1

Introduction

Contents

1.1 Introduction

Over the last two decades, the web has become the primary source of information for humans living in developed countries. With the recent efforts to bring high-bandwidth, affordable access internet to the whole world [2–5], everything indicates that in a short period, the first statement will become right for everyone.

A study [6] by the Reuters Institute for the Study of Journalism at the University of Oxford, conducted in 2016, reveals that the primary source of information used by UK citizens is a combination between online (including social networks) and TV. However, if age is taken into account, in the age group 18-44, online becomes the primary source, ranging from 53%, in older people, to 84%, in younger people, while TV drops to between only 9%, in younger people, and 34%, in older people. The study shows a clear trend about the importance online media will have in the coming years. US studies arrived at similar conclusions [7–9].

Parallel to this paradigm change in the consumption of information we have become more and more aware of often lack of credibility, falsehood and incompleteness of that very same information. More often than not this is achieved via a combination of sponsored censorship and media bias.

China is probably the most famous example, exerting censorship on many levels, be that through a massive firewall designed to isolate its population from the outside world [10], to creating legislation that financially incentivises individuals and organisations to self-censorship [11] , to flat-out deletion of posts and news articles [12]. However, this is a global issue [13, 14]. The UK is currently pushing legislation to have more significant powers of censorship over its citisens [15], the US also has a history of state-sponsored censorship. The W. Bush administration censored reports about the presence of weapons of mass destruction in Iraq to justify the invasion and suppressed climate studies analysed the impacts and causes of global warming [16]. The Society of Professional Journalists criticised the Obama administration for blacklisting news outlets which criticised the administration. More recently, the Trump administration has prevented state scientists from speaking publically about climate change [17], and forbid the use of specific abortion-related terms in Central For Disease Control's reports [18, 19].

Multiplying the effect of state censorship is media bias, which is characterised by having journalists' and editors' personal views affecting the way they disseminate information. This can be manifested in many ways [20], by differentiated visibility (more or less than normal) given to selected stories, to the stories that are selected to be reported on or by having the reporters and pundits covering specific issues or people favourably [21]. Looking only at the United State's 2016 election it is easy to find occurrences of all three types [22]. It is important to point out that media bias is not always caused by personal or political reasons, it can also be caused by advertiser pressure [23] - a news organisation may be motivated to not report on stories that are harmful to their advertisers for fearing the lost revenue- or corporate bias - not reporting on a story that can damage the news outlet's parent companies.

The importance of state-sponsored censorship and media bias should not be understated. Democracies are fed by public opinion which in turn is shaped by the information individuals receive. Campaigns of disinformation have the power to start wars [24], influence government elections [25], endanger human lifes [26] and even jeopardise the future of the human species [27].

It is necessary to find a way to provide more credibility to information that is accessible to the public.

One of the approaches to address the problems mentioned above is to fact-check information. Fact-checking organisations [28], such as PolitiFact[1], FactCheck.Org[2] and WikiTribune[3] constantly analyse news articles and issue classifications about the truthfulness of statements. However, these fact-checking organisations are as centralised as regular news outlets, they can suffer from media bias, be censored by governments and even if none of that happens, it will be hard for them to scale at a rate capable of facing the enormous expansion of digital media.

Social media platforms revealed to be useful tools to fight against censorship and state-controlled organisations [29], with several examples of journalists, live reporting on events that governments try to contain, as was the case, in 2016, of the failed attempt to topple the Turkish governement [30]. Unfortunately, just as news outlets and fact-checking organisations, social networks are not immune to state-sponsored censorship. There is evidence of very aggressive efforts by governments (at least 13 countries including China, Egypt, Turkey, United Kingdom, Russia and France) to block access to these platforms or censor user-generated content [31, 32][4].

More recently a new way of interacting with information on the web has surged, which allows for users to make and share annotations of web pages, just like they would annotate a physical notepad. This technology is called web annotations and empowers people to highlight text, create sticky notes or comment specific parts of a web page. Imagine a user visiting a news webpage article which showed portions of the text highlighted by her friends (or anyone chosen by her) and when she places her mouse over the highlighted portions she sees comments made by the users about the highlighted text. What web annotations allow is for the creation of a new layer of information, on top of the existing web resources. This layer can be used to provide context, clarification, additional information about the resource a person is viewing.

Consequently, there is an exciting use case for using web annotations to enhance the information available on the web which has the potential to revolutionise the way people access and consume information. While sharing web pages on social media platforms spreads information, web annotations concentrate the information, which makes it much easier for people to access it. A single web page can be shared 1000 times by 1000 different users on three different social networks. Each user adds some

[1] http://www.politifact.com/
[2] http://factcheck.org/
[3] https://www.wikitribune.com/
[4] These organisations seem to be transparent about the censorship being exerted on them by governments by publishing transparency reports [33–35].

useful information to the page, but all that information is scattered across all these different profiles on different platforms. With web annotations, all the context, clarification and additional information supplied by people are concentrated in one place - the website where that information is relevant.

Currently, there are several web annotation services available to the general public [36], the most widespread one being Genius Web Annotator[5] which allows for the annotation of any webpage and offers a vibrant feature set. The problem with platforms such as Genius, and other similar ones is that they are proprietary and closed, meaning that they are not compatible with each other. An annotation made by a user using service A cannot be used by service B. We call these services silos because the information they hold and services they offer are only useful and relevant to their ecosystem.

A siloed architecture which is characterised by low interoperability, low control of the data by the users, offering no guarantees of permanent storage of information and propensity to censorship, just like the examples listed earlier.

An excellent way to combat siloed platforms is by the creation of clear standards which achieve broad adoption. This moves users to search for services that use these standards. Early in 2017, a W3C Working Group finished their proposal for a Web Annotation standard [37]. The standard includes guidelines for the Web Annotation data model and transport protocol. The standardisation of the data format improves on the previously existing systems by enabling interoperability between platforms. They also take some steps towards decentralisation.

W3C's Web Annotations is a step in the right direction but is still lacking in some aspects since the storage and authentication of data is delegated to a third party - a service provider. That means that when using such a service, users need to place a considerable amount of trust in it. A user needs to trust that the service: correctly authenticates its users; stores the annotations correctly; does not delete annotations; does not tamper with the annotations and shows you all the annotations that have been created, omitting none.

More precisely, the shortcomings of current web annotations platforms are:

- **Weak Verification:** The web annotations are not self-verifiable. This means that for a user to verify their authenticity and integrity, it needs to contact and trust a central server.

- **No Permanence Assurances:** The data is stored in servers controlled by the service provider. As a consequence, a user has no guaranteed that the provider will maintain the file stored. Moreover, data is addressed via HTTP URLs, this means that if a file is moved from one server to another, the link might get broken and access to the file, lost. This also implies that a user can not maintain copies of its data to solve the permanence issue. Because the HTTP link of the file would change, no one would be able to access data.

[5] https://genius.com/web-annotator

- **Not Censorship Resistant:** This comes as a natural consequence of the previous aspects. Since the service provider has full control of the data, it can choose to suppress annotations that are harmful to it.

We argue that the only way to overcome censorship is by eliminating central points of failure where a powerful actor (such as a government, or powerful media organisation) can exert pressure. To fight the problem of state-censorship and media bias we propose a novel system with a decentralised architecture that brings together three key technologies, web annotations, the Ethereum blockchain[6] and IPFS[7] .

As presented earlier, web annotations are a powerful way of bringing an additional layer of information to the web. Offering the possibility for users to annotate websites, making the information much more relevant to a higher number of people. The Web Annotation standard by W3C is a good starting point by defining a data model upon which applications and services can be built, but as noted, has some shortcomings.

We use the Ethereum blockchain to keep a permanent, canonical record of all the annotations made. This blockchain is entirely decentralised and cannot be controlled by any government authority or media outlet. The records that are assured to be fresh and ordered. These properties guarantee that every time a user queries the system, he is receiving the latest information, unfiltered and uncensored.

As will be explained later, storing data on the blockchain is prohibitively expensive. For that reason, we only use Ethereum to save pointers to the data. Now if we chose to store the data in typical cloud storage services such as Google Drive[8] or Dropbox[9] we would be introducing a point of centralisation in the system, precisely what we want to avoid. That is where IPFS enters - a distributed file system where data is stored and scattered all across the web. IPFS has been used as a tool against state censorship by granting access to Wikipedia to Turkish citizens, following the government's blockage of the website [38]. Besides decentralisation, IPFS offers strong integrity assurances of the files it stores. That happens because the link to an IPFS file depends on the content of the file rather than the file's location. Typical HTTP links point to a server on the web that is hosting a given file. The hash of the file partially forms an IPFS link it is addressing. That means that if a file gets tampered with, the link changes. An IPFS link will always point to the same file.

To marry the feature set of web annotations, with the integrity and censorship resistance assurances of IPFS and the ordered registry and freshness of Ethereum, our system stores web annotations on IPFS and records the IPFS links of these files on Ethereum.

Here is an example of how the platform works: A user, Alice, visits the Fox News website and opens an article about global warming. She notices that the chart that plots the rise in temperature in the

[6] https://www.ethereum.org/
[7] https://ipfs.io/
[8] https://www.google.com/drive/
[9] https://www.dropbox.com

last 50 years contains inaccurate information. The makes an annotation on the website, detailing the inaccuracies in the chart. That web annotation gets stored on IPFS, and a registry is made on the Ethereum blockchain. Later that day, Bob visits the same news article and sees Alice's annotation. Bob is assured that the annotation was made by Alice and that it has not been tampered with. Bob is also assured that no annotation is being withheld from him. Now Bob has all the information he needs to form an opinion about the issue being discussed in the article.

Developing a platform with a blockchain based architecture comes with its own set of challenges. The Ethereum blockchain works as a giant computer, a single machine being run by a community of nodes scattered all across the globe. Every node runs all the operations and maintains a copy of the state of the chain. This results in a high cost of computing operations and storage when compared to traditional cloud computation. For that reason, designing this system meant developing a scalable and cost-effective blockchain architecture. We achieved this by the creation of a network of proxies called Publishers.

We present DClaims, a protocol for web annotations that offers the following features:

- **Censorship resistance** The system is built on top of decentralised platforms, namely IPFS and Ethereum, which are beyond the control of state actors and media conglomerates.
- **Authenticity and Integrity Assurances:** The verifiable claims layer allows for a more flexible and complete approach to how annotations' integrity and authenticity is verified. Furthermore, it allows for revocation how annotations.
- **Data Permanence:** The way our network of Publishers is implemented offers permanence assurances of the stored data through an accountability mechanism for bad actors. Adding to that, one of IPFS' properties is permanent links, meaning that a link to a file never gets broken.
- **Financial Cost Efficiency:** Blockchain based applications have two main sources of cost. Storage and transaction fees. Our architecture moves storage out of the blockchain and minimises the number of transactions.
- **Scalability:** The distribution of content is done with a network of Publishers, similar to the one proposed in the original Web Annotation protocol.
- **Future compatibility with standards:** The data model we use can accommodate not only the current version of Web Annotations but is also future proof. To achieve this, we build a layer on top of the traditional protocol, one which has the ability to encapsulate any data format, including web annotations. This is called the verifiable claims layer.

Bibliography

[1] S. Nakamoto, “Bitcoin: A Peer-to-Peer Electronic Cash System,” *self-published paper*, pp. 1–9, Mar. 2009.

[2] (2018) Watch SpaceX Launch the First of Its Global Internet Satellites. [Online]. Available: https://www.wired.com/story/watch-spacex-launch-the-first-of-its-global-internet-satellites/

[3] (2014) **Google reportedly launching 180 satellites for global internet service**. [Online]. Available: https://www.theverge.com/2014/6/2/5771322/google-reportedly-launching-180-satellites-for-worldwide-internet

[4] (2018) Tech Giants Are Battling It Out to Supply Internet Infrastructure. [Online]. Available: http://www.news.com.au/technology/online/tech-giants-are-battling-it-out-to-supply-internet-infrastructure-heres-why-thats-a-problem/news-story/f7626b7cd81fd082f2da0c6820532c9b

[5] (2016) Inside Facebook’s Ambitious Plan to Connect the Whole World. [Online]. Available: https://www.wired.com/2016/01/facebook-zuckerberg-internet-org/

[6] S. Cushion, A. Kilby, R. Thomas, M. Morani, and R. Sambrook. (2016, May) Newspapers, Impartiality and Television News. [Online]. Available: https://medium.com/oxford-university/where-do-people-get-their-news-8e850a0dea03

[7] (2006) Where Do We Get Our News? [Online]. Available: https://www.pbs.org/wgbh/pages/frontline/newswar/part3/stats.html

[8] (2012) **How People Get Local News and Information in Different Communities**. [Online]. Available: http://www.pewinternet.org/2012/09/26/how-people-get-local-news-and-information-in-different-communities/

[9] (2014, Mar.) How Americans get their news - American Press Institute. [Online]. Available: https://www.americanpressinstitute.org/publications/reports/survey-research/how-americans-get-news/

[10] X. Xu, Z. M. Mao, and J. A. Halderman, “Internet Censorship in China - Where Does the Filtering Occur?” *PAM*, vol. 6579, no. Chapter 14, pp. 133–, 2011.

[11] “Freedom House Condemns Latest Act of Media Repression in China,” Jan. 2006.

[12] (2017) China is deleting posts about a kindergarten allegedly abusing its toddlers. [Online]. Available: https://mashable.com/2017/11/30/weibo-social-media-china/#KoT8ru1J2kqU

[13] S. Canaves, “China’s social networking problem,” *IEEE Spectrum*, vol. 48, no. 6, pp. 74–77, 2011.

[14] (2016) **10 Most Censored Countries**. [Online]. Available: https://web.archive.org/web/20180417194953/https://cpj.org/2015/04/10-most-censored-countries.php

[15] (2015) Why Government Censorship [in No Way at All] Carries Greater Risks Than Benefits. [Online]. Available: Whygovernmentcensorship[innowayatall]carriesgreaterrisksthanbenefits

[16] (2017) Trump is copying the Bush censorship playbook. Scientists aren’t standing for it — Dana Nuccitelli. [Online]. Available: https://www.theguardian.com/environment/climate-consensus-97-per-cent/2017/jan/31/trumps-copying-the-bush-censorship-playbook-scientists-arent-standing-for-it

[17] “Donald Trump’s decision to stop US government scientists speaking publicly is ’chilling’,” *The Independent*, Jan. 2017.

[18] “CDC gets list of forbidden words: Fetus, transgender, diversity,” 2017.

[19] (2017) Trump Administration Reportedly Instructs CDC on Its Own Version of 7 Dirty Words. [Online]. Available: https://www.npr.org/2017/12/16/571329234/trump-administration-reportedly-instructs-cdc-on-its-own-version-of-7-dirty-word

[20] J.-M. Eberl, H. G. Boomgaarden, and M. Wagner, “One Bias Fits All? Three Types of Media Bias and Their Effects on Party Preferences,” *Communication Research*, vol. 44, no. 8, pp. 1125–1148, 2017.

[21] K. Lazaridou, R. Krestel, and F. Naumann, “Identifying Media Bias by Analyzing Reported Speech,” in *2017 IEEE International Conference on Data Mining (ICDM)*. IEEE, Oct. 2017, pp. 943–948.

[22] (2018) **Is the Media Biased Toward Clinton or Trump? Here Is Some Actual Hard Data.** [Online]. Available: https://web.archive.org/web/20180105065615/https://www.washingtonpost.com/news/monkey-cage/wp/2016/09/20/is-the-media-biased-toward-clinton-or-trump-heres-some-actual-hard-data/?utm_term=.6c55f7be2bf2

[23] J.-M. Eberl, M. Wagner, and H. G. Boomgaarden, "**Party Advertising in Newspapers**," *Journalism Studies*, vol. 19, no. 6, pp. 782–802, Oct. 2016.

[24] "Official's Key Report On Iraq Is Faulted," 2007.

[25] H. A. Gentzkow and Matthew, "Social Media and Fake News in the 2016 Election," pp. 1–26, Apr. 2017.

[26] "Crisis Pregnancy Centers: Last Week Tonight with John Oliver (HBO)," 2018.

[27] (2018) **Here's What the EPA's Website Looks Like After a Year of Climate Change Censorship**. [Online]. Available: https://web.archive.org/web/20180319165205/http://time.com/5075265/epa-website-climate-change-censorship/

[28] (2016) **Fact-Checking Organizations Around the Globe Embrace Code of Principles**. [Online]. Available: https://web.archive.org/web/20180424193829/https://www.washingtonpost.com/news/fact-checker/wp/2016/09/15/fact-checking-organizations-around-the-globe-embrace-code-of-principles/?utm_term=.f2f14a4e4f2c

[29] "Facebook, Twitter Help the Arab Spring Blossom," 2013.

[30] "At least 256 dead as Turkish president calls coup attempt 'treason'," 2016.

[31] Censorship of Facebook. [Online]. Available: https://web.archive.org/web/20180424201915/https://en.wikipedia.org/wiki/Censorship_of_Facebook

[32] Censorship of Twitter. [Online]. Available: https://web.archive.org/web/20180424201603/https://en.wikipedia.org/wiki/Censorship_of_Twitter

[33] Facebook transparency report. [Online]. Available: https://govtrequests.facebook.com/

[34] Google transparency report. [Online]. Available: https://transparencyreport.google.com/

[35] Twitter transparency report. [Online]. Available: https://transparency.twitter.com/

[36] Making web annotations persistent over time. [Online]. Available: https://web.archive.org/save/https://en.wikipedia.org/wiki/Web_annotation

[37] (2017) Web Annotation Working Group. [Online]. Available: https://web.archive.org/web/20180425000721/https://www.w3.org/annotation/

[38] (2017) Turkey Can't Block This Copy of Wikipedia. [Online]. Available: https://web.archive.org/web/20180425002528/http://observer.com/2017/05/turkey-wikipedia-ipfs/

[39] R. S. Tanash, Z. Chen, T. Thakur, D. S. Wallach, and D. Subramanian, “Known Unknowns,” in *the 14th ACM Workshop*. New York, New York, USA: ACM Press, 2015, pp. 11–20.

[40] J. Bonneau, A. Miller, J. Clark, A. Narayanan, J. A. Kroll, and E. W. Felten, “SoK: Research Perspectives and Challenges for Bitcoin and Cryptocurrencies,” in *2015 IEEE Symposium on Security and Privacy (SP)*. IEEE, 2015, pp. 104–121.

[41] S. Barber, X. Boyen, E. Shi, and E. Uzun, “Bitter to Better - How to Make Bitcoin a Better Currency.” *Financial Cryptography*, vol. 7397, no. Chapter 29, pp. 399–414, 2012.

[42] A. Gervais, G. O. Karame, K. Wüst, V. Glykantzis, H. Ritzdorf, and S. Capkun, “On the Security and Performance of Proof of Work Blockchains,” in *the 2016 ACM SIGSAC Conference*. New York, New York, USA: ACM Press, 2016, pp. 3–16.

[43] S. Dziembowski, S. Faust, V. Kolmogorov, and K. Pietrzak, “Proofs of Space,” in *Advances in Cryptology – CRYPTO 2015*. Berlin, Heidelberg: Springer, Berlin, Heidelberg, Aug. 2015, pp. 585–605.

[44] M. Ali, J. Nelson, R. Shea, and M. J. Freedman, “Blockstack: A global naming and storage system secured by blockchains,” *2016 USENIX Annual Technical*, 2016.

[45] S. King and S. Nadal, “Ppcoin: Peer-to-peer crypto-currency with proof-of-stake,” *self-published paper*, 2012.

[46] E. Wustrow and B. VanderSloot, “DDoSCoin: cryptocurrency with a malicious proof-of-work,” *of the 10th USENIX Conference on*, 2016.

[47] M. Ghosh, M. Richardson, B. Ford, and R. Jansen, “A TorPath to TorCoin: Proof-of-Bandwidth Altcoins for Compensating Relays,” *Yale University*, Jul. 2014.

[48] M. Milutinovic, W. He, H. Wu, and M. Kanwal, “Proof of Luck,” in *the 1st Workshop*. New York, New York, USA: ACM Press, 2016, pp. 1–6.

[49] S. Bowe and et al., “Zcash Protocol Specification,” *self-published paper*, pp. 1–53, Mar. 2017.

[50] A. Kosba, A. Miller, E. Shi, Z. Wen, and C. Papamanthou, “Hawk: The Blockchain Model of Cryptography and Privacy-Preserving Smart Contracts,” in *2016 IEEE Symposium on Security and Privacy (SP)*. IEEE, Apr. 2016, pp. 839–858.

[51] G. Wood, “Ethereum Yellow Paper,” *self-published paper*, 2014.

[52] V. Buterin, “Ethereum white paper,” *self-published paper*, 2013.

[53] J. Benet, “IPFS-content addressed, versioned, P2P file system,” *arXiv.org*, 2014.

[54] R. C. Merkle, “A Digital Signature Based on a Conventional Encryption Function,” in *Advances in Cryptology — CRYPTO ’87*. Berlin, Heidelberg: Springer, Berlin, Heidelberg, Aug. 1987, pp. 369–378.

[55] I. Marty, I. Miers, and E. Wustrow, “Proof-of-Censorship: Enabling centralized censorship-resistant content providers,” *Financial Cryptography and Data Security*, pp. 1–18, Jan. 2018.