

Sussex Research

Fake news: a technological approach to proving the origins of content, using blockchains

Steve Huckle, Martin White

Publication date

01-12-2017

Licence

This work is made available under the [CC BY 4.0](#) licence and should only be used in accordance with that licence. For more information on the specific terms, consult the repository record for this item.

Document Version

Published version

Citation for this work (American Psychological Association 7th edition)

Huckle, S., & White, M. (2017). *Fake news: a technological approach to proving the origins of content, using blockchains* (Version 1). University of Sussex. <https://hdl.handle.net/10779/uos.23450429.v1>

Published in

Big Data

Link to external publisher version

<https://doi.org/10.1089/big.2017.0071>

Copyright and reuse:

This work was downloaded from Sussex Research Open (SRO). This document is made available in line with publisher policy and may differ from the published version. Please cite the published version where possible. Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners unless otherwise stated. For more information on this work, SRO or to report an issue, you can contact the repository administrators at sro@sussex.ac.uk. Discover more of the University's research at <https://sussex.figshare.com/>

BIG DATA

Big Data

Fake News - a Technological Approach to Proving the Origins of Content, Using Blockchains

| | |
|-------------------------------------|--|
| Journal: | <i>Big Data</i> |
| Manuscript ID | BIG-2017-0071.R2 |
| Manuscript Type: | Original Article |
| Date Submitted by the Author: | 31-Oct-2017 |
| Complete List of Authors: | Huckle, Steve; University of Sussex, School of Engineering and Informatics White, Martin; University of Sussex, School of Engineering and Informatics |
| Keywords: | Big data analytics, Big data infrastructure design, Data mining, data protection, privacy, and policy, Semi-structured data |
| Manuscript Keywords (Search Terms): | Fake News, Ethereum, Blockchain, Cryptography, PREMIS, Hash functions |
| | |

SCHOLARONE™
Manuscripts

Fake News - a Technological Approach to Proving the Origins of Content, Using Blockchains

Steve Huckle ^{1*}, Martin White ¹

1 University of Sussex, Sussex House, Falmer, Brighton, BN1 9RH, United Kingdom

* Direct comments to: s.huckle@sussex.ac.uk; Tel.: +44 (0)1273 606755

Abstract

In this paper, we introduce a prototype of an innovative technology for proving the origins of captured digital media. In an era of fake news, when someone shows us a video or picture of some event, how can we trust its authenticity? It seems the public no longer believe that traditional media is a reliable reference of fact, perhaps due, in part, to the onset of many diverse sources of conflicting information, via social media. Indeed, the issue of 'fake' reached a crescendo during the 2016 US Presidential Election, when the winner, Donald Trump, claimed that the New York Times was trying to discredit him by pushing disinformation. Current research into overcoming the problem of fake news does not focus on establishing the ownership of media resources used in such stories - the blockchain-based application introduced in this article is technology that is capable of indicating the authenticity of digital media. Put simply; by using the trust mechanisms of blockchain technology, the tool can show, beyond doubt, the provenance of any source of digital media, including images used out of context in attempts to mislead. Although the application is an early prototype and its capability to find fake resources is

1
2
3 somewhat limited, we outline future improvements that would overcome such
4
5 limitations. Furthermore, we believe our application (and its use of blockchain
6
7 technology and standardised metadata), introduces a novel approach to overcoming
8
9 falsities in news reporting and the provenance of media resources used therein.
10
11 However, while our application has the potential to be able to verify the originality
12
13 of media resources, we believe technology is only capable of providing a partial
14
15 solution to fake news. That is because it is incapable of proving the authenticity of a
16
17 news story as a whole. We believe that takes human skills.
18
19
20
21

22
23 **Keywords:** Fake News, blockchain, big data, Ethereum, hash functions,
24
25 cryptography, public-key cryptography, digital signatures, Preservation Metadata
26
27

28 29 **Introduction**

30
31 The issue of fake news hit the headlines when Donald Trump, the winner of the
32
33 2016 U.S. Presidential Election, accused various media outlets of mounting a
34
35 concerted effort to discredit him¹ by publishing hoaxes and propaganda². Even
36
37 before the President's accusations, one of the implicated newspapers, The New York
38
39 Times, printed a story asserting that one of Trump's prominent supporters was
40
41 spreading disinformation³. After, presumably, much journalistic investigation, the
42
43 newspaper claimed falsehood by showing that a photograph (illustrated in Figure
44
45 1), which was used on the Christian Times website to suggest that the US President's
46
47 opponents were rigging votes, was, in fact, a picture from the UK's Birmingham Mail.
48
49 The picture showed ballot boxes used in a UK election, not fraudulent Clinton votes
50
51 found in an Ohio Warehouse, as the website claimed. However, what if such
52
53
54
55
56
57
58
59
60

1
2
3 detective work was unnecessary? What if it were trivial to ascertain the provenance
4 of a picture or video? Not only could we trust that material, but we could distrust
5 of a picture or video? Not only could we trust that material, but we could distrust
6 any material that was not validated that way.
7
8
9



36 *Figure 1. Birmingham Mail picture of the delivery of ballot boxes used in a UK election.*
37 *The picture was misappropriated by a Trump supporter, who (falsely) claimed that the*
38 *image showed that the Clinton campaign team were rigging votes⁴*
39

40 The primary aim of this article is to introduce a blockchain-based distributed
41 application that we are calling Provenator (intended as the agent noun of the verb
42 form of provenance, which means establishing the origin of something), a tool that
43 helps prove the originator of media sources. Before describing Provenator, we
44 provide some background by introducing the motivation for this work - fake news.
45 Then we present big data's role in technological attempts to counter false reporting.
46 Next, we describe the technologies underlying Provenator - blockchains and a data
47 schema for recording metadata about media resources. Then we discuss Provenator
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 in detail, including its use, current limitations and future improvements that might
4
5 address those limitations, before concluding.
6
7

8 9 **Fake News**

10
11 Fake News is, quite simply, invented information⁵. Unfortunately, it is often difficult
12
13 to spot invented from real. For instance, in a recent survey, when the UK's Channel 4
14
15 News showed three real and three fake stories to 1,684 adults, only 4% of the
16
17 respondents were able to identify all the stories correctly, and nearly half thought
18
19 that at least one of the fakes was real⁶.
20
21
22

23
24 While the Channel 4 survey may not appear, at first glance, to raise a major issue, a
25
26 somewhat more nuanced interpretation of fake news is that they are stories that are
27
28 distorted or decontextualised and deliberately designed to deceive. Often, such
29
30 stories have an undeclared political bias⁵. Thus, fake news is a synonym for
31
32 propaganda, a term which has sinister connotations. As an example, during the
33
34 recent annexation of the Crimea, NATO accused Russia of using fake news to spread
35
36 disinformation about their actions there⁷. Moreover, in a follow-up to their survey,
37
38 Channel 4 ran a news series on fake news, in which they interviewed Janis Sartis, the
39
40 Director of the NATO Strategic Communications Centre. During the interview, Sartis
41
42 said: "You don't need tanks. You might actually achieve your goals if you change the
43
44 perception of a given society in a way that corresponds to your interests and the
45
46 society starts to act how you want them to act"⁸.
47
48
49
50
51

52
53 Social media companies have come under political pressure for not providing tools
54
55 to counter the problem of fake news. Consequently, politicians have accused those
56
57
58
59
60

1
2
3 companies of having an undue influence on elections both in the UK and U.S.⁹.
4
5 Indeed, analysis has shown that, during the final three months of the U.S.
6
7 presidential campaign, Facebook's fake news stories about the U.S. presidential
8
9 election generated much more interest than stories from traditional news outlets¹⁰.
10
11 Indeed, Facebook admitted that: "more and more...debate is mirrored online on
12
13 platforms like Facebook, leading to an increase in individual access and agency in
14
15 political dialogue...as well as the diversity of influences on any given
16
17 conversation"¹¹. To counter this issue, Facebook placed advertisements in UK
18
19 newspapers, giving tips to its users on how to spot fake news items¹². The company
20
21 also implemented several design features on its platform's user interface; measures
22
23 included stronger automated detection of fakes, convenient user reporting of
24
25 suspicious content and third-party verification of news items¹³. The founder of
26
27 Wikipedia, James Wales, also announced a new initiative for countering fake news¹⁴.
28
29
30 The criticisms of social media platforms and fake news suggest that the issue is a
31
32 new phenomenon. However, propaganda has a long history.
33
34
35
36
37
38
39
40

41 **A Brief History of Fake News**

42
43 During a recent TED talk, Yuval Noah Harari said: "I think fake news has been with
44
45 us a long time; just think of the Bible!"¹⁵. Indeed, the earliest example of propaganda
46
47 is considered to be the Behistun Inscription, authored around 515 BC, which is an
48
49 inscription in three different cuniform dialects on a cliff at Mount Behistun in
50
51 Kermanshah Province, Western Iran. It details the rise to the throne of the Persian
52
53 Empire of Darius I and his success in quelling multiple rebellions¹⁶. However, Pope
54
55
56
57
58
59
60

1
2
3 Gregory XV was the first to use the term 'propaganda', when in 1622, he formed the
4
5 'Congregatio de Propaganda Fide', or "congregation for propagating the faith." The
6
7 word itself comes from the Latin word 'propagare', meaning propagation. Hence,
8
9 propaganda is understood to mean the propagation of an ideology¹⁷.
10
11

12
13 A more modern example of propaganda, yet still one-hundred years old, was
14
15 described by Dr David Clarke in a recent piece for the BBC¹⁸. Dr Clarke tells how, in
16
17 1917, the British Government, in an ultimately successful attempt to bring China
18
19 onto the Allied side in The Great War, fabricated a gruesome story about the German
20
21 military, whom they (falsely) accused of extracting glycerine from human corpses.
22
23 Apparently, Conservative MP John Charteris, Head of Intelligence at the time of the
24
25 story's fabrication, transposed captions from a photograph that showed a train of
26
27 dead horses that were to be rendered onto another showing a train taking dead
28
29 soldiers for burial. Unfortunately, the story was later used by the Nazi Party as proof
30
31 of British lies during the Great War, and it may have led to doubts about news of
32
33 Nazi atrocities during the Second World War; as Dr Clarke comments: "lies have
34
35 consequences"¹⁸. The Nazi Party, realising the importance of war propaganda,
36
37 formed the Reich Ministry of Public Enlightenment and Propaganda. The Ministry's
38
39 head, Joseph Goebbels, used his control of the press to help reinforce Nazi ideology
40
41 through fake news: "If you tell the same lie enough times, people will believe it; and
42
43 the bigger the lie, the better"¹⁹.
44
45
46
47
48
49
50

51
52 Much like Nazi Germany, Stalinist Russia, in an attempt to convince its people that
53
54 the Soviet Union enjoyed much higher living standards than those in the Capitalist
55
56 West, used propaganda extensively²⁰. During the lead-up to the Second World War,
57
58
59
60

1
2
3 the Soviet media suppressed heretical opinion through the censorship of dissonant
4 voices. Newspaper headlines took a standard form: “all workers greeted the policy
5 (of the Russian Government) with satisfaction.” They repeated the message often,
6 giving credence to Goebbels’ mantra that if you tell a lie often enough, people will
7 believe it. Soviet propaganda continued after the war too, with books heavily
8 censored and newspapers propagating ideolised reality²¹. Television and radio gave
9 that reality a degree of formality. Meanwhile, cinematography took a triumphalist
10 tone, depicting happy lives and the fulfilment of the ‘Soviet dream’²¹.

11
12
13 Despite increasing press freedom in the 1990s, following Glasnost (a Soviet policy of
14 open discussion of political and social issues), Russian authorities appear to
15 continue propagating fake news stories. Indeed, on February 22nd, 2017, the
16 Russian Minister of Defence, Sergei Shoigu, admitted that four years prior, the
17 Russian Government had established ‘Voyska Informatsionnykh Operatsiy,’ a
18 dedicated information warfare force, because: “Our propaganda needs to be clever,
19 smart and efficient”²². For instance, they may deliberately take images out of context
20 so that they support the state narrative²³. For example, to refute the Western
21 narrative that the passenger aircraft MH17 was shot down by Russian-backed
22 Ukrainian Separatists, Russian state television has reported on an ariel photograph
23 of a jet fighter firing a missile at the downed plane. However, an organisation called
24 StopFake has gone to great lengths to debunk the picture, citing evidence such as
25 the incorrect placement of the Malaysian Airlines logo and the lack of aircraft
26 vapour trails²³.

1
2
3 The Russian State has not been the sole purveyors of fake news in the modern
4
5 world. In 1928, Cornell Graduate Edward Bernays published a book called
6
7 *Propaganda*, which has become, essentially, a manual of mass manipulation²⁴. The
8
9 book opens with the following paragraph: “The conscious and intelligent
10
11 manipulation of the organized habits and opinions of the masses is an important
12
13 element in a democratic society. Those who manipulate this unseen mechanism of
14
15 society constitute an invisible government which is the true ruling power of our
16
17 country.” In fact, before the First World War, the term propaganda was not used
18
19 negatively, but the public began to mistrust the term once they realised the extent to
20
21 which the Anglo-American political machinery had deployed propaganda in an
22
23 attempt to demonise “The Hun”²⁴. Its use by the Nazi Party in the Second World
24
25 War²⁵, and later by Communist Russia, appears to have sealed the term’s fate; now
26
27 propaganda has extremely negative connotations. However, that does not mean that
28
29 its use in the West has diminished. Immediately after the war, U.S. President
30
31 Truman instigated NSC/10, a policy to contain the Soviet state using wide-ranging
32
33 covert operations, including propaganda²⁶. During the 1960s and 1970s, the media
34
35 corporations of Western nations were instrumental in promoting neo-colonialism
36
37 (the practice of exerting influence or control over less developed countries by using
38
39 trade policies, economic or financial means) and incapacitating attempts at self-
40
41 determination by third world countries²⁵. There are recent examples of Western
42
43 propaganda too; in 2005, the U.S. Government tried to sway public opinion as to the
44
45 benefits of the Iraq War by spending US\$300 million on an initiative to propagate
46
47 ‘positive news’²⁷.
48
49
50
51
52
53
54
55
56
57
58
59
60

Democracy and the Free Press

Perhaps the most famous example of fake news from the literary fiction is George Orwell's *Nineteen Eighty-Four*²⁸. The book depicts the Inner Party, a tyrannical organisation who govern a Super State. One of the novel's main themes is censorship through the Inner Party's modification of records, such as photographs. The protagonist is Winston Smith, who works for the Ministry of Truth; it is his job to re-write past newspaper articles and thereby distort records so that they correspond to the party's propoganda. By depicting a state that enforces suppression through historical revisionism, Orwell demonstrates that press freedom is core to the healthy functioning of a democratic nation. Undoubtedly, a free press plays a pivotal role in a democracy's political culture because it relies upon a "healthy and vibrant" media system, who keep its citizens adequately informed²⁵. Indeed, the media's ownership, management and funding directly affect its capacity to serve the democratic process²⁵.

The President of the United States is the 'Leader of the Free World'. The 'Free World' includes nations who espouse certain freedoms, such as those based on a free press, and it is formed primarily by the countries who opposed both Fascism in the Second World War and Soviet Communism during the Cold War. Hence, the U.S. is at the zenith of all the supposed free, democratic nations. The First Amendment to the U.S. Constitution guarantees individual and press freedom by prohibiting government from impinging on those freedoms: "Congress shall make no law respecting an establishment of religion, or prohibiting the free exercise thereof; or abridging the freedom of speech, or of the press; or the right of the people peaceably to assemble,

1
2
3 and to petition the government for a redress of grievances”²⁹. It is, then, somewhat
4
5 disconcerting when the US President starts to undermine the free press by accusing
6
7 them of spreading fake news. Coleen Christie, the host of Canada’s CTV News,
8
9 believes that the President’s fake news accusation is merely a symptom of the
10
11 explosion of digital media, which has changed our legacy news platforms and
12
13 undermined our trust in such platforms³⁰. Indeed, she warns that: “in this modern
14
15 news age, information is power, yet never has our ability to leverage that power
16
17 been more at risk.” As we have already seen, social media outlets are coming under
18
19 increasing political pressure to ensure the integrity of the items published on their
20
21 platforms, so they have started to implement measures to help counter the
22
23 phenomenon. Could new digital technologies help rather than hinder? Might
24
25 blockchains provide methods for circumventing the issue of fake news by
26
27 establishing the credentials (or not) of media resources used in such stories? Much
28
29 of the rest of this paper discusses such a possibility. However, first, we describe
30
31 some ongoing research into big data and fake news.
32
33
34
35
36
37
38
39

40 **Big Data and Fake News**

41
42 Big data refers not just to the large quantities of digital data, but also to the quality
43
44 of the data and the relationships formed³¹. In other words, big data is networked,
45
46 and recognising patterns therein creates value. Unfortunately, as we have shown
47
48 above, the data may not always reflect the truth³². Hence, even if big data has the
49
50 potential to transform our understanding of world events³¹, there are dangers
51
52 presented by inaccuracies and/or (deliberate) falsities³³. Indeed, news, in its purest
53
54 sense, is meant to convey truthful, unbiased and informative facts about issues
55
56
57
58
59
60

1
2
3 affecting the world. Hence, gathering reliable information is an important part of a
4
5 journalist's skills³⁴; they must take a critical perspective on all information collected
6
7 because their stories must stand up to later scrutiny.
8
9

10
11 Library and information science is adapting to the challenges of big data news
12
13 streams, by attempting to use automated methods for analysing text and verifying
14
15 online information³³: "separating the news from the noise is key to the verification
16
17 of digital information"³⁵. We take a look at some such initiatives next.
18
19
20

21 **Fake News Detection Technologies**

22
23 City University has instigated a project, sponsored by Google, with the goal of
24
25 helping journalists identify fake news by analysing relationships in large, complex
26
27 news-based datasets³⁵. City is developing a web-based tool that combines machine
28
29 learning and artificial intelligence technologies to visualise those relationships³⁵.
30
31 They are aiming to test their product with European-based news organisations, such
32
33 as the UK's Telegraph media group and the Guardian, as well as Ireland's national
34
35 broadcaster, RTE.
36
37
38
39

40
41 As we have already shown, nowadays, users don't get their news solely from
42
43 traditional print and broadcast media; they also get it from social media sources.
44
45 Hence, both Narwal et al.³⁶ and Jin et al.³⁷ focus their attention on overcoming fake
46
47 news on platforms such as Twitter. Jin et al. describe a tool that analyses messages
48
49 and creates a hierarchical graph optimisation of the relationship between news
50
51 events. By so doing, their application propagates the credibility of those events³⁷.
52
53
54
55 Narwat et al. have developed a tool called UnbiasedCrowd, whose purpose is to,
56
57
58
59
60

1
2
3 first, identify bias, second, identify images that are used out of context to support a
4 particular opinion, and third, create a call to action, whereby activists are urged to
5 expose the inherent bias³⁶.
6
7
8
9

10
11 The application developed in this article, Provenator, stores provenance metadata
12 on a blockchain, thus enabling content creators to prove, unequivocally, the origins
13 of their media resources. Because of the properties of blockchains (which we will
14 describe later), that also means users can trust the authenticity of the metadata
15 about those resources. Additionally, Provenator provides an interface whereby
16 users can check the provenance of media resources used in news stories. However,
17 that supposes that Provenator was used to document the resource in the first place;
18 in reality, that functionality is only useful given wide-scale deployment of our
19 application. Of course, since we are at the prototype stage, that has yet to happen.
20
21 However, such wide-scale use is possible, so later in the article, when we describe
22 such a scenario, we feel justified in doing so.
23
24
25
26
27
28
29
30
31
32
33
34
35
36

37
38 Before we detail Provenator itself, we first describe the technologies it uses to help
39 facilitate data integrity and authenticity.
40
41
42

43 **Methods for Trust and Authenticity**

44
45 As we have already discussed, it is crucial that reporters trust the integrity and
46 authenticity of the media resources contained within their news stories. For
47 example, suppose Alice is the Birmingham Mail Photographer who was responsible
48 for the picture of the UK ballot boxes, which we discussed in the introduction.
49
50
51
52

53
54
55 Imagine that Bob is her Picture Editor, who must be satisfied with the image's
56
57
58
59
60

1
2
3 integrity and authenticity. For instance, he has to be sure that, without Alice's
4
5 knowledge, someone has not swapped the picture for another (or that any
6
7 modifications have a verifiable provenance trail). We will show, to achieve such
8
9 confidence, Bob requires methods from the field of cryptography.
10
11

12 13 14 **Cryptography**

15
16 Cryptography is the mathematics of information security³⁸, a field of study that
17
18 investigates the confidentiality, integrity, authenticity and non-repudiation of data³⁹.
19
20 Next, we describe some tools that apply techniques from cryptography; namely,
21
22 public-key cryptography, cryptographic hash functions and digital signatures.
23
24

25 26 27 **Public-key Cryptography**

28
29 Data encryption is a process that produces ciphertext by combining some original
30
31 text (to be kept secret, for whatever reason), with a much shorter key. Later, it is
32
33 possible to use the key to transform the ciphertext back into the original text, a
34
35 process known as decryption⁴⁰.
36
37

38
39 Public-key cryptography (PKC) is a particular form of encryption that uses a pair of
40
41 asymmetric keys; a private key that is known only to the owner and a public key
42
43 that is widely shared³⁸. The basic idea is that encryption is achieved using the public
44
45 key and decryption using the private key³⁹. Figure 2 shows how Alice could use
46
47 public-key cryptography to send a secure message to Bob about her picture; she
48
49 uses Bob's public key to encrypt the message, and subsequently, only Bob can
50
51 decrypt Alice's message since he is the only person who has the paired private key.
52
53
54
55
56
57
58
59
60

Thus, the security of public-key cryptography systems relies upon the secrecy of the private key.

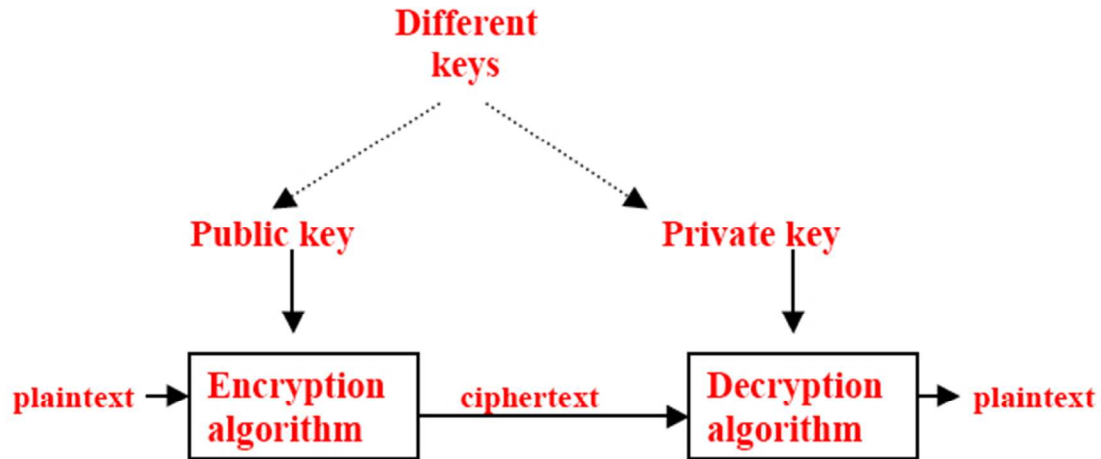


Figure 2. Public Key Cryptography³⁹

Figure 3 shows the process Bob uses to generate his private and public keys; he feeds a random number into a key generation program, from which it produces the required keys.

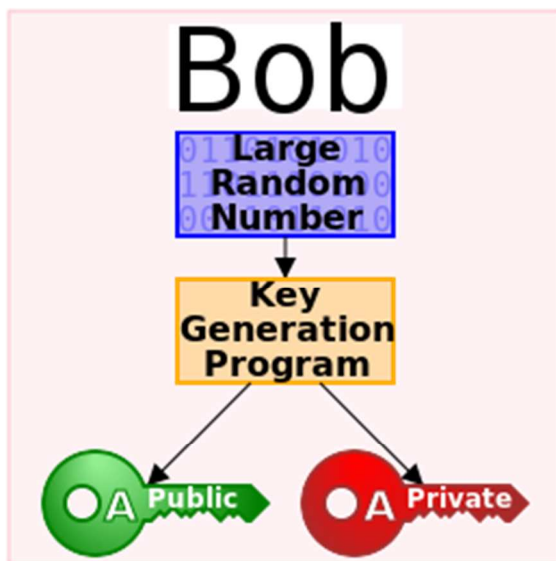


Figure 3. Key Generation⁴¹

1
2
3 In PKC systems, it is trivial (computationally) to generate public and private keys,
4
5 but once the public key is known, it is infeasible to find the private key. That is a
6
7 result of a class of mathematical problems that have no efficient solution. One such
8
9 problem is the discrete logarithm, which uses the modular exponentiation of large
10
11 prime numbers that are easy to compute, but practically impossible to invert³⁹.
12
13
14
15

16 **Cryptographic Hash Functions**

17
18 When Alice sends her photograph, Bob must be satisfied that, while in transit, it has
19
20 remained unaltered. Cryptographic hash functions can help there. The basic idea is
21
22 that Alice computes a cryptographic hash of the picture, which she then sends to
23
24 Bob alongside the image itself. Bob then calculates the cryptographic hash value of
25
26 the received photo and checks that the hash matches the value Alice sent.
27
28
29
30

31 A cryptographic hash is a one-way function that maps arbitrary data to a fixed-size
32
33 string. They are mathematical algorithms that are infeasible to invert (much like
34
35 their public-key cryptography counterparts). The ideal cryptographic hash function
36
37 has five main properties:
38
39

- 40 1 Deterministic - the same message results in the same hash.
- 41 2 Fast - for any message, it is quick to calculate the hash.
- 42 3 One-way - it is practically impossible to generate the message from its hash.
- 43 4 No correlation - a small change to a message will drastically modify the hash.
- 44 5 Collision resistance - it is computationally infeasible to find any two distinct
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60 inputs, M and M^* , which hash to the same value³⁹.

Figure 4 shows a hash function that converts an arbitrary length block of data into a unique, fixed-length, 'hash-value' that serves as a compact representation of the original data³⁹.

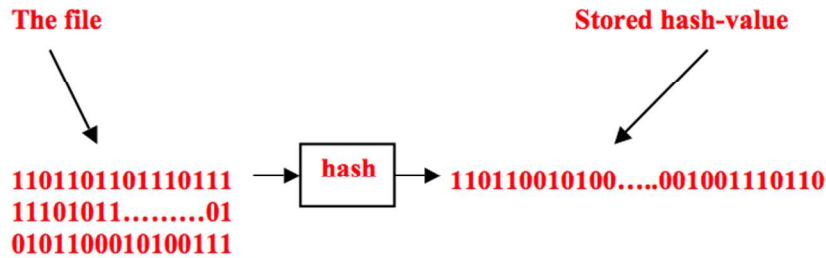


Figure 4. A Hash Function³⁹

Figure 5 shows that, after receiving Alice's photograph, the hash Bob computes must be unique to a given input⁴². In other words, if the hash is the same as the original, then Alice's image must have remained unaltered.

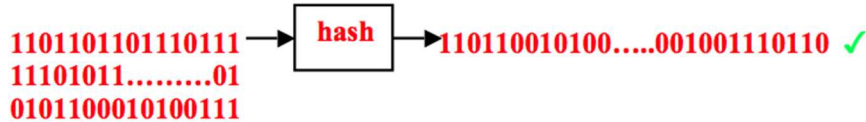


Figure 5. The Validated Hash³⁹

Similarly, Figure 6 shows that if the hash generated by Bob does not match that sent by Alice, then the picture must have been modified.

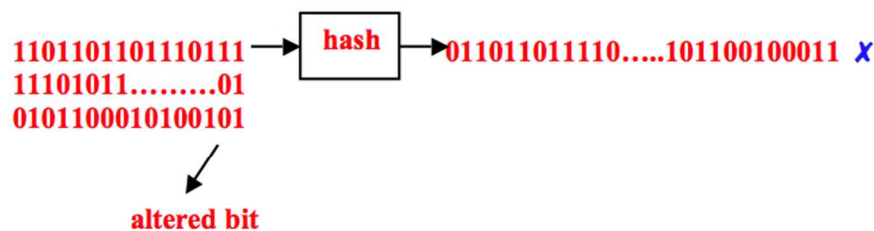


Figure 6. An Altered File³⁹

An example hash function is SHA-256, which produces many fixed-size 256-bit (32-byte) hashes. For all practical purposes, finding collisions is beyond the capabilities

of present-day computing. It is an iterated hash function, a process shown in Figure 7; its design ensures the use of all message bits in the final hash value H_k . It works by splitting the input into a sequence of fixed-size blocks $M_1, M_2, M_3, \dots, M_k$, with some padding rule for the last block M_k . Input blocks are processed in order, using a one-way compression function that gives a set of intermediate hash values $H_0, H_1, H_2, \dots, H_k$. H_0 is a predefined initialising value, and H_k is the hash value output of the SHA-256 function.

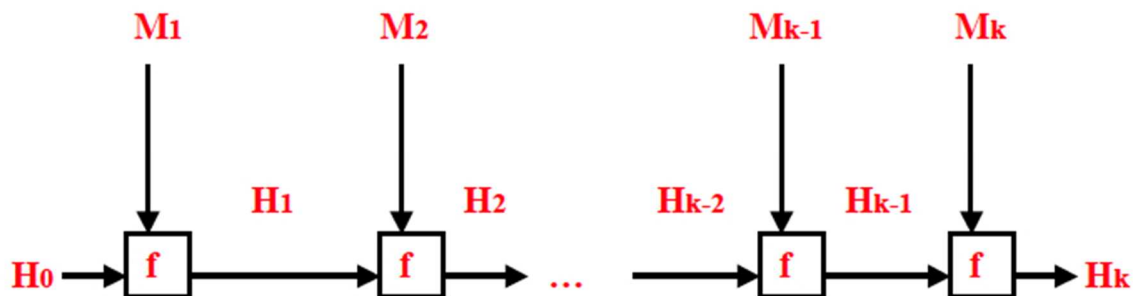


Figure 7. An Iterated Hash Function³⁹

Earlier, while giving an overview of hashing functions, we showed that a computed hash must match that of the origin. However, that raises the problem of ensuring the validity of the original hash. In other words, Bob may question whether it was Alice who sent the hash of the picture in the first place. Digital signatures can help there.

We discuss those next.

Digital Signatures

From an early age, we learn the importance of a written signature as it serves to identify, authorise and validate³⁸. In the electronic world, it is trivial to append to a document a signature that does not belong to the originator, so cryptography has developed advanced digital signature techniques that would allow Alice to bind her

identity to her photograph. The process would involve Alice executing a transform so that the final message she sends to Bob combines the original image together with some secret information held only by Alice³⁸.

An overview of the digital signature process is shown in Figure 8. To allow Alice to share information with Bob (in a manner that guarantees the data's authenticity), she creates a signature that Bob can use to validate her message. Moreover, Alice would be unable to deny that it was she who shared the information, due to the non-repudiation properties of digital signatures.

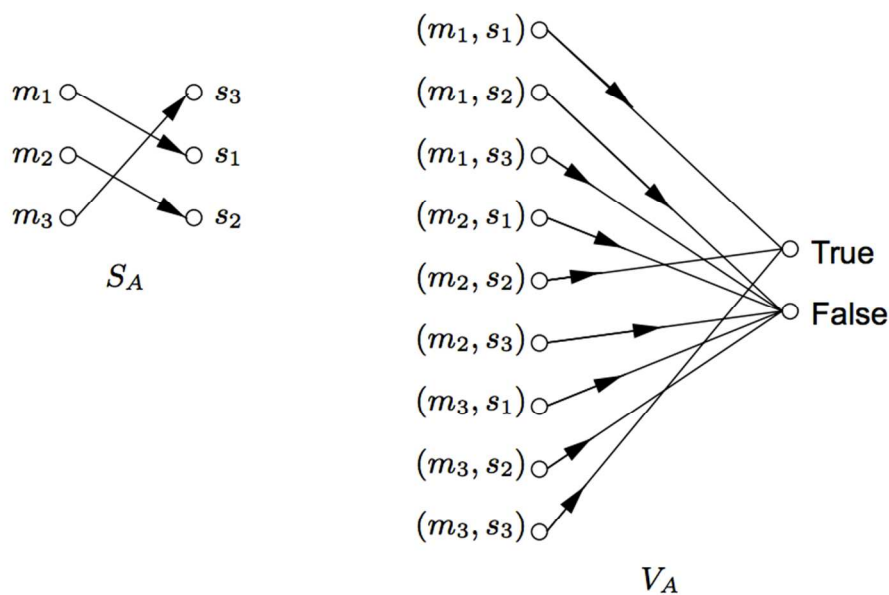


Figure 8. The Digital Signature Process.

- M - the set of messages to be signed by Alice.
- S - a collection of Alice's signatures.
- S_A - a secret signing transformation that will be used by Alice to create signatures from messages M .
- V_A - a verification transformation, from the set $M \times S$ to the set $\{true, false\}$, for Alice's signatures. V_A is publicly known, so Bob can use it to verify signatures created by Alice, thereby authenticating the messages they share³⁸.

1
2
3 A typical usage of a digital signature is to sign a cryptographic hash of a message
4
5 (the information that must be signed)³⁸, using the signees private key⁴³. The
6
7
8 signature then takes the form of a number that proves the signing operation took
9
10
11 place.

14 **Technologies Used by Provenator to Prove Authenticity**

16 The application we are about to describe, Provenator, uses technologies that employ
17
18 methods from cryptography to help determine the authenticity of media resources.
19
20 Additionally, it uses a schema to record and retrieve metadata describing those
21
22
23 media resources. We describe those technologies next.

27 **Blockchains**

29 Blockchains have capabilities resulting in their suitability for determining integrity
30
31 and authenticity because they are, essentially, an immutable database technology⁴⁴
32
33 with inbuilt trust mechanisms⁴⁵. They include cryptographic algorithms and digital
34
35 signatures that allow secure electronic collaboration, without requiring any
36
37
38 centralised authority⁴⁶. Blockchains also have the ability to execute smart contracts,
39
40
41 which are verifiable scripts that automate a system's rule set⁴⁷. In essence, then,
42
43
44 blockchains are a trusted ledger capable of running application logic⁴⁷. Furthermore,
45
46 they cannot be controlled by any single entity⁴⁸. Those mechanisms mean we can
47
48
49 use a blockchain to record data about our media resources and any entity that views
50
51 those records will be satisfied that the information conveyed is authentic. However,
52
53
54 we still require an appropriate schema for recording data on the blockchain. We
55
56
57 discuss that next.

Provenance Metadata

PREMIS stands for “Preservation Metadata: Implementation Strategies”; it outlines a provenance schema which helps identify a resource⁴⁹. The PREMIS data model⁵⁰, shown in Figure 9, describes four separate preservation entities: 1) Objects, 2) Events, 3) Agents and 4) Rights.

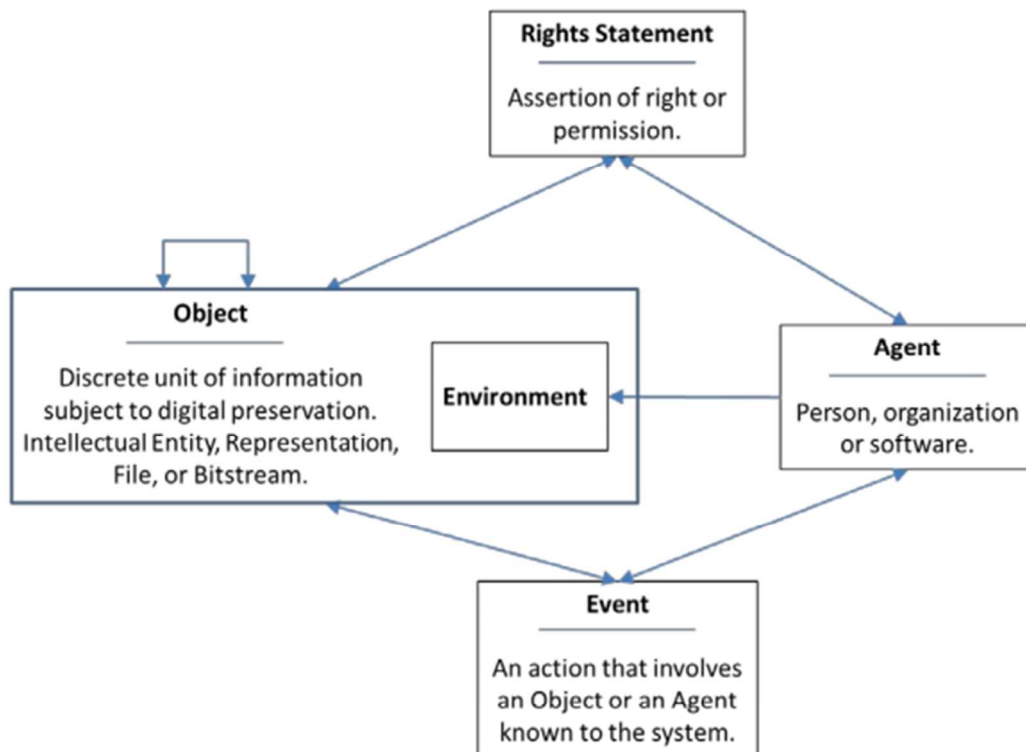


Figure 9. The PREMIS 3.0 data model⁵⁰

Provenator uses PREMIS metadata definitions to record the provenance of digital media items on the blockchain, using smart contracts. That ensures the data conforms to an open standard, which should ‘future-proof’ the information held and help facilitate further interactions with different users⁵¹. It also develops some of the ideas of Mannens et al.⁵², who propose using metadata, alongside descriptions,

1
2
3 to accompany news items because that would facilitate transparency and trust
4
5 estimation.
6
7

8 9 **The Provenator Application**

10
11 The general principle of Provenator is that a content creator should be able to prove
12
13 the provenance of the resources they create. To do so, Provenator gives creators the
14
15 ability to store relevant authentication information about their creations on the
16
17 blockchain so that it can be retrieved easily later and used to verify those same
18
19 resources.
20
21
22

23 24 **Requirements of the Provenator Application**

25
26 We are almost in a position to discuss Provenator in detail. However, we still need to
27
28 consider the steps required to prove the provenance metadata of media resources.
29
30 Thankfully, we need not think of those steps ourselves, because a similar ‘trust’
31
32 process is employed when distributing new releases of the Ubuntu operating system
33
34 software, which we describe next.
35
36
37
38

39 40 **Distributing the Ubuntu Operating System Software**

41
42 The steps for distributing Ubuntu, shown below, involve combining digital
43
44 signatures with PKC to help ensure that the software downloaded and installed can
45
46 be trusted. The process is as follows:
47
48

- 49 1 Download the operating system’s disk image, together with a file of checksums
50 and the signature used to sign the checksums file.
51
- 52 2 Fetch the public key used for the signature.
53
- 54 3 Use the key to verify the checksums file’s signature.
55
56
57
58
59
60

- 1
 - 2
 - 3
 - 4 Run a command that generates a SHA-256 cryptographic hash on the operating
 - 5 system disk image.
 - 6
 - 7
 - 8
 - 9
 - 10
 - 11
 - 12
 - 13
 - 14
 - 15
 - 16
 - 17
 - 18
 - 19
 - 20
 - 21
 - 22
 - 23
 - 24
 - 25
 - 26
 - 27
 - 28
 - 29
 - 30
 - 31
 - 32
 - 33
 - 34
 - 35
 - 36
 - 37
 - 38
 - 39
 - 40
 - 41
 - 42
 - 43
 - 44
 - 45
 - 46
 - 47
 - 48
 - 49
 - 50
 - 51
 - 52
 - 53
 - 54
 - 55
 - 56
 - 57
 - 58
 - 59
 - 60
- 4
 - 5

Hence, by following the process above, if the hashes match, a user can install the operating system and trust that they have an official Ubuntu release. Indeed, Alice could use a similar process to share her image with Bob.

Operations of the Provenator Application

Borrowing from the Ubuntu process for verifying the Ubuntu software, Provenator should do the following:

- 1 Get a cryptographic hash of the digital media resource.
- 2 Create the PREMIS metadata of the digital resource.
- 3 Sign the transaction that stores the cryptographic hash of the digital resource, and its associated metadata, on the blockchain.

By following that process, subsequent users of the data will be able to trust the integrity and authenticity of the digital media metadata because of the immutability of blockchain records. Below shows how Provenator will allow such users to check a digital resource's provenance data on the blockchain:

- 1 Get a cryptographic hash of the digital resource.
- 2 Check whether that hash exists on the blockchain.
- 3 If the hash exists, retrieve the associated metadata.

1
2
3 Next, we will look in more detail at Provenator's architecture.
4
5

6 7 **Provenator's Architecture**

8
9 Provenator consists of the following architecture:
10

- 11 • An Ethereum blockchain⁵⁴, that stores the provenance metadata about media
12 resources.
13
- 14 • Ethereum smart contracts, written in the language Solidity⁵⁵, which reads and
15 writes PREMIS metadata about media objects.
16
- 17 • A JavaScript web application, written in React⁵⁶, used for creating and accessing
18 the PREMIS data stored in the Ethereum smart contracts.
19
20
21
22
23
24
25
26

27 A working prototype of Provenator, as well as its source code, is available via the
28 source code repository GitHub.¹
29
30
31

32 33 **The Working Prototype**

34
35 The working prototype of Provenator exists on the network of the InterPlanetary
36 File System (IPFS). IPFS is a peer-to-peer, content-addressed file system that forms
37 the final component of our application's architecture; by publishing there, it means
38 that the application is wholly distributed because, as discussed above, its underlying
39 database, the blockchain, is also distributed. Furthermore, IPFS deploys
40 cryptographic tools to ensure the authenticity of resources stored on its network.
41
42
43
44
45
46
47
48
49 Thus, it is a good match for our technology. Below is a brief description of IPFS.
50
51
52
53

54
55
56 1 The address of the GitHub repository is
57 <https://github.com/glowkeeper/Provenator>.
58
59
60

The InterPlanetary File System

IPFS deploys a generalisation of a Merkle directed acyclic graph (DAG) to establish a decentralised network of trusted data. Applying cryptographic hashes to a graph was Ralph Merkle's solution for transferring reliable information over an untrusted network⁵⁷. The idea was profound; many systems that rely on trust employ Merkle DAGs - IPFS and Bitcoin⁵⁸ are just two examples among many. The fundamental principle behind a Merkle DAG is that if you have the hash of the root node, and the hash came from a trusted source, then, as long as the hashes match that of the root, you can trust all leaf nodes⁴². IPFS deploys a Merkle DAG to represent links between objects, which are cryptographic hashes of target blocks on the filesystem⁵⁹, a concept it has borrowed from the version control system Git⁶⁰. Figure 10 shows the representation of an image on IPFS. Hence, any file stored under IPFS is guaranteed to be unique. Moreover, as long as the file forms a Merkle DAG of objects, it can be trusted, too. Furthermore, because new objects hash differently, objects on IPFS are, essentially, immutable⁵⁹.

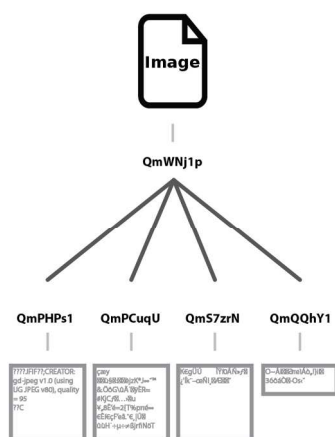


Figure 10. A Hash Tree⁶¹

1
2
3 Nodes on the IPFS network, which connect to one another to transfer and store
4
5 objects, can be considered as trusted sources since they use public-key
6
7 cryptography to establish their identity; they do so by using a cryptographic hash of
8
9 the public half of their public and private key pair. When two nodes connect, they do
10
11 so by exchanging those public keys, which are then used to encrypt subsequent
12
13 communication. IPFS nodes generate their key pairs using the asymmetric
14
15 cryptographic algorithm RSA⁶², which uses random numbers via entropy sources of
16
17 the IPFS nodes themselves. RSA's security relies on the properties of the integer
18
19 factorisation problem (IFP):
20
21
22

23
24
25 Given $n = pq$, find p and q , where p and q are primes.
26
27

28
29 IFP looks deceptively simple. However, provided that p and q are sufficiently large,
30
31 solving it is, actually, computationally infeasible³⁹.
32
33

34 **Not So Smart Contracts**

35
36 At the time of writing, the working prototype of Provenator uses the Ethereum
37
38 Testnet, Ropsten⁶³. However, we hope to produce a viable production release, so it
39
40 may be that, by the time of publication, the application is running on the Ethereum
41
42 blockchain itself. If that is the case, then Ethereum transactions that update the
43
44 blockchain cost Ether (the unit of currency on Ethereum), so there will be a fee for
45
46 storing metadata about digital resources.
47
48
49

50
51 Appendix G of the Ethereum yellow paper details some reasonably complex
52
53 calculations for determining the fee schedule of Ethereum transactions⁶⁴. However,
54
55 the essence of those fees is less code leads to less cost. Furthermore, retrieving
56
57
58
59
60

1
2
3 information from the blockchain is free. That leads to some important design
4
5 decisions when building a distributed application (dApp); not least is that the
6
7 JavaScript web application, which serves as the user interface, should do much of
8
9 the heavy lifting and the smart contracts should only set and get, rendering them not
10
11 so smart, after all. An example will serve to illustrate - when adding a media
12
13 resource to Provenator, the user must also input the agent, or content creator, who
14
15 owns that resource. A reasonable application design would be to send that agent
16
17 information to the smart contracts and have them check whether the agent already
18
19 exists in the database. However, that check, if it leads to a blockchain update, could
20
21 be prohibitively expensive. A less costly design is to have the smart contracts expose
22
23 a simple accessor method for retrieving agent data from an index of agents - an
24
25 operation that can be carried out for free. That way, the web application can use the
26
27 accessor method to perform the same check for nothing and only pay for agent data
28
29 to be stored on the blockchain if the agent does not already exist.
30
31
32
33
34
35
36
37

38 **Use of Provenator**

39
40 Consider the situation we described in the introduction to this paper, whereby the
41
42 supporter of the then-Republican candidate for the U.S. Presidency published a
43
44 photograph of a man behind some ballot boxes as an accompaniment to a claim that
45
46 the Democrats were rigging votes. Figure 11 below shows a screenshot from of the
47
48 Christian Times website making that claim.
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41

MONDAY, OCTOBER 3, 2016

ChristianTimes
Newspaper

NEWS ABOUT US POLITICS VIDEO RELIGION CHRISTIAN TIMES NEWS

Home > News > BREAKING: "Tens of thousands" of fraudulent Clinton votes found in Ohio warehouse

NEWS POLITICS

BREAKING: "Tens of thousands" of fraudulent Clinton votes found in Ohio warehouse

By *admin1* - September 30, 2016 46021 0

SHARE  Facebook  Twitter  G+  Pinterest



42 *Figure 11. A Snapshot of the Christian Times Website, Where it was Claimed the*
43 *Clinton's Were Rigging Votes. Picture Courtesy of The New York Times*³

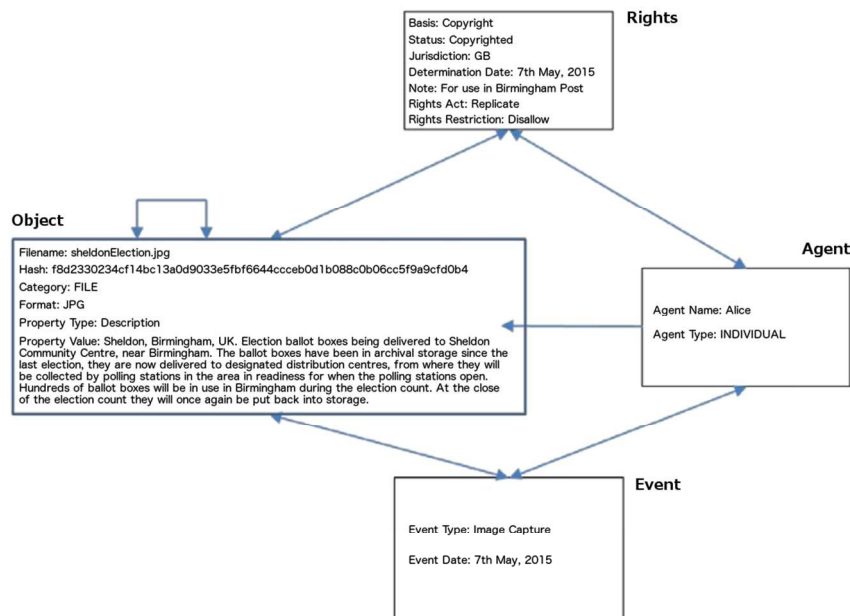
44
45 The exchangeable image file format (Exif), is a standard for specifying information
46 about image files⁶⁵, including data such as descriptions and copyright information.
47
48 Unfortunately, such data is easily changed⁶⁶. Presumably, the editor of the Christian
49 Times did just that, and therefore, the New York Times had to go to great lengths to
50 prove out of context use of the image. Now imagine that Alice was the photographer
51 who took that photograph and that she used Provenator to record data about the
52
53
54
55
56
57
58
59
60

1
2
3 picture on the blockchain. Under that circumstance, proving that the Christian
4
5 Times had used Alice's picture falsely would be a simple matter of using Provenator.
6
7 Thus, the New York Times could have saved itself much bother.
8
9

10
11 Next, we discuss the schema Alice uses to register herself, using Provenator, as the
12
13 creator of that photograph.
14
15

16 Provenator's PREMIS Metadata

17
18 Figure 12 below shows Alice using Provenator's PREMIS data model⁵⁰ to create
19
20 information about her photo, which she stores on the blockchain. She records a
21
22 cryptographic hash of her picture, along with associated metadata (such as a
23
24 description of the image), as a PREMIS object. She also records the date the photo
25
26 was taken, as a PREMIS event. The PREMIS agent describes Alice herself. The
27
28 PREMIS rights detail the image's license.
29
30
31
32



33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
Figure 12. The PREMIS 3.0 data model⁵⁰ Applied to Alice's Picture of the Sheldon Election Ballot Boxes

1
2
3 The implementation of the metadata, which we show above, describes a single
4
5 object - Alice's picture of the ballot boxes used in the Sheldon election. That object
6
7 has a single agent - Alice herself. It has a single event - the date when the picture was
8
9 taken, and a single right - the Birmingham Post's copyright. However, the
10
11 implementation of the PREMIS metadata used by Provenator is more complex. It
12
13 describes a PREMIS object that can have many properties, as well as many agents,
14
15 events and rights (for example, the licensing rights may be different in the UK to
16
17 those in the U.S.). Similarly, although an event may only belong to a single agent, an
18
19 agent may record multiple events, own many objects and deploy many different
20
21 rights. Finally, specific rights belong to a single object and a single agent.
22
23
24
25
26

27 28 **MetaMask**

29
30 MetaMask⁶⁷ is a tool able to run an Ethereum dApp in a browser. When using
31
32 Provenator, Alice can use MetaMask to sign the transactions she creates for storing
33
34 the PREMIS metadata about her picture on the blockchain. By doing so, anyone
35
36 accessing that data is confident that it was Alice herself who recorded the
37
38 information.
39
40
41
42

43 44 **Viewing The PREMIS Data**

45
46 Now Alice has recorded information about her photograph, Bob, her eEditor, can
47
48 use the image Alice sends to generate a cryptographic hash and retrieve information
49
50 about that hash from the blockchain. Figure 13 shows a screenshot of Provenator,
51
52 after Bob has recovered data about the picture Alice sent to him.
53
54
55
56
57
58
59
60

Select a File Object for Hashing

Select file:

Filename: sheldonElection.jpg

Hash: f8d2330234cf14bc13a0d9033e5fbf6644cceeb0d1b088c0b06cc5f9a9cfd0b4

Object Informationf8d2330234cf14bc13a0d9033e5fbf6644cceeb0d1b088c0b06cc5f9a9cfd0b4 - Category: FILE
f8d2330234cf14bc13a0d9033e5fbf6644cceeb0d1b088c0b06cc5f9a9cfd0b4 - Format: JPG

No. Properties: 1

f8d2330234cf14bc13a0d9033e5fbf6644cceeb0d1b088c0b06cc5f9a9cfd0b4 - Properties: Description - Sheldon, Birmingham, UK. Election ballot boxes being delivered to Sheldon Community Centre, near Birmingham. The ballot boxes have been in archival storage since the last election, they are now delivered to designated distribution centres, from where they will be collected by polling stations in the area in readiness for when the polling stations open. Hundreds of ballot boxes will be in use in Birmingham during the election count. At the close of the election count they will once again be put back into storage.

Object Event Information

No. Events: 1

f8d2330234cf14bc13a0d9033e5fbf6644cceeb0d1b088c0b06cc5f9a9cfd0b4 - Event ID: de37af3ee670897b9c05526a43f7ae7e6f85f5b11f80a6303fc64e9d4df68bf
de37af3ee670897b9c05526a43f7ae7e6f85f5b11f80a6303fc64e9d4df68bf - Event Object: f8d2330234cf14bc13a0d9033e5fbf6644cceeb0d1b088c0b06cc5f9a9cfd0b4
de37af3ee670897b9c05526a43f7ae7e6f85f5b11f80a6303fc64e9d4df68bf - Event Type: Image Capture
de37af3ee670897b9c05526a43f7ae7e6f85f5b11f80a6303fc64e9d4df68bf - Event Agent: fc0e76852d86642cf1425c0a75ba07e54228124f83bbb563d06f614dda4e47e5
de37af3ee670897b9c05526a43f7ae7e6f85f5b11f80a6303fc64e9d4df68bf - Event Time: 7th May, 2015

Object Agent Information

No. Agents: 1

f8d2330234cf14bc13a0d9033e5fbf6644cceeb0d1b088c0b06cc5f9a9cfd0b4 - Agent ID: fc0e76852d86642cf1425c0a75ba07e54228124f83bbb563d06f614dda4e47e5
fc0e76852d86642cf1425c0a75ba07e54228124f83bbb563d06f614dda4e47e5 - Agent Name: Alice
fc0e76852d86642cf1425c0a75ba07e54228124f83bbb563d06f614dda4e47e5 - Agent Type: INDIVIDUAL

Object Rights Information

No. Rights: 1

f8d2330234cf14bc13a0d9033e5fbf6644cceeb0d1b088c0b06cc5f9a9cfd0b4 - Rights ID: fe88001afc07abfba6452b40fc2ba449310a4a0fb3aff602b8e8ae450d4db
fe88001afc07abfba6452b40fc2ba449310a4a0fb3aff602b8e8ae450d4db - Rights Object: f8d2330234cf14bc13a0d9033e5fbf6644cceeb0d1b088c0b06cc5f9a9cfd0b4
fe88001afc07abfba6452b40fc2ba449310a4a0fb3aff602b8e8ae450d4db - Rights Basis: Copyright
fe88001afc07abfba6452b40fc2ba449310a4a0fb3aff602b8e8ae450d4db - Rights Copyright Status: Copyrighted
fe88001afc07abfba6452b40fc2ba449310a4a0fb3aff602b8e8ae450d4db - Rights Copyright Jurisdiction: GB
fe88001afc07abfba6452b40fc2ba449310a4a0fb3aff602b8e8ae450d4db - Rights Copyright Determination Date: 7th May, 2015
fe88001afc07abfba6452b40fc2ba449310a4a0fb3aff602b8e8ae450d4db - Rights Copyright Note: For use in Birmingham Post
fe88001afc07abfba6452b40fc2ba449310a4a0fb3aff602b8e8ae450d4db - Rights Granted Act: Replicate
fe88001afc07abfba6452b40fc2ba449310a4a0fb3aff602b8e8ae450d4db - Rights Granted Restriction: Disallow
fe88001afc07abfba6452b40fc2ba449310a4a0fb3aff602b8e8ae450d4db - Rights Agent: fc0e76852d86642cf1425c0a75ba07e54228124f83bbb563d06f614dda4e47e5

Finished fetching records from the blockchain.

Figure 13. Screenshot of Bob Using Provenator to Retrieve Information About Alice's Picture. Source: Authors' own work, whereby the scenario depicted in this paper has been recreated.

Due to the deterministic and collision resistance properties of cryptographic hashes, by retrieving the data above, Bob is confident as to the authenticity of the image Alice sent. He can also apply edits and record information about those changes, thus creating a provenance chain for the picture. Hence, rather than going to great investigative lengths to prove out of context use of Alice's image, the New York Times would have been able to check the validity of the picture simply by uploading the British Times' copy to Provenator. Then they would have retrieved the same metadata as Bob, which would have shown the picture to be fake.

However, although that would have shown the image itself was fake, it would not have proved that the article as a whole was fiction. Proving that might take a little more than technology. We consider that issue, next.

Validating News

The BBC has had many difficulties in providing accurate news stories from behind the frontlines of the Syrian conflict⁶⁸. Indeed, journalists have lost their lives there, so it has become common practice to source stories from ordinary Syrian citizens. However, ensuring the validity of such 'user-generated content' (UGC) has been "a skill journalists have had to learn"⁶⁸. To that end, the BBC has become proficient at developing new practices that ensure the validity of UGC content. Apparently, such methods involve technology, but also common sense and fostering healthy relationships with reliable Syrians⁶⁸. Augmenting big data news stream technology with a 'human touch' to verify items is a common theme⁶⁹. For example, one project argues for the formation of a fake news corpus to aid deception detection, and to that end, when collecting the data, qualified participants will be required to spot the fakes³³. In fact, all of the big data technologies we mentioned above require some form of human action - either through visualising graphs or acting upon some visual data. Therein lies the crucial point; when the BBC check the validity of stories given by users behind the Syrian front lines, technology can only go so far. A good deal of human skill is required, too. Moreover, while technology, such as Provenator, will make it possible to prove the validity of media resources used within news, proving the authenticity of fake news stories as a whole often takes good journalistic practices. Another good example is the experience of Facebook; while countering propaganda in the run-up to the 2016 U.S. Election, the company found that their algorithms were not always up to the job of spotting fake stories. Instead, to curate the news items appearing on their site, they had to fall back on human editors⁷⁰.

Current Limitations of Provenator and Future Work

1
2
3
4
5
6 A strength of Provenator is also a weakness. The strength is that the same digital
7
8 media resource will always generate the same cryptographic hash. Thus, if two
9
10 hashes match, it is certain that it is the same object. Therefore, we can retrieve
11
12 provenance data and trust that it accurately reflects the object's origins. To put that
13
14 another way, changing a single pixel in a digital resource will generate an entirely
15
16 different cryptographic hash. Therein lies the weakness - it would not have been
17
18 difficult for the Christian Times to alter the image of the Sheldon Election ballot
19
20 boxes, thus, as it stands, defeating our tool.
21
22
23

24
25 However, that weakness in our early prototype of Provenator is not
26
27 insurmountable. For example, it may be possible to use some form of mathematical
28
29 filter to remove or reduce the 'noise' of an object, thus rendering two seemingly
30
31 disparate resources, identical⁷¹. There may be better approaches than filtering,
32
33 however. Narwal et al. describe how they classify similar images using fisher vectors
34
35 and k-means clustering³⁶. Indeed, object classification via fisher vectors appears to
36
37 be an active area of computer vision research⁷². Hence, if Provenator used such
38
39 techniques, users may be able to classify images, discover similarities and find fakes
40
41 that way. Furthermore, fisher vectors are used for classifying videos, too⁷³, so
42
43 Provenator's scope could broaden beyond images. That could be true for another
44
45 technique, too - perceptual hashes⁷⁴, which establish object matches based on
46
47 perceived content⁷⁵. Whereas any change in two multimedia resources will generate
48
49 vastly different cryptographic hashes, perceptual hashes produce comparable
50
51 results if the resources are similar. Hence, if future versions of Provenator extend its
52
53
54
55
56
57
58
59
60

1
2
3 resource metadata to include a perceptual hash, that single pixel change above
4
5 would render a complimentary perceptual hash that can be matched against the
6
7 original by calculating their hamming distance⁷⁶. Indeed, perceptual hashing is
8
9 already used by organisations such as Shazam, Google and also by Youtube to detect
10
11 copyright infringement across a broad range of digital objects, such as audio, video
12
13 and images⁷⁵. Indeed, although this paper uses the example of a picture to help
14
15 explain the application's functionality, Provenator can be used to prove the
16
17 provenance of any media objects, even the news stories themselves. In fact,
18
19 improvements in future versions, using methods such as fisher vectors and or
20
21 perceptual hashes, would make it even more suitable as a tool for helping to prove
22
23 the origins of different media resources.
24
25
26
27
28
29
30

31 **Conclusion**

32
33 Fake news has hit the headlines recently. Indeed, Donald Trump has continued to
34
35 accuse various media outlets of distributing falsehoods that undermine him⁷⁷. We
36
37 have not examined the reasons for his doing so; such an examination would be
38
39 interesting, but it is not the focus of this paper. Moreover, although we have given
40
41 some background history on the issue of fake news, it is beyond the scope of this
42
43 article to discuss propaganda itself. Additionally, although we discuss the issue of
44
45 social media platforms and fake news, we do not examine the methods and
46
47 processes for distributing fake news items on such platforms or the efficacy of the
48
49 measures taken by those platforms to counter the problem. Instead, the purpose of
50
51 this paper has been to propose a technological solution to the problem of proving
52
53 the validity of media resources used in fake news. Various research groups are
54
55
56
57
58
59
60

1
2
3 investigating technologies capable of overcoming the problem of verifying big data
4 news streams. However, the application we have developed, Provenator, is uniquely
5 capable of recording metadata about digital media on blockchain technology so that
6 it becomes trivial to prove their authenticity in a manner that can be trusted. The
7 ultimate aim of our tool is to make content creators accountable for the resources
8 they create.
9

10
11 Unfortunately, as it stands, although Provenator works well for recording the
12 origins of a media resource, it is easy to defeat the 'find fake' capabilities of this early
13 prototype, simply by changing a single pixel of a misappropriated image. That may
14 be addressed in future version, since there are techniques available, such as fisher
15 vectors and perceptual hashes, which can improve future versions of the application
16 and make it much more capable.
17

18
19 However, while Provenator may become more proficient at verifying the
20 authenticity of media resources used within a story, the application will only ever be
21 capable of providing a partial solution to the problem of fake news. Unfortunately,
22 we do not think technology will ever be wholly capable of proving the truth of the
23 story as a whole. We believe, currently, that takes human skills. Certainly, while it
24 might take some sophisticated mathematics to determine the similarity between
25 two media resources that only differ by a single pixel, the same complexity does not
26 apply to the human eye, which would quickly decide that those resources are the
27 same.
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 Although we have reservations about the possible limitations of technology in
4 combating fake news, we believe the trust mechanisms of blockchains make them
5 better positioned than other technologies for proving the authenticity of media
6 resources. Indeed, organisations are investigating using blockchains for purposes
7 such as transparency and publicly auditable content ranking⁷⁸. Moreover, our
8 application is an example of a tool that can help fight 'fakeness'. Indeed, in our
9 supposed scenario, where Alice was the photographer who took the image used by
10 the Christian Times, The New York Times would have had a much easier job of
11 proving falsehood.
12
13
14
15
16
17
18
19
20
21
22
23
24

25 **Acknowledgements**

26 The idea for this paper came after a discussion over tea with colleagues at the
27 University of Sussex; namely, Phil Watten and Patrick Holroyd. Thank you also to
28 Ian Wakeman, Head of Informatics at the University of Sussex, who provided
29 feedback on the first draft of the paper and, in particular, provided useful references
30 on image capture. Also thank you to Konstantin Blyuss, reader in Mathematics at the
31 University of Sussex, who provided insight about cryptography. Finally, we're
32 grateful to the anonymous reviewers, as well as the editors, who gave suggestions
33 that improved the paper immeasurably.
34
35
36
37
38
39
40
41
42
43
44
45
46
47

48 **References**

49 1. Watts AW. How Russia Dominates Your Twitter Feed to Promote Lies (And,
50 Trump, Too) [Internet]. The Daily Beast. 2016–2016-08-06T04:03:00.000Z [cited
51 2017 Feb 5]. Available from:
52 <http://www.thedailybeast.com/articles/2016/08/06/how-russia-dominates-your-tweet-feed-to-promote-lies-and-trump-too.html>
53
54
55
56
57
58
59
60

- 1
2
3 2. Morin R. Trump: New York Times is 'fake news' [Internet]. POLITICO. 17AD [cited
4 2017 Feb 5]. Available from: <http://politi.co/2jAdgwn>
5
6
- 7 3. Shane S. From Headline to Photograph, a Fake News Masterpiece. The New York
8 Times [Internet]. 2017 Jan 18 [cited 2017 Feb 5]; Available from:
9 [https://www.nytimes.com/2017/01/18/us/fake-news-hillary-clinton-
10 harris.html](https://www.nytimes.com/2017/01/18/us/fake-news-hillary-clinton-cameron-harris.html)
11
- 12 4. Walker J. Birmingham Mail photo used in a 'fake news' Trump story shared with 6
13 million [Internet]. birminghammail. 2017 [cited 2017 Sep 27]. Available from:
14 [http://www.birminghammail.co.uk/news/midlands-news/
15 birmingham-mail-photo-used-fake-12476900](http://www.birminghammail.co.uk/news/midlands-news/birmingham-mail-photo-used-fake-12476900)
16
17
- 18 5. Hunt E. What is fake news? How to spot it and what you can do to stop it. The
19 Guardian: Media [Internet]. 2016 Dec 17 [cited 2017 Feb 27]; Available from:
20 <https://www.theguardian.com/media/2016/dec/18/what-is-fake-news-pizzagate>
21
22
- 23 6. Jessica Goodfellow. Only 4% of people can distinguish fake news from truth,
24 Channel 4 study finds [Internet]. The Drum. 2017 [cited 2017 Mar 23]. Available
25 from: [http://www.thedrum.com/news/2017/02/06/only-4-people-can-
26 distinguish-fake-news-truth-channel-4-study-finds](http://www.thedrum.com/news/2017/02/06/only-4-people-can-distinguish-fake-news-truth-channel-4-study-finds)
27
- 28 7. Foo Yun Chee. NATO says it's seen a sharp rise in Russian fake news since the
29 Kremlin seized Crimea [Internet]. Business Insider. 2017 [cited 2017 Apr 20].
30 Available from: [http://uk.businessinsider.com/r-nato-says-it-sees-sharp-rise-in-
31 russian-disinformation-since-crimea-seizure-2017-2](http://uk.businessinsider.com/r-nato-says-it-sees-sharp-rise-in-russian-disinformation-since-crimea-seizure-2017-2)
32
33
- 34 8. Fake news series part two [Internet]. Channel 4 News. 19:00 00:41:21-00:55:54:
35 Channel 4; 2017. Available from:
36 <https://learningonscreen.ac.uk/ondemand/index.php/clip/90422>
37
- 38 9. MacIntyre D. Facebook - the secret election weapon. BBC News: UK [Internet].
39 20172017-05-08T02:52:38+01:00 [cited 2017 May 8]; Available from:
40 <http://www.bbc.co.uk/news/uk-39830727>
41
42
- 43 10. Silverman C. This Analysis Shows How Viral Fake Election News Stories
44 Outperformed Real News On Facebook [Internet]. BuzzFeed. 2016 [cited 2017 May
45 9]. Available from: [https://www.buzzfeed.com/craigsilverman/viral-fake-election-
46 news-outperformed-real-news-on-facebook](https://www.buzzfeed.com/craigsilverman/viral-fake-election-news-outperformed-real-news-on-facebook)
47
- 48 11. Jen Weedon, William Nuland, Alex Stamos. Facebook and Information
49 Operations [Internet]. Facebook; 2017 [cited 2017 May 8]. Available from:
50 [https://fbnewsroomus.files.wordpress.com/2017/04/facebook-and-information-
51 operations-v1.pdf](https://fbnewsroomus.files.wordpress.com/2017/04/facebook-and-information-operations-v1.pdf)
52
53
- 54 12. Cellan-Jones R. Facebook publishes fake news ads in UK papers. BBC News:
55 Technology [Internet]. 20172017-05-08T05:00:25+01:00 [cited 2017 May 8];
56 Available from: <http://www.bbc.co.uk/news/technology-39840803>
57
58
59
60

- 1
2
3 13. Mark Zuckerberg. What we're doing about misinformation [Internet]. 2016
4 [cited 2017 May 9]. Available from:
5 <https://www.facebook.com/zuck/posts/10103269806149061?pnref=story>
6
7
8 14. reporter AHT. Wikipedia founder to fight fake news with new Wikitribune site.
9 The Guardian: Technology [Internet]. 2017 Apr 25 [cited 2017 May 8]; Available
10 from: [https://www.theguardian.com/technology/2017/apr/25/wikipedia-founder-](https://www.theguardian.com/technology/2017/apr/25/wikipedia-founder-jimmy-wales-to-fight-fake-news-with-new-wikitribune-site)
11 [jimmy-wales-to-fight-fake-news-with-new-wikitribune-site](https://www.theguardian.com/technology/2017/apr/25/wikipedia-founder-jimmy-wales-to-fight-fake-news-with-new-wikitribune-site)
12
13 15. Nationalism vs. globalism: The new political divide [Internet]. 2017 [cited 2017
14 Mar 3]. Available from:
15 [http://www.ted.com/talks/yuval_noah_harari_nationalism_vs_globalism_the_new_p](http://www.ted.com/talks/yuval_noah_harari_nationalism_vs_globalism_the_new_political_divide)
16 [olitical_divide](http://www.ted.com/talks/yuval_noah_harari_nationalism_vs_globalism_the_new_political_divide)
17
18 16. Nagle DB, Burstein SM, editors. The ancient world: Readings in social and
19 cultural history. 4th ed. New York: Prentice Hall; 2010. 280 pp.
20
21 17. Online Etymology Dictionary. Propaganda [Internet]. 2017 [cited 2017 Mar 27].
22 Available from: <http://www.etymonline.com/index.php?term=propaganda>
23
24 18. Dr David Clarke. The corpse factory and the birth of fake news. BBC News:
25 Entertainment & Arts [Internet]. 2017 2017-02-17T08:29:53+00:00 [cited 2017 Feb
26 25]; Available from: <http://www.bbc.co.uk/news/entertainment-arts-38995205>
27
28 19. A-Z Quotes. Joseph Goebbels Quote [Internet]. A-Z Quotes. 2017 [cited 2017 Mar
29 4]. Available from: <http://www.azquotes.com/quote/1419276>
30
31 20. Davies S. Popular opinion in Stalin's Russia: Terror, propaganda, and dissent,
32 1934-1941. Cambridge ; New York: Cambridge University Press; 1997. 236 pp.
33
34 21. Zasurskiĭ I. Media and power in post-Soviet Russia. Armonk, N.Y: M.E. Sharpe;
35 2004. 269 pp.
36
37 22. Jane's 360. Acknowledgement of Russia's information warfare capability
38 indicates its strategic importance and impracticability of maintaining plausible
39 [Internet]. 2017 [cited 2017 May 4]. Available from:
40 [http://www.janes.com/article/68267/acknowledgement-of-russia-s-information-](http://www.janes.com/article/68267/acknowledgement-of-russia-s-information-warfare-capability-indicates-its-strategic-importance-and-impracticability-of-maintaining-plausible-deniability-policy)
41 [warfare-capability-indicates-its-strategic-importance-and-impracticability-of-](http://www.janes.com/article/68267/acknowledgement-of-russia-s-information-warfare-capability-indicates-its-strategic-importance-and-impracticability-of-maintaining-plausible-deniability-policy)
42 [maintaining-plausible-deniability-policy](http://www.janes.com/article/68267/acknowledgement-of-russia-s-information-warfare-capability-indicates-its-strategic-importance-and-impracticability-of-maintaining-plausible-deniability-policy)
43
44 23. Khaldarova I, Pantti M. Fake News: The narrative battle over the Ukrainian
45 conflict. Journalism Practice [Internet]. 2016 Oct 2 [cited 2017 Mar 2];10(7):891–
46 901. Available from:
47 <https://www.tandfonline.com/doi/full/10.1080/17512786.2016.1163237> (DOI:
48 [10.1080/17512786.2016.1163237](https://www.tandfonline.com/doi/full/10.1080/17512786.2016.1163237))
49
50 24. Bernays EL, Miller MC. Propaganda. Brooklyn, N.Y: Ig Publishing; 2005. 168 pp.
51
52
53
54
55
56
57
58
59
60

- 1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
25. McChesney RW, Wood EM, Foster JB, editors. *Capitalism and the information age: The political economy of the global communication revolution*. New York, NY: Monthly Review Press; 1998. 254 pp.
26. Office of the Historian. *Foreign Relations of the United States, 1945–1950, Emergence of the Intelligence Establishment* [Internet]. 1948 [cited 2017 May 4]. Available from: <https://history.state.gov/historicaldocuments/frus1945-50Intel/d292>
27. Robert Love. *Before Jon Stewart* [Internet]. *Columbia Journalism Review*. 2007 [cited 2017 Mar 1]. Available from: http://www.cjr.org/feature/before_jon_stewart.php
28. Orwell G. *1984: A novel; revised and updated bibliography*. New York [u.a.: New American Library; 1985.
29. Staff LII. *First Amendment* [Internet]. LII / Legal Information Institute. 2010 [cited 2017 Mar 1]. Available from: https://www.law.cornell.edu/constitution/first_amendment
30. *Fixing the News* [Internet]. 2015 [cited 2017 May 4]. (TEDxVancouver). Available from: <https://www.youtube.com/watch?v=NwmGTM5Py8Y>
31. boyd danah, Crawford K. *Six Provocations for Big Data*. *SSRN Electronic Journal* [Internet]. 2011 [cited 2017 Sep 20]; Available from: <http://www.ssrn.com/abstract=1926431> (DOI: [10.2139/ssrn.1926431](https://doi.org/10.2139/ssrn.1926431))
32. Labrinidis A, Jagadish HV. *Challenges and opportunities with big data*. *Proceedings of the VLDB Endowment* [Internet]. 2012 Aug 1 [cited 2017 Sep 20];5(12):2032–3. Available from: <http://dl.acm.org/citation.cfm?doid=2367502.2367572> (DOI: [10.14778/2367502.2367572](https://doi.org/10.14778/2367502.2367572))
33. Rubin VL, Chen Y, Conroy NJ. *Deception detection for news: Three types of fakes: Deception Detection for News: Three Types of Fakes*. *Proceedings of the Association for Information Science and Technology* [Internet]. 2015 [cited 2017 Sep 21];52(1):1–4. Available from: <http://doi.wiley.com/10.1002/pra2.2015.145052010083> (DOI: [10.1002/pra2.2015.145052010083](https://doi.org/10.1002/pra2.2015.145052010083))
34. QUENTIN VIERGE. *Journalism: Gathering Information and Writing Your Story* [Internet]. 2017 [cited 2017 Apr 12]. Available from: <https://writingcommons.org/open-text/research-methods-methodologies/empirical-research/interviews/journalism-gathering-information-and-writing-your-story>
35. Ed Grover. *City journalism academics to lead European big data and fake news project* [Internet]. City, University of London. 7th July 2017 [cited 2017 Sep 21].

1
2
3 Available from: [https://www.city.ac.uk/news/2017/june/google-digital-news-](https://www.city.ac.uk/news/2017/june/google-digital-news-initiative-dminr)
4 [initiative-dminr](https://www.city.ac.uk/news/2017/june/google-digital-news-initiative-dminr)
5

6
7 36. Narwal V, Salih MH, Lopez JA, Ortega A, O'Donovan J, Höllerer T, et al. Automated
8 Assistants to Identify and Prompt Action on Visual News Bias. In ACM Press; 2017
9 [cited 2017 Sep 22]. pp. 2796–801. Available from:
10 <http://dl.acm.org/citation.cfm?doid=3027063.3053227> (DOI:
11 [10.1145/3027063.3053227](https://doi.org/10.1145/3027063.3053227))
12

13
14 37. Jin Z, Cao J, Jiang Y-G, Zhang Y. News Credibility Evaluation on Microblog with a
15 Hierarchical Propagation Model. In IEEE; 2014 [cited 2017 Sep 22]. pp. 230–9.
16 Available from: <http://ieeexplore.ieee.org/document/7023340/> (DOI:
17 [10.1109/ICDM.2014.91](https://doi.org/10.1109/ICDM.2014.91))
18

19
20 38. Menezes AJ, Van Oorschot PC, Vanstone SA. Handbook of applied cryptography.
21 Boca Raton: CRC Press; 1997. 780 pp. (CRC press series on discrete mathematics
22 and its applications).
23

24 39. Dr Konstantin Blyuss. Cryptography - Lecture Notes. 2016.
25

26 40. Ethereum. Ethereum Glossary [Internet]. GitHub. 2016 [cited 2017 Mar 22].
27 Available from: <https://github.com/ethereum/wiki>
28

29 41. KohanX. Public Key Cryptography [Internet]. 2010. Available from:
30 <https://commons.wikimedia.org/wiki/File:Public-key-crypto-1.svg>
31

32 42. Tyler Cipriani. Visualizing Git's Merkle DAG with D3.js - Tyler Cipriani [Internet].
33 2016 [cited 2017 Mar 13]. Available from:
34 [https://tylercipriani.com/blog/2016/03/21/Visualizing-Git-Merkle-DAG-with-](https://tylercipriani.com/blog/2016/03/21/Visualizing-Git-Merkle-DAG-with-D3.js/)
35 [D3.js/](https://tylercipriani.com/blog/2016/03/21/Visualizing-Git-Merkle-DAG-with-D3.js/)
36

37
38 43. Bitcoin Wiki. Elliptic Curve Digital Signature Algorithm [Internet]. 2015 [cited
39 2017 Mar 27]. Available from:
40 [https://en.bitcoin.it/wiki/Elliptic Curve Digital Signature Algorithm](https://en.bitcoin.it/wiki/Elliptic_Curve_Digital_Signature_Algorithm)
41

42 44. Swan M. Blockchain: Blueprint for a New Economy. "O'Reilly Media, Inc." 2015.
43 149 pp.
44

45
46 45. Umeh J. Blockchain Double Bubble or Double Trouble? ITNOW [Internet]. 2016
47 Mar [cited 2017 Feb 5];58(1):58–61. Available from:
48 <https://academic.oup.com/itnow/article-lookup/doi/10.1093/itnow/bww026>
49 (DOI: [10.1093/itnow/bww026](https://doi.org/10.1093/itnow/bww026))
50

51
52 46. Huckle S, White M. Socialism and the Blockchain. Future Internet [Internet].
53 2016 Oct 18 [cited 2016 Nov 7];8(4):49. Available from:
54 <http://www.mdpi.com/1999-5903/8/4/49> (DOI: [10.3390/fi8040049](https://doi.org/10.3390/fi8040049))
55
56
57
58
59
60

- 1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
47. Eris Industries. Explainer | Smart Contracts [Internet]. Eris Industries Documentation. 2016 [cited 2016 Mar 19]. Available from: https://docs.erisindustries.com/explainers/smart_contracts/
48. The Economist. The trust machine. The Economist [Internet]. 2015 Oct 31 [cited 2016 Feb 17]; Available from: <http://www.economist.com/news/leaders/21677198-technology-behind-bitcoin-could-transform-how-economy-works-trust-machine>
49. Priscilla Caplan. Understanding PREMIS [Internet]. The Library of Congress; 2009 [cited 2017 Sep 23]. Available from: <https://www.loc.gov/standards/premis/understanding-premis.pdf>
50. PREMIS Editorial Committee. The PREMIS Data Dictionary Version 3.0 [Internet]. 2015 [cited 2017 Jun 29]. Available from: <https://www.loc.gov/standards/premis/v3/premis-3-0-final.pdf>
51. W3C. Data on the Web - Best Practices [Internet]. 2015 [cited 2017 May 11]. Available from: <https://www.w3.org/TR/2015/WD-dwbp-20150224/>
52. Mannens E, Coppens S, Verborgh R, Hauttekeete L, Van Deursen D, Van de Walle R. Automated Trust Estimation in Developing Open News Stories: Combining Memento and Provenance. In IEEE; 2012 [cited 2017 Sep 22]. pp. 122–7. Available from: <http://ieeexplore.ieee.org/document/6341562/> (DOI: [10.1109/COMPSACW.2012.32](https://doi.org/10.1109/COMPSACW.2012.32))
53. Ubuntu. VerifyIsoHowto - Community Help Wiki [Internet]. 2016 [cited 2017 Mar 11]. Available from: <https://help.ubuntu.com/community/VerifyIsoHowto>
54. Ethereum. Ethereum Project [Internet]. 2017 [cited 2017 Jan 10]. Available from: <https://www.ethereum.org/>
55. Solidity. Solidity — Solidity 0.4.8-develop documentation [Internet]. 2016 [cited 2017 Jan 13]. Available from: <https://solidity.readthedocs.io/en/develop/>
56. React. React - A JavaScript library for building user interfaces [Internet]. 2017 [cited 2017 Jun 29]. Available from: <https://facebook.github.io/react/>
57. Merkle RC. A Digital Signature Based on a Conventional Encryption Function. In: Pomerance C, editor. Advances in Cryptology — CRYPTO '87 [Internet]. Berlin, Heidelberg: Springer Berlin Heidelberg; 1988 [cited 2017 Mar 13]. pp. 369–78. Available from: http://link.springer.com/10.1007/3-540-48184-2_32 (DOI: [10.1007/3-540-48184-2_32](https://doi.org/10.1007/3-540-48184-2_32))
58. Bitcoin. Bitcoin - Open source P2P money [Internet]. 2015 [cited 2015 Nov 27]. Available from: <https://bitcoin.org/en/>
59. Juan Benet. IPFS - Content Addressed, Versioned, P2P File System (DRAFT 3) [Internet]. 2017 [cited 2017 Mar 13]. Available from:

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

<https://ipfs.io/ipfs/QmR7GSQM93Cx5eAg6a6yRzNde1FQv7uL6X1o4k7zrJa3LX/ipfs.draft3.pdf>

60. Git. Git [Internet]. 2017 [cited 2017 Mar 13]. Available from: <https://git-scm.com/>

61. flyingzumwalt. The Decentralized Web Primer - Lesson: Turn a File into a Merkle Tree [Internet]. 2016 [cited 2017 Mar 13]. Available from: <https://flyingzumwalt.gitbooks.io/decentralized-web-primer/content/ipfs-dag/lessons/files-as-dags.html>

62. Dell EMC. RSA Laboratories - PKCS #1: RSA Cryptography Standard [Internet]. 2017 [cited 2017 May 23]. Available from: <https://www.emc.com/emc-plus/rsa-labs/standards-initiatives/pkcs-rsa-cryptography-standard.htm>

63. Ethereum. Ropsten: Ropsten public testnet PoW chain [Internet]. ethereum; 2017 [cited 2017 Jul 16]. Available from: <https://github.com/ethereum/ropsten>

64. Gavin Wood. ETHEREUM - A SECURE DECENTRALISED GENERALISED TRANSACTION LEDGER. EIP-150 REVISION [Internet]. 2013 [cited 2017 Jan 16]. Available from: <http://gavwood.com/paper.pdf>

65. EXIF.org. EXIF and related resources [Internet]. 2017 [cited 2017 Sep 26]. Available from: <http://www.exif.org/>

66. Mauro Huculak. How to edit image metadata on Windows 10 [Internet]. Windows Central. 2017 [cited 2017 Sep 26]. Available from: <https://www.windowscentral.com/how-edit-picture-metadata-windows-10>

67. MetaMask. MetaMask [Internet]. 2017 [cited 2017 Mar 28]. Available from: <https://metamask.io/>

68. Lisette Johnston. How is citizen journalism transforming the BBC's Newsroom practices? [Internet]. 2017. Available from: https://www.academia.edu/20427847/How_is_citizen_journalism_transforming_the_BBC_s_Newsroom_practices

69. Sethi RJ. Crowdsourcing the Verification of Fake News and Alternative Facts. In ACM Press; 2017 [cited 2017 Sep 21]. pp. 315–6. Available from: <http://dl.acm.org/citation.cfm?doid=3078714.3078746> (DOI: [10.1145/3078714.3078746](https://doi.org/10.1145/3078714.3078746))

70. McHugh M. Facebook Can Try to Fix Fake News, but It Can Never Be an Arbiter of Truth [Internet]. The Ringer. 2016 [cited 2017 May 9]. Available from: <https://theringer.com/facebook-can-try-to-fix-fake-news-but-it-can-never-be-an-arbiter-of-truth-75d4d0ca0176#.9pdaniqlk>

71. Datta R, Joshi D, Li J, Wang JZ. Image retrieval: Ideas, influences, and trends of the new age. ACM Computing Surveys [Internet]. 2008 Apr 1 [cited 2017 Jul

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

24];40(2):1–60. Available from:

<http://portal.acm.org/citation.cfm?doid=1348246.1348248> (DOI:
[10.1145/1348246.1348248](https://doi.org/10.1145/1348246.1348248))

72. Liu L, Wang P, Shen C, Wang L, van den Hengel A, Wang C, et al. Compositional Model Based Fisher Vector Coding for Image Classification. IEEE Transactions on Pattern Analysis and Machine Intelligence [Internet]. 2017 [cited 2017 Sep 24];1–1. Available from: <http://ieeexplore.ieee.org/document/7812753/> (DOI: [10.1109/TPAMI.2017.2651061](https://doi.org/10.1109/TPAMI.2017.2651061))

73. Sun C, Nevatia R. Large-scale web video event classification by use of Fisher Vectors. In IEEE; 2013 [cited 2017 Sep 24]. pp. 15–22. Available from: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6474994> (DOI: [10.1109/WACV.2013.6474994](https://doi.org/10.1109/WACV.2013.6474994))

74. Joe Bertolami. Perceptual Hashing [Internet]. Wednesday, May 28th, 2014 [cited 2017 Sep 24]. Available from: <http://bertolami.com/index.php?engine=blog&content=posts&detail=perceptual-hashing>

75. rframow. Adventures in Perceptual Hashing [Internet]. AAPB National Digital Stewardship Residency. 2017–2017-04-20T14:27:45+00:00 [cited 2017 Sep 24]. Available from: <https://ndsr.americanarchive.org/2017/04/20/adventures-in-perceptual-hashing/>

76. Hamming RW. Error Detecting and Error Correcting Codes. Bell System Technical Journal [Internet]. 1950 Apr [cited 2017 Sep 27];29(2):147–60. Available from: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6772729> (DOI: [10.1002/j.1538-7305.1950.tb00463.x](https://doi.org/10.1002/j.1538-7305.1950.tb00463.x))

77. Nikki Schwab. Trump calls NBC and ABC 'totally biased and fake news' AGAIN [Internet]. Mail Online. 2017 [cited 2017 Mar 28]. Available from: <http://www.dailymail.co.uk/~/article-4342518/index.html>

78. Jill Richmond. Using Blockchain Technology to Fight Fake News [Internet]. Distributed. 2017 [cited 2017 Jul 16]. Available from: [/news/how-blockchain-can-fight-fake-news/](https://news/how-blockchain-can-fight-fake-news/)