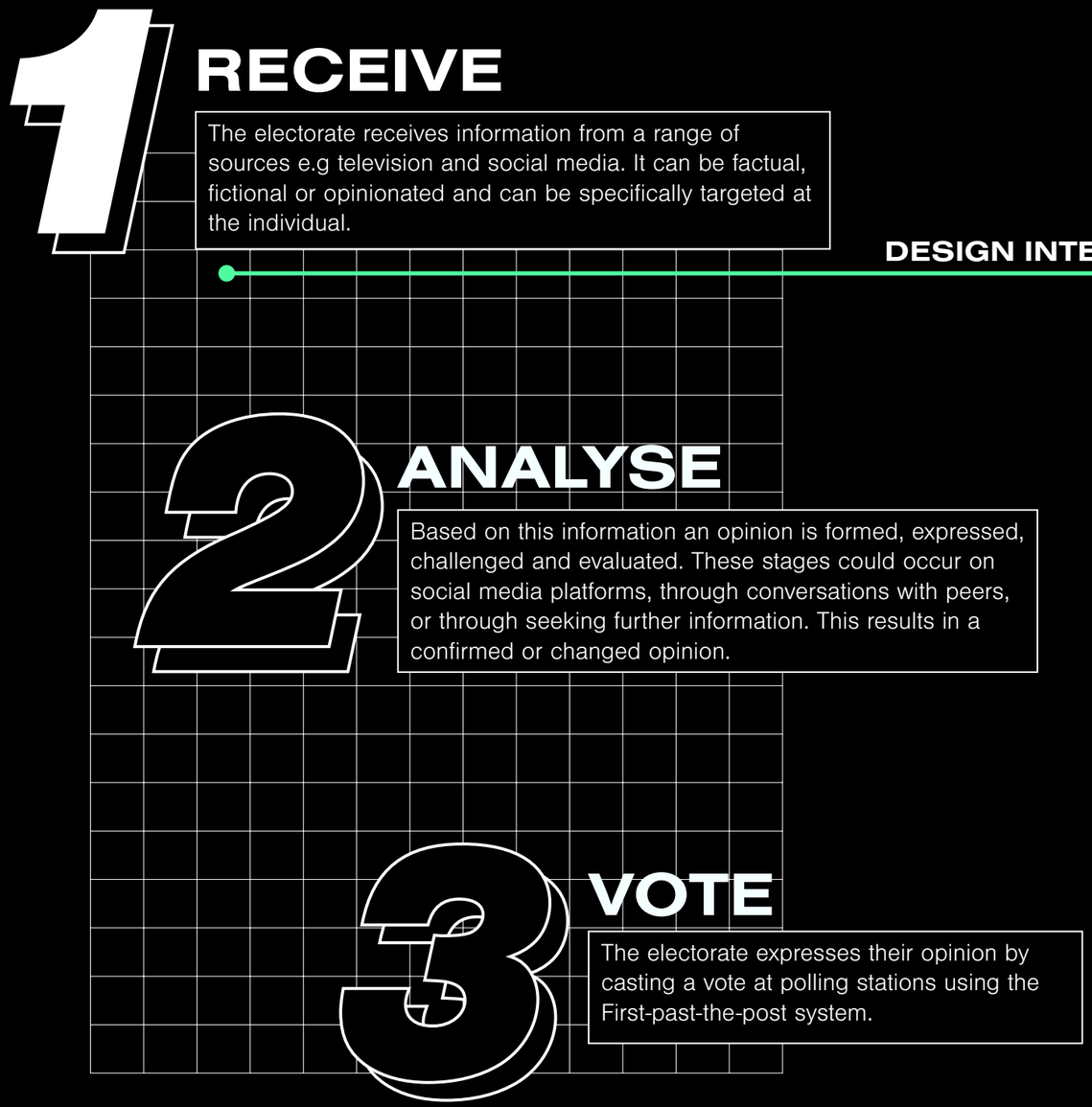
By Ella Cope, Saym Hussain, Esther Maltby, Tomáš Kňaze, and Jordan Kotler



Executive Summary

The future of democracy depends on having informed voters. Therefore, the electorate need to be aware of the validity of the information they receive in the media. The design intervention is an online aggregate news platform which uses a decentralised algorithm to assess the validity of information. The platform is to be implemented in the UK in 2030.

Authentication is executed by randomly selecting a parked car and running an open source algorithm on its internal computer. This protects the system from cyber attacks while encryption ensures biases do not influence the authentication. The algorithm identifies information that can be proven true or false, as well as opinions that cannot be proven either way.

Anyone can submit content to be published on the platform. Each publisher has an associated ‘credibility score’ which evolves and is affected by the authentication results of their previous submissions. This score influences the number of users who consume the publisher’s content.

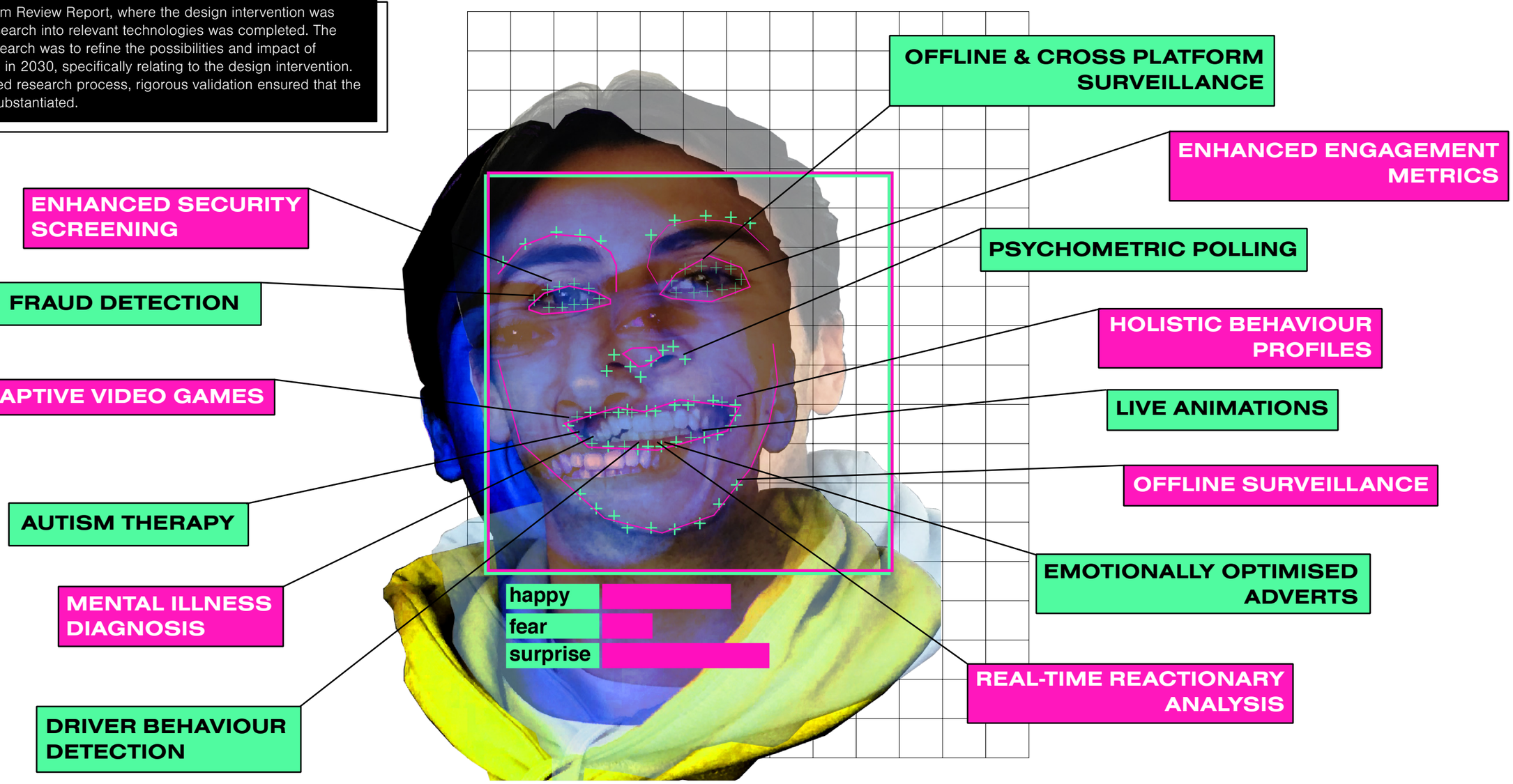
Aims

1. Effectively regulate fake news to ensure consumers are aware of which information is validated. Enable people to assess the information they receive, distinguishing between statements which can and cannot be proven to be true.
2. Provide personalisation of technology without infringing upon data rights or restricting a user to only consuming information from predetermined sources.
3. Encourage people to seek other viewpoints with the intention to understand them, without necessarily accepting them as their own.
4. Empower people to make individual and autonomous informed decisions in a climate of targeted and tailored information streams.
5. Protect vulnerable groups* from personalised persuasive advertising as a result of increased surveillance.

**Vulnerable groups - groups who face distinct disadvantages and need extra protections due to factors such as geographic location, ethnicity, gender, age, ability and citizenship.*

Extended Research

Following the Interim Review Report, where the design intervention was defined, further research into relevant technologies was completed. The purpose of this research was to refine the possibilities and impact of these technologies in 2030, specifically relating to the design intervention. During this extended research process, rigorous validation ensured that the predictions were substantiated.



2030 Possibilities

Predicting the form of social media in 2030 is difficult given the unpredictable popularity of apps and devices. However, key features such as efficient communication through visuals (e.g. memes) and engaging personal storytelling (e.g. the Kardashian family) are likely to be present in a new form. The platform would allow these popular features to be combined. Companies such as Affectiva have developed AI capable of measuring users' emotional responses to content on social media in order to optimise marketing campaigns[1]. By 2030, similar software may become mainstream, analysing text, speech and facial expressions to create actionable data in real-time.

Impact

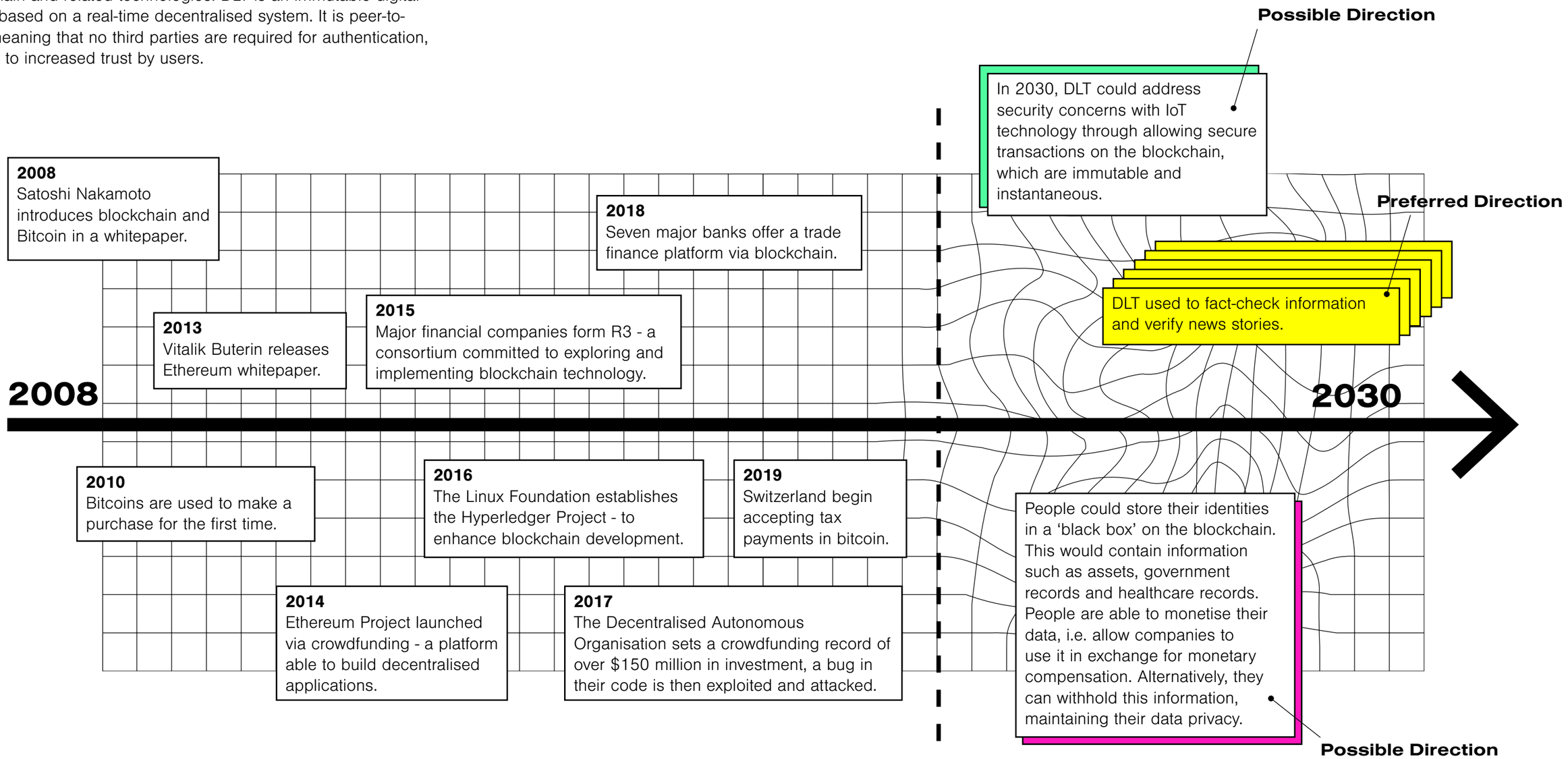
Data from emotion recognition provides huge scope for improvements in microtargeting of products, services and systems and offering dynamic and adaptive experiences for users. However, this data has the potential for misuse through behaviour manipulation. For example, data on location and mood could be used to predict depression, resulting in advertisements for medication or counselling. This poses ethical questions regarding emotional privacy.

Validation

Some psychologists doubt the capabilities of facial recognition due to differences in the individual, context and culture (i.e. the same facial expression between two people does not mean they are feeling the same thing)[2]. However, large corporations such as Microsoft[3], IBM[4] and Amazon[5] are all currently selling facial recognition software, suggesting this field is likely to expand in the future. The smart wearable market size valued at 216.18 million units in 2019 is expected to almost triple by 2025[6] validating the prediction that IoT devices and their data will be more widely used by 2030.

(1) Solutions - Affectiva : Affectiva [Internet]. [cited 2020 Feb 28]. Available from: https://www.affectiva.com/what/products/ ,(2) Emotional Expressions Reconsidered: Challenges to Inferring Emotion From Human Facial Movements – Association for Psychological Science – APS [Internet]. [cited 2020 Feb 28]. Available from: https://www.psychologicalscience.org/publications/emotional-expressions-reconsidered-challenges-to-inferring-emotion-from-human-facial-movements.html ,(3) Facial Recognition | Microsoft Azure [Internet]. [cited 2020 Feb 28]. Available from: https://azure.microsoft.com/en-gb/services/cognitive-services/face/ ,(4) Integrating IBM Intelligent Video Analytics with IBM i2 Facial Recognition [Internet]. [cited 2020 Feb 28]. Available from: https://www.ibm.com/support/knowledgecenter/en/SS8BXH_1.6.1/iva/int_i2frs_intro.html ,(5) Amazon Rekognition – Video and Image - AWS [Internet]. [cited 2020 Feb 28]. Available from: https://aws.amazon.com/rekognition/ ,(6) Smart Wearable Market | Growth, Trends, Forecast (2020-2025) [Internet]. [cited 2020 Feb 28]. Available from: https://www.mordorintelligence.com/industry-reports/smart-wearables-market

Distributed Ledger Technology (DLT) is an umbrella term for blockchain and related technologies. DLT is an immutable digital ledger based on a real-time decentralised system. It is peer-to-peer, meaning that no third parties are required for authentication, leading to increased trust by users.



2030 Possibilities

As a distributed technology, it is more robust than a centralised payment system and transactions can be carried out without a fee. DLT also provides a higher level of privacy when undertaking transactions than most forms of digital payments (which are usually associated with a person, location and time). These attributes make DLT an attractive innovation and it is predicted that 10-20% of global economic infrastructure will be running on blockchain-based systems by 2030[7].

Impact

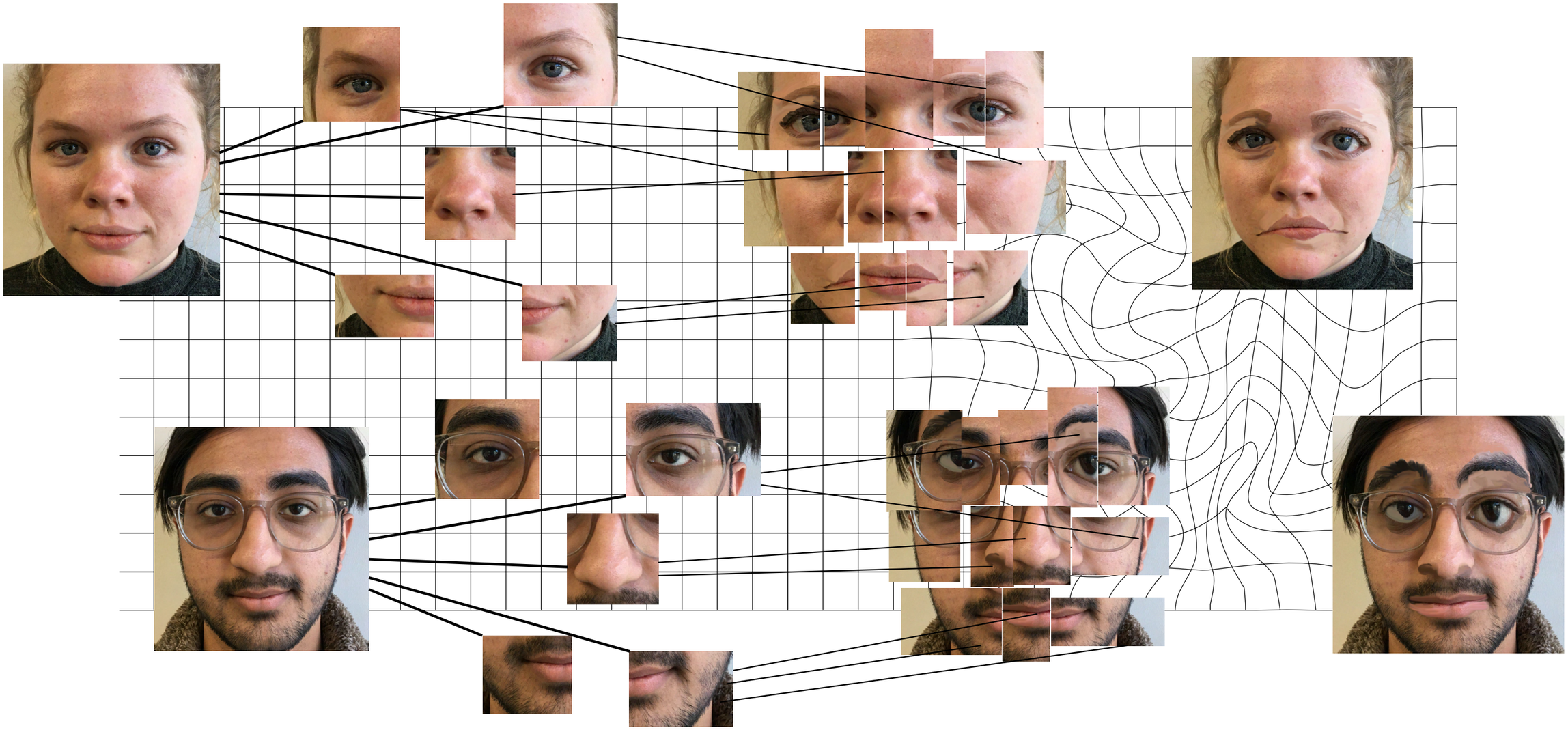
Blockchain technology may advance from cryptocurrencies to support transactions that require personal identification, peer review, elections and other types of democratic decision-making and audit trails.

Validation

Some argue the use of DLT is all hype. However, there has been huge global investment in the technology, including from large corporations such as IBM and Facebook[8]. This suggests that despite some scepticism, it is widely accepted that DLT will be very disruptive and will have widespread implications for society.

(7) PwC Global Blockchain Survey: PwC [Internet]. [cited 2020 Feb 28]. Available from: https://www.pwc.com/gx/en/issues/blockchain/blockchain-in-business.html , (8) - Ferraro P, King C, Shorten R. Distributed Ledger Technology for Smart Cities, the Sharing Economy, and Social Compliance. IEEE Access. 2018;6:62728–46. (Timeline Reference 1) - blockchain-timeline_final.pdf [Internet]. [cited 2020 Mar 5]. Available from: https://www.grantthornton.global/globalassets/1.-member-firms/global/insights/blockchain-hub/blockchain-timeline_final.pdf , (Timeline Reference 2) - History_of_Blockchain_Technology.png (1600×1000) [Internet]. [cited 2020 Mar 5]. Available from: https://101blockchains.com/wp-content/uploads/2018/10/History_of_Blockchain_Technology.png , (Timeline Reference 3) - DL8_9NeX4AMsn8f.jpg (638×359) [Internet]. [cited 2020 Mar 5]. Available from: https://pbs.twimg.com/media/DL8_9NeX4AMsn8f.jpg

Deepfakes are an emerging technology that use automatic algorithms to synthesise fake video and audio. The basic premise is that one person's face can be mapped onto another person's body by splicing existing footage. Deepfakes pose a threat to the validity of video content, a major media format shared online.



2030 Possibilities

By 2030, Deepfakes will be indistinguishable from real videos, even to the trained eye. This is based on developments in artificial neural network (ANN) research as well as improved access to cheap computational power. Free software such as DeepFaceLab and FakeApp are capable of making convincing videos and there are numerous online groups dedicated to sharing techniques to improve their Deepfakes. This will enable anyone to create convincing and misleading videos with little expertise.

Impact

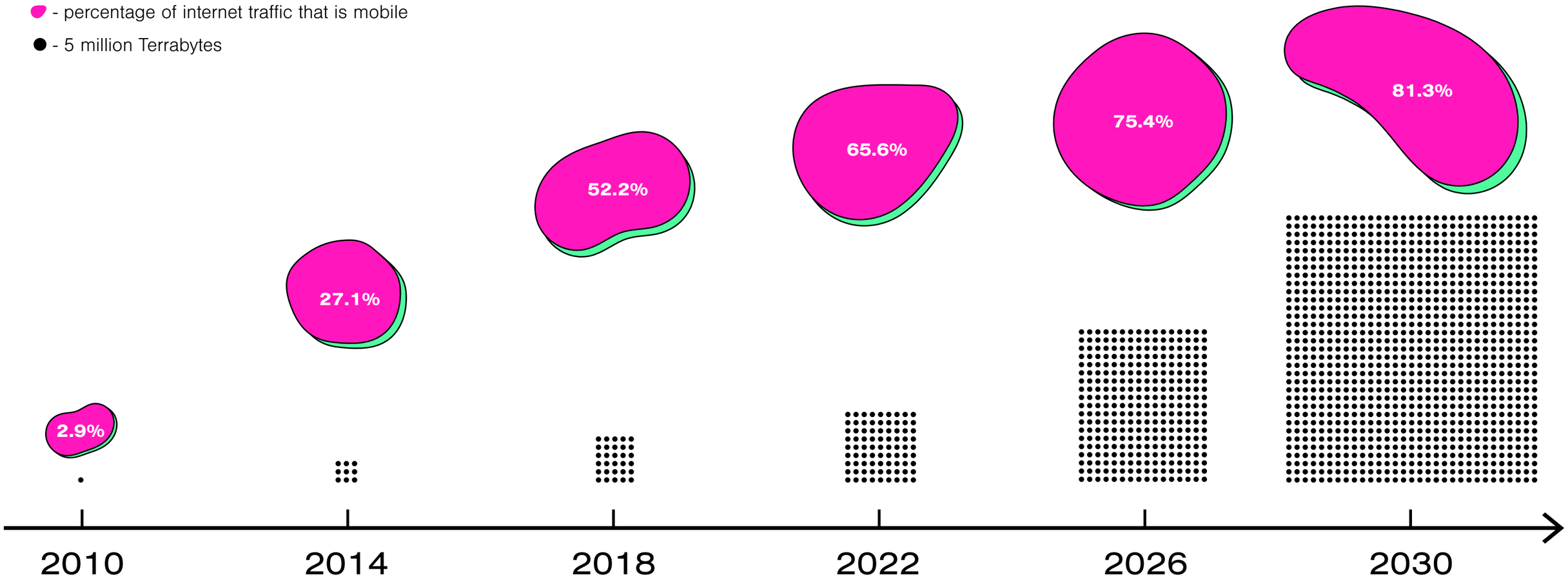
While there are many potential adverse scenarios that Deepfakes enable, one of the most concerning is regarding politics. Deepfakes make it possible to put words in any politician's or influential individual's mouth. This extremely convincing form of false information can have detrimental and irreversible effects on public opinion. People may then also start to doubt real videos of politicians, passing them off as a Deepfakes to support their existing beliefs. This is the concept of the Liars Dividend[9].

Validation

Deepfakes of Boris Johnson and Jeremy Corbyn, created in 2019 by Future Advocacy, highlight the potential threat of this technology. Major governments are developing ways to reduce the impact of Deepfakes on society, evidencing its topicality.
For example, the "DEEP FAKES Accountability Act" by New York's Democratic representative, Yvette Clark, states that watermarks must be placed on altered media and violators would face criminal penalties[10]. This is unlikely to work as an effective solution as it is becoming harder to identify people online with ubiquitous tools such as virtual private networks (VPNs) which circumvent online accountability.

(9) Deepfakes - Future Advocacy [Internet]. [cited 2020 Feb 28]. Available from: https://futureadvocacy.com/deepfakes/ , (10) - Clarke YD. Text - H.R.3230 - 116th Congress (2019-2020): Defending Each and Every Person from False Appearances by Keeping Exploitation Subject to Accountability Act of 2019 [Internet]. 2019 [cited 2020 Mar 5]. Available from: https://www.congress.gov/bill/116th-congress/house-bill/3230/text

5G is the 5th generation mobile network standard that is 100 times faster than current 4G technology. The graph below shows the percentage of internet traffic that is mobile and the amount of data that will be generated by all the users on the internet.



2030 Possibilities

By 2025, 34% of the global population are expected to have access to the 5G network and the number of connections is predicted to reach 2.6 billion[11]. This number is expected to more than double by 2030[12]. Mobile data traffic reached 28 exabytes per month in 2019 and this is predicted to double by 2021[12].

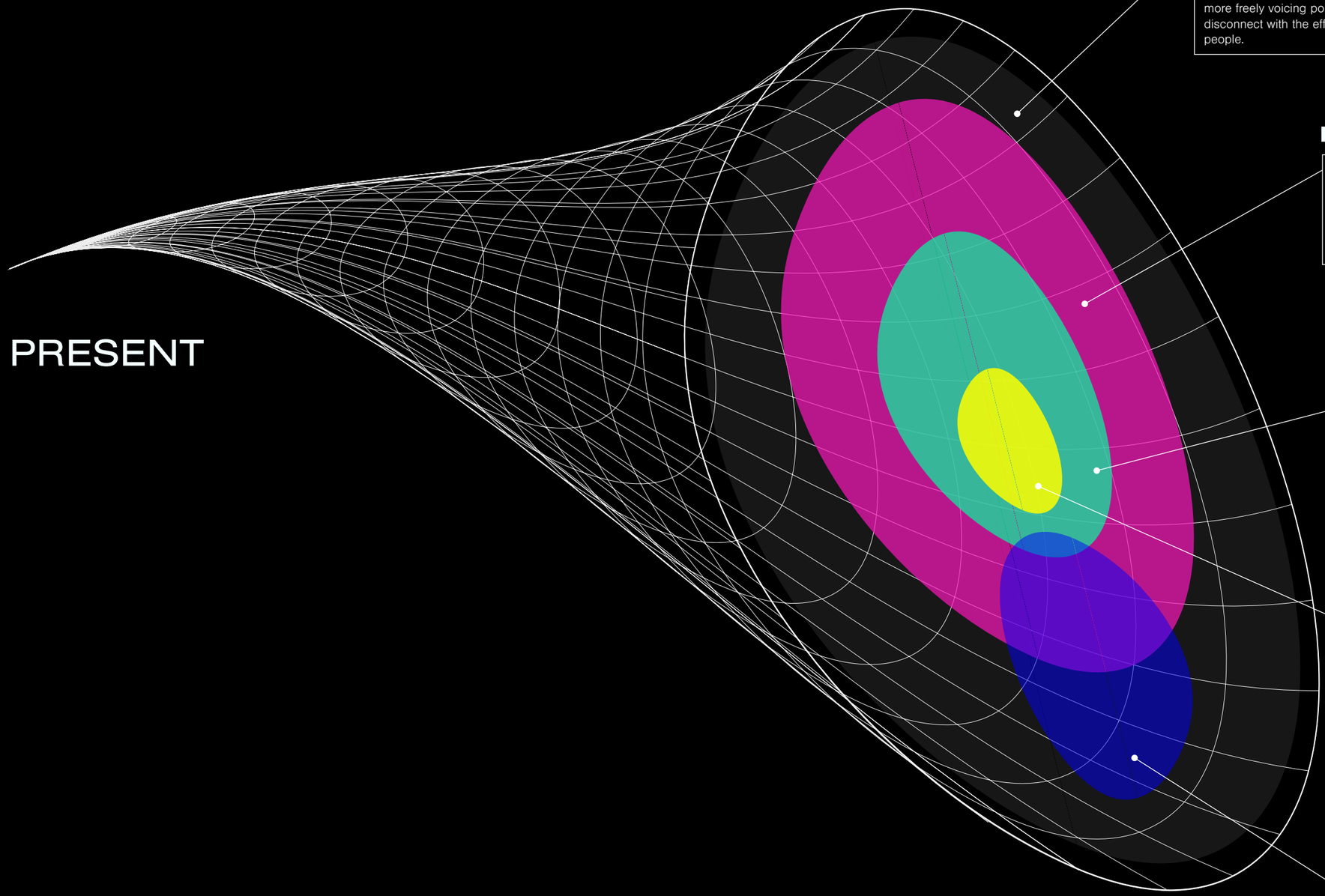
Impact

The increase in both speed and bandwidth will enable an even more rapid spread of information. More people will be able to access large quantities of information and content in new forms like high definition video and AR/VR content. Conclusively, misinformation will be able to propagate more rapidly and in these new forms.

Validation

3G connectivity enabled full internet access on mobile devices, which in turn accelerated content consumption. Video streaming over cellular networks was enabled by the increased speed of 4G. A shift towards on demand content followed. Audio and video generated 89% of mobile network traffic in 2019[13]. The amount of data consumed increases rapidly with increased data speeds. It is safe to assume that more content will be consumed in 2030 due to the adoption of 5G.

11) - CCS Insight [Internet]. CCS Insight. [cited 2020 Mar 5]. Available from: https://www.ccsinsight.com/ , (12) Statista - The Statistics Portal for Market Data, Market Research and Market Studies [Internet]. [cited 2020 Feb 28]. Available from: https://www.statista.com/ , (13) Deloitte | Audit, Consulting, Financial, Risk Management, Tax Services [Internet]. [cited 2020 Feb 28]. Available from: https://www2.deloitte.com/global/en.html



POSSIBLE

All information received online is targeted, further compounding existing opinion and widespread manipulation of opinion by utilising data. Significantly decreased physical social interaction leads to people more freely voicing polarised opinions online and a disconnect with the effect their words have on other people.

PLAUSIBLE

AR and VR form the platform for popular social media and news apps where content is consumed through a fully immersive experience. Fiction and reality are almost indistinguishable. Therefore, voters are irrationally critical of any new information and mostly vote based on news which aligns with their prior judgements.

PROBABLE

Data harvesting is used to a greater extent to target groups in order to sway elections e.g. widespread emotion recognition cameras combined with IoT wearable data lawfully identify people's moods and vulnerabilities. This data influences the media that people are exposed to. Those utilising the data aim to create profit by influencing people's future behaviour.

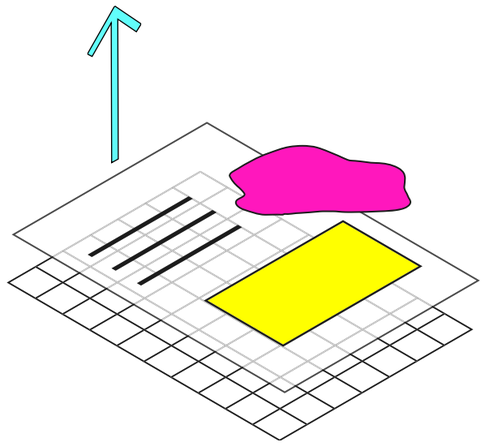
PROJECTED

Targeted advertising increases as a result of increased surveillance and use of personal data. Social media is increasingly popular and is largely unregulated, providing a platform for fictitious stories and videos to spread virally. Wealthier corporations can fully utilise the persuasive power of data. Events similar to the Facebook-Cambridge Analytica data scandal continue to occur.

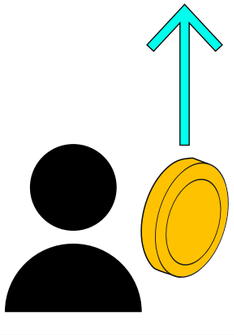
PREFERABLE

Authenticated information is accessible, allowing the electorate to reach voting decisions based on factual news. The presentation of information results in an appreciation of different viewpoints on a topic. Freedom of speech is maintained.

While the intention for the platform to meet the initial five aims remained, the concept needed to be supported by a more tangible back-end system. To help realise this, Robert Shorten, Professor of Cyber-Physical Systems, was consulted to explain the key requirements of a robust system. Below is the first system prototype.



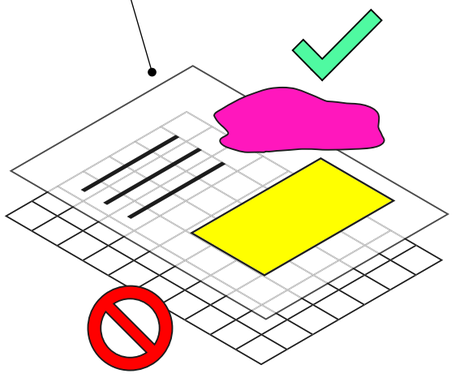
Journalist produces a MediaBite.



Token is purchased and deposited when MediaBite is submitted for publishing.

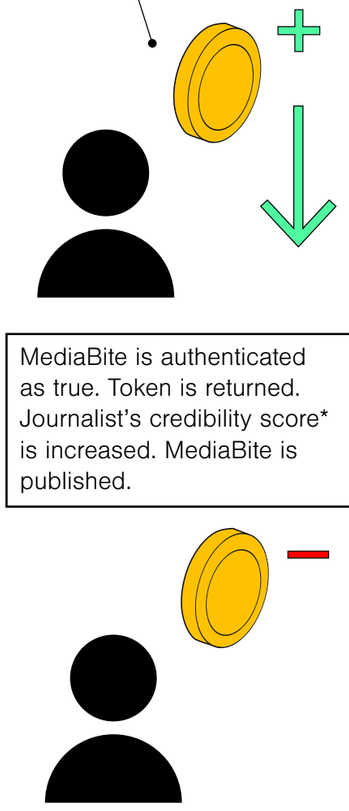
Cost of token is pegged to a currency to avoid issues with price volatility. If there is low compliance in the system, i.e. people post fake news, then the number of tokens that need to be deposited rises - increasing the punishment for non-compliance.

A computer is selected to run the algorithm based on a lottery system. This increases the security of the network as it decentralises fact checking.



MediaBite is checked by algorithm for validity.

A publisher's posts will appear higher in a user's feed, resulting in more readers if they have a better credibility score



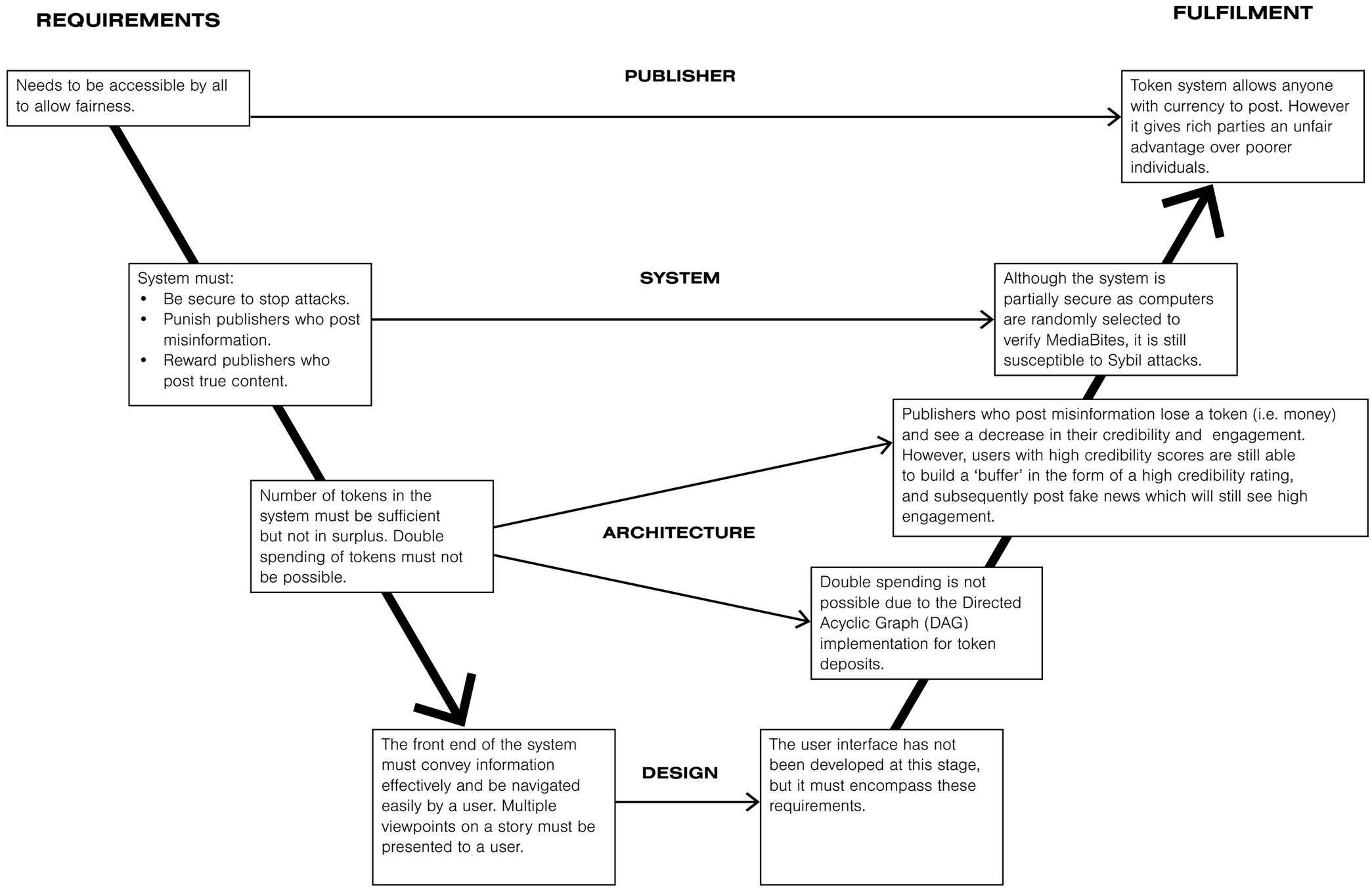
MediaBite is authenticated as true. Token is returned. Journalist's credibility score* is increased. MediaBite is published.

MediaBite is not authenticated as true. Token is withheld as punishment. Credibility score decreases. MediaBite is published.

All MediaBites are published regardless of truthfulness. This is to avoid the system becoming a tool for censorship.

**Each publisher has their own credibility score between 0 and 1, which is affected by whether or not they submit MediaBites containing fake news. A publisher's credibility score is important because it affects the number of user's feeds that their Mediabites will appear on. So a higher score results in more readers.*

The V model was used to verify and validate the initial system by identifying system requirements and testing whether or not they had been met. The traditional V model has been adapted as the system cannot be physically tested given its speculative nature. Instead, the team had another meeting with Prof. Robert Shorten along with two of his PhD students to find flaws in the design. The traditional application of this model is on software engineering projects and is created by Donald Firesmith of Carnegie Mellon[14]. However, it can be useful to verify and validate many different systems at initial development stages.



(14) - Using V Models for Testing [Internet]. [cited 2020 Mar 5]. Available from: https://insights.sei.cmu.edu/sei_blog/2013/11/using-v-models-for-testing.html

EVENT

In July 2030, following another recent tax increase on meat, scientists find traces of a new drug resistant bacteria in produce from the UK's largest artificial meat lab. This results in widespread hysteria and creates headlines across the country. It is assumed that 25 million users exist on the platform.

REACTION

A *"This will be the source of a global health pandemic - the death toll will be huge!"*

B *"We think fake meat must be banned to prevent any further spread of this bacteria"*

C *"Our political manifesto contains a comprehensive set of policies on lab grown meat to mitigate events like this."*

D *"Scientists prove the bacteria is not harmful to humans."*

E *"Tests are shown to have been fabricated by the government in order to detract from the backlash of pulling out of the 2026 International Climate Agreement"*

AUTHENTICATION SYSTEM

Information from each MediaBite which can be authenticated as true or false are identified, as well as matters of opinion which cannot be authenticated. MediaBites containing information found to be false will be pushed down on a user's feed, resulting in a reduced readership.

RESULT

D **Authenticated**

The facts are verified. 10 million users consume this content. The publisher's already high credibility score is marginally increased.

B **Authenticated**

Statements are subjective and cannot be proved or disproved. Their high credibility score causes 500,000 users to consume this content.

C **Authenticated**

No false information identified. Medium-to-low credibility score results in 10,000 users consuming the content.

A **Not Authenticated**

Confirmed as deepfake video. 200 users consume this content. In addition to a series of previous unverified MediaBites, this causes a reduction in credibility score of the publisher so low they are suspended from the platform.

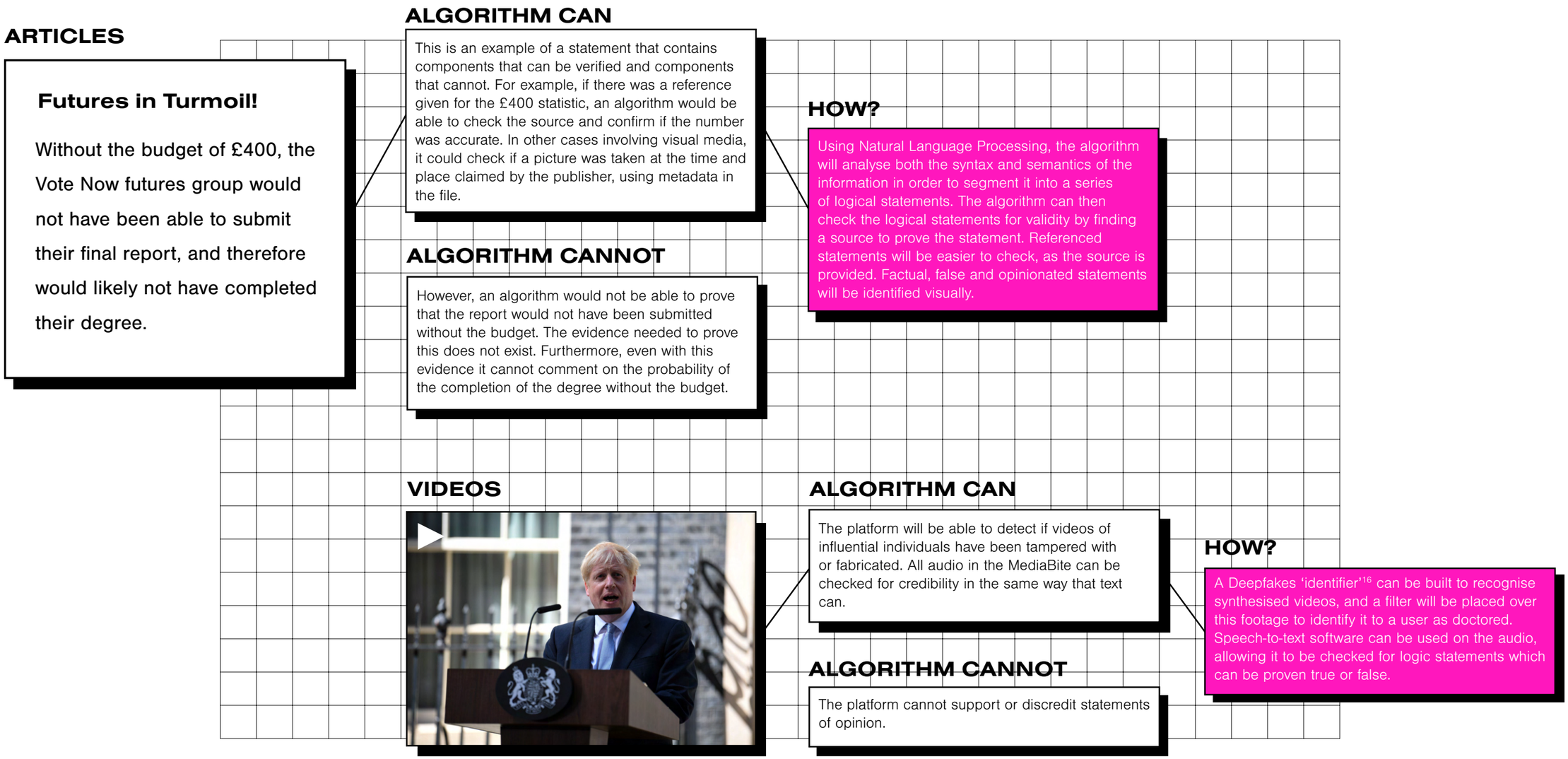
E **Not Authenticated**

Confirmed as deepfake video. 200 users consume this content. In addition to a series of previous unauthenticated MediaBites, this causes a reduction in reputation score of the publisher so low they are banned from the platform.

USER FEED

MediaBites are posted on the platform for users to view. Each user has a personal optimised feed, combining their personal content preferences with content currently proving popular on the platform and breaking stories.

In an ideal future scenario a piece of software would scan a MediaBite and classify whether it is true or not. However, the challenge is not as straightforward as this, given that not all publications contain only content which can be proven to be true or false. Instead, many publications contain subjective statements that cannot be strictly confirmed or denied, but still make implications that inform people's opinions. As a result, the extent to which information could be validated in 2030 has been researched.



Natural Language Processing

The authentication algorithm will use Natural Language Processing to analyse the content of Mediabites and understand it. Early examples of text analysis include spam email detection. NLP was used to learn specific punctuation marks and keywords commonly used in spam emails (such as "free"), to automatically detect and filter them[15].

Present uses of NLP are in virtual assistants such as Amazon's Alexa which converts a users speech into words, sounds and ideas. NLP uses syntactic analysis which derives meaning from words by applying grammatical rules to small groups of them. Semantic analysis is also attempts to understand the meaning conveyed by text, for example by assigning a definition to a word based on the context.

Sentiment analysis tries to interpret and classify opinions and emotion within text. It is already used at scale in customer service, for example to identify if reviews are positive, negative or neutral, or to identify escalating issues on social media in real time. However, there are still limitations to be addressed such as challenges with interpreting irony, sarcasm, tone and the effect of context.

(15) Hinds R. How Natural Language Processing Is Shaping The Future of Communication [Internet]. [cited 2020 Mar 3]. Available from: https://www.affinity.co/blog/natural-language-processing (Image - Boris Johnson) - CNN B <a href="https://twitter com/robpicheta" target="_blank">Rob P and KF. Secretary of State for International Trade and Business Secretary are out [Internet]. CNN. 2019 [cited 2020 Mar 3]. Available from: https://edition.cnn.com/uk/live-news/boris-johnson-prime-minister-wednesday-intl-gbr/h_a7cec8e0c46f6a7a8ade959fb0205d8 , (16) - Vincent J. Deepfake detection algorithms will never be enough [Internet]. The Verge. 2019 [cited 2020 Mar 5]. Available from: https://www.theverge.com/2019/6/27/18715235/deepfake-detection-ai-algorithms-accuracy-will-they-ever-work

Accounts are created using biometric identification to prevent single parties from creating multiple accounts. Each MediaBite must be submitted with biometric authentication to prevent account transfer.

The MediaBite that is sent to the car is encrypted, so that only the sender and the verifier (car) can access the information. This avoids potential bias influencing the verification result of a MediaBite.

Verifying MediaBites must be decentralised from the publishing platform to ensure trust among publishers and users.

CREATION

An account is created and the initial credibility score is assigned to all new publishers.

SUBMISSION

A publisher creates a MediaBite and submits it to be authenticated.

AUTHENTICATION

The authentication algorithm is run on the computer of a randomly selected car. This checks if the MediaBite contains false information. If there is none, it will receive an ‘authenticated.’ marker.

PUBLICATION*

The MediaBite contains no false information and an ‘authenticated’ marker is attached. The publisher’s credibility score is increased accordingly. The system promotes MediaBites from publishers with higher credibility scores to more users, resulting in an increased readership.

The MediaBite contains false information so the ‘authenticated’ marker is not attached. The publisher’s credibility score is decreased accordingly. The system demotes MediaBites from publishers with lower credibility scores, resulting in a decreased readership.

If a publisher’s score falls too low then they will be suspended for a year from the platform.

WHERE DOES THE MONEY COME FROM?

Users pay a fee to subscribe to the platform.

Media outlets pay to use authenticated MediaBites on their platform.

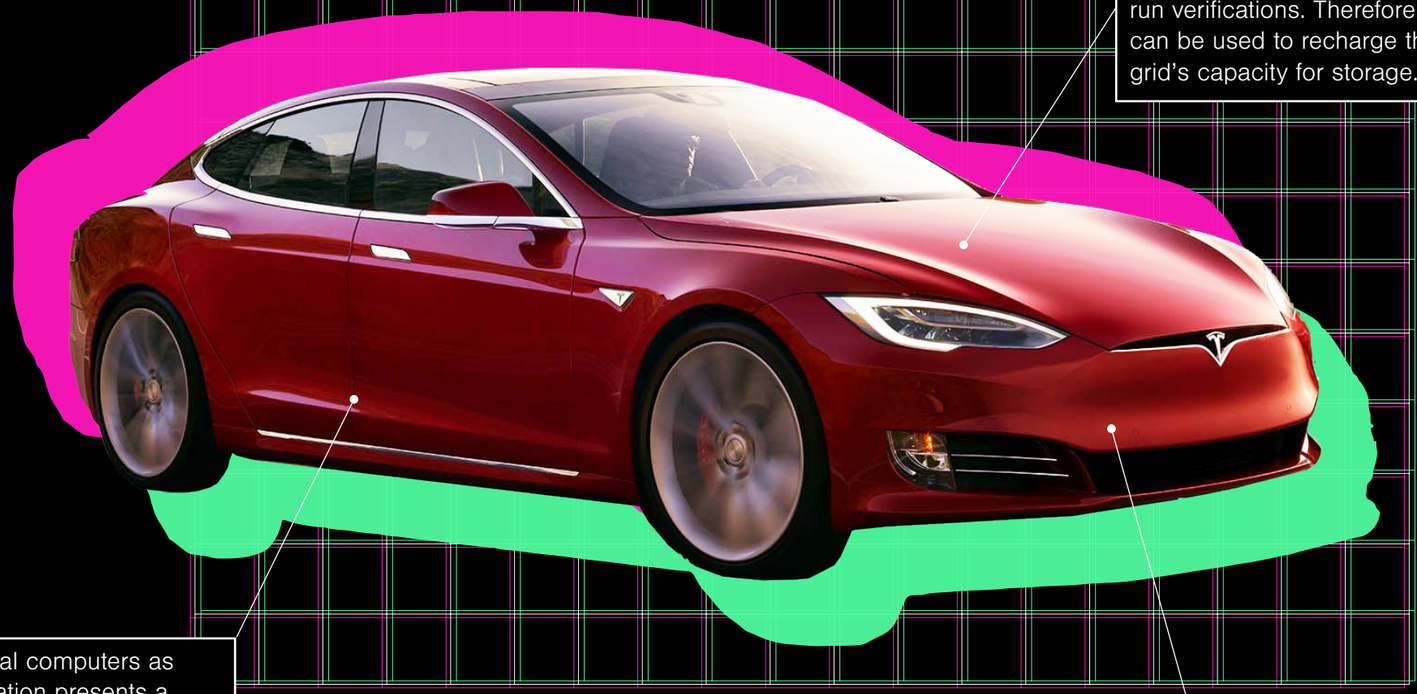
The owner of the car is compensated by the system for their service based on the amount of computation that was necessary for authenticated.

From a business model perspective, the system cannot exist without a monetary input, as car owners need to be compensated for their service.

**All MediaBites are published, regardless of verification decision*

The two main constraints on conventional compact computing are that the devices must be small and able to be cooled efficiently. Cars have ample space to fit a computing device and these devices are likely to further reduce in size, based on current trends. The cooling issue is also solved as heat is able to freely dissipate around the volume of the car.

CARS?



UNDER UTILISED

Cars are parked 95% of the time[17] and can be utilised as secure computers without affecting their performance. There will be a small drain on the battery but the owner of the car is compensated for this. Cars in certain areas of the country can also be used to store excess renewable energy from the grid. If there is a surge of this intermittent energy in a specific region, cars close to this region can be targeted to run verifications. Therefore, surplus renewable energy can be used to recharge their batteries, increasing the grid's capacity for storage.

EXPENSIVE

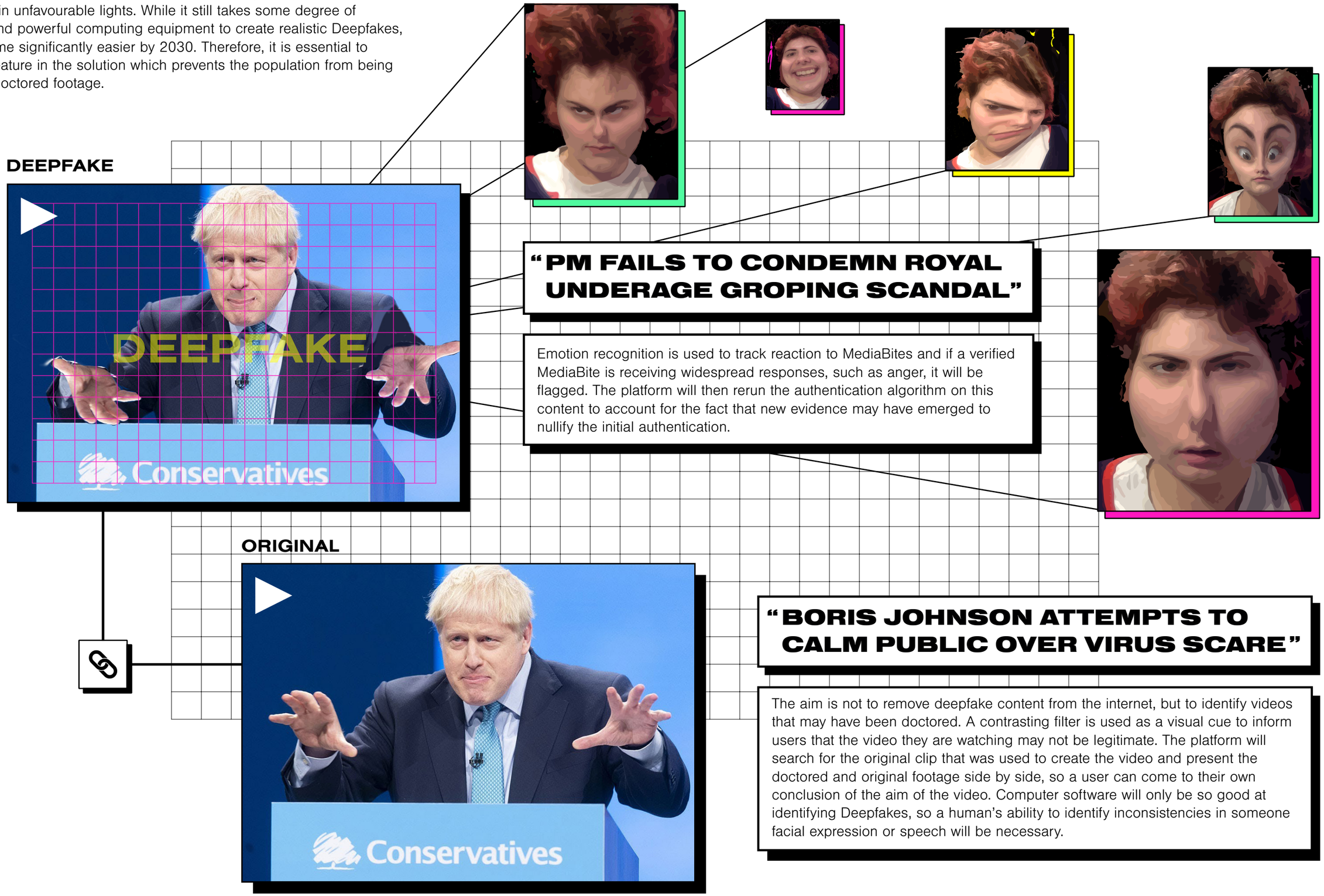
Using cars instead of personal computers as the device performing verification presents a large security advantage; cars are inherently expensive. This removes the main point of failure, Sybil or 51% attacks[18], as it would not be feasible to purchase the necessary number of cars to attack the network.

POWERFUL

The latest Tesla chips run at 2GHz and are capable of performing 36 trillion operations per second[19]. In 2030, these chips will be more powerful and capable of performing computation at a similar (if not higher) level compared to current PCs. This improvement in car computing power is driven by the demand for autonomous cars which perform high levels of computation.

(17) Want to know why Uber and automation really matter? Here's your answer. [Internet]. Fortune. [cited 2020 Mar 3]. Available from: https://fortune.com/2016/03/13/cars-parked-95-percent-of-time/ , (18) lainfinity. Solution to Sybil attacks and 51% attacks in Decentralized Networks [Internet]. Let's Talk Bitcoin. [cited 2020 Mar 3]. Available from: https://letstalkbitcoin.com/blog/post/solution-to-sybil-attacks-and-51-attacks-in-decentralized-networks, (19) Self-Driving Cars Use Crazy Amounts of Power, and It's Becoming a Problem. Wired [Internet]. [cited 2020 Mar 3]; Available from: https://www.wired.com/story/self-driving-cars-power-consumption-nvidia-chip/ ,(Tesla Image) - Model S | Tesla [Internet]. [cited 2020 Mar 3]. Available from: https://www.tesla.com/en_gb/models

As previously validated, Deepfakes can be used to portray influential individuals in unfavourable lights. While it still takes some degree of expertise and powerful computing equipment to create realistic Deepfakes, it will become significantly easier by 2030. Therefore, it is essential to include a feature in the solution which prevents the population from being misled by doctored footage.



(Image - Boris Johnson) - (16) editor HSP. Boris Johnson elected new Tory leader. The Guardian [Internet]. 2019 Jul 23 [cited 2020 Mar 3]; Available from: https://www.theguardian.com/politics/2019/jul/23/boris-johnson-elected-new-tory-leader-prime-minister

Feed: Personal
Feed: Public
Topic: Brexit Negotiations
Topic: Artificial Meat
The Labour leadership contest has exposed new factions in the party
Sajid David
Coronavirus deaths in Iran rise to 54 as global outbreak worsens
Michael Waters
Speculators raise bets against US oil and gas companies
Sajid David
Italy unveils 3.8bn stimulus to tackle coronavirus
Sajid David
Greece blocks 10,000 migrants at Turkey border
Sajid David
Coronavirus crisis: UK cities could be CLOSED to contain virus
Michael Waters
First coronavirus death in the US reported in Washington state
Article information:
Publisher:
Guardian
Author's reputation:
76%
Bias Score:
21%
Verified
Article information:
Publisher:
FT
Author's reputation:
88%
Bias Score:
9%
Verified
Article information:
Publisher:
Vox
Author's reputation:
84%
Bias Score:
13%
Verified
Article information:
Publisher:
Daily Mail
Author's reputation:
74%
Bias Score:
32%
Verified
Articles you've seen on this topic:
Artificial Meat

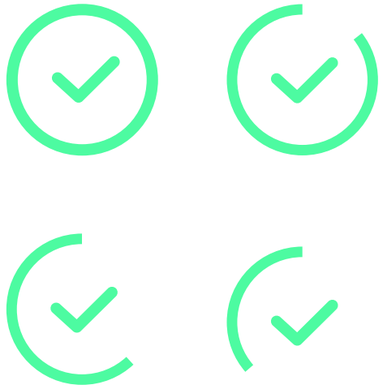
Verification

The platform promotes the truth through positive reinforcements. Authentication takes place when a MediaBite is submitted by a journalist. On a user's feed, MediaBites that contain no false information are presented with a tick. Within a MediaBite, the algorithm will also identify information that can be proved to be true or false, as well as opinionated statements which cannot be proven either way. These three classifications of statements are highlighted to a user when they are viewing the MediaBite, so they are aware of the validity of what they are reading. These classifications will slowly fade over time, to account for new contradictory information that may emerge after the fact, and to show that the authentication was most valid at the time of publishing.

Millions of Americans have been saddled with a total of $1.6 trillion in debt.

Millions of Americans have been saddled with a total of $1.6 trillion in debt.

Millions of Americans have been saddled with a total of $1.6 trillion in debt.



Information proven to be true

Information proven to be false

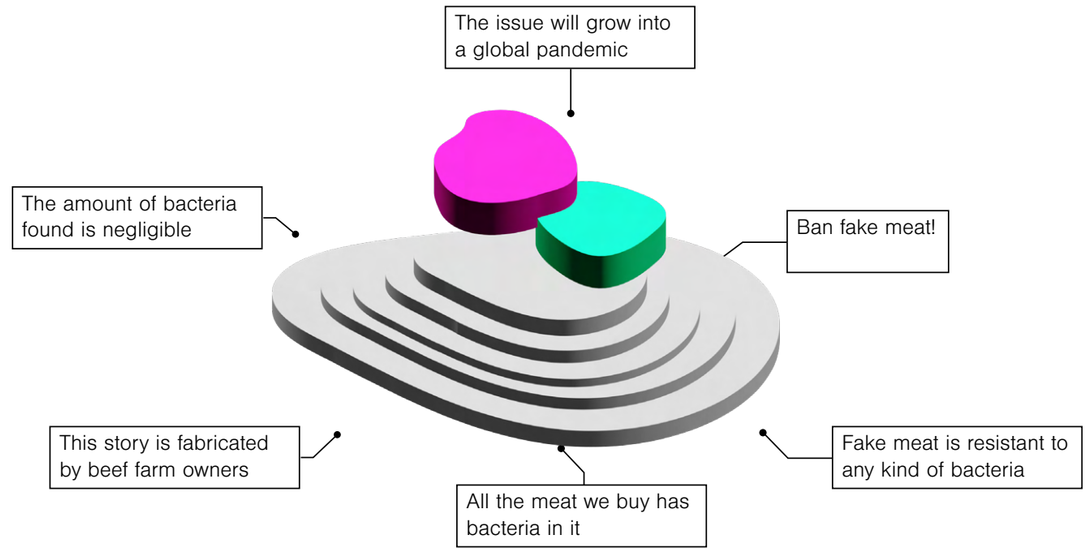
Opinion

Plastic pollution: Snowdon research is a 'wake-up call'

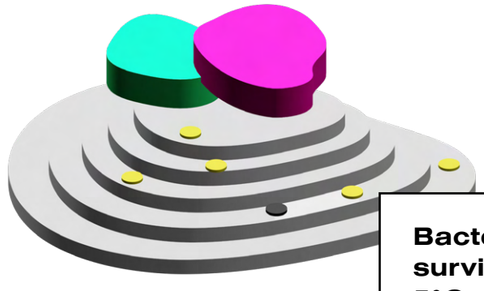
The discovery of microplastic pollution near the top of the highest mountain in Wales is a "scary wake-up call", environmentalists have said. Traces of plastic have been found in samples collected from Llyn Glaslyn - a remote lake near the summit of Snowdon.
A more detailed analysis would almost certainly find more plastic. Microplastics at these levels in drinking water have been proven to cause significant damage to tissue in the human digestive system.

Opinion Map

When a controversial event happens in the world, there will be lots of MediaBites submitted by different journalists. When a user interacts with one MediaBite covering this event they will be presented with an Opinion Map. It combines all of the MediaBites on this event and will contain many different points of view. The MediaBites are assessed using Natural Language Processing and visually grouped by their standpoint, be that positive, negative or neutral.



The user can see a red bubble, which represents the parts of the map they have interacted with, and a green bubble for the areas popular among other users. For some users, this may nudge them to consider interacting with new MediaBites from alternative viewpoints.



- Recommended MediaBites
- Public opinion area
- Personal opinion area

Bacteria found in meat cannot survive temperatures below 5°C

An article that was verified, out of the personal opinion area.

[redacted]

(Snowdon Article) - Plastic pollution: Snowdon research is a 'wake-up call' - BBC News [Internet]. [cited 2020 Mar 3]. Available from: https://www.bbc.co.uk/news/uk-wales-51546978

MediaBites containing no misformation will appear higher on a user's feed than those with unauthenticated content. This arrangement of a user's feed is optimised to prevent the virality of misinformation.

Promoting the Truth

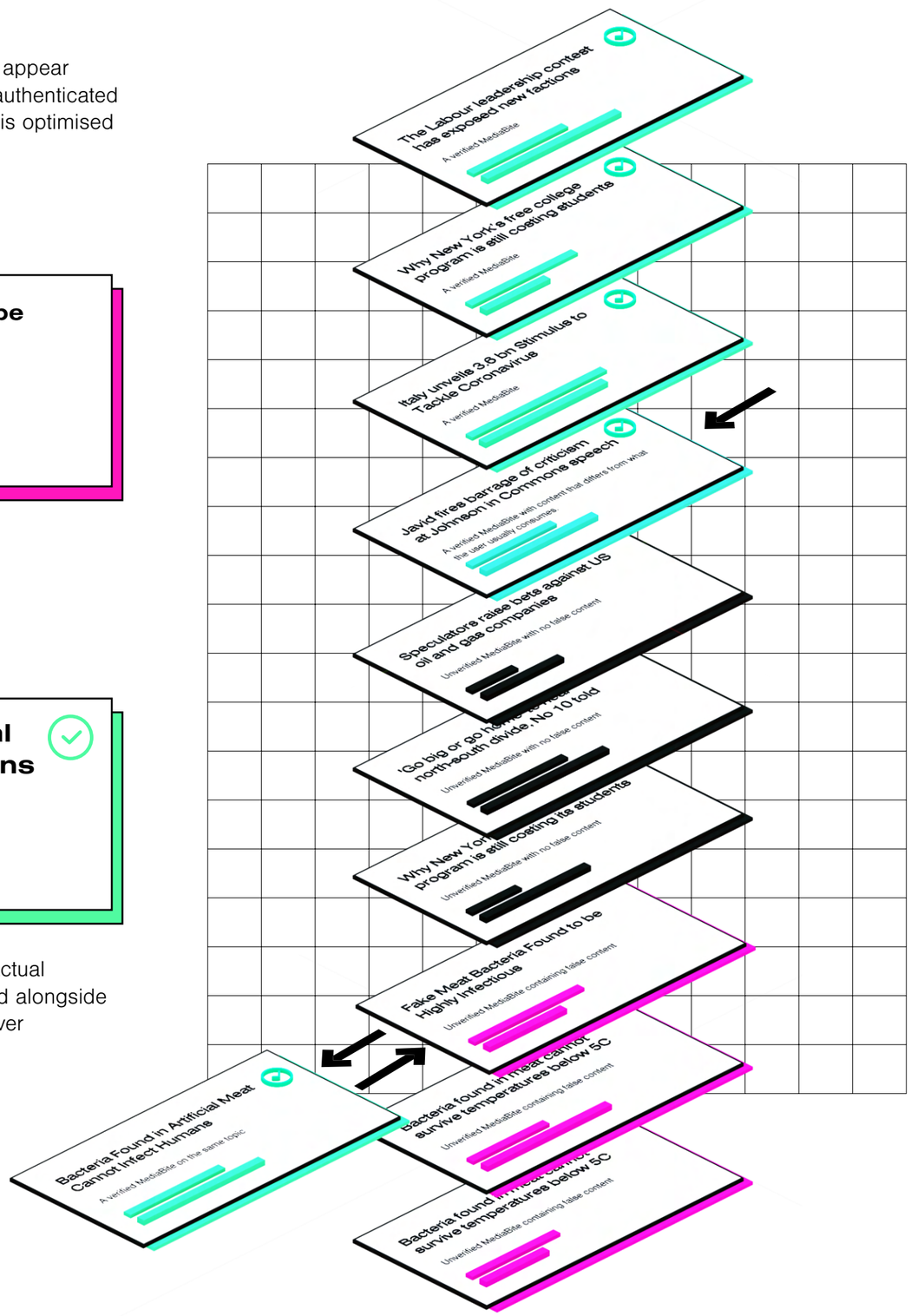
Fake Meat Bacteria Found to be Highly Infectious

A Media Bite containing misinformation.

Bacteria Found in Artificial Meat Cannot Infect Humans

A verified Media Bite on the same topic

If a MediaBite contains misinformation, a factual MediaBite on the same subject is presented alongside it, to ensure that the truth is promulgated over misinformation.



Diversification of content

Javid fires barrage of criticism at Johnson in Commons Speech

Authenticated MediaBite with content that differs from what the user usually consumes.

A user has a personalised feed based on their topical preferences, presented alongside trending MediaBites and breaking stories. However, to ensure users are not confined to a closed loop of information, the platform attempts to return some serendipity to the media.

It is currently hard to stumble upon something online by chance as everything is highly optimised to maximise engagement. Therefore, the platform will also populate a user's feed with a few 'random' stories from topic areas that the user does not regularly engage with.

FUTURE WITH DESIGN INTERVENTION

Jared, 19, Student.

Browsing social media over lunch, the artificial meat story is the first thing he sees. He's intrigued as he only ate a vegan burger from this brand last week and thinks this may affect the availability of lab grown meat in his scheduled food delivery. The authenticated stories put his mind at rest - it appears most of the extreme articles are just scaremongering.

EVENT

In July 2030, following another recent tax increase on meat, scientists find traces of a new drug resistant bacteria in produce from the UK's largest artificial meat lab. This results in widespread hysteria and creates headlines across the country. It is assumed that 25 million users exist on the platform.

AUTHENTICATION SYSTEM

USER FEED

FUTURE WITHOUT DESIGN INTERVENTION

Alex, 34, Office Worker.

Alex hears about the meat scandal and impending pandemic in the canteen at lunch, he is not going to eat artificial meat ever again. There's a state of panic in the office as people reach for their face masks. Videos are spreading of hospitalised consumers of the meat. The Prime Minister has made a statement and attempts to diminish any fears - a ridiculous response considering the scale of this news.

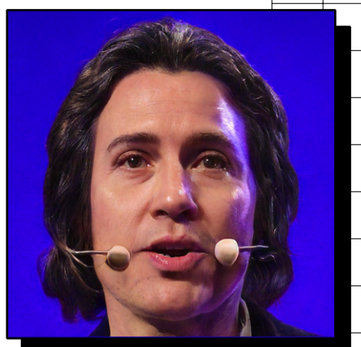
FUTURE WITH DESIGN INTERVENTION

Eve, 67, Retiree.

She sees that the bacteria is proven not to be harmful to humans, but she is disappointed the government she voted for are not taking more precautions anyway - it is only going to get worse as the industry continues to grow. Eve learns about an alternative party's manifesto and their policy on lab grown meat.

(Images) - This Person Does Not Exist [Internet]. [cited 2020 Mar 3]. Available from: https://www.thispersondoesnotexist.com/

A design tool used to assess the robustness of the system was hypothesising scenarios of corporations or individuals attempting to attack it. Platforms that archive large volumes of actionable data, such as Facebook, are a valuable entity making them continual targets of attack. An example of this was the Cambridge Analytica scandal, where millions of individual's personal data was harvested and used without consent[21].



INFLUENTIAL AND WEALTHY INDIVIDUAL / CORPORATION

ATTEMPT 1

Attempt to create numerous accounts in order to flood the platform with fake news, pushing their personal agenda.

OUTCOME

Biometric identification and unique personal information are required to create an account and submit MediaBites. This means each individual can only have a single account, making this approach impractical. Journalists could still be paid to publish MediaBites on the individual's behalf. However, any false information would be unauthenticated and therefore see low engagement. Eventually, these accounts would be suspended from the platform due to extremely low credibility scores.

ATTEMPT 2

Attempt to influence the MediaBite authentication process, by dishonestly biasing the authentication results surrounding a particular topic, to support their personal agenda.

OUTCOME

The decentralised authentication system makes widespread infiltration almost impossible. This is because authentication is carried out on the computer of a randomly selected parked car. The individual would need to purchase many cars to effectively influence the probability that one of their cars would be chosen to run the algorithm. This would be extremely expensive. Furthermore, the MediaBite sent to the car is encrypted so the individual is unable to identify the content they are authenticating.

JOURNALIST WITH LIMITED RESOURCES

ATTEMPT 1

A journalist increasing their credibility score in order to post a MediaBite containing fake news and still receive high engagement.

OUTCOME

Building up a buffer in the form of a high credibility score does not work as changes in credibility score are not linear; there is a maximum score of 1 which cannot be exceeded regardless of how many authenticated MediaBites they publish. Large penalties to already high credibility scores are justified as journalists have the opportunity to run the authentication algorithm themselves before submission, as it is open-source.

(21) - Cadwalladr C, Graham-Harrison E. Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach. The Guardian [Internet]. 2018 Mar 17 [cited 2020 Mar 5]; Available from: https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election , (Images) - This Person Does Not Exist [Internet]. [cited 2020 Mar 3]. Available from: https://www.thispersondoesnotexist.com/

Limitations

1. Unverified news may still be consumed and influence public opinion. The platform will not be used by everyone and some may be sceptical of its efficacy. However, it is assumed that an increase in fake news and information targeting over the next 10 years will result in an increase in perceived value of factual stories. Consequently, the population will actively seek out the truth.

2. A personal computer digitally disguised as a car computer could be used to run a separate algorithm to the one that is provided.

3. Emotion recognition is utilised by the platform. However, some psychologists are not supportive of the the widespread assumption that “a person’s emotions can be readily inferred from his or her facial expressions”, due to differences in the individual, context and culture. Development of computational models will require further research into these areas as well as facial movements. (https://www.psychologicalscience.org/publications/emotional-expressions-reconsidered-challenges-to-inferring-emotion-from-human-facial-movements.html). Users may also not be willing to be subject to emotion recognition. This limits the ability of the platform to gauge the collective reaction to MediaBites.

4. Currently, most machine learning based algorithms take in existing data sets and are trained using supervised learning techniques. This existing input data and any improvements made by a programmer may be biased and it is not certain that this will be improved in the future. To mitigate this, the authentication algorithm’s code will be available to the public for complete transparency.

5. There may be individual cases in which unforeseen circumstances lead to a user’s reputation score being reduced and them being suspended from the platform in error. This prediction is based on current case studies such as speculation on issues with YouTube’s algorithm promoting sensational videos. (https://thenextweb.com/google/2019/06/14/youtube-recommendations-toxic-algorithm-google-ai/)

6. Parts of the system will be refined as a result of initial testing. For example, the assigned initial credibility score, the amount by which this score is increased or decreased after it is assessed by the authentication algorithm and the time for which users are suspended.

Management

All documentation for the project was stored in a group Google Drive folder. This worked well in Term 1 and was continued for Term 2. Working on shared documents allowed for easy collaboration and communication. The overall progress for the project and tasks for each member were recorded in the Management Log spreadsheet. This term, each member had more specific roles, dictated by their strengths e.g. graphics, overall strategy and content writing. This meant that each member knew which areas they were responsible for and allowed for effective time management. Regular group meetings were maintained which ensured clear communication throughout the project.

Link to Google Drive with Management Log included:

https://drive.google.com/drive/folders/1LJhp4WTCRon-k9rB_xEVYXDK-S0g3EXg?usp=sharing

Aims

1. Effectively regulate fake news to ensure consumers are aware of which information is validated. Enable people to assess the information they receive, distinguishing between statements which can and cannot be proven to be true.

2. Provide personalisation of technology without infringing upon data rights or restricting a user to only consuming information from predetermined sources.

3. Encourage people to seek other viewpoints with the intention to understand them, without necessarily accepting them as their own.

4. Empower people to make individual and autonomous informed decisions in a climate of targeted and tailored information streams.

5. Protect vulnerable groups* from personalised persuasive advertising as a result of increased surveillance.

Evaluation

5/5 Appropriate validation markers are present when information is proved to be true or false, as well as highlighting opinionated statements which cannot be proven.

4/5 The platform promotes engagement with multiple viewpoints on a story to break the echo chamber. However, a user’s feed will still be partially personalised based on their topical preferences.

5/5 The opinion map was created to directly address this aim.

4/5

Users will be able to evaluate information more easily, but other free platforms will still be available where targeting and persuasion are easily achieved.

3/5

Protect more vulnerable groups from personalised persuasive advertising as a result of increased surveillance.

Group Contribution

Ella was in charge of overall planning, coordinating the group, creating internal deadlines and scheduling meetings. She contributed written content for the report and presentation and full final report proofread.

Jordan was responsible for overseeing the progression of the main system, organising meetings with Robert Shorten and his PhD students. He also contributed written content for the report and full final report proofread.

Saym was responsible for the construction of the report slides, presentation slides, and the visual style across the project. He also assisted with the exhibition and report content.

Tomas was responsible for constructing and visualising the UI of the platform and the AR demo for the exhibition. He took part in creating the visual style of the project and contributed content for the report.

Esther was responsible for minute taking in meetings, planning and making of the video, complimentary AR filters, exhibition curation, illustrations for the report.