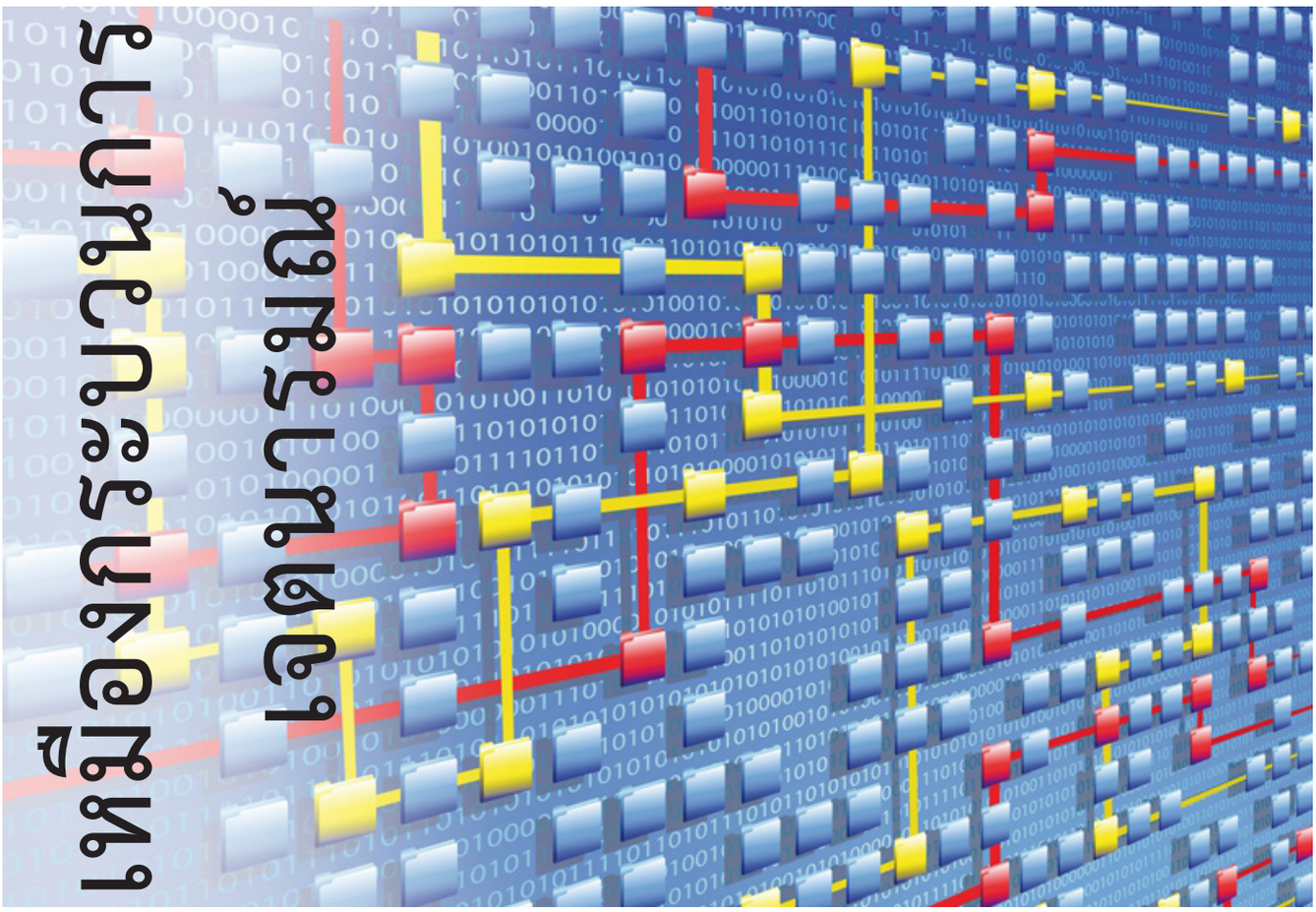


เหมืองกระบวนการ

เจตนารมณ์

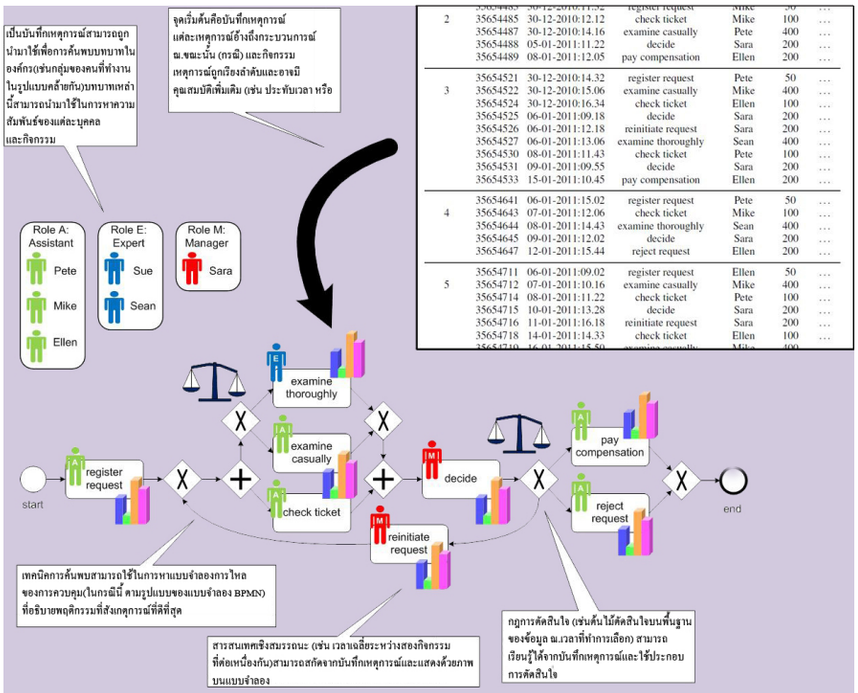
เจตนารมณ์ คือ “คำประกาศต่อสาธารณะถึงหลักการและความตั้งใจ” ของคนกลุ่มหนึ่ง เจตนารมณ์นี้เขียนขึ้นโดยสมาชิกและผู้สนับสนุนของ IEEE Task Force on Process Mining เป้าหมายของคณะทำงานเฉพาะกิจนี้ คือ เพื่อส่งเสริมการวิจัย การพัฒนา การศึกษา การทำให้เกิดผล วิวัฒนาการ และความเข้าใจในเรื่องเหมืองกระบวนการ

เหมืองกระบวนการ เป็นสาขาการวิจัยที่อยู่ในระยะเริ่มแรก ซึ่งในมุมหนึ่งมองว่าอยู่ระหว่างการคำนวณอัจฉริยะและเหมืองข้อมูล ส่วนในอีกมุมหนึ่งมองว่าอยู่ระหว่างแบบจำลองกระบวนการและการวิเคราะห์ แนวคิดของเหมืองกระบวนการคือ เพื่อการค้นพบ การเฝ้าสังเกตุ และการปรับปรุงกระบวนการจริง (ไม่ใช่กระบวนการสมมุติ) ด้วยการสกัดความรู้จากบันทึกเหตุการณ์ที่มีอยู่ในระบบ (สารสนเทศ) ที่มีอยู่แล้วในปัจจุบัน เหมืองกระบวนการรวมถึง การค้นพบกระบวนการ (โดยอัตโนมัติ) (การสกัดแบบจำลองกระบวนการจากบันทึกเหตุการณ์) การตรวจสอบการสอดคล้อง (เช่น การเฝ้าสังเกตุความเบี่ยงเบนโดยการเปรียบเทียบระหว่างแบบจำลองกับบันทึก) เครือข่ายสังคม / เหมืององค์กร การสร้างแบบจำลองโดยอัตโนมัติ การขยายแบบจำลอง ซ่อมแซมแบบจำลอง การพยากรณ์กรณี และการให้คำแนะนำโดยใช้ความเป็นมาเป็นพื้นฐาน

เนื้อหา:

เหมืองกระบวนการ - ความทันรู้สมัย	3
หลักการแนะนำ	6
ความท้าทาย	10
ปัจฉิมบท	13
อภิธานศัพท์	14

เทคนิคของเหมืองกระบวนการสามารถสกัดความรู้จากบันทึกเหตุการณ์ที่มีอยู่โดยทั่วไปในระบบสารสนเทศในปัจจุบัน เทคนิคเหล่านี้ให้วิธีใหม่ในการค้นพบ, การเฝ้าสังเกตุ และการปรับปรุงกระบวนการในการประยุกต์ใช้ในหลากหลายโดเมน โดยมีตัวขับเคลื่อนหลักสองประการที่ทำให้เกิดความสนใจเหมืองกระบวนการเพิ่มขึ้น ในมุมหนึ่งมีการบันทึกเหตุการณ์เพิ่มมากขึ้น ทำให้มีสารสนเทศรายละเอียดที่เกี่ยวกับประวัติของกระบวนการ ในอีกมุมหนึ่งมีความต้องการปรับปรุงและสนับสนุนกระบวนการธุรกิจในสภาวะแวดล้อมที่มีการแข่งขันและการเปลี่ยนแปลงอย่างรวดเร็ว เจตนารมณ์นี้สร้างสรรค์ขึ้นโดย IEEE Task Force on Process Mining และตั้งใจส่งเสริมเรื่องเหมืองกระบวนการ ยิ่งไปกว่านั้น ด้วยการกำหนดชุดของหลักการแนะนำ และรายการความท้าทายที่สำคัญ เจตนารมณ์นี้หวังที่จะทำหน้าที่เป็นคู่มือสำหรับนักพัฒนาซอฟต์แวร์ นักวิทยาศาสตร์ ที่ปรึกษา ผู้จัดการธุรกิจ และผู้ใช้ เป้าหมายคือเพื่อที่จะเพิ่มวุฒิภาวะของเหมืองกระบวนการในฐานะเครื่องมือใหม่เพื่อปรับปรุงการออกแบบ (ใหม่) ควบคุมและสนับสนุนกระบวนการในการดำเนินการทางธุรกิจ



2	35654485	30-12-2010:12.12	check ticket	Mike	100
	35654487	30-12-2010:14.16	examine casually	Pete	400
	35654488	05-01-2011:11.22	decide	Sara	200
	35654489	08-01-2011:12.05	pay compensation	Ellen	200
3	35654521	30-12-2010:14.32	register request	Pete	50
	35654522	30-12-2010:15.06	examine casually	Mike	400
	35654524	30-12-2010:16.34	check ticket	Ellen	100
	35654525	06-01-2011:09.18	decide	Sara	200
	35654526	06-01-2011:12.18	reinitiate request	Sara	200
	35654527	06-01-2011:13.06	examine thoroughly	Sean	400
	35654530	08-01-2011:11.43	check ticket	Pete	100
	35654531	09-01-2011:09.55	decide	Sara	200
	35654533	15-01-2011:10.45	pay compensation	Ellen	200
4	35654641	06-01-2011:15.02	register request	Pete	50
	35654643	07-01-2011:12.06	check ticket	Mike	100
	35654644	08-01-2011:14.43	examine thoroughly	Sean	400
	35654645	09-01-2011:12.02	decide	Sara	200
	35654647	12-01-2011:15.44	reject request	Ellen	200
5	35654711	06-01-2011:09.02	register request	Ellen	50
	35654712	07-01-2011:10.16	examine casually	Mike	400
	35654714	08-01-2011:11.22	check ticket	Pete	100
	35654715	10-01-2011:13.28	decide	Sara	200
	35654716	11-01-2011:16.18	reinitiate request	Sara	200
	35654718	14-01-2011:14.33	check ticket	Ellen	100

รูปที่ 1 : เทคนิคเหมืองกระบวนการสกัดความรู้จากบันทึกเหตุการณ์เพื่อการค้นพบการเฝ้าสังเกตุ และการปรับปรุบกระบวนการให้ดีขึ้น

เหมืองกระบวนการทำให้เกิดการเชื่อมต่อที่สำคัญระหว่างเหมืองข้อมูลและแบบจำลองกระบวนการธุรกิจและการวิเคราะห์ ภายใต้ธุรกิจอัจฉริยะ (BI) มีการใช้คำศัพท์เฉพาะหลายคำเพื่ออ้างถึงรายงานอย่างง่ายๆและเครื่องมือแดชบอร์ด การเฝ้าสังเกตุกิจกรรมธุรกิจ (BAM) หมายถึงเทคโนโลยีที่ทำให้เกิดความเป็นไปได้ในการเฝ้าสังเกตุกิจกรรมธุรกิจแบบเวลาจริง การประมวลผลเหตุการณ์ที่ซับซ้อน (CEP)

วัตถุประสงค์ที่เป็นรูปธรรมของคณะกรรมการเฉพาะกิจ

1) ทำให้ผู้ใช้, ผู้พัฒนา, ที่ปรึกษา, ผู้จัดการธุรกิจและนักวิจัย ตระหนักถึงความรู้ทันสมัยของการทำเหมืองกระบวนการ
2) ส่งเสริมการใช้เทคนิคการทำเหมืองกระบวนการและเครื่องมือและกระตุ้นให้เกิดการประยุกต์ใช้ในงานใหม่ๆ
3) แสดงบทบาทความพยายามในการกำหนดมาตรฐานสำหรับการบันทึกข้อมูลเหตุการณ์
4) จัดการสอนพิเศษ, กลุ่มพิเศษ, การประชุมเชิงปฏิบัติการ, กลุ่มอภิปราย
5) ตีพิมพ์บทความ, หนังสือ, วิดีโอ และปัญหาเฉพาะในวารสาร

หมายถึง เทคโนโลยีที่ใช้ในการประมวลผลเหตุการณ์จำนวนมากๆ เพื่อนำมาใช้ประโยชน์ในการเฝ้าสังเกตุ นำทาง และทำธุรกิจให้เหมาะสมที่สุดแบบเวลาจริง การบริหารสมรรถภาพองค์กร (CPM) เป็นอีกคำศัพท์เฉพาะหนึ่งที่ใช้ในการวัดสมรรถภาพของกระบวนการหรือองค์กร ทั้งยังมีความสัมพันธ์กับวิธีการบริหารอีกด้วย เช่น การปรับปรุงกระบวนการอย่างต่อเนื่อง (CPI), การปรับปรุงกระบวนการธุรกิจ (BPI), การบริหารคุณภาพโดยรวม (TQM), และ Six Sigma วิธีการเหล่านี้มีสิ่งที่ร่วมกันคือ กระบวนการ “ถูกนำไปวางไว้ภายใต้กล้องจุลทรรศน์” เพื่อดูความเป็นไปได้ว่ายังปรับปรุงเพิ่มเติมได้อีกหรือไม่ เหมืองกระบวนการเป็นเทคโนโลยีที่ทำให้เกิดความเป็นไปได้สำหรับ CPM, BPI, TQM, Six Sigma, และวิธีการอื่นๆในลักษณะนี้

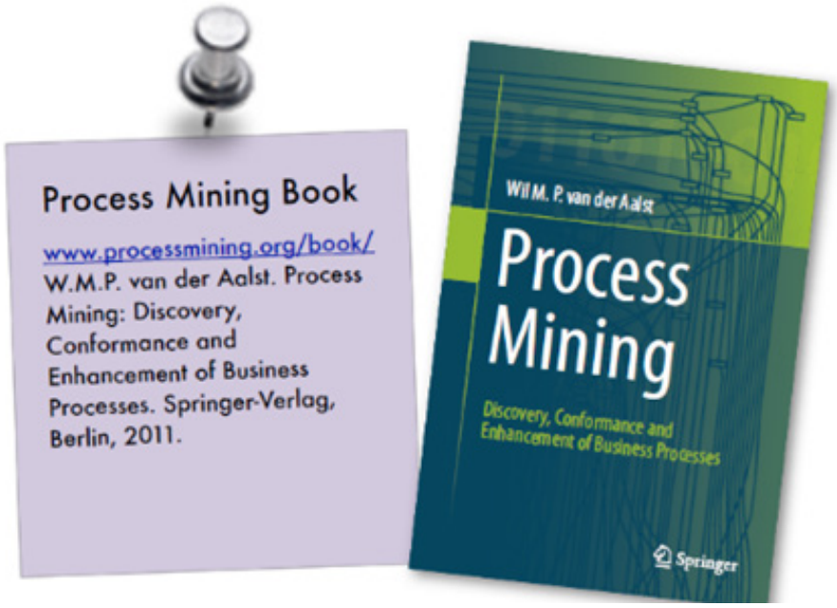
ในขณะที่เครื่องมือของ BI และวิธีการบริหาร เช่น Six Sigma และ TQM ตั้งใจที่การเพิ่มสมรรถนะการดำเนินงาน เช่น ลดเวลาและข้อบกพร่อง องค์กรมีการเน้นที่บรรษัทภิบาล ความเสี่ยง และการปฏิบัติตามข้อกำหนดมากขึ้น การปฏิบัติตามกฎหมาย เช่น Sarbanes-Oxley Act (SOX) และ the Basel II Accord แสดงถึงการเน้นประเด็นของการปฏิบัติตามข้อกำหนด เทคนิคเหมืองกระบวนการให้วิธีการตรวจสอบการปฏิบัติตามข้อกำหนดอย่างเข้มงวดและสมเหตุสมผล และมีความเชื่อถือได้ของสารสนเทศที่เกี่ยวกับกระบวนการหลักขององค์กร

ในทศวรรษที่ผ่านมา ข้อมูลเหตุการณ์ได้กลายมาเป็นสิ่งที่พร้อมใช้งาน และเทคนิคของเหมืองกระบวนการได้เติบโตเต็มที่ ยิ่งไปกว่านั้นดังที่ได้กล่าวมาแล้ว แนวโน้มการบริหารเกี่ยวกับการปรับปรุงกระบวนการให้ดีขึ้น (เช่น Six Sigma, TQM, CPI, และ CPM) และ การปฏิบัติตามข้อกำหนด (SOX, BAM, อื่นๆ) สามารถได้รับประโยชน์จากเหมืองกระบวนการ โชคดีที่อัลกอริทึมของเหมืองกระบวนการได้นำไปใช้ให้เกิดประโยชน์ทั้งในระบบการศึกษาและการพาณิชย์ที่หลากหลาย ทุกวันนี้ได้มีกลุ่มของนักวิจัยที่กระตือรือร้นกำลังทำการวิจัยในเรื่องเหมืองกระบวนการ และกำลังเป็นประเด็นที่ได้รับความนิยมในการวิจัยของการบริหารกระบวนการธุรกิจ (BPM) ยิ่งไปกว่านั้น เหมืองกระบวนการยังได้รับความสนใจอย่างมากจากภาคอุตสาหกรรมผู้จำหน่ายซอฟต์แวร์ได้เพิ่มฟังก์ชั่นของเหมืองกระบวนการเป็นเครื่องมือเพิ่มเติมมากขึ้น ตัวอย่างของผลิตภัณฑ์ซอฟต์แวร์ที่มีความสามารถของเหมืองกระบวนการ ได้แก่ ARIS Process Performance Manager (Software AG), Comprehend (OpenConnect),Discovery Analyst (StereoLOGIC), Flow (Fourspark), Futura Reflect (Futura Process Intelligence), Interstage Automated, Process Discovery (Fujitsu), OKT Process Mining suite (Exeura), Process Discovery Focus (Iontas/Verint), ProcessAnalyzer (QPR), ProM (TU/e), Rbminer/Dbminer (UPC), และ Reflect one (Pallas Athena) ความสนใจที่เพิ่มขึ้นในการวิเคราะห์กระบวนการจากบันทึกเป็นแรงผลักดันให้มีการจัดตั้งคณะกรรมการเฉพาะกิจของเหมือนกระบวนการขึ้น

คณะกรรมการเฉพาะกิจของเหมืองกระบวนการถูกจัดตั้งขึ้นในปี 2009 ในบริบทของ Data Mining Technical Committee (DMTC) ของ the Computational Intelligence Society (CIS) ของ the Institute of Electrical and Electronic Engineers (IEEE) ปัจจุบัน คณะทำงานเฉพาะกิจนี้มีสมาชิกที่เป็นตัวแทนของผู้จำหน่ายซอฟต์แวร์ (ตัวอย่างเช่น, Pallas Athena, Software AG, Futura Process Intelligence, HP, IBM, Infosys, Fluxicon, Businesscape, Iontas/Verint, Fujitsu, Fujitsu Laboratories, Business Process Mining, Stereologic), บริษัทที่ปรึกษา/ผู้ใช้ (ตัวอย่างเช่น, ProcessGold, BusinessProcess Trends, Gartner, Deloitte, Process Sphere, SiavSpA, BPM Chile, BWI Systeme GmbH, Excellentia BPM, Rabobank), และสถาบันวิจัย (ตัวอย่างเช่น, TU/e, University of Padua, UniversitatPolitecnica de Catalunya, New Mexico State University, IST - Technical University of Lisbon, University of Calabria, Penn State University, University of Bari, Humboldt-Universitatzu Berlin, Queensland University of Technology, Vienna University of Economics and Business,

Stevens Institute of Technology, University of Haifa, University of Bologna, Ulsan National Institute of Science and Technology, Cranfield University, K.U. Leuven, Tsinghua University, University of Innsbruck, University of Tartu)

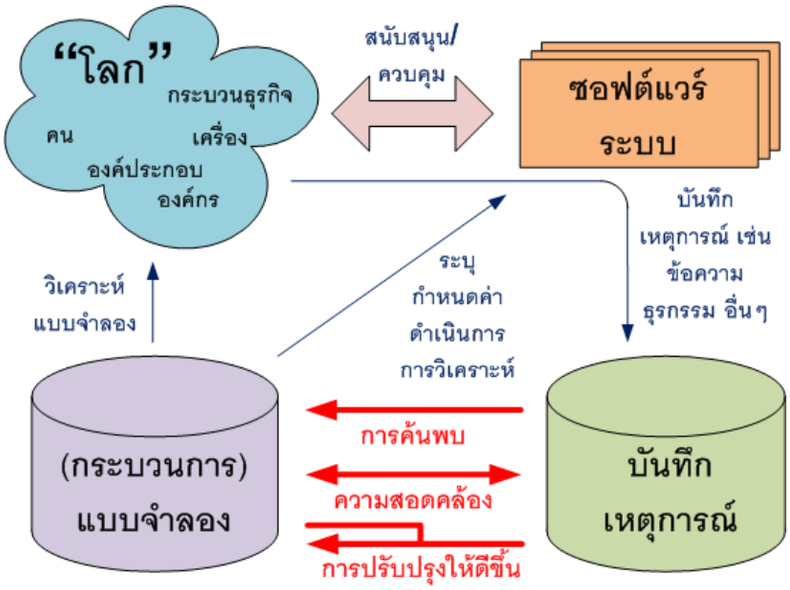
ตั้งแต่ก่อตั้งมาในปี 2009 ได้มีการจัดกิจกรรมหลายอย่างที่เกี่ยวกับวัตถุประสงค์ข้างต้น เช่น, คณะทำงานเฉพาะกิจนี้ได้เป็นผู้จัด (ผู้ร่วมจัด) การประชุมเชิงปฏิบัติการและกลุ่มเฉพาะ ตัวอย่างเช่น การประชุมเชิงปฏิบัติการ Business Process Intelligence (BPI'09, BPI'10, and BPI'11) และกลุ่มเฉพาะของการประชุมวิชาการของ IEEE (เช่น. CIDM'11) ความรู้ได้แพร่กระจายไปผ่านทางการสอนพิเศษ (เช่น. WCCI'10 and PMPM'09), โรงเรียนภาคฤดูร้อน (ESSCaSS'09, ACPN'10, CICH'10, etc.), วิดีโอ (www.processmining.org), และเอกสารต่างๆ รวมถึงหนังสือเล่มแรกในเรื่องเหมืองกระบวนการที่ตีพิมพ์โดย Springer คณะทำงานเฉพาะกิจนี้ยังเป็นผู้จัด the first Business Process Intelligence Challenge (BPIC'11): เป็นการแข่งขันที่ผู้เข้าร่วมแข่งขันต้องสกัดความรู้ที่เป็นประโยชน์จากบันทึกเหตุการณ์ขนาดใหญ่และซับซ้อน ในปี 2010 คณะทำงานเฉพาะกิจนี้ยังได้กำหนดมาตรฐาน XES (www.xesstandard.org), มาตรฐานรูปแบบของบันทึกที่ขยายได้และสนับสนุนโดย OpenXES library (www.openxes.org) และเครื่องมือ เช่น ProM, XESame, Nitro, เป็นต้น



ขอเชิญผู้อ่านเข้ามาเยี่ยมชมที่ www.win.tue.nl/ieeetfpm สำหรับสารสนเทศเพิ่มเติมและกิจกรรมของคณะทำงานเฉพาะกิจนี้

2. เหมืองกระบวนการ: ความรู้ทันสมัย

ขีดความสามารถที่เพิ่มขึ้นของระบบสารสนเทศและระบบอื่นๆที่ขึ้นอยู่กับการคำนวณ สามารถแสดงลักษณะได้ด้วยกฎของ Moore's (Moore's law) Gordon Moore ผู้ร่วมก่อตั้งบริษัท Intel ได้พยากรณ์ไว้ในปี 1965 ว่าจำนวนองค์ประกอบในวงจรรวม จะเพิ่มขึ้นเป็นสองเท่าทุกปี โดยแท้จริงในช่วง 50 ปีหลังนี้มีการเจริญเติบเป็นแบบเอกซ์โพเนนเชียล แม้ว่าจะเป็นไปอย่างช้าๆ



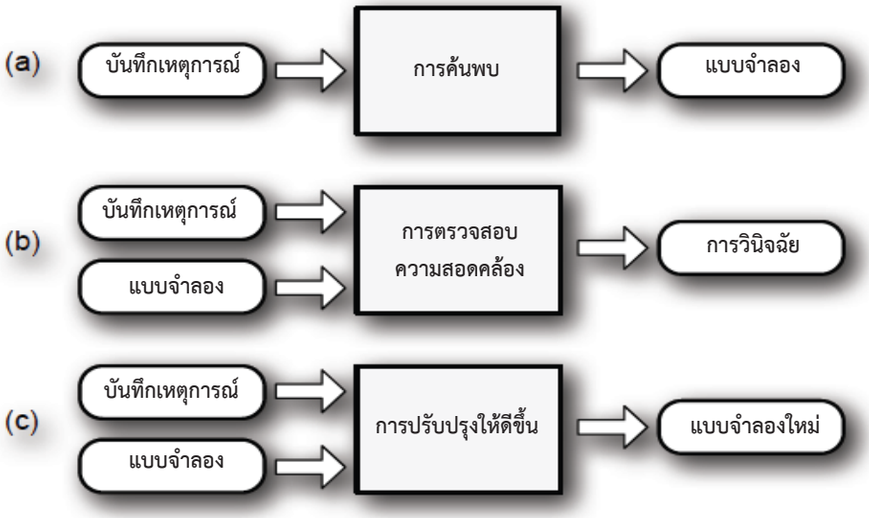
รูปที่ 2: การวางตำแหน่งของทั้งสามประเภทหลักของการทำเหมืองกระบวนการ (ก) การค้นพบ (ข) การตรวจสอบความสอดคล้องและ (ค) การปรับปรุงให้ดีขึ้น

ผลของความก้าวหน้าอย่างน่าประทับใจของ "โลกดิจิทัล" (ข้อมูลทั้งหมดถูกจัดเก็บ และ/หรือแลกเปลี่ยนในรูปของอิเล็กทรอนิกส์) นอกจากนั้นดิจิทัลและโลกของความจริงยังมีความเป็นไปในทิศทางเดียวกันมากขึ้น

ความเจริญเติบโตของโลกดิจิทัลที่เป็นไปในทิศทางเดียวกันกับกระบวนการในองค์กรเป็นอย่างดีทำให้มีความเป็นไปได้ที่จะบันทึกและวิเคราะห์เหตุการณ์ เหตุการณ์อาจมีพิสัยตั้งแต่การถอนเงินสดจากตู้ ATM แพทย์ทำการปรับแต่งเครื่อง X-ray ประชาชนขอรับใบอนุญาตขับขี่ การแจ้งภาษี และใบเสร็จรับเงินของหมายเลขตั๋วอิเลกทรอนิกส์จากผู้เดินทาง ความท้าทายคือการแสวงหาผลประโยชน์จากข้อมูลเหตุการณ์อย่างมีความหมาย ตัวอย่างเช่น เพื่อให้มีความเข้าใจอย่างลึกซึ้ง การระบุปัญหาคอขวด ปัญหาที่เกี่ยวข้อง บันทึกการละเมิดนโยบาย ให้คำแนะนำมาตรการโต้ตอบ และการปรับปรุงกระบวนการให้มีประสิทธิภาพมากยิ่งขึ้น เหมืองกระบวนการตั้งใจที่จะทำตรงกับสิ่งที่กล่าวมาแล้วข้างต้น

จุดเริ่มต้นของเหมืองกระบวนการคือบันทึกเหตุการณ์ เทคนิคของเหมืองกระบวนการทั้งหมดสมมุติว่ามันมีความเป็นไปได้ที่จะบันทึกเหตุการณ์อย่างลำดับ แต่ละเหตุการณ์จะอ้างถึงกิจกรรม (ขั้นตอนที่กำหนดอย่างชัดเจนในบางกระบวนการ) และสัมพันธ์กับกรณีใดกรณีหนึ่งเป็นการเฉพาะ (เช่น กระบวนการ ณ ขณะนั้น) บันทึกเหตุการณ์อาจมีการจัดเก็บสารสนเทศอื่นๆเพิ่มเติมอีกก็ได้ ในความเป็นจริง เมื่อมีความเป็นไปได้ เทคนิคของเหมืองกระบวนการจะใช้สารสนเทศเพิ่มเติมเหล่านี้ เช่น ทรัพยากร (บุคคลหรือเครื่องมือ) ที่ทำหรือริเริ่มกิจกรรม ประทับเวลา ของเหตุการณ์ หรือองค์ประกอบข้อมูลที่ถูกบันทึกกับเหตุการณ์ (เช่น ขนาดของการสั่งซื้อ)

ดังแสดงในรูปที่ 2 บันทึกเหตุการณ์สามารถนำไปใช้ในการทำเหมืองกระบวนการ 3 ชนิด ชนิดแรกของเหมืองกระบวนการคือ การค้นพบ เทคนิคการค้นพบจะรับเอาบันทึกเหตุการณ์และ



รูปที่ 3: สามประเภทพื้นฐานของการทำเหมืองกระบวนการที่อธิบายในรูปของ ข้อมูลเข้า และ ข้อมูลออก (a) การค้นพบ (b) การตรวจสอบความสอดคล้องและ (c) ปรับปรุงให้ดีขึ้น

สร้างแบบจำลองโดยไม่ต้องรู้สารสนเทศอื่นใดล่วงหน้า การค้นพบกระบวนการเป็นเทคนิคที่มีชื่อเสียงที่สุดของเทคนิคเหมืองกระบวนการ หลายองค์กรได้รับความประหลาดใจเมื่อได้เห็นว่าเทคนิคที่มีอยู่นี้สามารถค้นพบกระบวนการจริง เพียงแต่ใช้ตัวอย่างการทำงานในบันทึกเหตุการณ์ ชนิดที่สองของเหมืองกระบวนการคือการตรวจสอบความสอดคล้อง โดยทำการเปรียบเทียบแบบจำลองกระบวนการที่มีอยู่กับบันทึกเหตุการณ์ของกระบวนการเดียวกัน การตรวจสอบความสอดคล้องสามารถใช้ในการตรวจสอบว่าในความเป็นจริงแล้ว ตามที่บันทึกไว้ในบันทึกเหตุการณ์ มีความสอดคล้องกันกับแบบจำลองหรือไม่ และสามารถทำในทางกลับกันได้เช่นเดียวกัน เนื่องจากอาจมีแบบจำลองหลายชนิดที่ใช้ในการพิจารณา การตรวจสอบความสอดคล้องสามารถนำมาประยุกต์ใช้กับ แบบจำลองขั้นตอน แบบจำลององค์กร, แบบจำลองกระบวนการแบบดีแคลเรทีฟ, กฎ/นโยบายทางธุรกิจ, กฎหมาย, และอื่นๆ ชนิดที่สามของเหมืองกระบวนการคือการปรับปรุงให้ดีขึ้น แนวคิดคือการขยายหรือปรับปรุงแบบจำลองกระบวนการที่มีอยู่แล้วด้วยการใช้สารสนเทศของกระบวนการจริงที่ถูกบันทึกไว้ในบันทึกเหตุการณ์ ในขณะที่การตรวจสอบความสอดคล้องนั้นวัดความเป็นไปในทิศทางเดียวกัน ระหว่างแบบจำลองกับความเป็นจริง เหมืองกระบวนการชนิดที่สามนี้ตั้งใจที่จะเปลี่ยนแปลงหรือขยายแบบจำลองที่มีอยู่ ตัวอย่างเช่น ด้วยการใช้ประทับเวลาในบันทึกเหตุการณ์ เราสามารถขยายแบบจำลองให้สามารถแสดงถึงปัญหาคอขวด ระดับการให้บริการ เวลาที่ใช้ทั้งหมด และความถี่ได้

รูปที่ 3 อธิบายเหมืองกระบวนการทั้งสามชนิดในรูปของข้อมูลเข้าและข้อมูลออก เทคนิคสำหรับการค้นพบจะรับเอาบันทึกเหตุการณ์และสร้างแบบจำลอง แบบจำลองที่ค้นพบนี้คือแบบจำลองกระบวนการที่ใช้กันทั่วไป (เช่น Petri net, BPMN, EPC, หรือ UML activity diagram) อย่างไรก็ตาม แบบจำลองอาจอธิบายในมุมมองอื่นๆได้ด้วย (เช่น เครือข่ายสังคม) เทคนิคของการตรวจสอบความสอดคล้องต้องการบันทึกเหตุการณ์และแบบจำลองเป็นข้อมูลเข้า ข้อมูลออกประกอบด้วยสารสนเทศที่ใช้ในการวินิจฉัยที่แสดงถึงความแตกต่างและความเหมือนกันระหว่างแบบจำลองกับบันทึก เทคนิคสำหรับการปรับปรุงแบบจำลองให้ดีขึ้น (การซ่อมแซมและการขยาย) ต้องการบันทึกเหตุการณ์และแบบจำลองเป็นข้อมูลเข้า ผลลัพธ์ที่ได้คือแบบจำลองที่ได้รับการปรับปรุงหรือขยายออกไป

เหมืองกระบวนการอาจจะครอบคลุมมุมมองที่แตกต่างกัน ในมุมมองของการไหลของการควบคุมเน้นไปที่การไหลของการควบคุม เช่น ลำดับของกิจกรรม เป้าหมายของการทำเหมืองในมุมมองนี้คือการหาลักษณะที่ดีของเส้นทางที่เป็นไปได้ทั้งหมด ผลลัพธ์ที่ได้โดยทั่วไปจะอยู่ในรูปแบบของ Petri net หรือสัญลักษณ์กระบวนการแบบอื่นๆ (เช่น EPCs, BPMN, หรือ UML activity diagrams) มุมมององค์กรเน้นที่สารสนเทศเกี่ยวกับทรัพยากรที่ซ่อนอยู่ในบันทึกได้แก่ นักแสดง (เช่น คน ระบบ บทบาท หรือ แผนก) ที่เข้ามาเกี่ยวข้องและมีความเกี่ยวข้องกันอย่างไร เป้าหมายคือโครงสร้างองค์กรโดยการจัดกลุ่มคนตามบทบาทหน้าที่และหน่วยขององค์กร หรือแสดงเครือข่ายสังคม มุมมองของกรณี เน้นที่คุณลักษณะของกรณี จริงๆแล้วจะเห็นได้อย่างชัดเจนว่าแต่ละกรณีสามารถแสดงคุณลักษณะได้ด้วยเส้นทาง ในกระบวนการ หรือ โดยนักแสดงที่ทำงานนั้น อย่างไรก็ตาม กรณีสามารถถูกแสดงด้วยลักษณะของคุณค่าขององค์ประกอบข้อมูลที่สอดคล้องกัน ตัวอย่างเช่น

คุณลักษณะของเหมืองกระบวนการ:

1.เหมืองกระบวนการ ไม่ได้ถูกจำกัดอยู่ที่การค้นพบการไหลของการควบคุม

การค้นพบแบบจำลองกระบวนการจากบันทึกเหตุการณ์ เป็นตัวกระตุ้นจินตนาการของทั้งนักปฏิบัติและนักวิชาการ ดังนั้นการค้นพบการไหลของการควบคุมจึงมักถูกมองว่าเป็นส่วนที่น่าตื่นเต้นที่สุดของเหมืองกระบวนการ แต่อย่างไรก็ตาม เหมืองกระบวนการไม่ได้ถูกจำกัดอยู่ที่การค้นพบการไหลของการควบคุมเท่านั้นในมุมหนึ่ง การค้นพบเป็นเพียงหนึ่งในสามของรูปแบบพื้นฐานของการทำเหมืองกระบวนการ (การค้นพบ, การสอดคล้องกัน, และการเพิ่มประสิทธิภาพให้ดีขึ้น) ในอีกมุมหนึ่ง ขอบเขตก็ไม่ได้ถูกจำกัดอยู่เฉพาะการไหลของการควบคุม มุมมองขององค์การ กรณี และเวลา ก็เป็นบทบาทที่สำคัญ

2. เหมืองกระบวนการไม่ใช่รูปแบบเฉพาะของเหมืองข้อมูล

เหมืองกระบวนการสามารถถูกมองว่าเป็น “ตัวเชื่อมที่หายไป” ระหว่างเหมืองข้อมูลและแบบจำลองดั้งเดิม BPM เทคนิคส่วนใหญ่ของเหมืองข้อมูลไม่ได้มีจุดศูนย์กลางอยู่ที่กระบวนการ แบบจำลองกระบวนการมีศักยภาพในการแสดงให้เห็นถึงการทำงานพร้อมๆกันซึ่งเปรียบไม่ได้กับโครงสร้างพื้นฐานของเหมืองข้อมูล เช่น ต้นไม้ตัดสินใจ และ กฎความสัมพันธ์ ดังนั้นจึงมีความต้องการรูปแบบใหม่ทั้งหมดของตัวแทนและอัลกอริทึม

3. เหมืองกระบวนการไม่ได้ถูกจำกัดอยู่เฉพาะการวิเคราะห์แบบออฟไลน์

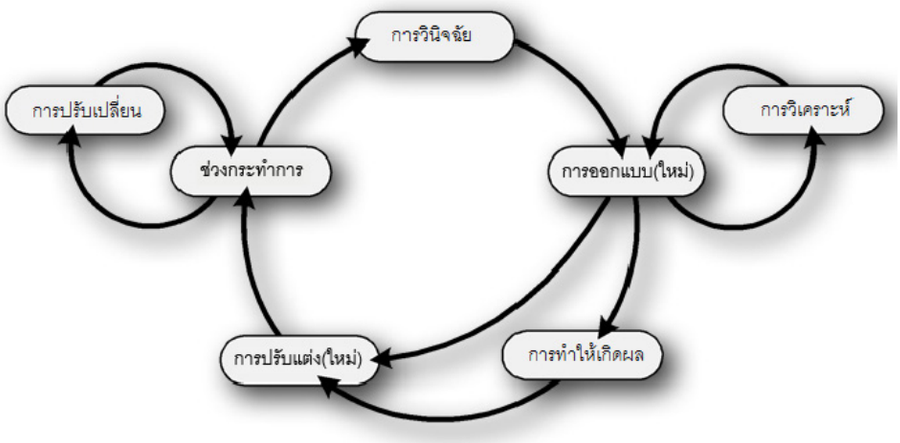
เทคนิคของเหมืองกระบวนการจะทำการสกัดความรู้จากข้อมูลเหตุการณ์ในอดีต ถึงแม้ว่าจะเป็นการใช้ข้อมูลในอดีตที่เกิดขึ้นและสิ้นสุดแล้ว ผลลัพธ์ที่ได้สามารถนำมาประยุกต์ใช้กับกรณีที่กำลังทำงานอยู่ได้ ตัวอย่างเช่น เวลาสิ้นสุดของการจัดการกับใบสั่งซื้อของลูกค้าที่ทำไปแล้วบางส่วนสามารถถูกพยากรณ์ได้โดยการใช้แบบจำลองกระบวนการที่ค้นพบ

ถ้ากรณีแทนคำสั่งซื้อเพิ่มเติม เราอาจสนใจต้องการรู้จักผู้ขาย หรือจำนวนผลิตภัณฑ์ที่สั่งซื้อไป มุมมองด้านเวลา จะเกี่ยวข้องกับเวลาและความถี่ของเหตุการณ์ เมื่อเหตุการณ์มีประทับเวลา มันมีความเป็นไปได้ที่จะค้นพบปัญหาคอขวด วัดระดับการให้บริการ การเฝ้าสังเกตุการใช้ทรัพยากร และพยากรณ์เวลาที่เหลืออยู่ในการประมวลผลของกรณีที่กำลังทำงานอยู่

มีความเข้าใจผิดบางอย่างร่วมกันเกี่ยวกับเหมืองกระบวนการของผู้ขาย นักวิเคราะห์ และนักวิจัยบางส่วนที่จำกัดขอบเขตของเหมืองกระบวนการว่าเป็นเทคนิคเฉพาะของเหมืองข้อมูลสำหรับกระบวนการค้นพบที่สามารถใช้ได้กับการวิเคราะห์แบบออฟไลน์เท่านั้น ซึ่งไม่ใช่สิ่งที่ถูกต้อง ดังนั้นเราจึงเน้นที่คุณลักษณะทั้งสามประการที่แสดงในหน้าที่แล้ว

เพื่อวางตำแหน่งของเหมืองกระบวนการเราใช้วงจรชีวิตของ BPM ที่แสดงในรูปที่ 4 วงจรชีวิตของ BPM แสดง 7 ขั้นตอนกระบวนการธุรกิจและระบบสารสนเทศที่สอดคล้องกัน ในขั้นตอนของการออกแบบ(ใหม่) แบบจำลองกระบวนการใหม่จะถูกสร้างขึ้น หรือแบบจำลองกระบวนการที่มีอยู่แล้วจะถูกปรับแต่งให้เหมาะสม ในขั้นตอนของการวิเคราะห์แบบจำลองตัวแทนและทางเลือกจะถูกวิเคราะห์ หลังจากขั้นตอนของการออกแบบ(ใหม่) แบบจำลองจะถูกนำไปทำให้เกิดผล (ขั้นตอนการนำไปให้เกิดผล) หรือระบบที่มีอยู่จะถูกปรับแต่ง (ใหม่) (ขั้นตอนการปรับแต่ง) ในขั้นตอนช่วงกระทำการ แบบจำลองที่ถูกออกแบบไว้แล้วจะถูกใช้งาน ระหว่างขั้นตอนช่วงกระทำการ กระบวนการจะถูกเฝ้าสังเกตุ ยิ่งไปกว่านั้นอาจมีการปรับแต่งเล็กๆน้อยๆโดยไม่ผ่านกระบวนการออกแบบใหม่ (ขั้นตอนการปรับเปลี่ยน) ในขั้นตอนการวินิจฉัย กระบวนการที่ใช้อยู่จะถูกวิเคราะห์และผลลัพธ์ของขั้นตอนนี้อาจกระตุ้นให้เกิดกระบวนการในขั้นตอนออกแบบใหม่ เหมืองกระบวนการเป็นเครื่องมือที่มีคุณค่าสำหรับขั้นตอนส่วนใหญ่ที่แสดงในรูปที่ 4 จะเห็นได้อย่างชัดเจนว่า ในขั้นตอนการวินิจฉัยสามารถได้รับประโยชน์จากเหมืองกระบวนการ แต่อย่างไรก็ตาม เหมืองกระบวนการไม่ได้จำกัดให้ใช้เฉพาะในขั้นตอนการวินิจฉัยเท่านั้น ตัวอย่างเช่นในขั้นตอนช่วงกระทำการ เทคนิคของเหมืองกระบวนการสามารถนำไปใช้ในการสนับสนุนการดำเนินการ การพยากรณ์และการให้คำแนะนำบนพื้นฐานของแบบจำลองที่เรียนรู้ด้วยการใช้สารสนเทศในอดีตสามารถถูกนำมาใช้ให้เกิดอิทธิพลต่องานในกรณีที่กำลังทำงานอยู่ได้ รูปแบบการสนับสนุนการตัดสินใจที่คล้ายคลึงกันสามารถนำไปใช้ในการปรับเปลี่ยนกระบวนการให้ดีขึ้นและแนะนำการปรับแต่งกระบวนการ

ขณะที่รูปที่ 4 แสดงภาพรวมวงจรชีวิตของ BPM รูปที่ 5 เน้นที่รูปธรรมของกิจกรรมเหมืองกระบวนการและตัวแปรต่างๆ รูปที่ 5 อธิบายถึงขั้นตอนที่เป็นไปได้ในโครงการของเหมืองกระบวนการ โครงการของเหมืองกระบวนการทุกโครงการเริ่มต้นด้วยการวางแผนและการให้เหตุผล (ขั้นที่ 0) หลังจากริเริ่มโครงการ ข้อมูลเหตุการณ์ แบบจำลอง วัตถุประสงค์ และคำถามจะต้องถูกสกัดจากระบบ, ผู้เชี่ยวชาญเฉพาะเรื่องและผู้บริหาร (ขั้นที่ 1) สิ่งเหล่านี้ต้องการความเข้าใจของข้อมูลที่มีอยู่ (“ข้อมูลอะไรบ้างที่สามารถถูกนำใช้ในการวิเคราะห์”) และเข้าใจโดเมน (“คำถามที่สำคัญคืออะไร”) และผลลัพธ์ในตัวแปรต่างๆที่แสดงในรูปที่ 5 (เช่น ข้อมูลในอดีต แบบจำลองที่สร้างด้วยมือ วัตถุประสงค์ และคำถาม) ในขั้นที่ 2 แบบจำลองการไหลของการควบคุมจะถูกสร้างขึ้นและเชื่อมโยงกับบันทึกเหตุการณ์ เทคนิคการค้นพบกระบวนการโดยอัตโนมัติสามารถถูกนำมาใช้ ณ.จุดนี้ แบบจำลองที่ถูกค้นพบอาจจะสามารถให้คำตอบสำหรับบางคำถามได้เลย และกระตุ้นให้เกิดการออกแบบใหม่ หรือการปรับเปลี่ยนการกระทำ



รูปที่ 4: วงจรชีวิต BPM ระบุถึงหลายขั้นตอนของกระบวนการธุรกิจ และ ระบบสารสนเทศที่สอดคล้องกันเหมืองกระบวนการ (มีแนวโน้มว่า) มีบทบาทในทุกๆขั้นตอน (ยกเว้นขั้นตอนการทำให้เกิดผล)

หลักการแนะนำ:

GP1: ข้อมูลเหตุการณ์ควรได้รับการปฏิบัติอย่างเช่นประชาชนชั้นหนึ่ง

GP2: การสกัดบันทึกควรถูกผลักดันจากคำถาม

GP3: การทำงานพร้อมๆกัน ทางเลือก และโครงสร้างการไหลของการควบคุมพื้นฐานควรได้รับการสนับสนุน

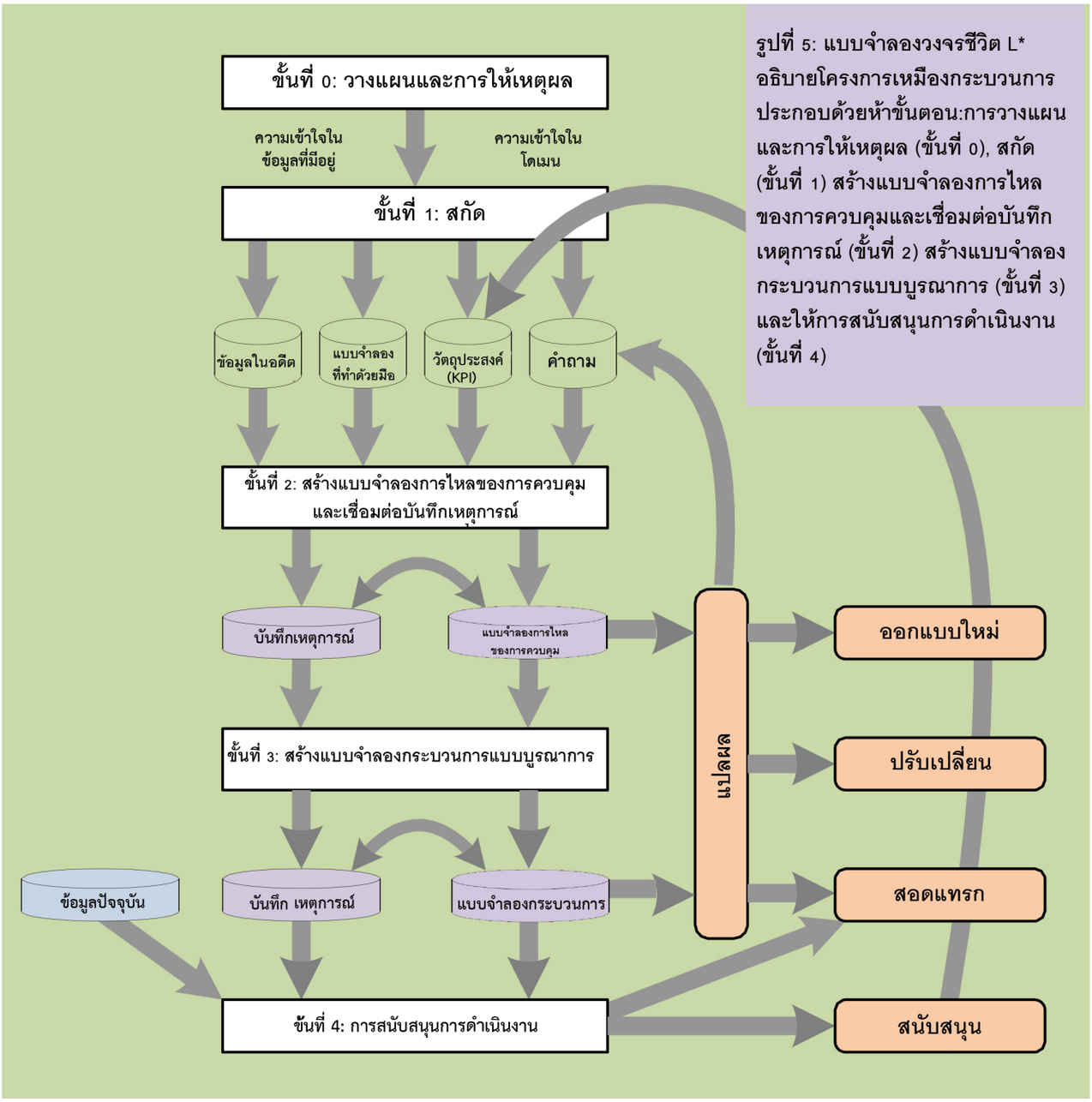
GP4: เหตุการณ์ควรมีความสัมพันธ์กับองค์ประกอบของแบบจำลอง

GP5: แบบจำลองควรได้รับการปฏิบัติในแบบที่เป็นนามธรรมที่มีความหมายตามความเป็นจริง

GP6: เหมืองกระบวนการควรเป็นกระบวนการที่ต่อเนื่อง

นอกจากนั้น บันทึกเหตุการณ์อาจถูกกรองหรือปรับให้เหมาะสมโดยใช้แบบจำลอง (เช่น การลบกิจกรรมที่เกิดขึ้นนานๆครั้งหรือกรณีที่ผิดปกติออกไป, และ การแทรกเหตุการณ์ที่ขาดหายไป) ในบางครั้งต้องใช้ความพยายามอย่างมากในการหาเหตุการณ์ที่มีความสัมพันธ์ร่วมกัน ของกระบวนการเดียวกัน ณ ขณะนั้น เหตุการณ์ที่เหลือมีความสัมพันธ์กับเอนทิตีของแบบจำลองกระบวนการ เมื่อกระบวนการค่อนข้างมีโครงสร้าง แบบจำลองการไหลของการควบคุมอาจถูกขยายออกไปใช้กับมุมมองอื่นๆได้ (เช่น ข้อมูล, เวลา, และ ทรัพยากร) ในระหว่างขั้นที่ 3 ความสัมพันธ์ระหว่างบันทึกเหตุการณ์และแบบจำลองที่สร้างขึ้นในขั้นที่ 2 จะถูกนำมาใช้เพื่อขยายแบบจำลอง (เช่น ประทับเวลาของเหตุการณ์ที่สัมพันธ์กัน จะถูกใช้ในการประมาณเวลาที่ต้องรอคอยสำหรับกิจกรรม) สิ่งเหล่านี้อาจใช้เป็นคำตอบของคำถามเพิ่มเติม และอาจจะกระตุ้นการกระทำอื่นๆเพิ่มเติมได้ ท้ายที่สุดแบบจำลองที่ถูกสร้างขึ้นในขั้นที่ 3 อาจถูกนำมาใช้สำหรับสนับสนุนการดำเนินการ (ขั้นที่ 4) ความรู้ที่สกัดมาได้จากข้อมูลเหตุการณ์ในอดีตจะถูกนำมารวมเข้ากับสารสนเทศของกรณีที่กำลังทำงานอยู่ สิ่งเหล่านี้อาจจะถูกนำไปใช้ในการสอดแทรก, พยากรณ์, และแนะนำ ขั้นที่ 3 และ 4 จะมาถึงได้ต่อได้ก็ต่อเมื่อกระบวนการมีความเสถียร และ มีโครงสร้างอย่างเพียงพอ

ปัจจุบันมีเทคนิคและเครื่องมือที่สามารถสนับสนุนขั้นตอนทั้งหมดที่แสดงในรูปที่ 5 อย่างไรก็ตาม เหมืองกระบวนการยังถือว่าเป็นกระบวนทัศน์ใหม่ และเครื่องมือที่มีอยู่ขณะนี้เกือบทั้งหมดยังค่อนข้างไม่สมบูรณ์ นอกจากนั้น ผู้ใช้ที่มุ่งหวังไว้ส่วนใหญ่ไม่ได้ตระหนักถึงศักยภาพที่สามารถพัฒนาได้และข้อจำกัดของการทำเหมืองกระบวนการ



รูปที่ 5: แบบจำลองวงจรชีวิต L* อธิบายโครงการเหมืองกระบวนการประกอบด้วยห้าขั้นตอน:การวางแผนและการให้เหตุผล (ขั้นที่ 0), สกัด (ขั้นที่ 1) สร้างแบบจำลองการไหลของการควบคุมและเชื่อมต่อบันทึกเหตุการณ์ (ขั้นที่ 2) สร้างแบบจำลองกระบวนการแบบบูรณาการ (ขั้นที่ 3) และให้การสนับสนุนการดำเนินงาน (ขั้นที่ 4)

ดังนั้นเจตนารมณ์นี้จึงได้แสดงรายการบางส่วนของ หลักการแนะนำ (ในหัวข้อถัดไป) และความท้าทาย (หน้าที่ 10) สำหรับผู้ใช้เทคนิคของเหมืองกระบวนการ นักวิจัยและนักพัฒนาที่สนใจความทันสมัย

3. หลักการแนะนำ

เช่นเดียวกันกับเทคโนโลยีใหม่ใดๆ มักจะมีความผิดพลาดที่เห็นได้อย่างชัดเจนที่เกิดขึ้นเมื่อนำเหมืองกระบวนการไปประยุกต์ใช้กับชีวิตจริง ดังนั้น เราจึงได้ให้รายการหลักการแนะนำ 6 ประการเพื่อป้องกัน ผู้ใช้/นักวิเคราะห์ จากการกระทำที่ผิดพลาดเหล่านี้

GP1: ข้อมูลเหตุการณ์ควรได้รับการปฏิบัติอย่างประชาชนชั้นหนึ่ง

จุดเริ่มต้นของกิจกรรมสำหรับเหมืองกระบวนการใดๆคือการบันทึกเหตุการณ์ เราเรียกชุดของเหตุการณ์ว่าบันทึกเหตุการณ์ อย่างไรก็ตาม มันก็ไม่ได้มีนัยว่าเหตุการณ์จะต้องถูกจัดเก็บไว้ในแฟ้มบันทึกโดยเฉพาะ เหตุการณ์อาจจะถูกจัดเก็บไว้ในตารางในฐานข้อมูล บันทึกข้อความ ที่จัดเก็บจดหมายอย่างถาวร, บันทึกธุรกรรม, และแหล่งข้อมูลอื่นๆ ที่สำคัญกว่ารูปแบบหน่วยจัดเก็บก็คือคุณภาพของบันทึกเหตุการณ์ คุณภาพของผลลัพธ์ของเหมืองกระบวนการขึ้นอยู่กับข้อมูลเข้าเป็นสำคัญ ดังนั้น บันทึกเหตุการณ์ควรจะได้รับการปฏิบัติอย่างประชาชนชั้นหนึ่งในระบบสารสนเทศที่สนับสนุนกระบวนการต่างๆที่ต้องการวิเคราะห์โชคไม่ดีที่ บันทึกเหตุการณ์ส่วนใหญ่เป็นเพียง “ผลพลอยได้” ที่ถูกใช้สำหรับแก้จุดบกพร่องหรือโพรไฟล์ลิ่ง ตัวอย่างเช่น เครื่องมือแพทย์ของ Philips Healthcare มีการบันทึกเหตุการณ์เพียงเพราะว่าผู้พัฒนาซอฟต์แวร์ได้แทรก “คำสั่งพิมพ์” เข้าไปในรหัสคำสั่ง ถึงแม้ว่าจะมีข้อแนะนำอย่างไม่เป็นทางการในการแทรกคำสั่งเหล่านี้เข้าไปในรหัสคำสั่ง แต่ยังคงมีความต้องการวิธีการที่เป็นระบบเพื่อเพิ่มคุณภาพของบันทึกเหตุการณ์ ข้อมูลเหตุการณ์จะต้องถูกมองว่าเป็นประชาชนชั้นหนึ่ง (มากกว่าเป็นประชาชนชั้นสอง)

ระดับ	คุณลักษณะ	ตัวอย่าง
★★★★★	ระดับสูงสุด: บันทึกเหตุการณ์มีคุณภาพยอดเยี่ยม (น่าไว้วางใจ และสมบูรณ์) และเหตุการณ์มีการกำหนดอย่างชัดเจน เหตุการณ์ถูกบันทึกอย่าง อัตโนมัติ เป็นระบบ น่าเชื่อถือ และปลอดภัย มีการพิจารณาความเป็นส่วนตัวและความมั่นคงอย่างเหมาะสม ยิ่งไปกว่านั้น เหตุการณ์ที่บันทึก (และแอตทริบิวต์ทั้งหมด) มีความหมายที่ชัดเจน บอกเป็นนัยว่ามีออนทอลโลจีอยู่หนึ่งหรือมากกว่า เหตุการณ์และแอตทริบิวต์ชี้ไปยังออนทอลโลจีนี้	บันทึกหมายเหตุประกอบทางความหมายของของระบบ BPM
★★★★	เหตุการณ์ถูกบันทึกอย่าง อัตโนมัติ เป็นระบบ และน่าเชื่อถือ บันทึกน่าไว้วางใจและสมบูรณ์ ไม่เหมือนกับการดำเนินการในระดับ ★★★ คือ แนวคิด เช่น กระบวนการ ณ ขณะนั้น (กรณี) และกิจกรรมได้รับการสนับสนุนอย่างเปิดเผย	บันทึกเหตุการณ์ของระบบ BPM/ workflow
★★★	เหตุการณ์ถูกบันทึกอย่างอัตโนมัติ แต่ไม่มีวิธีการที่เป็นระบบในการบันทึกเหตูการณ์ อย่างไรก็ตาม ไม่เหมือนกับบันทึกในระดับ ★★ คือมีการรับประกันในบางระดับว่าเหตุการณ์ที่บันทึกตรงกับความเป็นจริง (บันทึกเหตุการณ์น่าไว้วางใจแต่ไม่จำเป็นว่าสมบูรณ์) ตัวอย่าง เช่น เหตุการณ์ที่บันทึกโดยระบบ ERP ถึงแม้ว่าเหตุการณ์ต้องถูกสกัดจากหลากหลายตาราง สารสนเทศสามารถถูกสมมุติว่าถูกต้องได้ (มันเป็นการปลอดภัยที่สมมุติว่าบันทึกการจ่ายเงินโดย ERP มีอยู่จริงและในทางกลับกันด้วย)	ตารางในระบบ ERP, บันทึกเหตุการณ์ของระบบ CRM, บันทึกธุรกรรมของระบบข่าวสาร, บันทึกเหตุการณ์ของระบบ high-tech เป็นต้น
★★	เหตุการณ์ถูกบันทึกอย่างอัตโนมัติ เป็นผลพลอยได้จากระบบสารสนเทศบางระบบ การครอบคลุมที่ผันแปร ไม่มีวิธีการที่เป็นระบบให้ติดตามในการตัดสินใจว่าเหตุการณ์ใดควรถูกบันทึก ยิ่งไปกว่านั้น มันมีความเป็นไปได้ที่จะเลี่ยงการใช้ระบบสารสนเทศ ดังนั้น เหตุการณ์อาจจะหายไปหรือไม่ได้ถูกบันทึกอย่างถูกต้อง	บันทึกเหตุการณ์ของเอกสารและระบบบริหารผลิตภัณฑ์ บันทึกความผิดพลาดของระบบฝังตัว ตารางทำการของวิศกรบริการ เป็นต้น
★	ระดับต่ำสุด: บันทึกเหตุการณ์มีคุณภาพต่ำ เหตุการณ์ที่บันทึกไว้อาจไม่สอดคล้องกับความเป็นจริงและเหตุการณ์อาจหายไป บันทึกเหตุการณ์ที่เหตุการณ์ถูกบันทึกด้วยมือโดยทั่วไปจะมีคุณลักษณะนี้	ร่องรอยที่เหลืออยู่ในเอกสารกระดาษที่ส่งต่อกันในองค์กร (“yellow notes”), บันทึกทางการแพทย์ด้วยกระดาษ.

ตารางที่ 1: ระดับวุฒิภาวะสำหรับบันทึกเหตุการณ์

มีหลายเงื่อนไขที่ใช้ในการตัดสินคุณภาพของบันทึกเหตุการณ์ เหตุการณ์ควรจะน่าไว้วางใจ มันควรจะปลอดภัยที่จะสมมุติว่าเหตุการณ์ที่ถูกบันทึกเกิดขึ้นจริงและแอตทริบิวต์ของเหตุการณ์ถูกต้อง บันทึกเหตุการณ์ควรจะสมบูรณ์ เช่น มีขอบเขตที่เฉพาะเจาะจง ไม่มีเหตุการณ์ที่อาจจะหายไป เหตุการณ์ใดๆที่ถูกบันทึกไว้ควรจะมีการระบุความหมายที่ชัดเจน มากไปกว่านั้น ข้อมูลเหตุการณ์ควรจะปลอดภัยด้วยความตระหนักถึงความเป็นส่วนตัว และ ความมั่นคง เมื่อมีการบันทึกเหตุการณ์ ตัวอย่างเช่น นักแสดง ควรจะตระหนักถึงชนิดของเหตุการณ์ที่กำลังถูกบันทึกและวิธีการใช้ข้อมูลเหล่านี้

ตารางที่ 1 กำหนดวุฒิภาวะของบันทึกเหตุการณ์ 5 ระดับ จากคุณภาพยอดเยี่ยม (★★★★★) ถึงคุณภาพต่ำ (★) ตัวอย่างเช่น บันทึกเหตุการณ์ของ Philips Healthcare อยู่ในระดับ ★★★ กล่าวคือเหตุการณ์ได้รับการบันทึกอย่างอัตโนมัติและพฤติกรรมที่ถูกบันทึกตรงกับ ความเป็นจริง แต่ไม่มีวิธีการที่เป็นระบบที่ใช้ในการกำหนดความหมายให้กับเหตุการณ์ และให้ความแน่ใจว่าจะครอบคลุมถึงในระดับที่ต้องการ เทคนิคของเหมืองกระบวนการสามารถประยุกต์ใช้กับบันทึกในระดับ ★★★★★, ★★★★ และ ★★★ โดยหลักการ มันมีความเป็นไปได้ที่จะประยุกต์ใช้เหมืองกระบวนการกับบันทึกเหตุการณ์ในระดับ ★★ หรือ ★ แต่อย่างไรก็ตาม การวิเคราะห์บันทึกเหล่านี้มักมีปัญหาและผลลัพธ์ที่ได้ขาดความน่าไว้วางใจ ในความเป็นจริง มันเป็นเรื่องที่ไม่สมเหตุสมผลนักที่จะประยุกต์ใช้เหมืองกระบวนการกับบันทึกในระดับ ★

เพื่อที่จะให้ได้รับประโยชน์จากเหมืองกระบวนการ องค์กรควรตั้งเป้าหมายที่บันทึกเหตุการณ์ที่ระดับสูงที่สุดเท่าที่จะเป็นไปได้

GP2: การสกัดบันทึกควรถูกผลักดันจากคำถาม

ดังที่แสดงในรูปที่ 5 กิจกรรมของเหมืองกระบวนการจำเป็นต้องถูกผลักดันด้วยคำถาม หากปราศจากคำถามที่ชัดเจน มันเป็นการยากมากที่จะสกัดข้อมูลเหตุการณ์ที่มีความหมายได้ ตัวอย่างเช่น ตารางนับพันตารางในฐานข้อมูลของระบบ ERP เช่น SAP หากปราศจากคำถามที่ชัดเจน มันเป็นไปไม่ได้เลยที่จะเลือกตารางที่มีความเกี่ยวเนื่องสำหรับสกัดข้อมูล

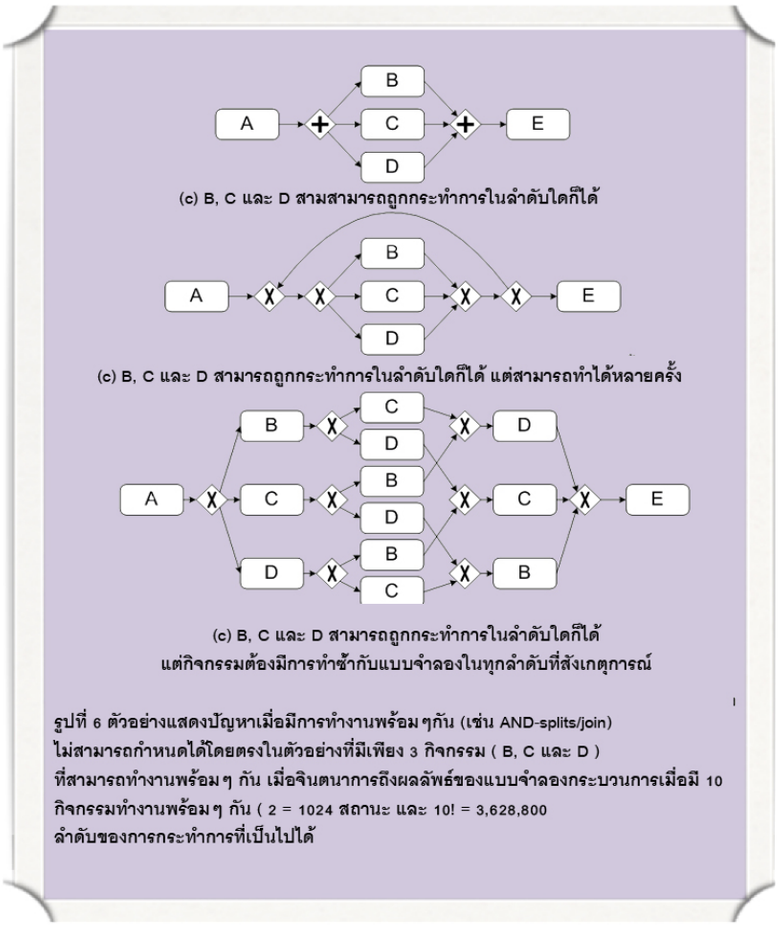
แบบจำลองกระบวนการดังเช่นที่แสดงในรูปที่ 1 อธิบายถึงวงจรชีวิตของกรณี (กระบวนการ ณ ขณะนั้น) ของชนิดหนึ่งโดยเฉพาะ ดังนั้น ก่อนที่จะประยุกต์เทคนิคของเหมืองกระบวนการใดๆ เราจะต้องทำการเลือกชนิดของกรณีที่ต้องการทำการวิเคราะห์ ทางเลือกนี้ควรได้รับการผลักดันจากคำถามที่ต้องการคำตอบ

และควรจะไม่ใช่เรื่องเล็กน้อย ตัวอย่างเช่น การจัดการคำสั่งซื้อของลูกค้า คำสั่งซื้อของลูกค้าแต่ละใบอาจประกอบด้วยหลายรายการ เนื่องจากลูกค้าอาจสั่งซื้อหลายผลิตภัณฑ์ในคำสั่งซื้อหนึ่งใบ คำสั่งซื้อของลูกค้าหนึ่งใบอาจมีการจัดส่งได้หลายครั้ง ในการจัดส่งหนึ่งครั้งอาจอ้างอิงถึงรายการสั่งซื้อในคำสั่งซื้อหลายใบ ดังนั้นจึงมีความสัมพันธ์แบบ many-to-many ระหว่างคำสั่งซื้อกับการจัดส่ง และความสัมพันธ์แบบ one-to-many ระหว่างคำสั่งซื้อกับรายการสั่งซื้อ ในฐานข้อมูลที่กำหนดประกอบด้วยข้อมูลเหตุการณ์ที่เกี่ยวกับคำสั่งซื้อ รายการสั่งซื้อ และการจัดส่ง มีแบบจำลองกระบวนการหลายแบบที่แตกต่างกันที่สามารถค้นพบได้ บางคนอาจสกัดข้อมูลโดยมีเป้าหมายที่จะอธิบายถึงวงจรชีวิตของคำสั่งซื้อแต่ละใบ อย่างไรก็ตาม มันก็ยังมีความเป็นไปได้ที่จะสกัดข้อมูลโดยมีเป้าหมายที่ต้องการค้นพบวงจรชีวิตของแต่ละรายการสั่งซื้อ หรือ วงจรชีวิตของการจัดส่งแต่ละครั้ง

GP3: การทำงานพร้อมๆกัน ทางเลือก และโครงสร้างการไหลของการควบคุมพื้นฐานควรได้รับการสนับสนุน

มีภาษาของแบบจำลองกระบวนการอยู่มากมาย (เช่น BPMN, EPCs, Petri nets, BPEL, และ UML activity diagrams) บางภาษามีองค์ประกอบของแบบจำลองอยู่มากมาย (เช่น BPMN มีมากว่า 50 องค์ประกอบด้านกราฟิกที่ต่างกัน) ในขณะที่ภาษาที่พื้นฐานมากๆ (เช่น Petri nets มีองค์ประกอบที่แตกต่างกันเพียง 3 องค์ประกอบ คือ places, transitions, และ arcs) การอธิบายด้วยการไหลของการควบคุมถือได้ว่าเป็นแกนหลักของแบบจำลองกระบวนการใดๆ โครงสร้างการไหลของงานพื้นฐาน (ที่รู้จักกันในชื่อที่เรียกว่า รูปแบบ) ที่สนับสนุนโดยภาษาหลักๆ ได้แก่ ลำดับ การจัดเส้นทางอย่างขนาน (AND-splits/joins), การเลือก (XORsplits / joins), และ ทำซ้ำ ซึ่งเห็นได้อย่างชัดเจนว่ารูปแบบเหล่านี้ควรจะได้รับการสนับสนุนจากเทคนิคเหมืองกระบวนการ อย่างไรก็ตาม บางเทคนิคไม่สามารถจัดการการทำงานพร้อมกันได้ และสนับสนุนเฉพาะ Markov chains/transition systems

รูปที่ 6 แสดงผลกระทบของการใช้เทคนิคของเหมืองกระบวนการที่ไม่สามารถค้นพบการทำงานพร้อมๆกัน (no AND-split/joins) ลองพิจารณาบันทึกเหตุการณ์ $L=\{\langle A,B,C,D,E\rangle, \langle A,B,D,C,E\rangle, \langle A,C,B,D,E\rangle, \langle A,C,D,B,E\rangle, \langle A,D,B,C,E\rangle, \langle A,D,C,B,E\rangle\}$.



รูปที่ 6 ตัวอย่างแสดงปัญหาเมื่อมีการทำงานพร้อม ๆกัน (เช่น AND-splits/join) ไม่สามารถกำหนดได้โดยตรงในตัวอย่างที่มีเพียง 3 กิจกรรม (B, C และ D) ที่สามารถทำงานพร้อม ๆ กัน เมื่อจินตนาการถึงผลลัพธ์ของแบบจำลองกระบวนการเมื่อมี 10 กิจกรรมทำงานพร้อม ๆ กัน (2 = 1024 สถานะ และ 10! = 3,628,800 ลำดับของการกระทำการที่เป็นไปได้

L ประกอบด้วยกรณีต่างๆที่เริ่มต้นด้วย A และจบด้วย E กิจกรรม B, C, และ D เกิดขึ้นระหว่าง A และ E ในลำดับแบบใดก็ได้ แบบจำลอง BPMN ในรูปที่ 6(a) แสดงการแทนกระบวนการเหล่านี้ด้วยการใช้ AND gateways 2 ตัว สมมุติว่าเทคนิคของเหมืองกระบวนการไม่สนับสนุน AND gateways ในกรณีนี้ แบบจำลอง BPMN อื่นอีกสองแบบในรูปที่ 6 จึงน่าจะเป็นตัวเลือกที่เห็นได้อย่างชัดเจน แบบจำลอง BPMN ในรูปที่ 6(b) กะทัดรัด แต่ยอมให้มีพฤติกรรมได้มากมายจนเกินไป (เช่นกรณีของ $\langle A,B,B,B,E\rangle$ เป็นไปได้ตามแบบจำลองแต่ไม่น่าเกิดขึ้นตามบันทึกเหตุการณ์). แบบจำลอง BPMN ในรูปที่ 6(c) อนุญาตสำหรับกรณีใน L แต่มีการเข้ารหัสของลำดับทั้งหมดอย่างชัดเจน จึงไม่เป็นตัวแทนที่กะทัดรัดในการใช้แทนบันทึก ตัวอย่างนี้แสดงให้เห็นว่าสำหรับแบบจำลองในชีวิตจริง ที่มีกิจกรรมที่ต้องทำงานพร้อมกันเป็นจำนวนมาก แบบจำลองผลลัพธ์ที่ได้จึงเป็นแบบจำลองที่หลวมเกินไป (underfitting) อย่างร้ายแรง (อนุญาตให้มีพฤติกรรมได้มากมายจนเกินไป) และ/หรือ มีความซับซ้อนมาก ถ้าไม่สนับสนุนการทำงานพร้อมๆกัน

ดังที่แสดงไว้ในรูปที่ 6 มันมีความสำคัญที่ต้องสนับสนุนอย่างน้อยที่สุดคือรูปแบบการไหลของงานพื้นฐาน นอกจากรูปแบบพื้นฐานที่ได้กล่าวมาแล้ว ยังต้องการให้สนับสนุน OR-splits/joins, เพราะว่าสิ่งเหล่านี้ให้ตัวแทนที่กะทัดรัดของการตัดสินใจอย่างครอบคลุม และการประสานเวลาบางส่วน

GP4: เหตุการณ์ควรสัมพันธ์กับองค์ประกอบของแบบจำลอง

ดังที่ได้ชี้ให้เห็นในส่วน 2 ว่ามีความเข้าใจผิดว่าเหมืองกระบวนการจำกัดอยู่เฉพาะการค้นพบการไหลของการควบคุม ดังแสดงในรูปที่ 1 แบบจำลองกระบวนการที่ค้นพบอาจครอบคลุมในหลายมุมมอง (มุมมองด้านองค์กร มุมมองด้านเวลา มุมมองด้านข้อมูล เป็นต้น) ยิ่งไปกว่านั้น การค้นพบยังเป็นเพียงหนึ่งในสามชนิดของเหมืองกระบวนการ ดังแสดงในรูปที่ 3 อีกสองชนิดของเหมืองกระบวนการ (การตรวจสอบความสอดคล้อง และ การปรับปรุงให้ดีขึ้น) ขึ้นอยู่กับความสัมพันธ์ระหว่างองค์ประกอบของ

แบบจำลองและเหตุการณ์ในบันทึกอย่างมาก ความสัมพันธ์นี้อาจถูกนำมาใช้ใน “การเล่นซ้ำ” ของบันทึกเหตุการณ์บนแบบจำลอง การเล่นซ้ำอาจถูกใช้เพื่อแสดงถึงความแตกต่างกันระหว่างบันทึกเหตุการณ์และแบบจำลอง เช่น บางเหตุการณ์ในบันทึกไม่สามารถเป็นไปได้ตามแบบจำลอง เทคนิคสำหรับตรวจสอบความสอดคล้องจะแสดงปริมาณและวินิจฉัยความแตกต่างนั้น ประทับเวลาในบันทึกเหตุการณ์สามารถนำมาใช้ในการวิเคราะห์พฤติกรรมเชิงเวลาในระหว่างการเล่นซ้ำได้ เวลาที่แตกต่างกันระหว่างสาเหตุที่เกี่ยวข้องกับกิจกรรมสามารถถูกนำมาใช้เพื่อเพิ่มเวลารอคอยที่คาดหวังในแบบจำลอง ตัวอย่างเหล่านี้แสดงว่าความสัมพันธ์ระหว่างเหตุการณ์ในบันทึกและองค์ประกอบในแบบจำลองทำหน้าที่เป็นจุดเริ่มต้นสำหรับชนิดของการวิเคราะห์ที่แตกต่างกัน

ในบางกรณีมันอาจไม่ใช่เรื่องเล็กน้อยที่จะสร้างความสัมพันธ์เหล่านั้น ตัวอย่างเช่น เหตุการณ์หนึ่งอาจมีการอ้างอิงถึงสองกิจกรรมที่แตกต่างกัน หรือมันอาจมีความไม่ชัดเจนว่าอ้างอิงถึงกิจกรรมใด ความกำกวมเหล่านี้ต้องถูกกำจัดออกไปเพื่อทำให้การแปลผลลัพธ์ของเหมืองกระบวนการถูกต้อง นอกจากปัญหาของความสัมพันธ์ระหว่างเหตุการณ์กับกิจกรรมแล้ว ยังมีปัญหาของความสัมพันธ์ของเหตุการณ์กับกระบวนการ ณ ขณะนั้น ที่เรียกว่า สหสัมพันธ์ของเหตุการณ์

GP5: แบบจำลองควรได้รับการปฏิบัติในแบบที่เป็นนามธรรมที่มีความหมายของความเป็นจริง

แบบจำลองที่ได้มาจากข้อมูลเหตุการณ์ให้มุมมองของความเป็นจริง เป็นมุมมองที่ให้นามธรรมที่มีความหมายของพฤติกรรมที่ตรวจจับได้ในบันทึกเหตุการณ์ บันทึกเหตุการณ์ที่ให้มา อาจมีมุมมองที่เป็นประโยชน์ได้หลายมุมมอง ยิ่งไปกว่านั้น ผู้มีส่วนได้ส่วนเสียที่หลากหลายอาจมีความต้องการมุมมองที่แตกต่างกัน ในความเป็นจริง แบบจำลองที่ค้นพบจากบันทึกเหตุการณ์ ควรถูกมองว่าเป็น “แผนที่” (เช่นเดียวกับแผนที่ทางภูมิศาสตร์) หลักการแนะนำนี้ให้ความเข้าใจที่ลึกซึ้งอย่างสำคัญสองสิ่งจะได้อธิบายในส่วนที่เหลือต่อไป

ก่อนอื่น มันมีความสำคัญที่จะต้องหมายเหตุว่าไม่มีสิ่งที่เหมือนกับ “แผนที่” สำหรับพื้นที่ทางภูมิศาสตร์โดยเฉพาะ ขึ้นอยู่กับความต้องการในการใช้งาน มีแผนที่ที่แตกต่างกัน: แผนที่ถนน, แผนที่การเดินเขา, แผนที่การขี่จักรยาน เป็นต้น แผนที่เหล่านี้ทั้งหมดแสดงมุมมองของสิ่งที่เป็นจริงเดียวกัน และมันอาจเป็นสิ่งไร้สาระที่จะสมมุติว่ามีสิ่งที่เรียกว่า “แผนที่ที่สมบูรณ์” เช่นเดียวกันกับแบบจำลองกระบวนการ: แบบจำลองควรจะเน้นเฉพาะสิ่งที่ตรงประเด็นกับชนิดของผู้ใช้อย่างเฉพาะเจาะจง แบบจำลองที่ค้นพบอาจเน้นในมุมมองที่แตกต่างกัน (การไหลของการควบคุม, การไหลของข้อมูล, เวลา, ทรัพยากร, ค่าใช้จ่าย เป็นต้น) และแสดงสิ่งเหล่านี้ในระดับความละเอียด และความแม่นยำ ที่แตกต่างกัน เช่น ผู้จัดการอาจต้องการเห็นแบบจำลองที่ไม่เป็นทางการอย่างคร่าวๆโดยเน้นไปที่ค่าใช้จ่าย ในขณะที่นักวิเคราะห์กระบวนการอาจต้องการเห็นแบบจำลองกระบวนการอย่างละเอียดโดยเน้นไปที่ความเบี่ยงเบนจากการไหลปกติ มีข้อสังเกตอีกว่า ผู้มีส่วนได้ส่วนเสียที่แตกต่างกันอาจต้องการมุมมองของกระบวนการที่ระดับที่แตกต่างกัน: ระดับกลยุทธ์ (การตัดสินใจในระดับนี้มีผลกระทบในระยะยาวและอยู่บนพื้นฐานของยอดรวมของข้อมูลเหตุการณ์ในช่วงระยะยาว), ระดับกลวิธี (การตัดสินใจในระดับนี้มีผลกระทบในระยะปานกลาง และส่วนใหญ่ขึ้นอยู่กับข้อมูลที่ได้มาเร็วๆนี้), และ ระดับดำเนินการ (การตัดสินใจในระดับนี้มีผลกระทบแบบทันทีทันใด และขึ้นอยู่กับข้อมูลที่เกี่ยวข้องกับกรณีที่กำลังทำงานอยู่)

ประการที่สอง มันมีประโยชน์ที่จะรับเอาแนวคิดจากการทำแผนที่ เมื่อมันสามารถสร้างแผนที่ๆสามารถเข้าใจได้ ตัวอย่างเช่น นามธรรมของแผนที่ถนนจากการลดนัยสำคัญของถนนและเมืองให้น้อยลง การลดนัยสำคัญของสิ่งต่างๆให้น้อยลงอาจทำโดยการเอาส่วนนั้นออกไป หรือ ด้วยการรวมกลุ่มให้อยู่ในรูปของผลรวม (เช่น ถนน และชานเมือง รวมกันเป็นเมือง) นักทำแผนที่ นอกจากกำจัดรายละเอียดที่ไม่ตรงประเด็นแล้ว ยังมีการใช้สีเพื่อเน้นข้อมูลที่สำคัญ ยิ่งไปกว่านั้นในองค์ประกอบทางกราฟิกมีการใช้ขนาดเพื่อบ่งบอกนัยสำคัญ (เช่น ขนาดของเส้น และจุดอาจแปรผัน) แผนที่ทางภูมิศาสตร์นั้นมีความชัดเจนในการแปลผลของแกน X และแกน Y การวางผังของแผนที่ ไม่ได้ทำโดยอำเภอใจ เนื่องจากพิกัดขององค์ประกอบมีความหมาย ทั้งหมดที่กล่าวมาเป็นสิ่งที่ตายตัว ซึ่งมีความแตกต่างกับกระแสหลักของแบบจำลองกระบวนการ ที่โดยทั่วไปจะไม่ใช้สี ขนาด และลักษณะของตำแหน่งมาสร้างแบบจำลองเพื่อให้มีความเข้าใจได้มากขึ้น แต่อย่างไรก็ตาม แนวคิดจากการทำแผนที่สามารถนำเข้ามารวมกันโดยง่ายในการสร้างแผนที่กระบวนการที่ค้นพบ ตัวอย่างเช่น ขนาดของกิจกรรมสามารถถูกใช้เพื่อสะท้อนถึงความถี่หรือคุณสมบัติอื่นๆที่บ่งชี้ถึงนัยสำคัญ (เช่น ค่าใช้จ่าย หรือ ทรัพยากรที่ใช้) ความกว้างของเส้นสามารถสะท้อนถึงความสำคัญของการขึ้นต่อกันเชิงสาเหตุที่สอดคล้องกัน และสีของเส้นสามารถใช้ในการเน้นจุดที่เกิดปัญหาคอขวด

ข้อสังเกตที่กล่าวมาแล้วข้างต้นแสดงให้เห็นว่า มันมีความสำคัญในการเลือกตัวแทนที่ถูกต้อง และการปรับแต่งอย่างละเอียด สำหรับกลุ่มผู้ฟังเป้าหมาย มันมีความสำคัญในการแสดงผลลัพธ์ด้วยมโนภาพสำหรับผู้ใช้ และ สำหรับการแนะนำอัลกอริทึมของการค้นพบไปสู่แบบจำลองที่เหมาะสม (ดูที่ความท้าทาย C5)

GP6: เหมืองกระบวนการควรเป็นกระบวนการที่ต่อเนื่อง

เหมืองกระบวนการสามารถช่วยให้มี “แผนที่” ที่มีความหมาย ที่เชื่อมต่อโดยตรงกับข้อมูลเหตุการณ์ทั้งข้อมูลเหตุการณ์ในอดีตและข้อมูลปัจจุบันสามารถฉายภาพลงในแบบจำลอง ยิ่งไปกว่านั้นอาจมีการเปลี่ยนแปลงกระบวนการในขณะที่กำลังทำการวิเคราะห์ ด้วยความมีพลวัติโดยธรรมชาติของกระบวนการ จึงไม่แนะนำให้มองเหมืองกระบวนการเป็นกิจกรรมที่ทำเพียงครั้งเดียว เป้าหมายไม่ควรเป็นการสร้างแบบจำลองที่คงที่ แต่ควรทำให้แบบจำลองกระบวนการมีชีวิตชีวา เพื่อให้ผู้ใช้และนักวิเคราะห์เข้าไปดูมันทุกวัน

เปรียบสิ่งนี้ได้กับการที่ mashups โดยใช้ geo-tagging มี mashups นับพันที่ใช้ Google Maps (เช่น แอปพลิเคชันของการฉายภาพสารสนเทศเกี่ยวกับสภาพการจราจร อสังหาริมทรัพย์, ร้านอาหารจานด่วน, หรือ เวลาของภาพยนต์ลงในแผนที่) ประชาชนสามารถซูมเข้าและซูมออกได้อย่างไร้รอยต่อโดยการใช้แผนที่และปฏิสัมพันธ์กับมัน (เช่น การจราจรติดขัดถูกฉายภาพลงบนแผนที่และผู้ใช้สามารถเลือกปัญหาเฉพาะเพื่อดูรายละเอียด) มันควรจะมีความเป็นไปได้ในการสร้างเหมืองกระบวนการจากข้อมูลเหตุการณ์แบบเวลาจริง ด้วยการใช้ “อุปมาของแผนที่” เราสามารถคิดถึงเหตุการณ์ที่มีพิกัด GPS ที่สามารถฉายภาพลงบนแผนที่แบบเวลาจริง คล้ายคลึงกันกับระบบนำทางของรถยนต์ เครื่องมือของเหมืองกระบวนการสามารถช่วยผู้ใช้ (a) โดยการนำทางผ่านกระบวนการ (b) โดยฉายภาพสารสนเทศพลวัติลงบนแผนที่กระบวนการ (ตัวอย่างเช่น แสดง “การติดขัดของการจราจร” ในกระบวนการธุรกิจ) และ (c) โดยการพยากรณ์กรณีที่กำลังทำงานอยู่ (เช่น การคาดคะเน “เวลาที่มาถึง” ของกรณีที่มีการล่าช้า) ตัวอย่างเหล่านี้แสดงให้เห็นว่ามันจะเป็นสิ่งที่น่าเสียดายหากไม่ใช้เหมืองกระบวนการอย่างกระตือรือร้นให้มากขึ้น ดังนั้น เหมืองกระบวนการควรถูกมองว่าเป็นกระบวนการอย่างต่อเนื่อง ที่ให้สารสนเทศที่สามารถนำไปทำงานได้ตามมาตรเวลาที่หลากหลาย (นาที ชั่วโมง วัน สัปดาห์ และ เดือน)

4. ความท้าทาย

เหมืองกระบวนการเป็นเครื่องมือที่สำคัญสำหรับองค์กรสมัยใหม่ที่ต้องการบริหารจัดการกระบวนการดำเนินการที่ไม่ใช่เรื่องเล็กน้อย ในมุมหนึ่งมีการเติบโตของข้อมูลเหตุการณ์เพิ่มขึ้นอย่างไม่น่าเชื่อ ส่วนในอีกมุมหนึ่ง กระบวนการและสารสนเทศต้องเข้ากันได้และเป็นไปในทิศทางเดียวกันอย่างสมบูรณ์เพื่อให้เป็นไปตามความต้องการที่เกี่ยวข้องกับการปฏิบัติให้เป็นไปตามข้อกำหนด ประสิทธิภาพ และ การบริการลูกค้า นอกจากประโยชน์ที่มีอย่างกว้างขวางของเหมืองกระบวนการแล้ว ยังมีความท้าทายที่สำคัญที่ต้องกล่าวถึง: สิ่งเหล่านี้แสดงให้เห็นว่าเหมืองกระบวนการเป็นศาสตร์ที่เกิดขึ้นใหม่ ในส่วนที่เหลือ ได้แสดงรายการของความท้าทายบางส่วน รายการที่แสดงนี้ไม่ได้ตั้งใจให้สมบูรณ์ และตามกาลเวลา ความท้าทายใหม่ๆอาจเกิดขึ้น หรือ ความท้าทายที่มีอยู่อาจหายไปเนื่องจากความก้าวหน้าของเหมืองกระบวนการ

C1: การค้นหา การรวมเข้าด้วยกัน และ การทำความสะอาดข้อมูลเหตุการณ์

มันยังต้องใช้ความพยายามพอสมควรที่จะสกัดข้อมูลเหตุการณ์ที่เหมาะสมสำหรับเหมืองกระบวนการ โดยทั่วไปมีสิ่งกีดขวางหลายอย่างที่จะต้องเอาชนะ:

- ข้อมูลอาจกระจายอยู่ตามแหล่งข้อมูลที่หลากหลาย สารสนเทศนี้ต้องถูกนำมารวมเข้าด้วยกัน แนวโน้มนี้ก่อให้เกิดปัญหาเมื่อมีการใช้ตัวระบุตัวตน ที่แตกต่างกันในแหล่งข้อมูลที่แตกต่างกัน ตัวอย่างเช่น ระบบหนึ่งใช้ชื่อและวันเกิดเป็นตัวระบุตัวตนของบุคคล ในขณะที่ระบบอื่นๆใช้เลขบัตรประจำตัวประชาชน
- ข้อมูลเหตุการณ์โดยทั่วไปจะมี “ออฟเจ็กเป็นศูนย์กลาง” มากกว่าที่จะเป็น “กระบวนการเป็นศูนย์กลาง” ตัวอย่างเช่น ผลิตภัณฑ์แต่ละชิ้น pallets และ containers อาจจะมี RFID tag และบันทึกเหตุการณ์ที่เกี่ยวกับ tag นี้ อย่างไรก็ตาม การเฝ้าสังเกตุคำสั่งซื้อของลูกค้าโดยเฉพาะเจาะจงเหตุการณ์ที่เป็นออฟเจ็กเป็นศูนย์กลางจำเป็นต้องถูกนำมารวมกันและต้องมีการประมวลผลล่วงหน้า
- ข้อมูลเหตุการณ์อาจไม่สมบูรณ์ ปัญหาที่พบโดยทั่วไปคือเหตุการณ์ไม่ได้ชี้ไปที่กระบวนการ ณ ขณะนั้น อย่างชัดเจน บ่อยครั้งที่มีความเป็นไปได้ที่จะได้สารสนเทศนี้จากการสรุปมาจากข้อมูลอื่น แต่อาจต้องใช้ความพยายามพอสมควร เช่นเดียวกับสารสนเทศเกี่ยวกับเวลาอาจหายไปสำหรับบางเหตุการณ์ อาจมีความจำเป็นต้องสอดแทรกประทับเวลาเพื่อให้สารสนเทศเวลาที่มีอยู่ยังคงใช้งานได้
- บันทึกเหตุการณ์อาจมีความผิดปกติ อยู่ด้วย ได้แก่ พฤติกรรมพิเศษที่เรียกว่าสัญญาณรบกวนจะค้นหาความผิดปกติได้อย่างไร เราจะต้องตอบคำถามนี้เพื่อทำความสะอาดข้อมูลเหตุการณ์
- บันทึกอาจประกอบด้วยเหตุการณ์ที่มีความละเอียดที่แตกต่างกัน ในบันทึกเหตุการณ์ของระบบสารสนเทศโรงพยาบาล เหตุการณ์อาจหมายถึงเพียงการตรวจเลือด หรือขั้นตอนที่ซับซ้อนของการผ่าตัด ประทับเวลา อาจมีระดับความละเอียดที่แตกต่างกันจากความแม่นยำระดับมิลลิวินาที (28-9-2011:h11m28s32ms342) ถึงสารสนเทศแบบหยาบๆระดับวัน (28-9-2011).
- เหตุการณ์เกิดขึ้นในบริบทหนึ่งโดยเฉพาะ (อากาศ, ปริมาณงาน, วันในสัปดาห์ เป็นต้น) บริบทนี้อาจอธิบายถึงปรากฏการณ์ใดโดยเฉพาะ เช่น เวลาในการตอบสนองที่ยาวนานกว่าปกติเพราะว่า เป็นงานที่อยู่ระหว่างการผลิตหรือเป็นวันหยุด สำหรับการวิเคราะห์ต้องรวมเอาบริบทเหล่านี้เข้าไปด้วย สิ่งนี้บอกเป็นนัยว่าถึงการรวมข้อมูลเหตุการณ์เข้ากับข้อมูลบริบท “ภัยพิบัติของมิติ” เกิดขึ้นเมื่อการวิเคราะห์กลายมาเป็นสิ่งที่ไม่สามารถติดตามได้เมื่อมีการเพิ่มตัวแปรเข้าไปมากเกินไป

เครื่องมือและวิธีการที่ดีขึ้นเป็นสิ่งที่ต้องการเพื่อจัดการกับปัญหาที่กล่าวมาแล้วข้างต้น ยิ่งไปกว่านั้น ดังที่ได้ชี้ให้เห็นมาก่อนหน้านี้ องค์กรจำเป็นต้องปฏิบัติกับบันทึกเหตุการณ์อย่างประชาชนชั้นหนึ่ง ไม่ใช่เป็นเพียงผลพลอยได้ เป้าหมายคือต้องการให้ได้บันทึกเหตุการณ์ในระดับ ★★★★★ (ที่แสดงในตารางที่ 1) บทเรียนที่ได้เรียนรู้จากบริบทของคลังข้อมูลมีประโยชน์ที่จะทำให้แน่ใจว่าบันทึกเหตุการณ์มีคุณภาพสูง ตัวอย่างเช่น การตรวจสอบอย่างง่ายๆ ระหว่างการนำข้อมูลเข้าสามารถช่วยลดอัตราส่วนของข้อมูลเหตุการณ์ที่ไม่ถูกต้องได้อย่างมีนัยสำคัญ

C2: การจัดการกับบันทึกเหตุการณ์ที่ซับซ้อนมีหลากหลายคุณลักษณะ

บันทึกเหตุการณ์อาจมีคุณลักษณะที่แตกต่างกันมาก บางบันทึกเหตุการณ์อาจมีขนาดใหญ่มากทำให้ยากต่อการจัดการ ในขณะที่บางบันทึกเหตุการณ์อาจมีขนาดเล็กจนไม่มีข้อมูลที่พอเพียงต่อการสรุปผลที่ทำให้เกิดความน่าเชื่อถือ

ในบางโดเมนอาจมีการบันทึกเหตุการณ์ที่ยากที่จะเข้าใจได้ ดังนั้น จึงต้องมีความพยายามเพิ่มเติมเพื่อเพิ่มสมรรถภาพและความสามารถในการปรับขนาด ตัวอย่างเช่น ASML มีการเฝ้าสังเกตุอย่างต่อเนื่องของ wafer scanners ทั้งหมด wafer scanners เหล่านี้มีการใช้งานใน หลายองค์กร (เช่น, Samsung and Texas Instruments) เพื่อผลิต chips (ประมาณ 70% ของ chips ถูกผลิตโดยใช้ ASML’s wafer scanners) เครื่องมือที่มีอยู่ในปัจจุบันนั้นมีความยุ่งยากที่จะจัดการกับข้อมูลขนาด petabytes ที่เก็บในโดเมนนั้น นอกจากจำนวนของเหตุการณ์ที่บันทึกแล้ว ยังมีคุณลักษณะอื่นๆ เช่น

ความท้าทาย:

C1: การค้นหา, การรวมเข้าด้วยกันและการทำความสะอาดข้อมูลเหตุการณ์

C2: การจัดการกับบันทึกเหตุการณ์ที่ซับซ้อนมีคุณลักษณะที่หลากหลาย

C3: การสร้างตัวแทน เกณฑ์เปรียบเทียบสมรรถนะ

C4: การจัดการกับมโนทัศน์ที่เบี่ยงเบน

C5: การปรับปรุงความโน้มเอียงของตัวแทนที่ใช้สำหรับกระบวนการค้นพบ

C6: ความสมดุลระหว่างเงื่อนไขคุณภาพได้แก่ ความเหมาะสม, ความเรียบง่าย, ความแม่นยำ และนัยทั่วไป

C7: การทำเหมืองข้ามองค์กร

C8: การให้การสนับสนุนการดำเนินการ

C9: การรวมเหมืองกระบวนการเข้ากับการวิเคราะห์ชนิดอื่นๆ

C10: เพิ่มความสามารถในการใช้งานสำหรับผู้ใช้ที่ไม่ใช่ผู้เชี่ยวชาญ

C11:เพิ่มความสามารถในการเข้าใจสำหรับผู้ที่ไม่ใช่ผู้เชี่ยวชาญ

จำนวนเฉลี่ยของเหตุการณ์ต่อกรณี ความคล้ายคลึงกันระหว่างกรณี จำนวนของเหตุการณ์ที่มีลักษณะเฉพาะ และจำนวนเส้นทางเฉพาะ ลองพิจารณาบันทึกเหตุการณ์ L1 ที่มีคุณลักษณะดังต่อไปนี้: มี 1000 กรณี โดยเฉลี่ยมี 10 เหตุการณ์ต่อกรณี และมีการแปรผันเล็กน้อย (หลายกรณีมีเส้นทางเดียวกัน หรือมีเส้นทางที่คล้ายกันมาก) บันทึกเหตุการณ์ L2 ประกอบด้วย เพียง 100 กรณี แต่โดยเฉลี่ยมี 100 เหตุการณ์ต่อกรณี และกรณีทั้งหมดมีเส้นทางเป็นหนึ่งเดียว จะเห็นได้อย่างชัดเจนว่า L2 นั้นวิเคราะห์ได้ยากกว่า L1 ถึงแม้ว่าบันทึกทั้งสองมีขนาดเท่ากัน (ประมาณ 10,000 เหตุการณ์)

เนื่องจากบันทึกเหตุการณ์ประกอบด้วยพฤติกรรมตัวอย่างเท่านั้น จึงไม่ควรสมมุติว่ามันสมบูรณ์ เทคนิคของเหมืองกระบวนการจำเป็นต้องจัดการกับความไม่สมบูรณ์โดยการใช้ “สมมุติฐานโลกเปิด”: ความจริงที่ว่าบางสิ่งที่ยังไม่ได้เกิดขึ้น ไม่ได้หมายความว่ามันไม่สามารถเกิดขึ้นได้

ทำให้เกิดความท้าทายในการจัดการกับบันทึกเหตุการณ์ขนาดเล็กที่มีความผันแปรมากมาย

ดังที่ได้กล่าวมาแล้ว บางบันทึกประกอบด้วยเหตุการณ์ที่มีนามธรรมอยู่ในระดับต่ำมาก บันทึกเหล่านี้มีแนวโน้มว่าจะมีขนาดใหญ่มาก และแต่ละเหตุการณ์ในระดับต่ำนี้ได้รับความสนใจจากผู้มีส่วนได้ส่วนเสียน้อย ดังนั้นบางคนอาจต้องการรวมเหตุการณ์ในระดับต่ำเข้าด้วยกันให้เป็นเหตุการณ์ในระดับสูง ตัวอย่างเช่น เมื่อวิเคราะห์การวินิจฉัยและกระบวนการดูแลกลุ่มเฉพาะของผู้ป่วย เราอาจไม่สนใจข้อมูลที่บันทึกผลการทดสอบรายบุคคลในระบบสารสนเทศของห้องปฏิบัติการในโรงพยาบาล

ปัจจุบันนี้ องค์กรจำเป็นต้องใช้วิธีการลองผิดลองถูกเพื่อดูว่าบันทึกเหตุการณ์มีความเหมาะสมต่อการทำเหมืองกระบวนการหรือไม่ ดังนั้นเครื่องมือควรให้มีการตรวจสอบความเป็นไปได้ของข้อมูลอย่างเร็ว การทดสอบนี้ควรบ่งชี้แนวโน้มของปัญหาด้านสมรรถภาพและมีการแจ้งเตือนสำหรับบันทึกที่ยังห่างไกลจากความสมบูรณ์ หรือมีรายละเอียดมากเกินไป

C3: สร้างตัวแทนเกณฑ์เปรียบเทียบสมรรถนะ

เหมืองกระบวนการเป็นเทคโนโลยีที่เกิดขึ้นใหม่ ซึ่งเป็นการอธิบายว่าทำไมเกณฑ์การเปรียบเทียบสมรรถนะที่ดีจึงยังขาดหายไป ตัวอย่างเช่น เทคนิคในกระบวนการค้นพบมีเป็นจำนวนมากและผู้ขายก็นำเสนอผลิตภัณฑ์ที่แตกต่างกัน แต่ไม่มีสิ่งที่เห็นพ้องกันในเรื่องของคุณภาพของเทคนิคเหล่านี้ มีความแตกต่างกันอย่างมากมายในเรื่องของฟังก์ชั่น และสมรรถภาพ มันจึงเป็นการยากที่จะเปรียบเทียบเทคนิคและเครื่องมือที่ต่างกัน ดังนั้นเกณฑ์การเปรียบเทียบสมรรถนะที่ดีที่ประกอบไปด้วยตัวอย่างชุดข้อมูล และตัวแทนเงื่อนไขทางคุณภาพจำเป็นต้องมีการพัฒนาขึ้น

สำหรับเทคนิคของเหมืองข้อมูลดั้งเดิม มีเกณฑ์การเปรียบเทียบสมรรถนะที่ดีอยู่หลายตัว เกณฑ์การเปรียบเทียบสมรรถนะเหล่านี้กระตุ้นให้ผู้ขายและนักวิจัยปรับปรุงสมรรถภาพของเทคนิคของเขาให้ดีขึ้น ในกรณีของเหมืองกระบวนการสิ่งนี้ยิ่งมีความท้าทายมากกว่า ตัวอย่างเช่น แบบจำลองเชิงสัมพันธ์ที่ Codd ได้นำเสนอในปี 1969 ง่ายและได้รับการสนับสนุนอย่างกว้างขวาง ผลก็คือมันใช้ความพยายามเพียงเล็กน้อยในการแปลงข้อมูลจากฐานข้อมูลหนึ่งไปยังอีกฐานข้อมูลหนึ่งและไม่มีปัญหาในการแปลผล แต่สำหรับกระบวนการยังไม่มีแบบจำลองที่ง่าย มาตรฐานที่นำเสนอสำหรับแบบจำลองกระบวนการนั้นมีความซับซ้อนกว่ามากและมีผู้ขายไม่กี่รายสนับสนุนชุดของหลักการเดียวกัน กระบวนการโดยทั่วไปมีความซับซ้อนมากกว่าข้อมูลที่เป็นตาราง

อย่างไรก็ตาม มันมีความสำคัญในการสร้างตัวแทนเกณฑ์การเปรียบเทียบสมรรถนะสำหรับเหมืองกระบวนการ บางงานก็เริ่มมีให้ใช้แล้ว ตัวอย่างเช่น มีหลายมาตรวัดที่ใช้สำหรับการวัดคุณภาพของผลลัพธ์ของเหมืองกระบวนการ(ความเหมาะสม, ความเรียบง่าย, ความแม่นยำ, และนัยทั่วไป ยิ่งไปกว่านั้น มีบันทึกเหตุการณ์หลายอันให้ใช้แบบสาธารณะ (www.processmining.org) ตัวอย่างเช่นบันทึกเหตุการณ์ที่ใช้สำหรับ Business Process Intelligence Challenge (BPIC'11) ครั้งที่ 1 ที่จัดโดยคณะกรรมการเฉพาะกิจ (cf. doi:10.4121/ uuid:d9769f3d-0ab0-4fb8-803b-0d1120ffcf54).

ในมุมหนึ่ง ควรมีเกณฑ์การเปรียบเทียบสมรรถนะซึ่งอยู่บนพื้นฐานของชุดข้อมูลในชีวิตจริง ส่วนในอีกมุมหนึ่ง มีความต้องการในการสร้างชุดข้อมูลสังเคราะห์ที่ใช้ตรวจจับคุณลักษณะเฉพาะ ชุดข้อมูลสังเคราะห์ช่วยในการพัฒนาเทคนิคของเหมืองกระบวนการที่ปรับให้เหมาะสมกับบันทึกเหตุการณ์ที่ไม่สมบูรณ์ บันทึกเหตุการณ์ที่มีสัญญาณรบกวน หรือ กลุ่มประชากรเฉพาะของกระบวนการ

นอกจากการสร้างตัวแทนเกณฑ์การเปรียบเทียบสมรรถนะยังมีความจำเป็นที่ต้องเห็นพ้องต้องกันมากขึ้นในเงื่อนไขที่ใช้ในการตัดสินคุณภาพของผลลัพธ์ของเหมืองกระบวนการ (ดูที่ความท้าทาย C6) ยิ่งไปกว่านั้น การตรวจสอบความสมเหตุสมผลข้ามกลุ่มจากเหมืองข้อมูลสามารถถูกนำมาปรับใช้ในการตัดสินผลลัพธ์ได้ ลองพิจารณาตัวอย่างการตรวจสอบ k-fold เราสามารถแยกบันทึกเหตุการณ์ออกเป็น k ส่วน โดยที่ k-1 ส่วนสามารถถูกนำมาใช้เพื่อเรียนรู้แบบจำลองกระบวนการ และ เทคนิคการตรวจสอบความสอดคล้องสามารถถูกนำมาใช้ในการตัดสินผลลัพธ์โดยเทียบกับส่วนที่ยังคงเหลืออยู่ ซึ่งสามารถทำซ้ำได้ k ครั้ง ทำให้มีความเข้าใจอย่างลึกซึ้งถึงคุณภาพของแบบจำลอง

C4: จัดการกับมโนทัศน์ที่เบี่ยงเบน

คำว่ามโนทัศน์เบี่ยงเบน หมายถึงสถานการณ์ที่กระบวนการกำลังเปลี่ยนแปลงในขณะที่กำลังถูกวิเคราะห์ ตัวอย่างเช่น ในตอนเริ่มต้นของบันทึกเหตุการณ์ 2 กิจกรรมอาจมีการทำงานพร้อมๆกัน แต่ต่อมาในบันทึกกิจกรรมเหล่านี้กลายมาเป็นมีการทำงานแบบลำดับ กระบวนการอาจเปลี่ยนแปลงเนื่องจากมีการเปลี่ยนเป็นช่วงๆ / เปลี่ยนแปลงตามฤดูกาล (เช่น "ในเดือนธันวาคมมีความต้องการมากกว่า" หรือ "บ่ายวันศุกร์มีพนักงานเพียงไม่กี่คน") หรือ เนื่องจากการเปลี่ยนแปลงเงื่อนไข (เช่น ตลาดมีการแข่งขันมากขึ้น) การเปลี่ยนแปลงเหล่านี้มีผลกระทบต่อกระบวนการ และมีความสำคัญมากที่จะต้องค้นหาและวิเคราะห์มัน มโนทัศน์เบี่ยงเบนในกระบวนการสามารถถูกค้นพบได้โดยการแบ่งบันทึกเหตุการณ์ออกเป็นบันทึกขนาดเล็กลง และวิเคราะห์ "รอยเท้า" ของบันทึกเล็กลงเหล่านี้ การวิเคราะห์ในระดับที่สองนี้ต้องการข้อมูลเหตุการณ์จำนวนมากขึ้น อย่างไรก็ตาม มีกระบวนการจำนวนน้อยที่อยู่ในสภาวะคงที่และความเข้าใจในเรื่องของมโนทัศน์เบี่ยงเบนคือสิ่งที่สำคัญที่สุดในการบริหารจัดการกระบวนการ ดังนั้น การวิจัยและเครื่องมือสนับสนุนเพิ่มเติมจึงมีความจำเป็นเพื่อให้เหมาะสมกับการวิเคราะห์มโนทัศน์เบี่ยงเบน

C5: การปรับปรุงความโน้มเอียงของตัวแทนที่ใช้สำหรับกระบวนการค้นพบ

เทคนิคการค้นพบกระบวนสร้างแบบจำลองโดยใช้ภาษาเฉพาะ (เช่น BPMN หรือ Petri nets) แต่อย่างไรก็ตาม มันมีความสำคัญในการแยกการนำเสนอผลลัพธ์ด้วยมโนภาพออกจากตัวแทนที่ใช้ระหว่างกระบวนการค้นพบจริงๆ การเลือกภาษาเป้าหมายบ่อยครั้งมักแวดล้อมด้วยสมมุติฐานโดยนัยหลายประการ มันจำกัดพื้นที่การค้นหา; กระบวนการที่ไม่สามารถแทนได้ด้วยภาษาเป้าหมาย จะไม่สามารถถูกค้นพบได้ สิ่งนี้เรียกว่า "ความโน้มเอียงของตัวแทน" ที่ใช้ระหว่างกระบวนการค้นพบควรเลือกด้วยความตระหนักและไม่ควรถูกผลักดันด้วยตัวแทนทางกราฟิกที่ชื่นชอบเท่านั้น

พิจารณาตัวอย่างในรูปที่ 6 การที่ภาษาเป้าหมายอนุญาตให้มีการทำงานพร้อมๆกันหรือไม่ อาจมีผลกระทบต่อทั้งการสร้างมโนภาพของแบบจำลองที่ค้นพบและคลาสของแบบจำลองที่พิจารณาจากอัลกอริทึม ถ้าความโน้มเอียงของตัวแทนนั้นไม่อนุญาตให้มีการทำงานพร้อมๆกัน (รูปที่ 6(a) ก็จะเป็นไปไม่ได้) และไม่อนุญาตให้หลายกิจกรรมมีชื่อเดียวกัน (รูปที่ 6(c) ก็จะเป็นไปไม่ได้) ดังนั้นแบบจำลองที่เป็นปัญหาเช่นที่แสดงในรูปที่ 6(b) จึงมีความเป็นไปได้ ตัวอย่างนี้แสดงให้เห็นว่าต้องมีการเลือกความโน้มเอียงของตัวแทนอย่างระมัดระวังและประณีต

C6: ความสมดุลระหว่างเงื่อนไขคุณภาพ ได้แก่ ความเหมาะสม, ความเรียบง่าย, ความแม่นยำ และนัยทั่วไป

โดยทั่วไป บันทึกเหตุการณ์ยังห่างไกลจากความสมบูรณ์ เช่น มีเพียงพฤติกรรมตัวอย่างเท่านั้น แบบจำลองกระบวนการโดยทั่วไปอนุญาตให้มีจำนวนร่องรอยเป็น เอ็กโพเน็นเชียล หรือ จำนวนอนันต์ (ในกรณีของการวนซ้ำ) ยิ่งไปกว่านั้น บางร่องรอยอาจจะมีโอกาสความน่าจะเป็นน้อยกว่าร่องรอยอื่นมาก ดังนั้น มันจึงไม่เป็นจริงที่จะสมมุติว่ามีทุกร่องรอยที่เป็นไปได้อยู่ในบันทึกเหตุการณ์ เพื่อเป็นการอธิบาย มันจึงไม่มีประโยชน์ที่จะทึกทักเอาเองว่าบันทึกนั้นสมบูรณ์ ลองพิจารณาจากกระบวนการที่ประกอบด้วย 10 กิจกรรมที่สามารถทำงานอย่างขนานได้และมีบันทึกที่มีสารสนเทศประมาณ 10,000 กรณี จำนวนที่เป็นไปได้ของการแทรกสลับทั้งหมดในแบบจำลองที่มี 10 กิจกรรมทำงานพร้อมๆกันคือ 10! = 3,628,800 ดังนั้นมันจึงเป็นไปไม่ได้ที่แต่ละการแทรกสลับในบันทึกที่มีจำนวนกรณี (10,000) ซึ่งน้อยกว่าจำนวนร่องรอยที่เป็นไปได้ (3,628,800) แม้กระทั่งมีจำนวนกรณีเป็นล้านในบันทึก มันก็ยังไม่น่าจะเกิดขึ้นได้อย่างมากที่มีทุกการเปลี่ยนแปลงอยู่ในนั้น ยังมีความซับซ้อนเพิ่มเติมคือการที่บางทางเลือกมีความถี่น้อยกว่าทางเลือกอื่น สิ่งเหล่านี้อาจพิจารณาว่าคือ “สัญญาณรบกวน” มันเป็นไปไม่ได้ที่จะสร้างแบบจำลองอย่างสมเหตุสมผลสำหรับพฤติกรรมที่มีสัญญาณรบกวน แบบจำลองที่ค้นพบต้องการรูปแบบทางนามธรรมของสิ่งนี้; มันจะเป็นการดีกว่าที่จะไปตรวจสอบหาความจริงกับพฤติกรรมที่มีความถี่น้อยด้วยการใช้การตรวจสอบความสอดคล้อง

สัญญาณรบกวนและความไม่สมบูรณ์ของบันทึกเหตุการณ์ทำให้ค้นพบกระบวนการเป็นปัญหาที่ท้าท้าย ในความเป็นจริง มีการแข่งขันด้านคุณภาพใน 4 มิติ ได้แก่ (a) ความเหมาะสม, (b) ความเรียบง่าย, (c) ความแม่นยำ, และ (d) นัยทั่วไป แบบจำลองที่มีความเหมาะสมที่ดีจะยอมรับพฤติกรรมในบันทึกเหตุการณ์เกือบทั้งหมด แบบจำลองมีความเหมาะสมอย่างสมบูรณ์ถ้าร่องรอยทั้งหมดในบันทึกสามารถถูกเล่นซ้ำได้ด้วยแบบจำลองตั้งแต่ต้นจนจบ แบบจำลองที่ง่ายที่สุดที่สามารถอธิบายพฤติกรรมที่เห็นในบันทึกได้คือแบบจำลองที่ดีที่สุด หลักการนี้เป็นที่รู้จักกันในชื่อ Occam’s Razor แต่ลำพังความเหมาะสมและ ความเรียบง่าย ยังไม่เพียงพอเพียงพอที่จะใช้ในการตัดติดคุณภาพของแบบจำลองกระบวนการที่ค้นพบ ตัวอย่างเช่น มันเป็นการง่ายมากที่จะสร้าง Petri net (“แบบจำลองดอกไม้”) ที่ง่ายที่สุดเพื่อให้สามารถเล่นซ้ำ ร่องรอยทั้งหมดที่มีในบันทึกเหตุการณ์ได้ (แต่บันทึกเหตุการณ์อื่นๆที่อ้างถึงชุดของกิจกรรมเดียวกันก็สามารถใช้ได้ด้วยเช่นกัน) ในทำนองเดียวกัน เราก็ไม่ต้องการแบบจำลองที่ยอมรับเฉพาะพฤติกรรมตรงตามที่มีอยู่ในบันทึกเหตุการณ์เท่านั้น พึงระลึกว่าบันทึกประกอบด้วยตัวอย่างของพฤติกรรมเท่านั้น และมีร่องรอยที่เป็นไปได้จำนวนมากที่ยังไม่พบ แบบจำลองมีความแม่นยำ ถ้ามันไม่ได้ยอมรับพฤติกรรมจำนวน “มากเกินไป” ซึ่งเห็นได้อย่างชัดเจนว่า แบบจำลองดอกไม้ ขาดความแม่นยำ แบบจำลองที่ไม่แม่นยำคือ “หลวมเกินไป (underfitting)” หลวมเกินไปเป็นปัญหาที่แบบจำลองยอมรับพฤติกรรมตัวอย่างทั่วไปในบันทึกมากเกินไป (ตัวอย่างเช่น แบบจำลองยอมรับพฤติกรรมที่แตกต่างอย่างมากกับที่เห็นในบันทึก) แบบจำลองควรมีความเป็นนัยทั่วไปและไม่จำกัดเฉพาะพฤติกรรมที่เห็นจากตัวอย่างในบันทึกเท่านั้น แบบจำลองที่ไม่มีนัยทั่วไปคือ “คับเกินไป” (overfitting) คับเกินไปคือปัญหาที่แบบจำลองจำเฉพาะเจาะจงถูกสร้างขึ้นมาสำหรับตัวอย่างพฤติกรรมที่เห็นในตัวอย่างในบันทึกอย่างชัดเจนเท่านั้น (เช่น แบบจำลองอธิบายตัวอย่างในบันทึกอย่างเฉพาะเจาะจง แต่ตัวอย่างในบันทึกถัดมาของกระบวนการเดียวกันอาจสร้างแบบจำลองกระบวนการที่แตกต่างกันอย่างสิ้นเชิง)

การหาสมดุลของความเหมาะสม, ความเรียบง่าย, ความแม่นยำ และ นัยทั่วไป เป็นความท้าทาย จึงเป็นเหตุผลที่เทคนิคของกระบวนการค้นพบที่มีประสิทธิภาพมากส่วนใหญ่จัดให้มีตัวแปรที่หลากหลาย อัลกอริทึมต้องถูกพัฒนาให้ดีขึ้นเพื่อสร้างสมดุลที่ดีขึ้นของการแข่งขันในมิติคุณภาพ นอกจากนั้น ตัวแปรที่ใช้ควรจะสามารถเข้าใจได้โดยผู้ใช้

C7: การทำเหมืองข้ามองค์กร

ตามที่ปฏิบัติกันมา เหมืองกระบวนการจะถูกนำไปประยุกต์ใช้ภายในองค์กรเดียว แต่อย่างไรก็ตาม เมื่อเทคโนโลยีการบริการ ห่วงโซ่อุปทานแบบบูรณาการ, และ การคำนวณบนก้อนเมฆ ได้กลายมาเป็นสิ่งที่แพร่หลาย ทำให้เกิดแผนการที่บันทึกเหตุการณ์ของหลายองค์กรมีพร้อมให้ใช้ได้สำหรับการวิเคราะห์ โดยหลักการ มีสองสภาวะแวดล้อมสำหรับการทำเหมืองกระบวนการข้ามองค์กร

ประการแรก เราควรพิจารณาที่สภาวะแวดล้อมของความร่วมมือกันที่ต่างองค์กรทำงานร่วมกันเพื่อจัดการกับกระบวนการ ณ ขณะนั้น เราสามารถนึกถึงกระบวนการข้ามองค์กรเช่นเดียวกับ “ภาพตัวต่อ” เช่น กระบวนการทั้งหมดถูกตัดออกเป็นส่วนๆ และกระจายไปทั่วองค์กรที่ต้องการความร่วมมือกันเพื่อให้กรณีนั้นประสพความสำเร็จอย่างสมบูรณ์ การวิเคราะห์บันทึกเหตุการณ์เพียงภายในหนึ่งในองค์กรที่เกี่ยวข้องเหล่านี้จึงไม่เพียงพอเพื่อค้นพบกระบวนการตั้งแต่ต้นจนจบบันทึกเหตุการณ์ของต่างองค์กรจำเป็นที่จะต้องถูกนำมารวมกัน ซึ่งเป็นงานที่ไม่ง่าย เมื่อเหตุการณ์ต้องมีความสัมพันธ์ข้ามขอบเขตองค์กร

ประการที่สอง เราอาจพิจารณาสภาวะแวดล้อมที่ต่างองค์กรจำเป็นต้องกระทำการกับกระบวนการเดียวกัน ในขณะที่ใช้ประสบการณ์ ความรู้และโครงสร้างพื้นฐานร่วมกัน ลองพิจารณาตัวอย่างของ Salesforce.com กระบวนการขายของหลายองค์กรถูกบริหารและสนับสนุนโดย Salesforce ในทางหนึ่ง องค์กรเหล่านี้ใช้โครงสร้างร่วมกัน (กระบวนการ ฐานข้อมูล เป็นต้น) ส่วนในอีกทางหนึ่ง ไม่มีการบังคับให้ทำตามแบบจำลองกระบวนการอย่างเคร่งครัด โดยที่ระบบสามารถถูกปรับแต่งให้สนับสนุนการแปรผันของกระบวนการเดียวกัน อีกตัวอย่างหนึ่ง พิจารณากระบวนการพื้นฐานในการดำเนินการของเทศบาล (เช่นการขออนุญาตสร้างอาคาร) ถึงแม้ว่าเทศบาลทั้งหมดในประเทศต้องสนับสนุนชุดของกระบวนการพื้นฐานเดียวกัน แต่อาจจะมีส่วนที่แตกต่างกันได้ เห็นได้อย่างชัดเจนว่ามันเป็นสิ่งที่น่าสนใจที่จะวิเคราะห์ความแปรผันท่ามกลางองค์กรที่ต่างกัน องค์กรเหล่านี้สามารถเรียนรู้จากกันและกันและผู้ให้บริการอาจจะปรับปรุงการบริการและให้คุณค่าเพิ่มแก่บริการจากผลลัพธ์ของเหมืองกระบวนการข้ามองค์กร

มีความต้องการเทคนิคการวิเคราะห์แบบใหม่สำหรับเหมืองกระบวนการแบบข้ามองค์กรทั้งสองประเภท เทคนิคเหล่านี้ควรพิจารณาถึงเรื่องของความเป็นส่วนตัวและความมั่นคงด้วย องค์กรอาจไม่ต้องการให้ใช้ข้อมูลร่วมกันด้วยเหตุผลของการแข่งขันหรือการขาดความไว้วางใจ ดังนั้นมันจึงมีความสำคัญในการพัฒนาเทคนิคของเหมืองกระบวนการที่ยังคงรักษาความเป็นส่วนตัวไว้ได้ด้วย

C8: ให้การสนับสนุนการดำเนินการ

ในตอนเริ่มแรกนั้น เหมืองกระบวนการเน้นที่ข้อมูลในอดีต อย่างไรก็ตาม ณ วันนี้ แหล่งข้อมูลหลายแหล่งมีการปรับข้อมูลให้เป็นปัจจุบันในแบบ(ใกล้)เวลาจริง และมีความสามารถในการคำนวณอย่างพอเพียงพร้อมใช้ในการวิเคราะห์เหตุการณ์เมื่อมันเกิดขึ้น ดังนั้น เหมืองกระบวนการจึงไม่ควรถูกจำกัดอยู่กับการวิเคราะห์แบบออฟไลน์ และสามารถใช้สำหรับสนับสนุนการดำเนินการการแบบออนไลน์ สามกิจกรรมที่สนับสนุนการดำเนินการมีดังนี้: การตรวจหา, การพยากรณ์, และ การแนะนำ ในขณะที่กรณี เบี่ยงเบนจากกระบวนการที่กำหนดไว้ล่วงหน้า สิ่งนี้สามารถถูกตรวจหาได้และระบบสามารถแจ้งเตือน บ่อยครั้งที่เราต้องการให้มีการแจ้งเตือนแบบทันทีทันใด

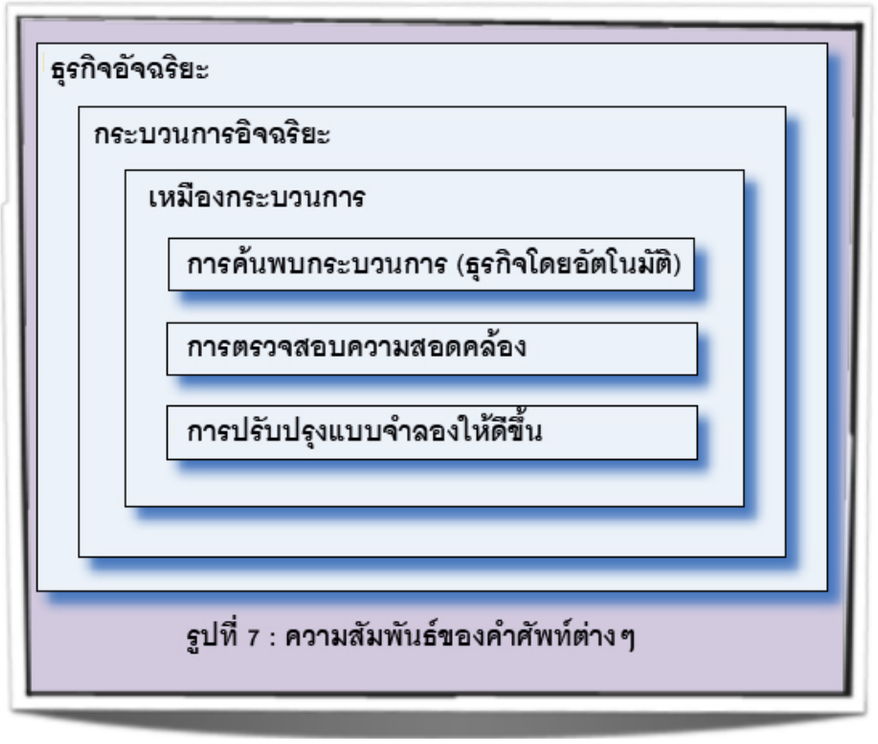
(เพื่อให้ยังสามารถมีอิทธิพลต่อสิ่งต่างๆได้) และ ไม่ได้ทำแบบออฟไลน์ ข้อมูลในอดีตสามารถถูกนำมาใช้ในการสร้างแบบจำลองพยากรณ์ได้ สิ่งเหล่านี้สามารถนำมาใช้ในการแนะนำกระบวนการ ณ ขณะนั้น ที่กำลังทำงานอยู่ ตัวอย่างเช่น มันมีความเป็นไปได้ที่จะพยากรณ์เวลาการประมวลผลที่เหลืออยู่ของกรณี จากพื้นฐานการพยากรณ์นี้ เราสามารถที่จะสร้างระบบให้คำแนะนำที่เสนอแนะการกระทำที่เป็นการเฉพาะเพื่อลดค่าใช้จ่ายหรือทำให้การไหลของเวลาสั้นลง การประยุกต์เทคนิคของทำเหมืองกระบวนการในสภาวะแบบออนไลน์ ทำให้เกิดความท้าทายเพิ่มขึ้นในเชิงของสมรรถนะในการคำนวณและคุณภาพข้อมูล

C9: การรวมเหมืองกระบวนการเข้ากับการวิเคราะห์ชนิดอื่นๆ

การบริหารดำเนินการ และโดยเฉพาะการวิจัยดำเนินงาน เป็นสาขาของวิทยาการการจัดการ ที่ขึ้นอยู่กับแบบจำลองอย่างมาก แบบจำลองทางคณิตศาสตร์ที่หลากหลายตั้งแต่ โปรแกรมเชิงเส้นและการวางแผนโครงการ ไปจนถึง แบบจำลองคิว, Markov chains, และ การจำลอง ได้ถูกนำมาใช้ เหมืองข้อมูลสามารถกำหนดว่า “การวิเคราะห์ชุดข้อมูล (โดยทั่วไปมีขนาดใหญ่) เพื่อหาความสัมพันธ์ที่คาดไม่ถึงและทำการสรุปผลข้อมูลออกมาในรูปแบบที่ทั้งสามารถเข้าใจได้และเป็นประโยชน์ต่อเจ้าของข้อมูล” หลากหลายเทคนิคได้ถูกพัฒนาขึ้นมา: การจัดหมวดหมู่ (เช่น decision tree learning), regression, clustering (เช่น k-means clustering) และ pattern discovery (เช่น association rule learning)

ทั้งสองแขนง (การบริหารการดำเนินการและเหมืองข้อมูล) ให้เทคนิคการวิเคราะห์ที่ทรงคุณค่า ความท้าทายก็คือการรวมเอาเทคนิคเหล่านี้เข้ากับเหมืองกระบวนการ ตัวอย่างเช่นการจำลอง เทคนิคของทำเหมืองกระบวนการสามารถถูกนำมาใช้ในการเรียนรู้แบบจำลอง บนพื้นฐานของข้อมูลในอดีต จากนั้นแบบจำลอง สามารถถูกใช้เพื่อให้การสนับสนุนการดำเนินการ เพราะว่าความเกี่ยวพันธ์กันอย่างใกล้ชิดของบันทึกเหตุการณ์และแบบจำลอง แบบจำลองสามารถถูกนำมาใช้ในการเล่นซ้ำ ข้อมูลในอดีต และเราสามารถเริ่มการจำลอง จากสถานะปัจจุบัน ทำให้มี “ปุ่มกดไปข้างหน้าอย่างรวดเร็ว“ ไปในอนาคตจากข้อมูลที่มีอยู่เดี๋ยวนี้

ในทำนองเดียวกัน มันมีความต้องการที่จะรวมเอาเหมืองกระบวนการเข้ากับการวิเคราะห์เชิงภาพ โดยที่การวิเคราะห์เชิงภาพรวมเอาการวิเคราะห์แบบอัตโนมัติกับการสร้างมโนภาพเชิงโต้ตอบ เพื่อให้เกิดความเข้าใจที่ดีขึ้นสำหรับชุดข้อมูลขนาดใหญ่และซับซ้อน การวิเคราะห์เชิงภาพเพิ่มขีดความสามารถของมนุษย์ได้อย่างน่าประหลาดใจในการมองเห็นรูปแบบของข้อมูลที่ไม่มีโครงสร้าง โดยการรวมเทคนิคของ



รูปที่ 7 : ความสัมพันธ์ของคำศัพท์ต่างๆ

เหมืองกระบวนการแบบอัตโนมัติเข้ากับการวิเคราะห์เชิงภาพแบบโต้ตอบ มันมีความเป็นไปได้ที่จะสกัดความเข้าใจอย่างลึกซึ้งมากขึ้นจากข้อมูลเหตุการณ์.

C10: เพิ่มความสามารถในการใช้งานสำหรับผู้ที่ไม่ได้เป็นผู้เชี่ยวชาญ

เป้าหมายหนึ่งของเหมืองกระบวนการคือการสร้าง “แบบจำลองที่มีชีวิต” นั่นคือ แบบจำลองกระบวนการที่ใช้ในชีวิตประจำวันได้มากกว่าแบบจำลองสถิต ที่จบลงด้วยการจัดเก็บอย่างถาวร ข้อมูลเหตุการณ์ใหม่สามารถใช้ในการค้นพบพฤติกรรมที่ที่เกิดขึ้นใหม่ได้ การเชื่อมโยงระหว่างข้อมูลเหตุการณ์และแบบจำลองกระบวนการอนุญาตให้การทำการฉายภาพของสถานการณ์ปัจจุบันและกิจกรรมเร็วๆนี้ไปยังแบบจำลองที่เป็นปัจจุบัน ดังนั้นผู้ใช้สามารถมีปฏิสัมพันธ์กับผลลัพธ์ของเหมืองกระบวนการแบบวันต่อวันได้ การมีปฏิสัมพันธ์เช่นนี้เป็นสิ่งที่มีค่ามาก แต่ก็จำเป็นต้องมีตัวต่อประสานกับผู้ใช้ที่สามารถเข้าใจได้ดี ความท้าทายคือการซ่อนความซับซ้อนของอัลกอริทึมของเหมืองกระบวนการไว้เบื้องหลังการเชื่อมต่อที่เป็นมิตรกับผู้ใช้ ที่มีการกำหนดชุดของพารามิเตอร์และแนะนำชนิดของการวิเคราะห์ที่เหมาะสมโดยอัตโนมัติ

C11: เพิ่มความสามารถในการเข้าใจของผู้ที่ไม่ใช่ผู้เชี่ยวชาญ

แม้กระทั่งผลลัพธ์อย่างง่ายๆที่เหมืองกระบวนการสร้างขึ้นมา ก็ไม่ได้หมายความว่ามันจะมีประโยชน์จริง ผู้ใช้อาจมีปัญหาในการทำความเข้าใจผลลัพธ์ เพื่อหลีกเลี่ยงปัญหาเหล่านี้ ผลลัพธ์ควรจะถูกนำเสนอโดยใช้ตัวแทน ที่เหมาะสม (ดูที่ GP5) นอกจากนั้น ความคู่ควรต่อการไว้วางใจ ของผลลัพธ์ควรมีการบ่งชี้อย่างชัดเจน อาจมีข้อมูลน้อยเกินไปกว่าที่จะทำการสรุปอย่างหนึ่งอย่างใดโดยเฉพาะได้ ในความเป็นจริงเทคนิคกระบวนการค้นพบที่มีอยู่ โดยทั่วไปไม่มีการแจ้งเตือนว่ามีความเหมาะสมต่ำ หรือ คับเกินไป แต่มีการแสดงแบบจำลองเสมอ แม้ว่าจะมีความชัดเจนว่ามีจำนวนข้อมูลน้อยเกินไปกว่าที่จะนำมาสรุปผลใดๆ ได้ก็ตาม

ปัจฉิมบท

คณะกรรมการเฉพาะกิจเหมืองกระบวนการของ IEEE มีเป้าหมายที่จะ (a) ส่งเสริมการประยุกต์ใช้เหมืองกระบวนการ (b) แนะนำผู้พัฒนาซอฟต์แวร์ ที่ปรึกษา ผู้บริหารธุรกิจ และผู้ใช้ เมื่อใช้เทคนิคที่ทันสมัย และ (c) กระตุ้นการวิจัยเหมืองกระบวนการ เจตนารมณ์นี้กล่าวถึงหลักการหลักและความตั้งใจของคณะกรรมการเฉพาะกิจ หลังจากแนะนำหัวข้อของเหมืองกระบวนการ เจตนารมณ์นี้ได้จำแนกหลักการแนะนำ (ส่วนที่ 3) และความท้าทาย (ส่วนที่ 4) บางประการ หลักการแนะนำสามารถใช้เพื่อหลีกเลี่ยงความผิดพลาดที่ชัดเจน รายการความท้าทายมีความตั้งใจที่จะให้ทิศทางการวิจัยและความพยายามในการพัฒนา ทั้งสองมีเป้าหมายที่จะเพิ่มระดับวุฒิภาวะของเหมืองกระบวนการ

สรุปคำศัพท์ที่ใช้บางคำ ต่อไปนี้คือคำศัพท์เฉพาะที่ใช้ในเรื่องเหมืองกระบวนการ: เหมืองการไหลของงาน, เหมืองกระบวนการ (ธุรกิจ), การค้นพบกระบวนการ (ธุรกิจ) โดยอัตโนมัติ,

และ กระบวนการ (ธุรกิจ) อัจฉริยะ ต่างองค์กรดูเหมือนว่าจะมีการใช้ศัพท์ที่ต่างกันสำหรับมโนภาพที่ทับซ้อนกัน ตัวอย่างเช่น Gartner ส่งเสริมการใช้คำว่า “Automated Business Process Discovery” (ABPD) และ Software AG ใช้ “Process Intelligence” เพื่ออ้างถึงแพลตฟอร์มที่ควบคุม คำว่า “เหมืองการไหลของงาน” ดูเหมือนว่าจะมีความเหมาะสมน้อยกว่า เนื่องจากการสร้างแบบจำลองการไหลของงาน เป็นเพียงหนึ่งในหลายๆสิ่งที่เป็นไปได้ในการประยุกต์ใช้เหมืองกระบวนการ ในทำนองเดียวกัน การเพิ่มคำว่าธุรกิจ เข้าไปทำให้ขอบเขตการประยุกต์ใช้เหมืองกระบวนการแคบลง เหมืองกระบวนการสามารถประยุกต์ใช้ได้หลายอย่าง (เช่น การวิเคราะห์การใช้ระบบที่ใช้เทคโนโลยีสูง หรือ การวิเคราะห์เว็บไซต์ โดยที่สิ่งที่เพิ่มเข้ามานี้ดูไม่น่าเหมาะสม ถึงแม้ว่าการค้นพบกระบวนการเป็นส่วนที่สำคัญอย่างหนึ่งของเหมืองกระบวนการ แต่มันก็เป็นเพียงหนึ่งในหลายสิ่งที่เหมืองกระบวนการทำได้ การตรวจสอบความถูกต้อง, การพยากรณ์, เหมืององค์กร, การวิเคราะห์เครือข่ายสังคม เป็นต้น คือตัวอย่างของการใช้ในกรณีอื่นๆที่ขยายออกไปมากกว่าการค้นพบกระบวนการ

รูปที่ 7 แสดงความสัมพันธ์ของบางคำที่ได้กล่าวไป ทุกเทคโนโลยีและวิธีการที่มีเป้าหมายที่จะให้ได้สารสนเทศที่สามารถนำไปกระทำการได้ ซึ่งสามารถใช้ในการสนับสนุนการตัดสินใจ นั้นสามารถวางอยู่ในตำแหน่งที่อยู่ภายใต้ธุรกิจอัจฉริยะ กระบวนการ (ธุรกิจ) อัจฉริยะสามารถถูกมองว่าเป็นการรวมกันของ BI และ BPM, นั่นคือ เทคนิคของ BI ถูกนำมาใช้เพื่อการวิเคราะห์และปรับปรุงกระบวนการ และการบริหารจัดการเหมืองกระบวนการสามารถถูกมองว่าเป็นรูปธรรมของกระบวนการอัจฉริยะ โดยมีบันทึกเหตุการณ์เป็นจุดเริ่มต้น การค้นพบกระบวนการ (ทางธุรกิจโดยอัตโนมัติ) เป็นเพียงหนึ่งในสามชนิดพื้นฐานของเหมืองกระบวนการ รูปที่ 7 อาจทำให้เข้าใจผิดได้เล็กน้อยในแง่ที่ว่า เครื่องมือของ BI ส่วนใหญ่ไม่มีฟังก์ชันของเหมืองกระบวนการ ที่กล่าวไว้ในเอกสารนี้ คำว่า BI บ่อยครั้งเพื่อความสะดวกถูกนำไปใช้สำหรับเครื่องมือที่เฉพาะเจาะจง หรือวิธีการที่ครอบคลุมเฉพาะบางส่วนของ BI โดยรวมเท่านั้น

อาจมีเหตุผลทางการค้าสำหรับการใช้คำศัพท์อื่น ผู้ขายบางรายต้องการที่จะเน้นในบางมุมมอง (เช่น การค้นพบ หรือ อัจฉริยะ) แต่อย่างไรก็ตามเพื่อหลีกเลี่ยงความสับสน มันจึงเป็นสิ่งที่ดีกว่าที่จะใช้คำว่า “เหมืองกระบวนการ” สำหรับศาสตร์ที่ครอบคลุมอยู่ในเจตนารมณ์นี้

อภิธานศัพท์

กิจกรรม (Activity): ขั้นตอนที่มีการกำหนดอย่างชัดเจนในกระบวนการ เหตุการณ์อาจหมายถึงเริ่มต้น (start), เสร็จสิ้น (completion), ยกเลิก (cancelation), เป็นต้น ของกิจกรรม สำหรับกระบวนการหนึ่ง ณ ขณะนั้น (a specific process instance)

การค้นพบกระบวนการธุรกิจโดยอัตโนมัติ (Automated Business Process Discovery): ดูที่ การค้นพบกระบวนการ

ธุรกิจอัจฉริยะ (Business Intelligence (BI)): ชุดของเครื่องมือและวิธีการที่ใช้ข้อมูลในการสนับสนุนการตัดสินใจ

กระบวนการธุรกิจอัจฉริยะ (Business Process Intelligence): ดูที่ กระบวนการอัจฉริยะ

การบริหารกระบวนการธุรกิจ (Business Process Management (BPM)): ศาสตร์ที่รวมความรู้ด้านเทคโนโลยีสารสนเทศและความรู้จากวิทยาการการจัดการ และประยุกต์ทั้งสองศาสตร์นี้เพื่อกระบวนการดำเนินการทางธุรกิจ

กรณี (Case): ดูที่ กระบวนการ ณ ขณะนั้น

มโนทัศน์ที่เบี่ยงเบน (Concept Drift): ปรากฏการณ์ที่กระบวนการมีการเปลี่ยนแปลงบ่อบครั้งตามเวลา กระบวนการที่เฝ้าสังเกตอาจค่อยๆ เปลี่ยนแปลง (หรือเปลี่ยนแปลงอย่างทันทีทันใด) ที่เกิดจากการเปลี่ยนแปลงตามฤดูกาล หรือการแข่งขันที่เพิ่มขึ้น ทำให้การวิเคราะห์มีความซับซ้อน

การตรวจสอบความสอดคล้อง (Conformance Checking): การวิเคราะห์ว่าความเป็นจริงที่ถูกบันทึกไว้ในบันทึก สอดคล้องกับแบบจำลองและในทางกลับกัน เป้าหมายคือตรวจหาความแตกต่าง และวัดความรุนแรง การตรวจสอบความสอดคล้องเป็นหนึ่งในสามชนิดพื้นฐานของเหมืองกระบวนการ

การทำเหมืองกระบวนการข้ามองค์กร (Cross-Organizational Process Mining): การประยุกต์เทคนิคการทำเหมืองกระบวนการกับบันทึกเหตุการณ์ที่มีจุดกำเนิดมาจากต่างองค์กร

เหมืองข้อมูล (Data Mining): การวิเคราะห์ชุดข้อมูล (ส่วนใหญ่มีขนาดใหญ่) เพื่อหาความสัมพันธ์ที่คาดไม่ถึง และเพื่อสรุปข้อมูลในลักษณะที่ให้ความเข้าใจที่ลึกซึ้งใหม่

เหตุการณ์ (Event): การกระทำที่ถูกบันทึกไว้ในบันทึก เช่น เริ่มต้น (start), เสร็จสิ้น (completion) หรือ ยกเลิก (cancelation), ของกิจกรรมสำหรับกระบวนการ ณ ขณะนั้นโดยเฉพาะ

บันทึกเหตุการณ์ (Event Log): ชุดของเหตุการณ์ใช้เพื่อเป็นข้อมูลเข้าของเหมืองกระบวนการ เหตุการณ์ไม่มีความจำเป็นที่ถูกจัดเก็บไว้ในแฟ้มบันทึกที่แยกเฉพาะ (เช่น เหตุการณ์อาจจะกระจัดกระจายอยู่ในตารางที่ต่างกันในฐานข้อมูล)

เจตนารมณ์นี้ถูกตีพิมพ์ครั้งแรกใน “Business Process Management Workshops 2011, Lecture Notes in Business Information Processing, Vol. 99, Springer-Verlag, 2011, “ และถูกแปลไปอยู่ในหลายภาษา ดูได้ที่ Home Page ของ IEEE Task Force on Process Mining: http://www.win.tue.nl/ ieeetfpm/ สำหรับข้อมูลเพิ่มเติม

ความเหมาะสม (Fitness): การวัดเพื่อตัดสินว่าแบบจำลองยอมรับพฤติกรรมที่พบในบันทึกเหตุการณ์อย่างไร แบบจำลองมีความเหมาะสมแบบสมบูรณ์ถ้าทุกร่องรอยในบันทึกสามารถเล่นซ้ำได้โดยแบบจำลองตั้งแต่ต้นจนจบ

นัยทั่วไป (Generalization): การวัดเพื่อตัดสินว่าแบบจำลองสามารถยอมรับให้พฤติกรรมที่ไม่เคยพบในบันทึกเหตุการณ์อย่างไร แบบจำลองที่คับเกินไปคือไม่สามารถมีนัยทั่วไปเพียงพอ

การปรับปรุงแบบจำลองให้ดีขึ้น (Model Enhancement): เป็นหนึ่งในสามของชนิดพื้นฐานของเหมืองกระบวนการ แบบจำลองกระบวนการถูกขยายหรือปรับปรุงให้ดีขึ้นโดยใช้สารสนเทศที่สกัดจากบันทึก ตัวอย่างเช่น ปัญหาคอขวดสามารถถูกบ่งชี้ได้โดยการเล่นซ้ำบันทึกเหตุการณ์บนแบบจำลองกระบวนการในขณะที่สำรวจประทับเวลา

MXML: รูปแบบที่อยู่บนพื้นฐานของ XML สำหรับการแลกเปลี่ยนบันทึกเหตุการณ์ XES เข้ามาแทนที่ MXML ในฐานะรูปแบบใหม่ที่ไม่ขึ้นกับเครื่องมือของเหมืองกระบวนการ

สนับสนุนการดำเนินการ (Operational Support): การวิเคราะห์แบบออนไลน์ของข้อมูลเหตุการณ์โดยมีเป้าหมายเพื่อเฝ้าสังเกตุและมีอิทธิพลต่อกระบวนการที่กำลังทำงานวิ่งอยู่ ณ ขณะนั้น สามกิจกรรมที่สนับสนุนการดำเนินการสามารถบ่งชี้ได้แก่: ตรวจหา (สร้างการแจ้งเตือนถ้าพฤติกรรมที่สังเกตเบี่ยงเบนจากพฤติกรรมของแบบจำลอง), พยากรณ์ (พยากรณ์พฤติกรรมอนาคตบนพื้นฐานของพฤติกรรมในอดีต เช่น พยากรณ์เวลาประมวลผลที่เหลืออยู่), และการแนะนำ (แนะนำการกระทำที่เหมาะสมที่จะนำมาซึ่งเป้าหมายที่เฉพาะเจาะจง เช่น ค่าใช้จ่ายต่ำสุด).

ความแม่นยำ (Precision): การวัดเพื่อตัดสินว่าแบบจำลองห้ามพฤติกรรมที่แตกต่างกันมากกับพฤติกรรมที่พบในบันทึกเหตุการณ์ได้เพียงใด แบบจำลองที่มีความแม่นยำต่ำเรียกว่า หลวมเกินไป

การค้นพบกระบวนการ (Process Discovery): หนึ่งในสามชนิดพื้นฐานของเหมืองกระบวนการ แบบจำลองกระบวนการเรียนรู้จากบันทึกเหตุการณ์ ตัวอย่างเช่น α อัลกอริทึมสามารถค้นพบ Petri net โดยการบ่งชี้รูปแบบกระบวนการในชุดของเหตุการณ์

กระบวนการ ณ ขณะนั้น (Process Instance): เอนทิตีที่กำลังถูกจัดการโดยกระบวนการที่ถูกวิเคราะห์ เหตุการณ์อ้างถึง กระบวนการ ณ ขณะนั้น ตัวอย่างของกระบวนการ ณ ขณะนั้น ได้แก่ คำสั่งซื้อของลูกค้า การเรียกร้องประกัน การขอสินเชื่อ

กระบวนการอัจฉริยะ (Process Intelligence): แขนงของธุรกิจอัจฉริยะที่เน้นที่ การบริหารกระบวนการธุรกิจ

เหมืองกระบวนการ (Process Mining): เทคนิค เครื่องมือ และวิธีการ ที่ค้นพบ ตรวจสอบ และปรับปรุงกระบวนการจริง (ไม่ใช่กระบวนการสมมุติ) โดยสกัดความรู้จากบันทึกเหตุการณ์ที่มีโดยทั่วไปในระบบสารสนเทศในทุกวันนี้

ความโน้มเอียงของตัวแทน (Representational Bias): ภาษาเป้าหมายที่ถูกเลือกสำหรับเป็นตัวแทนและสร้างผลลัพธ์ของเหมืองกระบวนการ

ความเรียบง่าย (Simplicity): วัดตาม Occam's Razor, แบบจำลองที่เรียบง่ายที่สุดที่สามารถอธิบายพฤติกรรมที่พบในบันทึก คือแบบจำลองที่ดีที่สุด ความเรียบง่ายสามารถบอกเป็นจำนวนได้หลายวิธี เช่น จำนวนโนด และ เส้นเชื่อมในแบบจำลอง

XES: อยู่บนพื้นฐานของ XML มาตรฐานสำหรับบันทึกเหตุการณ์ มาตรฐานนี้ได้รับความเห็นชอบโดยคณะกรรมการเฉพาะกิจเหมืองกระบวนการของ IEEE ให้เป็นรูปแบบที่ใช้ในการแลกเปลี่ยนบันทึกเหตุการณ์โดยปริยาย (ดูที่ www.xes-standard.org)

แปลโดย รองศาสตราจารย์ ดร. วิเชียร เปรมชัยสวัสดิ์ และนักศึกษาบัณฑิตวิทยาลัย สาขาเทคโนโลยีสารสนเทศ มหาวิทยาลัยสยาม โดยการสนับสนุนของสมาคมคอมพิวเตอร์แห่งประเทศไทยในพระบรมราชูปถัมภ์

ผู้เขียน

Wil van der Aalst
Arya Adriansyah
Ana Karla Alves de Medeiros
Franco Arcieri
Thomas Baier
Tobias Blickle
Jagadeesh Chandra Bose
Peter van den Brand
Ronald Brandtjen
Joos Buijs
Andrea Burattin
Josep Carmona
Malu Castellanos
Jan Claes
Jonathan Cook
icola Costantini
Francisco Curbera
Ernesto Damiani
Massimiliano de eoni
Pavlos Delias
Boudewijn van Dongen
Marlon Dumas
Schahram Dustdar
Dirk Fahland
Diogo R. Ferreira
Walid Gaaloul
Frank van Geffen
Sukriti Goel
Christian Günther
Antonella Guzzo
Paul Harmon
Arthur ter Hofstede
John Hoogland
Jon Espen Ingvaldsen
Koki Kato
Rudolf Kuhn
Akhil Kumar
Marcello La Rosa
Fabrizio Maggi
Donato Malerba
Ronny Mans
Alberto Manuel
Martin McCreesh
Paola Mello
Jan Mendling
Marco Montali
Hamid Motahari Nezhad
Michael zur Muehlen
Jorge Munoz-Gama
Luigi Pontieri
Joel Ribeiro
Anne Rozinat
Hugo Seguel Pérez
Ricardo Seguel Pérez
Marcos Sepúlveda
Jim Sinur
Pnina Soffer
Minseok Song
Alessandro Sperduti
Giovanni Stilo
Casper Stoel
Keith Swenson
Maurizio Talamo
Wei Tan
Chris Turner
Jan Vanthienen
George Varvaressos
Eric Verbeek
Marc Verdonk
Roberto Vigo
Jianmin Wang
Barbara Weber
Matthias Weidlich
Ton Weijters
Lijie Wen
Michael Westergaard
Moe Wynn