

อภินันทนาการ

รายงานการวิจัย



ชื่อโครงการ

(ภาษาไทย) การปรับปรุงระบบตรวจสอบชื่อไทยโดยใช้หลักการ Statistical Ontology

(ภาษาอังกฤษ) An Improvement of Web-based Thai Name Checking System
Using Statistical Ontology

ผู้วิจัย

หัวหน้าโครงการวิจัย

ดร. จักรกฤษณ์ เสน่ห์

สังกัดภาควิชาภysics คณิตศาสตร์และเทคโนโลยีสารสนเทศ

คณะวิทยาศาสตร์ มหาวิทยาลัยแม่ฟ้าฯ

กิตติกรรมประกาศ

โครงการที่ได้รับทุนสนับสนุนจาก เงินงบประมาณรายได้ประจำปี 2550 กองทุนวิจัย แผนงานวิจัย
งานวิจัย พัฒนาและถ่ายทอดเทคโนโลยี หมวดเงินอุดหนุน ทุนอุดหนุนการวิจัย ของสำนักงาน
เลขานุการคณะวิทยาศาสตร์ จึงขอขอบคุณไว้ ณ ที่นี่ด้วย

สำนักหอสมุด มหาวิทยาลัยแม่ฟ้าฯ
วันลงทะเบียน..... 13 JUL 2011
เลขทะเบียน..... 15678365
เลขเรียกหนังสือ..... ๒ ๖๘
๑๙
พ.ศ.
๒๕๕๐

บทคัดย่อ

ชื่อคนไทยเริ่มมีวิถีตามนากการขึ้นมาเรื่อยๆจนถึงปัจจุบัน การตั้งชื่อจึงถือเป็นสิ่งสำคัญอย่างหนึ่งในชีวิตประจำวันของแต่ละบุคคล โดยเฉพาะผู้ที่เป็นบิดา มารดา เมื่อได้บุตรแล้วก็ต้องการได้ชื่อบุตรที่มีความไฟแรง มีความหมายที่ดีและเป็นสิริมงคล เพราะมีความเชื่อที่ว่าชื่อที่มีความหมายดีจะนำโชคดีตามหลักการตั้งชื่อ ย่อมเป็นสิริมงคลที่จะทำให้ชีวิตมีความสุขความเจริญยิ่งขึ้น หลักการตั้งชื่อมีอยู่หลายวิธี อาทิ เช่น หลักทักษะปกรณ์ หลักตัวเลข และหลักให้ราศีสัตร์ แต่ปัจจุบันการตั้งชื่อที่ดีเป็นสิริมงคลและนิยมกันอย่างแพร่หลายก็คือการตั้งชื่อด้วยใช้หลักทักษะปกรณ์

ในงานวิจัยครั้งนี้เราได้ปรับปรุงระบบการตรวจสอบชื่อเดินแบบทักษะปกรณ์ที่ใช้คำความดีทางสถิติเป็นตัวตรวจสอบและวิเคราะห์ชื่อดีของ กนกกาญจน์ และจักรกฤษณ์ (กนกกาญจน์ และจักรกฤษณ์, 2006a) (จักรกฤษณ์และ กนกกาญจน์, 2007a) โดยระบบที่ปรับปรุงใหม่นี้จะใช้หลักการอ่อนโน้มใจเชิงสถิติเป็นตัวช่วยในการตรวจสอบชื่อดีให้มีความน่าเชื่อถือมากยิ่งขึ้น โดยหลักการอ่อนโน้มใจเชิงสถิตินั้นเป็นตัวกำหนดขอบเขตการแบ่งกลุ่มค่าความดีทางสถิติ (กนกกาญจน์และจักรกฤษณ์, 2006a) และเป็นตัวช่วยในการวิเคราะห์และตรวจสอบชื่อดีให้มีความน่าเชื่อถือมากยิ่งขึ้น โดยระบบปรับปรุงขึ้นใหม่นี้ได้ถูกนำไปทดสอบระบบ โดยใช้ฐานข้อมูลของนิสิตมหาวิทยาลัยนรศวร จำนวน 11,658 ชื่อ

ABSTRACT

Naming has been continuously developed from the past to the present and has evolved into a variety of patterns. In many cultures naming is not only important because every individual needs to have a name but also to have helpful names or names with a good sound. The basic goal of naming is to provide a good fortune and progress during life. Thai parents always try to choose names which they feel will bring good luck to their offsprings and to the family. Three methodologies of Thai naming use Thai astrology, a new numeric methodology, or traditional calendar. The principal naming using Thai astrology is widely used since former times as it involves the birthday in order to form the name.

In this research we propose an improvement of a web-based Thai name checking system based on the value of a statistical model proposed by Kanokkarn and Chakkrit (Kanokkarn and Chakkrit, 2006a) (Chakkrit and Kanokkarn, 2007a). The new approach uses the concept of a statistical ontology to help in defining a group of statistical values and to improve the reliability of Thai name checking. Furthermore, we present the improvement of Thai name checking system which is tested with a database of 11,658 student names of Naresuan University.

การตีพิมพ์เผยแพร่ผลงานวิจัย

1. Kanokkarn Namahoot and Chakkrit Snae "An Analysis of Thai Names using Name Checking System with Clustering Based on Statistical Values", The 1st Naresuan Science Conference (NSC1)"15-16 March, Phitsanulok, Thailand, 2007.
2. Kanokkarn Namahoot and Chakkrit Snae "An Addition of Realiable Values in Name Checking System Using Statistical Ontology The Third Naresuan Research Conference, 28-29 July, Phitsanulok. Thailand. 2007.
3. Chakkrit Snae and Kanokkarn Namahoot, "WEb-Based Thai Name Checking System using Thai Astrology Naming and Clustering Techniques (To appear) Naresuan University Science Journal



An Analysis of Thai Names using Name Checking System with Clustering Based on Statistical Values

Kanokkarn Namahoot and Chakkrit Snae*

Department of computer Science and Information Technology

Naresuan University, Phitsanulok, 65000, Thailand

Naming has been continuously developed from the past to the present and has evolved into a variety of patterns. In many cultures naming is not only important because every individual needs to have a name but have helpful names or names with a good sound. Thai parent always try to choose names which they feel will bring good luck to their offsprings and to the family. The choice of appropriate names bases on the rules of available letters that can influence the destiny of the individuals as described in principal naming using Thai astrology.

In this paper we propose a web-based Thai name checking system which uses the Thai astrology naming concept and clustering techniques based on a statistical model and probability to check and analyze how good names in the Thai cultural environment with respect to fortune and sound are. The system is developed to help analyzing recent Thai naming habits and the naming conventions in the past 20 years.

The system was implemented and designed using a rapid prototyping approach. Written specifications were implemented and rough prototypes were assessed. Evolutionary prototyping will also be used in the user interface design, which is based on the idea of developing an initial implementation of a simple command subset and refining through many stages until an adequate system and final implementation will be developed. This process will include the use of iterative enhancement techniques and iterative procedure development, which underwent several iterations between prototyping and user testing before final implementation.

We concentrate on the analysis of how names have been adopted as "good names", how much they followed the concept of Thai astrology and whether or not they contain letters out of the misfortune attribute set. In this research the Thai name checking system is tested with a database of 3,000 student names of Nakhon Sawan province, Thailand. From the testing and analysis we find that most of the students (approximately 67%) in Nakhon Sawan province have good names. Therefore, we can conclude that most parents are concerned with Thai naming and choose good names for their children.

Keywords- Thai naming system, statistical model, clustering, probability

*Corresponding author. Tel.: 0-5526-1000 ext 3223; Fax 0-5526-1025
E-mail: chakkrits@nu.ac.th

การเพิ่มค่าความเชื่อมั่นในระบบการตรวจสอบชื่อตีโดยการใช้หลักการอนโนทัลจีเจิงสถิติ

กนกกาญจน์ นามนุด และจักรกฤษณ์ เสน่ห์
ภาควิชาจิตยารักษาระบบทิพย์ เศรษฐศาสตร์ มหาวิทยาลัยนเรศวร จังหวัดพิษณุโลก 65000

*Corresponding author. E-mail: chakkrits@nu.ac.th

บทคัดย่อ: ชื่อคนไทยเริ่มนิรภัยและการชื่อ命名เรื่อยๆ จนถึงปัจจุบัน การตั้งชื่อจึงถือเป็นสิ่งสำคัญอย่างหนึ่งในชีวิตประจำวันของแต่ละบุคคล โดยเฉพาะผู้ที่เป็นบิดา มารดา เมื่อได้บุตรแล้วก็ต้องการให้ชื่อบุตรที่มีความหมายอะไร นี้ความหมายที่ดีและเป็นศิริมงคล เพราะมีความเชื่อว่าชื่อที่มีความหมายดีจะถูกต้องตามหลักการตั้งชื่อ ย่อมเป็นศิริมงคลที่จะทำให้ชีวิตมีความสุขความเจริญยิ่งขึ้น หลักการตั้งชื่อที่บุญล้ำวิธี อาทิเช่น หลักทักษะปกรณ์ หลักดัวเลข และหลักในราศีศาสตร์ แต่ปัจจุบันการตั้งชื่อที่ดีเป็นศิริมงคลและนิยมกันอย่างแพร่หลายก็คือการตั้งชื่อด้วยใช้หลักทักษะปกรณ์

ในงานวิจัยครั้งนี้เราได้ปรับปรุงระบบการตรวจสอบชื่อตามแบบทักษะปกรณ์ที่ใช้คำว่า “ดี” ทางสถิติเป็นตัวกรองและวิเคราะห์ที่ดีของogn กัญจน์และจักรกฤษณ์ (กนกกาญจน์และจักรกฤษณ์, 2006) (จักรกฤษณ์และกนกกาญจน์, 2007) โดยระบบที่ปรับปรุงใหม่นี้จะใช้หลักการอนโนทัลจีเจิงสถิติเป็นตัวช่วยในการตรวจสอบชื่อตีให้มีความน่าเชื่อถือมากยิ่งขึ้น โดยหลักการอนโนทัลจีเจิงสถิตินั้นเป็นตัวกำหนดขอบเขตการแบ่งกลุ่มค่าความดีทางสถิติ (กนกกาญจน์และจักรกฤษณ์, 2006) และเป็นตัวช่วยในการวิเคราะห์และตรวจสอบชื่อตีให้มีความน่าเชื่อถือมากยิ่งขึ้น โดยระบบปรับปรุงนี้ในขณะนี้ได้ถูกนำไปทดสอบระบบ โดยได้รับการชื่อชุมชนนิพัฒนาวิทยาลัยนเรศวร จำนวน 11,658 ชื่อ ค่าสำคัญ: ระบบการตั้งชื่อ อนโนทัลจีเจิงสถิติ กฎพื้นฐาน การจัดกลุ่ม

Abstract: Naming has been continuously developed from the past to the present and has evolved into a variety of patterns. In many cultures naming is not only important because every individual needs to have a name but also to have helpful names or names with a good sound. The basic goal of naming is to provide a good fortune and progress during life. Thai parents always try to choose names which they feel will bring good luck to their offsprings and to the family. Three methodologies of Thai naming use Thai astrology, a new numeric methodology, or traditional calendar. The principal naming using Thai astrology is widely used since former times as it involves the birthday in order to form the name. In this research we propose an improvement of a web-based Thai name checking system based on the value of a statistical model proposed by Kanokkarn and Chakkrit (Kanokkarn and Chakkrit, 2006a) (Chakkrit and Kanokkarn, 2007a). The new approach uses the concept of a statistical ontology to help in defining a group of statistical values and to improve the reliability of Thai name checking. Furthermore, we present the improvement of Thai name checking system which is tested with a database of 11,658 student names of Naresuan University.

Key words: naming system, statistical ontology, rule base, clustering

บทที่ 1 บทนำ (Introduction)

1.1 ความสำคัญและที่มาของปัญหาที่ทำการวิจัย

ในสมัยก่อนสามัญชนชาวไทยตั้งชื่อลูกกันง่ายๆ ด้วยพยางค์เดียว หรืออย่างมากก็ไม่เกินสองพยางค์ เช่น นายทอง นายเข้ม นางเจ้ม นางอิน นางนา โดยมากซึ่งอ่อนกว่าจะให้ได้ทั้งหนูน และชาย และมีชื่อกันอยู่จำนวนมาก ส่วนชื่อที่ยาวและໄเพเราะส่วนใหญ่แล้วมาจากภาษาบาลีและสันสกฤต ซึ่งมักจะใช้กันในหมู่เจ้าชุมชนนายหรือผู้ที่มีศรีบวรราศีก็ด้วย ความนิยมตั้งชื่อยาวๆ ໄเพเราะในหมู่คนทั่วไปเริ่มในสมัยรัชกาลที่ 5 และต่อมาชื่อคนไทยเริ่มยาวขึ้นเป็นสองพยางค์ ถือว่ามีความทันสมัยกว่าคนรุ่นบิดามารดา อย่างเช่น ชื่อของลูกๆ คือ มนติรา ณัฐสุดา มนติรุ ชุติมา ต่อมาได้มีการเปลี่ยนแปลงครั้งใหญ่ ปีพุทธศักราช 2482 โดยรัฐบาลกำหนดให้ผู้ชายและผู้หญิงมีชื่อเหมาะสมสมกับเพศของตน การตั้งชื่อมีความสำคัญมากสำหรับบุคคลทั่วไป และมีการพัฒนามาตั้งแต่สมัยอดีตถึงปัจจุบันอย่างต่อเนื่อง ซึ่งจะมีความแตกต่างกันออกไปขึ้นอยู่กับวัฒนธรรม หรือภูมิภาคที่ในการตั้งชื่อ ของแต่ละกลุ่มชนหรือประเทศนั้นๆ

สำหรับคนไทยการตั้งชื่อถือว่าเป็นสิ่งที่จำเป็นอย่างยิ่ง เพราะว่าคนไทยส่วนใหญ่มีความเชื่อว่าการมีชื่อที่ดีทำให้มีโชคชะตาและชีวิตที่ดียิ่งขึ้น ซึ่งส่วนใหญ่การตั้งชื่อที่ดีนั้นจะใช้ความหมายของชื่อเป็นหลัก ดังนั้นชื่อแรก (ที่ไม่ใช่ชื่อสกุล) ส่วนใหญ่จะเป็นชื่อที่มีความหมายและวิธีที่นิยมและใช้กันอย่างแพร่หลายก็คือ การตั้งชื่อโดยยึดหลักทักษะปกรณ์ โดยใช้หลักการตั้งชื่อตามวันเกิดและใช้ภูมิของหลักทักษะปกรณ์เป็นตัวกำหนดตัวอักษรในชื่อ ซึ่งในหลักทักษะปกรณ์นี้จะแบ่งกลุ่มของตัวอักษรออกเป็นภูมิต่างๆ และผู้ที่เกิดในวันแต่ละวันของสัปดาห์นั้นๆ ก็จะมีภูมิของอักษรที่ต่างกัน ส่วนใหญ่แล้วจะเลือกเอาเฉพาะอักษรที่ดี เป็นสิริมงคลตามวันเกิดนั้นๆ ซึ่งมีความเชื่อท่ว่าจะทำให้เป็นการเสริมดวงและโชคชะตาให้ดีขึ้น (Snae and Brueckner, 2006a)

การตั้งชื่ออาจทำได้โดยบุรุษ พร. นักโบราณค术 หรือผู้ที่มีความรู้และความเชี่ยวชาญในการตั้งชื่อเป็นผู้ตั้งชื่อให้ แต่ทั้งนี้จะต้องมีเวลาและเสียค่าใช้จ่ายโดยไม่จำเป็น อีกทั้งอาจจะไม่ได้ชื่อดีตามที่ตนเองต้องการ เนื่องจากมีตัวเลือกของชื่อเป็นจำนวนน้อยมาก ซึ่งชื่อเหล่านี้อาจจะเป็นชื่อซ้ำกับผู้อื่นได้ ในปัจจุบันนี้ระบบการตั้งชื่อที่ดีและตรงตามหลักทักษะปกรณ์นั้นสามารถค้นหาได้ตามเว็บไซด์ต่างๆ แต่ระบบส่วนใหญ่จะอยู่ในรูป สเตติคเว็บ(Static web) ซึ่งไม่มีความยืดหยุ่นที่เพียงพอที่จะทำให้ผู้ใช้สามารถนำมาเป็นทางเลือกที่ดีที่สุดในการตั้งชื่อ เช่น ไม่มีการตรวจสอบชื่อที่ตั้งมาว่าดีหรือไม่ยั่งไว

ดังนั้นผู้วิจัยจึงได้ศึกษาและได้นำหลักการและแนวคิดของวรรณกรรมรวมทั้งงานวิจัยที่เกี่ยวข้องกับระบบการตั้งชื่อแบบทักษะปกรณ์มาออกแบบและพัฒนาระบบการตรวจสอบชื่อตามแนวความคิดที่ได้จากการวิเคราะห์วรรณกรรมของระบบการตั้งชื่อ อาทิเช่น ระบบการตั้งชื่อไทยแนวคิดและกฎพื้นฐาน (Rule-based) ของระบบการตั้งชื่อ (Snae and Brueckner, 2006a) กฎพื้นฐานของการตั้งชื่อไทยโดยใช้เทคนิคไฮบริดเดนเมทชิ่ง (Hybrid name matching) (Snae and Brueckner, 2006b) การสร้างกฎพื้นฐานของการตั้งชื่อไทยโดยใช้หลักการ Name Ontology (Snae, 2006b) ระบบการตรวจสอบชื่อไทยโดยใช้หลักการทักษะปกรณ์ (กนกกาญจน์ และ จักรกฤษณ์, 2006) เป็นต้น ซึ่งผู้วิจัยจะได้นำระบบที่พัฒนาขึ้นมาใหม่นี้มาช่วยในการวิเคราะห์ว่า การตั้งชื่อในปัจจุบัน หรือย้อนไปในอดีตที่ผ่านมาประมาณ 20 ปีมันได้ถูกต้องตามหลักทักษะปกรณ์หรือไม่ ซึ่งที่ตั้งนานี้ดีหรือไม่ดีอย่างไร

1.2 วัตถุประสงค์ของโครงการ

1.2.1 เพื่อศึกษาและวิเคราะห์ระบบการตั้งชื่อแบบทักษะปกรณ์ และระบบการตรวจสอบชื่อ

ในงานวรรณกรรม

1.2.2 เพื่อนำแนวคิดและ หลักการ Ontology มาปรับปรุงและพัฒนาระบบการตรวจสอบชื่อด้วยการตั้งชื่อแบบทักษะปกรณ์ ให้มีความนำร่องมากยิ่งขึ้น

1.3 ครอบแนวคิดหรือทฤษฎี และงานวิจัยที่เกี่ยวข้อง

1.3.1 Rule Base Expert System (Giarratano, 1998) ใช้ในการสร้างกฎการตรวจสอบ

1.3.2 Web Base Technology (Snae and Diaz, 2002) เพื่อสร้างระบบโดย บริการผ่าน Web Site โดยใช้ภาษา PHP (เจริญโภคานนท์, 2548) ในการเด้งข้อมูลมาแสดงบน อินเตอร์เน็ต ได้อย่างมีประสิทธิภาพ

1.3.3 Data Base System (Silberschatz, et. all, 2002) เพื่อสร้างและพัฒนาระบบฐานข้อมูลซึ่งของนิสิตมหาวิทยาลัยนเรศวรเพื่อใช้ทดสอบกับระบบที่พัฒนา

1.3.4 Statistical model (กนกกาญจน์ มะหุต และ จักรกฤษณ์ เสน่ห์ 2006) เพื่อคำนวณค่าความดีของชื่อว่าชื่อที่ผู้ใช้สามารถเป็นชื่อดีกี่เปอร์เซ็นต์ในแต่ละภูมิทักษะ และค่าผลกระทบของความนำร่องของชื่อดี

1.3.5 Statistical ontology (Snae and Diaz, 2002) เพื่อนำค่าความดีของชื่อจากชื่อ

1.5.4 มา กำหนดขอบเขตทางสถิติ โดยแบ่งตามช่วงของค่าความดี และอักษรกลกนี

1.3.7 Software Engineering (Sommerville, 2001) ใช้ในการออกแบบระบบและ ตรวจสอบระบบ ทุกขั้นตอนอย่างเป็น ระบบและถูกต้องตาม ทฤษฎี

1.3.8 หลักการตั้งชื่อแบบทักษะปกรณ์ ใช้ในการสร้างกฎการตั้งชื่อและกฎการตรวจสอบชื่อดี

1.3.9 งานวิจัยที่เกี่ยวข้อง

คำว่า Ontology ได้กล้ายมาเป็นที่รู้จักและใช้กันอย่างแพร่หลายในปัจจุบัน และหลายๆ ปีที่ผ่านมา เช่น ในสาขาวิชาปัญญาประดิษฐ์ (Artificial Intelligence: AI) สาขาวิชาคอมพิวเตอร์ (computer) และสาขาวิชาพัฒนาสารสนเทศ (information science) โดยเฉพาะอย่างยิ่ง ในขอบเขตสาขาวิชาทางด้านระบบสารสนเทศ (information system) การรวมสารสนเทศทางปัญญา (intelligence information integration) การสืบค้นข้อมูล (information retrieval and extraction) การแทนที่ความรู้ (knowledge representation) การออกแบบฐานข้อมูล (database design) การจัดการพื้นฐานองค์ความรู้ (knowledge base management) และระบบการจัดการฐานข้อมูล (database management system) (Guarino, 1998, Andrade and Saltz, 2000) ในสาขาทั้งหมดที่กล่าวมานี้ ความหมายของ Ontology ก็ยังแตกต่างกันอีกด้วย และไม่มีนิยามเฉพาะตัว ดังนั้นในหัวข้อนี้เราได้นำเสนอความหมายของ Ontology ที่แตกต่างกันและคำว่า Ontology ที่ได้ถูกกล่าวถึงในวรรณกรรมต่างๆ เช่น ในสาขาวิชาปัญญา สาขาวิชาปัญญาประดิษฐ์ เป็นต้น

Gruber (1993) ได้นิยามความหมายของ Ontology ซึ่งเป็นที่รู้จักกันอย่างกว้างขวางมากที่สุด ก็คือ Ontology เป็นรายละเอียดที่ชัดเจนແเนื่องจากความคิด ("An ontology is an explicit specification of a conceptualization") ดังนั้นเมื่อเราต้องการแสดงหรือระบุความคิดของวัตถุหรือสิ่งของนั้นๆ เราควรจะแสดงรายละเอียดที่ชัดเจนและแน่นอน เช่น การอธิบาย Ontology ของโปรแกรม โดยการนิยามกลุ่มของคำที่แสดงและความสัมพันธ์ระหว่างกลุ่มคำนั้นๆ การนิยามนี้อาจจะเกี่ยวข้องกับชื่อของสิ่งที่มีอยู่ในโปรแกรม ซึ่งอาจประกอบไปด้วย classes, relations, functions หรือ objects อื่นๆ

Stoffel et al. (1997) ยังได้นำเสนอแนวความคิดของการสร้างดัชนีที่เกี่ยวกับความหมายของคำ (semantic indexing) ในสร้าง Ontology ในรูปแบบดัชนีที่สามารถให้ระบบสืบค้นหาข้อมูลที่จัดกลุ่มไว้แล้วด้วยหลักการ Ontology ซึ่งโดยความสำคัญนี้ จะทำให้การสืบค้นเป็นไปอย่างมีประสิทธิภาพซึ่งจะได้ผลลัพธ์ของข้อมูลที่สามารถอธิบายความหมายที่เกี่ยวข้องของข้อมูลที่สืบค้นได้อย่างชัดเจน ตัวอย่างเช่น Miller (1990) ได้จัดกลุ่มรากและประวัติของคำศัพท์ โดยใช้ Ontology มาจัดระบบโครงสร้างของภาษา ซึ่งเรียกว่า WordNet โดยจะทำการนิยามแนวความคิดให้เป็นหมวดหมู่ (Cluster) ของคำซึ่งจะเรียกกลุ่มคำว่า synsets (เป็นกลุ่มของคำศัพท์ที่มีความหมายคล้ายคลึงกันหรือเหมือนกัน)

วิธีการสืบค้นข้อมูลอีกวิธีการหนึ่งได้แก่ การทำเหมืองข้อมูล (Data mining) วิธีการทำเหมืองข้อมูลส่วนใหญ่นี้สร้างกฎพื้นฐาน (Rule-based) อย่างเดียวบนสารบัญของฐานข้อมูล แต่อย่างไรก็ตาม การใช้ความรู้ที่มีอยู่ในการรวมสามารถสนับสนุนกระบวนการสืบค้น และสร้างกฎด้วยความหมายของคำในรูปแบบ Ontology ได้อย่างมีประสิทธิภาพ

หลักการอนโนทेशัน สามารถนำมาประยุกต์ใช้กับหลักสถิติได้เป็นอย่างดี โดยจะถูกเรียกว่า อนโนทेशันสถิติ (Statistical ontology) ซึ่งส่วนใหญ่จะถูกนำไปใช้ในการวิเคราะห์ข้อมูล ทำเหมืองข้อมูล หรือใช้ในการจัดกลุ่มของข้อมูลที่มีการสัมพันธ์กันโดยผลลัพธ์จะถูกแสดงในรูปแบบสถิติ (Denk et al, 2002 ; Hert and Hass, 2003)

Marchinonini et al, (2003) ได้พัฒนา Statistical ontology ในการหาความสัมพันธ์ของหลักสถิติ และเพื่อมโยงหลักการที่เกี่ยวข้องกันเข้าด้วยกัน โดยเริ่มจากการสร้างนิยามศัพท์ และพัฒนาการอธินายเชิงกราฟ ซึ่งอนโนทेशันจะถูกนำไปใช้ในการสนับสนุนการสร้างการอธินาย โดยจะอนุญาตให้ผู้ใช้สืบค้นหลักการสถิติต่างๆ และความสัมพันธ์ที่เกี่ยวโยงหรือเกี่ยวข้องซึ่งกันและกัน

จักรกฤษณ์ (Snae, 2006b) ได้ใช้หลักการอนโนทेशัน ontology (Guarino, 1998 and Gruber, 1998) ช่วยในการตั้งชื่อ โดยการแบ่งกลุ่ม (clustering) และสร้างอินเด็กซ์ (indexing) (Stoffel et al., 1997) ของชื่อที่สามารถระบุความหมายว่าเป็นเพศชาย หรือเพศหญิงได้ โดยใช้กลุ่มคำจากความหมายของชื่อเป็นหลักในการระบุเพศ ตัวอย่างเช่น เราสามารถระบุเพศชายได้จาก กลุ่มคำที่มีความหมายว่า "แข็งแรง" "แข็งแกร่ง" "กล้าหาญ" เช่นชื่อ วชิรพล วรพล พลาธิป เกรียงไกร วีรวัฒน์ อิรศานต์ ระบุเพศหญิงได้จาก กลุ่มคำที่มีความหมายว่า "สวยงาม" "น่ารัก" เช่นชื่อ กันตพงษ์ ปิยาพร บริยาภรณ์ พิมพ์วิไล พิจิตร จิตรวินด์ วิมพ์วิภา

ในปัจจุบันระบบการตั้งชื่อด้วยใช้หลักทักษะการณ์นี้ได้ถูกสร้างขึ้นในรูปของเว็บเบล (Web Base) และพบเห็นได้ทั่วไปบนอินเตอร์เน็ต ส่วนใหญ่จะอยู่ในรูป Static Web โดยจะสร้างอินเด็กซ์ (index) ของชื่อตามวันเกิด รวมทั้งความหมายของชื่อแล้วจัดเก็บลงในระบบฐานข้อมูล ข้อเสียของระบบเหล่านี้คือ ระบบไม่มีการตรวจสอบชื่อแบบ statistic model เช่น ตรวจสอบว่าชื่อนี้ดีกับปีร์เซ่นต์ หรือตรวจสอบปีร์เซ่นต์วันเกิดของชื่อ จักรกฤษณ์ (Snae and Brueckner, 2006a, 2006b) ได้สร้างและพัฒนาระบบการตั้งชื่อไทยตามหลักทักษะการณ์แบบไดนามิกเก็บ (dynamic web) โดยใช้รูปแบบ statistic model, rule-based, name matching algorithm และหลักการ ontology ช่วยในการพัฒนาระบบการตั้งชื่อ ซึ่งมีคุณสมบัติดังต่อไปนี้

- ตรวจสอบเพศจากความหมายของชื่อ (Snae, 2006b)
- ตั้งชื่อถูกตามชื่อพ่อแม่ เช่นชื่อ ฤชาดา มาจากพ่อชื่อ สุชาติ และแม่ชื่อ วิภาดา (Snae and Brueckner, 2006b) โดยใช้ LIG3 (Levenshtein, Index of Similarity Group (called ISG), and Guth) (Snae and Diaz, 2002)

1.4 ขอบเขตการวิจัย

การวิจัยครั้งนี้เราได้ออกแบบและพัฒนาระบบต่อจากการตรวจสอบชื่อตามหลักการตั้งชื่อแบบทักษะปกรณ์ โดย กนกกาญจน์และจักรกฤษณ์ (2006) และการตรวจสอบชื่อด้วย ให้หลักการ Clustering (จักรกฤษณ์ และ กนกกาญจน์, Naresuan Science Journal 2007) ซึ่งในงานวิจัยครั้งนี้ ráได้นำสถาปัตยกรรมการคำนวณหาค่าความดีของชื่อใน (กนกกาญจน์และจักรกฤษณ์ 2006) และแนวคิดทางด้าน Ontology (Gruber, 1993; Guarino, 1998) มาประยุกต์ใช้กับหลักสถิติ (Statistic) เพื่อได้ระบบตรวจสอบชื่อที่มีประสิทธิภาพและมีความนำเอามากยิ่งขึ้น

1.4.1 ออกแบบและพัฒนาระบบฐานข้อมูลในการเก็บข้อมูลชื่อของนิสิตมหาวิทยาลัยนเรศวรที่จะนำมาใช้ในส่วนของการทดสอบระบบ

1.4.2 พัฒนาโปรแกรมคอมพิวเตอร์ด้วยระบบ Expert System โดยใช้ Rule Base ในการสร้างกฎการหลักการตั้งชื่อแบบทักษะปกรณ์ ซึ่งเป็นการตั้งชื่อตามวันเกิดโดยใช้ภูมิหลักของทักษะปกรณ์ (ดังตารางที่ 1) “ได้แก่ เดช อายุ ศรี มูละ อุดสาหะ มโนธิ บริหาร และ กาลกิจ” ซึ่งไม่แต่ละภูมิจะมีตัวอักษรประจำภูมิที่แตกต่างกันที่ใช้เป็นหลักในการตั้งชื่อและจะถูกออกแบบในระบบโดยใช้หลัก rule base ส่วนใหญ่จะใช้ภูมิเด่นนำหน้าเพศชายและภูมิศรีนำหน้าเพศหญิง แต่มักจะหลีกเลี่ยงการใช้ตัวอักษรที่อยู่ในกาลกิจ เพราะคนไทยเชื่อว่าเป็นสิ่งที่ไม่ดีและอับมงคล ถ้าใช้หลักการนี้ในการตั้งชื่อ เราจะได้ชื่อที่มีความหมายดีเป็นสิริมงคลและถูกต้องตามหลักทักษะปกรณ์

ตารางที่ 1 แสดงชื่อวันในแต่ละสัปดาห์ซึ่งไม่แต่ละวันจะมีตัวอักษรประจำภูมิ(8 ภูมิทักษะ) จะเห็นว่ามีวันพุธ 2 วัน เพราะว่าการนับวันของไทยแตกต่างจากนับแบบสามกอล โดย การนับแบบไทยจะเริ่มนับวันใหม่ตอน 06.00 นาฬิกา (6 โมงเช้า) และส่วนวันพุธก็จะแบ่งได้ 2 ช่วงคือวันพุธกลางวัน(06.00-17.59) และวันพุธกลางคืน(18.00-05.59) เราชาระบบน้ำอาอักษรประจำวันในแนวนอนมาตั้งชื่อตามกฎเกณฑ์ตามหลักทักษะได้ ยกเว้นอักษรในคอลัมน์สุดท้ายในตารางไม่ควรจะนำมาใช้ในการตั้งชื่อ เพราะเป็นอักษรกาลกิจ

ตารางที่ 1 แสดงตัวอักษรตามวันเกิด

วันเกิด	นิริเวช	อาทิตย์	เดือน	ครึ่ง	บุคคล	ฤกษ์สุขุมวัฒน์	นาครี	ราศี
อาทิตย์	ธ. อริชี ดูโอเอ โอ	ท.น.ก จ.ส.	ธ.ล.ช ช.ก.ญ	อ.ภ.ช ก.ก.ณ	ต.ต.ก ก.ช.น	บ.ป.ก. พ.ท.ก.น	บ.ร.ก.ว	ก.ส.น ก.พ.ช
จันทร์	ก.น.ก จ.ส.	อ.ล.ช ช.ก.ญ	อ.ภ.ช ก.ก.ณ	ต.ต.ก ก.ช.น	บ.ป.ก. พ.ท.ก.น	บ.ร.ก.ว	ก.ส.น ก.พ.ช	อ.อ.อิช ดูโอเอ โอ
พุธ	อ.ล.ช ช.ก.ญ	อ.ภ.ช ก.ก.ณ	ต.ต.ก ก.ช.น	บ.ป.ก. พ.ท.ก.น	บ.ร.ก.ว	ก.ส.น ก.พ.ช	อ.อ.อิช ดูโอเอ โอ	ก.น.ก จ.ส.
พฤหัสบดี	บ.ป.ก. พ.ท.ก.น	บ.ร.ก.ว	ก.ก.น ก.พ.ช	อ. อ.อ. อิช ดูโอเอ โอ	ก.น.ก จ.ส.	บ.ป.ก. พ.ท.ก.น	บ.ร.ก.ว	อ.ล.ช ช.ก.ญ
ศุกร์	ก.ส.น ก.พ.ช	อ.อ.อ. อิช ดูโอเอ โอ	ก.ก.น.ก.ช.	อ.ภ.ช ช.ก.ญ	ต.ต.ก ก.ช.น	บ.ป.ก. พ.ท.ก.น	บ.ร.ก.ว	อ.ภ.ช ช.ก.ญ
เสาร์	ต.ต.ก ก.ช.น	บ.ป.ก. พ.ท.ก.น	บ.ร.ก.ว	ก.ส.น ก.พ.ช	อ.อ.อ. อิช ดูโอเอ โอ	ก.ก.ก.ก.ช.	อ.ภ.ช ช.ก.ญ	อ.ภ.ช ช.ก.ญ
อาทิตย์	บ.ร.ก.ว	ก.ส.น ก.พ.ช	อ.อ.อ. อิช ดูโอเอ โอ	ก.ก.ก.ก.ช.	อ.ภ.ช ช.ก.ญ	ต.ต.ก ก.ช.น	บ.ป.ก. พ.ท.ก.น	บ.ร.ก.ว

1.4.3 ปรับปรุงและพัฒนาระบบการตรวจสอบชื่อเดียวกันลักษณะ Statistical Ontology ในรูปแบบ Web base Technology โดยใช้ภาษา PHP และการใช้เทคโนโลยี Database System ในการเก็บฐานข้อมูลชื่อของนิสิตมหาวิทยาลัยนเรศวร เพื่อให้ทดสอบกับระบบที่พัฒนาและดึงข้อมูลมาแสดงบนอินเตอร์เน็ตได้อย่างมีประสิทธิภาพ

ระบบสามารถตรวจสอบชื่อเดียวกันลักษณะ Statistical Ontology โดย Ontology ในที่นี้ทำหน้าที่ ในการกำหนดขอบเขตทางสถิติ โดยแบ่งตามช่วงของค่าความดี และอักษรกลกนี ได้ดังต่อไปนี้

1.4.3.1 ค่าทางสถิติโดยแบ่งตามช่วงของค่าความดี

ชื่อที่มีค่าความดีมากกว่า 79% ขึ้นไป จะถูกกำหนดให้มีค่าทางสถิติเท่ากับ 2

ชื่อที่มีค่าความดีมากกว่า 49 % แต่น้อยกว่า 79 % จะถูกกำหนดให้มีค่าทาง

สถิติเท่ากับ 1.5

ชื่อที่มีค่าความดีมากกว่า 20 % แต่น้อยกว่า 49 % จะถูกกำหนดให้มีค่าทาง

สถิติ เท่ากับ 1

ชื่อที่มีค่าความดีมากกว่า 0 % แต่น้อยกว่า 20 % จะถูกกำหนดให้มีค่าทาง

สถิติเท่ากับ 0.5

1.4.3.2 ค่าทางสถิติโดยแบ่งจากอักษรกลกนี

ชื่อที่ไม่มีอักษรกลกนี จะถูกกำหนดให้มีค่าทางสถิติเท่ากับ 1

ชื่อที่มีอักษรกลกนี จะถูกกำหนดให้มีค่าทางสถิติเท่ากับ -1

ผลของการตรวจสอบจะออกแบบในรูปของการคำนวนเชิงสถิติ โดยใช้สูตรผลรวมของค่าทางสถิติหารด้วย ผลรวมของค่าสูงสุดจากการวัดค่าทางสถิติ (จากข้อ 1.4.3.1 และ 1.4.3.2 ซึ่งก็คือ 3 นั่นเอง) เช่น ชื่อ “กนกกาญจน์” มีค่าความดีเท่ากับ 89% (คิดจากสูตร (1) หัวข้อ 1.5) และไม่มีตัวกลกนี ดังนั้น ค่าการคำนวนเชิงสถิติจะมีค่าเท่ากับ $(2+1)/3$ เท่ากับ 100% ซึ่งหมายความว่า ชื่อนี้เป็นชื่อที่ได้เติมร้อยหรือสมบูรณ์แบบตามหลักทักษะปกรณ์

1.5 ขั้นตอนในการดำเนินการ

1.5.1 วิเคราะห์และศึกษางานวิจัยที่เกี่ยวข้องกับระบบตั้งข้อแบบทักษะปกรณ์ ศึกษา Rule Base Expert System (Giarratano, 1998) ใช้ในการสร้างกฎการตั้งข้อตามวันเกิดโดยใช้ภูมิหลักของทักษะปกรณ์ และกฎตรวจสอบซึ่งก็คือ ศึกษาหลักการ Ontology (Gruber, 1993; Guarino, 1998) ว่าสามารถนำมาประยุกต์ใช้กับหลักสถิติ (Statistic) เพื่อให้ระบบตรวจสอบซึ่งที่มีประสิทธิภาพ และมีความน่าเชื่อถือมากยิ่งขึ้น ได้อย่างไร

1.5.2 วิเคราะห์หาความต้องการของระบบและออกแบบรูปแบบการทำงานของระบบ โดยรวมออกแบบระบบฐานข้อมูลในการเก็บข้อมูลข้อของนิสิตมหาวิทยาลัยราชภัฏที่จะนำมาใช้ในส่วนของการทดสอบระบบ ออกแบบกฎการกฎตรวจสอบซึ่งก็คือโดยใช้ Statistical Ontology และออกแบบหน้าจอของระบบ

1.5.3 นำความรู้ที่ศึกษาจากข้อ 1.10.1 การออกแบบจากข้อ 1.10.2 และ ส่วนของการตั้งข้อแบบทักษะปกรณ์และระบบตรวจสอบซึ่งก็คือโดยใช้ statistical model ของ กนกกาญจน์ มะหุต และ จักรกฤษณ์ เสน่ห์ (2006) มาพัฒนาระบบ และเขียนโปรแกรม โดยแบ่งเป็น 2 ส่วน คือ ระบบตรวจสอบซึ่งก็คือโดยใช้ Statistical Ontology และ ส่วนของการทดสอบระบบ

1.5.4 ทดสอบ แก้ไข โดยนำระบบไปให้กลุ่มตัวอย่างทดสอบ แล้ววัดผลของระบบบว่าให้ผลตามที่ต้องการหรือใหม่หรือมีส่วนใดที่ผิดพลาดหรือไม่ และสามารถเพิ่มส่วนที่ทำให้ระบบสมบูรณ์มากยิ่งขึ้น

1.5.5 ทดสอบและปรับปรุงระบบตามที่ได้รับข้อผิดพลาดหรือส่วนเพิ่มเติมจากข้อ 1.10.4 เป็นครั้งสุดท้าย

1.5.6 สรุปผล ประเมินโครงการ

1.6 ประโยชน์ที่คาดว่าจะได้รับ

1.6.1 ระบบตรวจสอบชื่อตี ที่มีประสิทธิภาพ และ ถูกต้องตามหลักการตั้งชื่อ

แบบทักษะปกรณ์

1.6.2 เป็นประโยชน์ต่อผู้ที่ต้องการตรวจสอบชื่อของตนเองว่าถูกต้องตามหลัก
ทักษะปกรณ์หรือไม่ ซึ่งที่ตั้งมานั้นดีหรือไม่อย่างไร



บทที่ 2

วิธีดำเนินงานวิจัย

1. อุปกรณ์ที่ใช้ในการดำเนินงานวิจัย

อุปกรณ์ที่ใช้ในการดำเนินงานวิจัย ประกอบด้วยเครื่องคอมพิวเตอร์จำนวน 1 เครื่อง ส่วนประกอบที่สำคัญ คือ

1.1 อุปกรณ์ด้านฮาร์ดแวร์ ได้แก่

1.1.1 เครื่องคอมพิวเตอร์ (Laptop) - Intel® Core™ 2 Duo Processor

1.66 GHz 512 MB of RAM

1.1.2 ฮาร์ดดิสก์ขนาด 80 GB

1.2 อุปกรณ์ด้านซอฟต์แวร์หรือโปรแกรม (Program) ประกอบด้วยโปรแกรม ดังนี้

1.2.1 โปรแกรม Windows XP เป็นระบบปฏิบัติการ

1.2.2 โปรแกรม Editplus v2.20 ใช้สร้างอัลกอริธึมทั้งหมด

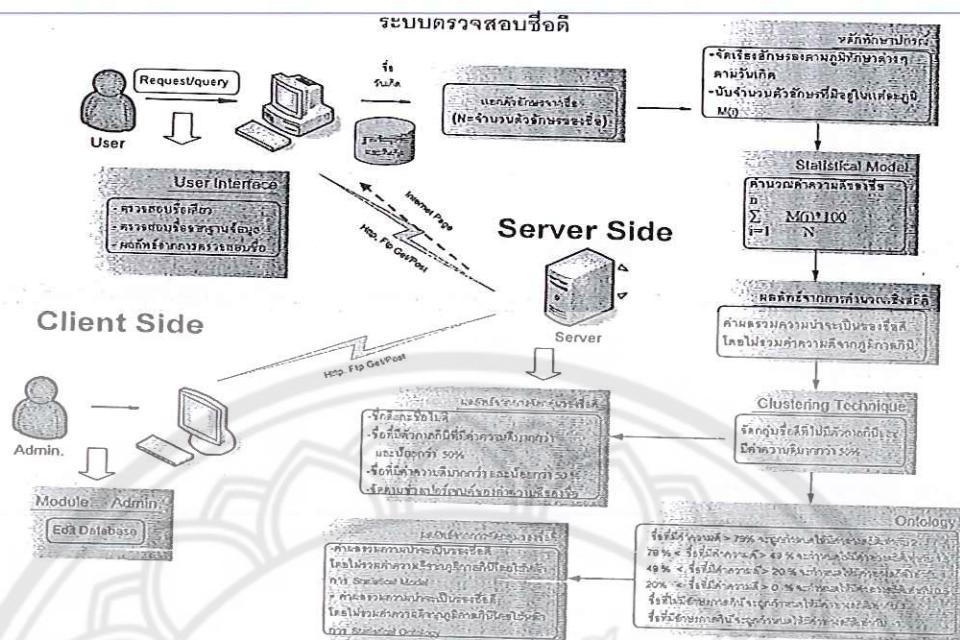
1.2.3 โปรแกรม MySQL-Front ใช้เก็บข้อมูลเชื่อมความหมาย

1.2.4 โปรแกรม appserv-win32-2.5.1 ใช้เป็น server จำลองในการรัน

1.2.5 โปรแกรม Internet Explorer ใช้เป็น Web Brower

2 การออกแบบและพัฒนาระบบ

สถาปัตยกรรมของระบบตรวจสอบชื่อใหม่นี้ ขึ้นอยู่จากการออกแบบของกนกภานุนันต์และจักรกฤษณ์ (กนกภานุนันต์และจักรกฤษณ์, 2006a) แต่ผู้วิจัยได้เพิ่มหลักการอนโนทัลจีเซิงส์ติดเข้าไปในระบบ ด้วย ซึ่งระบบสถาปัตยกรรมจะใช้หลักการและแนวคิดทางเทคโนโลยีทั้งต่อไปนี้ (ดูรูปที่ 1)



รูปที่ 1 สถาปัตยกรรมของระบบตรวจสอบชื่อคิ้ว

2.1 หลักทักษะปกรณ์

เป็นการตั้งชื่อตามวันเกิดโดยใช้ภูมิหลักของทักษะปกรณ์ (ดังตารางที่ 1) ได้แก่ เดช อายุ ศรี มูล อุดสาหะ มนตรี บริวาร และ กาลกิน ซึ่งในแต่ละภูมิจะมีตัวอักษรประจำภูมิที่แตกต่างกัน ที่ใช้เป็นหลักในการตั้งชื่อและจะถูกออกแบบในระบบโดยใช้หลัก rule base ส่วนใหญ่จะใช้ภูมิเดช นำหน้าเพศชายและภูมิศรีนำหน้าเพศหญิง แต่มักจะหลีกเลี่ยงการใช้ตัวอักษรที่อยู่ในกาลกิน เพราะคนไทยเชื่อว่าเป็นสิ่งที่ไม่ดีและอัปมงคล ถ้าใช้หลักการนี้ในการตั้งชื่อ เราอาจจะได้ชื่อที่มีความหมายดีเป็นสิริมงคลและถูกต้องตามหลักทักษะปกรณ์ ซึ่งความหมายของภูมิทักษะสามารถอธิบายได้ดังต่อไปนี้

ภูมิเดช หมายถึง อำนาจความสามารถ เกียรติคุณ ได้รับการยก ย่องสรรเสริญ

ภูมิอายุ หมายถึง มีอายุยืนร่างกายแข็งแรง

ภูมิศรี หมายถึง เสน่ห์ โชคความ喜บีถึงความสำเร็จ

ภูมิมูล หมายถึง ทรัพย์สินหลักฐานบ้านช่อง

ภูมิอุดสาหะ หมายถึง ความยั่นหยันมั่นเพียร

ภูมิมนตรี หมายถึง ตำแหน่งหน้าที่การงานดี มีผู้อุปถัมภ์ให้ความก้าวหน้า

ภูมิบริวารหมายถึง บุตรภรรยาสามี เพื่อนและลูกน้องที่จะช่วยเกื้อหนุน

ส่วนภูมิกาลกิน เป็นตัวอัปมงคล ซึ่งควรหลีกเลี่ยงไม่เอามาไว้ในชื่อ

ตารางที่ 1 แสดงชื่อวันในแต่ละสัปดาห์ซึ่งในแต่ละวันจะมีตัวอักษรประจำภูมิ (8 ภูมิ) จะเห็นว่ามีวันพุธ 2 วัน เพราะว่าการนับวันของไทยแตกต่างจากนับแบบตากล โดย การนับแบบไทยจะเริ่มนับวันใหม่ตอน 06.00 นาฬิกา (6 โมงเช้า) และส่วนวันพุธจะแบ่งได้ 2 ช่วงคือ วันพุธกลางวัน (06.00-17.59) และวันพุธกลางคืน (18.00-05.59) เรายสามารถนำเอาอักษรประจำวันไปแนบกับมาตั้งที่ตามกฎเกณฑ์ตามหลักทักษาราดี ยกเว้นอักษรในคอลัมน์สุดท้ายในตารางไม่ควรจะนำมาใช้ในการตั้งที่ เพราะเป็นอักษรากาศกินี ตัวอย่างเช่น

ตารางที่ 1 แสดงตัวอักษรตามวันเกิด

วันเกิด	บวชวาร	อาทิตย์	เดช	พีรี	บุตร	อุคลาภ	นาครี	ภาคภูมิ
อะกิลล์	อะ อารี อุ ฐะ เอ โอ	ก ข ก ฯ ฯ	ช ล ช ช ก ญ	ก ช ช ช ช ช ช ช	ก ด ด ก ช ช ช	บ ป บ บ ช ช ช	บ ร บ บ ร บ	ก ศ บ ก ช ช ช
ธีราทร	ก ข ก ฯ ฯ	จ ล ช ช ก ญ	ก ช ช ช ช ช ช ช	ก ด ด ก ช ช ช	บ บ บ บ ช ช ช	บ ร บ บ ร บ	ก ศ บ ก ช ช ช	ธ อา อี ธ ฐ ฐ เอ โอ
อัจฉรา	อ ล ช ช ก ญ	ก ช ช ช ช ช ช ช	ก ด ด ก ช ช ช	บ ป บ บ ช ช ช	บ ร บ	ก ศ บ ก ช ช ช	บ อา อี บ ฐ ฐ เอ โอ	ก ช ค ก ช ช ช
พุช กิตาภรณ์	พ ช ช ช ก ช ช ช	ก ด ด ก ช ช ช	บ บ บ บ ช ช ช	บ ร บ บ ร บ	ก ศ บ ก ช ช ช	บ อา อี บ ฐ ฐ เอ โอ	ก ข ก ก ช ช ช	อ ล ช อ ล ช ช
ราธยาสันต์	บ ป บ พ ช ช ช	บ ร บ บ ร บ	ก ศ บ ก ช ช ช	บ อา อี บ ฐ ฐ เอ โอ	ก ข ก ก ช ช ช	อ ล ช อ ล ช ช	บ ร บ บ ร บ	ด ต ด ก ช ช ช
อุ๊ก	ก ศ บ ก ช ช ช	อ ล ช อ ล ช ช	ก ข ก ก ช ช ช	ก ช ช ช ก ช ช ช	ก ด ด ก ช ช ช	บ ป บ บ ช ช ช	บ ร บ บ ร บ	ก ช ค ก ช ช ช
สุมาร์	ด ต ด ก ช ช ช	บ ป บ พ ช ช ช	บ ร บ บ ร บ	ก ศ บ ก ช ช ช	บ อา อี บ ฐ ฐ เอ โอ	ก ข ก ก ช ช ช	อ ช ช ช อ ช ช ช	อ ล ช อ ล ช ช
พุช กิตาภินี	บ ร บ บ ร บ	ก ศ บ ก ช ช ช	บ อา อี บ ฐ ฐ เอ โอ	ก ข ก ก ช ช ช	อ ล ช อ ล ช ช	ก ด ด ก ช ช ช	บ ป บ บ ช ช ช	บ ป บ บ ช ช ช

ที่มา กนกกาญจน์และจักรกฤษณ์ (กนกกาญจน์และจักรกฤษณ์, 2006a)

ถ้าเกิดวันอาทิตย์ เป็นชัยนิยมให้ ภูมิเดช นำหน้า ก็ควรจะขึ้นต้นที่อักษรตัว “ก” ณ ญ ในกรณี ตั้งที่ชื่อ คือให้อักษรเหล่านี้มีในชื่อตัวได้ตัวหนึ่ง ไม่จำเป็นต้องนำมาทั้งหมด เพื่อ เสริมดวงทาง อาชญากรรม ภารมี เกียรติยศ และชื่อเสียงทำแท่นหน้าที่การงาน แต่ถ้าเป็นหญิงก็นิยม เอาตัวอักษรในภูมิศรี นำหน้า คือ ภ ู ม ิ ศ ร ี ณ (อาจะหาชื่อยาก) นี้ประกอบในชื่อเพื่อเสริมโชคทาง ตัวอักษรในภูมิศรี นำหน้าก็ไม่เป็นไร ต่อไปก็เลือกเอาอักษรในภูมิบุตร คือ ต ด ท ช น เพื่อ เสริมดวงทางทรัพย์สิน ภารมี หลักฐานความเป็นอยู่ หรือ อักษร ย ร ด ว จากภูมิมนต์ในกรณี เสริมดวงทางความเป็นใหญ่ เพื่อความสำเร็จหรือการได้รับความอุปถัมภ์ค้ำชู ถ้าอย่างเช่นเสริมดวง ทางการดำเนินชีวิต เช่น เพื่อให้มีอายุยืนยาว หรือมีอายุมั่นคงยืนยัน ให้ใช้ตัวอักษรในภูมิ อายุ คือ ก ช ช ช ง หรือจะเพิ่มตัวอักษรในภูมิอุดสาeras คือ บ บ พ ก น ด้วยก็ได้ แต่ไม่ควรใช้อักษรใน

ภูมิภาคภินีคือ ศ. ษ.ส. พ.ช. ใน การตั้งข้อ กล่าวคืออย่าให้อักษรเหล่านี้มีอยู่ในข้อเพื่อหลีกเลี่ยง ความไม่เป็นมงคล ความอับโชคและอุปสรรคต่างๆ

บางคนตั้งข้อโดยใช้ตัวอักษรให้ครบ 7 ภูมิ ผลลัพธ์ของข้อที่ได้อาจจะไม่มีความหมาย สะกด ยก หรือบางครั้งไม่พบในพจนานุกรม นอกจากนี้ บางคนหลีกเลี่ยงตัวภูมิภินี ก็เลยต้องสะกด หรือเปลี่ยนข้อใหม่ อย่างเช่น พ.ช.นา พนิภา ศรัณญา ซึ่งทำให้เป็นปัญหาอย่างมากในการใช้หรือ อ้างอิงข้อทั้งนี้ทั้งนั้นเราอนุโลม การใช้ตัวการันต์คู่กับตัวอักษรที่เป็นภูมิภินีได้ เช่น กนกการ์ด (เกิดวันพุธ) ถือว่าเป็นข้อดี

2.2 หลักการของโมเดลทางสถิติ และค่าความน่าจะเป็น (Statistical model and Probability)

ในการคำนวณหาค่าความดีทางสถิติ ระบบจะใช้รูปแบบเชิงสถิติและความน่าจะเป็นในการ ตรวจสอบข้อที่ได้ตั้งมาแล้วว่ามีตัวอักษรในข้อตรงตามภูมิหลักทักษะหรือไม่ โดยผู้ใช้ใส่ข้อ และวัน เกิด ระบบจะนำข้อมาทำการแยกตัวอักษร และนับจำนวนตัวอักษรของข้อ (N) นำวันเกิดไปเทียบ กับภูมิทักษะต่างๆโดยใช้หลักทักษะปกรณ์ (ตารางที่ 1) และจัดเรียงตัวอักษรลงไปเรื่องตามภูมิทักษะ ตามวันเกิด จากนั้นนับจำนวนตัวอักษรที่มีอยู่ในแต่ละภูมิ ($M(i)$) หลังจากนั้นจะทำการคำนวณค่า ความดีของข้อว่าข้อที่ผู้ใช้ใส่มาบันทึกเป็นข้อดีกี่เปอร์เซ็นต์ในแต่ละภูมิทักษะ จากนั้นค่าความดีในแต่ ละภูมิจะถูกนำมารวมกันเป็นค่าความดีเชิงสถิติของข้อ (G) ซึ่งไม่รวมภูมิภิกนี โดยใช้สูตร คำนวณดังต่อไปนี้

$$G = \sum_{i=1}^n \frac{M(i) * 100}{N} \quad \dots \dots \dots (1)$$

ตัวอย่าง การใช้สถาปัตยกรรมของระบบตรวจสอบข้อดี โดยใช้สมการที่ 1 ข้างต้นนั้นมาทดสอบกับ จักรภูมิ ภิกนี เกิดวันจันทร์ สามารถอธิบายได้ดังนี้

หลังจากระบบแยกตัวอักษรแล้ว ระบบจะจัดเรียงตัวอักษรลงตามภูมิทักษะปกรณ์ ตามวัน จันทร์ โดยจากตารางที่ 1 จะเห็นได้ว่า

๑ อัญญายภูมิ อายุ

ก อัญญายภูมิ บริวาร

ร อัญญายภูมิ อุตสาหะ

ก อัญญายภูมิ บริวาร

ช อัญญายภูมิ มนตรี

ณ อัญญายภูมิ เดช

จะเห็นได้ว่า

จำนวนอักษรที่อยู่ในภูมิบริวารมี 2 ตัว คือ ก , ก มีค่าความดีของชื่อคิดเป็น $(9/2)*100 = 22.22\%$

จำนวนอักษรอยู่ในภูมิอาชญา มี 1 ตัว คือ ฯ

ค่าความดีของชื่อคิดเป็น $(9/1)*100 = 11.11\%$

จำนวนอักษรอยู่ในภูมิเดช มี 1 ตัว คือ ณ

ค่าความดีของชื่อคิดเป็น $(9/1)*100 = 11.11\%$

จำนวนอักษรอยู่ในภูมิอุตสาหะ มี 1 ตัว คือ ร

ค่าความดีของชื่อคิดเป็น $(9/1)*100 = 11.11\%$

จำนวนตัวอักษรอยู่ในภูมินัทรี มี 1 ตัว คือ ชะ

ค่าความดีของชื่อคิดเป็น $(9/1)*100 = 11.11\%$

ดังนั้นเราสามารถสรุปได้ว่าชื่อ จักรกฤษณ์ เกิดวันจันทร์ มีค่าความดีทางสถิติของชื่อ (G) คิดเป็น 66.67%

2.3 หลักการอนโนท็อกซี (Ontology) และอนโนท็อกซีเชิงสถิติ (Statistical ontology)

คำว่า Ontology ได้ถูกนำมาเป็นที่รู้จักและใช้กันอย่างแพร่หลายในปัจจุบัน และหมายความว่า ผ่านมา เช่น ในสาขาวิชาปัญญาประดิษฐ์ (Artificial Intelligence: AI) สาขาวิชาคอมพิวเตอร์ (computer) และสาขาวิชาวิทยาศาสตร์สารสนเทศ (information science) โดยเฉพาะอย่างยิ่งใน ขอบเขตสาขางานด้านระบบสารสนเทศ (information system) การรวบรวมสารสนเทศทางปัญญา (intelligence information integration) การสืบค้นข้อมูล (information retrieval and extraction) การแทนที่ความรู้ (knowledge representation) การออกแบบฐานข้อมูล (database design) การจัดการพื้นฐานองค์ความรู้ (knowledge base management) และระบบการจัดการฐานข้อมูล (database management system) (กนกกาญจน์และจักรกฤษณ์, 2006a) (Guarino N, 1998) (Halkidi M et al., 2001) ในสาขางานด้านความหมายของ Ontology ก็ยังคงมีความแตกต่างกันอีกด้วย และไม่มีนิยามเฉพาะตัว ขึ้นอยู่กับการนิยามความหมายของผู้ใช้และ วัตถุประสงค์

Ontology ก็คือการเรียนรู้หรือการเกี่ยวข้องกับชนิดของสิ่งใดๆตามที่มีอยู่ในโลกนี้ (Blackburn S, 1996) Gruber (Gruber T.R., 1998) "ได้นิยามความหมายของ Ontology ซึ่งเป็น ที่รู้จักกันอย่างกว้างขวางมากที่สุด ก็คือ Ontology เป็นรายละเอียดที่ชัดเจนแน่นอนของ แนวความคิด ("An ontology is an explicit specification of a conceptualization") ดังนั้นมี"

เราต้องการแสดงหรือระบุแนวความคิดของวัตถุหรือสิ่งของนั้นๆ เราควรจะแสดงรายละเอียดที่ชัดเจนและแม่นยำ เช่น การอธิบาย Ontology ของโปรแกรม โดยการนิยามกสุ่มของคำที่แสดง และความสัมพันธ์ระหว่างกลุ่มคำนั้นๆ การนิยามนี้อาจจะเกี่ยวข้องกับชื่อของสิ่งที่มีอยู่ในโปรแกรม ซึ่งอาจประกอบไปด้วย Classes, Relations, Functions หรือ Objects อื่นๆ

หลักการอนโนโลจี สามารถนำมาประยุกต์ใช้กับหลักสถิติได้เป็นอย่างดี โดยจะถูกเรียกว่า อนโนโลจีเชิงสถิติ (Statistical Ontology) ซึ่งส่วนใหญ่จะถูกนำไปใช้ในการวิเคราะห์ข้อมูล ทำให้มีองข้อมูล หรือใช้ในการจัดกลุ่มของข้อมูลที่มีการสัมพันธ์กันโดยผลลัพธ์จะถูกแสดงในรูปแบบสถิติ (Denk, M et al., 2002) (Hert, C.A. & Haas, S., 2003)

Marchinonini et al., (Marchinonini et al., 2003) ได้พัฒนา Statistical ontology ในการหาความสัมพันธ์ของหลักสถิติ และเชื่อมโยงหลักการที่เกี่ยวข้องกันเข้าด้วยกัน โดยเริ่มจากการสร้างนิยามคำพิพากษา และพัฒนาการอธิบายเชิงกราฟ ซึ่งอนโนโลจีจะถูกใช้ในการสนับสนุนการสร้างการอธิบาย โดยจะอนุญาตให้ผู้ใช้สืบค้นหลักการสถิติต่างๆ และความสัมพันธ์ที่เกี่ยวโยงหรือเกี่ยงข้องซึ่งกันและกัน

Pasquier et al., (Pasquier et al., 2004) ได้สร้างเครื่องมือสำหรับการวิเคราะห์การทดลองงานที่มีปริมาณมาก ซึ่งเรียกเครื่องมือนี้ว่า THEA (Tool for High-throughput Experiments Analysis) โดยใช้หลักเกณฑ์ทางสถิติ และ Ontology สำหรับการสร้างความหมายทั่วไปจากความรู้เพิ่มฐานที่ได้จากการจัดกลุ่มและการสืบค้นโดยผ่านกระบวนการใช้คำอธิบายประกอบ ซึ่งหลักการอนโนโลจีจะใช้ในการสืบค้นข้อมูลหรือสาขาเฉพาะขององค์ความรู้ทางชีววิทยา และหลักการเหมือนข้อมูลนั้นได้ถูกนำมาใช้ในการจัดกลุ่มและแสดงผลข้อมูลและความรู้ให้อยู่ในรูปแบบต้นไม้ (Tree) อย่างอัตโนมัติ

ในการใช้หลักการอนโนโลจีเชิงสถิติมีช่วยวิเคราะห์และตรวจสอบซึ่งกันและกัน ได้โดยการใช้สูตรจากสมการที่ (1) มาคำนวณค่าความดีเชิงสถิติของชื่อ จากนั้นผลลัพธ์ (G) จะถูกนำไปใช้ในการแบ่งช่วงค่าความดีเชิงอนโนโลจี (O) ได้ดังต่อไปนี้

2.3.1 เกณฑ์กำหนดค่าความดีเชิงอนโนโลจี (O) จากค่าความดีทางสถิติ (G)

If $G > 79\%$ Then

$$O = 2$$

If $49\% < G > 79\%$ Then

$$O = 1.5$$

If $20\% < G > 49\%$ Then

$$O = 1$$

If $0 \% < G > 20 \% \text{ Then}$

$$O = 0.5$$

2.3.2 ค่าความดีทางสถิติโดยแบ่งจากอักษรกลกนี

ชื่อที่ไม่มีอักษรกลกนี จะถูกกำหนดให้มีค่าความดีเชิงอนโนโลจี (O) เท่ากับ 1

ชื่อที่มีอักษรกลกนี จะถูกกำหนดให้มีค่าความดีเชิงอนโนโลจี (O) เท่ากับ -1

จากนั้นระบบจะทำการคำนวณค่าความเชื่อมั่นของชื่อโดยใช้สมการดังต่อไปนี้

$$R = \sum_{i=1}^n \frac{O}{MaxN} \quad \dots \dots \dots \quad (2)$$

จากตัวอย่างข้างต้น ชื่อ จักรกฤษณ์ เกิดวันจันทร์มีค่าความดีของชื่อ (G) คิดเป็น 66.67 % ซึ่งมีค่าความดีทางสถิติ (G) อยู่ในช่วง มากกว่า 49 % แต่น้อยกว่า 79 % จะได้ค่าความดีเชิงอนโนโลจี (O) เท่ากับ 1.5 และ จักรกฤษณ์ เกิดวันจันทร์ ซึ่งไม่มีอักษรอยู่ในภูมิภาคกนี จะถูกกำหนดให้มีค่าความดีเชิงอนโนโลจี (O) เท่ากับ -1

$$\text{จะได้ } \frac{1.5 + 1}{3} = 83.33 \%$$

ดังนั้นเราสามารถสรุปได้ว่าชื่อ จักรกฤษณ์ เกิดวันจันทร์มีค่าความเชื่อมั่น (R) คิดเป็น 83.33 %

2.4 การตรวจสอบชื่อด้วยการจัดกลุ่มด้วยค่าความเชื่อมั่น (R)

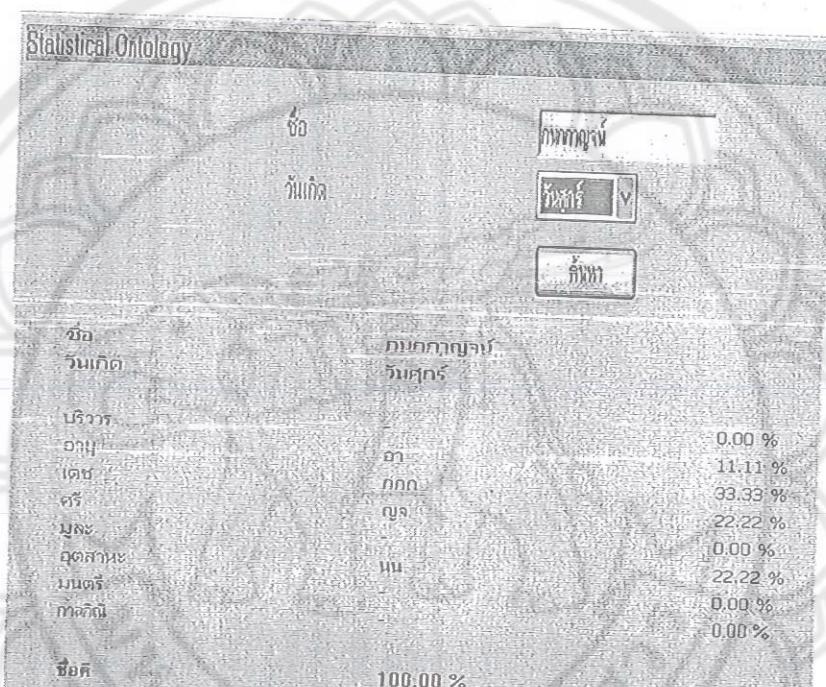
การจัดกลุ่มถือเป็นเทคนิคเป็นเทคโนโลยีที่มีการวิเคราะห์ข้อมูลทางสถิติ ซึ่งถูกใช้ในสาขาต่างๆ รวมทั้ง Machine Learning (การเรียนรู้ของเครื่องจักรกล) Data Mining (เหมืองข้อมูล) และเป็นการจำแนกหมวดหมู่ของสิ่งที่คล้ายกันให้อยู่ด้วยกัน หรือเป็นการแบ่งหมวดหมู่ของข้อมูลให้อยู่เป็นกลุ่มย่อยๆ ซึ่งเรียกว่า Cluster ดังนั้นข้อมูลในแต่ละกลุ่มย่อยๆ จะมีคุณสมบัติที่คล้ายกันหรือเหมือนกัน เทคนิค Clustering นั้นถูกใช้กันอย่างแพร่หลายในหลายสาขาวิชา เช่น Statistics, Machine learning และ Data Mining

จากการตรวจสอบชื่อด้วยตัวเอง จากร้านข้อมูลทั่วไป ระบบสามารถทำการจัดกลุ่มชื่อด้วยแบบอัตโนมัติ (ได้ด้วยตัวเอง) โดยใช้ค่าความดีของชื่อที่มากกว่า 50 เปอร์เซ็นต์ (โดยจัดเรียงลำดับชื่อที่มีค่าความดีมากสุดมาท้าย) และวันเกิดเป็นตัวระบุในการแบ่งกลุ่มได้อีกทั้งผู้ใช้สามารถกำหนดคุณเขตในการแบ่งกลุ่มชื่อด้วยหรือไม่ได้โดยผู้ใช้สามารถเลือกเงื่อนไขต่างๆ ได้ เช่น ผู้ใช้เลือกดูชื่อด้วยผู้ที่เกิดวันอาทิตย์หรือชื่อดีและไม่ตีทั้งหมดโดยไม่เลือกวัน

2.5 การทดสอบและผลลัพธ์

ในหัวข้อนี้เราได้แสดงการทดสอบวันเกิดจากชื่อโดยระบบจะคำนวณความเป็นไปได้ของชื่อวันเกิดวันไหน ในส่วนการทดสอบระบบตรวจสอบชื่อด้วยเราได้ใช้ฐานข้อมูลชื่อนิสิตมหาวิทยาลัยนเรศวร เป็นจำนวน 11658 ชื่อ ว่ามีจำนวนนักเรียนที่ตั้งชื่ออยู่กต้องตามหลักทักษะปกรณ์ หรือไม่อย่างไร

3.1 ผลลัพธ์ของระบบตรวจสอบชื่อด้วยใช้ หลักการอ่อนโน้มโลจี



รูปที่ 2 ผลลัพธ์การตรวจสอบชื่อด้วยใช้ กฎกาณจน์ เกิดวันศุกร์ (Screenshot)

จากรูปที่ 2 จะเห็นได้ว่า ชื่อ "กานกาณจน์" เป็นชื่อที่ดี 100% ซึ่งจะเห็นได้ว่าชื่อนี้ถูกต้องตามหลักทักษะปกรณ์ค่อนข้างสูง และไม่มีอักษรของชื่ออูญในภูมิภาคกินเลย โดยมี ระหว่าง อูญในภูมิภาค อักษร ก อูญในภูมิเดช อักษร ญ และ ฯ อูญในภูมิศรี และอักษร ន อูญในภูมิอุตสาหะ ผลของการตรวจสอบจะออกมากในรูปของการคำนวณเชิงสถิติ โดยใช้สูตรผลรวมของค่าความดีเชิงอ่อนโน้มโลจี (O) หารด้วย ผลรวมของค่าสูงสุดจากการค่าความดีเชิงอ่อนโน้มโลจี (O) (จากชื่อ 2.3.1 และ 2.3.2 ซึ่งก็คือ 3 นั้นเอง) เช่น ชื่อ "กานกาณจน์" มีค่าความดีทางสถิติ (G) เท่ากับ 89% และไม่มีอักษรภาคกิน ดังนั้นค่าความเชื่อมั่น (R) จะมีค่าเท่ากับ $(2+1)/3$ เท่ากับ 100% ซึ่งหมายความว่า ชื่อนี้เป็นชื่อที่ดีเต็มร้อยหรือสมบูรณ์แบบตามหลักทักษะปกรณ์

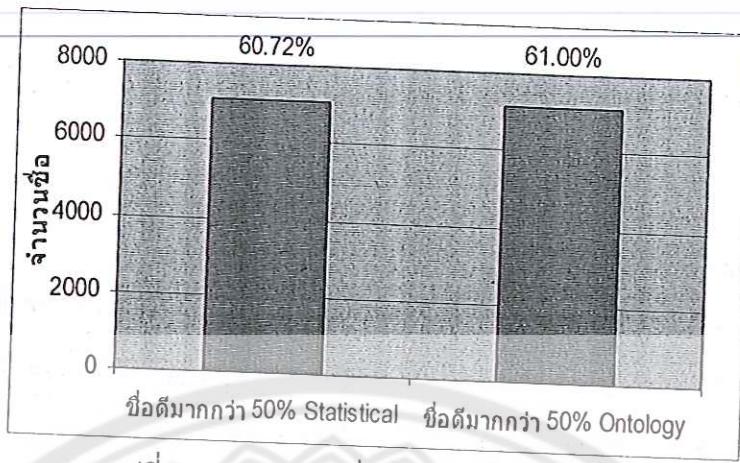
จะเห็นได้ว่าค่าความน่าเชื่อถือหรือความเชื่อมั่นของชื่อ "กานกาณจน์" เพิ่มขึ้นเป็น 11% (89% จากระบบเดิมเพิ่มเป็น 100% ในระบบใหม่)

จากการประมวลผลของระบบที่ปรับปรุงใหม่นี้ ผลการจัดกลุ่มโดยใช้บิสิตมทางวิทยาลักษณ์ เศรษฐจำนวน 11658 ชื่อ แสดงในรูปที่ 3

ชื่อ	วันเกิด	น้ำร้าว	ชาย	เพศ	หญิง	อายุ	อกซ้ำ	น้ำร้าว	กวนตัว	ช่องที่	ช่องที่	Status/ontology
												เด็ก
รวมชาย	ชายทั้งหมด	51%	50.00%	0.00%	17.00%	0.00%	17.00%	0.00%	0.00%	0.00%	0.00%	100.00%
วัฒนธรรมไทย	ชายไทย	74%	50.00%	0.00%	12.00%	12.00%	12.00%	0.00%	0.00%	0.00%	0.00%	100.00%
วัฒนธรรมไทย	ชายเชื้อสายไทย	43%	0.00%	0.00%	0.00%	14.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%
วัฒนธรรมไทย	ชายเชื้อสายไทย	37%	0.00%	0.00%	0.00%	25.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%
วัฒนธรรมไทย	ชายเชื้อสายไทย	33%	0.00%	0.00%	0.00%	0.00%	17.00%	0.00%	0.00%	0.00%	0.00%	100.00%
วัฒนธรรมไทย	ชายเชื้อสายไทย	43%	0.00%	14.00%	0.00%	14.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%
วัฒนธรรมไทย	ชายเชื้อสายไทย	50%	0.00%	12.00%	0.00%	12.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%
รวมหญิง	หญิงทั้งหมด	49%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%
รวมหญิง	หญิงเชื้อสายไทย	59.03%	0.00%	0.00%	0.00%	0.00%	0.00%	0.01%	0.00%	0.00%	0.00%	100.00%
จำนวนคนที่ร่วมชื่อเด็ก 50 % มีจำนวน 0 คน จำนวนคนที่ร่วมชื่อเด็ก มากกว่า 50 % มีจำนวน 7105 คน จำนวนคนที่ร่วมชื่อเด็กต่ำกว่า 50 % มีจำนวน 4553 คน คิดเป็น 39.05 เปอร์เซ็นต์												
จำนวนคนที่ร่วมชื่อเด็ก 0 - 19.99 เปอร์เซ็นต์ มีจำนวน 690 คน จำนวนคนที่ 20 - 49.99 เปอร์เซ็นต์ มีจำนวน 3854 คน จำนวนคนที่ 50 - 79.99 เปอร์เซ็นต์ มีจำนวน 9 คน จำนวนคนที่ 80 - 100.00 เปอร์เซ็นต์ มีจำนวน 20 คน รวมคนที่ร่วมชื่อเด็ก 7105 คน คิดเป็น 39.05 เปอร์เซ็นต์ จำนวนคนที่ร่วมชื่อเด็ก มากกว่า 50 % มีจำนวน 11650 คน												

รูปที่ 3 ผลลัพธ์การตรวจสอบชื่อเด็กโดยใช้หลักการอ่อนไหวให้เลือกฐานข้อมูล

จากการนิยามชื่อเด็กของจักรกรุงษณ์และนกกาญจน์ (2007) ชื่อที่ได้ก็คือชื่อที่ตั้งถูกต้องตามหลักทักษะ ต้องไม่มีอักษรกลากลืนอยู่ในชื่อและต้องมีค่าผลรวมของความน่าจะเป็นของชื่อเด็กมากกว่าหรือเท่ากับ 50 เปอร์เซ็นต์ ซึ่งจากผลลัพธ์ในรูปที่ 3 เราสามารถนำผลลัพธ์จากระบบเดิม (นกกาญจน์และจักรกรุงษณ์, 2007) มาเปรียบเทียบกับผลลัพธ์จากระบบที่ปรับปรุงได้ดังรูปที่ 4 และรูปที่ 5.



รูปที่ 4 แสดงจำนวนซึ่ง超出了 Statistical และ Ontology ของชื่อเดี๋ม

จำนวนนิสิตที่มีชื่อเดี๋ม (ถูกต้องตามหลักทักษะ ไม่มีอักษรกลกนิอยู่ในชื่อและมีค่าผลรวมของความน่าจะเป็นของชื่อเดี๋มมากกว่าหรือเท่ากับ 50%) ของการตรวจสอบชื่อแบบ Statistical มีจำนวน 7079 คน (คิดเป็น 60.72 %) แต่การตรวจสอบชื่อแบบ Statistical Ontology มีจำนวน 7105 คน (คิดเป็น 61.00 %) ซึ่งจะเห็นได้ว่าการตรวจสอบชื่อที่เราพัฒนาขึ้นมาใหม่นั้นมีความน่าเชื่อถือมากขึ้นกว่าเดิมประมาณ 0.28%



รูปที่ 5 แสดงจำนวนซึ่ง超出了 Statistical และ Ontology ของชื่อเดี๋ม

จากผลลัพธ์ในรูปที่ 5 เราสามารถสรุปได้ว่าจำนวนนิสิตที่มีชื่อเดี๋ม (ถูกต้องตามหลักทักษะ ไม่มีอักษรกลกนิอยู่ในชื่อและมีค่าผลรวมของความน่าจะเป็นของชื่อเดี๋มมากกว่าหรือเท่ากับ 80%) ของการตรวจสอบชื่อแบบ Statistical มีจำนวน 5321 คน (คิดเป็น 45.64 %) แต่การตรวจสอบชื่อแบบ Statistical Ontology มีจำนวน 7035 คน (คิดเป็น 60.35 %) จากการใช้ Fuzzy Logic ใน การแบ่งระดับประเมินผลค่าความเชื่อมั่นนั้น เราใช้ค่าความดีของชื่อที่ 80 % (ของทั้ง Statistical Model และ Statistical Ontology) เป็นตัวชี้วัดซึ่งจะเห็นได้ว่าการตรวจสอบชื่อที่เราพัฒนาขึ้นมา

ใหม่นั้นความน่าเชื่อถืออยู่ในระดับที่ดีมากคิดเป็น 60.35 % ซึ่งมีค่าความน่าเชื่อถือมากกว่าระบบ
เก่าประมาณ 15%

2.6 ข้อวิจารณ์ (Discussion)

เราได้เพิ่มค่าความเชื่อมั่นในระบบการตรวจสอบชื่อดี โดยได้ใช้หลักการอนโกลอจีเชิงสถิติมา
ปรับปรุงระบบเดิมให้มีความน่าเชื่อถือมากยิ่งขึ้น ผลลัพธ์จากการทดสอบระบบสามารถสรุปได้ว่า
ระบบที่ปรับปรุงขึ้นมาใหม่นี้มีความน่าเชื่อถืออยู่ในเกณฑ์ดี และมีความเชื่อถือมากขึ้นกว่าระบบ
เดิม 15% โดยประมาณ

ผู้วิจัยหวังว่าระบบจะได้รับการตอบรับและเป็นประโยชน์ที่สำคัญยิ่งในการตรวจสอบชื่อ
ในอนาคต อย่างไรก็ตามระบบที่ปรับปรุงใหม่นี้สามารถแสดงผลของชื่อดีที่ผู้ใช้งานนำไป
ตัดสินใจในการเปลี่ยนชื่อได้เลย (ถ้าชื่อมีความน่าเชื่อมั่นน้อยหรือประกอบไปด้วยอักษรกลีนในชื่อ)
หลังจากระบบได้ทำการตรวจสอบชื่อแล้ว

ในอนาคตผู้วิจัยได้ทำการตรวจสอบระบบและตรวจสอบชื่อนิสิตมหาวทยาลัยนเรศวรว่า
ชื่อที่ประกอบไปด้วยอักษรในภูมิอุตสาหะจะมีผลกระทบของการตั้งชื่อต่อการเรียนหรือไม่อย่างไร
อีกทั้งการตรวจสอบระบบกับชื่อของบุคคลที่เป็นคนมีชื่อเสียง และรวมทั้งชื่อวิเคราะห์ว่าการตั้ง
ชื่อมีผลเกี่ยวกับชีวิตความเป็นอยู่และอนาคตของผู้ใช้หรือไม่อย่างไร จะเป็นงานวิจัยที่จะทำการ
ดำเนินการอย่างต่อเนื่องในภายภาคหน้า

บทที่ 3

บทสรุปและข้อเสนอแนะ

ในงานวิจัยนี้เราได้สร้างระบบการตรวจสอบชื่อ โดยได้นำหลักการ Ontology มาประยุกต์ใช้กับหลักสถิติ โดยเรียกว่า Statistical ontology ซึ่งจะถูกนำมาพัฒนาในการปรับปรุงระบบการตรวจสอบชื่อตี ซึ่งจะทำให้ได้ระบบที่มีมาตรฐานและมีความนำ้หน้ามากยิ่งขึ้น

จากการวิเคราะห์วรรณกรรมและผลการทดสอบระบบในงานวิจัยนี้เราได้เอา concept ของอนโกลิจ มานิยามชื่อตีได้ดังนี้ ชื่อตีคือ ชื่อที่มีค่าความดีมากกว่า 50 เปอร์เซนต์ และต้องไม่มีอักษรากลกิณอยู่ในชื่อ ซึ่งในระบบผู้ใช้มีแนวทางเลือกในการตั้งชื่อที่หลากหลาย พร้อมทั้งสามารถตรวจสอบชื่อที่ตั้งมานั้นดีหรือไม่อย่างไร ซึ่งประโยชน์สูงสุดที่จะได้รับคือจะลดค่าใช้จ่าย ประหยัดเวลา สำหรับผู้ใช้ในการที่จะต้องไปค้นหาชื่อที่ต้องการจะตั้งตามหนังสือการตั้งชื่อ หรือ ปรึกษาพ่อ เป็นต้น

ข้อเสนอแนะ

จากข้อมูลของการวัดค่าทางสถิติจะเห็นได้ว่า ระบบเนื้อความนำ้หน้าเชื่อถือค่อนข้างสูงและประมวลผลได้รวดเร็ว และตรงตามความต้องการตามข้อบ่งชี้ที่ได้กำหนดในบทที่ 1 แต่อย่างไรก็ตามข้อเสนอแนะที่ควรปรับปรุงระบบให้ดียิ่งขึ้นนั้นมีดังต่อไปนี้

- ควรนำระบบผู้เชี่ยวชาญ (Expert system) มาใช้ในการตรวจสอบชื่อและการตั้งชื่อ เช่น ถ้าชื่อที่ได้ทำการตรวจสอบมีค่าความดีทางสถิติน้อยหรือมีอักษรากลกิณอยู่ในชื่อ ระบบจะแสดงหรือแนะนำชื่อที่มีค่าความดีทางสถิติสูงกว่า และชื่อนั้นต้องไม่มีอักษรากลกิณ
- ควรมีการเพิ่มหรือนำเข้าระบบการแปลงชื่อ อังกฤษเป็นไทย ที่พัฒนาโดยจักรกฤษณ์ (Chakkrit et.al., 2006; Chakkrit and Pupong, 2007) ลงในระบบเพื่อเป็นผลดีต่อชาติต่างชาติที่ต้องการตรวจสอบชื่อ โดยใช้หลักการตรวจสอบชื่อไทยแบบทักษะปกรณ์ ได้อิกหั้งควรมีระบบเวอร์ชันภาษาอังกฤษเพื่อผู้ใช้ต่างชาติสามารถใช้ระบบตรวจสอบชื่อและตั้งชื่อผ่านระบบเครือข่ายได้
- ควรมีการเพิ่มฐานข้อมูลชื่อพร้อมทั้งความหมายหรือฐานข้อมูลของพจนานุกรมเข้าไปในฐานข้อมูลของระบบ เพื่อทำให้ชื่อมีความหลากหลาย และมีความนำ้หน้าเชื่อถือมากยิ่งขึ้น สำหรับการตรวจสอบชื่อด้วยวิธีการจัดกลุ่ม
- ควรมำหลักการตรวจสอบชื่อตีไปตรวจสอบคุณภาพชีวิตของนิสิต หรือตรวจสอบดูว่า นิสิตที่มีชื่อตีนั้นผลการเรียนจะเป็นอย่างไร

บรรณานุกรม(Bibliography)

- กนกกาญจน์ นมะหุต และ จักรกฤษณ์ เสน่ห์. (2006). ระบบตราสารบชื่อไทยโดยใช้หลักการตั้งชื่อแบบทักษะปกรณ์. ในการประชุมทางวิชาการระดับชาติด้านเทคโนโลยีสารสนเทศ ครั้งที่ 1 (NCIT2006). 2-3 พฤษภาคม, กรุงเทพมหานคร. หน้า 206-209
- จักรกฤษณ์ เสน่ห์ และ กนกกาญจน์ นมะหุต ระบบตราสารบชื่อด้วยใช้หลักการตั้งชื่อแบบทักษะปกรณ์และการจัดกลุ่ม in review Naresuan University Science Journal, 2007
- H. Andrade and J. Saltz, Query Optimization in Kess – An Ontology-Based KBMS. Proceedings of the 15th Brazilian Symposium on Databases (SBBD'2000). João Pessoa, Brazil, 2000.
- G. Bouchard and C. Pouyez, "Name Variations And Computerised Record Linkage", *Historical Methods*, 13 (2), 1980, pp119-125.
- M. Denk, K.A. Froeschl, and W. Grossmann, "Statistical Composites: A Transformation-bond Representation of Statistical Data. In Proceedings 14th Conf, ACM SIGMOD, Los Alamitos, 219-226, 2002.
- J. Giarratano and G. Riley, *Expert System: Principles and Programming* (3rd edition), PWS Publishing Company , Boston , 1998.
- T.R. Gruber, *A Translation Approach to Portable Ontology Specification Knowledge Acquisition* vol.5 No.2,1998.
- N. Guarino, *Formal Ontology and Information System*. In: N Guarino Ed, PWS Publishing Company , Amsterdam , 1998.
- S. Guha, R. Rastogi, K. Shim. Cure: "An Efficient Clustering Algorithm for Large Database", Published in the Proceedings of the ACM SIGMOD Conference, 1998.
- M. Halkidi, Y. Batistakis, and M. Vazirgiannis. "Cluster algorithms and validity measures", *tutorial paper in the proceedings of the SSDBM*, 2001.
- C. A. Hert and S. Haas, Support End-users of Statistical Information: The Role of Statistical Metadata in the Statistical Knowledge Network" International Conference Proceeding Series.130 , 2003.
- V. I. Levenshtein, "Binary codes capable of correcting deletions, insertions and reversals". Doklady Akademii Nauk SSSR 163, 845-848 (trans. Soviet Physics Doklady 10, 1965), pp707-710.
- G. Marchionini, S. Haas, C. Plaisant, B. Shneiderman, & C. Hert, Towards a statistical knowledge network. Processdings of the National Conference on Digital Government Research 2003, 20(16), 2636-2643, 2003.
- G. A. Miller, Nouns in WordNet: a lexical inheritance system. In International Journal of Lexicography, 3 (4), 245-264, 1990.
- R. S. Pressman, Software Engineering, McGrawHill, International Edition 2001.
- C. Snae, Constructing a Rule Based Naming System for Thai Names Using the Concept of Ontologies , in EACL 2006 workshop 10: "Learning Structured Information in Natural Language Applications", 2006, Trento, Italy.
- C. Snae and B. M. Diaz, An Interface for Mining Genealogical Nominal Data Using the Concept of linkage and a Hybrid Name Matching Algorithm , *Journal of 3D-Forum Society*, 16 (1), 2002, pp. 142-147.
- C. Snae, and Michael Brueckner, Concept and Rule Based Naming System, The Information Universe: Journal of Issues in Informing Science and Information Technology, Volume 3, 2006, pp. 619-634.



ดำเนินการโดย

C. Snae, and Michael Brueckner, Hybrid Name Matching Methods for Rule Based Thai Naming System, *Naresuan University Science Journal*, 2 (2), 2006, pp. 139-150.

C. Snae, N. Singhadech, B. Emapana, and M. Brueckner, Interactive Transliteration Tools for Explanation Level Language System (IT-TELLS), In Proceeding of ITC-CSCC 2006: "The 21st International Technical Conference on Circuits/Systems, Computers and Communications, 2006, Chiang Mai, Thailand.

A. Silberschatz, H. F. Korth, S. Sudarshan, Data Base System Concepts, McGrawHill, 4th Edition 2002.

M. Sipser, Introduction to the theory of computation, PWS Publishing Company 1997.

I. Sommerville, Software Engineering, Addison-Wesley, 6th Edition 2001.

K. Stoffel, J. Saltz, J. Hendler, J. Dick, W. Merz, and R. Miller. "Semantic indexing for complex patient grouping", In *American Medical Informatics Association (AMIA '97)*, 1997.

13 JUL 2011

