

03

มนุษยศาสตร์ดิจิทัล
บทความปริทัศน์Digital Humanities
Review Articleธิตี สุวรรณทัต¹

Thiti Suwannathat

¹ นักศึกษาปริญญาเอก หลักสูตรปรัชญาดุษฎีบัณฑิต สาขาวิชาสหวิทยาการ วิทยาลัยสหวิทยาการ

บทคัดย่อ

โดยจารีตดั้งเดิม มนุษยศาสตร์ (Humanities) เป็นการศึกษาเรื่องราวที่เกี่ยวข้องกับมนุษย์ การสำนึกคุณค่าในความเป็นมนุษย์ เป็นวิชาที่พยายามศึกษาให้เข้าใจ เข้าถึงความรู้ลึกของมนุษย์ ผลงานสร้างสรรค์ของมนุษย์ (Artifacts) ที่มนุษย์ได้สร้างสรรค์ไว้โดยวิธีการศึกษาผ่านความเป็นวิชา (Disciplines) ที่เอามาใช้เฉพาะทาง เช่น วรรณกรรมศึกษา ภาษาศาสตร์ ประวัติศาสตร์ ปรัชญา ดนตรี ศิลปะ ศาสนา วิธีการศึกษาหลักของมนุษยศาสตร์คือการตีความ การประเมินคุณค่าและการวินิจฉัยประสบการณ์ โดยมีจุดประสงค์เพื่อแสวงหาความหมายคุณค่าของประสบการณ์มนุษย์ ลักษณะทั่วไปดังกล่าวทำให้แต่ดั้งเดิมมาวิธีการศึกษาในแขนงวิชามนุษยศาสตร์มีลักษณะที่เน้นวิธีการเชิงอัตนัย

ปัจจุบันเมื่อโลกพัฒนาเข้าสู่ยุคดิจิทัล (Digital) การศึกษาโดยผ่านวิธีการหาความรู้แบบอาศัยความก้าวหน้าจากวิวัฒนาการของเทคโนโลยีคอมพิวเตอร์ ทั้งซอฟต์แวร์และฮาร์ดแวร์โดยเฉพาะรูปแบบการประมวลผลเชิงดิจิทัล (Computation) ได้ทำให้เกิดการศึกษาเชิงสหวิทยาการ (Interdisciplinary) หรือการบูรณาการระหว่างสาขาวิชามากขึ้น และด้วยวิธีการศึกษา (Method) ดังกล่าวทำให้มีแขนงวิชาใหม่ๆ เกิดขึ้นเช่น มนุษยศาสตร์ดิจิทัล (Digital Humanities) อันเป็นการบูรณาการความรู้ทางด้านวิทยาการคอมพิวเตอร์เข้ากับการศึกษาทางด้านมนุษยศาสตร์

คำสำคัญ

สหวิทยาการ, วิธีการศึกษา, มนุษยศาสตร์, มนุษยศาสตร์ดิจิทัล
และการประมวลผลเชิงดิจิทัล

Abstract

Traditionally, the Humanities is study of human, how people appraise and to feel sense of the human being. It also can be described as the study of how people process and document the human experience and artifacts. The traditional humanists have used literature, linguistics, history, philosophy, art, music, religion as discipline to understand and record the world. The principal method of humanities is interpretation of meaning, evaluation of value or scrutinization of human experiences. Therefore, the humanities methodology typically emphasizes the domain of subjectivity.

In the 21st Century, the world moves towards the digital societies. The development of advanced technology, particularly computation has been leading to a novel interdisciplinary approach such as Digital Humanities. The Digital Humanities attempts to integrate the digital technology knowledge with the study of humanities.

Keywords

Interdisciplinary, Method, Humanities, Digital Humanities, Humanities Computing

● บทนำ

ในปัจจุบันเทคโนโลยีดิจิทัลและคอมพิวเตอร์ได้มีการพัฒนาไปอย่างมาก การศึกษาโดยผ่านวิธีการหาความรู้แบบอาศัยความก้าวหน้าจากวิวัฒนาการของเทคโนโลยีคอมพิวเตอร์ทั้งซอฟต์แวร์และฮาร์ดแวร์โดยเฉพาะวิธีการคำนวณและประมวลผลแบบดิจิทัล (Computation) กับข้อมูลขนาดใหญ่ (Big Data) ได้ทำให้เกิดการศึกษาเชิงสหวิทยาการ (Interdisciplinary) หรือการบูรณาการระหว่างสาขาวิชามากขึ้น ด้วยวิธีการศึกษา (Methods) ดังกล่าวทำให้มีแขนงวิชาใหม่ๆ เกิดขึ้นมากมาย เช่น มนุษยศาสตร์ดิจิทัล (Digital Humanities) สังคมและวัฒนธรรมแบบดิจิทัล (Digital Culture & Society) การจัดการทรัพย์สินทางดิจิทัล (Digital Asset Management) ประวัติศาสตร์เชิงดิจิทัล (Digital History) วรรณคดีดิจิทัล (Digital Literature) สิ่งพิมพ์ดิจิทัล (Digital Publishing) จินตทัศน์เชิงดิจิทัล (Digital Visualization) การศึกษาซากพืชและสัตว์ดึกดำบรรพ์เชิงดิจิทัล (Digital Paleontology) นิเวศวิทยาเชิงดิจิทัล (Digital Ecology) การประมวลผลสำหรับสื่อผสม (Computing for Multimedia) สังคมศาสตร์เชิงประมวลผล (Computational Social Science) ระบบสารสนเทศภูมิศาสตร์ในทางโบราณคดี (GIS in Archaeology) และการประมวลผลภาพเชิงดิจิทัล (Digital Image Processing) บางองค์ความรู้ใหม่เหล่านี้ เช่น “มนุษยศาสตร์ดิจิทัล” ยังถูกเรียกแตกต่างกันออกไปอีกหลายชื่อ คือ การประมวลผลทางมนุษยศาสตร์ (Humanities Computing) มนุษยศาสตร์สารสนเทศ (Humanist Informatics) การประมวลผลเชิงภาษาและวรรณคดี (Literary and Linguistic

Computing) ทรัพยากรดิจิทัลในมนุษยศาสตร์ (Digital Resource in the Humanities) และมนุษยศาสตร์อิเล็กทรอนิกส์ (e-Humanities) เป็นต้น

มนุษยศาสตร์ดิจิทัลเป็นการบูรณาการความรู้ทางด้านวิทยาการคอมพิวเตอร์ เข้ากับการศึกษาทางด้านมนุษยศาสตร์เป็นเรื่องเกี่ยวกับการใช้เทคโนโลยี คอมพิวเตอร์และรูปแบบการคำนวณหรือประมวลผลแบบดิจิทัลเข้าไปช่วยในการ ศึกษาวิชาแขนงมนุษยศาสตร์ เดิมรู้จักกันในชื่อของ *การประมวลผลทาง มนุษยศาสตร์* (Humanities Computing) โดยในระยะแรก เริ่มจากความสนใจ ของนักมนุษยศาสตร์ที่จะนำเทคโนโลยีคอมพิวเตอร์เข้ามาใช้ประโยชน์ในการ ศึกษาทางมนุษยศาสตร์เช่นใช้คอมพิวเตอร์ช่วยในการจัดเก็บข้อมูลงานเขียนและ ตัวบททางวรรณกรรมต่างๆ ที่ต้องการศึกษาด้วยการแปลงข้อมูลเหล่านั้นให้ รูปแบบอิเล็กทรอนิกส์เพื่อง่ายและสะดวกต่อการสืบค้น ต่อมาก็ได้ประยุกต์นำมา ใช้ในการศึกษาตัวบทต่างๆ โดยใช้เครื่องคอมพิวเตอร์เข้ามาช่วยในการประมวล ผล (automated text analysis) ตั้งแต่นำมาใช้ศึกษาพระคัมภีร์ ตัวบท วรรณกรรม โคลง ฉันท์ กาพย์ กลอน ข้อเขียนทางการเมืองไปจนถึงบทบัญญัติ ทางกฎหมายหรือใช้วิธีทางสถิติเพื่อหารูปแบบความสัมพันธ์ของการเลือกใช้ ถ้อยคำกับสไตล์การเขียนของนักเขียนแต่ละคนเพื่อแก้ปัญหาเรื่องความกำกวมใน เรื่องผู้แต่ง

ปัจจุบันงานศึกษาและวิจัยด้านมนุษยศาสตร์ดิจิทัลได้ขยายขอบเขตไป อย่างมากตามความก้าวหน้าทางเทคโนโลยี มีการใช้ประโยชน์จากวิวัฒนาการด้าน คอมพิวเตอร์และสารสนเทศไปจนถึงวิทยาการข้อมูลและการจัดการข้อมูลขนาดใหญ่เข้าไปช่วยในการศึกษามนุษย์ในสาขาต่างๆ เกือบทุกด้านไม่ว่าจะเป็นทาง ด้าน วรรณกรรมศึกษา ภาษาศาสตร์ วัฒนธรรม ประวัติศาสตร์ มานุษยวิทยา ปรัชญา ดนตรี นาฏศิลป์ บทละคร สุนทรียศาสตร์ ไปจนถึงภูมิศาสตร์และ โบราณคดีซึ่งเป็นสาขาวิชาที่ได้ประโยชน์อย่างมากจากการใช้เทคโนโลยีดิจิทัลเป็น

เครื่องมือ เพื่อขยายความรู้ในศาสตร์ของตนเองให้กว้างขวางยิ่งขึ้น²

ในขณะที่ชุมชนวิชาการตะวันตกได้มีการนำแนวคิดเรื่องมนุษยศาสตร์เข้า มาในวงวิชาการอย่างจริงจังตั้งแต่ทศวรรษที่ 1990s (Gold, 2012 : 87) แต่คำว่า “Digital Humanities” จะยังคงเป็นคำที่ใหม่อยู่ในวงวิชาการมนุษยศาสตร์ไทย ข้อบ่งชี้ประการหนึ่งคือยังไม่ได้มีการบัญญัติคำภาษาไทยอย่างทางการจากราชบัณฑิต แต่กลุ่มนักวิชาการไทยที่ใช้คำนี้เป็นกลุ่มแรกๆ ในภาษาไทย คือ กลุ่ม วิจัยมนุษยศาสตร์ดิจิทัล มหาวิทยาลัยขอนแก่น (Khon Kaen University - Digital Humanities Research Group : KKU-DHRG) ที่ก่อตั้งอย่างเป็นทางการ ในปี 2558 โดย ดร.กุลธิดา ท้วมสุข รองศาสตราจารย์ประจำสาขาวิชาสารสนเทศ และการสื่อสาร คณะศึกษาศาสตร์และสังคมศาสตร์³

จากข้อเท็จจริงดังกล่าวแสดงให้เห็นว่า มนุษยศาสตร์ดิจิทัลมีความสำคัญ มากขึ้นในปัจจุบันและจะเป็นหนึ่งในแนวทางการศึกษาแบบประสานความรู้ใน สาขาต่างๆ เข้าด้วยกัน (interdisciplinary approach) ที่สำคัญในอนาคต ดังนั้น บทความขึ้นนี้จะชี้ให้เห็นว่ามนุษยศาสตร์ดิจิทัลคืออะไร พัฒนาการของแนวคิด

² ตัวอย่างงานวิจัยที่นำเทคโนโลยีดิจิทัลและวิธีประมวลผลแบบคอมพิวเตอร์เข้าไปใช้ในการศึกษาด้านมนุษยศาสตร์ในแขนงต่างๆ เช่น งานของ Thomas Rommel, “Literary Studies”; John Burrows, “Textual Analysis”; William Winder, “Robotic Poetics”; Jan Hajic, “Linguistics Meets Exact Sciences”; Russon Wooldridge, “Lexicography”; William G. Thomas III, “Computing and the Historical Imagination”; Michael Greenhalgh, “Art History”; Greg Crane, “Classic and the Computer: An End of the History”; Ichiro Fujinaga and Susan Forscher Weiss, “Music”; David Z. Saltz, “Performing Arts”; Johanna Drucker, “Speculative Computing : Aesthetic Provocation in Humanities Computing”; Harrison Eiteljorg II, “Computing for Archaeologists” ใน A Companion to Digital Humanities, ed. Susan Schreibman, Ray Siemens, and John Unsworth (Blackwell, 2004)

³ <https://dh.kku.ac.th/th/background/> [Retrieved: August 10, 2017]

และองค์ความรู้มีความเป็นมาอย่างไร บทความนี้เป็นการศึกษาเชิงเอกสาร (documentary Research) ผู้เขียนได้ทำการสำรวจความเป็นมาและแนวคิดต่างๆ ที่เกี่ยวกับมนุษยศาสตร์ดิจิทัลโดยเน้นใช้เอกสารและงานวิจัยจากภาษาอังกฤษ รวมทั้งข้อมูลจากสิ่งพิมพ์อิเล็กทรอนิกส์และเว็บไซต์ของศูนย์วิจัย สถาบันความคิดด้านมนุษยศาสตร์ดิจิทัลในต่างประเทศเป็นหลัก แหล่งข้อมูลทั้งหมดจะได้รับการตรวจสอบอย่างละเอียด เพื่อความมั่นใจในความน่าเชื่อถือของเนื้อหาของข้อมูลที่สืบค้นได้ โดยทำการวิเคราะห์ทั้งในลักษณะเชิงพรรณนาความ (Descriptive Analysis) และอธิบายความ (Explanatory Analysis) ทั้งนี้ผู้เขียนได้แบ่งเนื้อหาออกเป็น 5 ส่วนคือ ส่วนแรกจะเป็นการเกริ่นนำ ส่วนที่สองจะเป็นการปูพื้นเกี่ยวกับมนุษยศาสตร์ดิจิทัลคืออะไร มีนิยามและ/หรือถูกให้คำจำกัดความว่าอย่างไร ส่วนที่สามเป็นฉายภาพให้เห็นถึงพัฒนาการและการก่อตัวของมนุษยศาสตร์ดิจิทัล ส่วนที่สี่กล่าวถึงการศึกษาวิจัยด้านมนุษยศาสตร์ดิจิทัลที่ใช้การบูรณาการระหว่างวิชาคอมพิวเตอร์ สถิติและคณิตศาสตร์เข้ากับวิชาในแขนงมนุษยศาสตร์ และส่วนสุดท้ายเป็นบทสรุป

● นิยามและคำจำกัดความ

แม้จะยังไม่มีคำตอบที่ชัดเจนว่ามนุษยศาสตร์ดิจิทัล (ต่อไปจะขอเรียกว่า “DH”) มีคำจำกัดความว่าอย่างไร มีสถานะอะไร อยู่ในตำแหน่งแห่งหนใดในสังคมวิทยาว่าด้วยความรู้ (Sociology of Knowledge)⁴ แต่ก็เป็นที่ยอมรับกันใน

⁴ ดูข้อถกเถียง เรื่องสถานะและคำนิยามของ Digital Humanities ได้ใน Debates in the Digital Humanities, ed. Matthew K. Gold (University of Minnesota Press, 2012); Debates in the Digital Humanities, ed. Matthew K. Gold and Lauren F. Klein (University of Minnesota Press, 2016); Defining Digital Humanities, ed. Melissa Terras, Julianne Nyhan and Edward Vanhoutte (Ashgate Publishing Company, 2013) และ A Companion to Digital Humanities, ed. Susan Schreibman, Ray Siemens, and John Unsworth (Blackwell, 2004)

เบื้องต้นว่า DH มีลักษณะเป็น (ชุดของ) เครื่องมือและวิธีวิทยาที่นำมาใช้ในการได้มาซึ่งความรู้มากกว่าการเป็นสาขาวิชา (Disciplinary) ดังที่ John Unsworth อดีตคณบดีวิทยาลัย วิทยาการข้อมูล มหาวิทยาลัยอิลลินอยส์ ให้นิยามไว้สั้นๆ ว่า “using computational tools to do the work of the humanities” (Terras, Nyhan and Vanhoutte, 2013 : 293) ซึ่งใกล้เคียงกับที่ Johanna Drucker ศาสตราจารย์ ภาควิชาข้อมูลศึกษา (Information Studies) มหาวิทยาลัยแคลิฟอร์เนีย ลอสแอนเจลิส (UCLA) กล่าวถึง DH ว่าเป็น “using digital tools to extend humanistic inquiry” หรือ “เป็นการใช้เครื่องมือทางดิจิทัลเพื่อช่วยในการได้มาซึ่งความรู้ทางมนุษยศาสตร์” (Drucker, 2009 : xi)

แต่ถ้าลองใช้ Google ค้นหาคำนี้ดูในช่วงปี 2015-2016 ก็จะพบคำนิยามที่ถูกเชื่อม (link) ไปในเว็บไซต์ของ Wikipedia ขึ้นมาเป็นลิงค์อันดับแรกสุดที่อธิบายว่า “Digital Humanities หรือที่รู้จักในชื่อเดิมว่า Humanities Computing เป็นสาขาวิชาหนึ่ง การวิจัยวิธีหนึ่ง การสอนรูปแบบหนึ่งและนวัตกรรมอย่างหนึ่งที่เป็นการใช้วิธีประมวลผลด้วยคอมพิวเตอร์เข้าไปผสมผสานกับการศึกษาในแขนงวิชาทางด้านมนุษยศาสตร์ โดยตัวของมันแล้ว DH ถูกจัดว่าเป็นวิธีการศึกษาแบบสหวิทยาการที่ใช้จัดเก็บ รวบรวม ตรวจสอบ วิเคราะห์ สังเคราะห์ ไปจนถึงการนำเสนอข้อมูลในรูปแบบดิจิทัล รวมถึงศึกษาผลกระทบที่ DH มีต่อวิธีการหาความรู้ทางมนุษยศาสตร์แบบดั้งเดิม”⁵

ต่อมาในปี 2018 Wikipedia ปรับเปลี่ยนคำจำกัดความของ DH ใหม่ว่า “เป็นการพบกันทางวิชาการระหว่างเทคโนโลยีการคำนวณแบบดิจิทัลกับวิชามนุษยศาสตร์ เป็นการประสานความรู้ระหว่างสาขาด้วยแนวทางการศึกษาแบบใหม่ ด้วยวิธีการวิจัยข้อมูลโดยการประมวลผลแบบดิจิทัล เป็นการนำ

⁵ http://en.wikipedia.org/wiki/Digital_humanities [Retrieved November 3, 2016]

เครื่องมือและวิธีวิทยาแบบดิจิทัลเข้ามามีใช้ในการศึกษามนุษยศาสตร์”⁶ คำจำกัดความใหม่ของ DH ใน Wikipedia นี้มาจากการผสมผสานในงานศึกษาเรื่อง DH ของ Terras (2012) Drucker (2013) และ Burdick, Drucker, Lunenfeld, Presner and Schnapp (2012) ตามลำดับ ส่วนสารานุกรมวิกิพีเดีย ภาาภาษาไทยให้คำนิยามของมนุษยศาสตร์ดิจิทัลไว้สั้นๆ ว่า “เป็นสาขาหนึ่งของการวิจัย การสอน และการสร้างสรรค์ ที่พิจารณาถึงการตัดข้ามของศาสตร์คอมพิวเตอร์ และสาขาวิชาของมนุษยศาสตร์ มนุษยศาสตร์ดิจิทัลพัฒนามาจากสาขาวิชาที่เรียกว่า คอมพิวเตอร์สำหรับมนุษยศาสตร์ ปัจจุบันสาขานี้ครอบคลุมหัวข้อหลากหลาย ตั้งแต่การถักทอผ้าออนไลน์ ไปจนถึงการทำเหมืองข้อมูลชุดข้อมูลวัฒนธรรมขนาดใหญ่”⁷

ขณะที่วารสาร Digital Humanities Quarterly⁸ ฉบับปฐมฤกษ์ ได้ให้คำนิยามของ มนุษยศาสตร์ดิจิทัลไว้ว่า “เป็นสาขาวิชาเกิดใหม่ที่มีความหลากหลายในตัวเองเพราะเป็นการทำงานวิจัยด้านมนุษยศาสตร์ในแขนงต่างๆ ผ่านทางเทคโนโลยีคอมพิวเตอร์ สารสนเทศและการประมวลผลแบบดิจิทัลเพื่อเปิดพรมแดนความรู้ระหว่างสาขาผ่านการหยิบยืมวิธีการศึกษาที่อาจใช้ร่วมกันได้เพื่อเอกภาพทางวิทยาการ”

⁶ http://en.wikipedia.org/wiki/Digital_humanities [Retrieved January 4, 2018]

⁷ <https://th.wikipedia.org/wiki/มนุษยศาสตร์ดิจิทัล> [Retrieved : June 13, 2018]

⁸ <http://www.digitalhumanities.org/dhq/vol/1/1/000007/000007.html>
[Retrieved: June 7, 2018]

Klein (2015 : 1-2) กล่าวถึง DH ในฐานะแขนงวิชาเกิดใหม่ที่กำลังมีการพัฒนาอย่างรวดเร็ว⁹ มีลักษณะเป็นสหวิทยาการ (MaCarty 1999; Rockwell 1999; Svensson 2009; Rosembloom 2012) ที่ประสานความรู้ด้านวิทยาการคอมพิวเตอร์เข้ากับการศึกษาทางมนุษยศาสตร์ วัฒนธรรมศึกษา นิเทศศาสตร์ ศึกษาศาสตร์ บรรณารักษศาสตร์และวิทยาการสารสนเทศ (Informatics) Klein ฉายภาพความป็นสหวิทยาการของ DH ต่อไปอีกว่า ความสัมพันธ์ระหว่างวิชา (Inter-Discipline) วิทยาการคอมพิวเตอร์กับสาขาวิชาดังกล่าวข้างต้นได้แตกแขนงไปสู่งานศึกษาวิจัยเรื่องต่างๆ เพื่อสร้างบูรณาการระหว่างสาขาวิชาอันนำไปสู่องค์ความรู้แบบใหม่ เช่น การตีความด้วยบททางดิจิทัล (Computer-Assisted Interpretation) การเล่าเรื่องเชิงปริมาณหรือการเล่าเรื่องด้วยตัวเลข (Quantitative Narrative Analysis) การวิเคราะห์ด้วยผ่านระบบอัตโนมัติ (Textual-Automated Analysis) ภาษาศาสตร์เชิงคำนวณ (Computational Linguistics) การประมวลผลด้านภาษา (Language Processing) การสร้างและผลิตด้วยบทดิจิทัล (Digital Text Production) วรรณกรรมอิเล็กทรอนิกส์ (Electronic

⁹ วิชาที่มีลักษณะแบบสหวิทยาการนั้นคือวิชาที่มีศาสตร์หลายแขนงรวมอยู่ด้วยกันในวิชาเดียว ศาสตร์แต่ละแขนงที่รวมอยู่ในตัววิชานั้นล้วนแล้วแต่มีความสัมพันธ์ซึ่งกันและกันตัวอย่างหลักสูตรสหวิทยาการซึ่งเป็นที่รู้จักกันดีโดยทั่วไปก็เช่น หลักสูตร B.A. (Philosophy Politics and Economics) ของ Oxford University, หลักสูตร B.A. (Government) ของ Harvard University ซึ่งผู้เรียนต้องเรียนวิชาที่หลากหลายและมีความสัมพันธ์ซึ่งกันและกันโดยต้องเรียนทั้ง ปรัชญา การเมือง การปกครอง เศรษฐศาสตร์และกฎหมายต่างๆ เหล่านี้เป็นต้น รวมทั้งมหาวิทยาลัยส่วนในฝรั่งเศสที่การเรียนกฎหมายและเศรษฐศาสตร์จะสอนอยู่ในคณะเดียวกัน (La Faculte De Droit Et Des Sciences Economiques) เจตนา นาควัชระ ตั้งข้อสังเกตว่า การแยกกลุ่มวิชาออกเป็น มนุษยศาสตร์ (Humanities) กับ สังคมศาสตร์ (Social Sciences) เป็นสิ่งที่พึงจะเกิดขึ้นไม่นาน แต่เดิมมาการบริหารงานวิชาการของมหาวิทยาลัยในตะวันตก จะจัดรวมวิชาทั้งที่เป็นมนุษยศาสตร์และสังคมศาสตร์เข้าไว้ด้วยกัน เช่น Faculty of Arts ในประเทศที่ใช้ภาษาอังกฤษ ส่วนประเทศที่ใช้ภาษาฝรั่งเศส จะเป็น Faculté des Lettres et Sciences Humaines และคำว่า Sciences

Literature) และ ห้องสมุดดิจิทัล (Digital Library) ที่อาศัยความก้าวหน้าจากเทคโนโลยีคอมพิวเตอร์และดิจิทัลช่วยในการจัดเก็บ ค้นหา หนังสือ เอกสาร บรรณานุกรม จดหมายเหตุในรูปแบบอิเล็กทรอนิกส์ (Electronic Forms) DH ยังครอบคลุมเรื่องวิธีการเรียน การสอนในรูปแบบใหม่โดยผ่านสื่ออิเล็กทรอนิกส์ และการสื่อสารทางดิจิทัล รวมไปถึงการทำงานวิจัยแบบ E-Research บนฐานความรู้และฐานข้อมูลจาก Cyber-Infrastructure ปัจจุบันบทความวิชาการและงานวิจัยทั้งด้านสังคมศาสตร์และมนุษยศาสตร์ได้ถูกจัดเก็บในรูปแบบดิจิทัล (Digital Form) มากขึ้นเช่น e-thesis, e-journal, e-book, e-magazine, etc. กิจกรรมทางวิชาการเหล่านี้ล้วนถูกจัดอยู่ภายใต้ปริมณฑลของ DH

กุลธิดา (2558) อธิบายว่า มนุษยศาสตร์ดิจิทัล (Digital Humanities)¹⁰ เป็นการบูรณาการความรู้ทางด้านวิทยาการคอมพิวเตอร์เข้ากับการศึกษาทางด้านมนุษยศาสตร์ การศึกษาวิจัยในทางมนุษยศาสตร์ดิจิทัลจะครอบคลุมการศึกษาเนื้อหาเกี่ยวกับการรวบรวมและจัดระบบความรู้ การเก็บและรักษาความรู้ในรูปแบบดิจิทัลทั้งที่เป็นข้อความและสื่อผสม และการอธิบายความรู้ในมิติเชิงความหมาย (Semantic) ตัวเลขและสถิติ (Statistics) และจินตทัศน์ (Visualization) ในเรื่องที่น่าสนใจในสาขาวิชามนุษยศาสตร์และ/หรืออาจครอบคลุมไปถึงผลงานใดๆของมนุษยชาติ มีการผนวกระเบียบวิธีการศึกษา (Methodology) ทั้งวิธีการดั้งเดิมในสาขามนุษยศาสตร์ เช่น ประวัติศาสตร์ ปรัชญา วรรณคดีศึกษา ภาษาศาสตร์ มานุษยวิทยาวัฒนธรรมศึกษา ศิลปะ ดนตรี ร่วมกับการใช้เครื่องมือและวิธีการทางคอมพิวเตอร์ อาทิ จดหมายเหตุดิจิทัล (Digital Archiving) เว็บเชิงความหมาย (Semantic Web) การทำเหมืองข้อมูล (Data Mining) การนำ

¹⁰ Humaines ในภาษาฝรั่งเศส รวมถึงทั้งสังคมศาสตร์และมนุษยศาสตร์ ส่วนในระบบการศึกษาของเยอรมันก็จัดแขนงวิชาทางมนุษยศาสตร์และสังคมศาสตร์รวมไว้ในคณะที่เรียกว่า Philosophische Fakultät (อ้างอิงใน เจตนา นาควัชระ, 2538 : 226)

เสนอข้อมูลหลายมิติ (Multi-media Presentation) การนำเสนอข้อมูลแบบจินตทัศน์ (Data Visualization) และการวิเคราะห์ตัวบท (Textual Analysis) เป็นต้น (กุลธิดา, 2558)

● พัฒนาการของมนุษยศาสตร์ดิจิทัล

การบูรณาการวิธีการคำนวณแบบดิจิทัลเข้ากับการศึกษาด้านมนุษยศาสตร์ ถูกบุกเบิกโดย โรแบร์โต บูซา (Roberto Busa : 1913-2011) นักเทววิทยาชาวอิตาลี ผู้ได้รับการยกย่องเป็นบิดาแห่ง Digital Humanities¹¹ (ในเวลานั้นยังถูกใช้ชื่อว่า Humanities Computing) ที่ทำดัชนีค้นคำในงานเขียนของ Thomas Aquinas ทั้งหมดชื่อ *Index Thomisticus* งานชิ้นนี้เป็นผลสืบเนื่องมาจากการทำวิทยานิพนธ์ปริญญาเอก¹² ของเขาในช่วงที่สงครามโลกครั้งที่ 2 ใกล้กำลังจะสิ้นสุด อันเป็นการศึกษาความคิดเรื่อง “Presence” ในงานเขียนของ Thomas Aquinas บูซาเริ่มต้นด้วยการศึกษารายละเอียดเกี่ยวกับความคิดเรื่อง “Presence” ด้วยการค้นหาคำว่า “Presence” ผ่านดัชนีตารางและหัวข้อเรื่องทั้งหมดในงานของ Aquinas และพบว่าการปรากฏขึ้นของคำว่า “Presence” ทุกครั้งจะต้องถูกเชื่อมกับคำบุพบท (preposition) “in” อยู่เสมอ ขั้นตอนต่อไป บูซาจึงเขียนประโยคทุกประโยคที่มีคำว่า “in” หรือ คำทุกคำที่อยู่ติดกับคำว่า “in” ลงในกระดาษขนาด 3*5 นิ้ว จำนวน 10,000 ใบที่เตรียมไว้ (Busa, 1980)

ด้วยวิธีดังกล่าวทำให้บูซาพบผลการศึกษาสองข้อ คือ 1) “ก่อนที่จะทำความเข้าใจหรือทำการตีความความคิดในงานของอะไควนัสหรือนักเขียนคนใดนั้น จำเป็นต้องศึกษา ที่มา ที่ไปของคำในเชิงนิรุกติศาสตร์ก่อนเพราะนักเขียนแต่ละคน

¹¹ Humaines ในภาษาฝรั่งเศส รวมถึงทั้งสังคมศาสตร์และมนุษยศาสตร์ ส่วนในระบบการศึกษา

¹² วิทยานิพนธ์ของบูซา มีชื่อในภาษาอิตาลีว่า “La Terminologia Tomistica dell'Interiorita: Saggi di metodoper una interpretazione della metafisica della presenza”

จะแสดงระบบความคิดของตัวเองผ่านระบบคำพูดของเขาเอง ฉะนั้นการที่ผู้อ่านจะค้นพบความรู้ที่อยู่ในระบบคิดของผู้เขียนนั้นจำเป็นที่จะต้องทำความเข้าใจความรู้สึกนึกคิดของผู้เขียนคนนั้นๆ ผ่านภาษาพูดที่เขาใช้ ผู้อ่านควรที่จะต้องพยายามค้นหานัยสำคัญของ/(ที่) คำเหล่านี้อยู่ในใจของผู้เขียน”

และ 2) บุชาคิดว่า “ที่มา ที่ไปของคำและตำแหน่งแห่งหนในประโยคของมัน แสดงให้เห็นถึงตรรกะเบื้องต้นที่สุดในสภาวะการดำรงอยู่ของมัน อันนำไปสู่การเกิดโครงสร้างพื้นฐานของคำพูดมนุษย์” กล่าวคือบุชาคิดว่า คำแต่ละคำนั้นล้วนมีที่มา และเมื่อถูกนำไปใช้ในต่างบริบทก็จะก่อให้เกิด พลิกผันทำให้ความหมายต่างกันไป (Busa, 1980 : 83-90) กรอบคิดดังกล่าวนี้เป็นการกลับไปใช้วิธีการศึกษาวรรณกรรมแบบดั้งเดิมในแนวนิรุกติศาสตร์ ขณะที่ช่วงเวลานั้นทฤษฎีวรรณกรรมแบบการวิจารณ์แนวใหม่ (New Criticism) กำลังมีอิทธิพลอย่างมากต่อการศึกษาวรรณกรรมในกลางศตวรรษที่ 20¹³ (ชูศักดิ์ ภัทรกุลวณิช, 2545 : 393)

ในปี ค.ศ. 1946 ภายหลังจากวิทยานิพนธ์ฉบับดังกล่าวเสร็จสิ้น บุชาเริ่มคิดถึงโครงการจัดทำดัชนีคำทั้งหมดในงานเขียนของอะไควนัสรวมถึงคำประเภท สรรพนาม สันธานและบุพบทด้วย เพื่อเป็นประโยชน์สำหรับการศึกษาเชิงเปรียบเทียบ อย่างไรก็ตามเขาตระหนักดีว่ากระบวนการแยกแยะ จัดกลุ่มคำดังกล่าวที่มีจำนวนมากกว่าสิบล้านคำนั้นไม่สามารถทำได้ด้วยแรงงานคน แต่ต้องอาศัยเครื่องจักรเข้ามาช่วย ดังนั้นในปี 1949 บุชาพร้อมด้วยคณะลูกศิษย์จำนวนหนึ่งจึงเดินทางไปยังมหาวิทยาลัยต่างๆ ทั่วประเทศสหรัฐอเมริกาเพื่อตรวจสอบว่า มีสถาบันการศึกษาแห่งไหนมีเครื่องจักรประมวลผลที่จะสามารถเอาเข้ามาช่วยเป็นเครื่องมือในโครงการวิจัยของเขาได้ บุชาเริ่มต้นด้วยการไปพบกับ

¹³ ดูรายละเอียดวิธีการศึกษาวรรณกรรมในแนวทางต่าง ๆ ใน “เลาะเลียบบเส้นทางวรรณกรรม” ใน อ่าน (ไม่)เอาเรื่อง, น. 359-374, ชูศักดิ์ ภัทรกุลวณิช, โครงการจัดพิมพ์คบไฟ 2548.

หัวหน้าฝ่ายกองกิจยุโรป หอสมุดรัฐสภา สหรัฐอเมริกา ซึ่งได้แนะนำให้เขารู้จักกับ Jerome Wiesner¹⁴ ศาสตราจารย์ด้านวิศวกรรมอิเล็กทรอนิกส์ สถาบันเทคโนโลยีแมสซาชูเซตส์ (MIT) และไวซ์เนอร์ก็พาบุชาไปพบกับ Thomas Watson¹⁵

ประธานและซีอีโอบริษัท IBM ผู้ซึ่งตกลงที่จะช่วยสนับสนุนเรื่องเครื่องจักรประมวลผลในงานของบุชา ซึ่งในเวลานั้นยังคงเป็นเครื่องคอมพิวเตอร์ที่รับอินพุต (Input) ข้อมูลหรือใช้วิธีอ่านโค้ดจากกระดาษแผ่นเจาะรู (Punched Card) ที่ออกแบบให้วิ่งผ่านช่องอ่านของเครื่องเพื่อให้อ่านข้อมูลหรือคำสั่งเข้าไปเพื่อทำการประมวลผลต่อไป ในโครงการจัดทำดัชนีคำ (concordance) ในงานเขียนทั้งหมดของโรมัส อะไควนัส หรือ *Index Thomisticus*¹⁶ บุชากับทีมงานใช้เวลาสามสัปดาห์ในการทำบัตรหรือกระดาษที่นำมาเจาะเป็นรูคำต่อคำ จำนวน 10,666,000 คำ จาก 1,700,000 บรรทัด เพื่อเป็นข้อมูลที่จะส่งเข้าสู่เครื่องคอมพิวเตอร์แบบเมนเฟรมประมวลผล แยกแยะ จัดหมวดหมู่คำต่างๆ เพื่อออกมาเป็นดัชนี (Busa, 1980 : 86-88)

Frederick Mosteller และ David Wallace เป็นผู้บุกเบิกรุ่นแรกๆ อีกกลุ่มหนึ่งที่ใช้การบูรณาการความรู้ด้านสถิติและการประมวลผลแบบคอมพิวเตอร์เข้ากับวิชาทางมนุษยศาสตร์ในงานศึกษาเรื่อง “*Inference in an Authorship Problem*” (Mosteller and Wallace, 1963) ที่ใช้วิธีการศึกษาแบบ DH เข้าไปช่วยหาคำตอบในประเด็นที่นักประวัติศาสตร์ถกเถียงกันว่าระหว่าง Alexander

¹⁴ ในช่วงรัฐบาล จอห์น เอฟ. เคนเนดี (1961-1963) เจอโรม ไวซ์เนอร์ได้รับการแต่งตั้งให้เป็นประธานคณะกรรมการที่ปรึกษาด้านวิทยาศาสตร์ของประธานาธิบดีเคนเนดี ต่อมาระหว่างปี 1971-1980 ไวซ์เนอร์ก็มารับตำแหน่งอธิการบดีของ MIT

¹⁵ โรมัส วัตสันเป็นประธานคนแรกของ IBM ปัจจุบันชื่อ Watson เป็นชื่อโครงการพัฒนาระบบปัญญาประดิษฐ์ของศูนย์วิจัย Thomas J. Watson ที่ตั้งชื่อตามนายวัตสัน

¹⁶ <http://www.corpusthomicum.org/it/index.agejsessionid=2D741DEBCE5DE21CE3CEFFE7D5A16CB6> [Retrieved: February 9, 2018]

Hamilton กับ James Madison ใครเป็นผู้เขียนบทความ 12 ชิ้นในเอกสารเฟเดอรัลลิสต์ (The Federalist Papers)¹⁷ ที่ทั้งคู่อ้างสถานะการเป็นผู้เขียนในบทความ 12 ชิ้นนั้น

ทั้งนักประวัติศาสตร์และนักรัฐศาสตร์พยายามหาวิธีพิสูจน์ประเด็นถกเถียงดังกล่าวมากกว่า 150 ปี เช่นด้วยวิธีพยายามแกะรอยตามอุดมการณ์ทางการเมืองที่ถูกนำเสนอออกมาในบทความ 12 ชิ้นที่ไม่สามารถระบุได้อย่างชัดเจนว่าใครคือผู้เขียน ด้วยการตรวจสอบว่าบทความไหนมีความสอดคล้องกับความคิดทางการเมืองของแฮมิลตันหรือเมดิสันมากกว่ากันหรือแม้กระทั่งจอห์น เจย์ ผู้ดำรงตำแหน่งประธานศาลสูงสุด คนแรกของประเทศสหรัฐอเมริกา แต่สุดท้ายก็ยังไม่ได้ข้อสรุปที่แน่ชัดในหมู่นักศึกษาเรื่องนี้

ในปี 1963 Mosteller และ Wallace ใช้วิชาสถิติเป็นเครื่องมือในทางวิทยาศาสตร์ โดยใช้วิธีง่าย ๆ ด้วยการนับความถี่ของคำบางคำ ที่อยู่ในบทความชิ้นที่ระบุได้แน่ชัดว่าแฮมิลตันหรือเมดิสันเป็นคนเขียน ขั้นตอนต่อมาคือใช้วิธีเดียวกันกับบทความ 12 ชิ้นที่ทั้งแฮมิลตันและเมดิสันต่างอ้างว่าตนเป็นผู้เขียน หลังจากนั้นก็เอาความถี่ของการใช้คำทั่วไปที่ค้นพบในขั้นตอนแรกมาเปรียบเทียบกับขั้นตอนที่สอง จากวิธีการศึกษาดังกล่าว Mosteller และ Wallace พบว่า

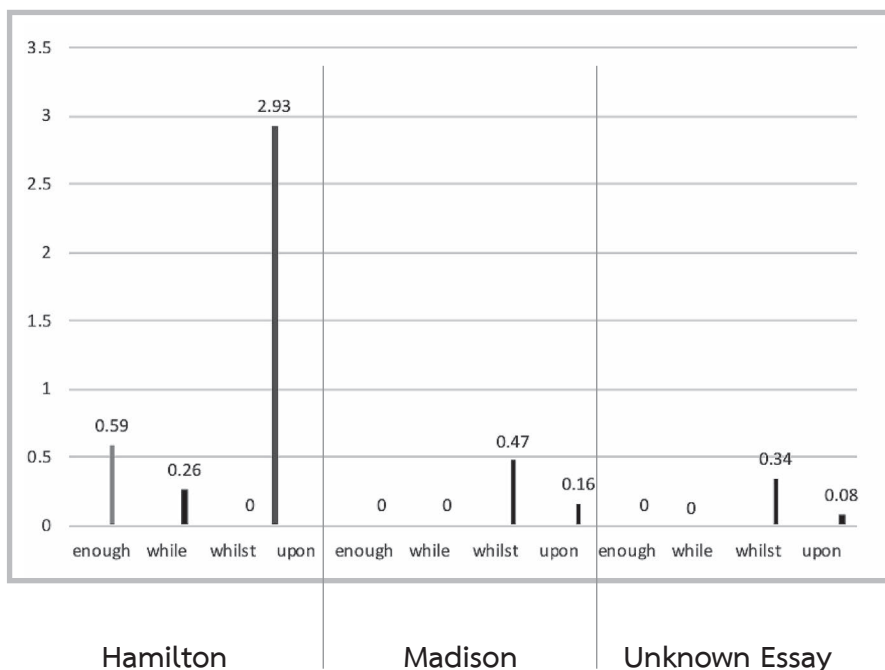
¹⁷ เอกสารเฟเดอรัลลิสต์ เป็น การรวบรวมบทความ 85 ชิ้นที่อเล็กซานเดอร์ แฮมิลตัน, เจมส์ เมดิสัน และจอห์น เจย์ ร่วมผลิตกันเขียนลงในหนังสือพิมพ์นิวยอร์ก เมื่อปี 1787-1788 บทความแต่ละชิ้นมีเนื้อหาเกี่ยวกับรัฐธรรมนูญฉบับใหม่สหรัฐอเมริกา เพื่อโน้มน้าวให้ชาวนิวยอร์กลงมติผ่านร่างรัฐธรรมนูญที่ทั้งสามเป็นผู้ร่วมร่าง โดยในเวลานั้นสาธารณะชนไม่มีใครรู้ว่าบทความแต่ละชิ้นถูกเขียนขึ้นโดยใคร เพราะทั้งสามต่างใช้นามปากกาเดียวกันว่า “Publius” ต่อมาในปี 1804 อเล็กซานเดอร์แฮมิลตัน รัฐมนตรีว่าการกระทรวงการคลังคนแรกของประเทศสหรัฐอเมริกาเขียนจดหมายหาเพื่อนของเขาชื่อ Egbert Benson พร้อมระบุผู้เขียนที่แท้จริงในแต่ละบทความทั้ง 85 ชิ้น ในปี 1817 ภายหลังที่หมดวาระจากการดำรงตำแหน่งประธานาธิบดี สหรัฐอเมริกาได้ไม่นาน เจมส์ เมดิสัน ก็ออกมาเปิดเผยรายชื่อผู้เขียนที่แท้จริงในแต่ละบทความเช่นเดียวกับที่แฮมิลตันทำไว้เมื่อ 13 ปีก่อน อย่างไรก็ตาม เมื่อเอารายชื่อของเมดิสันไปเทียบกับของแฮมิลตันปรากฏว่ามีอยู่ 12 บทความที่ทั้งคู่อ้างว่าตนเองเป็นผู้เขียน

1. เมดิสันใช้คำว่า *also* บ่อยมากกว่าแฮมิลตันประมาณสองเท่า
2. แฮมิลตันใช้คำว่า *according* ถี่มากกว่าเมดิสัน
3. ในบทความที่ถูกระบุชัดว่าเมดิสันเป็นผู้เขียน มีคำว่า *whilst* ในบทความกว่าครึ่งหนึ่ง ในขณะที่เดียวกัน ทุกบทความที่ถูกระบุชัดว่าเมดิสันเขียนไม่มีคำว่า *while* อยู่เลย
4. บทความที่ถูกระบุชัดว่าแฮมิลตันเป็นผู้เขียน ไม่มีคำว่า *whilst* อยู่เลย แต่ในขณะเดียวกัน กลับมีคำว่า *while* อยู่ถึงหนึ่งในสาม (Mosteller and Wallace, 1963 : 275-309)

ด้วยวิธีการหาคำตอบโดยใช้หลักการพื้นฐานทางสถิติด้วยการนับความถี่ของคำต่างๆ ที่ทั้งเมดิสันและแฮมิลตันใช้เพื่อค้นหารูปแบบการเขียนของทั้งคู่แล้วนำไปเปรียบเทียบกับเอกสารเฟเดอรัลลิสต์ 12 ดังกล่าว Mosteller และ Wallace พบว่าเอกสารปริศนา 12 ชิ้นนั้นมีรูปแบบการเขียนและความถี่ของคำที่ถูกเลือกใช้มีความคล้ายคลึงกับลีลาการเขียนของ เจมส์ เมดิสัน และในปัจจุบัน อีก 50 ปีต่อมา ทั้งนักวิชาการด้านประวัติศาสตร์ นักรัฐศาสตร์รวมถึงนักสถิติที่ทำงานวิจัยเพื่อค้นหาผู้เขียนเอกสารปริศนา 12 นี้เพื่อพิสูจน์ ตรวจสอบ ผลการศึกษาของ Mosteller และ Wallace ต่างได้ข้อสรุปเดียวกันว่า เจมส์ เมดิสัน คือผู้เขียนเอกสาร 12 ชิ้นนั้น เช่นเดียวกับงานของ Reynold and Saxonhouse (1995) ที่ใช้วิธีวิทยาเชิงปริมาณเพื่อไขปัญหาความกำกวมเรื่องผู้แต่งด้วยการวิเคราะห์การเลือกใช้ถ้อยคำที่นักเขียนแต่ละคนชอบใช้ซึ่งเป็นการบ่งบอกถึงความเป็นตัวตนของนักเขียนแต่ละคน ซึ่งจากผลการศึกษาพบว่าโรนัลด์ ฮีออบส์ (1588-1679) นักปรัชญาการเมือง ชาวอังกฤษ เป็นผู้เขียนบทความ 3 เรื่องที่ไม่ได้ถูกระบุผู้แต่งใน Horae Subsecivae (1620)

รูปภาพที่ 1.1

ภาพ : เปรียบเทียบ ระหว่าง แฮมิลตันกับเมดิสันและ 12 unknown essay



ที่มา : จาก Blatt (2017 : 60)

ปัจจุบัน ความก้าวหน้าของเทคโนโลยีดิจิทัลและคอมพิวเตอร์ทำให้คนทั่วไปสามารถเข้าถึงเอกสารและหนังสือจำนวนมากที่ถูกทำให้อยู่ในรูปแบบอิเล็กทรอนิกส์ (Digital Forms) และด้วยแนวทางการศึกษาแบบ DH ทำให้ผู้ศึกษาสามารถดึงเอกสารเพเดิร์ลลิสต์ในรูปแบบดิจิทัลฟอร์มออกมาจากอินเทอร์เน็ตหรือในคลาวด์ (Cloud) ที่ไหนสักแห่งในโลกใบนี้ ไม่ว่าจะเป็นรูปแบบ PDF (Portable Document Format) หรือ Microsoft Word Document หลังจากนั้นก็นำไปเปิดในโปรแกรม Microsoft Word แล้วเลือกใช้เมนูในฟังก์ชัน “ค้นหา”

คำว่า also, according, whilst, while แทนที่จะเป็นการนับความถี่ของแต่ละคำด้วยแรงงานของมนุษย์เหมือนกับที่ Mosteller และ Wallace หรือ Busa เคยทำในยุคสมัยของพวกเขา แต่ด้วยการประมวลผลแบบดิจิทัลซึ่งใช้เวลา ภายในเสี้ยววินาที คอมพิวเตอร์ก็จะแสดงความถี่ของคำดังกล่าวออกมาให้เห็นเพื่อให้ผู้ศึกษาวิจัยได้ทำการค้นหาความหมายและตีความต่อไป

ในปีเดียวกันกับที่ Mosteller และ Wallace ตีพิมพ์ผลงานที่ถือว่าเป็นงานวิจัยชิ้นแรกๆ ในวงวิชาการด้าน DH มหาวิทยาลัยเคมบริดจ์ได้ก่อตั้ง Centre for Literary and Linguistic Computing และได้พัฒนากลายเป็นสมาคมคอมพิวเตอร์และมนุษยศาสตร์ หรือ Association for Computer and the Humanities (ในปี 1970) ต่อมาปี 1964 บริษัท IBM ได้จัดสัมมนาเกี่ยวกับเรื่อง Humanities Computing เป็นครั้งแรกที่เมือง Yorktown Heights รัฐนิวยอร์ก สหรัฐอเมริกา ภายใต้หัวข้อ *Literary Data Processing Conference Proceedings* และปี 1966 ได้มีการจัดพิมพ์วารสารชื่อ *Computer and the Humanities* และต่อมาเมื่อเทคโนโลยีดิจิทัลและคอมพิวเตอร์ได้เจริญรุดหน้าขึ้นจึงทำให้เริ่มมีการนำแนวคิดเรื่อง DH มาใช้ในวงวิชาการอย่างจริงจังตั้งแต่ทศวรรษที่ 1990s พร้อมกับปรับเปลี่ยนชื่อใหม่จาก Humanities Computing เป็น Digital Humanities (Gold, 2012 : 5,87; Schreibman et al. 2004 : xxiii : 3-17)

ในทศวรรษที่ 90 อันเป็นยุคที่อินเทอร์เน็ตเว็บไซต์ถือกำเนิดขึ้นประกอบกับขีดความสามารถที่เพิ่มขึ้นของเครื่องคอมพิวเตอร์ส่วนบุคคลหรือ PC (Personal Computer) ทำให้การศึกษา ค้นคว้า วิจัยด้าน DH หรือการบูรณาการระหว่างวิทยาการคอมพิวเตอร์เข้ากับมนุษยศาสตร์ได้รับอานิสงค์เป็นอย่างมาก โดยในระยะแรกๆ นักมนุษยศาสตร์จะเริ่มต้นจากการใช้ประโยชน์ความสามารถในการจัดเก็บข้อมูลจำนวนมหาศาลของระบบคอมพิวเตอร์เข้ามาจัดเก็บผลงานของนักคิดนักเขียน ศิลปิน กวี จิตรกรคนสำคัญๆ ในรูปแบบอิเล็กทรอนิกส์ (Digital

Forms) เช่น โครงการ Transcribe Bentham ของ University College London โครงการ The Rossetti Archive หรือภายใต้การสนับสนุนของสถาบันเพื่อเทคโนโลยีขั้นสูงเพื่อมนุษยศาสตร์มหาวิทยาลัยเวอร์จิเนียและ The Walt Whitman Archive โดยความร่วมมือของมหาวิทยาลัยเนบราสก้ากับมหาวิทยาลัยไอโอวา เป็นต้น การเกิดขึ้นของแขนงวิชา DH ยังนำไปสู่การจัดตั้งศูนย์ DH ขึ้นตามมาในระยะ 10 ปีที่ผ่านมาทั้งในหน่วยงานภาครัฐ เอกชน องค์กรอิสระ สถาบันความคิด (Think Tank) และสถาบันการศึกษาต่างๆ มากกว่า 100 ศูนย์ในกว่า 25 ประเทศทั่วโลก

ปัจจุบันหน่วยงานหรือศูนย์วิจัยด้านมนุษยศาสตร์ดิจิทัลเหล่านี้ต่างได้รับทุนวิจัยทางด้านนี้กันอย่างมากมาย เช่น ในปี 2007 สำนักงานมนุษยศาสตร์ดิจิทัล (Office of Digital Humanities) ของสหรัฐอเมริกาได้รับเงินทุนสนับสนุนการวิจัยระยะเวลา 5 ปี จากกองทุนแห่งชาติทางมนุษยศาสตร์ จำนวน 15,268,130 ดอลลาร์ใน 250 โครงการวิจัยทางมนุษยศาสตร์ดิจิทัล เช่นเดียวกับ สภาวิจัยทางมนุษยศาสตร์และศิลปะ สหราชอาณาจักรให้เงินอุดหนุนงานวิจัยทาง DH จำนวน 121.5 ล้านปอนด์ ใน 330 โครงการวิจัย ระหว่างปี ค.ศ. 1998-2008 รวมทั้ง มูลนิธิแอนดริว เมลอน อันเป็นองค์กรเอกชนที่ไม่แสวงหากำไรก็ได้ให้เงินจำนวนกว่า 30 ล้านดอลลาร์ ในปี 2010 เพื่อการวิจัยทางด้านมนุษยศาสตร์ดิจิทัล¹⁸

¹⁸ <http://www.ucl.ac.uk/infostudies/melissa-terras/DigitalHumanitiesInfographic.pdf>
[Retrieved: June 13, 2018]

● ข้อมูลขนาดใหญ่ เหมือนข้อมูลและมนุษยศาสตร์ดิจิทัล

ทุกวันนี้เทคโนโลยีทางคอมพิวเตอร์และดิจิทัลได้ช่วยให้สามารถจัดเก็บข้อมูลขนาดใหญ่¹⁹ (Big Data) ไว้ในรูปแบบอิเล็กทรอนิกส์เทคโนโลยีดังกล่าวยังช่วยให้สามารถตรวจสอบ ติดตาม ค้นหาข้อมูลจำนวนมากมหาศาลทั้งที่เป็นเอกสาร หนังสือ รูปภาพ เสียง ภาพยนตร์และอื่นๆ ที่อยู่ในรูปแบบดิจิทัลได้อย่างมีประสิทธิภาพและสะดวกรวดเร็วยิ่งกว่าที่เคยเป็นมาก่อน แต่เดิมการจัดเก็บข้อมูลขนาดมหึมาจะทำเพื่อตอบโจทย์ในวงการธุรกิจเพื่อใช้ประกอบการตัดสินใจของผู้บริหารต่อมาระบบหรือฐานข้อมูลที่ถูกออกแบบโดยเฉพาะเพื่อ

¹⁹ Big Data คือข้อมูลที่มีปริมาณระดับ Terabyte ขึ้นไป 1 Terabytes มีค่าเท่ากับ 1,024 Gigabytes ข้อมูลเหล่านี้จะอยู่ในรูปแบบที่หลายหลายในรูปแบบอิเล็กทรอนิกส์ ไม่ใช่เฉพาะที่เป็นตัวหนังสือหรือสมการตัวเลขทางคณิตศาสตร์อย่างเดียว แต่ยังรวมถึงข้อมูลที่ไม่เป็นโครงสร้าง (unstructured data) เช่น ข้อความ ภาพ เสียง วิดีโอ นอกจากนี้ยังเปลี่ยนแปลงไปอย่างรวดเร็วอีกด้วย เช่น ข้อมูลบนโลกออนไลน์ ข้อมูลการใช้สมาร์ทโฟน

แหล่งข้อมูลสี่เสาหลักที่หลอมรวมกันเป็น Big Data ก็คือ 1) ข้อมูลแบบเปิดทั่วไป (Open Data) เป็นข้อมูลที่ถูกจัดทำขึ้นทั้งจากหน่วยงานราชการและองค์กรเอกชนต่างๆ เพื่อให้สาธารณะรับทราบและสามารถนำไปใช้ประโยชน์เชิงการวิเคราะห์ต่างๆ โดยเน้นการเข้าถึงข้อมูลได้โดยง่ายแบบออนไลน์ สามารถนำข้อมูลไปใช้ได้อย่างทันที 2) ข้อมูลจากระบบอัตโนมัติ (Automated Data) อันมาจากโครงข่ายโทรคมนาคมที่มีความครอบคลุม การเชื่อมต่อกับเครือข่ายการสื่อสารไร้สาย เช่น ระบบโทรศัพท์เคลื่อนที่ อินเทอร์เน็ตกับระบบอุปกรณ์เครื่องตรวจวัดอัตโนมัติ (automation sensor) หรืออุปกรณ์ประเภทที่เรียกว่า Internet of Things (IoT) ซึ่งส่งข้อมูลแบบ real time เช่น กล้องโทรทัศน์วงจรปิด (CCTV) การวัดปริมาณการจราจร การตรวจสอบความหนาแน่นหรือความเร็วของระบบขนส่งสาธารณะ 3) ข้อมูลจากพลเมือง (Crowdsourcing Data) เป็นข้อมูลจากระบบอัตโนมัติและ IoT คล้ายกับข้อ 2) แต่ได้มาจากพลเมืองที่ใช้อุปกรณ์ที่เป็น smart device เช่น smart phone หรือ tablet และ 4) ข้อมูลจากสื่อสังคมออนไลน์ (Social Data) อันได้แก่ข้อมูลที่มาจาก Facebook, Twitter, Instagram, Snapchat หรือระบบ (platform) ของสื่อสังคมออนไลน์อื่น ๆ ซึ่งผู้ใช้มันจะโพสต์เรื่องที่พบเจอในชีวิตประจำวันลงไป, ดูรายละเอียดเพิ่มเติมใน Zikopoulos et al. 2012, Understanding Big Data : Analytics for Enterprise Class Hadoop and Streaming Data, (McGraw-Hill, 2012)

ใช้ในการจัดเก็บข้อมูลขนาดใหญ่หรือที่เรียกว่า คลังข้อมูล (Data Warehouse) ได้ถูกนำไปใช้ในด้านต่างๆ เช่น ด้านภูมิศาสตร์ อุตุนิยมวิทยา วิศวกรรมวิทยาศาสตร์ การแพทย์ การเกษตร การเงิน เป็นต้น อย่างไรก็ตามการจัดการกับข้อมูลจำนวนมากมหาศาลเพื่อทำให้เป็นสารสนเทศที่มีประโยชน์และมูลค่าต้องอาศัยเทคนิควิธีการที่เรียกว่า การทำเหมืองข้อมูล (Data Mining)

การทำเหมืองข้อมูลเป็นกระบวนการสกัดความรู้ที่น่าสนใจจากข้อมูลปริมาณมาก ความรู้ที่ไม่เห็นเด่นชัด ความรู้ที่บ่งบอกเป็นนัย ความรู้ที่ไม่ทราบมาก่อน ที่มีศักยภาพในการนำไปใช้ประโยชน์ การทำเหมืองข้อมูลยังมีความหมายรวมถึงกระบวนการหาความหมายของรูปแบบ ความสัมพันธ์ของข้อมูลที่ซ่อนอยู่หรือความรู้ใหม่ๆ จากฐานข้อมูลขนาดใหญ่เพื่อนำมาใช้ในการตัดสินใจ²⁰ (Ye, 2014 : xiii) ในปัจจุบันการทำเหมืองข้อมูลได้เริ่มถูกนำไปประยุกต์ใช้ในวงวิชาการด้านสังคมศาสตร์²¹ และมนุษยศาสตร์มากขึ้นอันนำไปสู่การศึกษาวิจัยเชิงสหวิทยาการหรือการศึกษาทางมนุษยศาสตร์ดิจิทัลดังที่กล่าวมาในตอนต้น

ปัจจุบันมีการวิจัยเชิงสหวิทยาการที่ทำงานกับข้อมูลขนาดใหญ่ด้วยการทำเหมืองข้อมูลผ่านการใช้เครื่องมือทางมนุษยศาสตร์ดิจิทัลและ/หรือรูปแบบการประมวลผลทางสังคมศาสตร์ เกิดขึ้นมากมายเช่น โครงการวิจัยที่ห้องทดลองการเล่าเรื่องด้วยตัวเลข (Computational Story Lab) มหาวิทยาลัยเวอร์มอนต์ สหรัฐอเมริกาที่เก็บข้อมูล Digital forms จากโซเชียลเน็ตเวิร์กที่คนอเมริกันใช้กัน

²⁰ นอกจากคำว่า Data mining คำที่ถูกใช้แทนในความหมายที่ใกล้เคียงกันก็ได้แก่ knowledge discovery in databases, knowledge mining from databases, knowledge extraction, knowledge-based system, data archaeology, data dredging, data analysis เป็นต้น

²¹ การนำเทคโนโลยีและรูปแบบการประมวลผลแบบเดียวกันไปใช้การศึกษาด้านสังคมศาสตร์จะถูกเรียกว่า Computational Social Science ดูเพิ่มเติมใน Hanna Wallach. "Computational Social Science: Toward Collaborative Future." *Data Science for Politics, Policy, and Government*. Cambridge University Press, 2015.

ในแต่ละวัน เป็น Big Data โดยมีวัตถุประสงค์เพื่อหาความหมายของรูปแบบความสัมพันธ์ของข้อมูลที่ซ่อนอยู่หรือความรู้ใหม่ๆ จากกองข้อมูลขนาดใหญ่เหล่านี้ผ่านการตีความ

ห้องทดลอง Computational Story เป็นแหล่งรวมนักวิจัยจากหลากหลายสาขา เช่น ฟิสิกส์เชิงทฤษฎี คณิตศาสตร์ วิศวกรรมอิเล็กทรอนิกส์ วิทยาการคอมพิวเตอร์ สังคมศาสตร์ มนุษยศาสตร์และอีกหลายๆ สาขามาทำงานร่วมกันเพื่อศึกษาวิจัยสังคมในยุคดิจิทัลหรือสังคมในศตวรรษที่ 21 ที่ถูกขับเคลื่อนโดยเทคโนโลยีคอมพิวเตอร์และดิจิทัลขั้นสูง (Socio-technical System)

เช่น การศึกษาเรื่อง *"Measuring the Happiness of Large-Written Expression: Songs, Blogs and President"* (Dodds and Danforth, 2009) ใช้วิธีการศึกษาแบบ Data Mining เป็นเครื่องมือในทางวิธีวิทยาเพื่อวัดหา "ระดับความสุข" ของมนุษย์ด้วยการศึกษาจากชุดข้อมูล (Data Set) 4 ชุด คือ

1. เนื้อเพลงจำนวน 232,574 เพลง จากศิลปิน 20,025 คน ระหว่างปี 1960-2007 ด้วยการดึงข้อมูล (download) จาก เว็บไซต์ <http://www.hotlyrics.net>
2. ชื่อเพลงดึงข้อมูลจาก เว็บไซต์ <http://www.freedb.org>
3. ข้อความในบล็อก (blog) ที่ถูกเขียนโดยคนๆ แรกในบล็อกและมีคำว่า "รู้สึก" (Feel) ดึงข้อมูลจากเว็บไซต์ <http://www.wefeelfine.org> ด้วยวิธี API²² จากวันที่ 25 สิงหาคม 2005 ถึง 30 มิถุนายน 2009 จำนวน 9,113,772 ประโยค จาก 2.3 ล้านบล็อก

²² API หรือ Application Programming Interface เป็นช่องทางการเชื่อมต่อระหว่างเว็บไซต์หนึ่งไปยังอีกเว็บไซต์หนึ่งหรือเป็นการเชื่อมต่อระหว่างผู้ใช้งานกับเครื่อง Server หรือระหว่างเครื่อง Server กับ Server หรือกล่าวอย่างง่าย ๆ ก็คือ API เปรียบเสมือนภาษาคอมพิวเตอร์ที่ทำให้เครื่องแต่ละเครื่องสามารถสื่อสารแลกเปลี่ยนข้อมูลกันได้ (ผู้เขียน)

4. คำแถลงประจำปีของประธานาธิบดี สหรัฐอเมริกา (State of the Union addresses) 43 คน ตั้งแต่จอร์จ วอชิงตัน ถึง จอร์จ ดับเบิลยู บุช ดึงข้อมูลจากเว็บไซต์ the American Presidency Project ที่ <http://www.presidency.ucsb.edu> และ เว็บไซต์ของ British National Corpus ที่ <http://www.natcorp.ox.ac.uk>

Dodds and Danforth (2009 : 2-4) เริ่มต้นด้วยการใช้งานวิจัยของ Bradley and Lang (1999) ที่ศึกษาการวัดบรรทัดฐานทางอารมณ์จากคำในภาษาอังกฤษ (The Affective Norms for English Words: ANEW) ที่เลือกมาจากทั้ง คำนาม คำกริยาและคำคุณศัพท์จำนวนทั้งหมด 1,034 คำ ด้วยการให้กลุ่มตัวอย่าง ให้คะแนนกับคำแต่ละคำแล้วตีเป็นคะแนนออกมาด้วยวิธีทางสถิติในสเกล 1-9 เป็นมาตรวัดความสุข Bradley and Lang (1999 : 4-45) พบว่า คำที่มีระดับคะแนนความสุขมากที่สุดห้าคำแรกคือ triumphant (8.82), paradise (8.72), love (8.72), luxury (7.88), trophy (7.78) และ glory (7.55) ส่วนคำที่มีระดับความทุกข์มากที่สุดห้าคำแรกคือ suicide (1.25), rape (1.25), funeral (1.39), trauma (2.10), hate (2.12), และ hostage (2.20)

Dodds and Danforth (2009: 2) ใช้ผลการศึกษาของ Bradley and Lang (1999) มาเป็นเกณฑ์ในการใช้จัดอันดับของระดับคำที่บ่งบอกถึงความสุข และคำที่แสดงออกถึงความสุข ด้วยการนำคำเหล่านี้ไปเทียบความถี่ที่กลุ่มคำใน ANEW ไปปรากฏอยู่ในชุดข้อมูล 4 ชุดข้างต้นในตารางที่ 1.1

ตารางที่ 1.1

จำนวนคำทั้งหมดในแต่ละชุดของข้อมูลเทียบกับจำนวนคำ
ใน ANEW ที่ปรากฏอยู่ในแต่ละชุดของข้อมูล

ชุดข้อมูล	เนื้อเพลง (คำ)	ชื่อเพลง (คำ)	บล็อก	คำแถลงประจำปี
จำนวนคำ	58,610,849	60,867,223	157,853,709	1,796,763
จำนวนคำ ใน ANEW	3,477,575	5,612,708	8,697,633	61,926

ที่มา : Dodds and Danforth (2009 : 5)

ยกตัวอย่าง ในเนื้อหาจากทวีตเตอร์ของบุคคลสองคนสามารถบ่งบอกถึงระดับความสุขได้แตกต่างกัน เมื่อคนหนึ่งกำลังมีความรู้สึกยินดี ก็อาจทวีตคำที่แสดงออกถึงความสุขออกมา เช่น “I love you” กับอีกคนหนึ่งที่อยู่ในห้วงภาวะที่ย่ำแย่ ก็อาจจะทวีตคำที่หม่นหมองออกมา เช่น “I hate you” ตัวอักษรสีแดงที่ขีดเส้นใต้เป็นคำที่อยู่ในบัญชีรายชื่อ ANEW คำว่า “love” มีค่าระดับความสุขที่ 8.74 ส่วน “hate” มีค่าระดับความสุขเพียงแค่ 2.1 ดังนั้นเนื้อหาที่บุคคลสองคนนี้ทวีตออกมาก็จะบ่งบอกถึงระดับความสุขหรือทุกข์ที่แตกต่างกัน

ด้วยหลักการเดียวกัน เพลง Billy Jean ของ Michael Jackson มีเนื้อร้องท่อนหนึ่งว่า “....She was more like a beauty queen from a movie scene...” ตัวอักษรสีแดงที่ขีดเส้นใต้เป็นคำที่อยู่ในบัญชีรายชื่อ ANEW และมีค่าระดับความสุขที่ 7.82, 6.44 และ 6.86 ตามลำดับ ดังนั้นระดับความสุขหรือ(ทุกข์) ของเพลง Billy Jean (เฉพาะท่อนนี้) ก็คือ $\frac{1}{3} * (7.82 + 6.44 + 6.86) = 7.04$

ในภาพใหญ่ เมื่อนำการใช้อันดับของคำที่มีความสุขไปจนถึงคำที่มีความทุกข์จากผลการศึกษาของ Bradley and Lang (1999) ไปจับคู่กับ Big Data 4 ชุด ก็พบรูปแบบที่ซ่อนอยู่ในกองข้อมูลขนาดใหญ่เหล่านี้ ในตาราง 1.2 และสถิติ

ที่น่าสนใจอีกมากมายเมื่อนำคำที่อยู่ใน ANEW List ไปจับคู่กับประโยคแรกในบล็อกที่มีคำว่า “feel” ที่ถูกหีตออกมาระหว่าง สิงหาคม 2005 ถึง มิถุนายน 2009พบว่า

- วันที่คนมีความสุขมากที่สุดในช่วงปี 2005-2009 คือวันที่โอบามาได้รับเลือกตั้ง (4 พฤศจิกายน 2008)
- วันที่คนมีความสุขที่สุดในแต่ละปีคือวันวาเลนไทน์และวันคริสต์มาส
- วันที่มีความสุขที่สุดในสัปดาห์คือวันอาทิตย์ ส่วนวันพุธเป็นวันที่ทุกข์ที่สุด
- วันที่มีความสุขทุกข์มากที่สุดคือวันที่ 11 กันยายนของแต่ละปีและวันที่ไมเคิล แจ็คสันเสียชีวิต (25 มิถุนายน 2009)

ตารางที่ 1.2

อันดับห้าคำแรกใน ANEW ที่ปรากฏขึ้นบ่อยที่สุดในแต่ละชุดของมูล

ลำดับ	เนื้อหาเพลง	ชื่อเพลง	บล็อก	คำแถลงประจำปี
1	Love (7.37%)	Love (7.39%)	Good (4.89%)	People (5.49%)
2	Time (4.18%)	Time (4.19%)	Time (4.72%)	Time (4.09%)
3	Baby (2.75%)	Baby (2.75%)	People (3.34%)	Present (3.45%)
4	Life (2.59%)	Life (2.60%)	Love (3.31%)	World (3.10%)
5	Heart (2.14%)	Heart (2.15%)	Life (3.13%)	War (2.98%)

ที่มา : Dodds and Danforth (2009 : 5)

ผลการศึกษาจากงานวิจัยชิ้นนี้ถูกนำไปพัฒนาต่อยอดเป็นตัวแบบในการสร้าง *มาตรวัดความสุข*(Hedonometer) อันเป็นกราฟแสดงความสุขด้วยการวัดระดับความสุขผ่านวิธีอ่านและตีความผ่านข้อมูลบนโซเชียลมีเดีย-เน็ตเวิร์กต่างๆ ของปัจเจกบุคคลด้วยการทำตำแหน่งของเหตุการณ์สำคัญๆ ที่มีผลต่อระดับความสุขหรือความทุกข์ของกลุ่มคนผู้ใช้โซเชียลมีเดียผ่านแพลตฟอร์มต่างๆ ลงบนกราฟ เช่น วันที่เครื่องบินมาเลเซียแอร์ไลน์ ไฟล์ 370 บินหายไปในมหาสมุทร วันที่โรบิน วิลเลียม ดาราตลกชื่อดังของฮอลลีวูด เสียชีวิต วันวาเลนไทน์ หรือ วันคริสต์มาส เป็นต้น ด้วยข้อมูลชุดเดียวกันนี้ คณะนักวิจัยจากห้องทดลองการเล่าเรื่องด้วยตัวเลขยังขยายผลการวัดระดับความสุขและความทุกข์ของมนุษย์จากการดูภาพยนตร์เรื่องต่างๆ พร้อมกับทำไทม์ไลน์ของความสุขและความทุกข์ตามการลำดับเรื่องและเหตุการณ์ในภาพยนตร์ และจากชุดของมูล (data set) เดียวกันนี้ โครงการวิจัยต่อไปของพวกเขาก็คือการใช้ข้อมูลจากโซเชียลมีเดีย-เน็ตเวิร์กทั้งหลายทำนายการเกิดวิกฤตเศรษฐกิจก่อนที่มันจะเกิดขึ้นจริงๆ หรือทำการแยกแยะข่าวปลอม (fake news) ออกจากข่าวจริง ในแพลตฟอร์มต่างๆ (Dodds and Danforth, 2009; Dodds et al. 2011; Klouman et al 2012; Bliss et al 2012; Mitchell et al 2013; Frank et al 2013)

โครงการ Google Book ก็เป็นอีกโครงการที่ทำงานกับข้อมูลขนาดใหญ่ด้วยการทำเหมืองข้อมูล ผ่านวิธีการศึกษาแบบมนุษยศาสตร์ดิจิทัลเพื่อการผลิตหรือสร้างองค์ความรู้ใหม่ขึ้นมาจากการทำวิจัยเชิงสหวิทยาการดังกล่าว บริษัท Google เปิดตัวโครงการดังกล่าวด้วยการสแกน (scan) หนังสือจำนวน 12 ล้านเล่ม มากกว่า 400 ภาษาเป็นจำนวนมากกว่า 5 พันล้านหน้าหรือ 2 ล้านล้านคำให้อยู่ในรูปแบบดิจิทัลฟอร์ม แล้วเก็บไว้ในระบบฐานข้อมูลของบริษัท ต่อจากนั้นในปี 2010 Google ให้ทุนวิจัยเรื่องมนุษยศาสตร์ดิจิทัล หรือ Google’s Digital Humanities Research Award จำนวน 1 ล้านดอลลาร์กับนักวิจัย 23 คนจาก

25 มหาวิทยาลัย ใน 12 โครงการวิจัยทางมนุษยศาสตร์ดิจิทัลโดยอนุญาตให้นักวิจัยทั้งหมดสามารถเข้าถึงระบบฐานข้อมูลและระบบคอมพิวเตอร์ของ Google เพื่อช่วยในการประมวลผลงานวิจัยดังกล่าวได้²³

เช่น งานวิจัยของ Daniel Cohen กับ Frederick Gibbs จาก George Mason University สหรัฐอเมริกา เรื่อง “Reframing the Victorian” อันเป็นส่วนหนึ่งใน 12 โครงการวิจัยที่ได้ Google’s Digital Humanities Award ได้ใช้วิธีแนวทางการศึกษาแบบ DH สร้างแผนภูมิ (Chart) แสดงความถี่การปรากฏขึ้นของคำทุกคำซึ่งเป็นส่วนหนึ่งของชื่อหนังสือจำนวน 1,681,161 เล่ม ที่ถูกพิมพ์เป็นภาษาอังกฤษ ในประเทศอังกฤษ ช่วงยุคสมัยวิคตอเรียน ระหว่างปี ค.ศ. 1789-1914 เพื่อช่วยให้นักวิคตอเรียนศึกษา (Victorian Studies) เข้าใจภูมิปัญญาและความคิดในยุคสมัยนั้นได้อย่างถ่องแท้ยิ่งขึ้น จากแผนภูมิดังกล่าว Cohen and Gibbs พบรูปแบบและความหมายที่ซ่อนอยู่ในข้อมูล ขนาดเกือบ 1.7 ล้านรายชื่อว่าความหมายของคำนั้นเปลี่ยนแปลงไปตลอดเวลา เช่นเดียวกับการเลือกใช้คำต่างๆ ที่เอามาตั้งเป็นชื่อเรื่อง เช่นคำว่า “Science” ในระยะแรกๆ จะถูกใช้ในความหมายโน้มเอียงว่า *ความรู้* และคำๆ นี้จะถูกใช้มากที่สุดในปี 1874 อันเป็นช่วงปีภายหลังของการปฏิวัติอุตสาหกรรม อันแสดงให้เห็นความสนใจของยุคสมัยในเรื่อง “Science” แต่ในขณะเดียวกัน ความหมายที่ถูกใช้ก็ถูกเปลี่ยนให้แคบลงไปจากความหมายว่า *ความรู้* ว่าหมายถึงเฉพาะ “วิทยาศาสตร์ธรรมชาติ” (Natural Science) อันสอดคล้องกับคำว่า “Industrial” ที่เผยแพร่ครั้งแรกขึ้นบนบรรณพิภพการเป็นส่วนหนึ่งของชื่อเรื่องในโลกหนังสือเมื่อปี 1832 และรักษาความถี่ของการถูกใช้เป็นส่วนหนึ่งของชื่อเรื่องหนังสืออย่างเสมอดันเสมอไปจนสิ้นสุดสมัยวิคตอเรียน ในขณะที่คำว่า “Evil” แทบจะไม่เคยถูกใช้เอามาตั้งเป็นชื่อหนังสือเลย (Cohen and Gibbs , 2011)

²³ <https://googleblog.blogspot.com/2010/07/our-commitment-to-digital-humanities.html> [Retrieved: March 8, 2018]

อย่างไรก็ตาม งานศึกษาเรื่องหนังสือสมัยวิคตอเรียนของ Cohen กับ Gibbs ขึ้นต่อไปหรือเฟสสองกำลังอยู่ระหว่างการดำเนินการ งานเฟสที่สองนี้เป็นการอินพุตส์เนื้อหาของหนังสือในสมัยวิคตอเรียนทั้งหมดใส่เข้าไปให้คอมพิวเตอร์ประมวลผลเพื่อสร้างแผนภูมิแสดงรูปแบบความถี่ของคำต่างๆ ทั้งหมด ซึ่งแผนภูมิของความถี่ของคำที่ได้ย่อมแตกต่างจากงานชิ้นแรกอย่างแน่นอนไม่ต้องสงสัยเพราะเป็นการอินพุตส์เฉพาะแต่ชื่อหนังสือ นอกจากนี้ Cohen กับ Gibbs ยังจะเลือกใช้เฉพาะบริบทที่อยู่รายล้อมอยู่รอบๆ คำแต่ละคำเพื่อที่จะเป็นกรอบในการวิเคราะห์ให้ละเอียดลึกลงไปอีกว่า *คำแต่ละคำ* ถูกใช้ในความหมายเชิงบวกหรือเชิงลบในแต่ละบริบทๆ หนึ่ง เป้าหมายของงานในเฟสที่สองนี้ คือการหาความสัมพันธ์เชื่อมโยงระหว่างชุดของข้อมูล 3 ชุดคือ ผลการศึกษาในเฟสหนึ่ง เนื้อหาหนังสือที่ถูกอินพุตเข้าไปในเฟสสองและบริบททางประวัติศาสตร์ของข้อมูลชุดหนึ่งและสอง

การนำคอมพิวเตอร์มาใช้ประมวลผลหรือช่วยในการอ่านหนังสือกองโตจำนวนเกือบ 1.7 ล้านเล่มนั้นไม่ได้หมายความว่าเครื่องจักรสมองกลเหล่านี้จะเข้ามาแทนที่การอ่านแบบละเอียดละไม (Closed Reading) อันเป็นการอ่านแบบขบเคี้ยวเนื้อหาเพื่อซาบซึ้งดื่มด่ำกับสุนทรียะของความหมายที่นักมนุษยศาสตร์นิยมกระทำกันมาแต่ดั้งเดิม แต่การอ่านหนังสือโดยเครื่องคอมพิวเตอร์จะเป็นการเสนอวิธีการอ่านอีกรูปแบบหนึ่ง เป็นการอ่านตัวบทองมหิมาที่อาจจะนำไปสู่การตั้งคำถามใหม่ๆ ในการอ่านตัวบทจำนวนมากที่มนุษย์ไม่สามารถอ่านได้หมดในช่วงชีวิตหนึ่งๆ (Cohen and Gibbs , 2011)



● unสรุป

มนุษยศาสตร์ดิจิทัลเป็นการบูรณาการความรู้ทางด้านวิทยาการคอมพิวเตอร์เข้ากับการศึกษาทางด้านมนุษยศาสตร์ เป็นเรื่องเกี่ยวกับการใช้เทคโนโลยีคอมพิวเตอร์และรูปแบบการคำนวณเชิงดิจิทัลเข้าไปช่วยในประมวลผลข้อมูลทางมนุษยศาสตร์ที่ปัจจุบันสามารถทำให้อยู่ในรูปแบบอิเล็กทรอนิกส์ได้เกือบทั้งหมด อย่างไรก็ตาม ในทัศนะผู้เขียน มนุษยศาสตร์ดิจิทัลไม่ใช่การพยายามที่จะกลับไปทำให้การศึกษามนุษยศาสตร์มีลักษณะเป็นปรนัยอีกครั้ง เหมือนเมื่อครั้งหนึ่งที่แนวคิดปรัชญาแบบปฏิฐานนิยม (positivism) ได้มีอิทธิพลครอบงำสังคมวิทยาว่าด้วยความรู้ในช่วงปลายศตวรรษที่ 19 ถึงกลางศตวรรษที่ 20 อันมีผลทำให้ศาสตร์ต่างๆ มีความพยายามจะนำวิธีการหาความรู้แบบวิทยาศาสตร์ (scientific method) มาใช้กับแขนงวิชาตัวเองเพื่อให้มีลักษณะเป็นปรนัยมากยิ่งขึ้นรวมทั้งมนุษยศาสตร์เองด้วย

วิธีการทำงานของนักมนุษยศาสตร์ดิจิทัลโดยทั่วไปมีสองลักษณะคือแบบแรกเป็นการออกแบบโปรแกรมและเขียนชุดคำสั่ง (coding) ซอฟต์แวร์คอมพิวเตอร์ด้วยตัวเองหรือทำงานร่วมกันกับนักคอมพิวเตอร์และ/หรือนักคณิตศาสตร์ นักสถิติ ในการพัฒนาโปรแกรมหรือเทคโนโลยีเฉพาะขึ้นมาเพื่อใช้ประมวลชุดของข้อมูล (data set) ที่ต้องการศึกษาไม่ว่าจะเป็น เพื่อวิเคราะห์ตัวบทต่างๆ เช่น วรรณกรรม หนังสือ เอกสาร จดหมายเหตุ ข้อความ ถ้อยคำสุนทรพจน์หรือคำสบลในแต่ละวัฒนธรรม หรือวิเคราะห์งานด้านประติมากรรม ศิลปกรรม จิตรกรรม เช่น ภาพวาด ภาพเขียน รูปปั้น รูปสลัก รูปถ่าย ไปจนถึงโบราณวัตถุหลากหลายชนิดและสิ่งอื่นๆ อีกมากมายอันเกิดจากผลงานสร้างสรรค์ของมนุษย์ ส่วนแบบที่สอง เป็นการนำเทคโนโลยีหรือโปรแกรมสำเร็จรูปที่ถูกพัฒนาโดยกลุ่มแรกไปใช้เป็นเครื่องมือในการศึกษาข้อมูลด้านมนุษยศาสตร์เพื่อหาความสัมพันธ์ในชุดของข้อมูลดังกล่าว โดยอาจจะกระทำผ่านกรอบคิดด้าน

มนุษยศาสตร์และ/หรือทฤษฎีทางสังคมศาสตร์ที่มีอยู่แล้วหรือนำไปสู่การตีความหาความหมาย วินิจฉัย ประเมินคุณค่าด้วยชุดของคำอธิบายแบบใหม่หรือการสร้างสรรค์สิ่งใหม่ (originality)

ด้วยวิธีการดังกล่าว ทำให้เห็นได้ว่ามนุษยศาสตร์ดิจิทัลเป็นวิธีการศึกษาที่ใช้ในการเข้าหาข้อเท็จจริงและเข้าถึงข้อมูลในอีกรูปแบบหนึ่ง เพื่อหาความสัมพันธ์ชุดของข้อมูล เป็นวิธีที่สามารถจะนำมาอธิบายข้อมูลและเอกสารต่างๆ ให้มีความละเอียดและมีความหลากหลายมากขึ้น ทำให้เห็นรูปแบบ (pattern) ของข้อมูลได้ชัดเจนยิ่งขึ้นและเห็นในรูปแบบที่แตกต่างออกไปจากเดิม อันนำไปสู่การสกัดความรู้ใหม่หรือความรู้บางอย่างที่ถึงแม้วิธีการของมนุษยศาสตร์แบบเดิมจะไม่สามารถทำได้ เช่นการอ่านแบบละเอียดละไม (closed reading) แต่ก็อาจจะยังต้องอาศัยกรอบคิดทางมนุษยศาสตร์แบบเดิมและ/หรือทฤษฎีสังคมศาสตร์รวมไปถึงบริบททางประวัติศาสตร์และวัฒนธรรมเพื่อช่วยจัดระเบียบข้อมูลที่ได้จากการประมวลผลด้วยวิธีการของมนุษยศาสตร์ดิจิทัล เช่น การวิเคราะห์ความหมายของคำว่า “ความรัก” ที่ปรากฏอยู่ในวรรณกรรมฝรั่งเศสและเยอรมันในยุคกลาง สึกกับยุคโรแมนติกทั้งหมดว่ามีความแตกต่างกันอย่างไร ซึ่งปัจจุบันตัวบทของวรรณกรรมทั้งหมดดังกล่าวได้อยู่ในรูปแบบหนังสืออิเล็กทรอนิกส์เกือบทั้งหมดแล้ว หรือ การหาความหมายของคำว่า “ความงาม” ในตัวบทที่อยู่ใน google book ทั้งหมดในรอบ 20 ปีที่ผ่านมา ด้วยการวิเคราะห์ข้อความและบริบทโดยวิธีการประมวลผลตามภาษาธรรมชาติ

ตัวอย่างดังกล่าวแสดงให้เห็นว่ามนุษยศาสตร์ดิจิทัลจะเป็นการเพิ่มขีดความสามารถในการยกระดับการเก็บและวิเคราะห์ชุดข้อมูลด้านมนุษยศาสตร์ที่ถูกจัดเก็บอยู่ในรูปแบบที่หลายหลายและไม่ใช่เพียงเฉพาะที่เป็นตัวหนังสือเท่านั้น แต่ยังรวมถึงข้อมูลที่ไม่เป็นโครงสร้าง (unstructured data) เช่น ข้อความ ภาพ เสียง วีดีโอ ข้อมูลที่อยู่บนโลกออนไลน์ ในโลกโซเชียลมีเดีย ข้อมูลการใช้สมาร์ตโฟน

ด้วยขนาดสเกลที่กว้างและลึกอย่างที่ไม่เคยทำได้มาก่อนในอดีต วิธีดังกล่าวทำให้นักมนุษยศาสตร์สามารถเข้าถึงชุมชนทรัพยากรข้อมูลขนาดใหญ่ที่ในสมัยก่อนไม่สามารถทำได้ เช่น บริษัทกูเกิลอนุญาตให้นักวิจัยผู้ได้รับทุน Google's Digital Humanities Research Award สามารถเข้าถึงระบบฐานข้อมูลและระบบเทคโนโลยีคอมพิวเตอร์ของ Google เพื่อช่วยในการประมวลผลงานวิจัยดังที่กล่าวมาข้างต้นในโครงการ Google Book ซึ่งไม่เพียงแต่จะเป็นการทำให้ความรู้ด้านมนุษยศาสตร์ถูกขยายออกไปเท่านั้น แต่ยังอาจจะนำไปสู่การเชื่อมโยง ต่อยอดสร้างไวยากรณ์ทางความรู้กับศาสตร์อื่นๆ ได้อีกเช่นกัน

สุดท้าย ผู้เขียนเห็นว่าวิธีการศึกษาแบบมนุษยศาสตร์ดิจิทัลไม่ใช่วิธีการศึกษาที่จะมาทดแทนวิธีการศึกษามนุษยศาสตร์แบบดั้งเดิมที่เน้นใช้วิธีการตีความการประเมินคุณค่าและวินิจฉัยประสบการณ์ แต่ทั้งสองแนวทางนี้จะต้องทำงานร่วมกัน เป็นการศึกษาแบบประสานแนวทางการศึกษาเข้าด้วยกันเพื่อนำไปสู่ความเชื่อมโยงระหว่างสาขาวิชามนุษยศาสตร์เข้ากับสาขาวิทยาการข้อมูล (Data Science) และ/หรือสาขาอื่นๆ อีกที่เกี่ยวข้อง ในอนาคตเมื่อเทคโนโลยีก้าวหน้าไปเรื่อยๆ และทุกอย่างเข้าสู่หรือกลายเป็นดิจิทัลอย่างเต็มตัว การผลิตความรู้ (Knowledge Production) ทางด้านมนุษยศาสตร์ใหม่ๆ อาจไม่ได้อยู่ที่ตัวนักมนุษยศาสตร์เพียงกลุ่มเดียวไม่ว่าจะเป็นนักมนุษยศาสตร์แบบดั้งเดิมหรือแบบดิจิทัล

ดังเช่น John Asmus นักฟิสิกส์ผู้เชี่ยวชาญเรื่องแสงเลเซอร์ แสงซินโครตรอน ประจำสถาบันฟิสิกส์ประยุกต์และบริสุทธิ์ (Institute for Pure and Applied Physical Science) มหาวิทยาลัยแคลิฟอร์เนีย ซานดิเอโก สหรัฐอเมริกา ได้ใช้เทคนิคทางด้านฟิสิกส์เรื่องการกำทอนของแม่เหล็กนิวเคลียส (Nuclear Magnetic Resonance) ในการสแกนเลเซอร์บนฝาผนังเพื่อค้นหาภาพจิตรกรรมฝาผนังที่ชื่อ The Battle of Ahghiar ของ Leonardo da Vinci ที่หายสาบสูญ

โดยในเบื้องต้น ภายหลังจากที่นักมนุษยศาสตร์ (นักจิตรกรรม) ได้ศึกษาเอกสารและข้อมูลจำนวนมากมาแล้วได้ตั้งสมมติฐานว่าภาพดังกล่าวอาจจะถูกเขียนทับโดยจิตรกรอีกคนหนึ่งคือ Giorgio Vasari ผู้มีพฤติกรรมที่ชอบเขียนภาพวาดของตัวเองไปทับภาพวาดของคนอื่นๆ แต่ด้วยข้อจำกัดของวิธีการศึกษาทางมนุษยศาสตร์ทำให้ไม่สามารถทำการวิจัยต่อไปเพื่อพิสูจน์สมมติฐานดังกล่าวได้ เพราะไม่ต้องการทำลายภาพเขียนของ Vasari เพื่อค้นหาภาพของ Da Vinci ที่อาจจะถูกเขียนทับไว้ ต่อจากนั้น Asmus จึงได้ใช้เทคนิคการประมวลผลภาพแบบเชิงดิจิทัลและแสงอัลตราโซนิกเพื่อดึงภาพวาดของ Da Vinci ออกมาจากฝาผนังและเก็บไว้ในคอมพิวเตอร์ (Asmus, 1978 : 209-213 และ เจตนา, 2538 : 251)

งานของ Asmus เป็นตัวอย่างหนึ่งที่แสดงให้เห็นความสัมพันธ์ระหว่างศาสตร์สาขาต่างๆ ด้วยการหยิบยืมวิธีการวิจัยซึ่งกันและกันระหว่างวิทยาศาสตร์และมนุษยศาสตร์ อันแสดงให้เห็นว่าการผลิตองค์ความรู้ใหม่ไม่ว่าจะเป็นในสาขาใด ไม่จำเป็นต้องถูกผูกขาดโดยวิธีการศึกษาหลักในสาขานั้นเสมอไปเพราะด้วยความก้าวหน้าของเทคโนโลยีในปัจจุบันนี้ โดยเฉพาะด้านคอมพิวเตอร์และดิจิทัลล้วนแต่จะทำให้เกิดประโยชน์ในการศึกษาในทุกวงการวิชาการทั้งด้านวิทยาศาสตร์ สังคมศาสตร์และมนุษยศาสตร์



เอกสารอ้างอิง

หนังสือและบทความในหนังสือ

- กุลธิดา ท้วมสุข. (2558). *ก้าวสู่ดิจิทัลมนุษยศาสตร์ ดิจิทัลเพื่อสังคมและทรัพยากรความรู้*. เอกสารนำเสนอในการสัมมนาความร่วมมือระหว่างห้องสมุดสถาบันอุดมศึกษา ครั้งที่ 31 ณ สำนัก หอสมุด มหาวิทยาลัยขอนแก่น วันที่ 2-4 ธันวาคม 2558.
- จรรยา สุวรรณทัต. (2539). *มิติการวิจัยทางมนุษยศาสตร์-สังคม-พฤติกรรมศาสตร์. ที่ระลึก ๓๐ ปี คณะมนุษยศาสตร์ มหาวิทยาลัยเชียงใหม่*. เชียงใหม่ : โรงพิมพ์เมือง.
- จุฑาทิพย์ อุมะวิชนี. (2541). *สหวิทยาการมนุษยศาสตร์ : มิติแห่งมนุษย์*. กรุงเทพมหานคร : สำนักพิมพ์มหาวิทยาลัยธรรมศาสตร์.
- เจตนา นาควัชร. (2532). *เพื่อความอยู่รอดของมนุษยศาสตร์*. กรุงเทพมหานคร : สำนักพิมพ์จุฬาลงกรณ์มหาวิทยาลัย.
- เจตนา นาควัชร. (2538). *ทางไปสู่วัฒนธรรมแห่งการวิจารณ์*. กรุงเทพมหานคร : สำนักพิมพ์ผู้จัดการ.
- ชูศักดิ์ ภัทรกุลวณิช. (2545). *อ่าน(ไม่)เอาเรื่อง : รวมบทความวรรณกรรมศึกษาและบทวิจารณ์ วรรณกรรมไทยและเทศ : โครงการจัดพิมพ์คบไฟ*.
- นิธิ เอียวศรีวงศ์. (2539). *มองมนุษยศาสตร์ในโลกอนาคต. ที่ระลึก ๓๐ ปี คณะมนุษยศาสตร์ มหาวิทยาลัยเชียงใหม่*. เชียงใหม่ : โรงพิมพ์เมือง.
- นันทนัย ประสานนาม และ ชีระ รุ่งธีระ. (2557). *เพื่อก้าวข้ามสะพานที่พาดข้าม : การศึกษารวบรวมบทคัดย่อและบรรณานุกรมงานวิจัยด้านวรรณกรรมไทยร่วมสมัย*. กรุงเทพมหานคร : โรงพิมพ์เดือนตุลาคม

Books and Book Articles

- Asmus, John Fredrich. (1978). "The Search for Leonardo da Vinci's The Battle of Aghajari". *In the Spirit of Enterprise*. Edited by Gregory B. Stone. San Francisco: W.H. Freeman.
- Barths, Roland. (1977). "The Photographic Message". *Image/Music/Text*. Stephen Heath trans. NY : Hill and Wang.
- Blatt, Ben. (2017). *Nabokov's Favorite Word is Mauve: What the Number Reveal About the Classic, Bestseller, and Our Own Writing*. Simon & Schuster: New York.
- Blaydes, Lisa, Grimmer, Justin and McQueen, Alison. (2013 October). *Mirror for Princes And Sultans: Advice on the Art of Governance in the Medieval Christian and Islamic Worlds*. Paper Presented to Association for Analytic Learning About Islam and Muslim Societies. Princeton University, NJ.
- Drucker, Johana. (2009). *SpecLab: Digital Aesthetics and Projects in Speculative Computing*. Chicago: The University of Chicago Press.-----(2014), *Graphesis: Visual Forms of Knowledge Production*. Harvard University Press.
- Gold, Matthew K. (2012). *Debate in the Digital Humanities*. University of Minnesota Press.
- Gold, Matthew K. and Klein F., Lauren. (2016). *Debate in the Digital Humanities*. University of Minnesota Press.

- Kirschenbaum, M.G. (2009). *Mechanisms: New Media and the Forensic Imagination*. Cambridge: MIT Press.
- Klein, Julie Thompson. (2015). *Interdisciplining Digital Humanities: Boundary Work in an Emerging Field*, Michigan: University of Michigan Press.
- McCarty, Willard. (2005). *Humanities Computing*. Basingstoke and New York: Palgrave Macmillan.
- Terras, Melisa., Nyhan, Julianne. And Vanhoutte, Edward. (2013). *Definition Digital Humanities.*, Ashagte Publishing Company. USA.
- Tobler, W. (1970). A computer movie simulating urban growth in the Detroit region. *Economic Geography*. 46(2), 234-240.
- Ye, Nong. (2014). *Data Mining: Theories, Algorithms, and Examples*. CRC Press.
- Zikopoulos P, Eaton C, deRoos D, et al (2012). *Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data*. McGraw-Hill.

Articles

- Boix, Mansilla and Duraisingh, Dawes (2007). Target Assessment of Students' Interdisciplinary Work: An Empirical Framework Proposed. *The Journal of Higher Education*. Vol. 78, No. 2, 215-237.
- Borrego, M. and Newswander, L.K. (2010). Definitions of Interdisciplinary Research: Toward Graduate-Level Interdisciplinary Learning

- Outcomes. *The Review of Higher Education*, Vol. 34, No. 1, 61-84.
- Bradley, M., & Lang, P. (1999). Affective norms for English words (anew): Stimuli, instruction manual and affective ratings. Technical report c-1, Gainesville, FL: University of Florida.
- Busa, R. (1980). The annals of humanities computing: The Index Thomisticus. *Computer and the Humanities* (14), 83-90.
- Chen, M S., Han, J. and Yu, P.S. (December,1996). Data Mining: An Overview from a Database Perspective” *IEEE Trans. Knowledge and Data Eng.*, 866-883.
- Cleave, W.S. (2001). Data Science: an action plan for expanding the technical areas of the field of the statistics. *International Statistics Review* Vol. 69, No.1, 21- 26.
- Cohen, Daniel J. and Gibbes, Frederick W. (2011). A Conversation with Data : Prospecting Victorian Words and Ideas. *Victorian Studies*, Vol. 54, No. 1 (Autumn 2011), 69-77.
- De Smith, Michael John., Goodchild, Michael F., and Longley, Paul (2007). *Geospatial Analysis: A Comprehensive Guide to Principles. Techniques and Software Tools*, Troubador Publishing Ltd.
- Mosteller, Federick. and David F.Wallace. (1963). Inference in an Authorship Problem. *Journal of the American Statistical Association*. Vol 58, No. 302, 275-309.

Reynolds, Noel B. and Saxonhouse, Arlene. (1995). *Hobbes and the Horae Subsecivae*. Chicago, 3-22.

Rosembloom, Paul. (2012). Toward a Conceptual Framework for the Digital Humanities. *Digital Humanities Quarterly*. Vol 6. No. 2.

Sudhaharand, S. and Cristianini, Nello (2013). Automated analysis of narrative content for Digital humanities. *International Journal of Advanced Computer Science*, 3(9).

Electronic Media

Dodds, Peter Sheridan and Chris Danforth. (2009). Measuring the Happiness of Large-Written Expression: Songs, Blogs and President. Research Paper. Published with open access at www.springerlink.com [Retrieved: January 24, 2018]

Dodds, Peter Sheridan and Chris Danforth..(2014), *Unearthinh Homo Narrativu*, Retrieved from http://www.internationalinnovation.com/build/wpcontent/uploads/2014/10/Peter_Dodds_Intl_Innovation_157_Research_Media.pdf [access: January 24, 2018]

Erlandsson, Fredrik, Piotr Brodka, Martin Boldt and Henric Johnson. (2017). Do We Really Need to Catch Them All? A New User-Guided Social Media Crawling Method. *Entropy* 2017,19,686 Published with open access at www.mdpi.com/journal/entropy [Retrieved: May 24, 2018]

McCarty, Willard. (1999). *Humanity Computing as interdiscipline*. In An Interdisciplinary Seminar at The University of Virginia.

Retrieved from <http://www.iath.virginia.edu/hcs> [Retrieved: January 5, 2018]

Rockwell, Geoffrey. (1999). *Is Humanities Computing as Academic Discipline?* An Interdisciplinary Seminar at the University of Virginia (1999-2000). Retrieved from <http://www.iath.virginia.edu/hcs> [access: January 5, 2018].

Spiro, Lisa. *Digital Research Tools (DiRT)* WiKi. Retrieved from <https://digitalreserachtools.pbworks.com/w/page/17801672/FrontagePage>. [access: October 31, 2017].

Svensson, Patrik. (2009). Humanities Computing as Digital Humanities. *Digital Humanities Quarterly* 3, no 3. Retrieved from <http://digitalhumanities.org/dhq/vol/3/3/000065/000065.html>. [access: October 3, 2017].

Unsworth, John. (November 2002) What Is Humanities Computing and What Is Not? *Graduate School of Library and Information Sciences*. Illinois Informatics Institute. University of Illinois, Urbana. Retrieved from <http://computerphilologie.uni-muenchen.de/jg02/unsworth.html>. [access: March 31, 2018].
----- (2002). *Is Humanities Computing an Academic Discipline?* An Interdisciplinary Seminar at The University of Virginia (1999-2000) Retrieved from <http://www.iath.virginia.edu/hcs> [access: January 5, 2018].