Aplikasi Pohon Merkle dan Fungsi Hash pada Blockchain untuk Menjamin Integritas Data Publik

Arifin Rais (13517067)
Program Studi Teknik Informatika
Sekolah Teknik Elektro dan Informatika
Institut Teknologi Bandung, Jl. Ganesha 10 Bandung 40132, Indonesia
13517067@std.stei.itb.ac.id

***Abstrak*—Data publik pada hakikatnya adalah milik publik, yang proses akumulasinya seharusnya menerapkan prinsip partisipasi publik. Saat ini, proses akumulasi tersebut tersentralisasi pada pihak-pihak tertentu yang seringkali tidak bebas dari bias. Aplikasi pohon merkle dan fungsi hash pada blockchain berpotensi untuk mendesentralisasikan daya akumulasi data publik tersebut kepada masyarakat luas.**

***Kata kunci*—fungi hash kriptografik, pohon merkle, blockchain, data publik.**

I. PENDAHULUAN

1.1 LATAR BELAKANG

Kepemilikan merupakan suatu konsep dalam tatanan filosofis dan hukum yang menjelaskan salah satu bentuk hubungan antara manusia dengan benda. Menurut LeFevre (1966: 2), benda sebagai objek dari keinginan manusia untuk memiliki, jauh lebih tidak penting dari keinginan untuk memiliki itu sendiri. Pernyataan tersebut secara tidak langsung mengimplikasikan bahwa semakin besar keinginan untuk memiliki suatu benda menjadikan benda tersebut lebih penting bagi manusia (tidak selalu terepresentasikan secara adil oleh uang) baik secara individu maupun kolektif. Idealnya, suatu benda dimiliki oleh mereka yang membutuhkannya, tetapi nyatanya, seringkali kepemilikan atas suatu benda tersebar tidak merata sehingga muncul konsep jual-beli, yang melahirkan konsep uang, perdagangan, ekonomi, kesenjangan, dst. Hal ini pula yang terjadi pada data publik.

Data publik—atau informasi publik—dalam UU RI No. 14 Tahun 2008 dijelaskan sebagai informasi yang dihasilkan, disimpan, dikelola, dikirim, dan/atau diterima oleh suatu **badan publik** yang berkaitan dengan penyelenggara dan penyelenggaraan negara dan/atau penyelenggara dan penyelenggaraan badan publik lainnya yang sesuai dengan Undang-Undang ini serta informasi lain yang berkaitan dengan **kepentingan publik**. Kemudian badan publik dijelaskan sebagai lembaga eksekutif, legislatif, yudikatif, dan badan lain yang fungsi dan tugas pokoknya berkaitan dengan penyelenggaraan negara, yang sebagian atau seluruh dananya bersumber dari Anggaran Pendapatan dan Belanja Negara dan/atau Anggaran Pendapatan dan Belanja Daerah, atau organisasi nonpemerintah sepanjang sebagian atau seluruh dananya.

Undang-undang di atas tidak mengakomodasi partisipasi aktif masyarakat umum dalam pengelolaan data publik. Sebagai *stakeholder* data publik paling utama, kondisi ini jelas tidak ideal bagi masyarakat baik dari segi filosofis maupun praksis, karena akumulator data publik tersentralisasi di badan-badan publik tertentu. Banyak contoh kasus yang menunjukkan penyalahgunaan daya akumulasi data publik untuk kepentingan kelompok, seperti lembaga survey sebagai alat politik, AMDAL (analisis dampak lingkungan) proyek pembangunan yang tidak komprehensif, dll.

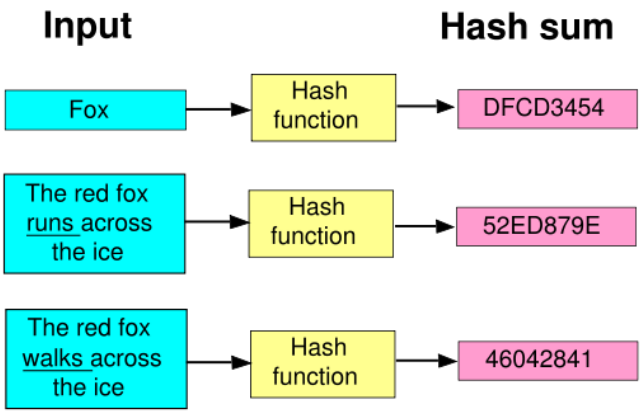
Masalah ini dapat diselesaikan dengan cara mendesentralisasi daya akumulasi data publik kepada gabungan antara badan-badan publik di tingkat yang lebih rendah dan komponen-komponen masyarakat yang kredibilitasnya terjamin, seperti akademisi dan tokoh masyarakat.

Dalam makalah ini, penulis akan membahas potensi penggunaan kombinasi fungsi hash dan pohon merkle yang digunakan pada blockchain untuk menjamin integritas data publik dari segi akumulasi data.

II. DASAR TEORI

2.1 FUNGSI HASH

Fungsi hash $\boldsymbol{H}$ memproyeksikan suatu nilai dari sebuah himpunan dengan dengan sejumlah elemen (bisa tak terbatas) ke suatu nilai pada himpunan yang berjumlah tetap. Fungsi hash sederhana dapat didefinisikan sebagai berikut, $\boldsymbol{H}(x) = \lfloor x \bmod m \rfloor$, $x \in D, m = n(R)$, dimana D adalah domain dan n(R) adalah jumlah elemen pada range.



Gambar 2.1 Diagram input-output dari fungsi hash.
(Sumber: Lecture Notes CSE 330 Data Structures, California State University, San Bernandino)

Kegunaan umum fungsi hash antara lain untuk menentukan apakah dua objek setara dan pengalamatan di memori untuk mencari entri database dengan suatu nilai kunci (pengaksesan menjadi lebih cepat). Contoh lainnya, UNIX c-shell (csh) menggunakan tabel hash untuk menyimpan lokasi program executable. Sehingga untuk menambah program executable baru pada search path pengguna membutuhkan regenerasi tabel hash menggunakan command rehash sebelum program-program tersebut bisa dieksekusi tanpa menspesifikkan path yang komplit.

2.2 FUNGSI HASH KRIPTOGRAFIK

Fungsi hash kriptografik merupakan salah satu jenis fungsi hash yang secara khusus diterapkan untuk enkripsi data. Fungsi hash kriptografik bersifat ireversibel—tidak memiliki invers, dan keseragamannya tinggi—setiap nilai hash memiliki kemungkinan yang sama untuk dihasilkan. Biasanya merupakan salah satu atau kombinasi dari fungsi-fungsi hash berikut:

fungsi hash satu-arah—fungsi searah yang untuk M dan ***H***(M) sulit ditemukan M' ≠ M yang memenuhi ***H***(M') = ***H***(M),

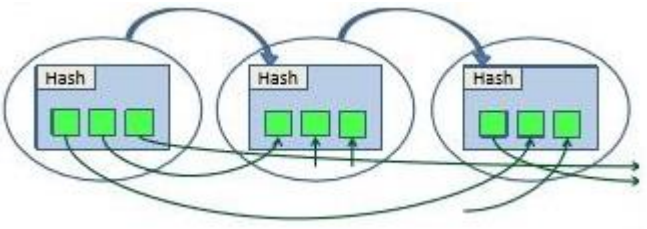
fungsi hash bebas-kolisi—fungsi hash satu-arah yang syarat M' ≠ M-nya tidak bisa ditemukan oleh algoritma manapun dalam waktu polinomial,

fungsi hash pintu-jebakan satu-arah—fungsi hash satu-arah yang memenuhi $f: \{0,1\}^{l(n)} \times \{0,1\}^{(n)} \rightarrow \{0,1\}^{m(n)}$,

atau salah satu jenis dari **fungsi hash universal**.

Pada penerapan blockchain di bitcoin, fungsi hash kriptografik yang digunakan adalah fungsi SHA256.

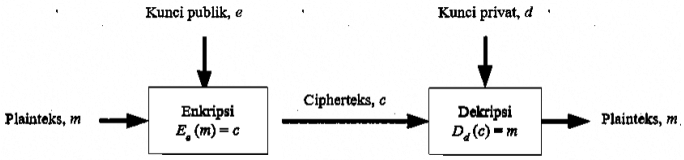
2.3 RANTAI-HASH



Gambar 2.3 Visualisasi rantai-hash.
(Sumber: Mazonka. Oleg. (2016). Blockchain: Simple Explanation.)

Rantai-hash adalah suatu barisan data homogen—atau disebut juga blok—yang dihubungkan oleh suatu fungsi hash. Setiap blok data tersusun oleh nilai hash dan isinya. Nilai hash dari masing-masing blok dihasilkan dari blok sebelumnya secara keseluruhan (nilai hash dan isinya), sehingga apabila terdapat data yang dimanipulasi pada blok manapun akan memengaruhi integritas blok-blok setelahnya (apabila isi blok pertama diubah, maka isi blok kedua harus diubah pula, lalu isi blok ketiga, keempat dst.)

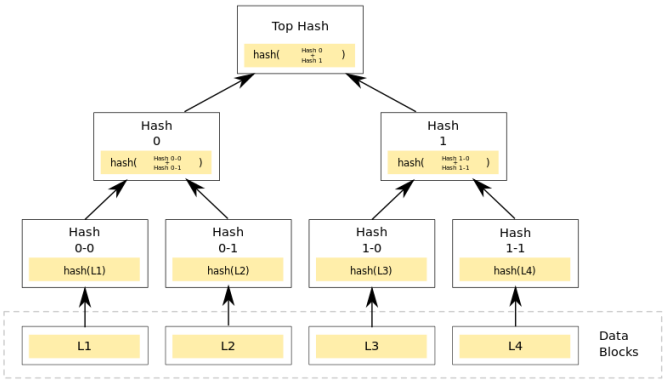
2.4 KRIPTOGRAFI KUNCI PUBLIK



Gambar 2.4 Diagram proses enkripsi-dekripsi pada kriptografi kunci publik.
(Sumber: Rinaldi M/IF2120 Matematika Diskrit, Institut Teknologi Bandung)

Kriptografi kunci publik adalah suatu tipe kriptografi asimetri (kunci enkripsi berbeda dengan kunci dekripsi) dimana kunci enkripsi bisa diperlihatkan tanpa membahayakan isi pesannya. Beberapa contoh metode yang cukup umum digunakan adalah masalah knapsack dan enkripsi RSA.

2.5 POHON MERKLE



Gambar 2.5 Pohon merkle dengan empat blok.
(Sumber: Wikimedia Commons)

Pohon merkle adalah sebuah struktur data berbasis hash, generalisasi dari senarai hash. Struktur yang setiap daunnya merupakan hash dari sebuah blok dan setiap simpul non-daunnya adalah hash dari anak-anaknya. Seringkali merkle tree memiliki faktor percabangan 2, atau setara dengan pohon biner.

Pohon merkle digunakan pada system terdistribusi untuk verifikasi data yang efisien. Efisiensi ini disebabkan oleh penggunaan hash yang jauh lebih cepat untuk diverifikasi dibandingkan isi bloknya secara keseluruhan. Penggunaannya saat ini antara lain, pada jaringan peer-to-peer seperti Tor, Bitcoin, dan Git.

2.6 BLOCKCHAIN

Blockchain adalah suatu mekanisme yang memanfaatkan teknologi-teknologi yang telah dibahas sebelumnya menjadi suatu proses yang secara terus-menerus menghasilkan blok baru dengan penanda digital (dari fungsi hash) dan *timestamp* yang terhubung ke suatu rantai-hash pada suatu sistem terdistribusi. Semakin besar jaringan, semakin terjamin pula integritas data yang disimpan.

Dalam perkembangannya, apabila rantai-hash sudah terlalu panjang, diperlukan peningkatan abstraksi data. Hal ini bisa dilakukan dengan cara mengorganisasikan rantai-hash di dalam rantai-hash lainnya. Rantai-hash eksternal harus berbasis pada

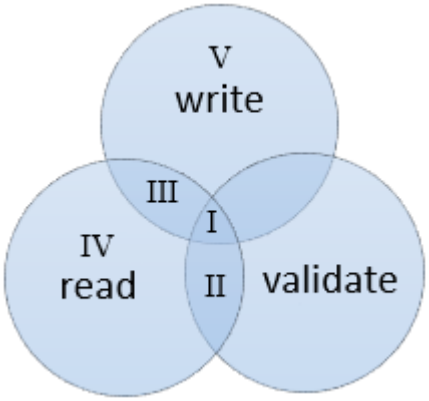
aturan yang seimbang antara kegunaan, kesederhanaan, insentif, kepercayaan, dan otoritas. Rantai-hash internal diharuskan untuk memiliki mekanisme otorisasi sehingga tiap pihak memiliki sesuatu yang bernilai.

III. PENERAPAN BLOCKCHAIN PADA PROSES AKUMULASI DATA PUBLIK

A. Derajat Otoritas, Pemetaan Dimensi Publik, dan Stakeholder Terkait

Pada penerapannya di bitcoin, nyatanya blokchain tidak bisa terlepas dari hierarki sepenuhnya, walaupun minim. Pada bitcoin, pengguna terbagi menjadi dua kelompok: kontributor sekaligus *co-owner* domain bitcoin.org (yang saat ini jumlahnya sudah banyak) dan kontributor biasa.

Untuk penerapannya pada data publik, penulis menawarkan sistem derajat otoritas yang direpresentasikan oleh diagram venn berikut:



Gambar 3.A Diagram otoritas pada sistem blockchain untuk data publik.

Pengguna dengan derajat I memiliki otoritas untuk membaca, menulis dan memvalidasi data, pengguna dengan derajat II memiliki otoritas untuk membaca dan memvalidasi data, pengguna dengan derajat III dapat membaca dan menulis data, pengguna dengan derajat IV hanya dapat membaca data, dan pengguna dengan derajat V hanya dapat menulis data. Setiap derajat otoritas berkorespondensi dengan suatu subhimpunan dari range fungsi hash kriptografik yang bersifat rahasia dan diubah secara berkala. Pengelempokkan ini dimaksudkan untuk menjamin integritas data yang disimpan, dan status derajat diberikan kepada stakeholder tertentu dari suatu dimensi publik.

Dimensi publik yang dimaksud antara lain:

Sosial dan Kependudukan—mencakup data kelahiran, kematian, ketergantungan hidup, persebaran usia, persebaran gender, kesuburan, pertanahan, migrasi, dll. Umumnya stakeholder pada dimensi ini adalah pemerintah dan masyarakat umum, dimana pemerintah memiliki derajat II dan masyarakat memiliki derajat IV-V tergantung topik dengan pertimbangan privasi.

Pembangunan Manusia—mencakup data literasi, partisipasi pendidikan, fasilitas pendidikan, sertifikasi, persebaran sarjana teknik, persebaran sarjana sosial, jenis pekerjaan, budaya daerah, budaya internet, hokum adat, dll. Umumnya stakeholder pada dimensi ini adalah pemerintah, akademisi, komunitas adat, dan masyarakat umum, dimana pemerintah dan akademisi memiliki derajat II, dan komunitas adat serta masyarakat memiliki derajat III.

Ekonomi dan Perdagangan—mencakup data harga eceran bahan-bahan pokok, harga perdagangan besar, harga produsen, industri besar dan menengah, industri mikro, ketenagakerjaan, inflasi, pendapatan dan pengeluaran, sektor-sektor ekonomi, transaksi, kewirausahaan, neraca sosial ekonomi, potensi ekonomi daerah, keuangan, dll. Pada dimensi ini, stakeholder bervariasi sehingga butuh penjabaran khusus. Namun umumnya terbagi menjadi empat dengan derajatnya masing-masing, yaitu pemerintah dengan derajat II, akademisi dengan derajat II, produsen dengan derajat III, konsumen dengan derajat III.

Ilmu Pengetahuan dan Teknologi—mencakup data riset, produk riset, hak paten, penerapan teknologi, masalah masyarakat, dll. Umumnya stakeholder pada dimensi ini adalah pemerintah, akademisi, dan masyarakat umum, dengen derajat I untuk akademisi, derajat II-IV untuk pemerintah tergantung konteks data, derajat IV untuk masyarakat umum.

Lingkungan Hidup—mencakup data satwa liar, tanaman, biodiversitas, kondisi aliran air, hutan, spesies terancam punah, iklim, geografi, geospasial, potensi sumber daya alam, jejak karbon, penggunaan listrik, pencemaran udara, perburuan illegal, penebangan liar, AMDAL, dll. Umumnya pada dimensi ini stakeholder terbagi menjadi empat kelompok utama, yaitu pemerintah, komunitas adat, dan masyarakat umum. Pemerintah memiliki derajat I-II pada topik tertentu, komunitas adat memiliki derajat I, akademisi memiliki derajat I, dan masyarakat umum memiliki derajat III.

Kesehatan—mencakup data *herd immunity*, persebaran penyakit, izin praktek dokter, izin peredaran obat, kesehatan masyarakat, kekurangan gizi, sanitasi lingkungan, kebiasaan hidup, dll. Umumnya pada dimensi ini stakeholder terbagi menjadi empat kelompok utama, yaitu pemerintah, dokter dan pemberi layanan kesehatan lainnya, industri farmasi, dan masyarakat umum. Pemerintah memiliki derajat II, pemberi layanan kesehatan memiliki derajat I, industri farmasi memiliki derajat IV, dan masyarakat umum memiliki derajat III.

Hukum dan HAM—mencakup data pelanggaran hukum dan HAM, dan isu-isu hukum lainnya. Umumnya stakeholder terbagi menjadi tiga: pemerintah, lembaga peradilan, dan masyarakat umum. Pemerintah memiliki derajat II, lembaga peradilan memiliki derajat I dan masyarakat umum memiliki derajat IV-V.

Keamanan, Ketahanan dan Pertahanan—mencakup data potensi bencana baik buatan maupun alam, potensi kejahatan, angka kriminalitas, ancaman eksternal, ancaman internal, batas negara, dll. Umumnya stakeholder utama pada dimensi ini terbagi menjadi tiga kelompok yaitu pemerintah, militer dan kepolisian, serta masyarakat sipil. Pemerintah memiliki derajat I-II tergantung konteks, militer dan kepolisian memiliki derajat I, dan masyarakat sipil memiliki derajat IV-V tergantung konteks pula.

Perlu dicatat bahwa ada protokol yang mengatur penggunaan otoritas untuk masing-masing derajat.

B. Infrastruktur dan Suprastruktur Penunjang.

Infrastruktur yang diperlukan untuk sistem ini antara lain: jaringan sistem blockchain yang terdistribusi merata ke seluruh stakeholder dan suatu platform data publik dengan sistem keamanan siber maksimal yang menjamin anonimitas bagi masyarakat umum, namun masih bisa divalidasi (misalnya dari data kependudukan). Sedangkan untuk suprastruktur diperlukan pembiasaan partisipasi publik melalui penyuluhan yang efektif dan kebermanfaatan data yang bisa dirasakan masyarakat, serta pelatihan-pelatihan yang dikhususkan untuk stakeholder lainya.

C. Protokol Pengelolaan Data Publik

Berdasarkan semua teknologi dan konsep yang dijelaskan sebelumnya, protocol dibentuk untuk memungkinkan system blockchain data public bekerja. Protokol ini diperlukan untuk menjamin integritas jaringan sehingga menjamin integritas data pula. Aturan untuk setiap titik pada system blockchain terdistribusi antara lain:

1. Untuk setiap simpul yang memiliki otoritas untuk menulis data baru, tiap data yang dihasilkan akan disiarkan ke seluruh simpul yang memiliki otoritas untuk memvalidasi pada dimensi publik terkait, **kecuali simpul itu sendiri**.
2. Data yang telah sampai pada simpul validator, akan divalidasi dengan ketentuan yang bergantung pada masing-masing topik data publik terkait, dan apabila memenuhi threshold kriteria tertentu akan diberikan rating berdasarkan tingkat kepercayaannya.
3. Setiap node dengan otoritas membaca kemudian bekerja untuk menemukan *proof-of-work* blok tersebut, analog dari sistem bitcoin.
4. Ketika simpul menemukan *proof-of-work*, hasilnya akan disiarkan ke seluruh node dengan otoritas membaca pada dimensi publik terkait
5. Simpul hanya menerima blok apabila data pada blok tersebut valid dan belum ada pada rantai-bloknya masing-masing.
6. Kemudian apabila blok diterima, blok baru akan dibentuk dengan menggunakan hash blok yang diterima sebagai hash sebelum blok baru.

Rantai terpanjang pada suatu dimensi publik akan dianggap sebagai rantai dengan isi data paling terpercaya, dan tiap simpul akan selalu berusaha memperpanjangnya. Apabila dua simpul menyiarkan data yang berbeda—walaupun kecil kemungkinannya jika menggunakan derajat otoritas—maka beberapa simpul akan mendapatkan satu atau versi lainnya dengan urutan blok berbeda. Pada kasus ini, simpul akan bekerja pada blok pertama yang diterima dan meyimpan versi lainnya pada cabang lain, apabila *proof-of-work* selanjutnya ditemukan maka salah satu cabang akan bertambah panjang, dan versi yang salah akan ditinggalkan.

D. Keamanan, Privasi, dan Transparansi

Potensi ancaman keamanan pada sistem ini terdapat pada platform yang masih berbasis pada satu sumber data yang tersentralisasi, yaitu data kependudukan. Namun ada kemungkinan suatu saat nanti data kependudukan pun bisa diintegrasikan, ke dalam sistem blockchain. Selain itu, validator untuk setiap dimensi publik harus berjumlah banyak untuk menghindari adanya konspirasi masif antar validator untuk mengubah data pada blockchain.

Privasi harus dikedepankan pada data-data yang bersifat sensitif seperti tanggal lahir, nama orang tua, pekerjaan, agama, dll. Hal ini bisa diperoleh apabila data-data tersebut sudah terintegrasi ke dalam system dan protokol menjamin kerahasiaan data-data tersebut.

Transparansi hanya diberikan untuk pengguna dengan otoritas membaca, karena untuk beberapa dimensi publik terdapat data yang sensitif yang tidak bisa disebarkan secara luas.

IV. ANALISIS DAMPAK

Apabila sistem ini diterapkan, potensi disrupsinya sangat tinggi dalam kehidupan bermasyarakat. Pada dimensi sosial dan kependudukan, sengketa lahan dapat dihindari, data kematian jelas sehingga tidak ada isu kecurangan DPT (daftar pemilih tetap), kesenjangan sosial maupun ekonomi lebih mudah dipetakan. Pada dimensi pembangunan manusia, tingkat literasi lebih mudah dipetakan, alokasi tenaga ahli ke daerah lebih mudah, pemerataan pendidikan lebih mudah dipetakan, ketahanan budaya menguat sehingga identitas dan nilai-nilai luhur bangsa terjaga. Pada dimensi ekonomi dan perdagangan potensi ekonomi daerah bisa terakumulasi lebih baik, pemerataan tenaga kerja lebih mudah dilakukan, pihak ketiga untuk transaksi tak langsung bisa ditinggalkan, dll. Pada dimensi IPTEK proses perkembangan ilmu pengetahuan akan terakselerasi dan lebih mudah divalidasi. Pada dimensi kesehatan, persebaran penyakit lebih mudah dipetakan sehingga *outbreak* penyakit dapat dihindari,. Pada dimensi hukum dan HAM, literasi hukum masyarakat akan meningkat dan harapannya masayarakat akan lebih taat hukum, sehingga kasus-kasus sepele bisa dihindari, dan proses hukum pun berjalan transparan. Pada dimensi keamanan, ketahanan dan pertahanan, potensi bencana buatan atau alam dapat tersosialisasikan dengan baik harapannya masyarakat akan mengalami perubahan pola pikir ke arah pencegahan dan mitigasi bencana.

Dari seluruh hal di atas, hal yang paling penting adalah adanya partisipasi publik dalam proses akumulasi data membantu menjamin keberpihakan data tersebut kepada publik itu sendiri.

V. SIMPULAN

Aplikasi pohon merkle dan fungsi hash pada blockchain memungkinkan terjaminnya integritas data. Sejauh ini blockchain baru diterapkan secara umum di bidang cryptocurrency, padahal potensi penerapan dan kebermanfaatannya sangat luas. Salah satu bentuk pemanfaatan teknologi ini adalah pada proses akumulasi data publik seperti yang dibahas pada makalah ini.

VI. UCAPAN TERIMAKASIH

Alhamdulillahirabbilalamin. Terimakasih saya ucapkan

kepada dosen pengampu saya dalam mata kuliah IF2120 Matematika Diskrit ini, Pak Rinaldi Munir yang selalu memberi nilai tambah pada setiap kuliahnya. Kalau saya boleh menggunakan terminologi dari sebuah film berjudul *3 idiots*, Pak Rinaldi adalah sosok edukator yang baik, bukan trainer. Selain itu saya ucapkan terimakasih kepada Bang Radja yang selalu melontarkan wacana-wacana yang segar dan patut direnungi, Kang Yorga yang selalu mengingatkan pentingnya mencintai ilmu pengetahuan, dan teman saya Mara yang selalu mengingatkan bahwa puncak dari ilmu adalah amal.

REFERENSI

[1] LeFevre, Robert. 1966. *The Philosophy of Ownership*. Auburn: The Ludwig von Mises Institute.
[2] Republik Indonesia. 2008. *Undang-Undang Keterbukaan Informasi Publik.* Jakarta*:* Sekretariat Negara.
[3] Skalaban, Andrew. 1988. *Do the Polls Affect Elections? Some 1980 Evidence.* Political Behavior Vol. 10, No. 2, hlm.136-150.
[4] Partow, A. *General Purpose Hash Function Algorithms*. http://www.partow.net/programming/hashfunctions/. (diakses 9 Desember 2018).
[5] Bakhtiari, S.; Safavi-Naini, R.; and Pieprzyk, J. *Cryptographic Hash Functions: A Survey*. Technical Report 95-09, Department of Computer Science, University of Wollongong, Juli 1995.
[6] Mazonka, Oleg. 2017. *Blockchain: Simple Explanation*.
[7] Zaghal, A. 2016. *Merkle Tree*. https://brilliant.org/wiki/merkle-tree/, (diakses 9 Desember 2018).

PERNYATAAN

Dengan ini saya menyatakan bahwa makalah yang saya tulis ini adalah tulIpin saya sendiri, bukan saduran, atau terjemahan dari makalah orang lain, dan bukan plagiasi.

Bandung, 9 Desember 2018

Arifin Rais
13517067