

# ANALISIS POLA KETIDAKSEMPURNAAN CATATAN KEJADIAN SEBAGAI MASUKAN DALAM PROCESS MINING

**Satrio Adi Priyambada dan Mahendrawathi ER**

Departemen Sistem Informasi, Fakultas Teknologi Informasi dan Komunikasi,  
Institut Teknologi Sepuluh Nopember  
Jl. Raya ITS, Surabaya, 60111  
Telp : (+6231) 5999944, Fax : (+6231) 5964965  
E-mail : [jurusan@is.its.ac.id](mailto:jurusan@is.its.ac.id)<sup>1)</sup>

---

## **Abstrak**

*Process mining merupakan teknik yang digunakan untuk memodelkan dan menganalisis proses bisnis berdasarkan catatan kejadian yang diekstrak dari sistem informasi organisasi. Data yang diektrak adalah catatan kejadian yang terdiri dari case, aktivitas dan waktu aktivitas itu dikerjakan. Salah satu kendala dari teknik process mining adalah mengenai isu kualitas serta pola ketidaksempurnaan dari catatan kejadian. Dalam penelitian ini dilakukan analisis dan identifikasi pola ketidaksempurnaan dari berbagai tipe perusahaan, yaitu manufaktur, healthcare dan edukasi. Dalam melakukan analisis diperlukan pengumpulan data berupa catatan kejadian dari tugas akhir mahasiswa. Data catatan kejadian yang didapatkan berasal dari berbagai jenis perusahaan seperti perusahaan manufaktur, healthcare maupun edukasi. Analisis dilakukan dengan mengidentifikasi adanya pola ketidaksempurnaan serta solusi penyelesaiannya. Identifikasi juga dilakukan dengan menemukan pola ketidaksempurnaan baru jika dimungkinkan serta solusinya. Penelitian ini dapat mengidentifikasi berbagai pola ketidaksempurnaan yang sebelumnya telah diajukan serta menemukan pola ketidaksempurnaan baru.*

**Kata kunci:** *Process Mining, Catatan kejadian, Pola Ketidaksempurnaan*

## **1. PENDAHULUAN**

Proses bisnis yang berjalan saat ini dapat tersimpan dalam sistem informasi yang digunakan. Data yang tersimpan dalam sistem adalah data berupa *timestamp* dari proses yang berjalan. *Process mining* adalah disiplin ilmu baru yang dapat menemukan, memantau dan memperbaiki proses sebenarnya dengan mengekstrak pengetahuan dari catatan kejadian yang didapat dari sistem informasi [1]. Potensial dari keberhasilan penerapan *process mining* telah banyak dibuktikan dengan penggunaan data nyata, namun masih terdapat kesulitan untuk mendapatkan catatan kejadian yang diinginkan dan dapat digunakan [2]. Kematangan dari log tidak menjamin bahwa log tersebut bebas dari *pola ketidaksempurnaan*. Sehingga perlu adanya perbaikan yang dilakukan untuk menjaga kualitas dari catatan kejadian yang digunakan untuk analisis dalam penerapan *process mining*.

Ayo dkk telah mengajukan sebuah penelitian mengenai penanganan catatan kejadian yang tidak lengkap. Dimana untuk mendeteksi catatan kejadian yang hilang digunakan *fuzzy logic* dan *bayesian scoring* [3]. Sedangkan Yang, dkk menyelesaikan permasalahan kelengkapan catatan kejadian menggunakan CPL+ atau *the Completeness of Process Logs*, yaitu mengenai peluang kemunculan *direct succession* yang teramati maupun tidak teramati [4]. Sedangkan, Tax, dkk membahas mengenai kualitas pemberian label pada catatan kejadian, dimana pemberian ulang label pada catatan kejadian sangatlah memberikan sudut pandang lain dalam *process discovery*. Metode ini dievaluasi menggunakan teori statistik dan informasi [5].

Isu lain yang telah dibahas mengenai kualitas catatan kejadian adalah adanya *noise* yang harus dibersihkan agar tidak mengurangi kualitas analisis dari *process mining*. Conforti, dkk membersihkan catatan kejadian dari *noise* dengan mendeteksi *outliers*, selanjutnya melakukan deteksi *noise* pada *outlier* yang ditemukan dan menghapusnya. Maruster, dkk menangani *noise* pada catatan kejadian dengan menggunakan *rule-based* - berdasarkan *relational metrics* [6]. Sedangkan Cheng, dkk membersihkan catatan kejadian dari *noise* menggunakan Algoritma PRISM yang merupakan algoritma rule-based. Hasil dari catatan kejadian yang telah dibersihkan dari *noise* tersebut dapat menghasilkan model proses yang lebih menggambarkan proses sesungguhnya daripada catatan kejadian yang masih terdapat *noise* didalamnya dengan melihat nilai *recall* dan *precision* [7].

Dari sekian banyak isu kualitas pada catatan kejadian, Bose dkk mengajukan isu kualitas data dari catatan kejadian yang dapat diidentifikasi dari catatan kejadian berupa 4 isu kualitas data yang dapat menghambat analisis menggunakan *process mining*. Isu kualitas data ini adalah *missing data*, *incorrect data*, *imprecise data* dan *irrelevant data*. Dimana empat isu kualitas ini dihubungkan dengan entitas yang terdapat dalam catatan kejadian sehingga menghasilkan 27 kelas dari isu kualitas yang dapat diidentifikasi.

Suriadi dkk melakukan penelitian mengenai *pola ketidaksempurnaan* yang merupakan pola dari catatan kejadian yang dapat mengurangi kualitas dari catatan kejadian. *Pola ketidaksempurnaan* ini juga dipetakan ke dalam isu kualitas dari Bose dkk. Dalam penelitian ini dijelaskan mengenai deskripsi, contoh dikehidupan nyata, efek, isu data kualitas, cara mendeteksi pola, cara memperbaiki log, efek samping dari perbaikan dan *indicative rule* dari masing-masing *pola ketidaksempurnaan*.

Dari pemetaan isu kualitas data ke dalam *pola ketidaksempurnaan* yang dilakukan, terdapat 9 dari 27 isu kualitas yang dipetakan ke dalam *pola ketidaksempurnaan*. Sehingga masih terdapat 18 isu kualitas yang belum dipetakan. Hal ini yang mendasari penelitian ini untuk menemukan kemungkinan adanya *pola ketidaksempurnaan* lain yang dapat mewakili isu kualitas data yang tersisa. Ditambah lagi, penelitian [8] dilakukan dengan menggunakan data dari healthcare saja, sehingga belum dapat diketahui kemunculan *pola ketidaksempurnaan* pada catatan kejadian yang berasal dari bidang manufaktur maupun edukasi. Tujuan dari penelitian ini adalah untuk menemukan *pola ketidaksempurnaan* pada catatan kejadian yang dimiliki, baik di bidang manufaktur, healthcare maupun education. Diharapkan dapat dilihat kemunculan dari *pola ketidaksempurnaan* pada catatan kejadian serta dimungkinkan untuk menemukan *pola ketidaksempurnaan* baru.

## 2. PROCESS MINING

*Process Mining* merupakan suatu disiplin ilmu yang menggabungkan antara disiplin ilmu komputasi intelegensia, data mining di satu sisi dan pemodelan proses bisnis serta analisis di sisi lain [1]. *Process Mining* dilakukan dengan melakukan analisis pada Catatan kejadian yang tersimpan dalam suatu sistem informasi. Analisis ini dilakukan untuk memperoleh sebuah pengetahuan yang didapatkan dari Catatan kejadian tersebut. Catatan kejadian (event log) merupakan sebuah catatan yang direkam pada suatu sistem untuk masing-masing user. Catatan kejadian berisi mengenai catatan sistem terkait dengan peristiwa/ alur kerja yang biasanya digunakan untuk melakukan pelacakan alur kerja [9].

Dengan menggunakan teknik *Process Mining*, dapat diketahui mengenai proses yang terjadi, kontrol, penggunaan data, pemanfaatan sumber daya dan berbagai performa yang berhubungan dengan bidang statistik. Manfaat dari teknik *Process Mining* adalah untuk dapat melihat bagaimana prosedur bekerja. *Process Mining* dapat diterapkan pada sistem informasi secara luas baik yang berupa sistem informasi per divisi hingga sistem *enterprise resource planning*. Proses bisnis operasional yang menggunakan sistem informasi akan menghasilkan Catatan kejadian. Catatan kejadian ini perlu diekstrak untuk mendapatkan pengetahuan. Dapat dilihat bahwa salah satu peran dari teknik *Process Mining* adalah menggambarkan data Catatan kejadian dari satu set data real.

Catatan kejadian yang didapatkan dapat digunakan untuk melakukan tiga tipe *Process Mining*, yaitu *discovery*, *conformance* dan *enhancement*. *Process Mining* yang bertujuan untuk menemukan model dari Catatan kejadian yang ada disebut dengan *Discovery*. Sementara itu *Process Mining* yang bertujuan untuk membandingkan proses model yang ada dengan proses bisnis yang ideal disebut dengan *conformance*. *Process Mining* yang bertujuan untuk memperbaiki proses model dengan memodifikasi proses model yang dihasilkan oleh Catatan kejadian disebut dengan *Enhancement*. Sedangkan berdasarkan persepektif dari *Process Mining*, teknik ini dibagi menjadi empat perspektif [1], yaitu:

- a. Control-flow Perspective, yaitu perspektif *Process Mining* yang berfokus pada kontrol dari aliran proses dengan tujuan untuk menemukan karakteristik yang baik untuk kemungkinan semua path pada petri net.
- b. Organizational Perspective, yaitu perspektif *Process Mining* yang berfokus pada sumber daya yang terlibat dalam log seperti aktor, role, departemen dll. Perspektif ini lebih menunjukkan aktor yang terlibat dengan mengaitkannya dalam analisis jejaring sosial.
- c. Case Perspective, yaitu perspektif *Process Mining* yang berfokus pada kasus yang dicirikan berdasarkan alur prosesnya. Setiap karakteristik data yang sama akan dikelompokkan dan dimasukkan ke dalam sebuah kasus yang sama yang disebut varian.
- d. Time Perspective, yaitu perspektif *Process Mining* yang berfokus pada waktu frekuensi kejadian/aktivitas. Tujuan dari perspektif ini adalah untuk menemukan bottleneck proses, mengukur tingkat layanan, memonitoring performa dan mengaplikasikan waktu yang tersisa dalam mengeksekusi pelaksanaan aktivitas tersebut.

### 3. ISU KUALITAS CATATAN KEJADIAN

Bose dkk. [8] telah mengidentifikasi 4 kategori isu yang mempengaruhi kualitas catatan kejadian yang digunakan dalam catatan kejadian. *Missing Data* yaitu isu kualitas yang mengacu pada hubungan dari skenario dimana berbagai macam informasi yang dimungkinkan. *Incorrect Data* yaitu isu kualitas yang mengacu pada informasi dari data yang dihasilkan pada log tidak terekam dengan benar. *Imprecise Data* yaitu isu kualitas yang mengacu pada kurangnya presisi dari data. Contohnya adalah tingkat kedetailan dari *timestamp*. *Irrelevant Data* yaitu isu kualitas yang mengacu pada data yang tidak relevan, yaitu tingkat agregasi dari log tidak sesuai dengan analisis yang ingin dilakukan. Dari keempat isu kualitas catatan kejadian tersebut, kemudian dipetakan terhadap beberapa objek dari catatan kejadian yaitu, *case*, *event*, *belongs to*, *case attributes*, *position*, *activity name*, *timestamp*, *resource* dan *event attribute*. Sehingga dari tabel tersebut dihasilkan 27 kelas dari isu kualitas yang dapat diidentifikasi yang dapat dilihat pada gambar berikut.

	case	event	belongs to	c_attribute	position	activity name	timestamp	resource	e_attribute
Missing Data	I1	I2	I3	I4	I5	I6	I7	I8	I9
Incorrect Data	I10	I11	I12	I13	I14	I15	I16	I17	I18
Imprecise Data			I19	I20	I21	I22	I23	I24	I25
Irrelevant Data	I26	I27							

Gambar 1. Tabel isu kualitas catatan kejadian

Dari isu kualitas tersebut menyebabkan berkurangnya penerapan teknik *process mining* yang dapat diaplikasikan terhadap catatan kejadian tersebut. Sehingga perlu adanya penelitian yang berfokus pada bagaimana cara untuk mengatasi isu kualitas tersebut.

### 4. POLA KETIDAKSEMPURNAAN

Penelitian berikutnya adalah mengenai *pola ketidaksempurnaan*, Suriadi, dkk mengajukan framework dalam menangani catatan kejadian yang tidak sempurna dengan memberikan deskripsi, contoh di kehidupan nyata, efek yang ditimbulkan, isu data kualitas, cara mendeteksi pola, cara memperbaiki log, efek samping dari perbaikan log dan *indicative rule* dari pola yang diajukan. Dalam penelitiannya ditemukan 11 *pola ketidaksempurnaan*. Suriadi, dkk memetakan *pola ketidaksempurnaan* terhadap isu kualitas yang diajukan Bose, dkk sebelumnya. Sehingga menghasilkan tabel 1.

Tabel 1. Pemetaan *Pola ketidaksempurnaan* ke dalam isu kualitas catatan kejadian

No	Name	Isu Kualitas	P06	Scattered Case	I12
P01	Form-based Event Capture	I16, I27	No	Name	Isu Kualitas
P02	Inadvertent Time Travel	I16	P07	Collateral Events	I27
P03	Unanchored Event	I23, I16	P08	Polluted Label	I15, I17
P04	Scattered Event	I2	P09	Distorted Label	I15
P05	Elusive Case	I3	P10	Synonymous Labels	I22
			P11	Homonymous Label	I22

Dari tabel tersebut dapat dilihat bahwa terdapat isu kualitas yang belum muncul ke dalam *pola ketidaksempurnaan*. Hal ini yang menjadi motivasi penulis untuk menemukan *pola ketidaksempurnaan* yang sudah ada pada catatan kejadian yang dimiliki. Hal ini dilakukan untuk melihat apakah terdapat *pola ketidaksempurnaan* baru yang dapat ditemukan dan mencari solusi lain berdasarkan catatan kejadian yang dimiliki.

### 5. METODOLOGI

Penelitian ini akan berfokus pada menemukan *pola ketidaksempurnaan* dari catatan kejadian yang dimiliki dan apabila dimungkinkan menemukan *pola ketidaksempurnaan* lain. Terdapat beberapa langkah dalam pengerjaan penelitian ini yaitu, studi literatur, pengumpulan data, analisis data dan kesimpulan. Studi literatur dilakukan dengan mencari dan memahami paper mengenai permasalahan-permasalahan yang ditimbulkan oleh catatan kejadian, isu kualitas dari catatan kejadian dan berbagai penanganannya. Selanjutnya pengumpulan data dilakukan dengan menggunakan data yang digunakan dalam tugas akhir mahasiswa Sistem Informasi di Institut Teknologi

Sepuluh Nopember Surabaya. Pengumpulan data dilakukan dengan mencari mahasiswa yang mengambil topik *Process mining* sebagai tugas akhirnya. Tahapan analisis data terdiri dari mendeteksi *pola ketidaksempurnaan* dan mencari solusi penyelesaian. Identifikasi adanya *pola ketidaksempurnaan* dilakukan dengan melihat ciri-ciri khusus pada data yang didapatkan. Untuk ciri-ciri dari masing-masing pola ketidaksempurnaan dapat dilihat pada tabel 2. Kemudian setelah didapatkan *pola ketidaksempurnaan* yang teridentifikasi, dilakukan review bagaimana cara penulis dari masing-masing tugas akhir dalam menyelesaikan permasalahan terkait.

Tabel 2. Ciri-ciri pola ketidaksempurnaan pada event log

No	Ciri-ciri
P01	Adanya sekelompok event dalam log yang memiliki <i>case</i> dan <i>timestamp</i> yang sama
P02	Adanya beberapa <i>case</i> yang urutannya berbeda sangat signifikan
P03	Adanya format <i>timestamp</i> yang tidak kompatibel pada beberapa sistem
P04	Adanya informasi yang tersembunyi pada attribute dalam event log, untuk mengetahuinya perlu adanya bantuan dari pihak yang mengetahui mengenai proses tersebut
P05	Kurangnya informasi pada event sehingga sulit untuk dibentuk menjadi <i>case</i>
P06	Event log kurang detail karena tersimpan pada beberapa sistem informasi
P07	Adanya event yang memiliki <i>timestamp</i> yang terlalu dekat antara satu dengan yang lain
P08	Adanya campuran format dalam atribut event log
P09	Adanya perbedaan yang sangat kecil dalam atribut, misal perbedaan huruf kapital
P10	Adanya beberapa nilai dari atribut yang memiliki nilai sama namun berbeda penulisan
P11	Adanya aktivitas yang menjadi pusat arah aktivitas termasuk adanya self-loop

## 6. DATA

Pengumpulan data yang akan digunakan untuk proses analisis dalam menemukan pola ketidaksempurnaan dikategorikan berdasarkan tipe perusahaan yaitu, manufaktur, healthcare dan edukasi. Hal ini bertujuan untuk mengetahui karakteristik dari pola ketidaksempurnaan yang sering muncul di masing-masing tipe perusahaan. Studi kasus yang termasuk dalam kategori manufaktur adalah studi kasus yang terkait perusahaan yang bergerak dalam bidang industri pembuatan barang. Dalam penelitian ini terdapat 12 catatan kejadian yang termasuk ke dalam kategori manufaktur diantaranya delapan perusahaan sepatu, satu perusahaan semen, satu perusahaan farmasi dan satu perusahaan pupuk. Studi kasus yang termasuk dalam kategori Healthcare adalah organisasi yang bergerak dalam bidang pelayanan kesehatan. Dalam penelitian ini terdapat 4 catatan kejadian yang termasuk ke dalam kategori Healthcare. Catatan kejadian yang digunakan dari rumah sakit adalah catatan kejadian pasien penyakit diabetes dan jantung. Sedangkan studi kasus yang termasuk dalam kategori *Education* adalah organisasi yang bergerak dalam bidang edukasi. Dalam penelitian ini terdapat 2 catatan kejadian yang termasuk ke dalam kategori *Education*. Catatan kejadian yang digunakan adalah catatan kejadian beberapa angkatan mahasiswa dari satu universitas

### 6.1 Analisis Data

Proses Analisis dari catatan kejadian dilakukan dengan mengamati pola dari setiap catatan kejadian yang digunakan pada studi kasus. Pengamatan dilakukan berdasarkan ciri-ciri yang telah disebutkan dalam [8] untuk masing-masing pola. Dari hasil temuan pada tabel 3, dapat diketahui bahwa pola yang muncul adalah *Form-based Event Capture*, *Unanchored Event*, *Elusive Case*, *Collateral Events* dan *Homonymous Label*.

Tabel 3. Hasil Identifikasi Pola Ketidaksempurnaan

Event Log	P01	P02	P03	P04	P05	P06	P07	P08	P09	P10	P11
	Manufacture										
(Arsad, 2013) [10]	√		√		√						√
(Yudananto, 2013) [11]	√		√		√						√
(Utami, 2013) [12]	√		√		√						√
(Alamuddin, 2014) [13]			√		√						
(Wardhani, 2014) [14]	√		√		√						
(Ningrum, 2014) [15]			√		√						
(Amaliyah, 2015) [16]	√		√		√						
(Savitrie, 2015) [17]			√								
(Lestari, 2015) [18]			√		√						
(Muhammad, 2015) [19]			√		√						
(Muhaimin, 2015) [20]			√		√						

(Zayin, 2017) [21]		√		√	
Healthcare					
(Riawan, 2016) [22]	√	√			√
(Pudyarstiani, 2016) [23]		√			√
(Syarifuddin, 2016) [24]		√			
(Suhermanto, 2016) [25]		√			
Education					
(Priyambada, 2016) [26]	√	√			√
(Arifin, 2016) [27]	√	√			√

**Form-based Event Capture** adalah pola yang diakibatkan karena sistem informasi terbentuk dari form-form [8]. Bentuk form akan mengakibatkan *timestamp* dari aktivitas yang berada dalam satu form sama. Hal tersebut terjadi pada 7 dari 16 catatan kejadian yang diamati. Pola ini terjadi pada setiap kategori, baik manufaktur, healthcare dan edukasi. Ketujuh catatan kejadian tersebut memiliki *timestamp* yang sama dengan aktivitas yang lain. Untuk menangani hal ini adalah dengan me-agregat aktivitas yang mengacu pada satu proses yang sama [8]. Pada studi kasus yang diamati sebagian besar tidak ada aktivitas yang di-agregat menjadi satu, namun solusinya adalah dengan membedakan *timestamp*-nya secara manual dengan melihat proses bisnis perusahaan.

Dapat dilihat pada tabel 4 yang merupakan salah satu contoh pola Form-based Event Capture yang ditemukan pada catatan kejadian manufaktur dimana aktivitas “RFQ Dibuat” dan “PO dibuat” memiliki *timestamp* yang sama. Hal ini akan mengakibatkan pemodelan proses bisnis tidak sesuai, *timestamp* yang sama akan digambarkan sebagai aktivitas paralel yang dilakukan bersamaan. Pada kenyataannya RFQ dibuat terlebih dahulu sebelum PO dibuat. Namun pada catatan kejadian manufaktur ditemukan pola bahwa semua aktivitas “RFQ dibuat” selalu memiliki *timestamp* yang sama dengan “PO dibuat”. Hal ini mengindikasikan bahwa pembuatan RFQ dan PO berada pada satu form yang sama.

Tabel 4. Contoh Form-based Event Capture pada catatan kejadian manufaktur

Case ID	Activity	Timestamp	Case ID	Activity	Timestamp
101-100-00201	PR Dibuat	12/12/2011 0:00	101-100-00201	PO dibuat	11/6/2012 0:00
101-100-00201	PR release	12/13/2011 0:00	101-100-00201	Bahan diterima	12/6/2012 0:00
101-100-00201	RFQ Dibuat	11/6/2012 0:00	101-100-00201	GR101	12/7/2012 10:09

Sedangkan untuk contoh dari Healthcare dapat dilihat pada tabel 5. Dapat terlihat jelas bahwa catatan kejadian yang dihasilkan dari sistem informasi rumah sakit memiliki bentuk form-form sehingga memiliki banyak aktivitas dengan *timestamp* yang sama. Solusi yang dapat dilakukan adalah dengan cara mengagregasi aktivitas yang mengacu pada satu aktivitas, contohnya adalah aktivitas “Administrasi Ruangan” dan “Sewa Kamar” yang mengacu pada administrasi rawat inap.

Tabel 5. Contoh Form-based Event Capture pada catatan kejadian healthcare

Nocm	Tgl Pelayanan	Nama Pelayanan
448617	1/1/2015 17:31	Karcis IGD Gawat
448617	1/1/2015 17:44	ADMINISTRASI RUANGAN
448617	1/1/2015 17:44	SEWA KAMAR
448617	1/1/2015 18:27	BUN
448617	1/1/2015 18:27	CREATININ SERUM
448617	1/1/2015 18:27	DARAH LENGKAP

**Unanchored Event** adalah pola yang diakibatkan karena format *timestamp* dari sistem informasi yang tidak biasa [8]. Dimana format yang biasa ditulis adalah format dd/mm/yyyy namun untuk kategori manufacture dan healthcare seluruhnya terekstrak dalam format mm/dd/yyyy. Sedangkan untuk education hanya terekstrak dalam bentuk tahun dan semester pengambilan saja. Untuk menangani hal ini adalah dengan memasukkannya pada aplikasi yang sudah biasa digunakan untuk melakukan *process mining*, yaitu Disco [8]. Disco dapat mengidentifikasi secara otomatis format dari *timestamp* pada catatan kejadian.

Contoh dari unanchored event dapat dilihat juga pada tabel 6, dimana terdapat tanggal “12/7/2011” yang menunjukkan tanggal 12 Juli 2011, setelah itu terdapat “12/13/2011”. Hal ini menyebabkan salah mengartikan

timestamp yang tersimpan dalam format mm/dd/yyyy, sehingga “12/7/2011” sebenarnya adalah 7 Desember 2011. Masalah ini dapat diselesaikan dengan memasukkan data pada aplikasi disco. Pada saat memasukkan data tersebut akan muncul peringatan “Your pattern does not match this value!”. Sehingga harus dipilih preset “MM/dd/yyyy HH:mm”.

**Elusive Case** adalah pola yang mengakibatkan kesulitan menentukan case id yang nantinya akan digunakan dalam catatan kejadian. Pada paper [8] ditunjukkan contoh catatan kejadian gps pada mobil, yang solusinya adalah membuat case id baru yang mengacu pada “perjalanan” bukan kendaraan.

Elusive Case ditemukan pada 10 dari 16 catatan kejadian yang mana seluruhnya merupakan catatan kejadian dari manufaktur. Pada studi kasus yang diamati elusive case yang didapatkan memiliki penyebab yang berbeda, yakni kesulitan menentukan case id karena terdapat banyak kolom yang dapat dijadikan identifikasi sebagai case id, seperti material, order id dan kolom lain. Untuk contoh dari pola ini dapat dilihat pada salah satu catatan kejadian manufaktur pada tabel 7. Terdapat beberapa kolom yang dapat dijadikan case.

Tabel 6 Contoh Elusive Case

Nomor Dokumen	Activity	No Material	Timestamp	Nomor Dokumen	Activity	No Material	Timestamp
1000066727	Ins Created	101-100-0020	12/7/2012 10:09	2100003870	RFQ Dibuat	101-100-0020	11/6/2012 0:00
2100003393	PR Dibuat	101-100-0020	12/12/2011 0:00	6700002343	PO dibuat	101-100-0020	11/6/2012 0:00
2100003393	PR release	101-100-0020	12/13/2011 0:00	5001039295	Bahan diterima	101-100-0020	12/6/2012 0:00

Untuk menentukan case perlu melihat dari sudut pandang mana analisis dari *process mining* yang akan dilakukan. Untuk data tersebut akan dilihat bagaimana jalannya kasus pengadaan barang dalam beberapa aktivitas, dimana satu case id menunjukkan satu kasus, dimana bahan tersebut terlebih dahulu dilakukan inspeksi. Sehingga untuk case id, perlu dilakukan “concatenate” beberapa kolom yaitu nomor material, nomor dokumen inspeksi dan nomor dokumen good receive, sehingga dihasilkan seperti pada tabel 7.

Tabel 7. Contoh Penyelesaian Elusive Case

Case ID	Activity	Timestamp
101-100-002010000667275001039295	Ins Created	12/7/2012 10:09
101-100-002010000667275001039295	PR Dibuat	12/12/2011 0:00
101-100-002010000667275001039295	PR release	12/13/2011 0:00
101-100-002010000667275001039295	RFQ Dibuat	11/6/2012 0:00
101-100-002010000667275001039295	PO dibuat	11/6/2012 0:00
101-100-002010000667275001039295	Bahan diterima	12/6/2012 0:00

Dari tabel tersebut dicontohkan untuk mendapatkan case ID “101-100-002010000667275001039295” adalah dengan menggunakan nomor material “101-100-0020” ditambah dengan nomor dokumen inspeksi “1000066727” dan ditambah nomor dokumen GR “5001039295”.

**Collateral Events** adalah terdapat nama aktivitas yang berbeda namun mengacu pada satu aktivitas yang sama. Pada paper [8], collateral events disolusikan dengan menggabungkan event yang mengacu pada satu aktivitas. Pada studi kasus, pola ini ditemukan pada kategori education.

Collateral events pada education memiliki sedikit perbedaan dengan yang muncul pada healthcare, yaitu pada education, collateral event terjadi karena penamaan aktivitas. Pada perkuliahan yang terekam pada sistem akademik, mahasiswa dapat mengambil mata kuliah “agama” yang dapat berupa mata kuliah agama Islam, Kristen, Katolik, Hindu dan Budha Sehingga sebaiknya penamaan dari kedua mata kuliah ini disamakan, yaitu mata kuliah “Agama”. Karena secara proses, dengan mengambil mata kuliah agama apapun mahasiswa tetap melewati aktivitas yang sama yaitu mata kuliah agama.

**Homonymous Label** adalah terdapat nama aktivitas yang sama namun memiliki arti aktivitas yang berbeda, hal ini mengakibatkan pola self-loop yang menjadikan model proses kurang bisa menggambarkan realita. Pada paper [8] disebutkan bahwa solusinya adalah melakukan *relabelling* yaitu mengubah nama dari aktivitas yang memiliki

nama yang sama. Untuk contoh education dapat dilihat pada tabel 8 menunjukkan mahasiswa mengambil mata kuliah Tugas Akhir yang terdapat pada kurikulum yang berbeda.

Tabel 8 Contoh Homonymous Label

Case ID	Kode MK	Mata Kuliah	SKS	Tahun	Case ID	Kode MK	Mata Kuliah	SKS	Tahun
520216	CF1380	Tugas Akhir	4	2008	520216	IG1311	Pengantar Technopreneurship	2	2009
520216	CF1380	Tugas Akhir	4	2008	520216	KS1320	Perencanaan Strategis SI/TI	3	2009
520216	KS1324	Etika Profesi	3	2009	520216	KS1336	Tugas Akhir	6	2009
520216	KS1316	Manajemen Rantai Pasok	3	2009	520216	KS1336	Tugas Akhir	6	2009
520216	IG1311	Pengantar Technopreneurship	2	2009	520216	KS1336	Tugas Akhir	6	2010

Dari tabel dapat dilihat bahwa mahasiswa mengulang mata kuliah Tugas Akhir lima kali, namun pada kurikulum yang berbeda. Mata kuliah Tugas Akhir dengan kode CF1380 merupakan kurikulum 2004, sedangkan kode KS1336 merupakan kurikulum 2009. Dengan jumlah sks yang berbeda, seharusnya keduanya merupakan sebuah aktivitas yang berbeda. Pada studi kasus dilakukan solusi dengan menggabungkan kode mata kuliah dengan nama aktivitas sehingga mata kuliah “Tugas Akhir” pada catatan kejadian memiliki perbedaan nama antara kurikulum satu dengan lainnya, begitu pula mata kuliah lain yang memiliki nama yang sama.

## 7. SIMPULAN DAN SARAN

Pada bagian ini akan dijelaskan mengenai kesimpulan dan saran yang didapatkan dari hasil penelitian yang telah dilakukan.

### 7.1 Simpulan

Permasalahan yang dihadapi dalam melakukan teknik *process mining* adalah masalah terkait catatan kejadian. Kualitas catatan kejadian yang buruk akan menghasilkan analisis terkait model proses menjadi buruk pula. Isu kualitas dari catatan kejadian telah didefinisikan menjadi empat isu utama yang mengasikkan 27 isu kualitas catatan kejadian. Dimana isu kualitas ini menjadi ciri-ciri yang membentuk *imperfection patter* dari catatan kejadian.

Penelitian ini melakukan analisis dan menemukan *pola ketidaksempurnaan* pada catatan kejadian dari berbagai perusahaan. *Pola ketidaksempurnaan* yang berhasil diidentifikasi adalah *Form-based Event Capture*, *Unanchored Event*, *Elusive Case*, *Collateral Events* dan *Homonymous Label*. *Pola ketidaksempurnaan* yang muncul pada semua catatan kejadian adalah *unanchored event*, dimana seluruh timestamp pada catatan kejadian tidak dalam format yang umum.

Bentuk baru dari *pola ketidaksempurnaan* yang ditemukan adalah bentuk lain dari *elusive case*, yaitu kesulitan untuk menentukan *case* dalam catatan kejadian. Dalam penelitian ini ditemukan *elusive case* yang berbeda, yaitu catatan kejadian memiliki beberapa kolom yang sebenarnya dapat dijadikan *case*, namun hal ini tergantung dari sudut pandang mana peneliti akan berfokus. Pada penelitian ini solusi dicontohkan dengan menggabungkan beberapa kolom untuk dijadikan *case id*.

Dengan pengamatan pada catatan kejadian dari berbagai tipe perusahaan terdapat beberapa *pola ketidaksempurnaan* yang dapat diidentifikasi serta dimungkinkan untuk menemukan *pola ketidaksempurnaan* baru.

### 7.2 Saran

Dalam penelitian ini jumlah catatan kejadian yang digunakan dalam masing-masing tipe perusahaan tidak memiliki jumlah yang sama sehingga tidak didapatkan mengenai kemunculan pola ketidaksempurnaan pada setiap tipe perusahaan. Dalam penelitian berikutnya dapat ditambahkan jumlah catatan kejadian dalam bidang edukasi yang memungkinkan adanya pola ketidaksempurnaan baru lainnya.

## 8. DAFTAR RUJUKAN

- [1] W. M. P. van der Aalst, *Process Mining: Discovery, Conformance and Enhancement of Business Process*. Berlin: Springer-Verlag, 2011.
- [2] R. P. J. C. Bose, “Wanna Improve Process Mining Results ? It ’ s High Time We Consider Data Quality Issues Seriously,” *Proc. 2013 IEEE Symp. Comput. Intell. Data Mining, CIDM 2013 - 2013*

- IEEE Symp. Ser. Comput. Intell. SSCI 2013*, no. 1, pp. 127–134, 2013.
- [3] F. E. Ayo, O. Folorunso, and F. T. Ibhralu, “A probabilistic approach to event log completeness,” *Expert Syst. Appl.*, vol. 80, pp. 263–272, 2017.
- [4] H. Yang, L. Wen, J. Wang, and R. K. Wong, “CPL+: An improved approach for evaluating the local completeness of event logs,” *Inf. Process. Lett.*, vol. 114, no. 11, pp. 607–610, 2014.
- [5] N. Tax, N. Sidorova, R. Haakma, and W. M. P. V. Der Aalst, “Log-based Evaluation of Label Splits for Process Models,” *Procedia Comput. Sci.*, vol. 96, pp. 63–72, 2016.
- [6] L. M??ru??ter, A. J. M. M. Weijters, W. M. P. Van Der Aalst, and A. Van Den Bosch, “A rule-based approach for process discovery: Dealing with noise and imbalance in process logs,” *Data Min. Knowl. Discov.*, vol. 13, no. 1, pp. 67–87, 2006.
- [7] H. J. Cheng and A. Kumar, “Process mining on noisy logs - Can log sanitization help to improve performance?,” *Decis. Support Syst.*, vol. 79, pp. 138–149, 2015.
- [8] S. Suriadi, R. Andrews, A. H. M. ter Hofstede, and M. T. Wynn, “Event log imperfection patterns for process mining: Towards a systematic approach to cleaning event logs,” *Inf. Syst.*, vol. 64, no. October 2015, pp. 132–150, 2017.
- [9] IBM-Corporation, “About event logs - IBM,” 2012. [Online]. Available: [https://www.ibm.com/support/knowledgecenter/en/SSNW2F\\_5.1.0/com.ibm.p8.pe.user.doc/bpfes000.htm](https://www.ibm.com/support/knowledgecenter/en/SSNW2F_5.1.0/com.ibm.p8.pe.user.doc/bpfes000.htm). [Accessed: 13-Jun-2017].
- [10] N. Arsad, “Pembuatan Model Proses Menggunakan Algoritma Heuristic Miner Untuk Analisis Interaksi Proses Bisnis Perencanaan Produksi Dan Pengadaan Material Di PT. XYZ,” Surabaya.
- [11] I. H. Yudananto, “Pembuatan Model Proses Bisnis Sap Erp Dalam Interaksi Antara Modul Materials Management Dan Production Planning Di PT.XYZ Dengan Algoritma Alpha++ Dan Algoritma Genetika,” Surabaya, 2013.
- [12] R. A. Utami, “Pemodelan dan Analisis Bottleneck Proses Bisnis Perencanaan Produksi di PT. XYZ pada SAP dengan Algoritma Genetik,” Surabaya, 2013.
- [13] M. A. S. Alamuddin, “Pemodelan dan Analisis Kinerja Proses Bisnis Distribusi Produksi dengan Algoritma Heuristic Miner pada Departemen Production Distribution Center di PT. XYZ,” Surabaya, 2014.
- [14] I. R. K. Wardhani, “Analisis Pergerakan Material untuk Meningkatkan Kinerja Proses di Gudang Material PT.XYZ Menggunakan Algoritma Heuristic Miner,” Surabaya, 2014.
- [15] S. Ningrum, “Analisis Pengaruh Volume dan Variasi Artikel terhadap Lead Time Penyelesaian Pengepakan di Producton Distribution Center PT. XYZ dengan Menggunakan Algoritma Duplicate Genetic,” Surabaya, 2014.
- [16] M. Amaliyah, “Pemodelan dan Analisis Kinerja Proses Bisnis Pengadaan Bahan di PT. XYZ dengan Teknik Penggalan Proses,” Surabaya, 2015.
- [17] R. A. Savitrie, “Analisa Tingkat Kinerja Proses Produksi Berdasarkan Ketepatan Waktu Penyelesaian Produksi Menggunakan Teknik Penggalan Proses pada Divisi Produksi PT. Farmasi,” Surabaya, 2015.
- [18] Y. E. Lestari, “Pemodelan dan Analisis Proses Pengadaan Aset Investasi Perusahaan di PT. XYZ dengan Algoritma Heuristic Miner,” Surabaya, 2015.
- [19] D. A. Muhammad, “Pembuatan Model dan Analisis Kinerja Proses Perencanaan Produksi hingga Ekspor Upper Shoes di PT. XYZ dengan Menggunakan Algoritma Heuristic Miner,” Surabaya, 2015.
- [20] M. N. Muhaimin, “Analisis Deviasi dan Bottleneck dalam Proses Bisnis Perencanaan Hingga Stok PT. XYZ dengan Process Mining,” Surabaya, 2015.
- [21] S. O. Zayin, “Evaluasi Penerapan Modul Manajemen Material (MM) SAP untuk Pengadaan Material di PT. Petrokimia Gresik,” Surabaya, 2017.
- [22] A. A. Riawan, “Analisis Hubungan Dokter yang Menangani Pasien Rawat Inap Diabetes Menggunakan Social Network Mining (Studi Kasus RSUD Haji Surabaya),” Surabaya.
- [23] D. Pudyarstiani, “Pemodelan dan Analisis Proses Pelayanan Pasien Rawat Jalan untuk Penyakit Jantung dengan Pendekatan Penggalan Proses (Studi Kasus: RS ABC Surabaya),” Surabaya, 2016.
- [24] M. R. Syarifuddin, “Analisis Pola Hubungan Kerja antar Tenaga Medis dalam Melayani Pasien Diabetes pada Rawat Inap RS XYZ Menggunakan Teknik Process Mining,” Surabaya, 2016.
- [25] S. Suhermanto, “Pemodelan dan Analisis Hubungan Antar Dokter yang Menangani Pasien Rawat Inap Pengidap Penyakit Jantung dan Komplikasinya Menggunakan Teknik Social Network Mining (Studi Kasus: Rumah Sakit XYZ),” Surabaya, 2016.
- [26] S. A. Priyambada, “Analisa Pola Pengambilan Mata Kuliah dan Kinerja Mahasiswa Tiap Angkatan dengan Menggunakan Teknik Process Mining (Studi Kasus : Jurusan Sistem Informasi ITS),” Surabaya, 2016.
- [27] F. K. Arifin, “Pemodelan dan Analisis Pengambilan Mata Kuliah Terhadap Kurikulum dengan Teknik Penggalan Proses (studi Kasus Jurusan Sistem Informasi ITS),” Surabaya, 2016.