



UNIVERSITÀ  
DEGLI STUDI  
DI PADOVA



DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE

CORSO DI LAUREA IN INGEGNERIA INFORMATICA

# **Blockchain e Learning Analytics: Una Nuova Frontiera per la Gestione dei Dati**

**Relatore**

Prof. Varagnolo Damiano

**Laureando**

Speranza Ludovico

**Correlatore**

Cognome Nome

ANNO ACCADEMICO 2024-2025

Data di laurea GG/MM/AAAA



*Apes. Together. Strong.*



# Abstract

La Blockchain è una tecnologia spesso definita "l'Internet del futuro", poiché rappresenta un vero e proprio stravolgimento infrastrutturale con potenziali ripercussioni in numerosi settori. La sua componente innovativa risiede nella possibilità di sviluppare applicazioni decentralizzate e sicure, che non richiedono la presenza di un intermediario. Grazie a queste caratteristiche, la Blockchain è destinata a trasformare molti aspetti della società moderna e si trova oggi al centro di discussioni tecnologiche e monetarie.

Questa tesi esplora l'evoluzione della Blockchain, analizzando inizialmente l'aspetto per cui è nata, ovvero la gestione delle criptovalute, soffermandosi in particolare su Bitcoin ed Ethereum e sui loro algoritmi di validazione. Successivamente, il lavoro si concentra sul tema della protezione dei dati personali, partendo dal Web 3.0 per arrivare ad affrontare le problematiche del Learning Analytics. Viene fornita un'analisi critica del sito FaceItTools, proponendo una sua possibile trasformazione blockchain-oriented al fine di migliorare la gestione dei dati, garantendo maggiore trasparenza e affidabilità. In aggiunta, vengono analizzati Blockcerts, EduCTX e Sony Global Education come esempi concreti di piattaforme basate sulla blockchain.

La ricerca proposta si pone l'obiettivo di offrire una visione completa e critica delle opportunità offerte dalla Blockchain, mettendo in evidenza, in particolare, le potenzialità che essa può avere nel trattamento dei dati sensibili degli studenti in ambito scolastico.



# Indice

<b>1</b>	<b>Blockchain</b>	<b>1</b>
1.1	Blokchain 1.0 . . . . .	4
1.1.1	Bitcoin . . . . .	4
1.1.2	Proof Of Work . . . . .	6
1.2	Blokchain 2.0 . . . . .	8
1.2.1	Ethereum e Proof Of Stake . . . . .	8
1.2.2	Smart Contracts . . . . .	8
1.2.3	Vantaggi e svantaggi del PoS . . . . .	9
<b>2</b>	<b>Sicurezza</b>	<b>11</b>
2.1	Privacy e sicurezza . . . . .	11
2.1.1	Il difficile rapporto tra trasparenza e privacy . . . . .	11
2.1.2	Blockchain Permissioned e Permissionless . . . . .	13
2.1.3	Crittografia End-to-End e responsabilità degli algoritmi . . . . .	13
2.2	Web 3.0 . . . . .	14
<b>3</b>	<b>Learning Analytics</b>	<b>17</b>
3.1	Limiti del Learning Analytics . . . . .	18
3.1.1	Collegare le storie di apprendimento . . . . .	18
3.1.2	Privacy, sicurezza e controllo degli accessi . . . . .	19
3.1.3	Integrare i sistemi di ricerca e produzione . . . . .	19
3.2	Blockchain come soluzione . . . . .	19
3.3	EduCTX . . . . .	21
3.3.1	Struttura . . . . .	21
3.3.2	Come avviene la registrazione degli studenti . . . . .	21
3.4	Sony Global Education . . . . .	21
3.5	Blockcherts . . . . .	21
	<b>Bibliografia</b>	<b>23</b>





# Elenco delle figure

1.1	Statua raffigurante Satoshi Nakamoto . . . . .	1
1.2	Esempio di catena . . . . .	3
1.3	Esempio di cifratura asimetrica . . . . .	5
1.4	Esempio di funzionamento del PoW . . . . .	6
2.1	Livelli del web 3.0 . . . . .	15
3.1	Dragan Gasevic . . . . .	17
3.2	EduCTX homepage . . . . .	21



# Capitolo 1

## Blockchain

Nonostante il grande successo della trasformazione digitale, molti ambiti delle nostre interazioni sono rimasti indietro e non si sono adattati al digitale; un esempio su tutti è quello del settore legale. Inoltre, la digitalizzazione presenta anche lati oscuri [1]: i nostri dati sono controllati da organizzazioni il cui obiettivo principale è il profitto, gli hacker tentano continuamente di sottrarre informazioni personali, e gli Stati monitorano le attività dei cittadini per mantenere l'ordine e il controllo. Tutto ha origine nel 2008, con la pubblicazione online dell'articolo "Bitcoin: A Peer-to-Peer Electronic Cash System" da parte di Satoshi Nakamoto, pseudonimo che potrebbe celare un individuo o un gruppo di persone. L'obiettivo del documento era quello di creare un sistema puramente peer-to-peer per il trasferimento di valore tra le parti, eliminando la necessità di intermediari come banche o istituti finanziari. Il concetto chiave della blockchain è la possibilità di creare una generazione di piattaforme decentralizzate e disintermedate, nelle quali la fiducia tra le parti non è garantita da un'organizzazione centrale, ma dalla piattaforma stessa, attraverso un meccanismo algoritmico essa si riferisce, in senso più ampio, all'insieme di tecnologie che rendono possibile la decentralizzazione, basandosi sul modello peer-to-peer, in sostanza consente di immaginare un mondo organizzato in modo completamente diverso. La blockchain è un database completamente decentralizzato, di cui ogni nodo possiede una replica e concordano, sulla base di un algoritmo comune, un'unica versione aggiornata delle informazioni. Si tratta di una struttura dati nella quale è possibile solo aggiungere nuovi dati, senza possibilità di cancellare quelli precedenti, garantendo così uno storico completo delle modifiche. Per chiarire meglio questo concetto, possiamo immaginare la blockchain come una barriera corallina: lo strato attivo corrisponde alla parte più recente, che viene continuamente modifi-



Figura 1.1: Statua raffigurante Satoshi Nakamoto

cata con nuove aggiunte, mentre gli strati più vecchi rimangono immutabili e accessibili solo per attività di consultazione. Può essere quindi rappresentata come una lista in continua crescita di “blocchi” collegati tra loro e protetti mediante crittografia. Oltre alla decentralizzazione e all’immutabilità, il terzo grande punto di forza è il consenso: nuove transazioni possono essere registrate solo quando la maggioranza dei partecipanti alla rete dà il proprio consenso.

I blocchi sopra citati sono composti da 7 campi principali:

*Block*: rappresenta il numero del blocco all’interno della blockchain. Ogni volta che un nuovo blocco viene aggiunto, questo numero aumenta progressivamente.

*Nonce*: è un valore, spesso composto da 32 bit (come nel caso di Bitcoin), fondamentale perché viene cercato dai miner per generare un hash valido che soddisfi i requisiti di difficoltà della chain. Serve a modificare l’input del calcolo dell’hash e viene aggiornato continuamente durante il processo di mining.

*Timestamp*: indica il momento esatto in cui il blocco è stato aggiunto alla blockchain.

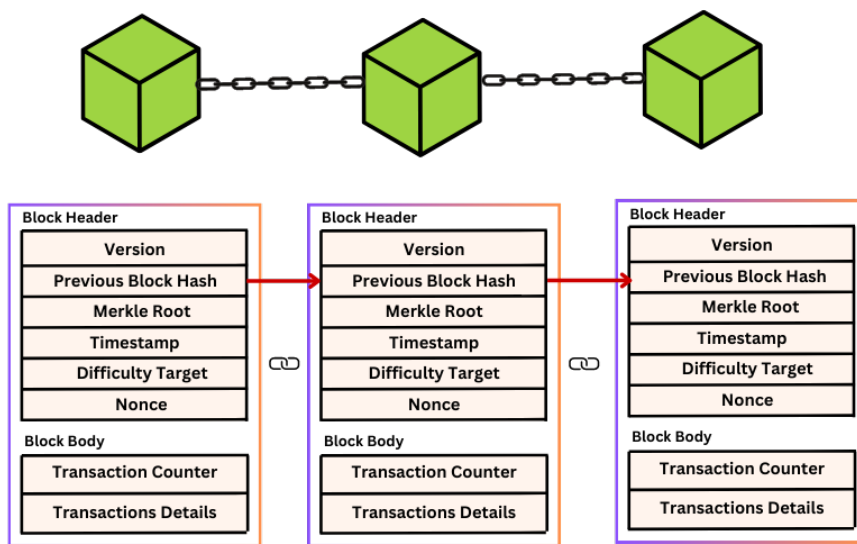
*Transaction*: contiene o un valore (ad esempio un importo monetario), o il corpo del messaggio, o uno smart contract.

*Transaction n*: si riferisce a una specifica transazione tra le diverse transazioni presenti all’interno del blocco.

*Prev Hash*: è l’hash del blocco precedente, che funge da chiave di collegamento tra i blocchi. Questo meccanismo garantisce che ogni blocco sia dipendente dal precedente, creando una catena ininterrotta.

*Hash*: rappresenta l’impronta digitale del blocco. È un valore alfanumerico unico calcolato applicando una funzione hash. L’hash è deterministico, nel senso che lo stesso input produrrà sempre lo stesso hash; tuttavia, non è possibile risalire ai dati originali partendo dall’hash. Ogni minima modifica ai dati genera un hash completamente diverso, garantendo così l’integrità e l’immutabilità delle informazioni.

Figura 1.2: Esempio di catena



## 1.1 Blockchain 1.0

### 1.1.1 Bitcoin

L'aspetto rivoluzionario di Bitcoin è la creazione di una forma monetaria completamente decentralizzata. La sua rete è composta da una serie di componenti chiamate nodi, che collaborano tra loro per raggiungere il consenso su una sequenza di transazioni. Essendo basato sulla blockchain, Bitcoin è completamente decentralizzato: non è gestito da un'entità centrale, come avviene per esempio con le banche. Questo comporta un vantaggio significativo, ovvero una maggiore resistenza a comportamenti anomali o malintenzionati. In un sistema centralizzato, infatti, il nodo centrale rappresenta un punto di vulnerabilità che potrebbe compromettere l'intero sistema. Un'altra caratteristica fondamentale di Bitcoin è la disintermediazione: ogni transazione coinvolge direttamente il mittente e il destinatario, che raggiungono il consenso tramite l'approvazione della rete. Bitcoin è quindi costituito da una rete di sistemi indipendenti, i nodi, capaci di mantenere un registro univoco e condiviso delle transazioni. Questo sistema garantisce che un singolo bitcoin non possa essere speso più di una volta.

In tal modo, Bitcoin soddisfa le funzioni critiche di ogni valuta:

- Immagazzinare valore.
- Essere un'entità contabile.
- Trasferire valore tra le parti.

La sua blockchain è di tipo permissionless, il che significa che chiunque può eseguire il software del nodo e collegarsi alla rete senza alcuna forma di autenticazione. In altre parole, chiunque può unirsi alla rete senza necessità di approvazioni o permessi da parte di un'autorità centrale. Un elemento fondamentale è il wallet che consente agli utenti di gestire i propri bitcoin, conservati in veri e propri conti. Ogni bitcoin è associato a un address, protetto da una coppia di chiavi crittografiche: La chiave pubblica rappresenta l'indirizzo del conto ed è accessibile a tutti. La chiave privata rappresenta l'identità del proprietario ed è necessaria per accedere e autorizzare le transazioni. Bitcoin utilizza la crittografia asimmetrica, un sistema inventato nel 1976 da Whitfield Diffie e Martin Hellman. Questo metodo permette di cifrare un messaggio con una chiave (pubblica) che può essere decifrato solo con l'altra chiave (privata), e viceversa, ciò determina due vantaggi principali. Il primo è Hashing dei dati che trasforma qualsiasi quantità di dati in una stringa cifrata di dimensioni fisse, chiamata digest.

L'hashing è:

*Deterministico*: ovvero lo stesso input produce sempre lo stesso digest.

*Lunghezza fissa*: il digest ha una lunghezza predeterminata, indipendentemente dalla dimensione dell'input.

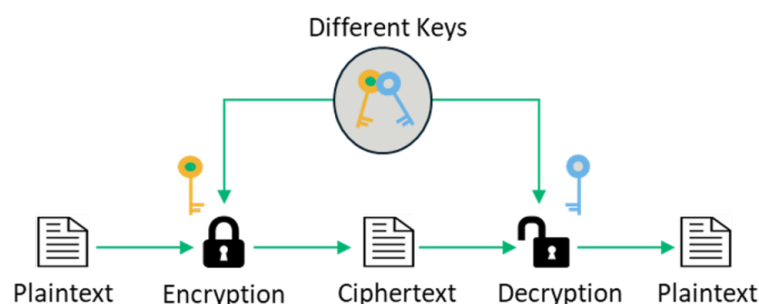
*Unico*: è estremamente improbabile che due input diversi generino lo stesso digest.

*Non invertibile*: è impossibile risalire all'input originale partendo dal digest.

*Instabile*: una minima modifica all'input genera un digest completamente diverso.

Il secondo grosso vantaggio è la firma digitale che garantisce unicità e non ripudiabilità. Quando un utente X invia a un utente Y un messaggio firmato digitalmente, X crea l'hash del messaggio e lo cifra con la propria chiave privata. Y, ricevendo il messaggio, può verificare l'hash con la chiave pubblica di X, garantendo così l'integrità del messaggio e la validità del mittente.

Figura 1.3: Esempio di cifratura asimmetrica



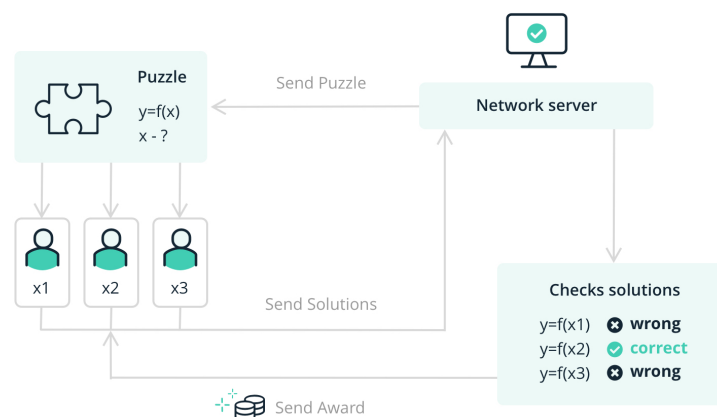
Bitcoin è open source, il suo codice sorgente è pubblico e rilasciato sotto licenza MIT, che permette la massima libertà di utilizzo per progetti futuri. Nonostante il grande successo riscosso, Bitcoin presenta diverse criticità che ne limitano l'efficacia e l'adozione su larga scala. Una delle principali problematiche riguarda la scalabilità: l'algoritmo Proof of Work (PoW) su cui si basa il sistema, pur garantendo sicurezza, tende a limitare l'efficienza complessiva, rendendo difficile gestire un alto numero di transazioni in tempi brevi. Un altro aspetto delicato è quello relativo a privacy e visibilità. Sebbene gli account su Bitcoin siano pseudonimi, tutte le transazioni effettuate sulla rete sono pubblicamente visibili. Questo permette, tramite un'analisi approfondita dei movimenti della valuta, di risalire potenzialmente all'identità dei possessori, mettendo in discussione la reale riservatezza del sistema. La difficoltà di aggiornamento rappresenta un'ulteriore sfida. Il complesso sistema di governance, che coinvolge minatori, sviluppatori e utenti, rende particolarmente arduo raggiungere un consenso per apportare modifiche al protocollo.

Ciò può rallentare l'implementazione di miglioramenti necessari. Infine, vi è il problema delle commissioni elevate, che si acuisce nei periodi di intenso utilizzo della rete. In queste circostanze, effettuare transazioni rapide può diventare molto costoso, specialmente se il tasso di cambio tra Bitcoin ed Euro è particolarmente alto.

### 1.1.2 Proof Of Work

Come accennato in precedenza, la sicurezza e la validità delle transazioni in Bitcoin sono garantite da un algoritmo di consenso chiamato Proof of Work (PoW) [2]. Un algoritmo di consenso è un protocollo condiviso da tutti i nodi della rete, che consente loro di concordare su una visione unica delle transazioni. In primo luogo, ogni nodo condivide con gli altri nodi della rete le transazioni ricevute nell'ultimo intervallo di tempo, firmate digitalmente. Successivamente, ciascun nodo costruisce in memoria un elenco delle transazioni ricevute dagli altri nodi con cui è in comunicazione. Queste transazioni vengono ordinate, e viene verificato che ogni account disponga dei fondi necessari per completare i pagamenti. A questo punto, i nodi partecipano a una competizione crittografica per individuare, in modo casuale, il nodo che avrà il diritto di pubblicare il blocco successivo. Quando un nodo vince la competizione, individua un blocco valido e lo comunica a tutti gli altri nodi. Questi ultimi verificano il blocco e lo aggiungono alla propria copia della blockchain.

Figura 1.4: Esempio di funzionamento del PoW



Nell'ambito blockchain si sente spesso parlare di fork [3], ma non è sempre chiaro cosa si intenda con fork e cosa effettivamente comporti. Sebbene il termine venga spesso utilizzato per indicare la divisione di una blockchain in realtà esso racchiude un insieme di diversi possibili scenari. Una fork è una situazione in cui accade una delle seguenti cose:



- Può accadere che due o più nodi trovino contemporaneamente una soluzione valida. In tal caso, si generano due blocchi con lo stesso genitore, creando una biforcazione nella catena, detta fork. Quando viene individuato il blocco successivo, l'algoritmo seleziona la catena più lunga (in termini di difficoltà complessiva) come valida. Questo caso viene denominato come fork regolare.
- Le regole della Blockchain sono cambiate in maniera retrocompatibile e tutti i nodi condividono la stessa cronologia delle transazioni, in questo caso si tratta di soft fork e non c'è una divisione della blockchain.
- Le regole della Blockchain sono cambiate in maniera non retrocompatibile, ma tutti i nodi si aggiornano alle nuove regole e condividono la stessa cronologia delle transazioni, anche in questo caso non si ha una divisione della blockchain, ma si tratta di un hard fork.
- Le regole della blockchain sono cambiate in maniera non retrocompatibile e nodi diversi hanno opinioni diverse sulle regole della blockchain, non condividendo la stessa cronologia delle transazioni, abbiamo quindi un hard fork con chain split, perché avviene una divisione della chain e il consenso sulla cronologia delle transazioni è perso definitivamente.

Questo criterio, basato sul "principio di maggior lavoro", permette a tutti i nodi di concordare su una versione univoca della blockchain e di risolvere le biforcazioni. Un blocco viene considerato finalizzato quando sono stati generati almeno sei blocchi successivi. Questo livello di profondità rende economicamente insostenibile, per una minoranza di nodi malintenzionati, modificare i dati della blockchain, poiché richiederebbe un effort computazionale estremamente elevato pari al 51% della potenza totale di calcolo del network e anche se ipoteticamente un miner riuscisse a raggiungere tale potenza non sarebbe comunque in grado di modificare le vecchie transazioni, poiché dovrebbe ricalcolare la PoW di tutti i blocchi successivi, mentre gli altri miner onesti continuano a minare sulla blockchain corretta. Un attacco di questo tipo richiederebbe l'utilizzo di una quantità incredibile di risorse per l'attacker. Se qualcuno effettivamente riuscisse a mettere insieme più del 51% della potenza di calcolo, sarebbe molto più redditizio per lui seguire le regole della blockchain. Nonostante i numerosi vantaggi che il Proof of Work offre, questo algoritmo di consenso presenta alcune significative criticità. La principale è il massiccio consumo di energia, un aspetto che, paradossalmente, costituisce anche uno dei suoi punti di forza, rendendo estremamente costoso e complesso attaccare la rete. A questo si aggiunge la scarsa scalabilità del sistema, che si traduce in una certa lentezza nell'elaborazione delle transazioni e nell'aumento delle commissioni, soprattutto nei periodi di maggiore attività sulla rete. Infine, il Proof of Work tende a creare una sorta di discriminazione geografica: attualmente, la mag-

gior parte dei miner si concentra in aree dove il costo dell'elettricità è più basso, limitando la partecipazione globale e accentuando disparità economiche e infrastrutturali [3].

## **1.2 Blockchain 2.0**

### **1.2.1 Ethereum e Proof Of Stake**

Bitcoin, per sua natura, non è stato concepito come un ambiente di sviluppo e offre funzionalità di programmazione molto limitate, mentre la community ha sempre voluto mantenere la piattaforma focalizzata sullo scambio di valori, lasciando spazio a progetti alternativi che miravano ad ampliare le possibilità della tecnologia blockchain. È in questo scenario che nasce Ethereum, la prima blockchain progettata specificamente per supportare lo sviluppo di applicazioni decentralizzate. Co-fondata nel 2013 da Vitalik Buterin, un programmatore russo-canadese che in passato aveva collaborato con Bitcoin Magazine, Ethereum si distingue da Bitcoin per molteplici innovazioni. Una delle principali novità introdotte è la possibilità di utilizzare due tipi di conti: i tradizionali conti posseduti tramite una coppia di chiavi pubblica e privata, come in Bitcoin, e i conti associati agli Smart Contract [2].

### **1.2.2 Smart Contracts**

Gli Smart Contract, concettualizzati per la prima volta da Nick Szabo nel 1994, sono programmi in grado di eseguire automaticamente azioni predeterminate una volta soddisfatte certe condizioni. Questi contratti decentralizzati garantiscono che le regole siano rispettate senza la necessità di intermediari, eliminando possibili interferenze. Un esempio pratico di utilizzo degli Smart Contract è il crowdfunding. Un utente potrebbe pubblicare un progetto stabilendo un obiettivo economico e un tempo limite per raggiungerlo. Lo Smart Contract, in questo caso, si occuperebbe di raccogliere i fondi dagli investitori e trattenerli fino al completamento della campagna. Se l'obiettivo viene raggiunto, i fondi vengono trasferiti automaticamente al creatore del progetto; in caso contrario, tornano ai donatori. Questa struttura elimina la necessità di intermediari e rende il processo completamente trasparente. Ethereum utilizza Solidity, un linguaggio di programmazione simile a JavaScript, per scrivere Smart Contract, la flessibilità offerta da questa piattaforma ha rivoluzionato il mondo delle blockchain, permettendo la creazione di applicazioni decentralizzate e inaugurando l'era delle cosiddette blockchain 2.0.

### 1.2.3 Vantaggi e svantaggi del PoS

Inizialmente veniva utilizzato il meccanismo di consenso Proof of Work (PoW), lo stesso di Bitcoin, però per migliorare efficienza e sostenibilità, il 15 settembre 2022, con l'aggiornamento noto come The Merge, Ethereum è passato al Proof of Stake (PoS). Questo cambiamento ha segnato un cambio epocale per la piattaforma ed ha portato con sé enormi vantaggi [4]. Il PoS riduce drasticamente il consumo energetico, abbassandolo del 99,9% rispetto al PoW, ed elimina la necessità di hardware costoso per il mining, rendendo la partecipazione al consenso più accessibile. Con l'introduzione del PoS, Ethereum ha reso possibile lo staking, un processo in cui i partecipanti bloccano una quantità di criptovaluta in uno Smart Contract per sostenere la sicurezza e il funzionamento della blockchain. Attraverso lo staking, gli utenti possono diventare validatori e partecipare attivamente alla convalida delle transazioni e alla creazione di nuovi blocchi. Lo staking offre numerosi vantaggi: aumenta la sicurezza del sistema rendendo gli attacchi costosi e difficili da eseguire, favorisce la decentralizzazione permettendo a più utenti di contribuire alla rete e riduce il consumo energetico rispetto al mining tradizionale. I partecipanti allo staking ricevono ricompense proporzionali alla loro attività, sotto forma di nuove criptovalute o una parte delle commissioni generate dalle transazioni. Le ricompense non si limitano ai validatori che propongono nuovi blocchi, ma includono anche coloro che verificano e confermano la validità dei blocchi proposti da altri. Questo sistema rende lo staking una delle principali motivazioni per partecipare al consenso su Ethereum, favorendo un equilibrio tra efficienza e incentivi economici. Inoltre, il sistema garantisce una maggiore scalabilità e permette ai validatori onesti di proteggere la rete da attacchi, penalizzando economicamente i nodi malevoli o inattivi per almeno il 50% del tempo attraverso un meccanismo chiamato slashing che risulterà totale per i primi andando a sottrarre in maniera permanente le risorse depositate nello stake, mentre sarà parziale per i secondi che si vedranno sottratti solo i guadagni ricevuti ma non saranno rimossi dalla chain. Nonostante questi vantaggi, il PoS presenta alcune criticità. Essendo una tecnologia più recente rispetto al PoW, è meno testata e potrebbe essere più vulnerabile a eventi imprevisti, i cosiddetti Black Swan. Inoltre, il fatto che il controllo della rete sia legato al possesso di criptovalute pone un rischio di centralizzazione: chi dispone di maggiori capitali potrebbe accumulare una quantità significativa di token, aumentando la propria influenza sul sistema [5].



# Capitolo 2

## Sicurezza

### 2.1 Privacy e sicurezza

L'adozione delle tecnologie basate sulla blockchain, si è progressivamente ampliata oltre il settore delle criptovalute, trovando applicazione in ambiti come quello della protezione della privacy [6]. La spinta verso la decentralizzazione deriva principalmente dalle crescenti preoccupazioni degli utenti riguardo alla perdita di controllo sui propri dati personali archiviati online [7] [8]. A tal riguardo, la struttura stessa della blockchain offre potenzialità per tutelare la riservatezza delle informazioni personali; tuttavia, l'analisi dei metadati può renderla vulnerabile in determinati contesti. Di conseguenza, senza un'adeguata progettazione, *«decentralized infrastructures intended to promote individual privacy and autonomy might turn out to be much more vulnerable to governmental or corporate surveillance than their centralized counterparts»* [9]. In questa prospettiva, non va dimenticato che la natura pseudonima di molte reti che si basano sulla blockchain consente agli individui la possibilità di condurre le proprie transazioni su base peer-to-peer, senza la necessità di rivelare la propria identità alle controparti.

#### 2.1.1 Il difficile rapporto tra trasparenza e privacy

Allo stesso tempo, la trasparenza derivante dalle distributed ledger technologies è tale che chiunque ha la possibilità di accedere alla cronologia di tutte le transazioni memorizzate sulla blockchain, affidandosi così all'analisi dei dati in essa contenuti per ricavare informazioni potenzialmente sensibili [10]. Se il sistema non viene progettato con la dovuta attenzione, la trasparenza potrebbe compromettere la tutela della privacy degli utenti, in quanto questi avrebbero accesso ai dati di tutti gli altri utenti presenti sulla piattaforma e quindi si avrebbe un miglioramento della trasparenza a discapito però della privacy [11] [12].

Pertanto, in assenza di soluzioni tecniche adeguate per salvaguardare la riservatezza delle co-

municazioni online, le infrastrutture decentralizzate, pensate per favorire privacy e autonomia, rischiano di risultare più esposte al controllo di governi o aziende rispetto ai sistemi centralizzati. Resta comunque una notevole incertezza riguardo alla capacità di soluzioni alternative e decentralizzate di affrontare efficacemente le problematiche legate alla protezione dei dati. Nonostante ciò, i sistemi decentralizzati, come la blockchain, hanno suscitato un crescente interesse nella comunità accademica, che sta approfondendo il potenziale utilizzo delle DLTs in ambito privacy.

Per comprendere meglio le questioni legate alla protezione dei dati personali, è fondamentale esaminare due aspetti principali:

- l'individuazione del soggetto responsabile della definizione delle modalità di trattamento dei dati personali.
- l'identificazione di chi controlla la conservazione e gestione dei dati.

La natura decentralizzata della blockchain non solo garantisce livelli più elevati di protezione dei dati personali, ma conferisce agli utenti un maggiore controllo sulle proprie informazioni, permettendo loro di gestirle autonomamente durante gli scambi.

Questa tecnologia è stata infatti descritta dalla dottrina come un vero e proprio «*sistema di cloud computing decentralizzato*» [13].

In tale contesto, il principale vantaggio offerto dai sistemi decentralizzati risiede nella possibilità per gli utenti di assumere un controllo diretto sulla gestione dei propri dati, inoltre, le informazioni generate, condivise e raccolte dagli utenti stessi potrebbero essere messe a disposizione o vendute per fini comuni, consentendo l'accesso anche a terze parti. Questo approccio ricorda quello degli open data, ma con meccanismi e logiche significativamente diversi.

Per quanto riguarda la struttura, gran parte delle architetture decentralizzate disponibili per gli utenti sono progettate con l'obiettivo di favorire la privacy, focalizzandosi su almeno uno dei due paradigmi fondamentali: la riservatezza dei dati e la "sovranità" su di essi.

In questa prospettiva, la decentralizzazione che caratterizza la blockchain offre un potenziale significativo per ridurre le asimmetrie informative che si verificano quando una delle parti coinvolte in una transazione possiede informazioni rilevanti che l'altra parte non conosce, le quali, nei sistemi centralizzati, tendono a creare vantaggi per gli operatori a discapito degli utenti.

Tuttavia, va evidenziato che, allo stato attuale, persistono numerose complessità nel comprendere appieno come privacy e trasparenza possano interagire tra loro. In una società "trasparente", qualsiasi parte interessata può facilmente accedere alle informazioni, questo implica che la trasparenza sociale possa compromettere il diritto alla privacy. È legittimo quindi associare una crescente trasparenza delle informazioni a un ridotto rispetto del diritto alla privacy. In tale contesto, la trasparenza offerta dalle DLTs non è né totale né incondizionata.

### **2.1.2 Blockchain Permissioned e Permissionless**

Le diverse tipologie di blockchain, infatti, possono garantire vari livelli di trasparenza. Esistono, in particolare, due tipologie di blockchain:

*permissioned*: in cui una sola autorità ha il permesso di scrivere sulla blockchain e detiene praticamente il controllo del sistema distribuito.

*permissionless*: in cui chiunque può partecipare come nodo, verificare transazioni e aggiungere blocchi al registro e tutti i dati delle transazioni sono visibili e accessibili a chiunque abbia accesso alla rete, anche se non tutti possono necessariamente scrivere nella blockchain (ma chiunque può leggere i dati).

Soprattutto nel caso delle blockchain permissioned, le transazioni (o scambi, scrittura delle informazioni) avvengono all'interno di un ecosistema chiuso, dove i dati registrati sono mantenuti riservati e le identità dei partecipanti sono note [3].

### **2.1.3 Crittografia End-to-End e responsabilità degli algoritmi**

Da un punto di vista pratico, i problemi di privacy legati al livello di trasparenza della blockchain possono essere attenuati mediante la crittografia “end-to-end” delle comunicazioni che utilizza chiavi private e pubbliche, anziché una chiave unica per crittografare e decrittografare. Il principio implica che, ove possibile, le operazioni del protocollo di comunicazione debbano essere eseguite ai punti finali di un sistema di comunicazione, il più vicino possibile alla fonte o al destinatario finale dei dati, o comunque alla risorsa da controllare. In particolare, le aspettative sociali riguardo la trasparenza e la supervisione degli algoritmi sono in forte crescita, con l'obiettivo di rendere i sistemi decisionali automatizzati più responsabili, trasparenti e governabili. Si sta cercando di dotare questi sistemi di nuovi strumenti tecnologici per verificare che le decisioni automatizzate rispettino standard di equità giuridica. Garantire la responsabilità attraverso valutazioni d'impatto degli algoritmi (AIA) [14], audit e certificazioni dovrebbe diventare parte integrante delle iniziative politiche e legali in questo ambito, considerando che la blockchain, al momento, non è stata ancora completamente implementata [15].

## 2.2 Web 3.0

L'idea del Web 3.0 è stata inizialmente utilizzata in stretta connessione con il concetto di web semantico, terminologia creata da Berners Lee in un articolo di *Scientific American* del 2001 per descrivere un nuovo web. La sua idea è che, così come il web 2.0 permette di collegare pagine web, a livello di visualizzazione, il web semantico deve permettere non solo di collegare pagine tra loro ma anche i dati contenuti in esse [16]. La visione del web semantico è una visione di dati interconnessi e navigabili che possono essere usati da chiunque. L'esigenza di socializzazione dei dati è ancora più importante oggi che l'intelligenza artificiale può costruire modelli a partire dai dati grezzi, sulla base di algoritmi generali.

L'obiettivo del web 3.0, riconosciuto da Gavin Wood, co fondatore di Ethereum, è quello di creare un web decentralizzato, dove i dati sono posseduti dagli utenti e non da poche aziende. Sta nascendo la consapevolezza di re-decentralizzare i servizi web [3]. Questa idea è diffusa da molto tempo, ma adesso finalmente esistono le tecnologie che permettono di realizzarla, come la blockchain e così si potrebbero ottenere numerosi vantaggi, tra i quali:

*centralizzazione:* Non è necessario alcun permesso da parte di un'autorità centrale per caricare qualcosa sul web. Questo fornisce una protezione contro qualsiasi forma di censura e controllo. Il web tornerebbe ad essere un sistema neutrale.

*democratizzazione:* E' possibile offrire un accesso a chiunque abbia una connessione a internet, senza discriminazioni su età, sesso, razza, religione e posizione geografica.

*uptime dei servizi:* Non essendoci nodi centrali, non esiste un punto di fallimento. Se un nodo va giù, il servizio è comunque disponibile.

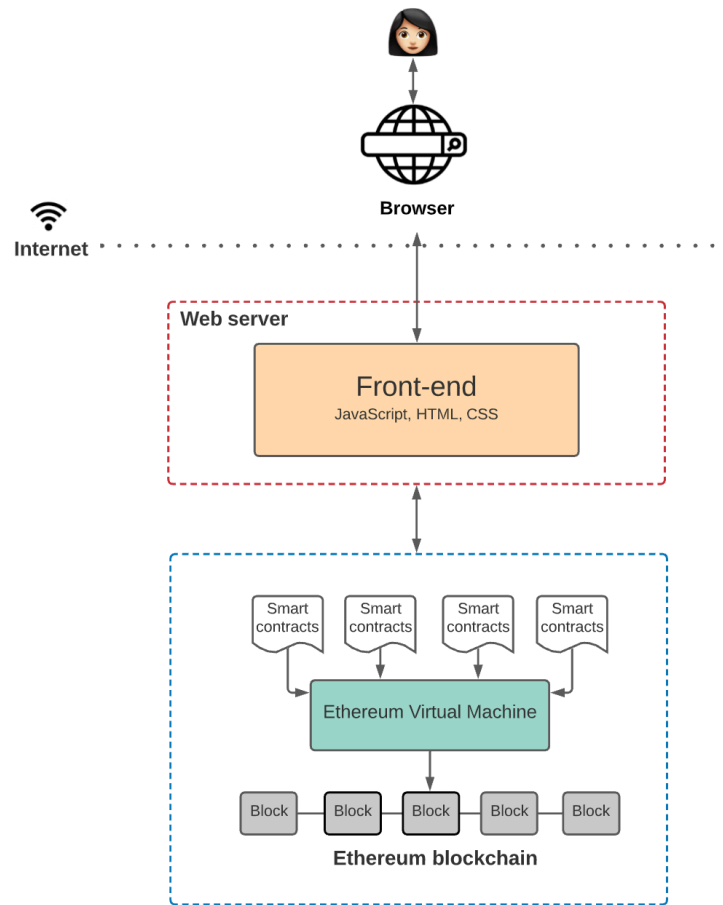
*Possesso dei dati:* Gli utenti riprenderebbero possesso dei propri dati potendo decidere con chi condividerli e in che modo, potendo potenzialmente guadagnare dalla vendita dei propri dati attraverso smart contracts alle grandi multinazionali come FaceBook e Google le quali hanno tantissime informazioni sugli utenti e gli advertiser che pubblicano le pubblicità sulle loro piattaforme pagano milioni per avere questi dati.

*persistenza dei dati:* I dati non possono essere cancellati, a meno che non venga cancellata l'intera blockchain, in quanto verranno salvati in maniera ridondante su diversi nodi distribuiti indipendentemente.

L'architettura del web 3.0 non è ancora stata definita in maniera chiara e ufficiale, ma la sua caratteristica chiave sarà certamente l'assenza di divisione netta tra utenti e fornitori di servizi. Per esempio quando ci colleghiamo a FaceBook siamo utenti e quest'ultimo agisce come provider offrendoci un servizio in cambio dei nostri dati personali. La prossima iterazione del web



Figura 2.1: Livelli del web 3.0



permetterà di eliminare questa contrapposizione netta, perchè gli utenti potranno essere anche fornitori di servizi, almeno nelle Blockchain permissionless [2].



# Capitolo 3

## Learning Analytics

La delicata tematica di privacy e sicurezza affrontata nel capitolo precedente, caratterizza fortemente il ramo del Learning Analytics. Questo particolare campo emergente nell'ambito dell'educazione si concentra nel raccogliere, misurare e analizzare i dati raccolti sugli studenti e sui loro corsi di studio, con l'obiettivo di ottimizzare l'esperienza educativa e migliorare non solo i loro risultati, ma anche i metodi di insegnamento.

Grazie alla crescente digitalizzazione e all'adozione di strumenti come i Learning Management System (LMS), i social media e i corsi online aperti e massivi (MOOCs), è possibile raccogliere una grande quantità di dati relativi al comportamento degli studenti, ai loro risultati e alle interazioni con i materiali didattici.

Il Learning Analytics offre un enorme aiuto a tutti gli stakeholder coinvolti nel processo educativo:

- Agli studenti che possono riflettere sui propri progressi e migliorare il proprio apprendimento attraverso feedback personalizzati e a loro volta dare feedback sui corsi per aiutare i docenti a migliorare i metodi di insegnamento.
- Ai docenti che possono adattare i contenuti dei corsi in base ai feedback e ai risultati degli studenti e identificare tempestivamente gli studenti in difficoltà.
- Agli amministratori accademici che possono utilizzare i dati per prendere decisioni basate sull'evidenza e sviluppare strategie efficaci per migliorare l'efficienza e la qualità dell'istruzione.



Figura 3.1: Dragan Gasevic

(a) Padre di Learning Analytics

Proprio per questo le principali applicazioni del Learning Analytics sono il monitoraggio delle prestazioni individuali, la prevenzione dell'abbandono scolastico, la personalizzazione dei percorsi educativi e l'analisi delle tecniche di valutazione e dei curricula [17].

### 3.1 Limiti del Learning Analytics

Nonostante la disponibilità di standard di riferimento per la gestione dei dati di apprendimento su un LRS (Learning Record Store), è ancora difficile raggiungere l'interoperabilità, ovvero la capacità di due piattaforme differenti di scambiare le informazioni in maniera indipendentemente, senza alcune limitazioni.

Questi problemi includono:

- Collegare le storie di apprendimento di uno studente su diverse piattaforme di apprendimento in un unico percorso immutabile, in modo tale che ogni studente abbia un'unica identità nel web e non diverse identità in base alle piattaforme su cui si collega.
- Garantire la privacy dei registri degli studenti con facilità nel controllo degli accessi.
- Integrare i sistemi di ricerca e produzione per migliorare l'apprendimento.

#### 3.1.1 Collegare le storie di apprendimento

Sebbene gli studenti spesso passino da una piattaforma di apprendimento di un provider a un'altra, i loro record di apprendimento vengono memorizzati separatamente in LRS distinti e in modo disconnesso. Di conseguenza, ogni sistema deve sostenere il costo di ricostruire i dati dell'apprendente da zero anche per casi molto semplici. Anche se questo potrebbe non rappresentare uno sforzo ripetuto per i principianti, è quasi impossibile determinare se uno studente sia realmente un principiante. Questo causa anche un problema di "*cold start*" nei sistemi di raccomandazione per la formazione, a causa della mancanza di azioni di apprendimento precedenti degli studenti [18]. I sistemi proposti dovrebbero consentire agli studenti di portare con sé i propri dati di apprendimento, nello stesso modo in cui possono facilmente trasferire i certificati da un'istituzione all'altra.

A tal proposito tornerebbe utile il design del Web 3.0 presentato nel capitolo precedente che permetterebbe a ogni studente di costruirsi un'identità digitale univoca indipendente dal provider di servizi e che consentirebbe di andare a risolvere il problema di cold start, in quanto non servirebbe più ricostruire ogni volta la storia di apprendimento di uno studente, ma basterebbe andare a leggere la sua identità digitale [19].

### **3.1.2 Privacy, sicurezza e controllo degli accessi**

Questo è un altro problema quando si condividono i risultati di apprendimento individuali degli studenti con terze parti.

Sebbene l'analisi dell'apprendimento aiuti a migliorare le prestazioni degli studenti, attraverso una domanda generale e quattro specifiche sulle condizioni per la privacy degli studenti, sostengono che, qualunque siano i guadagni dalle analisi dell'apprendimento, questi devono tener conto sempre al rispetto della privacy e dei diritti degli studenti, quindi non si potrà andare a violare la privacy dello studente per avere una maggiore efficienza dalle analisi dell'apprendimento, questo perché il trauma psicologico che potrebbe derivare da una singola violazione della privacy può essere devastante, poiché è possibile rivelare informazioni riservate e dunque sensibili.

I sistemi proposti devono quindi garantire in primis la priorità al rispetto alla privacy degli studenti, lasciando a loro il controllo decisionale dei propri dati di apprendimento [19].

### **3.1.3 Integrare i sistemi di ricerca e produzione**

La disponibilità di dati di apprendimento per la ricerca favorisce l'innovazione. Nei casi in cui i dati di apprendimento siano raccolti da sistemi di produzione e/o di ricerca, i ricercatori di learning analytics spesso si trovano a dover affrontare il difficile compito di anonimizzare le informazioni personali con lo scopo di mantenere protetta la privacy degli stakeholder, facendo così sì che ha però un'immensa dispersione di risorse impiegate nel processo di anonimizzazione e inoltre si ha un impatto negativo sui risultati, perché si va a perderne la loro caratterizzazione e personalizzazione.

Poiché i dati di apprendimento in tempo reale diventano sempre più desiderabili per la ricerca in learning analytics, è cruciale sviluppare nuove idee su come integrare in modo fluido e interoperabile i sistemi di ricerca e produzione, garantendo al contempo la privacy di tutti gli stakeholder coinvolti [19].

## **3.2 Blockchain come soluzione**

Ogni stakeholder deve quindi avere diritti diversi sui dati e metadati, di conseguenza, i dati devono essere circondati da diversi livelli di accesso degli utenti, ciascuno con permessi specifici e unici, in modo tale che gli studenti possano avere accesso solamente ai loro dati, gli insegnanti ai dati di tutti i loro studenti, mentre i genitori solo ai dati relativi ai propri figli.

La natura di questi diritti e il modo in cui vengono applicati devono essere concordati in modo esplicito e chiaramente comprensibile da tutti, inoltre potrebbero essere validi solo per periodi di

tempo limitati e vincolati al soddisfacimento di determinate condizioni, in modo tale da garantire un'adeguata anonimizzazione dei dati per rispettare la privacy di ciascun stakeholder.

Ha quindi senso archiviare tutti i dati e metadati relativi allo studente rilevanti per le Learning Analytics in un decentralized distributed ledger basato su un sistema Blockchain [20].

Grazie a questo sistema si potranno andare a soddisfare i seguenti requisiti:

*Autenticità:* I *"Data and metadata About Learning Processes"* (DALP) avranno la prova di essere validi e di essere stati aggiunti da utenti autorizzati, questo è molto importante perché dire *"Uno studente sostiene di aver preso un 30 all'esame"* non è la stessa cosa di dire *"L'insegnante sostiene che lo studente abbia preso 30"*.

*Integrità:* I dati non potranno essere modificati.

*Controllo:* Solo gli stakeholder autorizzati avranno il diritto di accedere ai dati, ma solo ai dati che gli corrispondono.

*Conoscenza:* Verrà sempre tenuto traccia sul registro di chi ha acceduto ai dati e il motivo per cui lo ha fatto.

*Sicurezza:* I dati saranno protetti da furti e attacchi informatici, grazie al design della Blockchain.

È facile notare come la nascita di una piattaforma che necessita questi requisiti sarebbe molto più spontanea e naturale e meno costosa se fosse inserita in un ecosistema che la favorisse soddisfacendo già a sua volta i requisiti, è importante quindi evidenziare ancora una volta l'importanza che avrebbe il passaggio del web da 2.0 a 3.0 e gli innumerevoli vantaggi che ciò porterebbe.

Nei paragrafi che seguiranno andremo a vedere nel dettaglio tre esempi tangibili di Learning Analytics platform Blockchain-based che sono al momento disponibili sul web.

### 3.3 EduCTX

La più importante piattaforma di Learning Analytics basata su Blockchain è EduCTX, sviluppata da un consorzio di università europee sfrutta la Blockchain 2.0 di Ethereum incentrata sugli smart contracts.

Al momento questa piattaforma presenta due versioni, la versione 1.2 sulla quale noi focalizzeremo il nostro interesse in quanto costruita su Blockchain sfruttando il portale di registrazione MetaMask e la versione 2.0 che invece si basa su Microsoft Azure Active Directory un servizio di gestione d'identità centralizzato basato su cloud.

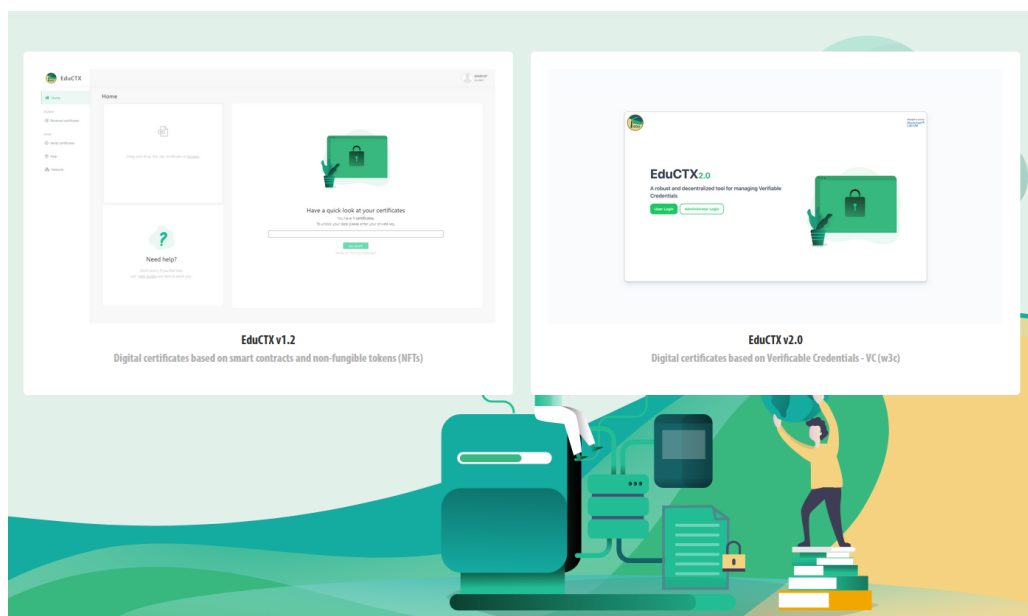


Figura 3.2: EduCTX homepage

#### 3.3.1 Struttura

#### 3.3.2 Come avviene la registrazione degli studenti

### 3.4 Sony Global Education

### 3.5 Blockcherts





# Bibliografia

- [1] klaus Schwab, *La quarta rivoluzione industriale*. FrancoAngeli, 2019, ISBN: 9788891743008.
- [2] N. Attico, *Blockchain, guida all'ecosistema*. Guerini Next, 2018, ISBN: 9788868962180.
- [3] R. B. Gianluca Chiap Jacopo Ranalli, *Blockchain, tecnologia e applicazioni per il business*. Hoepli, 2019, ISBN: 9788820389253.
- [4] T. C. G. I. in Criptovalute, *Ethereum: The Merge Explained*, Accessed: 2023-10-01, 2022. indirizzo: [https://www.youtube.com/watch?v=EOeNrCzVvfs&t=2598s&ab\\_channel=TheCryptoGateway-InvestireinCriptovalute](https://www.youtube.com/watch?v=EOeNrCzVvfs&t=2598s&ab_channel=TheCryptoGateway-InvestireinCriptovalute).
- [5] Kraken, *Documentazione su Kraken*, Accessed: 2023-10-01, 2023. indirizzo: <https://kraken.docsend.com/view/58b6xidjxk44xedc>.
- [6] A. Borroni, «Blockchain: Uses and Potential Value,» in *Legal Perspective on Blockchain Theory, Outcomes, and Outlooks*, A. Borroni, cur., Pubblicazioni del Dipartimento di Scienze Politiche Jean Monnet dell'Università degli Studi della Campania Luigi Vanvitelli, ESI, 2019.
- [7] S. Rodotà, *Quattro paradigmi per l'identità*. Bari: Vivere la democrazia, 2018, 20 ss.
- [8] G. Alpa, «L'identità digitale e la tutela della persona. Spunti di riflessione,» *CONTR. IMPR.*, p. 725, 2017.
- [9] P. D. Filippi, *The Interplay Between Decentralization and Privacy: The Case of Blockchain Technologies*, 1, Paris, 2016.
- [10] B. Marr, *A Very Brief History of Blockchain Technology Everyone Should Read*, Forbes, Accessed: 2023-10-01, 2018. indirizzo: <https://www.forbes.com/sites/bernardmarr/2018/02/16/a-very-brief-history-of-blockchain-technology-everyone-should-read>.
- [11] F. Zambardino, «La blockchain nel mercato del lavoro italiano con particolare enfasi sulla questione privacy e il rapporto con il General Data Protection Regulation,» *TSL*, pp. 23–38, 2022.

- [12] G. R. Mayes, *Privacy and Transparency*, 2018.
- [13] C. Bridge, «Blockchain's Next Frontier: Cloud Computing?» *Inv. Mark't Bus. Res.*, 2018, Concettualmente, in particolare, «[c]loud storage allows the user to store data and information online. This serves as a backup in case the data is lost and could be used to secure large amounts of data». Ibid.
- [14] D. Reisman, J. Schultz, K. Crawford e M. Whittaker, *Algorithmic Impact Assessment: A Practical Framework for Public Agency Accountability*, Consultato il 08 giugno 2022, 2018. indirizzo: <https://ainowinstitute.org/aiareport2018.pdf>.
- [15] L. Zhang, *End to end architecture*, cit., 3.
- [16] TED, *Tim Berners-Lee on the Next Web*, Accessed: 2023-10-01, 2009. indirizzo: [https://www.youtube.com/watch?v=OM6XIICm\\_qo&ab\\_channel=TED](https://www.youtube.com/watch?v=OM6XIICm_qo&ab_channel=TED).
- [17] Wikipedia contributors, *Learning Analytics*, Accessed: 2023-10-01, 2023. indirizzo: [https://en.wikipedia.org/wiki/Learning\\_analytics](https://en.wikipedia.org/wiki/Learning_analytics).
- [18] T. Barnes e J. Stamper, «Toward automatic hint generation for logic proof tutoring using historical student data,» in *International Conference on Intelligent Tutoring Systems*, Springer, Berlin, Heidelberg, 2008, pp. 373–382.
- [19] P. Ocheja, B. Flanagan e H. Ogata, «Connecting Decentralized Learning Records: A Blockchain Based Learning Analytics Platform,» *Proceedings of the 8th International Conference on Learning Analytics and Knowledge*, pp. 265–269, 2018, Author's version - not the published version. Cite published version. DOI: 10.1145/3170358.3170365. indirizzo: <http://hdl.handle.net/2433/231940>.
- [20] M. A. Forment, D. A. Filvà, F. J. García-Peñalvo, D. F. Escudero e M. J. Casany, «Learning Analytics' Privacy on the Blockchain,» in *Proceedings of the 6th International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM 2018)*, F. J. García-Peñalvo, cur., New York, NY, USA: ACM, 2018, p. 5. DOI: 10.1145/3284179.3284231. indirizzo: <https://doi.org/10.1145/3284179.3284231>.