Web 1.0 to Web 3.0 - Evolution of the Web and its Various Challenges

Keshab Nath
Department of
Information Technology
Assam University
Silchar-788011, India
keshabnath@live.com

Sourish Dhar
Department of
Information Technology
Assam University
Silchar-788011, India
dharsourish@gmail.com

Subhash Basishtha
Department of
Information Technology
Assam University
Silchar-788011, India
subhash.cse08@gmail.com

***Abstract*— Now a day's people can hardly conceive of life without internet. The World Wide Web (WWW) as the largest global information media through which user can share, read, and writes data through computers connected with internet. WWW has had much progress since its advent. This paper provides a brief idea of the evolution of the web from web1.0 to web3.0. Web1.0 was about connecting and getting information on the net. Web2.0 is about connecting people and Web 3.0 as a web of knowledge. This paper also gives an overview of security and challenges present in web1.0 to web 3.0.**

***Keywords*— Web1.0, Web2.0, Web3.0, Security, Challenges.**

I. INTRODUCTION

The internet and the web is not synonymous both are two separate but related thing. Internet is simply a network of networks where millions of computer are globally connected forming a network in which any computer can communicate with any other computer. World Wide Web is a way of accessing information over the medium of the internet by displaying web pages on a browser, information are connected by hyperlinks ,can contains text, graphics, audio, video.

Web1.0 is the first generation of the web, also known as informational web. User only can read and share information over web pages.

Web2.0 is the read write networking platform, where the user can communicate among each other.

Web3.0 could be define as semantic web, personalization like my yahoo, iGoogle etc. It changes the web into a language that can be read and categorized by the system rather than human.

II. WEB 1.0 (PUSH)

It's the origins of web, invented by Tim Berners-Lee and it represented as read only web where there are small amount of producer create web pages (interlinked) and a large number of customers access those web pages through browser via internet. Here user can only read information, user cannot interact with the content of the pages (like comment, answers etc). Technologies used in Web1.0 are HTML, HTTP, URL these are core web protocols, some newer protocol are also in use like XML, XHTML and CSS. In web1.0 both server side and client side scripting are used such as ASP, PHP, JSP, CGI, PERL as server side scripting and JavaScript, VB script, flash as client side.

A. Where Web1.0 Went Wrong

Problems related to web1.0 [1] was its slow and chunky nature and every time when new information entered to the web pages, it needs to be refresh every time.

Web1.0 doesn't support two-way communications [Fig1], it was purely base on client-pull model (HTTP) that can be initiated by client only.

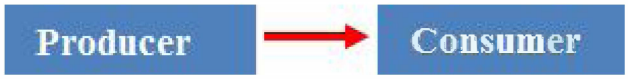

Fig 1: Web1.0 is a one-way platform [3]

Search Technologies used in web1.0 was seen as hopeless, it basically focused on size of the index, ignored the relevance and it cannot find itself.

The most wrong idea behind web1.0 was that it ignores the power of network effects, web1.0 consists of few writers and a large number of readers, and it causes the network slow and makes user starving for resources. If the more people use a networked service, then it becomes more useful for every one using that network, but web1.0 ignored this concept by allowing web1.0 as read only.

It assumes the web as a publishing not as participation, where only information can be read and no interaction can be made with the web pages. It misunderstood the web's dynamics, use software as an application not as a service. Web1.0 relied on old software business model.

III. WEB 2.0 (SHARE)

Web2.0 is known as read-write web [Fig 2]. It is basically a new way to use existing internet technologies. In web2.0 the web user cannot only read the content but also write, modify and update the content online, it supports collaboration and help to gather collective intelligence rather web1.0 [2].

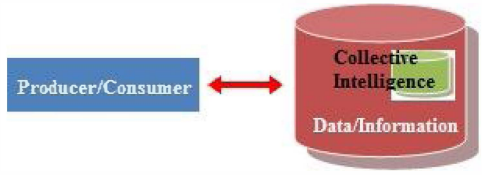

Fig.2 Web2.0 is a two-way Platform

Web2.0 as the next generation of networking services evolved and transferred the network in to a platform by supporting an new idea to exchanges as well as share the content through applications such as wiki, web blogs,widgets and mashups etc.

A.Ideas Behind Web 2.0

Web2.0 come up with six powerful ideas that totaly change the way of interaction among peoples. This ideas can be pointed out as [4].

1)Individual Production and User Generated Content: This idea is concern about the contribution of each single individual to create an usefull information through the use of online technologies such as wikis and blogs.

2)Harness The Power of The Crowd: The idea is concern with the re-use of the collective information or contribution provided by the participants and crowdsourcing.

3)Data on an Epic Scale: This idea related to user contributed data, which can be collected indirectly and aggregated in new ways.

4)Architecture of Participation: Architecture refers a way to design a online technology such that it facilitates the participants and helpful for collaborative knowledge construction.

5)Network Effects: This idea related to the increase in usefulness of a system when more an more user join the system.

6)Openness: This idea is mainly concern with open access,open software and the use and re-use of free data.

B.Technology Behind Web 2.0

Web2.0 still used most of the technologies present behind internet such as XHTML standards,style sheets,content syndication, AJAX and flash etc.

According to wikipedia technology infrastructure of web 2.0 includes.

1)Content Syndication: Content syndication such as RSS,Atom,RDF are used for the creation of web 2.0 services.

2)Ajax-based Internet Technology: Ajax stands for asynchronous JavaScript and XML. It makes web pages more interactive, faster and easier than traditional HTML-based websites by reducing the amount of data needs to be reloaded each time there is a request from the client to the server.

3) *DOM:* Document Object Model(DOM) represent HTML or XML document in atree structure.

4) *REST:* Representational State Transfer(REST) is a approach for getting information content from a web page.

5) *XML and CSS:* XML is basically used to manage information. It construct custom markup languages which can be used to describe any type of data.

Cascading Style Sheet(CSS) is a mechanism for adding style to the web pages.

C. Web2.0 Security Challenges

Today's web2.0 [6] applications are openly accessible and dynamically generated, this feature of web2.0 makes more interesting but it causes bigger security risk.

For example User/Hacker may upload content, which can run code or carry malware to perform some malicious task. Sometimes hacker may upload software like free anti-virus to social sites like facebook (Now-a-days people are too much addicted to facebook or other social networking sites and user are blindly click each and every link and every application and hacker takes the advantage of this stupidity) or any other web sites that is supposed to be virus removal software but that instead load a Trojan horse. Hackers may upload harmful code that could include key loggers that capture victims' keystrokes-including victims' credit card information, password and send them back to hacker.

D . Common Web2.0 Vulnerabilities

Web 2.0 has lots of Vulnerabilities [5][6], some of them are listed below.

1) *Cross Site Scripting*
2) *Cross Site Request Forgery.*
3) *SQL Injection.*
4) *Authentication and Authorisation Flaws.*
5) *Information Laekage.*

Cross Site Scripting (XSS): In XSS attack, hackers inject their own executable code in to legitimate, dynamically generated web pages. Whenever someone visit or download that particular web page, the code that reside inside the webpage execute on the victims' computer and make a way for the hacker to take control over the victims' computer.

Cross Site Request Forgery (CSRF): In a CSRF attack, a hacker uses either the victims' IP address or cookies to gain access or to by-pass firewall protection to an e-commerce, company intranet or other web sites to which the victim has been access/authenticated.

This enables the hacker to act as the computer owner and perform harmful action such as withdraw money from persons bank account, buying stuff from e-commerce sites, stealing data from company intranet or changing configuration on router or local firewall.

SQL Injection: In SQL Injection a hacker insert or “inject” his own SQL query via input data from the client to the application. Attacker can access not only client web app and database; depending on the database an attacker can access operating system also.

In Xpath injection, attacker alters an XML query to achieve his goals. When client supply the input to an websites, Xpath query is created for the XML data, so an attacker can send some data intentionally to find out how the XML data is structured or access data that he may not normally have access .

In JSON injection, attacker inject milicious JavaScript code

in to JSON on the client site, when client login to the target sites (sites that attacker wants to hack) as an authenticated user, attacker send a link to the victims' and convince her to visit that infected sites (sites that is created by the hacker).If the client visit that infected sites while she is already login to the target sites, then all information present in that target sites will be send to the hacker.

XSS Worms: Hacker injects self propagating XSS code in to web pages/app which will spread when users visit that page.

Mashups: The main idea behind mashups is that it combines all the resources or sevices from multiple websites into one single user experience. It can connect dynamically to websites not necessarily under the provider's control, which becomes a security risk for the content providers.

Authentication and Authorisation Flaws: There are various weaknesses in authentication and authorisation [5] technique such as there is no maximum age limit for passwords, sometimes password lengths and complexity are fixed, password can be guess (lack of brute force protection), session ID's can be predictable, most of the times there is no time limits for timeouts and lifetimes for session ID's and only one ID is used for the whole session, character or number used in CAPTCHA systems are broken.

Information Leakage: Sometimes applications unintentionally leak information about their internal working, configuration such as by adding small change on the URL, we can have several information every time.
For example "*http://www.mywebsites.com/home.html*" and instead of "*home.html*" if we try "*admin.html"* (which supposed to be not present in the web site interface) sometimes it may display that particular page if it is present in the server.

E. How to Prevent Web2.0 Vulnerabilities

To reduce vulnerabilities [5][7] we need to consider both technical safeguards and business processes for safe and effective use of web2.0 features.

Here are some of the things [Fig. 3] that may avoid becoming victims of identity theft, fraud and other criminal activities.

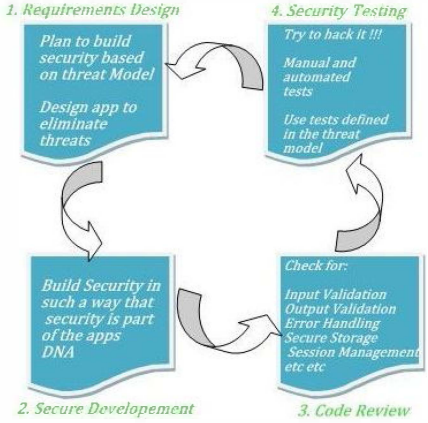

Fig: 3 Steps for Secure Development [5]

1)Validate all Websites: Security Vendors,anti-malware services, new industry standards such as secure socket layer(SSL) certificates, authenticate valid websites and provide a simple visual indicator within the browser navigation bar (riskred, trustgreen etc) will prevent client from hijacking or redirect traffic to bogus websites lurking to inject malware.
2) Risk Reporting Tools: Since criminal activities, fraud, stealing of personal information is increasing day by day, hence every web2.0 user should become aware of this risk and to mitigate these kind of risk web2.0 can provide tools like Google Alerts or credit monitoring services that report on each activities to the user.
3) Validate and Escape all Inputs: All data that has been input or downloaded from web2.0 applications and services should treat as untrusted and validate all downloaded information and applications.
4) Add Authorization: Enterprises should include new security policies and procedures with finer-grained authorization controls to complement authentication and access.

IV. WEB 3.0 (LIVE)

It is more difficult to answer the question "what is web3.0??", different Internet experts has their different approaches and opinions to the future web. Major IT experts consider web3.0[Fig 4] as a semantic web and personalization.
According to *Conrad Wolfram* web3.0 is where computer will generating and thinking new information rather than humans.

Fig.4 Web3.0(Semantic Web)

Google CEO, *Eric Schmidt* says web 3.0 will be "applications which are pieced together – relatively small, the data are in the cloud and it can be run on any device(pc or mobile), very fast, very customizable and distributed virally(social network,email,etc)"

Yahoo founder, *Jerry Yang* thinks that web3.0 is a collection of tools and techniques for creating programs and online application,which blur the distinction between professional, semi-professional and consumers. Yang stated[8].
"..you don't have to be a computer scientist to create a program. We are seeing that manifest in Web2.0 and 3.0 will be a great extension of that, a true communal medium...the distinction between professinal , semi-professional and consumers will get blured,creating a network effect of business and applications" – Jerry Yang[8].

Nova Spivak from Radar Network believes that web3.0 will be "The Semantic web" which will play a central role in the new generation.

A. Key Elements of Web 3.0

Web3.0 is build on the kinds of applications and services that makes it so popular in past few years. Now a days search engines are able to produce much more complete and targeted information, users are even more tightly connected with friends and businesses via social media applications , greater ability to record and store the information makes web3.0[8][11] more precise and helpful for the web users.In this section we present

some of the key elements that might become the building blocks of the next generation of the web.

1) The Social Web: In past several decades social networks has a huge popularity amoung like minded peoples and community groups, they share their feelings,thoughts,ideas using web3.0 technologies instead of linking documents only. Social web are considered as an efficient and attractive way of connecting people around the globe.

2) The Semantic Web: Semantic web is an evolving extension of the web3.0, that allow people to find the information much deeper level the meaning of the search terms and the context in which they are used.The information are structured in such a way that machines can read it and understand it as much as humans can,with out ambiguousness.

3) Web 3D: *In past few years virtual 3D world such as Second life[9],Red Light[10] etc have gained huge popularity in public.Web3D allow people to live in a virtual world as an* avatar on behalf of him and can explore,meet other residents, participate in individual and/or group activities etc etc as people do in their real life.All the activities occuring between avtars are reside in the virtual world only and there is nothing to do with the real life in real time.

4) The Media Centric Web: According to media centric web approach , in near future search engines are able to take media such as audio,video,image etc as an input element and be able to search for similar media objects.

For example if we want to search images about cars ,all we need to provide an car image as an input to the search engine and based on the features present in that car image, engine should able to retrieve images of cars with similar features.

B. Security and Challenges of Web 3.0

In the evolution of the web from web1.0 to web3.0 the various issues related to scalability,security and performance present in web1.0 and web2.0 are also propagate to web3.0 and create a big challenging task for IT expert. Because of the huge collaboration of public and private data make web2.0& web3.0 [12] more interactive and popular among web users and as well as for hackers also.

There is a lack of data standard for controlling over metadata and data privacy.RDF schema(RDFS) and Web Ontology Language(OWL) used URI(Unified Resource Identifiers) to represent data which can be held in database and/or interchanged without specifying any access policy or trust bounderies. It make web3.0 vulnerable ,attackers may falsify the data intentionally and can create false services.

Data privacy in web3.0 is one of most security issue for the IT professional.Producers and customers are creating new contents,techniques day by day and publish it for the world for anyone. They make deals,share their data and ideas among each others. If someone gives you the full control over his private data(like Online-games,) ,considering you as a trusted and capable of good control of his data.What happen when you betrayed him, you have the full control of his data so you can modify and publish it for the world by mistake or intentionally. Illegal and manipulated forms of the same type of data will be available on the web, which may create multiplications of error for anyone.

V. WEB EVOLUTION

A details comparison [11] among web1.0, web2.0 and web3.0 are discussed below in [Table 1].

TABLE 1: Comparison among the Web's

Web1.0	Web2.0	Web3.0
Read-only Static web	Read-write interactive web	Read-write intelligent web
Company-oriented	Community-oriented	Individually oriented
Low-portability (computing equipment)	Medium portability(mobile)	High portability(mobile and consumer electronics)
Professionally developed stand-alone applications	User-developed open applications	User-developed smart applications
Syntax-aware basic browsing and search capabilities	Syntax-aware advanced browsing and search capabilities	Content(semantic)-aware and context-aware next-generation browsing and search capabilities
Low data richness(HTML)	Medium data richness(XML)	High data riclmess(RDF)
Point-to-point/hub & spoke architecture	Service-oriented architecture(SOA)	Web oriented architecture (WOA) andinternet of things
Sliced data	Light interlinked data	Worldwide database

REFERENCES

[1] Dr Mike Evans."The Evolution of the Web-From Web1.0toWeb4.0".

[2] San, Murugesan (2007), "Understanding Web 2.0", Journal IT Professional.

[3] Akhilesh Dwivedi ,SureshKumar, Abhishek Dwivedi Dr. Manjeet Sing "Current Security Considerations for Issues and Challenges of Trustworthy Semantic Web" Int. J. Advanced Networking and Applications Volume: 03, Issue: 01, Pages: 978-983 (2011).

[4] http://en.wikibooks.org/wiki/Web_2.0_and_Emerging_ Learning_Technologies/Learning_Theory#Emerging_Web_2. 0_Related_Learning_Theory.

[5] David Rook "The Security Risks of Web 2.0"DefCon 17, Las Vegas.[6]George Lawton "Web 2.0 Creates Security Challenges" Published by the IEEE Computer Society October 2007.

[7] McAfee "White Paper The Security Implications of Web2.0"http://www.ingrammicro.com/visitor/servicesdivision /McAfee-SaaS-Web-2-0-White-paper.pdf.

[8] Juan M. Silva , Abu Saleh Md. Mahfujur Rahman , Abdulmotaleb El Saddik "Web 3.0: A Vision for Bridging the Gap between Real and Virtual CommunicabilityMS '08 Proceedings of the 1st ACM international workshop on Communicability design and evaluation in cultural and ecological multimedia system.

[9] Second Life Official Website: http://secondlife.com

[10] Red-light Official Website: http://redlightcenter.com.

[11] Karim Sabbagh Olaf Acker Danny Karam Jad Rahban Designing theTranscendent WebThe Power of Web 3.0 Booz & Company.

[12] Malik Muhammad Imran Pattal, LI yuan, ZENG Jianqiu "Web 3.0: A real personal Web! More opportunities & more threats" 2009 Third International Conference on Next Generation Mobile Applications, Services and Technologies, 2009 IEEE.