

การเข้าถึงฐานข้อมูลโดยภาษาธรรมชาติ



นายวิเชียร ศรีพระจันทร์

ศูนย์วิทยทรัพยากร

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

ภาควิชาวิศวกรรมคอมพิวเตอร์

บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย

พ.ศ. 2539

ISBN 974-8247-34-1

ลิขสิทธิ์ของบัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย

DATABASE ACCESS BY A NATURAL LANGUAGE



MR.WICHAIN SRIPRACHAN

**A Thesis Submitted in Partial Fullfillment of the requirements
for The Degree of Master of Science
Department of Computer Engineering**

Graduate School

Chulalongkorn University

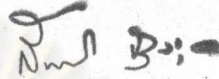
1996

ISBN 974-8247-34-1

หัวข้อวิทยานิพนธ์ การเข้าถึงฐานข้อมูลโดยภาษาธรรมชาติ
โดย นายวิเชียร ศรีพระจันทร์
ภาควิชา วิศวกรรมคอมพิวเตอร์
อาจารย์ที่ปรึกษา รองศาสตราจารย์เดือน สินธุ์พันธ์ประทุม

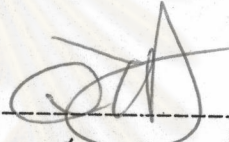


บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้บัณฑิตวิทยาลัยดำเนินการศึกษาตามหลักสูตรปริญญาโทบัณฑิต

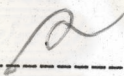


----- คณบดีบัณฑิตวิทยาลัย
(รองศาสตราจารย์ ดร. สันติ ฤงสุวรรณ)

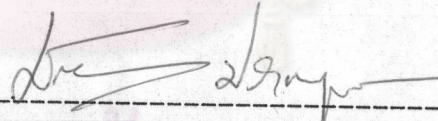
คณะกรรมการสอบวิทยานิพนธ์



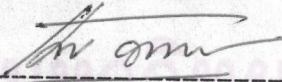
----- ประธานกรรมการ
(รองศาสตราจารย์ ดร. วันชัย รั้วไพบุลย์)



----- อาจารย์ที่ปรึกษา
(รองศาสตราจารย์เดือน สินธุ์พันธ์ประทุม)



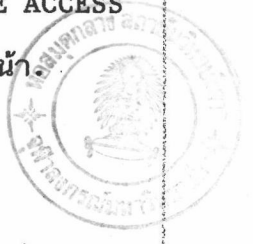
----- กรรมการ
(รองศาสตราจารย์มณฑนา ปราการสมุทร)



----- กรรมการ
(ผู้ช่วยศาสตราจารย์ชูชีพ ฉิมวงษ์)

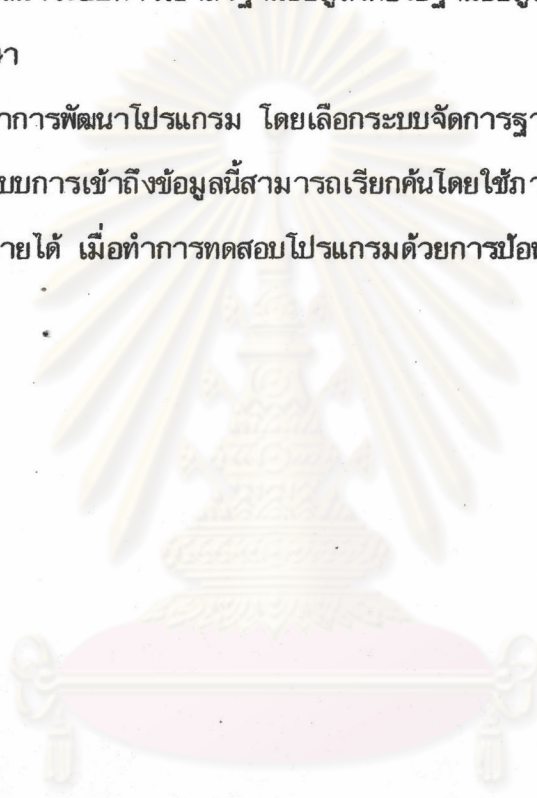
พิมพ์ต้นฉบับบทคัดย่อวิทยานิพนธ์ภายในกรอบสี่เหลี่ยมนี้เพียงแผ่นเดียว

นายวิเชียร ศรีพระจันทร์ : การเข้าถึงฐานข้อมูลโดยภาษาธรรมชาติ (DATABASE ACCESS BY A NATURAL LANGUAGE) อ.ที่ปรึกษา : รศ.เดือน สิริพันธุ์ประทุม, 98 หน้า
ISBN 974-8247-34-1





วิทยานิพนธ์ฉบับนี้มีวัตถุประสงค์เพื่อพัฒนาการเข้าถึงฐานข้อมูลโดยภาษาธรรมชาติ การพัฒนาระบบจะเริ่มศึกษาขั้นต้น การประมวลผลภาษาธรรมชาติ วิเคราะห์ปัญหา และการเรียกค้นข้อมูลที่กระทำอยู่ในปัจจุบัน การพัฒนาระบบการเข้าถึงฐานข้อมูลโดยใช้ฐานข้อมูลการลงทะเบียนของสถาบันราชภัฏสวนสุนันทา เป็นกรณีศึกษา

การวิจัยนี้ ได้ทำการพัฒนาโปรแกรม โดยเลือกระบบจัดการฐานข้อมูลแบบสัมพันธ์พัฒนาภายใต้ระบบปฏิบัติการวินโดวส์ ระบบการเข้าถึงข้อมูลนี้สามารถเรียกค้นโดยใช้ภาษาธรรมชาติที่เป็นคำ วลี หรือประโยคภาษาอังกฤษอย่างง่ายได้ เมื่อทำการทดสอบโปรแกรมด้วยการป้อนข้อความให้โปรแกรมทำงานได้ผลลัพธ์เป็นที่น่าพอใจ



ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ภาควิชา วิศวกรรมคอมพิวเตอร์
สาขาวิชา วิทยาศาสตร์คอมพิวเตอร์
ปีการศึกษา 2538

ลายมือชื่อนิสิต 
ลายมือชื่ออาจารย์ที่ปรึกษา 
ลายมือชื่ออาจารย์ที่ปรึกษาร่วม

C4:7841 : MAJOR COMPUTER SCIENCE

KEY WORD : ACCESS.DATABASE,NATURAL LANGUAGE

WICHIAN SRIPRACHAN : DATABASE ACCESS BY NATURAL
LANGUAGE

THESIS ADVISOR : DUAN SINTUPANPRATUM,Associate Professor, 98 pp.

ISBN 974-8247-34-1

The objectives of this thesis were to develop Database Access by Natural Language . Steps of developing natural language processing, the access methods, and the data query were studied and analysed Rajbhat Institute Suansunandha was used for the case study.


A package program was developed based on relational database approach and operated under windows environment. The database access system can be queried with words, phrase, or English simple sentences by natural language. The program was tested with some sentences input, the results were satisfactory.

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ภาควิชา วิศวกรรมคอมพิวเตอร์

สาขาวิชา วิทยาศาสตร์คอมพิวเตอร์

ปีการศึกษา 2538

ลายมือชื่อนิสิต 

ลายมือชื่ออาจารย์ที่ปรึกษา 

ลายมือชื่ออาจารย์ที่ปรึกษาร่วม



กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงไปได้ด้วยความช่วยเหลือทางด้าน วิชาการ ข้อมูล เอกสาร และคำแนะนำเป็นอย่างดีของคุณเมธี เสรีอรุณ ซึ่งเป็นนักวิเคราะห์ระบบงาน คอมพิวเตอร์ บริษัทวิทยุการบินแห่งประเทศไทย และพร้อมกันนี้ขอขอบพระคุณรองศาสตราจารย์ เดือน สินธุ์พันธุ์ประทุม อาจารย์ที่ปรึกษาวิทยานิพนธ์ ซึ่งท่านได้ให้คำปรึกษา ตรวจสอบ และ ให้ข้อคิดเห็นในการจัดทำวิทยานิพนธ์มาโดยตลอด รวมถึงสถาบันราชภัฏสวนสุนันทาที่ได้ อนุเคราะห์ สนับสนุน ในเรื่องการพัฒนาบุคลากรในสาขาที่ขาดแคลน ท้ายนี้ผู้วิจัยใคร่ขอกราบ ขอบพระคุณบิดามารดา ซึ่งได้เลี้ยงดูและให้การศึกษามาโดยตลอด อีกทั้งภรรยาและลูกที่ช่วยให้ กำลังใจ จนวิทยานิพนธ์ฉบับนี้ สำเร็จลุล่วงไปด้วยดี

วิเชียร ศรีพระจันทร์

ศูนย์วิทยุทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญ

	หน้า
บทคัดย่อไทย.....	ง
บทคัดย่ออังกฤษ.....	จ
กิตติกรรมประกาศ.....	ฉ
สารบัญตาราง.....	ญ
สารบัญภาพ.....	ฎ
บทที่	
1. บทนำ	
ความเป็นมาและความสำคัญของปัญหา.....	1
แนวทางแก้ไข.....	1
วัตถุประสงค์.....	1
ขอบเขตในการวิจัย.....	1
ขั้นตอนและวิธีการดำเนินงาน.....	3
ประโยชน์ที่คาดว่าจะได้รับ.....	3
2. การเข้าถึงฐานข้อมูลและการลงทะเบียนของนักศึกษา	
ฐานข้อมูลที่ใช้เรียกค้น.....	4
ตัวแบบความสัมพันธ์ของข้อมูล.....	9
การใช้ภาษาสอบถามเชิงโครงสร้างในระบบงานเดิม.....	10
ความสามารถของภาษาสอบถามเชิงโครงสร้าง.....	11
ตัวอย่างผังไวยากรณ์คำสั่ง Select และเงื่อนไขในการค้นหา.....	12
ตัวอย่างการใช้คำสั่งภาษาสอบถามเชิงโครงสร้างในระบบงานเดิม	13
โปรแกรมเคลฟ.....	15
หลักการเขียนโปรแกรมวิซวล.....	17

สารบัญ (ต่อ)

	หน้า
3. แนวคิดและทฤษฎีชาติ	
ยุคของการประมวลข้อมูลข่าวสาร.....	22
ระบบภาษาธรรมชาติ.....	23
ตัวแบบระบบภาษาธรรมชาติ.....	24
การประมวลผลภาษาธรรมชาติในระดับต่าง ๆ.....	25
เทคนิคต่าง ๆ ในการประมวลผลภาษาธรรมชาติ.....	27
4. การออกแบบและพัฒนาโปรแกรม	
การพัฒนาโปรแกรมการเข้าถึงระบบข้อมูลโดยภาษาธรรมชาติ...	42
ส่วนประกอบของเพจที่ใช้ป้อนข้อมูลภาษาธรรมชาติ.....	44
การออกแบบส่วนควบคุมเพจ.....	45
ส่วนของการออกจากโปรแกรม.....	50
คำแนะนำสำหรับวิธีการใช้โปรแกรม.....	50
ส่วนของกระบวนการในโปรแกรม.....	51
5. การทดสอบโปรแกรม	
ขั้นตอนการเข้าสู่โปรแกรม.....	57
ส่วนประกอบของเพจ 1.....	59
ส่วนประกอบของเพจ 2.....	61
การทำงานในเพจ 1.....	62
การประมวลผลภาษาธรรมชาติแบบมีเงื่อนไข.....	65
6. สรุปผลและเสนอแนะ	
สรุปผลการวิจัย.....	73
ข้อเสนอแนะ.....	73

สารบัญ (ต่อ)

	หน้า
รายการอ้างอิง.....	75
ภาคผนวก ก.....	77
ภาคผนวก ข.....	88
ประวัติผู้เขียน.....	98



ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญตาราง

		หน้า
ตารางที่ 1.	แสดงโครงสร้างของตารางอาจารย์.....	5
ตารางที่ 2.	แสดงโครงสร้างของตารางนักศึกษา.....	5
ตารางที่ 3.	แสดงโครงสร้างของตารางวิชา.....	6
ตารางที่ 4.	แสดงโครงสร้างของตารางอาจารย์-นักศึกษา.....	6
ตารางที่ 5.	แสดงโครงสร้างของตารางอาจารย์-วิชา.....	6
ตารางที่ 6.	แสดงโครงสร้างของตารางนักศึกษา-วิชา.....	7
ตารางที่ 7.	แสดงโครงสร้างของตารางวิชาเอก.....	7
ตารางที่ 8.	แสดงโครงสร้างของตารางวิชาโท.....	7
ตารางที่ 9.	แสดงโครงสร้างของตารางคุณวุฒิของอาจารย์.....	8
ตารางที่ 10.	แสดงโครงสร้างของตารางคณะวิชา.....	8
ตารางที่ 11.	แสดงคำและชนิดของคำ.....	28

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญรูปภาพ

		หน้า
รูปที่ 1.	ตัวแบบความสัมพันธ์ของข้อมูลการลงทะเบียน.....	9
รูปที่ 2.	ผังไวยากรณ์คำสั่ง Select.....	12
รูปที่ 3.	ตัวแบบระบบภาษาธรรมชาติ.....	24
รูปที่ 4.	วิธีการวิเคราะห์แบบกระจาย.....	29
รูปที่ 5.	การวิเคราะห์ประโยค.....	31
รูปที่ 6.	แสดงการใช้กฎการผลิต.....	32
รูปที่ 7.	การกระจายแบบต้นไม้.....	33
รูปที่ 8.	รูปแบบไวยากรณ์แบบการกำจัดคำออกไป.....	37
รูปที่ 9.	ขั้นตอนการออกแบบการประมวลผลภาษาธรรมชาติ....	42
รูปที่ 10.	ผังงานออบเจกต์EDIT7CHANGE.....	45
รูปที่ 11.	ผังงานออบเจกต์SpeedButton5Click.....	46
รูปที่ 12.	ผังงานSpeedButton8Click.....	47
รูปที่ 13.	ผังงานออบเจกต์Combox2Change.....	48
รูปที่ 14.	ผังงานออบเจกต์Combobox3Change.....	48
รูปที่ 15.	ผังงานออบเจกต์FormCreate.....	49
รูปที่ 16.	ผังงานออบเจกต์BitBtn1Click.....	50
รูปที่ 17.	ผังงานออบเจกต์BitBtn2Click.....	50
รูปที่ 18.	ผังงานกระบวนการAlltrim.....	51
รูปที่ 19.	ผังงานกระบวนการChktable.....	52
รูปที่ 20.	ผังงานกระบวนการChkfield.....	53
รูปที่ 21.	ผังงานกระบวนการStrUcase.....	54
รูปที่ 22.	ผังงานกระบวนการChkfdall.....	55
รูปที่ 23.	แสดงสัญรูปของโปรแกรม.....	57
รูปที่ 24.	แสดงกลุ่มของโปรแกรมที่จะใช้งาน.....	58
รูปที่ 25.	แสดงส่วนประกอบของเพจ 1.....	60
รูปที่ 26.	แสดงส่วนประกอบของเพจ 2.....	61

สารบัญรูปร่าง (ต่อ)

	หน้า
รูปที่ 27. แสดงการประมวลผลภาษาธรรมชาติแบบไม่มีเงื่อนไข...	63
รูปที่ 28. แสดงการประมวลผลภาษาสอบถามเชิงโครงสร้าง.....	64
รูปที่ 29. แสดงการประมวลผลแบบมีเงื่อนไข.....	65
รูปที่ 30. แสดงการประมวลผลแบบมีเงื่อนไข ที่เป็นภาษาสอบถามเชิงโครงสร้าง.....	66
รูปที่ 31. แสดงการประมวลผลภาษาธรรมชาติ แบบมีเงื่อนไขชนิด "OR".....	67
รูปที่ 32. แสดงการประมวลผลภาษาสอบถามเชิงโครงสร้าง แบบมีเงื่อนไข "OR".....	68.
รูปที่ 33. แสดงการประมวลผลภาษาธรรมชาติ แบบมีเงื่อนไขชนิด "OR" โดยการสลับขอบเขตข้อมูล	69
รูปที่ 34. แสดงการประมวลผลภาษาธรรมชาติ แบบมีเงื่อนไขชนิด "AND"	70
รูปที่ 35. แสดงการประมวลผลแบบสอบถามเชิงโครงสร้าง แบบมีเงื่อนไขชนิด "AND".....	71
รูปที่ 36. แสดงการประมวลผลภาษาธรรมชาติ แบบมีเงื่อนไขชนิด "AND" โดยการสลับขอบเขตข้อมูล	72