

Pendekatan historis dengan standar PROV: Kajian Konseptual Penerapan Lintas Domain

Belfrit Victor Batlajery^{1*}, Goldy Valendria Nivaan², Hennie Tuhuteru³, Dessy Gloria Palyama⁴, Golda Tomasila⁵, Marthevienty H T Soumokil⁶

^{1,2,3}Informatika, Universitas Kristen Indonesia Maluku, Indonesia

^{4,5,6}Sistem Informasi, Universitas Kristen Indonesia Maluku, Indonesia

^{1*}belfritvictorbatlajery@ukim.ac.id, ²valendria17@gmail.com,

³hannytuhuteru@gmail.com, ⁴dessypalijama70@gmail.com, ⁵tomasilagolda@gmail.com,

⁶marthevienty.hts@gmail.com

Abstrak: Penelusuran asal-usul atau riwayat modifikasi dari suatu entitas seringkali menjadi tantangan tersendiri. Ini disebabkan oleh keterlibatan banyak proses dan pihak dalam rantai produksi dan distribusi yang kompleks, sehingga akuntabilitas dan kepercayaan terhadap data yang dihasilkan menjadi sulit untuk diverifikasi. Penelitian ini merupakan kajian konseptual yang bertujuan mengenalkan standar PROV dari W3C yang mencakup PROV-DM, PROV-O, dan komponen lainnya sebagai mekanisme umum dalam mencatat dan mempertukarkan informasi *provenance* lintas domain. Dengan demikian pertukaran informasi *provenance* akan lebih mudah pada ekosistem yang lebih heterogen. Kajian difokuskan pada tiga komponen inti PROV-DM (*Entity, Activity, Agent*) beserta representasinya dalam bentuk *Provenance Graph* (PG), serta pemetaannya ke PROV-O untuk penerapan pada ranah spesifik. Pembahasan disertai sintesis dari sejumlah studi terdahulu yang menerapkan PROV pada berbagai domain (misalnya dalam ranah pangan, kesehatan, *marketplace*, pengambilan keputusan berbasis risiko, dll). Hasil kajian ini menunjukkan serangkaian tahapan yang sistematis dan saling menguatkan, meliputi standarisasi informasi, representasi dalam bentuk *graph*, dan integrasi PROV ke dalam ranah yang spesifik. Tahapan-tahapan ini dibuktikan melalui investigasi dari penelitian terdahulu pada berbagai ranah. Pendekatan ini terbukti dapat diterapkan pada domain yang berbeda dalam pemodelan *provenance*. Dengan penelitian ini, dapat disimpulkan bahwa PROV memang memiliki fleksibilitas yang baik untuk diterapkan lintas domain sehingga dapat menjadi langkah penting dalam implementasi sistem yang lebih mudah untuk diaudit dan ditelusuri.

Kata Kunci: Provenance; PROV; Pendekatan Historis; Standarisasi; Pemodelan;

Abstract: The tracing of the origin or modification history of an entity is often a complex challenge. This is due to the involvement of multiple processes and parties in complex production and distribution chains, making it difficult to verify accountability and trust in the resulting data. This study is a conceptual study aimed at introducing the W3C PROV standards, which include PROV-DM, PROV-O, and other components as a general mechanism to record and exchange *provenance* information across domains. Thus, the exchange of *provenance* information can be made easier in the more heterogeneous ecosystems. Our study focuses on the three core components of the PROV-DM (*Entity, Activity,*

Agent), along with the *Provenance Graph* (PG) as their representation, and the mapping to the PROV-O for implenting it in spesific domains. The discussion comprises a synthesis of previous studies around PROV implemetation across various domains (e.g., food, healthcare, marketplace, risk-based decision-making, etc.). The results of this research demonstrate a series of systematic stages, including information standardization, graph representation, and the integration of *provenance* into specific domains. Our approach has been proven applicable across different domains for modelling *provenance* information. Finally, this study concludes that PROV has good flexibility for cross-domain implementation, making it an important step toward developing systems that are easier to audit and trace.

Keywords: Provenance; PROV; Historical Approach; Standardization; Modeling;

1. PENDAHULUAN

Dalam pembentukan “sesuatu” (benda, hal, entitas, dll), proses yang melibatkan produksi dan distribusi “sesuatu” tersebut merupakan faktor yang penting dalam memahami asal-muasalnya. Kesulitan sering dialami ketika sumber utama dari “sesuatu” tersebut ingin diketahui dikarenakan berbagai proses perubahan atau modifikasi yang dialami dalam pembuatan dan distribusi hingga sampai pada bentuk akhirnya. Disamping itu, keterlibatan banyak orang dalam berbagai proses membuat rumit untuk mengidentifikasi siapa yang tanggung jawab pada proses yang mana. Sebagai contoh, sebuah paragraf yang ditulis oleh seseorang mungkin telah berubah bentuk oleh orang lain melalui proses pengeditan (misalnya menambah atau mengurangi kata). Dikarenakan paragraf tersebut akan selalu mengalami perubahan, hal ini menjadi sulit untuk menelusuri setiap perubahan.

Oleh sebab itu, konsep dari *provenance* diperkenalkan dengan tujuan untuk memperjelas hal-hal apa yang sudah mengalami perubahan, siapa saja yang bertanggung jawab untuk perubahan-perubahan tersebut, layanan (*service*) apa saja yang digunakan sehingga perubahan terjadi, dan bagaimana hasil akhir diperoleh. Secara khusus, *World Wide Web Consortium* (W3C) mendefinisikan *provenance* sebagai catatan informasi yang mendeskripsikan orang, institusi, entitas dan aktifitas dalam memproduksi, mempengaruhi, atau menyampaikan data atau entitas di dunia [16]. Catatan tersebut berisi data dan semua proses yang terlibat selama data tersebut ada, seperti bagaimana data tersebut diterima, siapa yang bertanggung jawab untuk aksi-aksi tertentu, konsekuensi apa saja yang akan terjadi untuk setiap aksi-aksi yang dibuat, resiko apa yang muncul selama data diproduksi, dsb.

Dengan kemampuan ini *provenance* sering kali dipandang sebagai bukti sejarah dari sebuah kejadian [6][20], dan dapat digunakan sebagai informasi yang penting untuk menentukan apakah segala sesuatu dapat dipercaya atau tidak [8][22], dan apakah hasil setelah proses komputasi masih relevan atau tidak. Kemampuan *provenance* inilah yang dapat membantu akuntabilitas dari sebuah sistem atau aplikasi yang dapat dimanfaatkan ketika melibatkan serangkaian proses, seperti rantai pasokan barang (*supply chain*). Bahkan, informasi *provenance* dapat mendukung penalaran berbasis catatan sejarah yang dapat membantu pengambilan keputusan atau untuk menemukan akar dari sebuah permasalahan.

Beberapa penelitian dilakukan sebagai upaya untuk menyimpan informasi *provenance* ini. Tang et. al. menyediakan pendekatan untuk mencatat provenance dari aktifitas pada basis data [21]. Pendekatan ini menyediakan mekanisme anotasi untuk memahami silsilah dari aktifitas pemrosesan data pada basis data relasional. Sementara itu, Jacques-Silva et al. mengembangkan sistem manajemen yang merekam silsilah data dan informasi secara

terpusat bernama Unified Lineage System [7]. Kemudian, Bhat juga memperkenalkan sebuah pendekatan untuk mencatat *provenance* pada sistem file dengan menelusuri eksekusi dari setiap file pada sistem *file* FAT [4]. Adapula *Data Provenance Initiative* yang diterapkan di ranah kecerdasan buatan untuk menelusuri asal-usul dan silsilah dari dataset yang digunakan [11]. Contoh lainnya adalah pendekatan data spasial yang mencatat *provenance* dari setiap datanya [5].

Namun permasalahannya, pendekatan-pendekatan tersebut diimplementasikan dengan cara yang berbeda tergantung pada bidang/ranahnya. Karena pendekatan-pendekatan tersebut sangat terfokus dan terbatas, maka dibutuhkan mekanisme sistematis untuk mencatat informasi *provenance* secara standar dan umum agar dapat diterapkan pada semua bidang. Selain itu, dari representasi dan format data pada pendekatan-pendekatan sebelumnya sangat berbeda-beda. Pencatatan *provenance* pada pendekatan-pendekatan sebelumnya cenderung menggunakan skema metadata dan struktur penyimpanan yang berbeda-beda sesuai dengan kebutuhan domainnya. Tanpa adanya keseragaman format, masalah interoperabilitas antar sistem muncul ketika informasi *provenance* perlu digabungkan atau dipertukarkan. Misalnya pada skenario dimana mengharuskan beberapa sistem yang berbeda untuk berkomunikasi (*supply chain*). Situasi sistem yang heterogen seperti ini banyak melibatkan proses-proses yang perlu diintegrasikan. Akibatnya, proses integrasi dan pertukaran informasi *provenance* menjadi tidak efisien dan rawan terhadap terjadinya ketidak-konsistenan data atau informasi.

Kendala lainnya adalah kemampuan proses informasi *provenance* yang otomatis oleh mesin. Dikarenakan sebagian besar pendekatan sebelumnya belum mengadopsi representasi informasi dengan PROV, maka proses analisa lanjutan untuk penelusuran informasi, deteksi anomali, maupun audit secara otomatis hanya dikembangkan secara terisolasi untuk masing-masing sistem. Dengan kata lain, terdapat kesenjangan antara kebutuhan untuk mekanisme *provenance* yang bersifat umum, berstandarisasi, dan dapat diproses secara otomatis oleh mesin, dengan kondisi pendekatan sebelumnya yang masih bersifat parsial dan spesifik. Oleh sebab itu, dibutuhkan mekanisme yang sistematis dalam menyelesaikan masalah diatas, yaitu keterbatasan cakupan domain untuk informasi *provenance*, ketidakberagaman format dalam pemrosesan informasi *provenance*, dan kurang optimalnya pemrosesan informasi *provenance* secara otomatis dalam sistem yang heterogen.

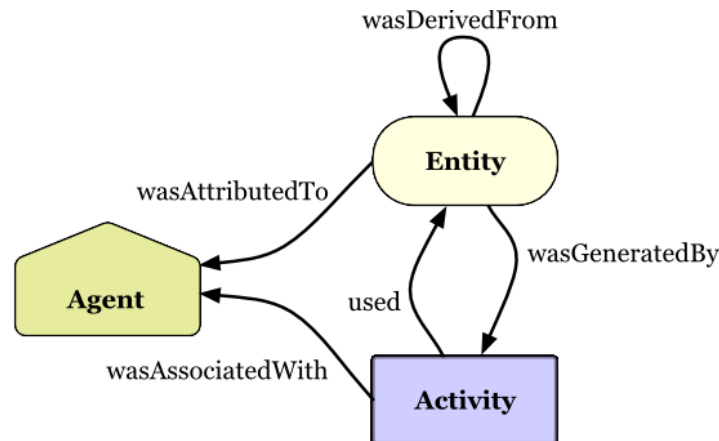
Pada penelitian ini, kami mengadopsi mekanisme standar pencatatan dan pertukaran informasi *provenance* yang dikenal dengan PROV. PROV dikembangkan oleh W3C dan ditujukan untuk memproses data *provenance* dengan mesin/komputer. PROV juga mempunyai setidaknya 3 keuntungan. Pertama, PROV merupakan mekanisme standar untuk merekam informasi *provenance*. Kedua, PROV bersifat abstrak dan dapat diekspansi ke masing-masing ranah yang lebih khusus sesuai kebutuhannya. Ketiga, PROV dapat direpresentasikan kedalam bentuk *graph* yang disebut *Provenance Graph* (PG), sehingga dapat dilakukan analisis lanjutan yang biasa diterapkan terhadap sebuah *graph*. Pada bab-bab selanjutnya, penelitian ini menyajikan Metode Penelitian yang digunakan di bab 2, Hasil dan Pembahasan di bab 3, dan Kesimpulan di bab 4.

2. METODE PENELITIAN

Pada bagian ini, kami memperkenalkan metode kami dalam melakukan pendekatan historis dengan PROV. Pertama-tama kami akan memperkenalkan konsep dari *provenance* beserta kerangka standar PROV yang digunakan untuk mencatat informasi *provenance*. Berikutnya adalah visualisasi dengan *Provenance Graph* (PG). Setelah itu akan dipaparkan pemetaan agar PROV dapat diterapkan juga kedalam bidang-bidang yang lebih khusus.

2.1. Standarisasi informasi historis

Tahapan pertama adalah upaya merekam *provenance* dengan mekanisme yang standar. Untuk itu, penelitian ini menggunakan PROV yang terdiri dari 3 konsep utama yaitu konsep entitas (*prov:Entity*), konsep aktifitas (*prov:Activity*), dan konsep agen (*prov:Agent*). Gambar 1 menggambarkan ketiga konsep tersebut beserta hubungan antar konsep, serta Tabel 1 dan Tabel 2 menjelaskan tiap-tiap konsep.



Gambar 1. Tiga konsep utama PROV

Tabel 1. Pengertian dari *Entity*, *Activity*, dan *Agent* menurut PROV

Konsep	Pengertian
<i>Entity</i>	Suatu hal yang bersifat fisik, digital, konseptual, atau jenis lainnya yang mempunyai aspek-aspek tertentu yang tetap. Hal tersebut dapat berupa sesuatu yang nyata atau bersifat imajinasi.
<i>Activity</i>	Segala sesuatu yang terjadi selama kurun waktu tertentu dan menindaki <i>Entity</i> . Cakupan <i>Activity</i> meliputi mengkonsumsi, memroses, mentransformasi, memodifikasi, memindahkan, atau menghasilkan <i>Entity</i> .
<i>Agent</i>	Segala sesuatu yang dibebankan tanggung-jawab terhadap terjadinya <i>Activity</i> , terciptanya <i>Entity</i> , atau aktifitas dari <i>Agent</i> lainnya.

Tabel 2. Hubungan antar konsep menurut PROV

Konsep	Notasi	Pengertian
<i>Derivation</i>	<i>WasDerivedFrom</i>	Terbentuknya <i>Entity</i> baru berdasarkan <i>Entity</i> yang sudah ada sebelumnya
<i>Generation</i>	<i>wasGeneratedBy</i>	Selesaiannya sebuah <i>Activity</i> untuk memproduksi <i>Entity</i> baru.
<i>Usage</i>	<i>used</i>	Penggunaan suatu <i>Entity</i> oleh sebuah <i>Activity</i> .
<i>Attribution</i>	<i>wasAttributedTo</i>	Pemberian kepemilikan/asal suatu <i>Entity</i> kepada suatu <i>Agent</i> .
<i>Association</i>	<i>wasAssociatedWith</i>	Pemberian tanggung jawab kepada suatu <i>Agent</i> untuk sebuah <i>Activity</i> .

2.2. Representasi dalam Graph

Belfrit Victor Batlajery: * Penulis Korespondensi

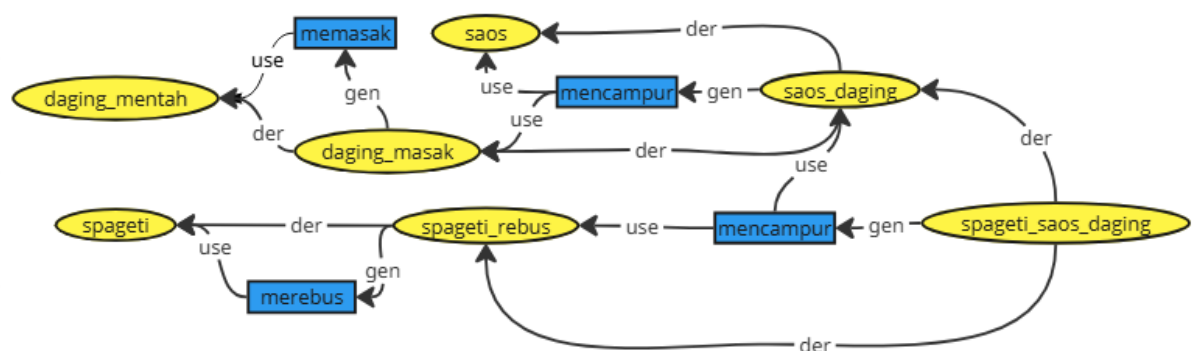


Copyright © 2026, Belfrit Victor Batlajery, Goldy Valendria Nivaan, Hennie Tuhuteru, Dessy Gloria Palyama, Golda Tomasila, Marthevienty H T Soumokil.

Pada penelitian ini, kami memilih PROV sebagai standar pemodelan karena kemampuannya untuk merepresentasikan informasi *provenance* sebagai *graph*, yang dikenal dengan *Provenance Graph* (PG) [15]. *Provenance Graph* diciptakan dengan menerjemahkan format standar dokumen *provenance* (PROV-O, PROV-XML, atau PROV-N [9]) kedalam bentuk *graph* asiklik berarah (*directed-acyclic graph*). Pada PG, hubungan antara ketiga konsep utama *provenance* (*prov:Entity*, *prov:Activity*, dan *prov:Agent*) direpresentasikan dengan panah **dari** dan/atau **ke** salah satu konsep tersebut. Sebagai contoh, Gambar 2 menampilkan bagaimana dokumen *provenance* direpresentasikan sebagai PG berdasarkan PROV-N pada Tabel 3.

Tabel 3. PROV-N dari sebuah spageti dengan saos daging

PROV-N
<pre> ... activity(mencampur_b,-,-,[prov:type="prFood:mencampur"%%xsd:string]) entity(spageti_rebus,[prov:type="prFood:spageti_rebus"%%xsd:string, prFood:stageDetails="spageti_rebus"%%xsd:string]) entity(saos_daging,[prov:type="prFood:saos_daging"%%xsd:string, prFood:stageDetails="saos_daging"%%xsd:string]) entity(spageti_saos_daging,[prov:type="prFood:spageti_saos_daging"%%xsd:string, prFood:stageDetails="spageti_saos_daging"%%xsd:string]) wasGeneratedBy(spageti_rebus,merebus,-) wasGeneratedBy(saos_daging,mencampur_a,-) wasGeneratedBy(spageti_saos_daging,mencampur_b,-) used(mencampur_b,saos_daging,-) used(mencampur_b,spageti_rebus,-) wasDerivedFrom(spageti_saos_daging,saos_daging) wasDerivedFrom(spageti_saos_daging,spageti_rebus) ... </pre>



Gambar 2. Provenance Graph dari sebuah spageti dengan saos daging

Gambar 2 menunjukan langkah demi langkah ketika sebuah spageti sudah dihidangkan. Dikarenakan terbatasnya tempat, kami menyingkat hubungan antara *prov:Entity* dan *prov:Activity* menjadi *use* (*prov:used*), *gen* (*prov:wasGeneratedBy*), dan *der* (*prov:wasDerivedFrom*). Proses *provenance* dari spageti tersebut dimulai dari **daging mentah** yang dimasak (melalui proses **memasak**) dan dicampur (melalui proses **mencampur**) dengan **saos** sehingga menjadi **saos daging**. Sementara itu, **spageti** direbus dan dicampur dengan **saos daging** sehingga menjadi **spageti saos daging** diakhir proses. Pemodelan seperti ini dapat diartikan juga sebagai rantai perjalanan asal-

usul dari **spageti saos daging**. Dengan kata lain, perjalanan historis dari sebuah entitas (yaitu **spageti saos daging**) dapat dimodelkan dengan PROV-DM sehingga merepresentasikan rantai perjalanan dari entitas tersebut.

2.3. Integrasi provenance pada ranah spesifik

Untuk merekam dan menstruktur secara detail informasi dari kejadian lampau, PROV menggunakan PROV-O. PROV-O merupakan sebuah ontologi sederhana yang menspesifikasikan kelas-kelas, properti-properti, dan batasan-batasan terhadap informasi yang dimodelkan dengan PROV-DM [15]. Sehingga tahapan berikutnya, yaitu proses pemetaan terhadap PROV-O, dapat dilakukan [3][18].

Sebagai contoh hewan dapat dikategorikan sebagai kelas dan seekor ikan sebagai sebuah individu dari kelas tersebut [14]. Kemudian properti-properti dari obyek (*object properties*) digunakan untuk membuat koneksi/hubungan antar kelas atau individu. Adapun properti-properti dari tipe data (*datatype properties*) digunakan untuk merekam nilai-nilai yang dimiliki oleh kelas atau individu tersebut. Sebagai contoh, seekor ikan dapat dimodelkan sebagai seekor individu yang mempunyai **usia** sebagai properti obyeknya dan **6 bulan** sebagai properti tipe datanya. Sebagai ontologi PROV-O menyediakan pengetahuan utama dalam merepresentasikan, mempertukarkan, dan mengintegrasikan aplikasi-aplikasi yang memproses informasi *provenance*.

Dengan demikian, PROV-O merupakan ontologi yang netral dan bebas dari domain apapun [18]. Akibatnya, ontologi spesifik lainnya dapat memperluas menggunakan PROV-O untuk tujuan khusus mereka. Beberapa penelitian yang menggunakan mekanisme ini antara lain, Pandit dan Lewis yang mengembangkan ontologi GDPRov yang memungkinkan penelusuran pada layanan pemrosesan data [17]. Kemudian penelitian Markovic et al. juga memperluas PROV-O (melalui P-PLAN) untuk mendokumentasikan rencana alur kerja yang dikenal dengan EP-PLAN [12]. Pada ranah sosial Riveni et al. menggunakan *provenance* untuk mengukur reputasi dan kepercayaan berbasis komunitas [19]. Pemetaan yang lainnya adalah pemetaan PROV-O terhadap *marketplace* (Batlajery, 2021) dan terhadap resiko untuk pengambilan keputusan (Batlajery et al., 2026).

3. HASIL DAN PEMBAHASAN

Pada penelitian ini, hasil yang kami peroleh adalah serangkaian tahapan yang sistematis dalam memanfaatkan standar PROV ketika melakukan pendekatan historis terhadap informasi. Ketiga tahapan tersebut—standarisasi informasi historis, representasi informasi historis dengan *graph*, dan integrasi *provenance* pada ranah spesifik—merupakan tahapan yang saling berkaitan membentuk alur kerja (*workflow*). Karena melibatkan *provenance*, alur kerja ini diawali dengan pemodelan informasi hingga kemampuannya untuk melakukan pengawasan (*monitoring*). Pendekatan ini juga bersifat cukup umum sehingga mudah diimplementasikan ke ranah-ranah yang lebih spesifik (contohnya, pangan, risiko, kesehatan, teknologi informasi, dll). Sifat generik ini sejalan dengan karakteristik PROV-DM sebagai model data yang memang dirancang agar dapat dipertukarkan antarsistem dalam ekosistem yang heterogen [15].

Pada tahap standarisasi, hasil penelitian ini menjelaskan bahwa penerapan PROV-DM tidak hanya terbatas pada konsep belaka, tetapi juga pada tahap implementasi. Hal ini dibuktikan dari implementasi sistem berbasis *provenance*, seperti yang dilakukan oleh Tang et al. dengan ProvPg [21], Jacques-Silva et al. dengan *Unified Lineage System* (ULS) [17], dan Bhat dengan pengolahan file FAT-nya [4]. Kemudian pada tahapan representasi *graph*, penelitian ini juga menunjukkan bahwa transformasi informasi *provenance* ke bentuk *Provenance Graph* (PG) memungkinkan penelusuran atas originalitas sebuah entitas. Ditambah lagi, bentuk dalam *graph* membuat algoritma-algoritma berbasis *graph*

dapat pula dilakukan lebih mendalam lagi untuk mengeksplorasi pola-pola yang tersembunyi.

Secara keseluruhan, ketiga tahapan hasil penelitian ini membentuk satu kesatuan alur kerja yang saling menguatkan. Standarisasi menyediakan data yang konsisten dan dapat dipertukarkan antar sistem, representasi graph dengan *Provenance Graph* (PG) mengubah fondasi tersebut menjadi struktur yang dapat ditelusuri dan dianalisis, dan integrasi PROV-O dengan ranah spesifik memastikan struktur tersebut relevan dengan kebutuhan. Temuan pada penelitian ini menunjukkan bahwa kerangka standar PROV memiliki fleksibilitas yang baik untuk diadopsi lintas domain/ranah tanpa kehilangan konsistensi dari inti permasalahan yang ingin diselesaikan pada ranah-ranah spesifik tersebut.

Fleksibilitas ini bukan hanya sebuah argumen teoritis semata, namun didukung oleh bukti empiris pada ranah yang beragam. Sebagai contoh, Batlajery et al. [3] memperluas PROV-O melalui prFood untuk memodelkan risiko kontaminasi makanan, Pandit dan Lewis [17] yang memperluas PROV-O melalui GDPRov untuk standar regulasi GDPR (*General Data Protection Regulation*). Keberagaman ini memperkuat temuan pada penelitian ini bahwa PROV-DM dan PROV-O memang dirancang sebagai kerangka generik yang tidak terikat pada satu domain saja. Hal ini menjadi landasan penting bagi implementasi sistem berbasis *provenance* lebih luas lagi, sekaligus membuka peluang penelitian lanjutan pada aspek otomatisasi pemetaan ontologi dan optimalisasi kinerja kueri pada PG berskala besar.

Walaupun PROV memiliki fleksibilitas dan sifat-sifat yang generik, ini tidak terlepas dari kelebihan dan kekurangannya. Kelebihan dari PROV sudah diketahui dan dipaparkan diatas meliputi kemampuannya untuk diadopsi tanpa harus merombak struktur data yang sudah ada pada masing-masing sistem. Ini terbukti dari penerapannya mulai dari ranah atau domain basis data [21], sistem file, pangan [3], sampai ke ranah regulasi [17]. Kekurangan yang tidak dapat dihindari adalah kompleksitas dari proses pemetaan akibat ekspansi dari PROV-O, seperti yang terlihat pada prFood [3] dan GDPRov [17]. Proses pemetaan ini tidak sederhana karena setiap domain memiliki semantik yang berbeda-beda berdasarkan ruang lingkungannya. Akibatnya, proses ini dapat berpotensi menghasilkan interpretasi yang berbeda pada level pengguna sehingga justru mengurangi interoperabilitas dari PROV itu sendiri. Disamping itu, representasi dari PG memang terlihat menarik apabila jumlah informasi *provenance* sangat banyak. Dalam skala yang besar, PG tidak dapat diinterpretasikan manual dan proses komputasinya bisa mahal apabila tidak disertai dengan mekanisme pengindexan atau penyimpanan *graph* yang efisien. Hambatan terakhir bersifat operasional. Ini dikarenakan pencatatan *provenance* merupakan pekerjaan tambahan diatas proses yang sudah ada sebelumnya. Dengan demikian perlu dipertimbangkan adanya penambahan tenaga dan biaya untuk pengembangan dan pemeliharaannya, diluar operasional utama yang sudah berjalan. Sehingga, implementasi dari PROV di lapangan tidak hanya bergantung pada kesiapan teknis saja, tetapi juga kesiapan tata kelola.

Dengan mempertimbangkan kelebihan, kekurangan dan tantangan diatas, dapat disimpulkan bahwa PROV memang memberikan dasar yang kuat dan fleksibel dalam mencatat informasi *provenance* lintas domain. Namun keberhasilan implementasinya tetap bergantung pada sejauh mana proses standarisasi serta integrasi teradap ranah yang spesifik secara tepat. Sehingga, hasil kajian ini menegaskan bahwa adopsi atau implementasi dari PROV bukanlah pekerjaan yang bersifat langsung pakai (*plug-and-play*), melainkan memerlukan penyesuaian ontologi dari ranah yang spesifik, pertimbangan beban kerja tambahan, dan kesiapan tata kelola.

4. KESIMPULAN

Pada penelitian ini, kami menggunakan standar PROV untuk melakukan pendekatan historis terhadap informasi. PROV dengan komponen-komponennya (PROV—DM, PROV-O, PROV-N, dll) mampu menjembatani pemodelan informasi ketikan histori atau rekam jejak merupakan aspek penting dalam pemodelan. Ini dikarenakan informasi yang dimodelakan dengan PROV-DM dapat direpresentasikan kedalam bentuk *graph*, *Provenance Graph* (PG) yang notabene merekam rekam-jejak atau histori dari perubahan-perubahan yang telah terjadi.

Pemodelan yang dilakukan dengan PROV-DM secara tidak langsung melibatkan juga proses pemetaan antar ontologi. Pemetaan ini merupakan penghubung antara konsep umum yang ditawarkan oleh PROV-O dengan konsep khusus dari ranah yang spesifik (*conceptual relationship*). Dengan demikian pendekatan pada penelitian ini menjembatani, dengan mengintegrasikan, kedua konsep tersebut. Hasilnya, *provenance* dapat dikontekstualkan secara spesifik kedalam domain yang berbeda-beda.

Kajian ini sekaligus menekankan akan 3 kontribusi utamanya. Pertama adalah perumusan sistematis alur kerja dengan PROV, yaitu standarisasi informasi historis, representasi dalam PG, dan integrasi generik *provenance* ke ranah yang spesifik. Kedua, kajian ini juga menegaskan bahwa PROV dapat menjadi penghubung konseptual antara konsep generik dengan konsep yang spesifik, walaupun implementasinya tidak mudah. Dan cukup kompleks. Ketiga, artikel ini mengkaji sifat-sifat generik dan fleksibilitas dari PROV, serta juga memaparkan kelebihan dan kekurangannya; sehingga, memberikan gambaran lengkap dibandingkan hanya sekedar implementasi secara langsung.

Akhir kata, kajian ini merupakan landasan dasar yang dapat digunakan dalam pengembangan implementasi sistem berbasis *provenance* diberbagai domain. Dengan sistem yang berbasis *provenance*, tindakan pengawasan dan pencegah juga dapat diterapkan dengan memanfaatkan algoritma-algoritma yang berjalan diatas PG. Kajian ini selanjutnya juga diharapkan dapat memperluas kontribusi yang sudah ada, terutama pada aspek otomatisasi dalam pemetaan ontologi dan optimalisasi kinerja kueri pada PG dalam skala besar. Dengan demikian dapat dilakukan analisa lanjutan, sehingga memungkinkan hubungan antara entitas dapat ditelusuri serta dianalisis secara sistematis dengan optimal.

5. REFERENCES

- [1] B. V. Batlajery, T. M. Tamtelahitu, J. C. Patty, and H. Tuhuteru, "Incorporate provenance for relevant decision making," in *Information Management: 11th Int. Conf., ICIM 2025, Revised Selected Papers, Part I*, S. Li, Ed., Communications in Computer and Information Science, vol. 2540. Cham, Switzerland: Springer, 2026, doi: 10.1007/978-3-031-99353-4_2.
- [2] B. V. Batlajery, "Modeling and generating marketplace activities with PROV-DM," in *2021 4th Int. Conf. Artificial Intelligence for Industries (AI4I)*, Laguna Hills, CA, USA, 2021, pp. 91–94.
- [3] B. V. Batlajery, M. Weal, A. Chapman, and L. Moreau, "prFood: Ontology principles for provenance and risk in the food domain," in *2018 IEEE 12th Int. Conf. Semantic Computing (ICSC)*, Laguna Hills, CA, USA, 2018, pp. 17–24, doi: 10.1109/ICSC.2018.00012.
- [4] W. A. Bhat, "Provenance maintenance of FAT file system for consumer electronic devices," *Arabian J. Sci. Eng.*, vol. 50, no. 23, pp. 19391–19401, 2025, doi: 10.1007/s13369-025-10201-1.
- [5] J. Fischer et al., "Approaches and tools for user-driven provenance and data quality information in spatial data infrastructures," *Int. J. Digit. Earth*, vol. 16, no. 1, pp. 1510–1529, 2023, doi: 10.1080/17538947.2023.2198778.

- [6] B. Glavic, "Data provenance: Origins, applications, algorithms, and models," *Found. Trends Databases*, vol. 9, no. 3–4, pp. 209–441, 2021.
- [7] G. Jacques-Silva et al., "Unified lineage system: Tracking data provenance at scale," in *Companion 2025 Int. Conf. Management of Data*, 2025, pp. 457–470, doi: 10.1145/3722212.3724458.
- [8] M. Johns, T. Meurers, F. N. Wirth, A. C. Haber, A. Müller, M. Halilovic, F. Balzer, and F. Prasser, "Data provenance in biomedical research: Scoping review," *J. Med. Internet Res.*, vol. 25, p. e42289, 2023, doi: 10.2196/42289.
- [9] M. A. Jarwar, A. Chapman, M. Elliot, and F. Raji, "Provenance, anonymisation and data environments: A unifying construction," *arXiv preprint*, arXiv:2107.09966, 2021.
- [10] T. Lebo, S. Sahoo, and D. McGuinness, Eds., *PROV-O: The PROV Ontology*. W3C Recommendation, World Wide Web Consortium (W3C), Apr. 30, 2013. [Online]. Available: <https://www.w3.org/TR/prov-o/>
- [11] S. Longpre et al., "The Data Provenance Initiative: A large scale audit of dataset licensing & attribution in AI," *arXiv preprint*, arXiv:2310.16787, 2023.
- [12] M. Markovic, D. Garijo, and P. Edwards, "Linking abstract plans of scientific experiments to their corresponding execution traces," in *Proc. 3rd Int. Workshop on Capturing Scientific Knowledge (Sciknow 2019)*, CEUR-WS, 2019.
- [13] M. Markovic, P. Edwards, and D. Corsar, "SC-PROV: A provenance vocabulary for social computation," in *Int. Provenance and Annotation Workshop (IPAW 2014)*, Cham, Switzerland: Springer, 2014, pp. 285–287, doi: 10.1007/978-3-319-16462-5_35.
- [14] K. D. Mogotlane and J. V. Fonou-Dombeu, "Automatic conversion of relational databases into ontologies: A comparative analysis of Protégé plug-ins performances," *arXiv preprint*, arXiv:1611.02816, 2016.
- [15] L. Moreau et al., *PROV-DM: The PROV Data Model*. W3C Recommendation, World Wide Web Consortium (W3C), Apr. 30, 2013. [Online]. Available: <https://www.w3.org/TR/prov-dm/>
- [16] L. Moreau and P. Missier, Eds., *PROV-N: The Provenance Notation*. W3C Recommendation, World Wide Web Consortium (W3C), Apr. 30, 2013. [Online]. Available: <https://www.w3.org/TR/prov-n/>
- [17] H. J. Pandit and D. Lewis, "Modelling provenance for GDPR compliance using linked open data vocabularies," in *Proc. 5th Workshop on Society, Privacy and the Semantic Web - Policy and Technology (PrivOn 2017)*, 2017.
- [18] T. Prudhomme, G. De Colle, A. Liebers, A. Sculley, P. K. Xie, S. Cohen, and J. Beverley, "A semantic approach to mapping the Provenance Ontology to Basic Formal Ontology," *Sci. Data*, vol. 12, p. 282, 2025, doi: 10.1038/s41597-025-04580-1.
- [19] M. Riveni, T. D. Nguyen, M. S. Aktas, and S. Dustdar, "Application of provenance in social computing: A case study," *Concurrency Comput.: Pract. Exp.*, vol. 31, no. 3, p. e4894, 2019, doi: 10.1002/cpe.4894.
- [20] B. Scheliga, M. Markovic, H. Rowlands, A. Wozniak, K. Wilde, and J. Butler, "Data provenance tracking and reporting in a high-security digital research environment," *Int. J. Population Data Sci.*, vol. 7, no. 3, 2022, doi: 10.23889/ijpds.v7i3.1909.
- [21] D. Tang, R. Zhao, Y. Lin, T. Zhang, and P. Zhang, "Modeling the data provenance of relational databases supporting full-featured SQL and procedural languages," *Appl. Sci.*, vol. 13, no. 1, p. 64, 2023, doi: 10.3390/app13010064.
- [22] W. Wu and G. Konstantinidis, "Trust and reputation in data sharing: A survey," *arXiv preprint*, arXiv:2508.14028, 2025.