูปแบบเลกซิคอล (Lexical Form) และเทอมคิวรี ส่วนประมวลผลคิวรีและวิธีการเปรียบเทียบ ในการออกแบบบนพื้นฐานของคุณสมบัติฐานข้อมูลและจุดมุ่งหมายของการใช้ระบบ สิ่งแรกที่ควรเริ่มกระทำคือขั้นการประยุกต์วิธีการต่าง ๆ ให้อยู่ในรูปแบบของโปรแกรม ในขั้นต่อมาคือเครื่องมือสำหรับเลกซิคอลที่สร้างขึ้นควรมีความน่าเชื่อถือไม่ว่าจะถูกสร้างขึ้นด้วยมือหรือใช้เครื่องมืออื่น ๆ สร้างขึ้น

2. อัลกอริทึมสเตมมิ่ง (Stemming Algorithms)

อัลกอริทึมสเตมมิ่งเป็นวิธีการที่พัฒนาขึ้นเพื่อเพิ่มความสามารถให้แก่ผู้ใช้ในการค้นหาคำให้มีประสิทธิภาพยิ่งขึ้น โดยหลักการ อัลกอริทึมจะพิจารณาคำที่มาจากรากศัพท์เดียวกันให้อยู่ในดัชนีเดียวกัน เช่น “Engineering” “Engineered” ”Engineer” ทั้งหมดมาจากรากศัพท์เดียวกัน ดังนั้นการเก็บข้อมูลในดัชนีจึงเก็บ "Engineer" เพียงคำเดียว มีขั้นตอนดังนี้

- ใช้วิธีการตัดคำ (Cutoff Method) เพื่อตัดคำขยายต่าง ๆ ของรากศัพท์ออกไป
- ใช้วิธีการปรับระดับคำ (Peak and Plateau Method) ปรับแต่งคำที่อาจมีเครื่องหมายพิเศษ หรือช่องว่างต่าง ๆ ออกไป
- ใช้วิธีการตรวจสอบคำ (Complete Word Method) ตรวจสอบความถูกต้องของคำที่ได้กับรากศัพท์จริง
- นำรากศัพท์ที่ได้มาสร้างเป็นดัชนี

สรุปสเตมเมอร์ (Stemmer) ถูกใช้เพื่อจัดกลุ่มของเทอม เพื่อปรับปรุงการค้นหาให้มีประสิทธิภาพ และ/หรือช่วยลดขนาดของดัชนี สเตมมิ่งที่ดีจะต้องมีการกำหนดรูปแบบ มีวิธีการที่แน่นอนไม่ว่าจะเป็นส่วนการเพิ่มหรือลดดัชนี และมีขนาดเป็นสัดส่วนต่อข้อมูลที่ร้อยละ 50 ซึ่งอัลกอริทึมนี้อาจสร้างขึ้นมาจากวิธีการต่าง ๆ ได้หลายวิธีขึ้นอยู่กับความสามารถของผู้ออกแบบระบบ เช่น ตารางเทียบ (Table Lookup) เอ็นแกรม (N-Gram) เป็นต้น

3. อัลกอริทึมสตริงค์เซิชชิ่ง (String Searching Algorithms)

การเปรียบเทียบสตริงค์ที่ไม่มีประสิทธิภาพอาจเป็นสาเหตุของปัญหามากมาย วิธีการเปรียบเทียบสตริงค์ขั้นพื้นฐานคือการเปรียบเทียบไปเรื่อยๆ สมมติให้ คำค้นมีขนาด m เพื่อเทียบกับข้อมูล n (ซึ่ง $n > m$) เวลาในการค้นหาจะเป็น $O(n)$ ในกรณีร้ายแรงที่สุด (ที่ค่า m คงที่) แต่อย่างไรก็ตามกรณีร้ายแรงที่สุดอาจค้นพบในเวลาน้อยที่สุดที่ค่า $n - m + 1$ อักษร ซึ่งวิธีนี้ถูกเรียกว่า การค้นหาเชิงเส้น (Linear Searching) ดังนั้นจึงมีผู้สนใจพัฒนาวิธีการค้นหาข้อมูลอัลกอริทึมนี้ให้มีความสามารถมากขึ้น

Baeza-Yates and Gonnet (1989 : 257 – 271) นำเสนออัลกอริทึมสเตมมิ่ง ด้วยการใช้วิธีการแบบชิฟออร์ (Shift-Or Algorithm) ซึ่งมีรูปแบบการทำงานเป็นการแทนค่าของสถานะด้วย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตัวเลข ขั้นตอนการค้นหาแต่ละครั้งจะใช้ตัวเลขที่ค่าน้อยและการเปรียบเทียบแบบออร์ (Or) ทางตรรกะและขยับ (Shift) ค่าสถานะ การจัดการขนาดตัวเลขให้เพียงพอต่อการแทนค่าสถานะของอักษรทั้งหมด ดังนั้น สมการที่ได้จาก $O(n)$ ครั้ง จะอยู่ในรูป $O(|\Sigma|)$ สำหรับค่าตัวเลขที่เพิ่มขึ้น และ $O(m + |\Sigma|)$ สำหรับเวลาในการจัดการก่อนการค้นหาที่ Σ กลุ่มของอักษร คุณสมบัติหลักของอัลกอริทึมนี้ คือ ง่ายเพราะไม่ว่าจะเป็นขั้นตอนการสร้างและค้นหาใช้ตัวเลขแทนค่าขนาด รวดเร็วเพราะเวลาสามารถควบคุมได้ และไม่มีความต้องการต่อการใช้เนื้อที่เนื่องจากไม่มีการบันทึกข้อมูลลงหน่วยความจำ

อัลกอริทึมสเตมมิ่ง ด้วยการใช้วิธีการแบบซิฟออร์นี้สร้างขึ้นโดยใช้พื้นฐานของทฤษฎีไฟไนท์ ออโต้มาตา (Finite Automata Theory) โดยใช้เวกเตอร์ m แทนสถานะที่ต่างกัน สถานะ i บ่งบอกถึงตำแหน่งของการค้นหาตั้งแต่ $1,\ldots,i$ ของรูปแบบและตำแหน่ง $(j - i + 1),\ldots,j$ ของข้อมูล สถานะ j บ่งบอกตำแหน่งที่ของอักษรในข้อมูล การแทนค่าจะใช้ 1 บิต แทนสถานะแต่ละสถานะ และเมื่อสถานะมีค่าเป็น 0 หรือ 1 ก็ตาม โดยค่า i เป็นตำแหน่งสุดท้ายของการค้นหาแล้ว สมการเวกเตอร์จะมีค่าเป็น

$$state = \sum_{i=0}^{m-1} s_{i+1} 2^i$$

เมื่อ s_i หมายถึงสถานะ การเทียบตรงจะพบถ้า s_m มีค่าเป็น 0 หรือมีค่าเท่ากัน หรือสถานะ $< 2^{m-1}$ การเทียบจะสิ้นสุดลงที่ตำแหน่งปัจจุบัน การปรับสถานะหลังจากอ่านค่าอักษรใหม่เข้ามามีขั้นตอนดังนี้ คือ

- เลื่อนสถานะเวกเตอร์ 1 บิต ทางด้านซ้ายออกและได้สถานะใหม่มากจากอักษรที่อ่านเข้ามา การตั้งค่าสถานะแรกของ s_i มีค่าเป็น 0 เสมอ
- การเปลี่ยนแปลงสถานะจะขึ้นอยู่กับค่าของอักษรใหม่ โดยใช้ตาราง T ในการ
- จัดการรูปแบบของอักษร และใช้วิธีการ *or* เทียบกับเวกเตอร์เก่าและค่าในตาราง T เพื่อให้ได้ค่าเวกเตอร์สถานะใหม่

ดังนั้นการค้นหาแต่ละครั้งจะมีค่า

$$state = (state << 1)\ or\ T[curr\ char]$$

ซึ่งค่า << แทนการเลื่อนไปด้านซ้าย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

นิยามของตารางคือ

$$T_x = \sum_{i=0}^{m-1} \delta(pat_{i+1} = x)2^i$$

สำหรับทุกสัญลักษณ์ x ของกลุ่มอักษร เมื่อ $\delta(C)$ มีค่าเป็น 0 ถ้าเงื่อนไข C เป็นจริง และ 1 เมื่อเป็นเท็จ ในส่วนนี้ต้องการเนื้อที่ขนาด m • $|\Sigma|$ การค้นหาจะใช้เวลาทั้งสิ้น $O(\left\lceil \frac{m}{w} \right\rceil (m+|\Sigma|))$ ครั้ง

ตัวอย่าง สมมติให้ {a,b,c,d} เป็นกลุ่มของอักษรและ ababc เป็นคำค้น เมื่อจัดเรียงข้อมูลเข้าตาราง (1 บิตต่อตำแหน่งในรูปแบบ) จะได้

T[x] T[a] = 11010 T[b] = 10101 T[c] = 01111 T[d] = 11111

เมื่อมีการค้นหาในครั้งแรกในดัชนี abdababababc จะเริ่มที่สถานะ 11111

ดัชนี	A	B	D	A	B	A	b	a	b	c
T[x]	11010	10101	11111	11010	10101	11010	10101	11010	10101	01111
สถานะ	11110	11101	11111	11110	11101	11010	10101	11010	10101	01111

จากตัวอย่างสถานะ 10101 หมายถึงในตำแหน่งปัจจุบันมีการเทียบตรงใน 2 ส่วนเกิดขึ้นทางด้านซ้าย ที่มีขนาด 2 และ 4 นับจากด้านขวา การเทียบตรงที่สุดท้ายของคำค้นแสดงว่าค่าที่ได้มีขนาด 0 หมายความค่าสิ้นสุดคำค้นแล้ว

2.5 ทฤษฎีที่ใช้เป็นพื้นฐานของแนวความคิดในวิธีการของผู้วิจัย

Richard Entlich และ Jan Olsen (1997) แห่ง Cornell University ได้ทำการศึกษาและพัฒนาการสร้างโปรแกรมห้องสมุดดิจิตอล โดยการพัฒนาระบบทรี (Tree) แต่ใช้โมเดลแบบ CORE (Chemical Online Retrieval Experiment) เข้ามาช่วย และพัฒนาบนระบบภายในองค์กรเป็น Unix-Based Client-Server เพื่อให้ผู้ใช้เข้ามาค้นหาข้อมูล โปรแกรมนี้มีชื่อว่านิวตัน (Newton) ซึ่งโปรแกรมนี้จะใช้วิธีการค้นหาข้อมูลแบบทรีกับข้อมูลที่มีความซับซ้อน ซึ่งประกอบด้วย ISSN Title Sentences Subheads ทำให้ทราบได้ว่าการสร้างวิธีการทางด้านการค้นหาข้อมูลโดยใช้โมเดลแบบ CORE ช่วยให้การค้นหามีประสิทธิภาพดีกว่าการค้นหาพื้นฐานแบบ Flat ASCII แต่ใช้เนื้อที่ในส่วนของดัชนีที่ต้องถูกโหลดขึ้นหน่วยความจำค่อนข้างมาก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Mark Lager (1996) แห่ง California Lutheran University ศึกษาถึงวิธีการค้นหาข้อมูลบนอินเตอร์เน็ต ทำให้ทราบได้ว่าเครื่องมือที่ใช้สำหรับค้นหาข้อมูลบนอินเตอร์เน็ตมี 2 ประเภทด้วยกันคือ Index และ Search Engines ซึ่งทั้ง 2 อย่างนี้ต้องทำงานไปด้วยกัน คือ Index จะสร้างดัชนีขึ้นมาเพื่อให้ Search Engines เข้ามาค้นหาข้อมูลที่เป็นดัชนี และวิธีการที่ AltaVista Excite Infoseek Lycos Magellan และ OpenText ใช้ค้นหาจะเป็นทั้งแบบ Depth-First search และ Breath-First Search ผสมผสานกัน

Ellman, J และ Tail, J. (1997) แห่ง Sunderland University ได้ศึกษาถึงการพิจารณาข้อมูลเพื่อใช้ในการค้นหาข้อมูลผ่านเว็ป พบว่าการใช้วิธีการธรรมชาติสำหรับมนุษย์ในการค้นหาข้อมูลนั้น จะช่วยให้ผู้ใช้สามารถกำหนดสิ่งที่ตนเองต้องการรู้ (Keyword) และค้นหาข้อมูลที่ตนเองต้องการได้เป็นที่น่าพอใจ ตรงกับความต้องการ และในงานลักษณะนี้ต้องอาศัยการสร้างโดยใช้โปรแกรมภาษาที่สามารถสนับสนุนการทำงานในสถาปัตยกรรมแบบเครือข่ายให้มากขึ้น ซึ่ง JAVA สามารถสนับสนุนการทำงานในสถาปัตยกรรมแบบเครือข่ายได้อย่างมีประสิทธิภาพ

G. Michael Youngblood (1999) ได้ทำการศึกษาและนำเสนอข้อมูลเรื่อง Web Haunting : Design of a Simple Intelligent Web Search Agent โดยงานชิ้นนี้เขาได้นำเสนอแนวทางในการสร้างอัลกอริทึมในการค้นหาและต้องการให้มีการศึกษาวิจัยเพิ่มที่ส่วนของวิธีค้นหาโดยใช้วิธีการของปัญญาประดิษฐ์ (Artificial Intelligent) สามารถสรุปได้ดังนี้

วิธีการค้นหาบนอินเตอร์เน็ตนั้น ควรจะมีการใช้วิธีการทางด้านปัญญาประดิษฐ์ให้มากขึ้นโดยคำนึงถึงโครงสร้างของข้อมูลที่จะนำเสนอ พิจารณาจากข้อมูลว่าเราสามารถรู้เป้าหมาย (Goal) ของผู้ใช้ได้หรือไม่ ถ้าไม่สามารถรู้ได้เลยว่าผู้ใช้ต้องการอะไรให้ใช้วิธีการของปัญญาประดิษฐ์ทางในส่วนวิธีการ Uninformed แต่ถ้าสามารถรู้ได้ว่าผู้ใช้ต้องการอะไรให้ใช้วิธีการในส่วน Informed

วิธีการค้นหาแบบ Uninformed เป็นวิธีการค้นหาที่โปรแกรมไม่สามารถรู้เป้าหมาย (Goal) ของผู้ใช้มาก่อนว่าผู้ใช้ต้องการค้นหาอะไร โปรแกรมจะสร้างโครงสร้างของกราฟเป็นรูปแบบทรี (Tree) ขึ้นมาเพื่อช่วยในการค้นหา โดยวิธีการประเภทนี้จะมีการค้นหาอยู่ 2 ประเภทหลัก คือ Depth-First Search และ Breath-First Search

วิธีการค้นหาแบบ Informed เป็นวิธีการค้นหาที่โปรแกรมสามารถรู้เป้าหมาย (Goal) ของผู้ใช้ได้ ได้แก่งานที่ประยุกต์ในการสร้างเกมส์ต่าง ๆ เช่น ทิคแทคโท (Tic-Tac-Toe) ที่มีช่องสำหรับเล่นอยู่ 9 ช่อง โปรแกรมสามารถสร้างเส้นทางได้ว่าก่อนเลยว่าควรจะเล่นอย่างไรจึงจะปลอดภัย และสามารถชนะได้เร็วที่สุด หรือหมากรุก เป็นต้น วิธีการค้นหานั้นโปรแกรมจะสร้างเส้นทางขึ้นมาตั้งแต่ต้นจนจบ ทำให้โปรแกรมทราบว่าเส้นทางที่ใช้สามารถเอาชนะคู่ต่อสู้ได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

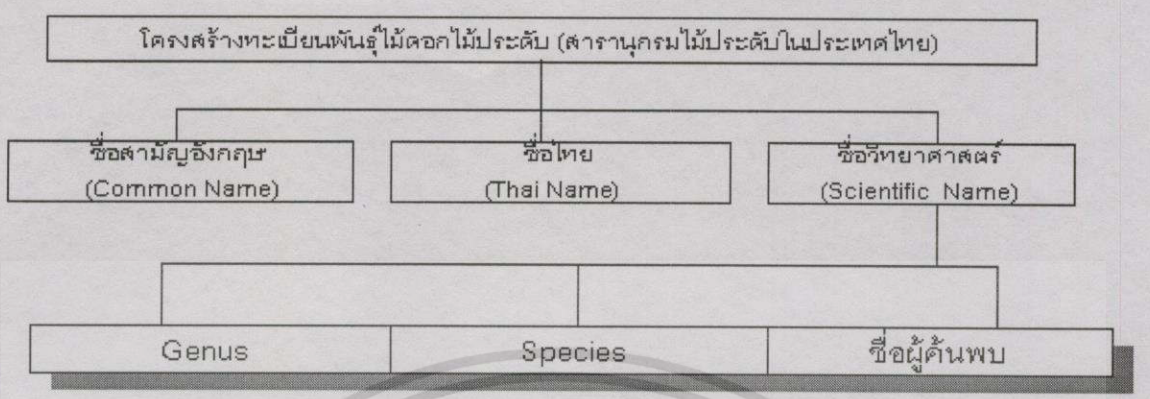
จากการศึกษางานวิจัยต่าง ๆ พอจะสรุปได้ว่า เนื่องจากในปัจจุบันยังไม่ค่อยมีผู้ให้ความสนใจกับการพัฒนาวิธีการค้นหาข้อมูล การค้นหาข้อมูลในปัจจุบันเราจะเห็นกันโดยทั่วไปแล้วว่าจะสร้างขึ้นมาโดยอาศัยเครื่องมือที่ใช้กันอยู่ทั่วไป เช่น ผ่านทาง ODBC DBE เป็นต้น การใช้ข้อมูลผ่านทางเครื่องมือเหล่านี้มีประสิทธิภาพสูงและเหมาะสมกับการทำงานในลักษณะขององค์กรมาก แต่เมื่อนำมาประยุกต์ใช้บนอินเตอร์เน็ตแล้ว จะมีข้อยุ่งยาก และน่ารำคาญอยู่มากเพราะจะต้องคำนึงถึงองค์ประกอบหลาย ๆ อย่างเช่น เนื้อที่โฮมเพจที่ได้รับมา หน่วยความจำของเครื่องเซิร์ฟเวอร์ที่ให้บริการ เนื้อที่ที่ใช้เก็บข้อมูลของเซิร์ฟเวอร์ที่ให้บริการ ซึ่งการค้นหาข้อมูลผ่านทางไดร์เวอร์นี้ก็ยังคงวิธีการพื้นฐานอยู่ คือการเทียบแบบ ASCII นั่นเอง ฉะนั้นทิศทางของงานวิจัยประเภทนี้ จึงจำเป็นที่จะต้องอาศัยวิธีการทางด้านปัญญาประดิษฐ์ (Artificial Intelligent) เพื่อช่วยพัฒนาประสิทธิภาพในการค้นหา โดยเฉพาะกับข้อมูลเทคนิคที่มีลักษณะเฉพาะ เช่น ข้อมูลทางด้านเคมี เกษตร แพทย์ หรือแม้กระทั่งข้อมูลทางด้านคอมพิวเตอร์เองก็ตาม ซึ่งวิธีการทางด้านนี้จะเริ่มจากการใช้วิธีแบบ Depth-First Search แต่มีข้อจำกัดอยู่มากจึงมีการพัฒนาวิธีการแบบ Breath-First Search ขึ้นมาเพื่อช่วยลดปัญหาของวิธีเดิมแต่วิธีการแบบ Breath-First Search นี้มีข้อจำกัดคือ ต้องการการเข้าถึงหน่วยความจำเป็นจำนวนมาก

2.6 ความหมายและความสำคัญของไม้ดอกไม้ประดับ

ประสิทธิ์ คำภูแสน (2520 : 1-2) กล่าวไว้ว่า ไม้ดอกไม้ประดับหมายถึง พันธุ์ไม้ที่มนุษย์เรานำเอามาเป็นเครื่องประดับบ้าน อาคารสถานที่ต่าง ๆ เพื่อชมความงามโดยคำนึงถึง ความงามของรูปร่างความสวยงามของทรวดทรง ดอก ผล ลำต้น ใบ เป็นสิ่งสำคัญ มนุษย์เป็นผู้ปลูกดูแลรักษา เข้าใจการผสมพันธุ์ใหม่ ๆ การขยายพันธุ์ การแยกไม้ดอกไม้ประดับออกเป็นหมู่เหล่า เพื่อให้เข้าใจตรงกัน เพราะมีความสำคัญและผูกพันกับมนุษย์เราจนแยกกันไม่ออก ตลอดจนปัจจัย 4 ซึ่งเป็นความสำคัญโดยตรงต่อมนุษย์เพื่อประกอบการดำรงชีพ ไม้ดอกไม้ประดับนั้นมีคุณค่าสูง เป็นที่อยู่อาศัย อาหาร เครื่องนุ่งห่ม ยารักษาโรค นอกจากนี้ต้นไม้ยังให้อากาศบริสุทธิ์แก่มนุษย์ โดยที่ใบไม้ดูดแก๊สคาร์บอนไดออกไซด์ที่มนุษย์หายใจออกมาในเวลากลางวัน เพื่อใช้ในการสังเคราะห์แสง แล้วคายออกซิเจนออกมาซึ่งก็กลายเป็นอากาศดีให้มนุษย์หายใจหมุนเวียนอยู่เช่นนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.7 โครงสร้างของทะเบียนพันธุ์ไม้ดอกไม้ประดับ



ภาพที่ 2.6 โครงสร้างทะเบียนพันธุ์ไม้ดอกไม้ประดับที่ใช้ในประเทศไทย

การใช้ชื่อพืชประดับ Species จะถือตาม "Binomial System" จะประกอบด้วย 2 คำเสมอ คำแรกเป็นชื่อสกุล (Genus) คำหลังเรียกว่า Specific Epithet ซึ่งมักเป็นคุณศัพท์แสดงคุณลักษณะสถานที่ และอื่น ๆ และขึ้นต้นด้วยตัวอักษรเล็ก เช่น Tagetes Tenuifolia Pumila สามารถอธิบายได้ดังนี้ ชื่อสกุลคือ "Tagetes" ชื่อ Specific Epithet คือ "Tenuifolia ผู้ค้นพบพืชนี้คือ "Pumila" (กองส่งเสริมพืชพันธ์. 2535 : 68)

การเรียกชื่อทางวิทยาศาสตร์จะประกอบด้วย Genus ก่อนตามด้วย Species การพิจารณาว่าพืชชนิดใดอยู่ในหมวดชนิดไหนนั้นจะต้องใช้ Key เป็นเครื่องพิจารณา การใช้ Key เปรียบเหมือนแนวทางให้เราพิจารณาไปสู่จุดหมายได้ ว่าพืชมีลักษณะอย่างไร โดยพิจารณารายละเอียดของ Genus และ Species ในบางครั้งจะพบชื่อตอนท้ายนั้น อาจเป็นชื่อ เมือง ถิ่นกำเนิดของพันธุ์ไม้นั้น หรือชื่อผู้ค้นพบ เช่น Pumila เป็นการให้เกียรติแก่ผู้พบ

ชื่อทางวิทยาศาสตร์นั้นเป็นที่ยอมรับกันในหมู่นักพฤกษศาสตร์ทั่วโลก เพราะ มีความหมายถึงพืช ชนิดเดียวกัน ทุกคนสามารถทราบได้ว่าชื่อนั้น หมายถึงพืชอะไร ชนิดใดได้ถูกต้องและตรงกันหมด (ประสิทธิ์ คำภูแสน, 2520 : 32-33)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3
วิธีดำเนินการวิจัย

ในการศึกษาวิจัยเรื่อง "การพัฒนาวิธีการค้นหาข้อมูลบนอินเตอร์เน็ตกับข้อมูลทะเบียนพันธุ์ไม้ดอกไม้ประดับ" ผู้วิจัยขอเสนอวิธีดำเนินการวิจัยดังนี้

3.1 ประชากรและกลุ่มตัวอย่าง

3.2 เครื่องมือที่ใช้ในการวิจัย

3.3 การดำเนินการทดลอง

3.4 การวิเคราะห์ข้อมูล

3.1 ประชากรและกลุ่มตัวอย่าง

1. ประชากรที่ใช้ในการวิจัยคือ รายชื่อพันธุ์ไม้ดอกไม้ประดับในเมืองไทย จากพจนานุกรมไม้ดอกไม้ประดับในเมืองไทย (วิทย์ เที่ยงบูรณธรรม. 2536) และสารานุกรมไม้ดอกไม้ประดับในประเทศไทย (สวัสดิ์ หรั่งเจริญ. 2525) จำนวน 2,000 ชื่อ

2. กลุ่มตัวอย่างที่ใช้ในการวิจัย คือ รายชื่อพันธุ์ไม้ดอกไม้ประดับในเมืองไทย จากพจนานุกรมไม้ดอกไม้ประดับในเมืองไทย (วิทย์ เที่ยงบูรณธรรม. 2536) และสารานุกรมไม้ดอกไม้ประดับในประเทศไทย (สวัสดิ์ หรั่งเจริญ. 2525) จำนวน 1,849 ชื่อ โดรเริ่มจาก 100 ชื่อ และทำการเพิ่มครั้งละ 20 เปอร์เซ็นต์ จำนวน 17 ครั้ง ได้ชื่อ 1849 ชื่อ

3.2 เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัยครั้งนี้ คือ

โปรแกรมการค้นหาข้อมูลบนอินเตอร์เน็ตกับข้อมูลทะเบียนพันธุ์ไม้ดอกไม้ประดับวิธีการของผู้วิจัย และวิธี Breath-First Search

3.2.1 การสร้างเครื่องมือที่ใช้ในการวิจัย

โปรแกรมการค้นหาข้อมูลบนอินเตอร์เน็ตกับข้อมูลทะเบียนพันธุ์ไม้ดอกไม้ประดับ

มีขั้นตอนการสร้างดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1. **วิเคราะห์คุณสมบัติของทะเบียนพันธุ์ไม้ดอกไม้ประดับ**

ชื่อไทยของพืชในทะเบียนพันธุ์ไม้ดอกไม้ประดับ จะใช้ตัวอักษรต่าง ๆ ในภาษาไทยประกอบกัน ซึ่งอักษรในภาษาไทยมี 3 ประเภท คือ พยัญชนะ สระ และวรรณยุกต์ โดยพยัญชนะประกอบไปด้วยตัวอักษร ก ถึง ฮ รวมทั้งหมด 44 ตัวอักษร แต่ในปัจจุบันมี 2 ตัวอักษรที่เลิกใช้แล้วคือ ฃ และ ฅ จึงเหลือที่ใช้อยู่จริง ๆ ทั้งหมด 42 ตัวอักษร ดังภาพที่ 3.1

ก ข ฃ ค ฅ ฆ ง จ ฉ ช

ซ ฌ ญ ฎ ฏ ฐ ฑ ฒ ณ ด

ต ถ ท ธ น บ ป ผ ฝ พ

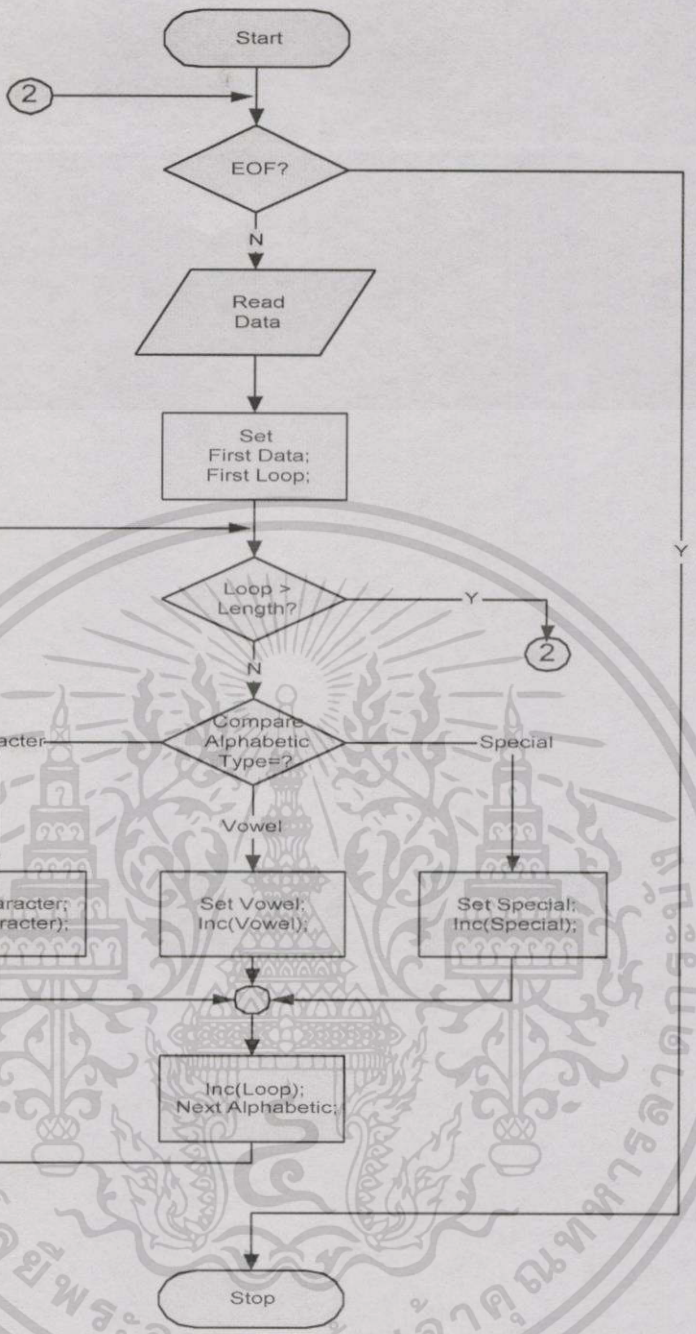
ฟ ภ ม ย ร ล ว ศ ษ ส

ห ฬ อ ฮ

ภาพที่ 3.1 แสดงตัวพยัญชนะภาษาไทย

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 3.2 แสดงขั้นตอนการวิเคราะห์คุณสมบัติของข้อมูลทะเบียนพันธุ์ไม้ดอกไม้ประดับ

เมื่อทราบถึงพยัญชนะไทยที่ถูกใช้อยู่จริงในปัจจุบันแล้ว จึงนำมาวิเคราะห์หาคุณสมบัติของชื่อไทยของพืชในการใช้พยัญชนะต่าง ๆ โดยวิธีการสุ่มอย่างง่ายมา 100 รายชื่อ ทำให้ทราบถึงคุณสมบัติของการใช้พยัญชนะไทย ได้ดังตารางที่ 3.1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 3.1 แสดงความถี่การใช้พยัญชนะในข้อมูลทะเบียนพันธุ์ไม้ดอกไม้ประดับ

ลำดับที่	พยัญชนะ	ความถี่	ร้อยละ
1	ก	50	8.61
2	ข	19	3.27
3	ฃ	-	-
4	ค	20	3.44
5	ฅ	-	0
6	ฆ	0	0
7	ง	45	7.75
8	จ	24	4.13
9	ฉ	4	0.69
10	ช	11	1.89
11	ซ	10	1.72
12	ฌ	0	0
13	ญ	3	0.52
14	ฎ	0	0
15	ฏ	0	0
16	ฐ	0	0
17	ฑ	0	0
18	ฒ	2	0.34
19	ณ	3	0.52
20	ด	30	5.16
21	ต	24	4.13
22	ถ	2	0.34
23	ท	12	2.07
24	ธ	4	0.69
25	น	43	7.40
26	บ	14	2.41
27	ป	6	1.03
28	ผ	3	0.52

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 3.1 (ต่อ)

ลำดับที่	พยัญชนะ	ความถี่	ร้อยละ
29	ฝ	4	0.69
30	พ	33	5.68
31	ฟ	2	0.34
32	ภ	1	0.17
33	ม	28	4.82
34	ย	20	3.44
35	ร	54	9.29
36	ล	18	3.10
37	ว	16	2.75
38	ศ	2	0.34
39	ษ	7	1.20
40	ส	13	2.24
41	ห	12	2.07
42	ฬ	0	0
43	อ	42	7.23
44	ฮ	0	0
รวม		581	100

ส่วนสระภาษาไทย แสดงในตารางที่ 3.2

ตารางที่ 3.2 แสดงสระในภาษาไทย

ะ	า	ิ	ี	ึ	ื	ุ	ู	เ-ะ	เ-	แ-ะ	แ-
โ-ะ	โ-	เ-าะ	ออ	ำ	ไ	ใ	เ-า	ฤ	ฤา	ฦ	ฦา

ซึ่งในสระทั้งหมดนี้ มีบางตัวที่เป็นการนำสระตั้งแต่ 2 ตัวขึ้นไปมาประกอบกัน เพื่อให้ได้ออกมาเป็นสระตัวใหม่ออกมา เช่น เ-ะ แ-ะ โ-ะ เ-าะ เ-า เป็นต้น ดังแสดงในตารางที่ 3.2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เมื่อทราบถึงสระที่ใช้ในภาษาไทยแล้ว จึงนำมาวิเคราะห์หาคุณสมบัติของชื่อไทยของพืชในการใช้สระต่าง ๆ โดยวิธีการสุ่มอย่างง่ายมา 100 รายชื่อ ทำให้ทราบคุณสมบัติของการใช้สระในภาษาไทย ได้ดังตารางที่ 3.3

ตารางที่ 3.3 แสดงความถี่การใช้สระในข้อมูลทะเบียนพันธุ์ไม้ดอกไม้ประดับ

ลำดับที่	สระ	ความถี่	ร้อยละ
1	ะ	24	10.86
2	า	57	25.79
3	ิ	22	9.95
4	ี	18	8.14
5	ึ	3	1.36
6	ื	8	3.62
7	ุ	7	3.17
8	ู	8	3.62
9	เ-ะ	การเก็บข้อมูลแยกเป็น 2 อักษร คือ เ และ ะ	
10	เ-	31	14.03
11	แ-ะ	การเก็บข้อมูลแยกเป็น 2 อักษร คือ แ และ ะ	
12	แ-	13	5.88
13	โ-ะ	การเก็บข้อมูลแยกเป็น 2 อักษร คือ โ และ ะ	
14	โ-	7	3.17
15	เ-าะ	การเก็บข้อมูลแยกเป็น 3 อักษร คือ เ , า และ ะ	
16	ออ	0	0
17	ำ	7	3.17
18	ไ	13	5.88
19	ใ	2	0.90
20	เ-า	การเก็บข้อมูลแยกเป็น 2 อักษร คือ เ และ า	
21	ฤ	0	0
22	ฤา	1	0.45
23	ฦ	0	0

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 3.3 (ต่อ)

ลำดับที่	สระ	ความถี่	ร้อยละ
24	ฦา	0	0
รวม		221	100

ส่วนสุดท้ายคือ วรรณยุกต์ และอักษรอื่น ๆ ซึ่งจะประกอบด้วย

ตารางที่ 3.4 แสดงวรรณยุกต์ และตัวอักษรอื่น ๆ ในภาษาไทย

่	้	๊	๋	์	๎	ั	ฯ	ๆ

นำวรรณยุกต์ และอักษรอื่น ๆ มาวิเคราะห์หาคุณสมบัติของชื่อไทยของพืชในการใช้วรรณยุกต์ และอักษรต่าง ๆ โดยวิธีการสุ่มอย่างง่ายมา 100 รายชื่อ ได้ดังนี้

ตารางที่ 3.5 แสดงความถี่การใช้วรรณยุกต์ในข้อมูลทะเบียนพันธุ์ไม้ดอกไม้ประดับ

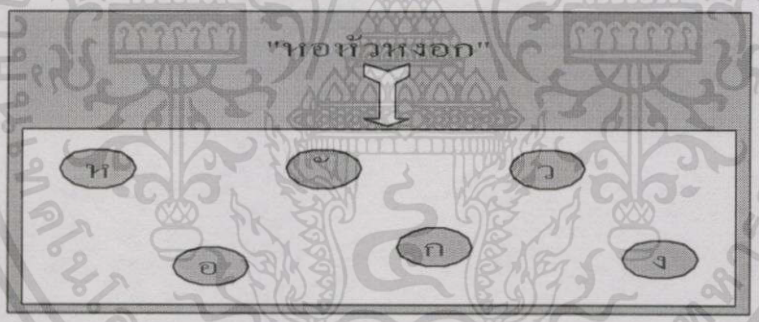
ลำดับที่	วรรณยุกต์ และอักษรอื่น	ความถี่	ร้อยละ
1	่	34	30.91
2	้	26	23.64
3	๊	2	1.82
4	๋	1	0.91
5	์	9	8.18
6	๎	7	6.36
7	็	2	1.82
8	ั	29	26.36
9	ฯ	0	0
10	ๆ	0	0
รวม		110	100

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. แนวความคิดในการค้นหาข้อมูลทะเบียนพันธุ์ไม้ดอกไม้ประดับ

หลังจากที่ได้ศึกษาถึงวิธีการค้นหาข้อมูลที่ใช้กันอยู่ในปัจจุบันแล้ว พบว่าขั้นตอนการค้นหาข้อมูลในแบบต่าง ๆ นั้น ถูกพัฒนาขึ้นโดยใช้พื้นฐานของการเปรียบเทียบอักษรกับข้อมูลภาษาอังกฤษเป็นหลัก แต่เมื่อนำวิธีการดังกล่าวมาประยุกต์ใช้กับข้อมูลทะเบียนพันธุ์ไม้ดอกไม้ประดับที่เป็นภาษาไทยจะมีข้อจำกัดเกิดขึ้น เนื่องจากข้อมูลทะเบียนพันธุ์ไม้ดอกไม้ประดับนั้นใช้อักษรภาษาไทยซึ่งมีลักษณะเฉพาะตัวที่ไม่เหมือนกับภาษาอื่น คือมีการใช้งานถึง 4 ระดับ ในขณะที่ในภาษาอังกฤษมีการใช้งานเพียงแค่ระดับเดียวเท่านั้น

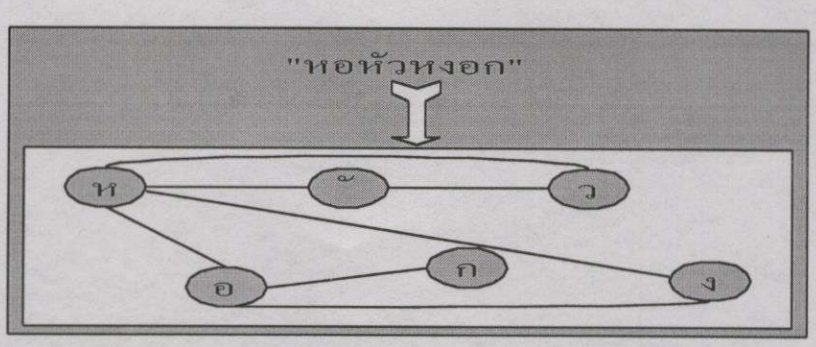
จึงทำให้เกิดแนวความคิดที่จะพัฒนาวิธีการค้นหาข้อมูลโดยพิจารณาจากข้อมูลทะเบียนพันธุ์ไม้ดอกไม้ประดับเป็นหลักขึ้นเพื่อให้การค้นหาข้อมูลมีประสิทธิภาพมากกว่าการนำวิธีการที่มีอยู่มาใช้โดยคุณสมบัติของข้อมูลไม่เหมาะสมกัน ซึ่งแนวความคิดนี้ก็คือ นำคุณสมบัติของอักษรภาษาไทยมาพิจารณาสร้างรูปแบบวิธีการค้นหาเฉพาะตัวขึ้น โดยพิจารณาพยัญชนะ สระ วรรณยุกต์ และอักษรอื่น ๆ ให้เป็นวัตถุอิสระต่อกัน ดังภาพที่ 3.3 ซึ่งมีค่าความสำคัญในระดับ ต่าง ๆ และเมื่ออักษรถูกมองเป็นวัตถุ ให้แต่ละวัตถุสามารถมีความสัมพันธ์กันได้อย่างอิสระต่อกัน ดังภาพที่ 3.4



ภาพที่ 3.3 แสดงแนวความคิดในการมองชื่อไทยของพืชเป็นวัตถุ

จากภาพที่ 3.3 แสดงให้เห็นว่า ชื่อไทยของพืช “หอหัวหงอก” มีอักษรซึ่งประกอบกันขึ้นจาก พยัญชนะ และสระ ใช้แสดงค่าทั้งหมด 9 อักษร แต่เมื่อถูกมองเป็นวัตถุแล้วจะมี 6 วัตถุ เนื่องจากมีอักษรซึ่งซ้ำกันอยู่ ถูกแสดงขึ้นมาหลายครั้ง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 3.4 แสดงแนวความคิดในการเชื่อมความสัมพันธ์ของวัตถุ

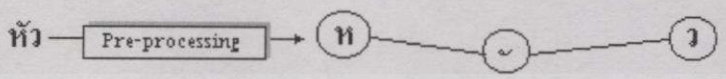
จากภาพที่ 3.4 แสดงให้เห็นว่าแต่ละวัตถุสามารถเชื่อมความสัมพันธ์กันได้อย่างอิสระ วัตถุ "ห" สามารถมีความสัมพันธ์กับวัตถุ "อ" ได้เนื่องจากคำว่า "หอ" เป็นตัวสร้างความสัมพันธ์นั่นเอง ในทำนองเดียวกัน วัตถุ "ห" "อ" "ก" "ง" "ว" และ "ั" จึงมีความสัมพันธ์กันจากคำว่า "หอหัวหงอก" นั่นเอง

เมื่อวัตถุต่าง ๆ มีความสัมพันธ์กัน ถูกสร้าง และเก็บบันทึกในโครงสร้างข้อมูลตามแนวความคิดแล้ว การกระทำการค้นหาจะเป็นส่วนต่อมา ซึ่งมีบทบาทสำคัญมีขั้นตอน และวิธีการดังต่อไปนี้ คือ เมื่อได้รับคำค้นหามาจากผู้ใช้แล้ว โปรแกรมจะทำการเตรียมการก่อนการค้นหา (Pre-processing) ก่อน โดยแปลงคำค้นหาให้เหมาะสมกับการค้นหาข้อมูลในแหล่งข้อมูลที่มีอยู่ จากนั้นก็จะมาพิจารณาว่าวัตถุใดควรจะเป็นราก (Root) เพื่อให้การค้นหามีประสิทธิภาพ และรวดเร็วมากที่สุด จากที่ได้กล่าวว่าการมองเป็นวัตถุจะทำให้วัตถุมีอิสระ ดังนั้นจึงสามารถเลือกวัตถุใด ๆ มาเป็นราก (Root) เพื่อความเหมาะสมต่อการค้นหาได้เสมอ เมื่อเลือกราก (Root) ได้แล้วก็เริ่มค้นหา (Query Process) ขั้นตอนนี้จะนำวัตถุที่ได้จากขั้น Pre-processing มาหาข้อมูลโดยการนำเอาวัตถุที่ได้เตรียมการแล้วมาพิจารณา จะทำให้ทราบว่าวัตถุดังกล่าวนี้มีคุณสมบัติตรงกับข้อมูลใดบ้าง ไม่ว่าคำดังกล่าวจะอยู่ในส่วนต้นของคำ ส่วนกลางของคำ หรือแม้กระทั่งอยู่ในส่วนท้ายของคำก็ตาม วัตถุใด ๆ ที่มีคุณสมบัติดังกล่าวอยู่ จะถือว่าข้อมูลนี้อยู่ในข่ายที่เป็นข้อมูลที่ผู้ใช้ต้องการด้วย เมื่อได้กลุ่มของข้อมูลที่อยู่ในข่ายที่ต้องการแล้ว ก็จะแสดงข้อมูลกลุ่มดังกล่าวเพื่อให้ผู้ใช้ได้เลือกข้อมูลที่ตนเองต้องการ (Selection Process) และเมื่อผู้ใช้ได้เลือกข้อมูลแล้ว ต่อมาก็จะเป็นการแสดงรายละเอียดต่าง ๆ ข้อมูลที่ผู้ใช้ได้เลือก แสดงผล (Output) ให้ผู้ใช้ได้รับทราบ

ตัวอย่าง สมมติว่า คำค้น (Keyword) ที่ผู้ใช้ต้องการค้นหาคือ "หัว" โดยมีจุดประสงค์เพื่อต้องการข้อมูล "หอหัวหงอก" มาเป็นผลลัพธ์ของการค้นหา จากข้อมูลทั้งหมด 10 ชื่อ คือ "กระโดงหางแบน" "ขันข้อ" "จันทน์" "ซงโคดอกเหลือง" "หอหัวหงอก" "ตะไลคงคา" "หัวใจแตก" "อโศก" "หนามแดงฝรั่ง" และ "แสยกฟิลิปปินส์" ซึ่งวิธีการค้นหาข้อมูลนั้นจะถูกแบ่งขั้นตอนเป็น 4 ขั้นใหญ่ ๆ ดังนี้ คือ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

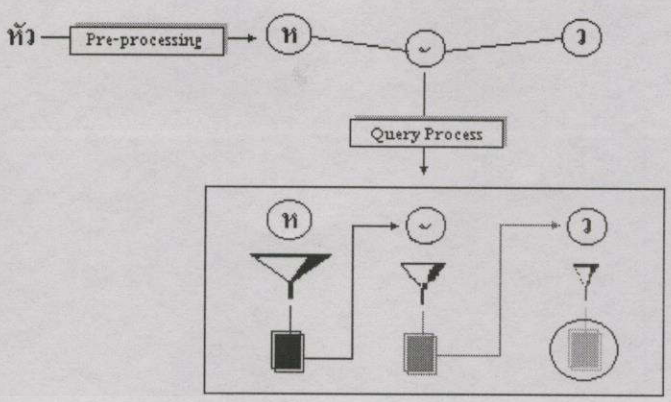
ขั้นตอน Pre-processing เป็นขั้นตอนการนำเอาคำค้นของผู้ใช้มาจัดรูปแบบให้เหมาะสมต่อการค้นหาข้อมูล ก่อนทำการค้นหาข้อมูลในฐานข้อมูล ดังนั้นคำว่า "หัว" จึงถูกแบ่งออกเป็นวัตถุ 3 วัตถุด้วยกัน และสร้างความสัมพันธ์ระหว่างวัตถุทั้ง 3



ภาพที่ 3.5 แสดงการเตรียมข้อมูลก่อนการค้นหา

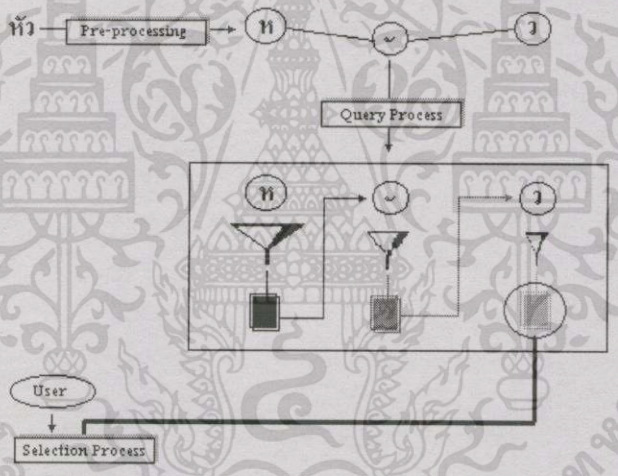
ขั้นตอน Query process ขั้นตอนนี้จะนำวัตถุที่ได้จากขั้น Pre-processing มาหาข้อมูลโดยการนำเอาวัตถุที่มีคุณสมบัติดังกล่าวมาพิจารณาจากวัตถุ "ว" "ั" และ "ห" โดยเริ่มจากการหาราก (Root) ก่อนโดยให้วัตถุ ห เป็นราก จากนั้นจึงพิจารณาว่าวัตถุ ห นั้น มีข้อมูลใดที่ประกอบด้วยวัตถุ ห ซึ่งผลที่ได้จะมี 5 ชื่อคือ "กระโดงหางแบน" "ชงโคดอกเหลือง" "หอหัวหงอก" "หัวใจแตก" "หนามแดงฝรั่ง" ซึ่งขั้นตอนนี้เราสามารถลดข้อมูลที่ไม่ต้องการได้ถึง 5 ชื่อ จากนั้นก็พิจารณาชื่อที่มีอยู่ชื่อใดมีวัตถุ "ั" และมีความสัมพันธ์กันกับคำค้นบ้าง ซึ่งผลที่ได้ก็จะมี 2 ชื่อคือ "หอหัวหงอก" และ "หัวใจแตก" ลองพิจารณาดูว่า "กระโดงหางแบน" "ชงโคดอกเหลือง" ไม่ใช่เนื่องจากไม่มีวัตถุ "ั" ส่วน "หนามแดงฝรั่ง" นั้นมีวัตถุ "ั" อยู่ แต่วัตถุทั้ง 2 ไม่มีความสัมพันธ์ซึ่งกันและกันทำให้ชื่อพืชชื่อนี้ไม่มีคุณสมบัติจึงไม่ใช่คำที่ผู้ใช้ต้องการ จากนั้นก็พิจารณาวัตถุ "ว" ผลที่ได้คือ มีอยู่ 2 ชื่อที่เข้าข่ายที่ผู้ใช้ต้องการคือ "หอหัวหงอก" และ "หัวใจแตก" ซึ่งจะเห็นได้ว่าทั้ง 2 ชื่อนี้คำว่า "หัว" อยู่คนละตำแหน่งกัน คือ "หอหัวหงอก" จะอยู่ในส่วนกลางของคำ และ "หัวใจแตก" นั้นอยู่ส่วนต้นของคำ ซึ่งความสามารถในส่วนนี้การค้นหาแบบเบรดเฟิร์สเซิชไม่มี จะต้องเทียบคำจากตัวแรกไปตัวสุดท้ายของคำเสมอ แต่วิธีนี้สามารถจะทำให้ทราบว่าวัตถุดังกล่าวนี้มีคุณสมบัติตรงกับข้อมูลใดบ้าง ไม่ว่าคำดังกล่าวจะอยู่ในส่วนต้น กลาง หรือแม้กระทั่งอยู่ส่วนท้ายของคำก็ตาม จะสังเกตได้ว่า "หอหัวหงอก" และ "หัวใจแตก" ก็มีคุณสมบัติดังกล่าวอยู่จึงทำให้ข้อมูลนี้อยู่ในข่ายที่เป็นข้อมูลที่ผู้ใช้ต้องการด้วย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 3.6 แสดงการค้นหาข้อมูลจากวัตถุที่มีคุณสมบัติเดียวกัน

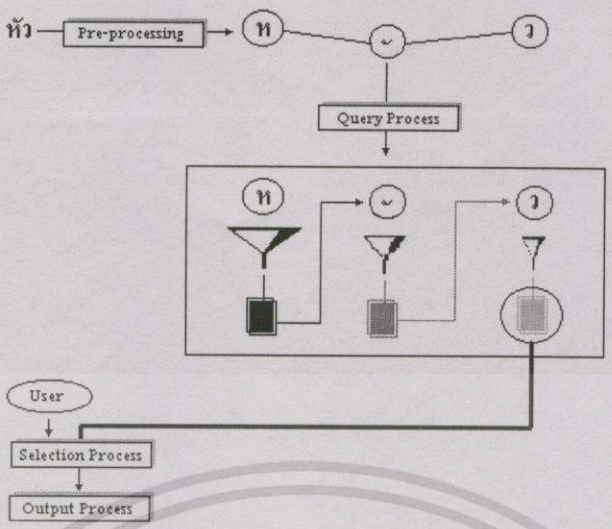
ขั้นตอน Selection process นำเอาชื่อไทยของพืชที่มีคุณสมบัติตรงกันขึ้นมาแสดงเพื่อให้ผู้ใช้ได้เลือกข้อมูลที่ตนเองต้องการ ในขั้นตอนนี้ผู้ใช้จะสามารถเห็น และเลือกข้อมูล "หอหัวหงอก" ที่ตนเองต้องการได้



ภาพที่ 3.7 แสดงกลุ่มของข้อมูลเพื่อให้ผู้ใช้ได้เลือกข้อมูลที่ตนต้องการ

ขั้นตอน Output process เมื่อผู้ใช้เลือกข้อมูล "หอหัวหงอก" ที่ตนเองต้องการแล้ว ก็จะมาถึงส่วนแสดงผลลัพธ์ คือจะแสดง ชื่อวิทยาศาสตร์ของพืช ซึ่งจะทำให้ผู้ใช้ได้ทราบถึง ข้อมูลพื้นฐานอย่างหยาบ ๆ ของพืชหอหัวหงอก และข้อมูลดังกล่าวนี้เอง ถ้าผู้ใช้ต้องการข้อมูล หรือรายละเอียดเพิ่มเติมก็สามารถค้นหาข้อมูลอย่างละเอียดต่อได้โดยไม่ยากนัก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

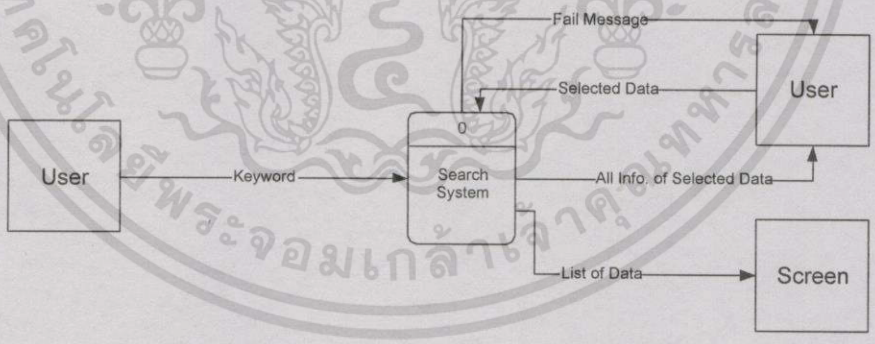


ภาพที่ 3.8 แสดงขั้นตอนในการแสดงข้อมูลที่ผู้ใช้ได้เลือกแล้ว

3. **การออกแบบโครงสร้าง**

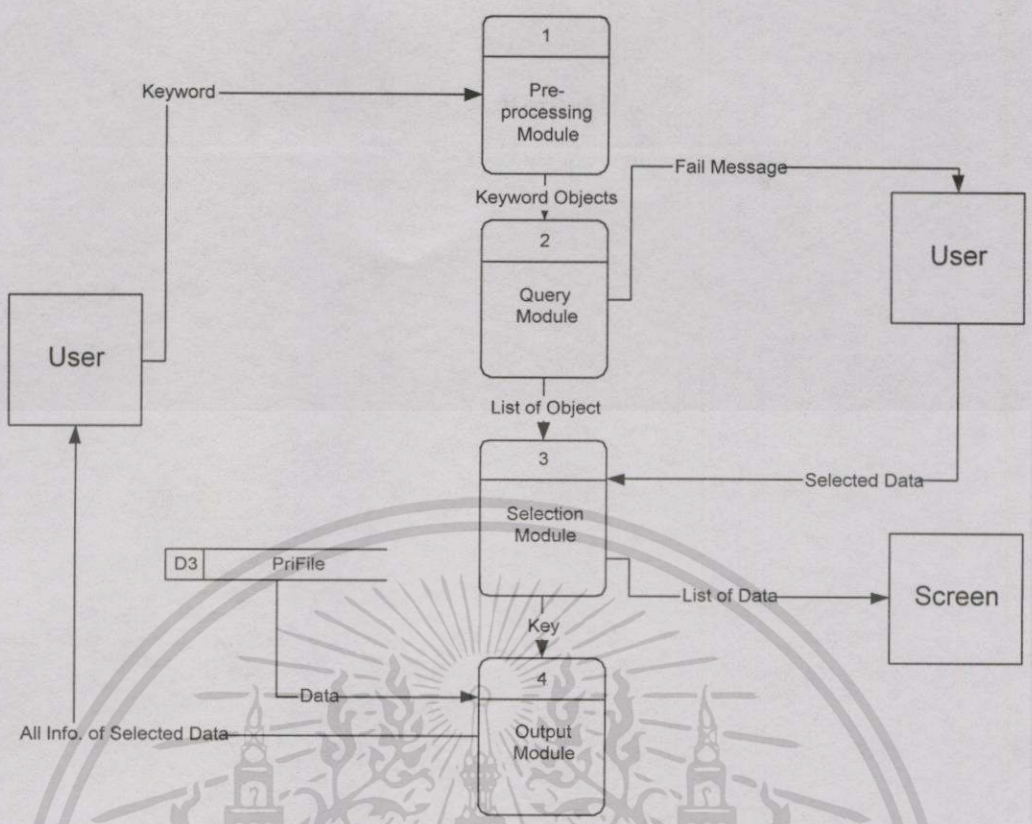
การออกแบบโครงสร้างตามแนวความคิดในการค้นหาข้อมูลที่ได้กล่าวมาแล้วข้างต้น มีลักษณะดังนี้

Data Flow Diagram (DFD)

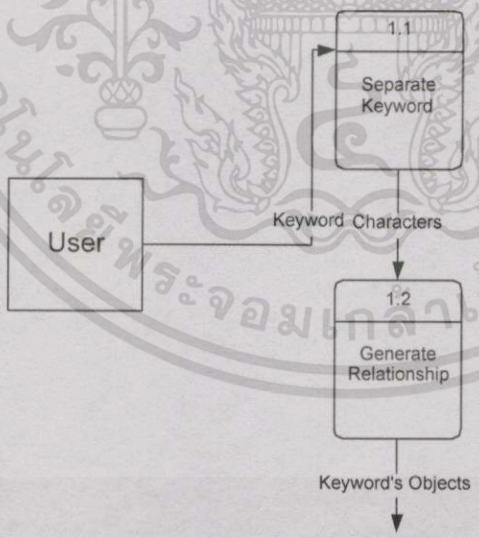


ภาพที่ 3.9 Data Flow Diagram (Context Diagram)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

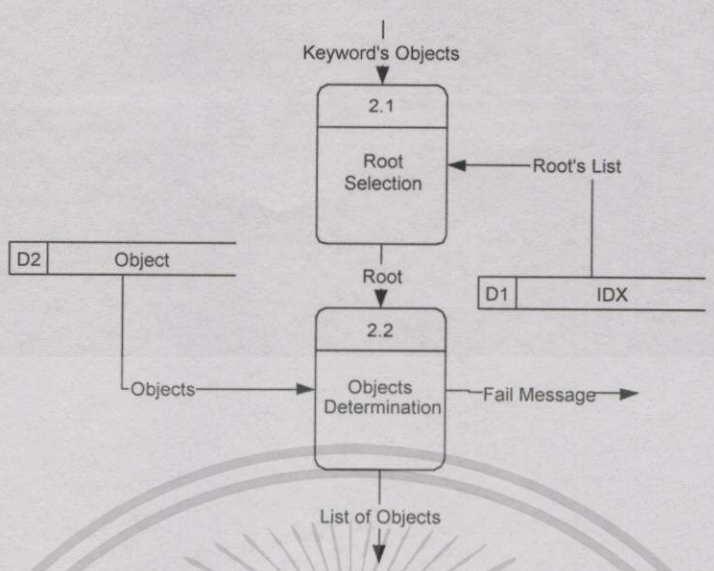


ภาพที่ 3.10 Data Flow Diagram (Level 0)



ภาพที่ 3.11 Data Flow Diagram (Level 1 of Pre-processing Module)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 3.12 Data Flow Diagram (Level 1 of Query Module)

ดังนั้น ฟังก์ชันในการค้นหาตามแนวความคิดนี้ ถูกกำหนดขึ้นให้มีลักษณะดังนี้

$$GS(N, A, GR, GG)$$

N = Graph ที่มี Set of Node

A = Set of Arcs ที่มีการ Connect กัน

GR = Graph Root State

GG = Graph Goal State

Graph Search Function (GS) จะใช้สำหรับค้นหาวัตถุจากรากที่เลือกขึ้นมาให้ใกล้เคียงกับเป้าหมายมากที่สุดเพื่อความรวดเร็วในการค้นหา ในการหาวัตถุเป้าหมายที่ต้องการไปตามโหนดต่าง ๆ ซึ่งเชื่อมกันด้วยตัวเชื่อม (ARC) ระหว่างวัตถุ 2 วัตถุเข้าด้วยกัน

N จะถูกสร้างขึ้นโดยโปรแกรมคือ จะเป็นตัวแบ่ง และสร้างข้อมูลโหนดแต่ละโหนด ซึ่งจะถูกเชื่อมด้วยกัน ถ้าโหนดนั้น ๆ มีความสัมพันธ์ต่อกัน

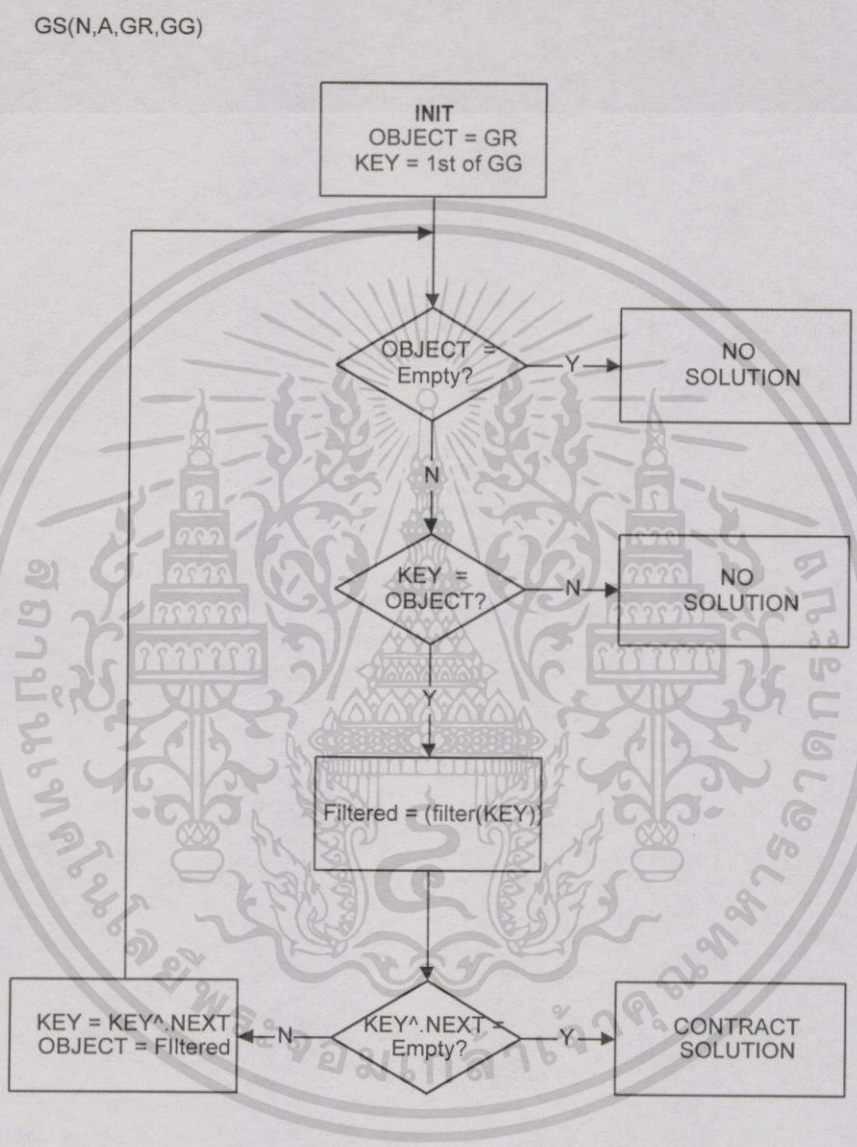
A ถูกสร้างขึ้นโดยโปรแกรม ส่วนข้อมูลที่แสดงจะเป็นกลุ่มข้อมูลซึ่งแสดงความสัมพันธ์ระหว่างโหนดต่าง ๆ

GR จะถูกกำหนดขึ้นโดยโปรแกรมเป็นผู้พิจารณาว่า โหนดใดสามารถที่จะไปถึงข้อมูลตามที่ผู้ใช้ต้องการได้ด้วยระยะทางที่สั้น และรวดเร็วที่สุด

GG จะเป็นข้อมูลคำค้น (Keyword) ซึ่งผู้ใช้จะเป็นผู้กำหนดขึ้นมาเอง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ฉะนั้น เมื่อผู้ใช้ต้องการค้นหาข้อมูลก็เพียงกำหนดคำค้น (Keyword) ของข้อมูลที่ต้องการทราบ จากนั้นโปรแกรมจะใช้ค่า N และ A ที่มีอยู่ทำการประมวลผล และกำหนด GR ขึ้นมาเอง หลังจากนั้นก็จะเป็นขั้นตอนการค้นหา ซึ่งโครงสร้างสำหรับการค้นหาข้อมูลจะมีลักษณะดังนี้



ภาพที่ 3.13 FlowChart แสดงโครงสร้างการค้นหาข้อมูล

1. วิธีการค้นหา (CGI CODE) ผู้วิจัยสร้างวิธีการค้นหา 2 รูปแบบ เพื่อทำหน้าที่เป็นสื่อกลางระหว่างคำค้นกับฐานข้อมูลทะเบียนพันธุ์ไม้ดอกไม้ประดับ
2. เขียนโปรแกรมตามวิธีการค้นหาแบบ Breath-First Search
3. เขียนโปรแกรมตามวิธีการค้นหาของผู้วิจัยเอง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4. ฐานข้อมูล (Database) เก็บบันทึกข้อมูลทะเบียนพันธุ์ไม้ดอกไม้ประดับเพื่อเป็นข้อมูลให้กับผู้ค้นหาที่ต้องการหาข้อมูลตามที่ตนเองต้องการ

5. นำโปรแกรมที่สร้างไปตรวจหาความผิดพลาดที่อาจเกิดขึ้นระหว่างการสร้างโดยการทดลองใช้โปรแกรมเพื่อตรวจสอบผลลัพธ์ที่ได้จากโปรแกรมกับข้อมูล

6. แก้ไขโปรแกรมจากความผิดพลาดที่ตรวจพบให้เรียบร้อย และนำไปทดสอบความถูกต้องของโปรแกรม จนแน่ใจว่าโปรแกรมสมบูรณ์ และพร้อมที่จะนำมาทดลอง

3.3 การดำเนินการทดลอง

ในการทดลองหาประสิทธิภาพของอัลกอริทึมโดยทั่วไปนั้นจะใช้วิธีการทางด้านคณิตศาสตร์ช่วยในการหาค่า ซึ่งจะพิจารณาจากเวลาที่ถูกใช้ไปในการค้นหา (Execution Time) เนื้อที่ที่ถูกใช้ (Memory Space) เป็นต้น ซึ่งการสร้างวิธีการค้นหาเพื่อทดลองในครั้งนี้ ผู้วิจัยแบ่งเป็นขั้นตอนดังนี้

1. เลือกเซิร์ฟเวอร์ที่ใช้สำหรับทดลอง โดยการทดลองครั้งนี้ผู้วิจัยเลือกเซิร์ฟเวอร์ของสถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง ซึ่งเป็นยูนิกซ์ (UNIX) โดยใช้เนื้อที่ของผู้วิจัยเองในการทดลองครั้งนี้ คือ "http://chaokhun.kmitl.ac.th/~s1064249/"

2. กำหนดภาษาที่ใช้ให้เหมาะสมกับเซิร์ฟเวอร์ที่ให้บริการ คือ CGI-Perl เนื่องจากภาษา Perl เป็นโปรแกรมประเภทอินเทอร์พรีเตอร์ (Interpreter) ข้อดีคือ ผู้เขียนโปรแกรมสามารถทดสอบโปรแกรม และ แก้ไขข้อผิดพลาดของโปรแกรมได้ทันทีโดยไม่จำเป็นต้องคอมไพล์ใหม่ และภาษา Perl มีความสามารถในการจัดการข้อมูลสตริงค์ได้ดี และสามารถนำไปใช้ได้กับเซิร์ฟเวอร์หลายประเภท เช่น ยูนิกซ์ วินโดว์สเอ็นที เป็นต้น ดังนั้น ภาษา Perl จึงเป็นภาษาที่ผู้วิจัยกำหนดขึ้นสำหรับงานวิจัยนี้

3. ลงมือสร้าง ซึ่งจะแบ่งเป็น 3 ส่วนดังนี้

- ส่วนเชื่อมกับผู้ใช้ (User Interface) ซึ่งเป็นส่วนที่ใช้ติดต่อกับ CGI เพื่อทำหน้าที่เป็นส่วนรับ และ แสดงผลระหว่างผู้ค้นหาและ CGI
- สร้างส่วนติดต่อเพื่อติดต่อกับวิธีการค้นหาของผู้วิจัย
- สร้างส่วนติดต่อเพื่อติดต่อกับวิธีการค้นหาแบบ breath-first search

ภายหลังจากที่โปรแกรมถูกสร้างขึ้นจากวิธีการค้นหาข้อมูลแล้ว จึงเป็นส่วนการทดลองหาประสิทธิภาพของวิธีการค้นหาข้อมูลของวิธีการทั้ง 2 วิธี โดยผู้วิจัยจะกำหนดคำค้น (Keyword) ต่าง ๆ จำนวน 100 คำขึ้นมา โดยกลุ่มของคำค้นที่กำหนดขึ้นมาเพื่อการทดลองนี้สามารถแบ่งประเภทของคำค้นเป็น 3 ประเภท ซึ่งจะมีลักษณะแตกต่างกันดังนี้

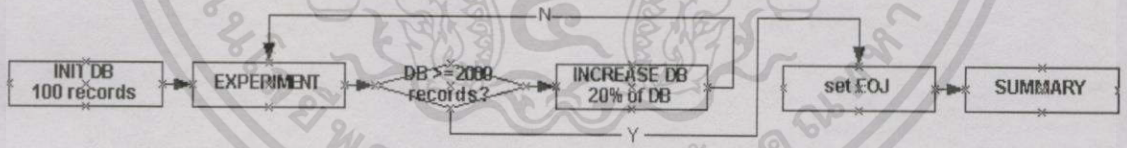
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 3.6 แสดงประเภทของคำค้น

ลำดับที่	ประเภทคำค้น	รายละเอียดเฉพาะ
1	คำค้นครบถ้วน	ผู้ใช้ระบุคำค้นครบคำซึ่งค้นหาแล้วปรากฏคำค้นดังกล่าวมีอยู่ในฐานข้อมูล
2	คำค้นบางส่วน	ผู้ใช้ระบุคำค้นเฉพาะบางส่วนซึ่งค้นหาแล้วปรากฏคำค้นดังกล่าวเป็นส่วนประกอบในข้อมูล 1 รายการ หรือมากกว่า ในฐานข้อมูล
3	คำค้นที่ไม่พบ	ผู้ใช้ระบุคำค้นซึ่งค้นหาแล้วไม่ปรากฏข้อมูลที่มีคำค้นดังกล่าวประกอบอยู่

กำหนดข้อมูลในฐานข้อมูล (ครั้งแรกที่ 100 รายการ)

- ทดลองค้นหาข้อมูลตามคำค้น และจดบันทึกค่าต่าง ๆ เพื่อใช้หาประสิทธิภาพของวิธีการค้นหา (ประกอบด้วย ค่าความครบถ้วนของข้อมูล ค่าความตรงต่อความต้องการ ค่าเวลาที่ใช้ และ ค่าของการใช้เนื้อที่)
  - โดยใช้วิธีการของผู้วิจัย จดบันทึกผลการทดลอง
  - โดยใช้วิธีการแบบ breath-first search จดบันทึกผลการทดลอง
- เพิ่มข้อมูลในฐานข้อมูลร้อยละ 20 ของฐานข้อมูล กลับไปทดลองค้นหาข้อมูลตามคำค้น จนกระทั่งข้อมูลมีมากกว่า 2,000 รายการ



ภาพที่ 3.14 แสดงขั้นตอนการทำการทดลอง

- ในการทดลองเริ่มทำการทดลองและเก็บข้อมูลตั้งแต่ 1 พฤษภาคม 2543 ถึง วันที่ 1 พฤศจิกายน 2543

3.4 การวิเคราะห์ข้อมูล

เมื่อดำเนินการทดลองและบันทึกผลการทดลองเรียบร้อยแล้ว ผู้วิจัยจะนำข้อมูลที่ได้มาวิเคราะห์หาประสิทธิภาพของวิธีการทั้ง 2 โดยนำข้อมูลที่ได้จากการทดลองในแต่ละครั้งมาหาประ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สิทธิภาพในด้านความครบถ้วนของข้อมูล ด้านความตรงต่อความต้องการ ด้านเวลาที่ใช้ในการค้นหาข้อมูล และด้านการใช้เนื้อที่ในการเก็บข้อมูลเพื่อการค้นหา และบันทึกลงในตารางแสดงประสิทธิภาพของวิธีการค้นหาของการทดลองแต่ละครั้ง

จากนั้น จึงนำเอาประสิทธิภาพของวิธีการที่หาได้มาเปรียบเทียบประสิทธิภาพของวิธีการค้นหา และทำการสรุป โดยนำเสนอในรูปของกราฟเปรียบเทียบประสิทธิภาพในด้านต่าง ๆ เพื่อศึกษาผลการทดลองตามแนวความคิดของผู้วิจัย ว่าควรมีแนวทางเป็นไปในลักษณะใด มีจุดบกพร่องตรงส่วนใด และนำมาใช้เป็นแนวทางในการพัฒนา และ แก้ไขข้อบกพร่องของวิธีการค้นหาข้อมูลทะเบียนพันธุ์ไม้ดอกไม้ประดับต่อไปในอนาคตต่อไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4
ผลการวิเคราะห์ข้อมูล

การวิจัยครั้งนี้เป็นการวิจัยเพื่อพัฒนาวิธีการค้นหาข้อมูลบนอินเตอร์เน็ตกับข้อมูลทะเบียนพันธุ์ไม้ดอกไม้ประดับให้มีประสิทธิภาพ และเปรียบเทียบประสิทธิภาพของการค้นหาข้อมูลบนอินเตอร์เน็ต ด้วยวิธีการของผู้วิจัย กับวิธี Breath-First Search และเพื่อให้เป็นการสอดคล้องกับวัตถุประสงค์ของการวิจัย ผู้วิจัยจึงขอเสนอผลการวิเคราะห์ข้อมูลที่ได้ทำการศึกษาดังนี้

4.1 เปรียบเทียบประสิทธิภาพของการค้นหาข้อมูลบนอินเตอร์เน็ตที่เป็นผลจากการทดลองด้วยวิธีการของผู้วิจัย กับวิธี Breath-First Search

4.2 สรุปผลการวิเคราะห์ข้อมูล

4.1 เปรียบเทียบประสิทธิภาพของการค้นหาข้อมูลบนอินเตอร์เน็ต ที่เป็นผลจากการทดลองด้วยวิธีของผู้วิจัย กับวิธี Breath-First Search

การเปรียบเทียบประสิทธิภาพของการค้นหาข้อมูลบนอินเตอร์เน็ต ด้วยวิธีการของผู้วิจัยกับวิธี Breath-First Search ผู้วิจัยได้ทำการทดลอง 17 ครั้ง เพื่อทำการเปรียบเทียบประสิทธิภาพของวิธีการค้นหาเมื่อข้อมูลในฐานข้อมูลกำหนดโดยการทดลองเปลี่ยนแปลงข้อมูลเพิ่มขึ้นในอัตราร้อยละ 20 และเนื่องจากครั้งที่ 18 จะต้องใช้ข้อมูลจำนวน 2,219 รายการ ซึ่งข้อมูลจริงมีจำนวนไม่ถึง 2,219 รายการแล้ว และจดบันทึกผลการทดลองที่แสดงในตารางที่ 4.1แสดงผลการทดลอง

ผู้วิจัยขอเสนอผลการวิเคราะห์ข้อมูลโดยแยกตามดัชนีที่ได้ทำการศึกษาดังนี้

1. ความครบถ้วนของข้อมูล
2. ความตรงต่อความต้องการ
3. เวลาที่ใช้ในการค้นหาข้อมูล
4. การใช้เนื้อที่ในการเก็บข้อมูลเพื่อการค้นหา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.1 แสดงผลการทดลองหาประสิทธิภาพการค้นหาข้อมูลบนอินเตอร์เน็ตกับข้อมูลทะเบียนพันธุ์ไม้ดอกไม้ประดับ โดยวิธีการของผู้วิจัย (S1) และ วิธี Breath-First Search (S2)

	ความครบถ้วนของข้อมูล		ความถูกต้องของข้อมูล		เวลาที่ใช้ในการค้นหาข้อมูล		การใช้เนื้อที่ในการเก็บข้อมูลเพื่อการค้นหา	
ครั้งที่	(%)		(%)		(1/1,000,000 วินาที)		(จำนวนไบต์)	
	S1	S2	S1	S2	S1	S2	S1	S2
1	100	100	100	100	245,702	607,452	201,800	367,988
2	100	100	100	100	265,841	607,560	243,160	443,409
3	100	100	100	100	269,071	608,228	291,142	530,906
4	100	100	100	100	302,684	608,048	353,710	645,001
5	100	100	100	100	302,966	608,055	421,952	769,443
6	100	100	100	100	326,753	608,069	506,943	924,425
7	100	100	100	100	331,465	615,054	604,075	1,101,548
8	100	100	100	100	334,689	614,118	724,887	1,321,853
9	100	100	100	100	360,815	617,416	868,815	1,584,310
10	100	100	100	100	401,658	617,494	1,047,286	1,909,756
11	100	100	100	100	434,091	619,395	1,254,772	2,288,114
12	100	100	100	100	445,798	617,608	1,509,391	2,752,419
13	100	100	100	100	479,052	620,517	1,839,269	3,353,961
14	100	100	100	100	498,673	624,679	2,174,158	3,964,641
15	100	100	100	100	522,416	630,562	2,601,947	4,744,728
16	100	100	100	100	554,988	635,624	3,131,289	5,709,998
17	100	100	100	100	560,214	645,658	3,796,396	6,922,840

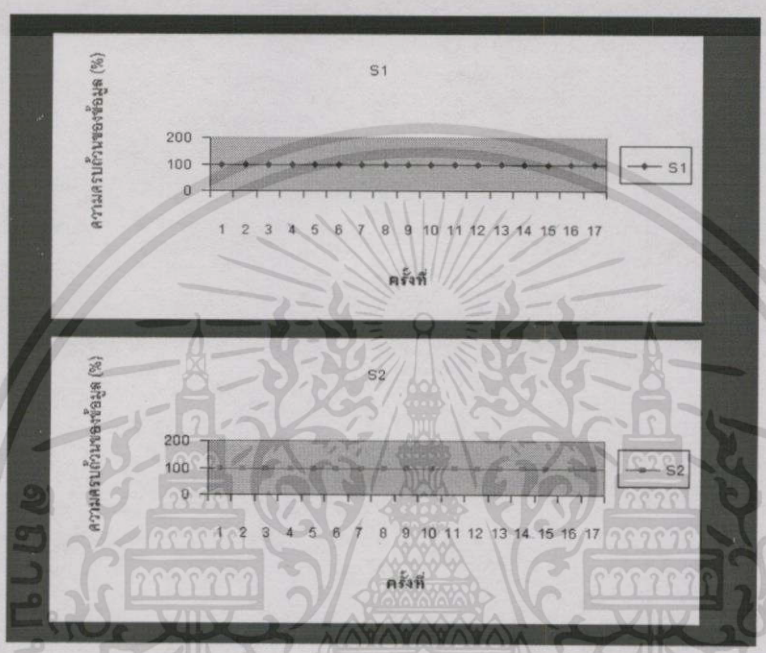
4.1.1 ผลการวิเคราะห์ความครบถ้วนของข้อมูล

ความครบถ้วนของข้อมูล คือ ความสามารถของวิธีการค้นหาข้อมูลที่สามารถค้นหา (Query) ข้อมูล และดึงข้อมูล (Retrieve) ข้อมูลทะเบียนพันธุ์ไม้ดอกไม้ประดับได้ครบถ้วนตามที่มีอยู่ในฐานข้อมูลจริง โดยผู้วิจัยทำการทดลองค้นหาชื่อพันธุ์ไม้ดอกไม้ประดับ จำนวน 17 ครั้ง โดยแต่ละครั้งจะค้นหาชื่อรายการพันธุ์ไม้ดอกไม้ประดับโดยวิธี การของผู้วิจัย จำนวน 100 ชื่อ และวิธี Breath-First Search จำนวน 100 ชื่อ ((17 ครั้ง* ครั้งละ 100) * 2 วิธี = 3,400 ครั้ง) ในแต่ละครั้งที่ทำการทดลอง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากผลการทดลองจำนวน 17 ครั้งปรากฏว่า ทั้งสองวิธี คือวิธีการของผู้วิจัยและ วิธี Breath-First Search ความครบถ้วนของข้อมูล 100 เปอร์เซ็นต์ เท่ากัน

สรุปว่าผลการวิเคราะห์ความครบถ้วนของข้อมูล วิธีการของผู้วิจัย และวิธี Breath-First Search นั้นได้ผลเท่ากัน ดังแสดงในกราฟแสดงผล โดยที่ S1 คือวิธีการของผู้วิจัย และ S2 คือวิธี Breath-First Search



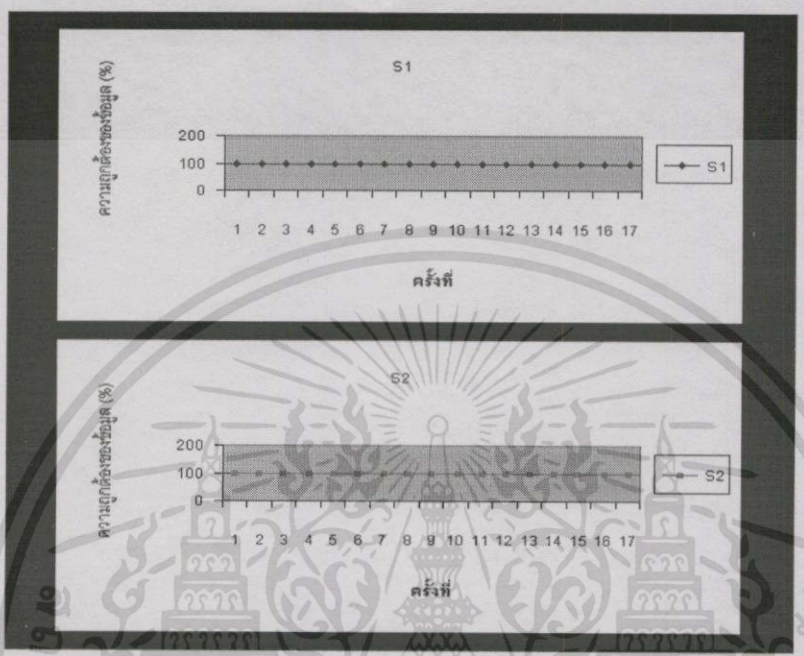
ภาพที่ 4.1 กราฟแสดงผลวิเคราะห์ในด้านความครบถ้วนของข้อมูล ทั้ง 2 วิธี

4.1.2 ผลการวิเคราะห์ความตรงต่อความต้องการ

ความตรงต่อความต้องการ (Relevance) คือ ความสามารถของวิธีการค้นหาข้อมูลที่สามารถเรียกข้อมูลทะเบียนพันธุ์ไม้ดอกไม้ประดับค้นคืนมาได้ถูกต้องตามที่ผู้ใช้ต้องการ โดยผู้วิจัยทำการทดลองค้นหาชื่อพันธุ์ไม้ดอกไม้ประดับ จำนวน 17 ครั้ง โดยแต่ละครั้งจะค้นหาชื่อรายการพันธุ์ไม้ดอกไม้ประดับโดยวิธีการของผู้วิจัย จำนวน 100 ชื่อ และวิธี Breath-First Search จำนวน 100 ชื่อ ((17 ครั้ง* ครั้งละ 100) * 2 วิธี = 3,400 ครั้ง) ในแต่ละครั้งที่ทำการทดลอง จากผลการทดลองจำนวน 17 ครั้งปรากฏว่า ทั้งสองวิธี คือ วิธีการของผู้วิจัย และวิธี Breath-First Search ได้ผลตรงต่อความต้องการ 100 เปอร์เซ็นต์ เท่ากัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สรุปว่าผลการวิเคราะห์ความตรงต่อความต้องการ วิธีการของผู้วิจัย และวิธี Breath-First Search ได้ผลเท่ากัน ดังแสดงในกราฟแสดงผล โดยที่ S1 คือ วิธีการของผู้วิจัย และ S2 คือวิธี Breath-First Search



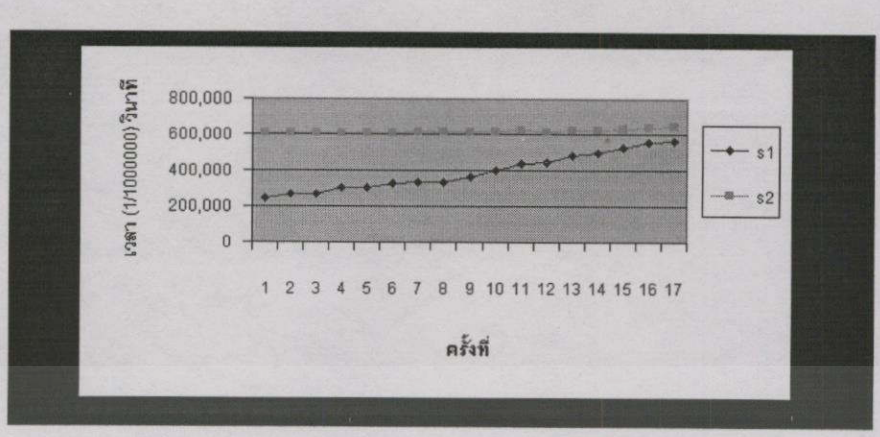
ภาพที่ 4.2 กราฟแสดงผลวิเคราะห์ในด้านความตรงต่อความต้องการ ทั้ง 2 วิธี

4.1.3 ผลการวิเคราะห์เวลาที่ใช้ในการค้นหาข้อมูล

เวลาที่ใช้ในการค้นหา (Execution Time) คือ จำนวนเวลาของวิธีการค้นหา ที่ใช้ไปในการค้นหาข้อมูลทะเบียนพันธุ์ไม้ดอกไม้ประดับ ซึ่งจะมีหน่วยเป็นวินาที โดยผู้วิจัยทำการทดลองค้นหาชื่อพันธุ์ไม้ดอกไม้ประดับ จำนวน 17 ครั้ง โดยแต่ละครั้งจะค้นหาชื่อรายการพันธุ์ไม้ดอกไม้ประดับโดยวิธีการของผู้วิจัย จำนวน 100 ชื่อ และวิธี Breath-First Search จำนวน 100 ชื่อ ((17 ครั้ง* ครั้งละ 100) * 2 วิธี = 3,400 ครั้ง) ในแต่ละครั้งที่ทำการทดลอง จากผลการทดลองจำนวน 17 ครั้งปรากฏว่า วิธีการของผู้วิจัย ใช้เวลาในการค้นหาข้อมูลน้อยกว่า วิธี Breath-First Search

สรุปว่าผลการวิเคราะห์เวลาที่ใช้ในการค้นหาข้อมูล วิธีการของผู้วิจัย ใช้เวลาในการค้นหาข้อมูลน้อยกว่า วิธี Breath-First Search ดังแสดงในกราฟแสดงผล โดยที่ S1 คือวิธีการของผู้วิจัย และ S2 คือวิธี Breath-First Search

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

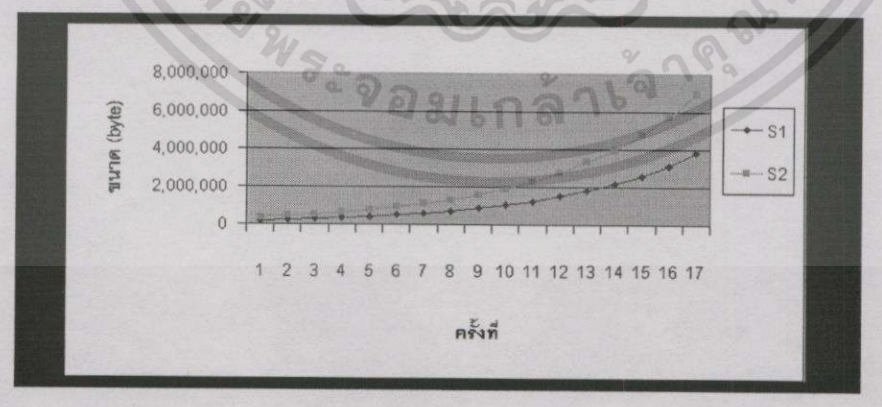


ภาพที่ 4.3 กราฟแสดงผลวิเคราะห์ในด้านเวลาที่ใช้ในการค้นหาข้อมูล ทั้ง 2 วิธี

4.1.4 ผลการวิเคราะห์การใช้เนื้อที่ในการเก็บข้อมูลเพื่อการค้นหา

เนื้อที่ (Memory Space) คือ จำนวนของเนื้อที่ที่ถูกใช้ในการเก็บข้อมูล ซึ่งจะมีหน่วยเป็นไบต์ (Byte) โดยผู้วิจัยทำการทดลองค้นหาชื่อพันธุ์ไม้ดอกไม้ประดับ จำนวน 17 ครั้ง โดยแต่ละครั้งจะค้นหาชื่อรายการพันธุ์ไม้ดอกไม้ประดับ โดยวิธีการของผู้วิจัย จำนวน 100 ชื่อ และวิธี Breath-First Search จำนวน 100 ชื่อ ((17 ครั้ง* ครั้งละ 100) * 2 วิธี = 3,400 ครั้ง) ในแต่ละครั้งที่ทำการทดลอง จากผลการทดลองจำนวน 17 ครั้งปรากฏว่า ใช้เนื้อที่สำหรับเก็บข้อมูลเพื่อการค้นหาน้อยกว่า วิธี Breath-First Search

ดังนั้นจึงสรุปว่า ผลการวิเคราะห์การใช้เนื้อที่ในการเก็บข้อมูลเพื่อการค้นหานั้น วิธีการของผู้วิจัย ใช้เนื้อที่น้อยกว่า วิธี Breath-First Search ดังแสดงในกราฟแสดงผล โดยที่ S1 คือ วิธีการของผู้วิจัย และ S2 คือวิธี Breath-First Search



ภาพที่ 4.4 กราฟแสดงผลวิเคราะห์ในด้านการใช้เนื่อที่ในการเก็บข้อมูลเพื่อการค้นหาทั้ง 2 วิธี

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

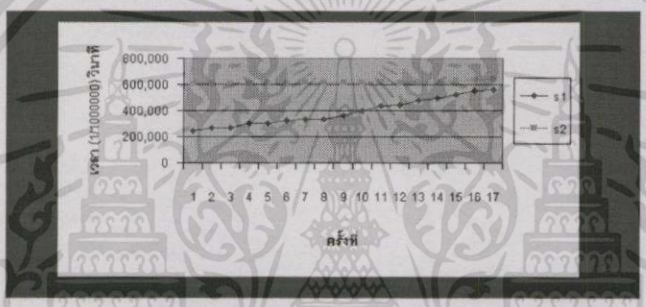
4.2 สรุปผลการวิเคราะห์ข้อมูล

เมื่อนำผลการทดลองมาวิเคราะห์แล้วสามารถสรุปได้ว่า

1. ด้านความครบถ้วนของข้อมูล วิธีการของผู้วิจัย และวิธี Breath-First Search ให้ความครบถ้วนของข้อมูลมีค่า 100 เปอร์เซ็นต์เท่ากันทั้งสองวิธี

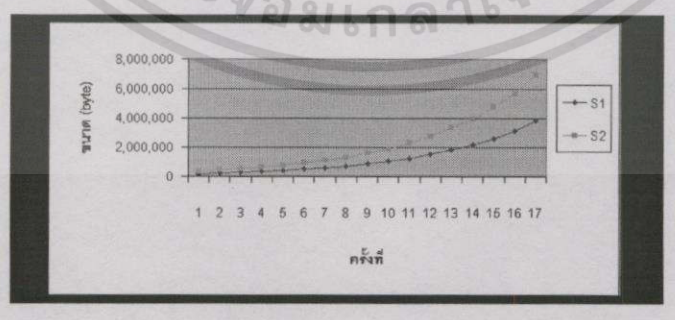
2. ด้านความตรงต่อความต้องการ วิธีการของผู้วิจัย และวิธี Breath-First Search ให้ความตรงต่อความต้องการมีค่า 100 เปอร์เซ็นต์เท่ากันทั้งสองวิธี

3. ด้านเวลาที่ใช้ในการค้นหาข้อมูล วิธีการของผู้วิจัย สามารถค้นหาข้อมูลโดยใช้เวลาน้อยกว่า วิธี Breath-First Search ดังนั้น วิธีการของผู้วิจัย ในด้านเวลาจึงมีประสิทธิภาพสูงกว่า วิธี Breath-First Search



ภาพที่ 4.5 กราฟแสดงผลวิเคราะห์ในด้านเวลาที่ใช้ในการค้นหาข้อมูล ทั้ง 2 วิธี

4. ด้านการใช้เนื้อที่ในการเก็บข้อมูลเพื่อการค้นหา วิธีการของผู้วิจัย ใช้เนื้อที่ในการเก็บข้อมูลเพื่อการค้นหาน้อยกว่า วิธี Breath-First Search ดังนั้น วิธีการของผู้วิจัยในด้านใช้เนื้อที่ในการเก็บข้อมูลเพื่อการค้นหา จึงมีประสิทธิภาพสูงกว่า วิธี Breath-First Search



ภาพที่ 4.6 กราฟแสดงผลวิเคราะห์ในด้านการใช้เนื่อที่ในการเก็บข้อมูลเพื่อการค้นหาทั้ง 2 วิธี

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5
สรุปผลการวิจัย อภิปรายผลและข้อเสนอแนะ

5.1 สรุปผลการวิจัย

5.1.1 วัตถุประสงค์ของการวิจัย

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อพัฒนาวิธีการค้นหาข้อมูลบนอินเตอร์เน็ตกับข้อมูลทะเบียนพันธุ์ไม้ดอกไม้ประดับให้มีประสิทธิภาพ และเปรียบเทียบประสิทธิภาพของการสืบค้นข้อมูลบนอินเตอร์เน็ต ด้วยวิธีการของผู้วิจัย กับวิธี Breath-First Search

5.1.2 ประชากรและกลุ่มตัวอย่าง

1. ประชากรที่ใช้ในการวิจัย คือ รายชื่อพันธุ์ไม้ดอกไม้ประดับในเมืองไทยจากพจนานุกรมไม้ดอกไม้ประดับในเมืองไทย (วิทย์ เที่ยงบูรณธรรม. 2536) และสารานุกรมไม้ดอกไม้ประดับในประเทศไทย (สวัสดิ์ หรั่งเจริญ. 2525) จำนวน 2,000 ชื่อ

2. กลุ่มตัวอย่างที่ใช้ในการวิจัย คือ รายชื่อพันธุ์ไม้ดอกไม้ประดับในเมืองไทยจากพจนานุกรมไม้ดอกไม้ประดับในเมืองไทย (วิทย์ เที่ยงบูรณธรรม. 2536) และสารานุกรมไม้ดอกไม้ประดับในประเทศไทย (สวัสดิ์ หรั่งเจริญ. 2525)จำนวน 1,849 ชื่อ

5.1.3 เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัยครั้งนี้คือ โปรแกรมการค้นหาข้อมูลบนอินเตอร์เน็ตกับข้อมูลทะเบียนพันธุ์ไม้ดอกไม้ประดับโดยวิธีการของผู้วิจัย กับวิธี Breath-First Search

5.1.4 การดำเนินการทดลอง

การดำเนินการทดลองการเปรียบเทียบประสิทธิภาพของการค้นหาข้อมูลบนอินเตอร์เน็ต ด้วยวิธีการของผู้วิจัย กับวิธี Breath-First Search ผู้วิจัยได้ทำการทดลอง 17 ครั้ง เพื่อทำการเปรียบเทียบประสิทธิภาพของวิธีการค้นหาเมื่อข้อมูลในฐานข้อมูลกำหนดโดยการทดลองเปลี่ยนแปลงข้อมูลเพิ่มขึ้นในอัตราร้อยละ 20 และเนื่องจากครั้งที่ 18 ซึ่งต้องใช้ข้อมูลจำนวน 2,219 รายการ ซึ่งข้อมูลจริงมีจำนวนไม่ถึง 2,219 รายการแล้ว และจดบันทึกผลการทดลอง การทดลองเริ่มทำการทดลองและเก็บข้อมูลตั้งแต่ 1 พฤษภาคม 2543 ถึง วันที่ 1 พฤศจิกายน 2543

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.1.5 การวิเคราะห์ข้อมูล

เมื่อดำเนินการทดลองและบันทึกผลการทดลองเรียบร้อยแล้ว ผู้วิจัยได้นำข้อมูลที่ได้มาวิเคราะห์หาประสิทธิภาพของวิธีการทั้ง 2 โดยนำข้อมูลที่ได้จากการทดลองในแต่ละครั้งมาหาประสิทธิภาพของวิธีการ และบันทึกลงในตารางแสดงประสิทธิภาพของวิธีการค้นหาของการทดลองแต่ละครั้ง จากนั้น จึงนำเอาประสิทธิภาพของวิธีการที่หาได้มาเปรียบเทียบประสิทธิภาพของวิธีการค้นหา โดยนำเสนอในรูปของกราฟเปรียบเทียบประสิทธิภาพในด้านต่าง ๆ และนำมาใช้เป็นแนวทางในการพัฒนา และ แก้ไขข้อบกพร่องของวิธีการค้นหาข้อมูลทะเบียนพันธุ์ไม้ดอกไม้ประดับต่อไปในอนาคตต่อไป

5.1.6 ผลการวิจัย

จากการวิเคราะห์เปรียบเทียบประสิทธิภาพของโปรแกรมการค้นหาข้อมูลบนอินเตอร์เน็ตกับข้อมูลทะเบียนพันธุ์ไม้ดอกไม้ประดับระหว่างวิธีการของผู้วิจัย กับวิธี Breath-First Search พบว่าวิธีการของผู้วิจัย มีประสิทธิภาพสูงกว่า วิธี Breath-First Search โดยที่ วิธีการของผู้วิจัยสามารถใช้ค้นหาข้อมูลพันธุ์ไม้ดอกไม้ประดับ ได้ครบถ้วนและตรงต่อความต้องการพอ ๆ กับ วิธี Breath-First Search แต่ใช้เวลาในการค้นหาข้อมูลและใช้เนื้อที่ในการเก็บข้อมูลเพื่อการค้นหาน้อยกว่าวิธี Breath-First Search

5.2 อภิปรายผลการวิจัย

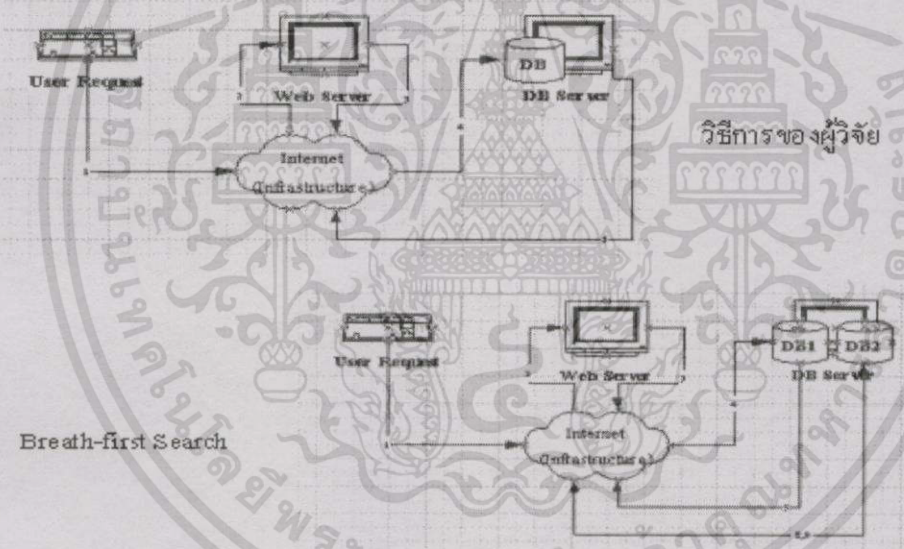
ในการวิจัยครั้งนี้ผู้วิจัยสร้างวิธีการค้นหาข้อมูล 2 แบบขึ้นมาด้วยกัน คือ วิธีการของผู้วิจัย และวิธี Breath-First Search จากนั้นนำมาทดลองค้นหาข้อมูลบนอินเตอร์เน็ตกับข้อมูลทะเบียนพันธุ์ไม้ดอกไม้ประดับ และทำการเปรียบเทียบประสิทธิภาพของการค้นหาข้อมูลบนอินเตอร์เน็ต ซึ่งข้อมูลที่นำมาใช้มาจากพจนานุกรมไม้ดอกไม้ประดับในเมืองไทย และสารานุกรมไม้ดอกไม้ประดับในประเทศไทย จำนวน 1,849 ชื่อ โดยผู้วิจัยทำการทดลองโดยค้นหาชื่อพันธุ์ไม้ดอกไม้ประดับ จำนวน 17 ครั้ง โดยแต่ละครั้งจะค้นหาชื่อรายการพันธุ์ไม้ดอกไม้ประดับโดยวิธี Breath-First Search จำนวน 100 ชื่อ และวิธีการของผู้วิจัย จำนวน 100 ชื่อในแต่ละครั้งที่ทำการทดลอง จากนั้นจึงนำผลที่ได้มาวิเคราะห์เปรียบเทียบใน 4 ด้านคือ ความครบถ้วนของข้อมูล ด้านความตรงต่อความต้องการ ด้านเวลาที่ใช้ในการค้นหาข้อมูล และ ด้านการใช้เนื้อที่ในการเก็บข้อมูลเพื่อการค้นหา

1. ผลการวิจัยในด้านความครบถ้วนของข้อมูลพบว่า วิธีการของผู้วิจัย กับวิธี Breath-First Search ได้ผลการทดลองเท่ากับ 100 เปอร์เซ็นต์เท่ากัน เนื่องจากผลจากการทดลองทั้งสองวิธีได้ข้อมูลครบถ้วน นั่นย่อมแสดงให้เห็นว่าทั้งสองวิธีมีประสิทธิภาพในความครบถ้วนของข้อมูลเท่ากัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. ผลการวิจัยในด้านความตรงต่อความต้องการพบว่า วิธีการของผู้วิจัย กับวิธี Breath-First Search ได้ผลการทดลองเท่ากับ 100 เปอร์เซ็นต์เท่ากัน เนื่องจากผลจากการทดลองทั้งสองวิธีได้ข้อมูลตรงต่อความต้องการ นั่นย่อมแสดงให้เห็นว่าทั้งสองวิธีมีประสิทธิภาพในความตรงต่อความต้องการเท่ากัน

3. ความน่าสนใจของผลการทดลองครั้งนี้อยู่ที่เวลาที่ใช้ในการค้นหาข้อมูล ผลการวิจัยในด้านเวลาที่ใช้ในการค้นหาข้อมูลพบว่า วิธีการของผู้วิจัยใช้เวลาโดยรวมน้อยกว่าวิธี Breath-First Search เนื่องจากวิธีการของผู้วิจัยจะสร้างดัชนีกับไฟล์ต่าง ๆ ทันทีเมื่อมีการร้องขอจากบราวเซอร์จะสังเกตได้ว่าวิธีการแบบนี้จะมีการติดต่อกับไฟล์ 1 ไฟล์เพียงแค่ครั้งเดียว คือเปิดออกมาสร้างดัชนีแล้วสามารถระบุได้ทันทีว่าไฟล์นี้มีคำที่กำลังค้นหาอยู่หรือไม่ แต่วิธี Breath-First Search จะมีการสร้างดัชนีเป็นตัวอักษรไว้ก่อนแล้วดังนั้นเมื่อมีการร้องขอจะต้องมีการเปิดไฟล์ดัชนีเพื่อนำค่าดังกล่าวไปเปิดไฟล์เพื่อตรวจสอบว่าเป็นคำที่กำลังค้นหาอยู่หรือไม่ จากนั้นจึงแสดงผล



ภาพที่ 5.1 แสดงขั้นตอนการค้นหาข้อมูลระหว่างวิธีการของผู้วิจัย และวิธี Breath-First Search

นั่นย่อมแสดงให้เห็นว่าวิธีการของผู้วิจัยมีประสิทธิภาพในด้านเวลาที่ใช้ในการค้นหาข้อมูลสูงกว่าวิธี Breath-First Search แต่เมื่อจำนวนข้อมูลมีมากขึ้นวิธีการของผู้วิจัยจะใช้เวลามากขึ้น เนื่องจากทุกไฟล์จะต้องถูกสร้างดัชนีก่อนเสมอก่อนจะสามารถค้นหาได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4. ผลการวิจัยในด้านการใช้เนื้อที่ในการเก็บข้อมูลเพื่อการค้นหาพบว่า วิธีการของผู้วิจัย ใช้เนื้อที่น้อยกว่าวิธี Breath-First Search เนื่องจากวิธีการของผู้วิจัยไม่จำเป็นต้องสร้างดัชนีทิ้งไว้ในหน่วยความจำ แต่วิธี Breath-First Search จำเป็นต้องสร้างดัชนีเก็บไว้ที่หน่วยความจำจึงจะสามารถค้นหาข้อมูลได้ นั่นย่อมแสดงให้เห็นว่าวิธีการของผู้วิจัยมีประสิทธิภาพในด้านการใช้เนื้อที่ในการเก็บข้อมูลเพื่อการค้นหาสูงกว่าวิธี Breath-First Search

5.3 ข้อเสนอแนะ

5.1.3 ข้อเสนอแนะเพื่อการนำผลการวิจัยไปใช้

สำหรับผู้ที่มีความสนใจจะนำวิธีการค้นหาที่ผู้วิจัยพัฒนาขึ้นไปใช้ในการค้นหาข้อมูล ควรคำนึงถึงขนาดของข้อมูลที่จะนำมาเป็นฐานข้อมูล เพราะจะมีผลต่อเวลาที่ใช้ในการค้นหาข้อมูล และมีปัจจัยอื่น ๆ อีกที่ควรคำนึงถึง เช่น ความเสถียรของระบบเครือข่าย อุปกรณ์ที่ใช้ในระบบเครือข่าย รวมทั้งปริมาณความหนาแน่นของข้อมูลบนระบบเครือข่าย เป็นต้น

5.3.2 ข้อเสนอแนะเพื่อการวิจัยต่อไป

สำหรับผู้ที่มีความสนใจที่จะศึกษาพัฒนาหรือนำไปประยุกต์ใช้ สิ่งที่มีความน่าสนใจคือ การพัฒนาต่อเพื่อให้สามารถเรียนรู้ข้อมูลที่เพิ่มเติมเข้าไปอัตโนมัติโดยไม่จำเป็นต้องมีการเรียนรู้ให้ทุกครั้ง (Out Learning) สามารถค้นหาข้อมูลที่มีลักษณะรูปแบบใกล้เคียงกับคำค้นของผู้ใช้ หรือนำไปพัฒนาเพื่อใช้กับภาษาอื่น ๆ ต่อไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บรรณานุกรม

กิดานันท์ มลิทอง. 2539. **อธิบายศัพท์คอมพิวเตอร์อินเทอร์เน็ตมัลติมิเดีย**. กรุงเทพฯ : เอดิสัน เพรสโพรดักส์.

กองส่งเสริมพืชพันธุ์, กรมวิชาการเกษตร. 2535. **คู่มือการผลิตไม้ตัดดอก**. กรุงเทพฯ : รุ่งศิลป์การพิมพ์.

ขนิษฐา รุจิโรจน์. 2537. "Internet เพื่อการศึกษา". **โครงการโสตฯ-เทคโนสัมพันธ์แห่งประเทศไทย**. ครั้งที่ 9. กรุงเทพฯ : มหาวิทยาลัยรามคำแหง.

ดวงแก้ว สวามิภักดิ์. 2535. **รู้จักกับคอมพิวเตอร์**. กรุงเทพฯ : ซีเอ็ดยูเคชั่น.

ต้น ตันฑ์สุทธิวงค์ และคณะ. 2539. **รอบรู้ INTERNET และ WORLD WIDE WEB**. กรุงเทพฯ : โปรวิชั่น.

ประสิทธิ์ คำภูแสน. 2520. **หนังสือเรียนวิชาไม้ดอกไม้ประดับ**. กรุงเทพฯ : บรรณกิจเทรดดิ้ง.

ปิฎฐะ บุนนาค. 2519. **ไม้ดอกไม้ประดับ**. กรุงเทพฯ : บรรณกิจเทรดดิ้ง.

พรทิพย์ โล่ห์เลขา. 2537. **การรับส่งจดหมายทางอิเล็คทรอนิกส์ Electronic Main (E-mail)**. กรุงเทพฯ : อุษาการพิมพ์.

พันธุ์ทิพย์ บริพัตร และคณะ. 2520. **ทะเบียนพันธุ์ไม้ดอกไม้ประดับ**. กรุงเทพฯ : บพิธการพิมพ์.

วันชัย ตันติวิทยาพิทักษ์. 2542. **ตะโกนก้องจากพงไพร รวมผลงานและความคิดของสีบ นาคะเสถียร**. พิมพ์ครั้งที่ 6. กรุงเทพฯ : วิริยะธุรกิจ.

วิจิตร วังใน และยิ่งยง ไพสุขศานติวัฒนา. 2538. **การจำแนกพืชสวน**. กรุงเทพฯ : มหาวิทยาลัยเกษตรศาสตร์.

วิทย์ เที่ยงบูรณธรรม. 2536. **พจนานุกรมไม้ดอกไม้ประดับในเมืองไทย**. กรุงเทพฯ : บรรณกิจเทรดดิ้ง.

สวัสดิ์ หรั่งเจริญ. 2525. **สารานุกรมไม้ประดับในประเทศไทย**. เล่มที่ 3. กรุงเทพฯ : อมรินทร์การพิมพ์

สัจจะ จรัสรุ่งวงวีวร และสมพร จิวรสกุล. 2542. Active Server Pages **และแอพพลิเคชันฐานข้อมูลสำหรับอินเตอร์เน็ต**. กรุงเทพฯ : ด่านสุทธาการพิมพ์.

สุรศักดิ์ สงวนพงษ์. 2538. **คู่มืออินเตอร์เน็ต**. กรุงเทพฯ : ซีเอ็ดยูเคชั่น.

Donald H. Kraft, Gloria Bordogna and Gabriella Pasi. 1998. **Information Retrieval Systems: Where is Fuzz?**. New York : IEEE Press.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Ellman, J and Tait, J. 1997. **Intelligent World Wide Web Agents Using Information Density to navigate the Web.** [CD-ROM]. Abstract from : Dissertation Abstracts Item : 5602561

G. Michael Youngblood. 1999. **Web Hunting : Design of a Simple Intelligent Web Search Agent.** [Online]. Available : http://www.acm.org/crossroads/xrds5-4/webhunting.html

George F. Luger and William A. Stubblefield. 1993. **Aritificial Intelligence.** Second Edition. Massachusetts : The Benjamin/Cummings Publishing Company, Inc.

Gilles Brassard and Paul Bratley. 1988. **Algorithmics Theory and Practice.** New Jersey : Prentice-Hall, Inc.

Greg W. Scragg. 1996. **Problem Solving with Computers.** Massachusetts : Jones and Bartlett Publishers.

Mark J. Fold and Bill Zoellick. 1992. **File Structures.** Second Edition. Massachusetts : Addison-Wesley Publishing Company, Inc.

Mark Lagar. 1999. **Spinning a Web Search.** [Online]. Available : http://www.library.ucsb.edu/untangle/lager.html

P. Kefalas and E. Tzelalis. 1995. **GRS-Prolog.** [Online]. Available : http://senanet.com/epy/09/www.grs_prol.html

Paul C. Jorgensen. 1995. **Software Testing.** New York : CRC Press, Inc.

Richard Entlich et. al. 1997. "Making a Digital Library: The Content of the CORE Project." **ACM Transaction on Information Systems.** 15(2) : 103-123.

Robert R. Korfhage. 1997. **Information Storage and Retrieval.** New York : John Wiley & Sons, Inc.

Thomas A. Standish. 1994. **Data Structure, Algorithms, and Software Principles.** Massachusetts : Addison-Wesley Publishing Company, Inc.

Thomas C. Royer. 1993. **Software Testing Management.** New Jersey : Prentice-Hall, Inc.

UC Irvine. 1996. **Design and Analysis of Algorithms.** [Online]. Available : http://www.ics.uci.edu/~eppstein/161/960215.html

William B. Frakes. 1992. **Information Retrieval Data Structure & Algorithms.** New Jersey : Prentice-Hall, Inc.

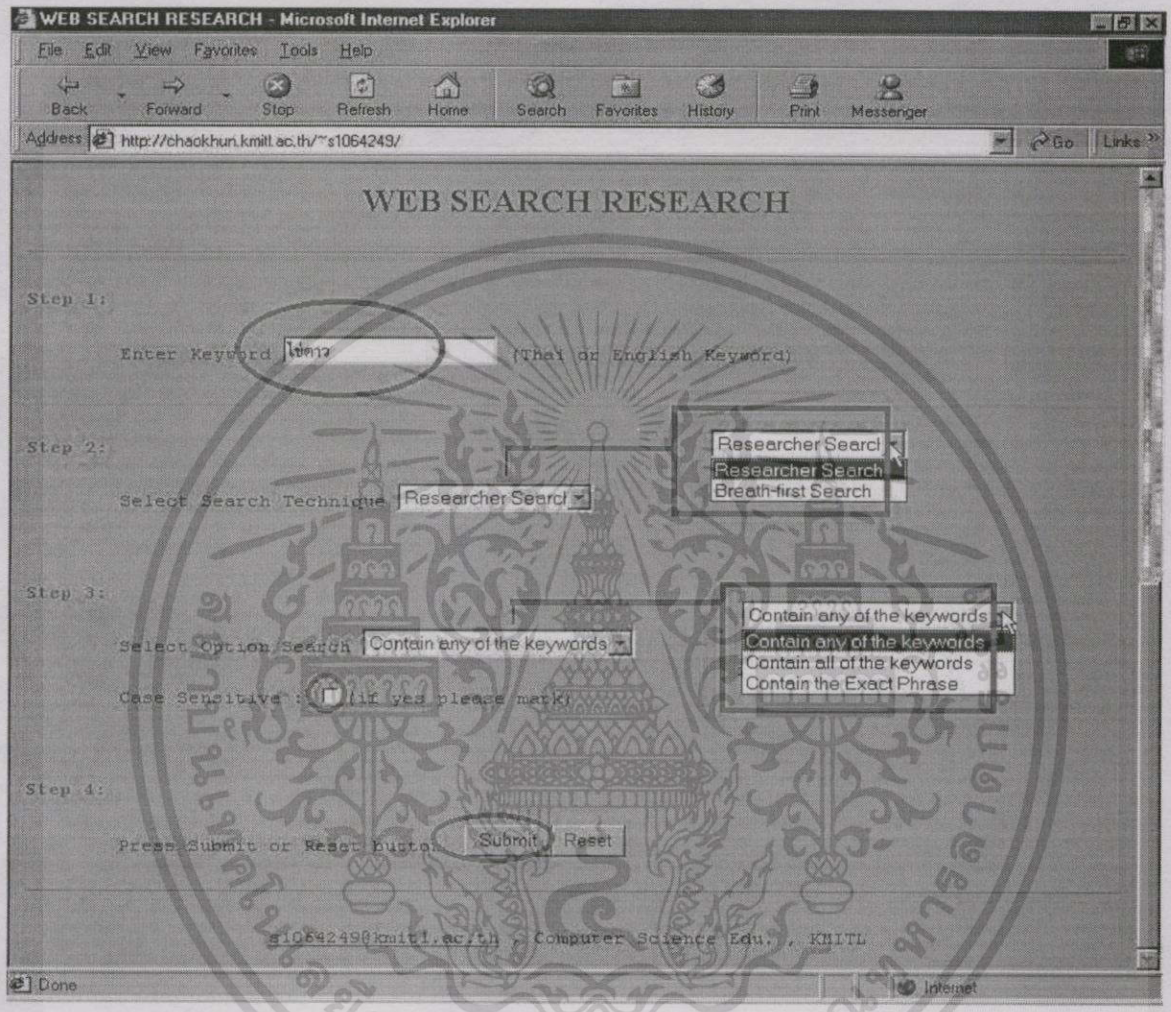
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ก คู่มือการใช้งาน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

คู่มือการใช้งาน

เมื่อผู้ใช้ใส่ URL : http://chaokhun.kmitl.ac.th/~s1064249 จะปรากฏหน้าจอดังนี้



หน้าหลักนี้จะถูกแบ่งออกเป็น 4 ส่วนให้ผู้ใช้สามารถปฏิบัติตามขั้นตอนดังต่อไปนี้

ขั้นแรก :- ให้ผู้ใช้ป้อนคำค้น ซึ่งเป็นคำ (ถ้ามีมากกว่า 1 คำ ให้เคาะเว้นวรรค 1 เคาะเพื่อพิมพ์คำต่อไป) ในภาพตัวอย่าง ป้อนคำค้น "ไข่ดาว" ลงไป (ในวงกลม)

ขั้นที่สอง :- ให้ผู้ใช้เลือกวิธีการค้นหาข้อมูล ซึ่งมี 2 วิธีด้วยกัน คือ Researcher Search และ Breath-first Search ในตัวอย่าง เลือกใช้ Researcher Search

ขั้นที่สาม :- ให้ผู้ใช้กำหนดรูปแบบการค้นหา มีให้เลือก 3 แบบด้วยกันคือ

- Contain any of the keywords
- Contain all of the keywords
- Contain the Exact Phrase

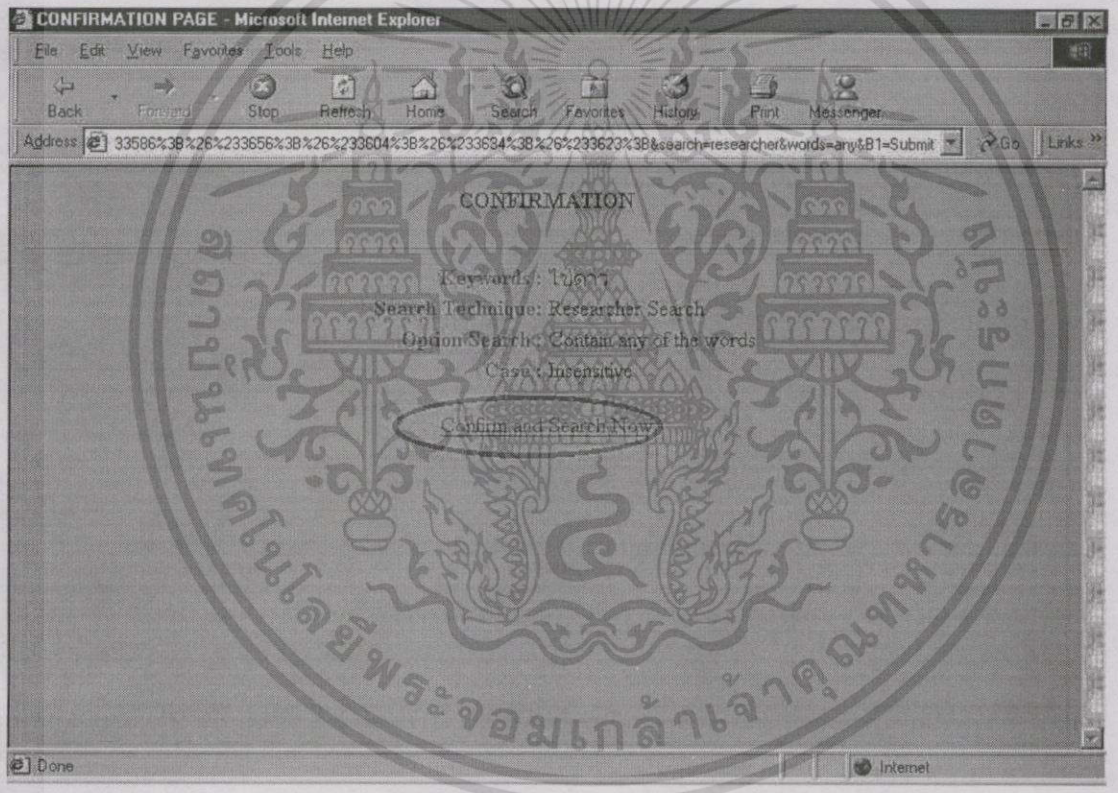
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หากต้องการให้การค้นหาตรวจสอบลักษณะของตัวอักษรให้ตรงกับ Keywords ก็ให้ check ที่ Case Sensitive (Default : Case Non-Sensitive)

ขั้นที่สี่ :- ลงมือค้นหา โดยกดปุ่ม Submit หรือ ยกเลิกข้อมูลที่กรอกไว้เพื่อค้นหากดปุ่ม Reset

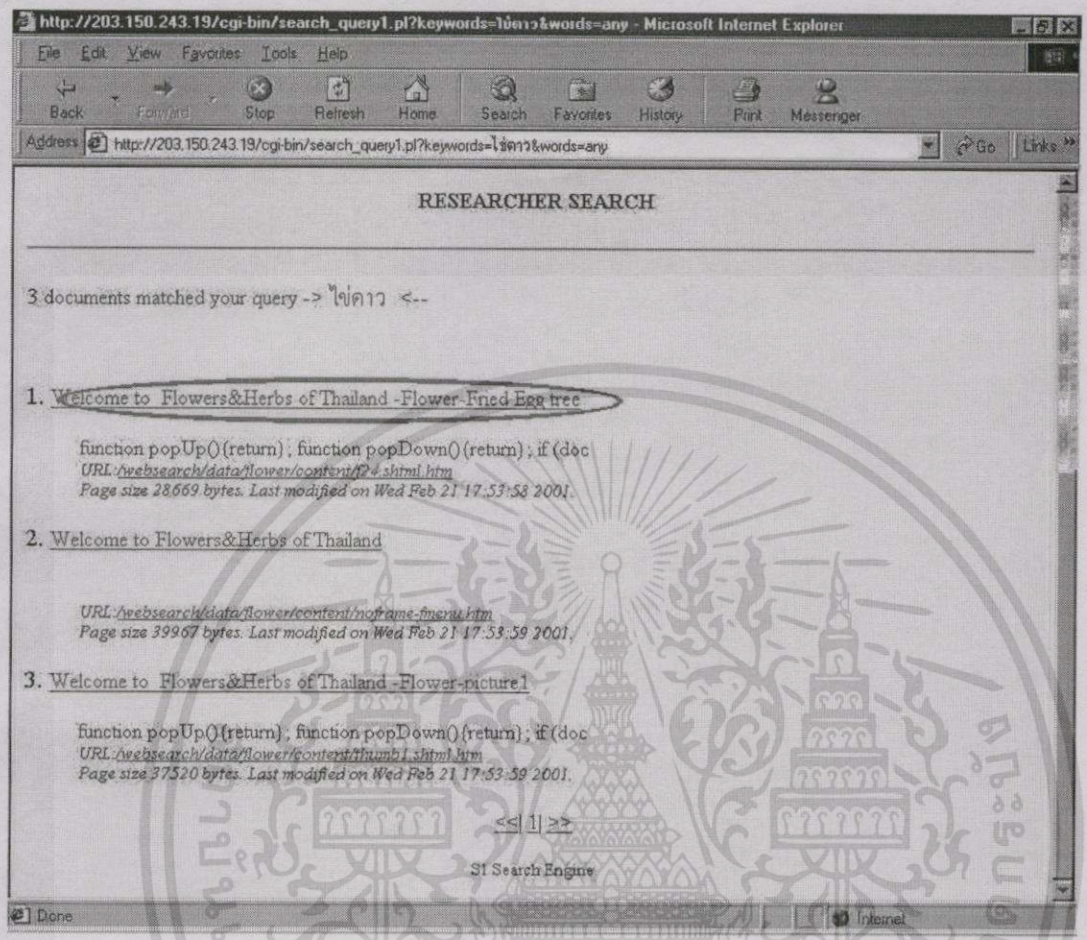
---

เมื่อคลิ๊กที่ปุ่ม Submit แล้วจะมีหน้าจอขึ้นมาเพื่อให้ Confirm ว่าข้อมูลถูกต้องตามที่ต้องการ ให้คลิ๊กที่ปุ่ม Confirm and Search Now หรือกดปุ่ม Back ที่บราวเซอร์เพื่อกลับไปกรอกข้อมูลเพื่อค้นหาใหม่ ดังรูป



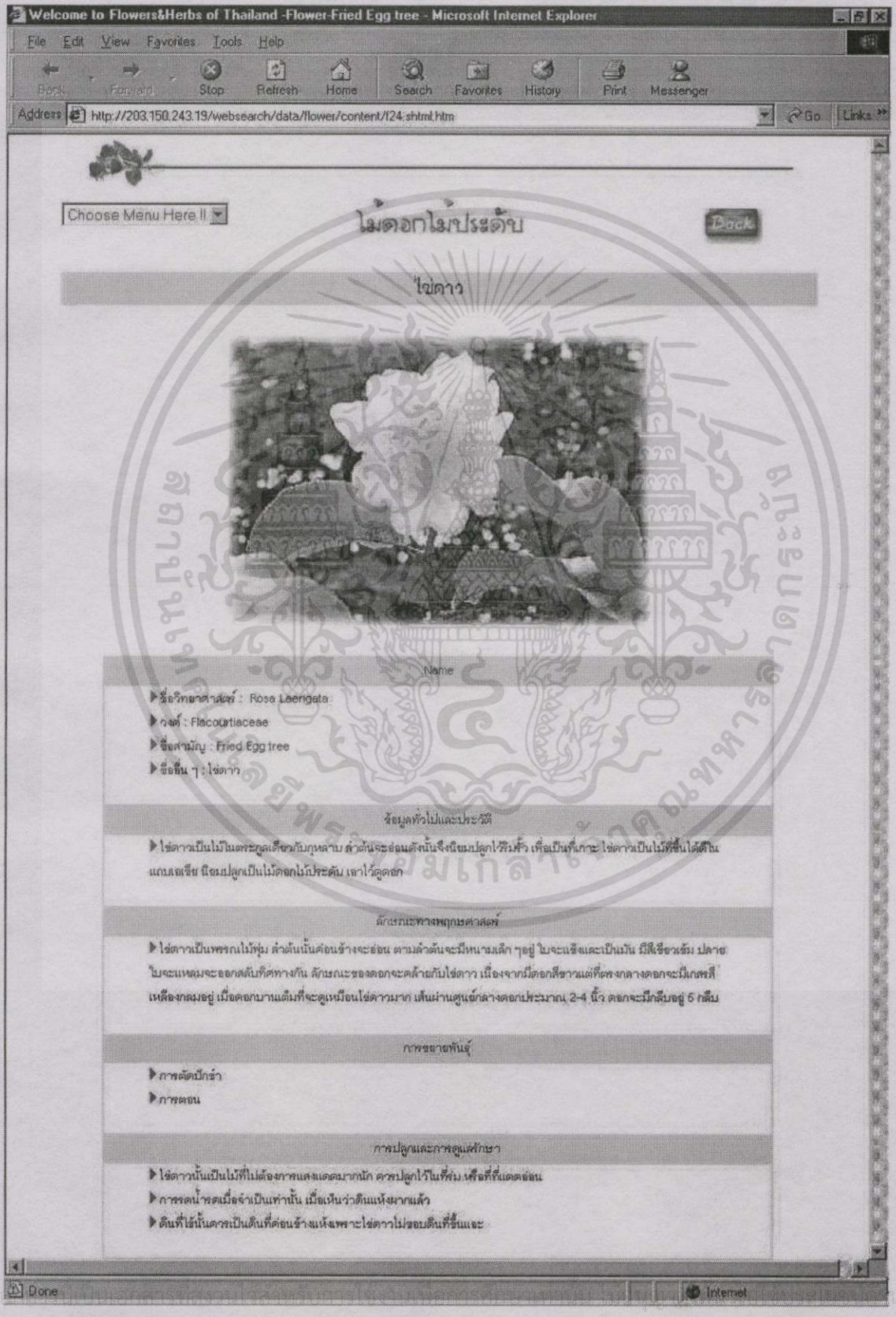
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หน้าจอแสดงผลการค้นหาจะปรากฏขึ้น ให้ผู้ใช้เลือกข้อมูลที่ต้องการดูโดยคลิ๊กที่ชื่อดังรูป



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เมื่อคลิ๊กเข้าไปแล้วก็จะเข้าสู่เพจรายละเอียดของข้อมูลที่ถูกค้นพบ ดังรูป



ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ประวัติผู้เขียน

ชื่อ – สกุล	นายพุธ นาฑีสุวรรณ
วัน เดือน ปี เกิด	27 ตุลาคม 2519
สถานที่เกิด	กรุงเทพมหานคร
สถานที่อยู่ปัจจุบัน	บ้านเลขที่ 48/38 ถนนกรุงเทพ-นนท์ บางซื่อ กรุงเทพมหานคร
สถานที่ทำงาน	ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ
ตำแหน่ง	ผู้ช่วยนักวิจัย
ประวัติการศึกษา	ปีการศึกษา 2540 สำเร็จการศึกษา วิทยาศาสตรบัณฑิต (วิทยาการคอมพิวเตอร์) จากมหาวิทยาลัยมหิดล ปีการศึกษา 2544 สำเร็จการศึกษา วิทยาศาสตรมหาบัณฑิต (การศึกษาวิทยาศาสตร์ คอมพิวเตอร์) จากสถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้