NIST Internal Report
NIST IR 8475 ipd

A Security Perspective on the Web3 Paradigm

Initial Public Draft

Dylan Yaga
Peter Mell

This publication is available free of charge from:
https://doi.org/10.6028/NIST.IR.8475.ipd

NIST NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY
U.S. DEPARTMENT OF COMMERCE

NIST Internal Report
NIST IR 8475 ipd

A Security Perspective on the Web3 Paradigm

Initial Public Draft

Dylan Yaga
Peter Mell
Computer Security Division
Information Technology Laboratory

This publication is available free of charge from:
https://doi.org/10.6028/NIST.IR.8475.ipd

April 2024

U.S. Department of Commerce
Gina M. Raimondo, Secretary

National Institute of Standards and Technology
Laurie E. Locascio, NIST Director and Under Secretary of Commerce for Standards and Technology

Publication History
Approved by the NIST Editorial Review Board on YYYY-MM-DD [Will be added in the final publication]

How to Cite this NIST Technical Series Publication:
Yaga D, Mell PM (2024) A Security Perspective on the Web3 Paradigm. (National Institute of Standards and Technology, Gaithersburg, MD), NIST Interagency or Internal Report (IR) NIST IR 8475 ipd. https://doi.org/10.6028/NIST.IR.8475.ipd

Author ORCID iDs
Dylan Yaga: 0000-0003-4058-3645
Peter Mell: 0000-0003-2938-897X

Public Comment Period
April 11, 2024 – May 27, 2024

Submit Comments
ir8475-comments@nist.gov

National Institute of Standards and Technology
Attn: Computer Security Division, Information Technology Laboratory
100 Bureau Drive (Mail Stop 8930) Gaithersburg, MD 20899-8930

All comments are subject to release under the Freedom of Information Act (FOIA).

Abstract

Web3 is a proposed vision for the future of the internet that is restructured to be more user-centric with an emphasis on decentralized data. Users would own and manage their personal data, and systems would be decentralized and distributed. Digital tokens would be used to represent assets, and web-native currencies (such as cryptocurrencies) would be used for payments. This document provides a high-level technical overview of Web3 and discusses the technologies that are proposed to implement it. The integration of these developing technologies may present novel security challenges, so this paper presents security considerations that should be addressed when considering Web3 technology and adoption.

Keywords

blockchain; cryptocurrency; data; decentralized; decentralized identity; non-fungible tokens; smart contracts; tokens; Web3.

Reports on Computer Systems Technology

The Information Technology Laboratory (ITL) at the National Institute of Standards and Technology (NIST) promotes the U.S. economy and public welfare by providing technical leadership for the Nation's measurement and standards infrastructure. ITL develops tests, test methods, reference data, proof of concept implementations, and technical analyses to advance the development and productive use of information technology. ITL's responsibilities include the development of management, administrative, technical, and physical standards and guidelines for the cost-effective security and privacy of other than national security-related information in federal information systems.

Audience

This publication is designed for readers with little or no knowledge of Web3 technology who wish to understand how it works at a high level. It is not intended to be a technical guide. The discussion of the technology provides a conceptual understanding, and some examples, figures, and tables are simplified to fit the audience.

Call for Patent Claims

This public review includes a call for information on essential patent claims (claims whose use would be required for compliance with the guidance or requirements in this Information Technology Laboratory (ITL) draft publication). Such guidance and/or requirements may be directly stated in this ITL Publication or by reference to another publication. This call also includes disclosure, where known, of the existence of pending U.S. or foreign patent applications relating to this ITL draft publication and of any relevant unexpired U.S. or foreign patents.

ITL may require from the patent holder, or a party authorized to make assurances on its behalf, in written or electronic form, either:

a) assurance in the form of a general disclaimer to the effect that such party does not hold and does not currently intend holding any essential patent claim(s); or

b) assurance that a license to such essential patent claim(s) will be made available to applicants desiring to utilize the license for the purpose of complying with the guidance or requirements in this ITL draft publication either:

    i. under reasonable terms and conditions that are demonstrably free of any unfair discrimination; or

    ii. without compensation and under reasonable terms and conditions that are demonstrably free of any unfair discrimination.

Such assurance shall indicate that the patent holder (or third party authorized to make assurances on its behalf) will include in any documents transferring ownership of patents subject to the assurance, provisions sufficient to ensure that the commitments in the assurance are binding on the transferee, and that the transferee will similarly include appropriate provisions in the event of future transfers with the goal of binding each successor-in-interest.

The assurance shall also indicate that it is intended to be binding on successors-in-interest regardless of whether such provisions are included in the relevant transfer documents.

Such statements should be addressed to: ir8475-comments@nist.gov

Table of Contents

List of Tables

Author Contributions

Author 1: Conceptualization, Investigation, Writing – Original draft

Author 2: Writing – Original draft

1. Introduction

Web3 is a proposed vision for the future of the internet. It is not a specific single design, architecture, or software but rather a goal for restructuring the internet to be more user-centric. Users would own and manage their personal data, acting as gatekeepers to other applications and services that need it. Systems would be implemented in a decentralized and distributed manner while also providing for direct user participation. Digital tokens would be used to represent assets, and web-native currencies (such as cryptocurrencies) would be used for payments.

This document provides a high-level technical overview of Web3 and enumerates its envisioned components. This paper also discusses the various technologies that are proposed components of Web3. Many of these technologies already exist in different stages of technical maturity. The concrete work in Web3 is largely in maturing these technologies and integrating them to create something greater than the sum of its parts. This integration may present novel security challenges, so this paper uses its Web3 technical description to present security considerations for Web3 technology and adoption.

Opinions and evaluations of the utility and feasibility of the Web3 vision are out of scope for this document, which takes no position on whether the Web3 vision can or should be implemented. For readers who are interested in learning about the case for Web3 adoption, a variety of resources are available [1][2][3]. This paper does not delve into the philosophies held by some Web3 proponents and does not take a position on them.

The remainder of this document is organized as follows.

- Section 2 provides a short history of the internet through a discussion of its early generations: Web 1.0 and Web 2.0.
- Section 3 discusses the vision for Web3 and its technical components.
- Section 4 considers the potential security and privacy issues that may arise.
- Section 5 provides a conclusion.

2. Background

The internet can be viewed in terms of "generations" of capabilities. These generations are often divided into the nascent Web 1.0, the current Web 2.0, and a conceptual next generation Web 3 (named Web3). This section provides an overview of Web 1.0 and Web 2.0. It concludes with a brief discussion of a separate concept called Web 3.0, which embodies a different vision than Web3 (while unfortunately sharing a similar name). This context is provided to highlight where Web3 diverges from the existing web and differs from the separately envisioned Web 3.0.

2.1. Web 1.0 – The Nascent Web

In the beginning, the internet hosted very basic websites that were mostly comprised of text (often just plain text but sometimes with simple formatting), images, and hyperlinks to other webpages. As a result, this era has since been dubbed the "static" or "read-only" web. Most websites were hosted by either large, tech-savvy organizations; government organizations; internet service providers; or tech-savvy users who were allotted a small portion of web storage from their internet service provider or other web-hosting provider to develop their own "home page." Websites eventually began to create more designs and styles via Hypertext Markup Language (HTML) tables, which allowed developers to change the format of their page.

At this time, there were also some "dynamic" pages on the web that used the Common Gateway Interface (CGI) to execute code on the server and generate a static webpage to be delivered to the end user. There were little to no client-side manipulatable websites. Online communications were done through email, bulletin boards, and forums, and there was almost no online shopping during this period. Since there was not a lot of user interaction, organizations hosted massive amounts of user data.

2.2. Web 2.0 – The Current Web

As the internet grew, so too did the number of use cases for it. Web servers continued to gain features and integrate more technologies, such as databases. The user interface of the internet – the web browser – also continued to evolve and gain new features. The development of multiple browsers gave users more freedom of choice.

Developers also found new methods for user interaction. In the beginning, these methods were largely closed source or proprietary technologies (e.g., Adobe Flash, Microsoft Silverlight, JAVA Applets) but eventually migrated to standardized and/or open-source technologies (e.g., HTML 5, JavaScript, and utilizing Document Object Model [DOM] manipulation). Communication methods expanded from forums and email to chatting, messaging, and social media. Many active web users also saw their "web presence" migrate through several genres over the lifespan of Web 2.0 – from personal home pages to online web journals and burgeoning social media platforms, such as MySpace, to more modern social media platforms, such as Facebook and Twitter [4]. The web also became more media rich as it evolved, creating spaces for sharing images and videos (e.g., Instagram, YouTube, and TikTok).

The development of websites also saw a major leap by splitting style and content into two portions. Style is now largely handled by Cascading Style Sheets (CSS), and content is handled by webpages. Developers no longer embed tables within tables to achieve specific designs. This split has allowed for easier manipulation of content on the client side. This era of the internet was dubbed the "interactive," "participative," or "social" web since websites became more interactive and responsive to user input, and users migrated toward social media websites.

There was a significant growth of organizations offering multiple interconnected services (e.g., Google's Gmail and Drive, Microsoft's Hotmail/Outlook and OneDrive, Apple's iCloud) free of charge. Eventually, these organizations became hosts to massive amounts of user data.

As mobile devices advanced in power and pervasiveness, organizations could collect significant data from them, and as the world began to "make an app" out of online services, organizations realized they could get more data from a smartphone than a website. Through smartphones, organizations had access to a myriad of sensor data, geolocation data, contact information, and stored media, all of which was made accessible through application permission requests.

As organizations continued to expand and collect user data, they also began to diversify their offerings. Many opened online storefronts that allowed users to purchase licenses to view digital media such as music, books, and films[1] . Users quickly found themselves becoming more attached to individual platforms. Users could not easily migrate away from their chosen platform since their licenses were specific to that platform. Some users found out too late that if they were removed from an organization's platform, they lost all access to the media they had purchased licenses for [5]. In many cases, users were unable to return digital content that they were unhappy with or transfer digital content to other users (either a temporary transfer, such as lending to another user, or permanent transfer, such as selling a digital item to another user). This change from the ownership of physical items to licenses to view digital content was seen by many as a step backwards. Proponents of Web3 saw the mass collection of user data, platform lock-in, and the inability to obtain and transfer the ownership of digital items as issues with Web 2.0.

2.3. Web3 vs. Web 3.0 – The "Semantic" Web

Web 3.0 is different from Web3, though they share a similar name. Web 3.0 is known as the "semantic" web. It is an effort to make the internet more machine-readable by adding additional metadata, such as tags and identifiers, to data hosted on websites. These tags would enable computers to process web data and allow for data to be shared and reused across different applications more easily. By utilizing specific tags, users can find similar resources that use the same tags instead of needing a direct hyperlink between the two sources. This change allows for faster discoverability of data. Currently, the Semantic Web has not reached widespread adoption or use. Web 3.0 will not be further discussed in this paper. For more information on the semantic web, see [6].

[1] Most online storefronts do not allow users to purchase the actual digital media for certain media types but rather a limited license to view the digital media through authorized applications. This license can be revoked and acts as a form of digital rights management or DRM

3. Web3 Overview

This section provides an overview of Web3. It discusses the Web3 vision, data model, and technological components and concludes with a discussion of Web3 benefits and challenges.

3.1. Web3 Vision

The definition provided below is intended to be descriptive and inclusive of all Web3 applications. It is not intended to define what is or what is not part of Web3, nor is it intended to limit future Web3 applications. The purpose of the definition and resultant characteristics is to enable the reader to understand the current proposed technology and to provide a foundation for an exploration of potential security and privacy issues.

> Web3 is a restructuring of the internet to place ownership and operation into the hands of users themselves, thus changing the structure from organization-centric to user-centric.
>
> Web3 proposes several changes to the existing web architecture:
>
> - Users own their data and are responsible for their data, data security, and data privacy.
> - Decentralized and distributed systems are used, and users can host and run applications.
> - Applications and organizations request data directly from users.
> - Users can supply applications and organizations with actual data or verifiable credentials/verifiable presentations of their data or choose to deny applications and organizations access to their data.
> - Applications and organizations may offer incentives for users to provide data.
> - Data can be tokenized and transferred directly between users.
> - Application execution and transaction fees are paid for with web-native currencies (e.g., cryptocurrencies).
> - Users who execute application logic and maintain the state of systems can receive payment in web-native currencies (e.g., cryptocurrencies) for doing so.

This description leads to several characteristics of Web3, which are documented in **Table 1**.

Table 1. Web3 characteristics

Characteristic	Description
Data Ownership	Web3 seeks to have users own their data. This can enable the portability of data and the transfer of data ownership. Users will need to securely store their data and manage requests for their data. Users will be able to determine the level of security to place on their data, as well as where, when, how, how long, and with whom they share their data.
Decentralized	Web3 is envisioned to be operated by those who use it and provide an infrastructure that anyone can build upon through blockchain technology. See [7] for more information on blockchain technology.
Distributed	Web3 applications are envisioned to deployed across the Web3 infrastructure and executed by multiple users with smart contracts deployed on a blockchain. See [7] Section 6, entitled "Smart Contracts," for more information.
Verifiable Credentials and Verifiable Presentations	Web3 users can either provide information directly or utilize verifiable credentials to prove information without providing the underlying data. W3C has a Verifiable Credentials Model that can provide verifiable credentials and verifiable presentations [8]
Incentives	Since users may be reluctant to give data away, organizations that require users' data may provide additional incentives, such as digital asset (e.g., tokens, cryptocurrency) or expanded application capabilities. Users may also choose, and be incentivized, to maintain the integrity of the networks, verify transactions, and execute applications.
Tokenization and Digital Assets	Web3 is envisioned to rely on both fungible and non-fungible tokens to represent data and digital assets that can be exchanged between users.
Web-Native Currency and Cryptocurrency	Web3 is envisioned to use web-native currencies, such as cryptocurrency, for the basis of purchases, money exchange between users, and the cost of executing distributed applications.

3.2. Web3 Data

Implementing the proposed vision of Web3 would require changes to data, data ownership, data location, and data access. Currently, much of the internet's data is proprietary, highly application-specific, and non-interoperable. In most cases, even user data is owned by the organization that provides the platform rather than the user. **Table 2** describes and compares the current data model with the proposed Web3 data model.

Table 2. Current web data model vs. Web3 data model

Data Aspect	Current Model	Web3 Model
Data	While there are many standardized data formats for various media (e.g., images, sound, video), non-media data is largely application specific. Interoperability between applications is cumbersome and often requires data translation and transformations. Often, a loss of data or data precision occurs.	Open standardized data formats for non-media data would allow for interoperability between organizations and greater user freedom. Some data can be replaced by verifiable credentials and verifiable presentations to help preserve private information.
Data Ownership	Most user data is owned by organizations. End-user agreement documents generally limit the rights of users over the data within applications. Users typically cannot give, trade, or sell their data to other users. While many organizations have a “Data Export” feature in their applications, few have a “Data Import” feature, meaning that the data itself is tightly bound to the application that created it. Data can also be perfectly copied an infinite number of times, meaning that there is no scarcity of the data, and provenance is quickly muddied.	Most user data is owned by users. Data ownership can be proven through use of digital signatures. For private information, users can elect to use trusted third parties to create verifiable credentials so that the information remains private but external organizations can obtain the results. For organizations that need access to private data, users can elect to allow access (e.g., stored off of a blockchain, in a secure data hub, or with a decentralized cloud service) at a granular level. Access to this data can be revoked after a set period or at the user’s discretion. Data itself can be tokenized on a blockchain, which allows for transfer of ownership and provides full provenance.
Data Location	Data is stored by the organization within databases that consist of many users’ data. User data is also redundantly contained across multiple different applications, as each one needs to maintain its own copy of user data, resulting in users needing to update each application whenever data changes.	Public data and verifiable credentials/verifiable presentations are posted on a blockchain. For large data, it may be necessary to utilize a decentralized online storage location with pointers to it posted on a blockchain [9]. Private information is stored on an external secure data hub.

Data Aspect	Current Model	Web3 Model
Data Access	Data contained within applications can be accessed, modified, removed, transferred, sold, or monetized at any time without user knowledge.	Public data that is stored on the blockchain itself is easily accessible by anyone. Data that is stored outside of the blockchain may require additional authorization to access. This authorization is done by the user and can be managed at a granular level (as opposed to wholesale access to all data) that is application specific. Access to data stored outside of the blockchain can be revoked after a set period or at the user's discretion.

3.3. Web3 Technology Components

Like Web 1.0 and Web 2.0, Web3 is not a single technology. Rather, Web3 combines longstanding existing technologies and recent technological advancements to accomplish a specific set of goals. Web3 combines use of mobile devices, new forms of digital identities, blockchains, tokens, smart contracts, and verifiable attestations of data. The discussion below is not comprehensive, and Web3 may use additional or alternative technologies.

Web3 utilizes existing internet technologies that make up much of the current web architecture, such as Transmission Control Protocol/Internet Protocol (TCP/IP), remote procedure call (RPC), and Transport Layer Security (TLS). It can also use existing web services, servers, databases, and webpages to act as an interface. Web3 applications can be designed to interact with (as both input to and utilize output from) existing systems. Like blockchain systems and cryptocurrency systems, Web3 leverages well-known technologies, such as public-key cryptography, digital signatures, and cryptographic hashing algorithms.

Web3 can take advantage of the ever-growing access to mobile technology. Mobile devices are highly personal devices that often contain more personal information than personal computers or laptops (which may be shared by multiple people). Mobile devices are not typically shared among multiple users and have a one-to-one relationship between device and user. Modern mobile devices are often equipped with hardware security modules, trusted compute modules, and other modern security features. This scenario sets mobile devices up to be an ideal portal into Web3 technologies. Web3 allows users to take control over their digital identities, decide how others access their personal information, revoke access at their discretion.

Related to this vision, the NIST National Cybersecurity Center of Excellence (NCCoE) is working with Department of Homeland Security Science and Technology Directorate (DHS S&T) on a project to *Accelerate the Adoption of Digital Identities on Mobile Devices* [10].

The NCCoE effort describes the stage for mobile digital identities:

> However, with the proliferation of mobile devices, new digital credentials are emerging that can support both greater individual

> control of identity attributes and more direct validation with issuing sources. This provides the potential for both improved usability and convenience for the end user and stronger assurance in identity for organizations [10].

Governments around the world have been researching methods to expand existing forms of identity into the digital space. Proponents of Web3 call for the use of decentralized digital identities along with verifiable credentials. NIST has investigated multiple emerging blockchain identity management systems [11] that may be utilized by Web3 systems. By employing mobile devices and integrating different types of digital identities, Web3 can help facilitate an identity hub that can incorporate government-issued identities, decentralized identities, and other forms of digital identities.

As part of a digital identity, Web3 proposes the use of Decentralized Identifiers (DIDs) [12]. These DIDs provide a method for a unique identifier to be issued without the need for a central authority and provide mechanisms to prove control of an identifier via cryptographic means. DIDs can be used on their own or as part of a larger system, such as the use of verifiable credentials.

Web3 also plans to enable users to utilize verifiable credentials and verifiable presentations of their data [8]. Verifiable credentials and verifiable presentations allow users to own identifying information about themselves that has been verified by a third party. Users can choose to present a subset of the characteristics of their verifiable credentials to others by generating a verifiable presentation. Others can then verify that the information has been digitally signed by a third party and choose whether or not to trust that third party.

With Web3, there may be an entire decentralized ecosystem of verifiable credential issuing organizations with varying levels of trust among users. To provide an example of verifiable credentials and presentations:

> A user requests an issuing authority to issue them a verifiable credential based on a piece of identifying information that the user provides (e.g., a driver’s license). The issuing authority then performs checks to validate the information and ensure that it belongs to the user before issuing the user a verifiable credential that is digitally signed by the issuing authority. The user can now use the verifiable credential to generate verifiable presentations of the credential in whole or in part (e.g., proof that they are older than 21 but not their birth date) to other users and organizations. These other users and organizations can verify that the presentation came from a verifiable credential and check that the digital signatures match. The verifying user can then accept the verifiable presentation as valid or deny it depending on the level of trust that they have in the issuing organization.

Much of the discussion surrounding Web3 focuses on blockchains, tokens, and smart contracts. These newer technologies are key to the underlying architecture of Web3 and allow for much of the desired features to be realized. Blockchains allow for the system to be decentralized, which affords ownership of digital data. Tokens, as part of a blockchain, allow for data to be transferred rather than simply copied. Smart contracts allow for these systems to automate procedures, perform more complex transactions, and record the results on the blockchain itself.

3.4. Web3 Discussion

One of the main goals of Web3 is to change the data ownership model of the internet. Today, many users give up certain rights to the data they generate within applications and platforms as part of agreeing to the terms of use for the platform. The data that users generate is a valuable resource to each application, and the organizations that run those applications can use the data to generate additional revenue. The sale of user information – or even access to the user via anonymized data – is often done without the user's knowledge and does not directly benefit the user.

Web3 proposes that rather than organizations owning and storing user data, users themselves should own and store their own data and provide organizations access to portions of that data when necessary (e.g., verifiable credentials and verifiable presentations of data). With this change, users would know exactly when an organization needed their data and what data was needed, which would allow the user to allow or deny an organization access to that data (potentially denying access could also result in the application failing to work properly).

Web3 facilitates the shift of organization-centric data ownership to user-centric data ownership by proposing a shift from centralization to decentralization of applications and data. Decentralized applications would take the form of smart contracts and be hosted and run on a blockchain. Users of these decentralized applications could publish art, documents, and other application-specific data by posting either the actual data or a cryptographic hash representation of the data to a blockchain or smart contract. However, sensitive data, such as personally identifiable information (PII), is not something that many users would want hosted on a blockchain (even if encrypted). Users would instead have some form of data storage hub where they stored their data off of a blockchain and have verifiable credentials issued and verifiable presentations of information posted to the blockchain.

The shift from centralized to decentralized would affect both users and organizations. For organizations, it would mean relinquishing much of the data ownership that they privately hold. Many organizations may see this as disadvantageous to their businesses, as the data would not be exclusively theirs to utilize, and thus, they may be reluctant to migrate to a Web3 application. However, there may be some beneficial trade-offs. Much – if not all – of the user data could be migrated away from organizations and into the hands of users themselves, which would reduce much of the burden that organizations face with securing private user data. Due to the reduced amount of data held, organizations would be less of a target for malicious attackers who seek to steal the data. Organizations could also utilize a much larger pool of data posted by other organizations and users within blockchain systems. Users may even choose to accept incentives from organizations to share data that they would have been reluctant to share in the past, allowing organizations to gain greater insight into their users.

Web3 could provide a shared data layer that applications could be designed to utilize. Since the focus of data would move from being application-centric to user-centric, users would be able to utilize their data across multiple applications without needing to reenter it into each new system or export/import it from somewhere else.

While Web3 could provide a shared data layer, it would not provide intrinsic interoperability. Even if the data is present within a smart contract or on a blockchain itself, some organizations may choose to implement proprietary data formats to facilitate lock in. To prevent this, open data format standards would need to be developed and adopted by communities, organizations, and users.

With the current Web 2.0 model, users often accept third-party hosting of their personal data to acquire a “free” service. Often, complex user agreements are in place that allow organizations to access, exchange, and potentially sell user information either directly, or by providing access to the user for advertising or marketing purposes, without directly notifying the user that a transaction has taken place. With Web3’s proposed changes, user data would need to be explicitly requested from the user. Once users are aware of how often an organization or application utilizes their data, they may be reluctant to allow it. Organizations may then need to provide greater incentives to access user data.

User incentives could be monetary (i.e., organizations pay users for access to their data) or offer increased capabilities within an application (e.g., premium features). With an incentive model in place, organizations could ask for data that users would otherwise be unlikely to share. For example, if an organization wishes to conduct research that requires a large sample pool, they may be able to access more user data by providing greater incentives to users for their data. This exchange benefits both organizations and users.

4. Web3 Security and Privacy

This section discusses some potential Web3 security and privacy challenges. Many Web3 security challenges arise from the increased need for users to be actively involved in protecting and managing their personal data. Others arise from data permanence, the mechanics of blockchains themselves, and the scalability issues of blockchain data. Privacy challenges can arise in the Web3 model due to the public accessibility and permanence of blockchain data.

4.1. Phishing, Scams and Trust in a Decentralized Ecosystem

With the current web architecture (Web 2.0) phishing attacks and scams have been very successful against users, and these malicious techniques will likely continue to be an issue on the internet, even with Web3. With Web3, phishing and scams may be more impactful to the user as an individual, depending on what data the scam seeks to obtain. Since users would be responsible for their data, they may be tricked into giving out far more than what is possible to do in legacy Web 2.0 applications. One of the worst scenarios would be a user giving away their private keys to a malicious actor and allowing them full access to all their data (like giving away a username/password combination in Web 2.0).

Scams are not limited to attempts to steal user data. Scammers may use stolen or “look alike” accounts, posing as someone with influence such as an administrator, support staff or celebrity, on social platforms to entice users to purchase ultimately worthless tokens (both fungible and non-fungible) or to utilize fraudulent websites and services.

There is a significant amount of trust built into the current Web 2.0 ecosystem. This trust has been built up over many years, and most well-known organizations have garnered some degree of trust from users. With Web3, many applications are likely to be developed by organizations that may not be well-known. Users would then need to rely on each other to determine the legitimacy of an application or organization. Malicious actors could use this lack of familiarity to their advantage to harvest user data or exploit a user’s lack of knowledge.

Chainabuse [13], a website where users can “report malicious crypto activity,” shows that phishing scams outnumber the other categories of scams combined. Numerous reports and articles have been posted about the extent of phishing scams and Web3/NFTs [14][15][16].

Chainabuse categorizes scams into three high level categories [17]. The descriptions of these categories by Chainabuse are included below.

- Blackmail

> During a blackmail scam, the scammer demands payment from their victim for not revealing damaging information the scammer claims to have about them or to unblock something their victim needs. Blackmail scams differ in the information scammers leverage to threaten their victims.

- Fraud

  During a crypto fraud, the scammer lures their victim either to have them:

  - Provide personal information associated with login information. Scammers use this login information to sign transactions and transfer funds on the victim's behalf.
  - Transfer crypto funds directly.
  - The scammer can lure their victim into pretending they are someone they are not, promising fake returns, and pretending they are associated with a fake project.

- Hack

  During a hack, the hacker exploits a vulnerability in a smart contract, protocol, infrastructure, or software, or steals information from their victims to gain unauthorized use of their device and transfer funds directly on their behalf.

The Department of Financial Protection & Innovation for the state of California also maintains a *Crypto Scam Tracker* that users can submit complaints to [18].

Phishing and scams will continue to plague the internet for the foreseeable future, and ultimately it is up to users to educate and prepare themselves for the tactics employed by malicious users. Many companies have begun to develop specific Web3 education, advice, glossaries and taxonomies for attacks, phishing and scams to help educate users [19][20][21][22]. Users and developers must adopt a continuous learning model since the threat landscape continues to change and adapt as well.

4.2. Increased User Responsibility and Access Recovery

The shift to users being fully responsible for their own data, security, and privacy may be seen as burdensome to some and beneficial to others. It could provide an opportunity for users to control and utilize their data in ways that they have not been able to in the past, and it could also come with increased responsibilities and complexities for those who are used to organizations maintaining their personal data. Non-technical users may not understand the implications behind the different security and privacy options available to them and may stick with default options in software. This complexity can be reduced with software that abstracts the underlying blockchain technology and has been designed with security and usability in mind. User options should be clearly presented with explanations of benefits and potential issues that may accompany those choices.

Software and hardware failures and loss can occur. With these failures comes the burden of users recovering access to the various systems with which they interact. With Web 2.0 applications, users can enter their credentials or utilize the application's built-in recovery features. For example, a surprising number of users frequently utilize the "Forgot Password" feature provided by many existing applications [23] to restore access.

Web3 applications will be different. Web3 user software will need to ease the burden of recovering and restoring account access. It is currently not computationally feasible to reverse-engineer or regenerate a private key (the underpinning technology behind Web3 accounts). Users will need to be proactive since the only option is for users to set up a recovery scheme ahead of time. It will be necessary to ensure that users have a robust backup system in place so that they can restore their access to accounts with as little friction as possible while also preventing unauthorized users from restoring someone else's account.

It is currently estimated that nearly 20% of the total amount of Bitcoin is "lost" due to users having lost access to their keys [24][25].

4.3. Data Persistence and Difficulty Removing Data

It is often said that the internet "never forgets" [26] and that anything posted to the internet is there forever, which is both true and false. Data posted to the current internet is largely ephemeral and can disappear at any moment. However, copies of the data may have been made and posted in numerous other locations.

Web3, which utilizes blockchains and distributed ledgers, is the inverse. Data posted to a blockchain is likely to remain, and copies of that data made outside of the blockchain will have reduced meaning because all context and provenance will have been removed. Some stand-alone data may be posted to a blockchain, so users and organizations should keep in mind that there may be some difficulty in removing data from such systems and should refrain from posting any sensitive information directly to a blockchain system.

Additionally, both organizations and users will likely make mistakes with Web3 and post sensitive data to the blockchain, and malicious actors may post sensitive data as a form of attack. The removal of this data from the blockchain (also known as rollbacks or reorgs) may not take place immediately if at all. Currently, there are no formalized procedures for seeking to have data removed from blockchain systems, and removal is largely decided by lengthy discussions between organizations and users.

The removal of data may also be costly. To rollback a series of confirmed transactions on a blockchain, the same amount of work must be redone from that block onward (e.g., if a rollback of a transaction is 10 blocks away from the latest block, then all 11 blocks must be remade after removing the confirmed transactions because each block is cryptographically linked to the previous block; see Section 3.7 in [7]). The further back the rollback must go, the more work must be done. This is especially costly for proof-of-work blockchain systems.

Often, the removal of data is controversial among the users of the system and may erode user trust in the system overall or even lead to a chain split. The chain split may occur before the data is removed from the system, meaning that it still exists on a copy of the blockchain.

There is also nothing to prevent individual users or organizations from caching data that they can access from a blockchain into some other database to ensure that the data is available for use or analysis even if it is removed from the blockchain itself.

4.4. User Security Through Decentralization

Compared to large, centralized data sources that malicious actors can attack to steal vast quantities of data on multiple users, Web3's change to users being responsible for their own data would require malicious actors to specifically target individual users. This mean that attacks would be less significant for the system but far more devastating for the individual user who was targeted. Placing data into the hands of users will require them to protect their own data, which would entail securing the data, managing external access to the data, and creating methods to restore their access to the data should their primary means be lost. Users may choose to utilize as much or as little security as desired, use verifiable credentials instead of their actual data, monitor the use of their data, and revoke access to it.

Increased user data control may also result in increased user privacy. Organizations would need to specify exactly what data they need access to and potentially provide users with data retention policies. Users can then decide whether to provide the requested data. In many cases, the user may only need to provide a verifiable attestation of the data rather than the data itself (e.g., proof of age over a specified value but not a specific birth date).

4.5. Errors and Bugs

No hardware or software is immune to errors and bugs. Extensive testing and review can help to prevent, and/or mitigate bugs and errors. Since Web3 is still in the early stages of development, domain-specific best practices have not been established. Web3 will need to rely on existing best practices of existing software development and build upon them. Web3 developers will also need to actively monitor for exploits, mitigate attacks, and quickly deploy fixes to reduce the impact of attacks.

Errors and bugs can be present at any technology layer within Web3, from the blockchain itself, to user interfaces, web servers, operating systems, smart contracts, data oracles, cross-chain bridges, wallet software, and even hardware. Since bugs in one layer of technology can have an adverse effect on another layer of technology, developers will need to monitor all layers for vulnerabilities. Testing, updating, and maintaining up-to-date information on current vulnerabilities and mitigations will help to reduce or eliminate the impact of bugs.

4.6. Inability to Refuse a Transaction

Currently, if a user has a digital asset and can pay the fees to send it to someone else, they can transfer ownership of the digital asset to any address they want. Current blockchain systems do not require a user to accept a transfer of digital assets to them, therefore recipients cannot refuse the transfer. As the use of Web3 systems grows, this inability to refuse assets may become an issue, as users could potentially send unsolicited spam, advertisement transactions, or more malicious digital assets.

A malicious actor could also post data to a blockchain that is illegal in another region and then send it to addresses of people known to be in those regions. The user cannot refuse receipt of the digital asset or even prove that it was unsolicited. Even if the user burns the digital asset, it can still be proven that they owned it at one time, and that fact may be used against them in a legal system.

4.7. Availability and Denial of Service

The choice of underlying blockchain platform for any given Web3 application will be an important decision in order to avoid availability issues and mitigate potential denial-of-service attacks. Most will likely target larger smart contract-capable blockchains to deploy their Web3 applications. However, there may be issues if a significant number of developers choose the same blockchain, such as execution cost increases, and longer wait times for execution. Scaling solutions are still being actively investigated and developed, so this may become an irrelevant discussion in the future.

Denial-of-service attacks may still occur, as malicious actors attempt to exploit flaws in smart contracts to overwhelm and hinder contract execution [27]. Identifying areas of a smart contract that would need to enforce limits and require additional authentication to prevent denial-of-service attacks will be critical for developers.

Additionally, developers may seek to deploy Web3 applications on multiple blockchain platforms to spread the execution load and potentially reduce operating costs – perhaps even temporarily during peak execution or cost times on their main blockchain platform of choice. To provide maximum benefit, the various deployments will need to interact with one another, so cross-chain bridges will need to be utilized. There have been many articles [28] about cross-chain bridge vulnerabilities, and this will remain a key aspect of security to improve for Web3.

4.8. Censorship Resistance

Since Web3 utilizes blockchain technologies (which are tamper-resistant, tamper-evident, decentralized, and likely distributed in many different geographical locations around the world), removing or censoring data will become more difficult. With the current Web 2.0 model, organizations can remove data at will (or when they are ordered to by law) with ease and without transparency. Since Web3 is used, owned, and operated by many different users where no single user can remove data on their own, a majority of blockchain operators who maintain the blockchain (often called miners) would need to agree to remove data from a blockchain. The operators would know exactly what data was being requested to be removed and could determine whether it was beneficial to them and the system overall.

Some operators may choose to remove the information, while others may not. In the past, decisions such as these have led to chain splits that result in dividing a single blockchain into separate and incompatible versions.

4.9. Chain Splits, Duplicated Applications and Data

A chain split, sometimes also called a hard fork, occurs when a technical modification is made to a blockchain that some users do not wish to adopt, thus making older versions incompatible with the changes[2]. In a chain split, everything (e.g., transactions, cryptocurrency, smart contracts, and smart contract states) up to the point of the split is present on all copies of the blockchain that result from the chain split.

A chain split may be triggered for many reasons, such as changes to the underlying codebase (e.g., fixing an exploit, upgrading cryptographic mechanisms), changes to the blockchain data itself (e.g., reversing a transaction, removing data), and even philosophical differences (e.g., a group of users disagrees with proposed changes). Chain splits do not typically occur out of nowhere, and changes that could lead to them are discussed, debated, and evolve over a long period of time. Most chain splits end up being temporary as users eventually migrate to the blockchain with more users, and the others are abandoned. This is not always the case, and a split chain can retain enough users to maintain its activities.

With Web3, this could lead to unforeseen issues that users and developers would need to address. Web3 smart contract applications would be affected and would continue running on all the different chains that split. For smart contracts built with the ability to self-destruct, the developer could determine which blockchain they wished to support and self-destruct the rest.

However, there are some smart contracts built without the ability to self-destruct to provide users with a sense of longevity in the application. Non-fungible token (NFT) smart contracts are often deployed without the ability to self-destruct. After a chain split, the smart contract and all its NFTs exist on all split chains. This may cause confusion for users and potential investors of those NFTs.

There may also be differences in choice between a Web3 application developer and the users. The developer may pick a specific chain to support after the split, while users may choose another. If the developer decides to only support one of the chains, the users of other chains could lose access to their preferred chain's application.

4.10. User Profiling

Even though one of the goals of Web3 is to move user data away from organizations into the hands of the users themselves, organizations may still choose to store data relating to a user and build a profile. These profiles could be built from a combination of Web3 data and metadata along with existing data about the user that the organization already possessed from existing applications and even public data available on the internet. Organizations could monitor blockchain activity so that they could record user transactions with other users and organizations. Organizations may even attempt to link online users with real world identities.

[2] See section 5.2 in [5] for more information on hard forks.

Since these profiles may be built with indirect data, there will be a level of uncertainty to their accuracy. Assumptions may need to be made by the organization when creating the profile, and attribution of multiple transactions to a single user may be tenuous at best.

Well written user software (such as wallets) could help mitigate this issue by implementing user privacy features, such as automatically (and transparently to the user) using new addresses for every transaction, clearly displaying what information is being requested and what information will be sent.

4.11. Privacy-Preserving Regulations

This paper does not focus on regulations. However, some regulations may conflict with the technical aspects of Web3 applications, so a brief discussion follows.

Some governments have passed privacy-preserving regulations to protect their citizens and enable individuals to request that their data be completely removed from an application. With the proposed Web3 architecture, this may become more difficult to accomplish. Web3 developers will then have to determine how they will accommodate such regulations and whether they are even technically possible to implement. There may also be conflicting regulations in different regions, so developers would need to determine which regulations to follow and what regions they could potentially lose business in. Some regulations may be passed after an application is deployed, so the developer must decide whether they will update the application to adapt to the new regulations. It may be possible for some governments to utilize this as a form of censorship, which is antithetical to Web3.

Alternatively, governments may find it difficult to enforce regulations on a decentralized and distributed system. Application developers may be anonymous, and the applications are hosted and run by the entire decentralized network (the network itself is resilient to disruption and tampering). It may be unclear whether a developer, users, or even an application falls within a regulator's jurisdiction.

5. Conclusion

The Web3 vision proposes significant changes to how the internet functions. As the community creates concrete designs and architectures, it is critical to consider security issues as early as possible. Security should be integrated into the design instead of being added later to a built solution. This paper enumerates a list of potential security and privacy concerns that should be kept in mind as Web3 continues to develop.

References

[1] Saito D (2023) *Creating the internet we deserve: The case for Web3.* Available at https://venturebeat.com/virtual/creating-the-internet-we-deserve-the-case-for-web3/

[2] Lisk (2023) *More Than a Meme – The Case for Web3.* Available at https://lisk.com/learn/about-web3/the-case-for-web-3

[3] Banerjee A, Byrne R, De Bode I, Higginson M (2022) *Web3 beyond the hype.* Available at https://www.mckinsey.com/industries/financial-services/our-insights/web3-beyond-the-hype

[4] Maryville University (2023) *The Evolution of Social Media: How Did It Begin, and Where Could It Go Next?* Available at https://online.maryville.edu/blog/evolution-social-media/

[5] Morrison, G (2021) *You Don't Really Own the Digital Movies You Buy.* Available at https://www.nytimes.com/wirecutter/blog/you-dont-own-your-digital-movies/

[6] W3C (2023) *Semantic Web*. Available at https://www.w3.org/standards/semanticweb/

[7] Yaga D, Mell P, Roby N, Scarfone K (2018), Blockchain Technology Overview. (National Institute of Standards and Technology, Gaithersburg, MD), NIST Interagency or Internal Report (IR) NIST IR 8202. https://doi.org/10.6028/NIST.IR.8202

[8] W3C (2022) *Verifiable Credentials Data Model v1.1.* Available at https://www.w3.org/TR/vc-data-model/

[9] STORE (2023) *Exploring Data Ownership in Web3 and Decentralized Cloud Storage Solutions.* Available at https://storecloud.org/blog/exploring-data-ownership-in-web3-and-decentralized-cloud-storage-solutions

[10] Mehta K, Vemury A, Prisby J, Finke J (2023) *Accelerate Adoption of Digital Identities on Mobile Devices*. Available at https://www.nccoe.nist.gov/sites/default/files/2023-03/mdl-project-description-draft.pdf

[11] Lesavre L, Varin P, Mell P, Davidson M, Shook J, (January 14, 2020) *A Taxonomic Approach to Understanding Emerging Blockchain Identity Management Systems*. https://doi.org/10.6028/NIST.CSWP.01142020

[12] W3C (2022) *Decentralized Identifiers (DIDs) v1.0*. Available at https://www.w3.org/TR/did-core/

[13] chainabuse (2023) Available at https://www.chainabuse.com/

[14] Rektify AI (2023) *NFT Discord Hacks Demystified*. Available at https://medium.com/@rektifyai/nft-discord-hacks-demystified-5412937326f4

[15] Moody R (2023) *Worldwide NFT heists tracker.* Available at https://www.comparitech.com/blog/vpn-privacy/nft-heists/

[16] TRM Insights (2022) *Analysis of Recent NFT Discord Hacks Shows Some Attacks Are Connected.* Available at https://www.trmlabs.com/post/trms-analysis-of-recent-surge-in-discord-hacks-shows-some-attacks-are-connected

[17] chainabuse (2023) *Scam Glossary.* Available at https://www.chainabuse.com/glossary

[18] Department of Financial Protection & Innovation (2023) *Crypto Scam Tracker.* Available at https://dfpi.ca.gov/crypto-scams/

[19] Rektify AI (2023) *Attack Playbook.* Available at https://github.com/RektifyAI/attack-playbook/tree/main

[20] Surge (2023) *Learn About Web3.* Available at https://www.surgewomen.io/learn-about-web3

[21] Cointelegraph (2023) *DeFi Scams 101: How to avoid the most common cryptocurrency frauds.* Available at https://cointelegraph.com/learn/defi-scams-101-how-to-avoid-the-most-common-cryptocurrency-frauds

[22] Department of Financial Protection & Innovation (2023) *Glossary*. Available at https://dfpi.ca.gov/crypto-scams/#Glossary

[23] ExpressVPN (2022) *Survey: How much time do you waste resetting your passwords?* Available at https://www.expressvpn.com/blog/survey-how-much-time-do-you-waste-resetting-your-passwords/

[24] The New York Times (2021) *Tens of billions worth of Bitcoin have been locked by people who forgot their key.* Available at https://www.nytimes.com/2021/01/13/business/tens-of-billions-worth-of-bitcoin-have-been-locked-by-people-who-forgot-their-key.html

[25] Chainalysis Team (2020) *60% of Bitcoin is Held Long Term as Digital Gold. What About the Rest?* Available at https://www.chainalysis.com/blog/bitcoin-market-data-exchanges-trading/

[26] Crockett M (2016), The Internet (Never) Forgets. *SMU Science and Technology Law Review* 19(2), Article 4:151-181. Available at https://scholar.smu.edu/scitech/vol19/iss2/4/.

[27] Darkrelay. *Web3 Security Vulnerabilities 2: Comprehensive Guide to Protecting Digital Assets.* Available at https://www.darkrelay.com/post/web3-security-comprehensive-guide-2

[28] Chainalysis Team (2022) *Vulnerabilities in Cross-chain Bridge Protocols Emerge as Top Security Risk*. Available at https://blog.chainalysis.com/reports/cross-chain-bridge-hacks-2022/