

**Whitepaper:**  
**How to build a Civilizational Social  
Media Platform and Enable 21st Century  
Democracy**

By Matt Chessen  
[matt.chessen@icloud.com](mailto:matt.chessen@icloud.com)

## Table of Contents

<b>I. Executive Summary: CivSocial: a 21st Century Institution of Democracy</b>	<b>2</b>
<b>II. Elements of a Civilizational Social Media Platform</b>	<b>6</b>
By the people, for the people: CivSocial's express purpose is to provide a platform for citizen interaction that is by the people and for the people.	6
Transparency: CivSocial's mission is to generate transparency in all elements of the operation, with reasonable protections for privacy and security.	7
Accountability: CivSocial is committed to re-introducing accountability into public discourse, while preserving the option for anonymous free speech.	7
Democracy: CivSocial's explicit mission is to support the principles and ideals of modern democracy.	8
<b>III. How CivSocial Works</b>	<b>10</b>
The Trust Engine	10
Democratic vs. Authoritarian reputation systems	13
Communities	14
CivSocial for government	16
Community governance	18
Data Escrow Service and Data Marketplace	18
Secure Messaging	18
<b>Appendix I. How we got into this mess: Modern Democracy, the Enlightenment and Disinformation</b>	<b>20</b>
<b>Appendix II: A Few of The Problems with Existing Media</b>	<b>24</b>

# I. Executive Summary: CivSocial: a 21st Century Institution of Democracy

Democratic nations need a social media platform that explicitly supports democratic civilization. Current social media platforms optimize for profit and are not necessarily aligned with democratic values. This results in a wide set of impacts that threaten democratic nations, ranging from rampant, viral disinformation, to political filter bubbles and polarization, to the generation and exposure of sensitive citizen data that can be used to manipulate perceptions and influence behaviors.

**Social media platforms exacerbate threats to democracy** through business models that are not aligned with the interests of democratic nations or their peoples. Social platform **users are the product** and platforms **sell or carelessly expose user data** to entities which use it to influence and manipulate. Terms of service and privacy **settings are not transparent and understandable** to the average person. Algorithms that optimize for engagement **promote sensationalized content** over truth, and their functions are rarely disclosed to users. The same algorithms **facilitate filter bubbles** that exacerbate political polarization and identity politics. Virality is critical to platform business models and this **enables disinformation to spread rapidly**. Anonymity and a lack of accountability **facilitate astroturfed consensus** by trolls, extremist groups and state sponsored agents, enabling malicious actors to **weaponize narratives and divide democratic populations**.

**These problems can not be solved by tactical fixes to the social platforms** because they are inherent to social business models. The entire social media paradigm must be transformed by the creation of a platform that is aligned with the interests of the people who use it, and the values of the democratic nations where they live and work. Democracy needs a new digital institution for the 21st Century that protects its users' data, promotes positive engagement through accountability, considers transparency and openness core goals, and expressly protects and promotes democratic values. Democracy needs a **Civilizational Social network**.

**CivSocial** is a transformational social media platform that expressly supports democracy. CivSocial is a civilizational platform because it isn't just about social interactions - it is designed with the express purpose of fostering civic engagement, incentivizing civil discourse, and bolstering modern democracy. CivSocial accomplishes these goals through four core principles:

**Principle I: CivSocial is by the people, for the people.** CivSocial is a **non-profit platform** that is designed, built, and **directed by the community** of users, for the benefit of the users. **CivSocial users own their own data.** They can choose to license their data through a blockchain based contracting system, and **users earn the revenue.** If they choose to participate in advertising, users own the revenue from any engagement they generate. CivSocial takes a cut of this revenue to pay for operating expenses, but **CivSocial never owns, trades, sells, or uses user data** without their express consent. **CivSocial is an open source platform.** Users can build additional functionality and customize their experience, and **external applications that support community values are welcome** and can be integrated into the experience using flexible APIs.

**Principle II: Accountability.** CivSocial promotes accountability through a **Trust Engine** that uses verified identities and a Reputation Scoring system to ensure civil interactions and elevate great ideas and expertise. CivSocial encourages users to have a **verified identity** which is disclosed to other users. The Trust Engine allows users to **build Influence in online communities** where they have expertise through the power of their ideas and their adherence to community standards. **Influence is not just an idea**, it is a quantified score that determines how much impact users can have on ideas, communities and other users Reputation Scores. Poor behavior reduces user influence, undermining the impact of trolls and extremists. Unverified, anonymous users can have a voice, but have little Influence. Similarly, bots are allowed, but must be disclosed as bots, owners must be verified and disclosed, and bots have zero Influence. This **Reputation Scoring** system, combined with norms against social engineering and rigorous policing of manipulation, solves the problems inherent to crowdsourcing credibility indicators and promotes a civil environment for democratic discourse.

**Principle 3: Transparency.** CivSocial operates on the principle that **everything should be open and transparent to the user community** to the maximum extent possible. CivSocial is **open and auditable** from top to bottom. Organization operations, strategy, and financial decisions are **directed by the user community**. Terms of service are determined by the community and **disclosed in plain language**. **Privacy settings are active by default** and modifications are understandable by the average user. **All algorithms must disclose what they are optimizing for and are user customizable**. External applications must comply with CivSocial community standards and will be subject to **regular data audits** and stiff penalties for violations. And **user data is portable**. Everything from Reputation Scores to networks of contacts are exportable and easily used by whatever websites and applications are useful to the user base.

**Principle 4: Democracy.** CivSocial's explicit mission is to **promote and support democratic nations, peoples, and values, as determined by the user community**. CivSocial can be used for hobbies and social interactions, but it is also designed to be a trusted platform where **citizens and government can work on complex problems together**. CivSocial is designed to give citizens a **voice in government policy and programs**, and to make it easy for elected representatives and government officials to **interact with the people in an efficient, civil, trusted environment**. Communities of interest are the heart of CivSocial and they drive its utility as a **platform for communication, knowledge building, and civic activism**. Users are rewarded and promoted in communities based on their **contributions and adherence to community norms**, while malicious actors are disincentivized and demoted. And CivSocial is truly a platform for democracy: its **open source environment** enables other pro-democracy tools and applications to seamlessly integrate and share functionality and users.

**Why CivSocial and why now?** The Founding Fathers of the United States created modern democracy based on the Enlightenment principles of the search for truth through reason. These principles underpin all rule of law systems and are critical for everything from evidence based trials to due process. These principles are under threat because modern digital tools hyper-empower individuals and groups, but have not introduced effective features for determining

truth or enabling accountability and responsibility. Our democratic institutions and their checks and balances were developed during the industrial age and are based on rule by elites and elected representatives. They weren't designed for an age of individual hyper-empowerment and aren't able to quickly or effectively adapt to the digital age. We need a new democratic institution for our digital lives. **Social media is not an institution of democracy, and right now it's all we have.** We need to evolve this paradigm and create a digital communication, information, and collaboration institution that is by the people, and for the people. Democracy needs a Civilizational Social Platform.

## II. Elements of a Civilizational Social Media Platform

CivSocial is a collection of applications and tools, but more than that is a set of democratic values build into a technology platform that is accessible by anyone or any application sharing those values. There are many, many tools available or under development that are designed to solve various aspects of the many challenges to democracy. What is missing is a framework for integrating these tools into a coherent platform that allows portability across applications for verified, trusted users and their data.

**By the people, for the people:** CivSocial's express purpose is to provide a platform for citizen interaction that is by the people and for the people.

- **Users own their data and any revenues from it:** CivSocial users own their data. User data is a treasured and highly protected resource that users should be empowered to use as they wish. Users can license their data to the platform or other applications for specified periods of time through a blockchain-based smart-contract system. Revenues from that data belong to the user. Similarly, any revenues gained through ad engagement, licensing of reputation scores, or other voluntary, user-driven revenues belong to the user. CivSocial only asks for a small fraction of these revenues to help pay for operating costs.
- **CivSocial is a non-profit organization:** CivSocial considers its users to be the heart and soul of the platform. CivSocial users are not a product for the platform to monetize. CivSocial funds operations through donations and by withholding a fraction of the revenue users earn on the platform. This avoids all of the negative incentives associated with for-profit social platforms, and allows CivSocial to operate in the best interests of the users, not shareholders.
- **CivSocial is run by the users:** CivSocial will be guided and governed by the community of users. Decisions on community rules, revenue generation, partnerships, and all other aspects of the platform will be determined by the users.

**Transparency:** CivSocial's mission is to generate transparency in all elements of the operation, with reasonable protections for privacy and security.

- **CivSocial discloses algorithmic optimizations:** CivSocial discloses what all algorithms are optimizing for, and whenever possible, provides users a choice so they can customize that optimization. The open source nature of the platform enables users to create their own optimization algorithms and make them available to the community. So if users want to optimize for engagement, they can. But if they want to optimize for truth, content from family members, or just information that is likely to make them happy, they can do so.
- **CivSocial discloses information in plain, easy to understand language:** CivSocial terms of service are in plain language, not legalese. Privacy controls are simple, easy to understand, and easy to manage. The average user can easily understand the choices they are making when they make them.
- **CivSocial's operations are open and auditable:** CivSocial's mission is to be an open-source platform in every respect. Users can develop new tools and capabilities for the platform. External applications that bolster democracy can interface with the platform through open APIs. CivSocial's operations and finances are open and disclosed to the community to the maximum extent possible. Independent auditing of finances, operations, and algorithms is built into the organizational DNA. Partnerships with academic organizations will be established to ensure researchers have access to data they can use to generate knowledge and hold the platform accountable.

**Accountability:** CivSocial is committed to re-introducing accountability into public discourse, while preserving the option for anonymous free speech.

- **CivSocial users are accountable for their actions online:** CivSocial operates using a Trust Engine where the concept of "Influence" determines how much weight your opinion carries on issues. Users acknowledged as



experts in a topic gain Influence in that topic through acknowledgements by the community. Users who don't conduct themselves according to community standards - for example, by trolling or posting hate speech - will be docked by other users and their Influence will decline. Users who uphold community standards will be acknowledged by other users, and their influence will grow.

- **CivSocial user history is long, but fades over time:** CivSocial encourages accountability by maintaining a long history of user actions on the platform. Users can't break community standards without consequences. But actions, both good and bad, fade over time. The good deeds of the distant past don't necessarily provide influence in the present, and the misdeeds of the past are not irredeemable.
- **CivSocial users know who is a real person and who isn't:** CivSocial users are strongly encouraged to verify their true identity. Only fully verified users carry full Influence on the platform. Anonymous users are permitted - after all, anonymity is an important element of free speech - but they have little influence in the Trust Engine. The power of their words alone must carry the day. Bots are permitted - since they can serve positive functions, like connecting and breaking down filter bubbles - but the identity of the bot and its owner must be fully disclosed, and bots carry zero Influence. These measures are designed to avoid the social influence engineering that plagues social platforms.

**Democracy:** CivSocial's explicit mission is to support the principles and ideals of modern democracy.

- **CivSocial communities support democratic engagement:** CivSocial is explicitly based on the idea that accountable and involved communities are the bedrock of democracy. Many of the problems we now see with citizens feeling disenfranchised and detached from government, toxic online users, and government difficulties solving complex problems, can be helped with strong communities of citizens working towards common goals.
- **CivSocial communities are designed to expose the best ideas and experts:** CivSocial is about contributing your knowledge, expertise, time, empathy, and love to the communities you care about. The structure of these communities is designed to surface the ideas and people who can

make a difference. Leadership in these communities isn't based on where you went to school, where you work, who you know, or who your parents are - it's based on the power of your words, ideas, and actions.

- **CivSocial will uphold democratic values, always:** CivSocial will never compromise its values. CivSocial will never sacrifice its core principals for market access, greater profitability, or any other goal that isn't fully aligned with user interests. CivSocial is designed for democracy-supporting publics in democratic nations. If authoritarian regimes like China and Russia want to ban CivSocial because it doesn't play by their rules, so be it. We will strive to make the platform available for democracy-minded activists in any country. But the core of our focus is bolstering democracy in existing democratic nations. CivSocial is an institution of democracy for the 21st Century.

### III. How CivSocial Works

#### **The Trust Engine**

During the 20th Century, the United States had a functioning collective intelligence system for determining truth from fiction. A combination of national news channels, national newspapers, a rich local media ecosystem, academia, government and the church enabled Americans to determine truth. Those institutions have been undermined and malicious actors have stepped into the gap, using new technologies and breakthroughs from cognitive psychology to craft monolithic collective intelligence systems designed to manipulate perceptions and influence behavior. Social platforms are a primary battleground for these malicious actors, because these platforms are designed to optimize engagement, not trust.

The Trust Engine is at the core of what makes CivSocial effective. It is also the hardest element of CivSocial to implement. Here is how it works.

**The Trust Engine uses verified identities and a reputation system to expose the best ideas, knowledge, and expertise.**

CivSocial users are expected to have a verified identity. This verification would be performed through validation of a government issued identification, or through a mix of other factors, such as financial information. Verification qualifies the user as a full member of the CivSocial community.

Non-verified and anonymous members are limited in their participation in the site. They can not form communities and their influence in the reputation system is extremely limited. Anonymous members who break community rules or who have reputation scores below a certain threshold will be banned from the platform.

Bots are allowed on the site, but with significant restrictions. Bots must be identified as such on their profile and in all posts. Their owner must have a verified identity and that identity must be disclosed. Bots are permitted because they can have significant value in sharing information among community

members and in breaking down informational barriers that trap people in filter bubbles.

The reputation system is the crux of the trust engine. It is a reputation score system where acknowledgements and dockings (roughly analogous to upvotes and downvotes) are weighted based on the user who is providing the feedback. This weighting is the missing element in most crowdsourced reputation scores, which do not distinguish between different types of users and treat all votes the same. CivSocial acknowledges that some users are experts on some topics - and their opinion should count more in those communities - but where they are novices, their opinion should count less.

**Example:** Linda has a PhD in international security, has worked in international affairs for twenty years, and is well regarded in the CivSocial foreign policy community, with a high influence score. John is an accomplished pediatric nurse, with fifteen years experience and a high influence score in the pediatric care community. Linda's opinion should count for more than John's when she provides an opinion related to foreign policy, but her opinion counts far less than John's when she provides opinions on pediatric issues, where she has a low/average influence score.

Within communities, experts can be determined in different ways. Some communities have curated groups of experts when the community is formed. Users can move into or out of the experts group based on other user feedback. Other communities start with no experts and the crowd determines expertise over time. Over time, expertise is always determined by merit, not by extrinsic factors.

The trust engine will be built to undermine the ability of trolls, special interests, or bots to manipulate communities and drive their pet ideas or people into the circle of expertise. Obviously these malicious actors will have an incentive to manipulate the system in order to influence user behavior. Rigorous and continuous testing and tweaking of the trust engine will be needed to prevent malicious users from gaming the system. AI tools will be used to identify suspicious patterns of behavior. And identified instances of manipulation - which violate CivSocial community standards - will be dealt with by lowering the influence scores of the users involved. Since users are tied to a verified and

singular identity, they will have a disincentive to behave poorly on the platform, since this behavior will be punished by the community with a loss of influence.

Coordinated manipulation of users is a serious offense that can result in indefinite or permanent bans from the platform, although typically redemption is possible after a period of time.

The Trust Engine will also provide users fidelity not just on how another user or piece of content is rated, but why it is rated as such. Initial iterations would display the characteristics of users who rate the person or content highly or poorly. For example, a piece of content might be highly rated overall, but it is rated poorly by a sub-community with characteristics similar to the user.

The Trust Engine also has APIs so that other credibility indicator tools can be integrated into its functionality. Default tools will be determined by the CivSocial community and additional tools will be available at the discretion of the user. These tools can be used to fight misinformation and disinformation and promote a more informed user base. One major problem with the current crop of counter-disinformation tools is that they don't have a unified platform where the public can access them all. CivSocial would provide that platform where the tools could be made available to users based on agreed open source standards.

The reputation indicators of the trust engine may be valuable to other organizations or companies. For example, an insurance company may find that a high influence score is correlated with safer drivers, or healthier behavior, and would be willing to offer lower premiums to users who disclose their score. Users could license access to their reputation score to outside organizations. And a user's reputation score would have cross-platform portability. So if an online dating site wanted to indicate reputation scores for CivSocial users, or integrate their own reputational functionality which would affect the user's CivSocial reputation score, this would be facilitated through an open-source standard and APIs.

The trust engine will be tweaked and developed over time so that it optimizes for identifying the best experts and ideas. Future improvements may include:

- Network effects: a users' network might determine a portion of their score. So rather than "friend" everyone, users will need to be thoughtful about who they connect to in their network. Additional enhancements may require them to disclose the strength of their relationships, and weight network effects based on the strength of those connections.
- Rating products and services: the online rating industry is a mess of low-integrity platforms, fake reviews, paid reviews, and other factors that detract from honest, transparent reviews. The Trust Engine could be used as a review platform. And the reputation score could be licensed to outside entities (like Amazon) to increase the trust in their own reviews and eliminate fraud.

### **Democratic vs. Authoritarian reputation systems**

The community of democracies must develop a viable alternative to the Chinese Social Credit System before it becomes a standard for much of the world. In 2014, China announced the creation of a Social Credit System which will rate citizens on their loyalty to the Chinese government and to Chinese brands. Citizens with high scores would gain access to better educational opportunities and jobs, obtain faster processing of bureaucratic paperwork, have more travel permissions, and other benefits. Low score citizens would have fewer opportunities, more red tape, and slower Internet speeds. Techniques from computational propaganda and new AI tools could be used to subtly shape the information environment of low-score citizens, manipulating their behavior so they raise their scores of their own accord.

The Chinese Social Credit System is ostensibly designed to increase public morality and provide an alternative to Western-style credit scores for determining willingness to repay loans. But the system is undeniably a tool of China's social control system, and will likely be fully integrated into the sophisticated and omnipresent surveillance network China is constructing across the country. This system will provide the Chinese Communist Party near total information awareness about its citizens actions and movements minute to minute, and will provide unprecedented powers to influence citizen perceptions and manipulate their behavior.

China is already exporting advanced surveillance technology to the developing world, and almost certainly will export its reputation system as an essential tool of the surveillance state and social control system. China's goals under the One Belt, One Road Initiative are to extend its economic influence and fold Eurasia under its economic umbrella. It is easy to envision China exporting its the Social Credit System in conjunction with infrastructure or other economic development projects. China would fold new countries into its Social Credit System, providing those countries new tools of social control, while China manages the system infrastructure and, by necessity, access to user data.

The risks from allowing the Chinese Social Credit System to develop without a democratic alternative should not be underestimated. If the Chinese system becomes the default standard for the developing world, it would provide radically enhanced capabilities for authoritarian governments to control their populations, and could provide China unimaginable access to extremely sensitive data on a large swath of the world's population. And when combined with sophisticated computational propaganda techniques, that data would give China the ability to influence perceptions and manipulate behavior on a scale never before seen in human history.

We must develop a democratic alternative to the Chinese Social Credit System. CivSocial can serve as that open, democratic, transparent, accountable, people-centered alternative.

## **Communities**

Communities are the focus for user interactions on CivSocial. Communities can be organized around any topic - climate change, abortion, windsurfing, Elon Musk - all are valid subjects for communities. But unlike traditional social media, CivSocial communities are focused on surfacing the best experts and ideas from within those communities, and generating collective action.

Communities are organized with rings of expertise. In the outer ring are the general public or novice community members. Novices are users who have average influence scores in the Community topic. As users generate content that is valuable to the community, other users acknowledge their efforts and their influence score grows. With a higher influence score, the novice can move into

the middle ring of the community, and they become an intermediary. Their role is to help identify quality ideas and experts from the novice group and help move their ideas upwards to the inner ring, where we find the community experts.

Merit driven expertise is the heart of every community. This inner ring can converse with each other, and novices and intermediates can view these interactions. But unlike a traditional social media platform where anyone can interact with anyone and conversations open to the public usually disintegrate into trolling and flame wars, CivSocial experts are shielded from external conversations. They don't have to participate in the wider discussion unless they so desire. In this way, community experts can interact, discuss and refine ideas without fear that their interactions will be interrupted by basic questions, trolling, flaming, distractions, manipulations, and other evils that plague social platform communities.

The circle of experts should not become the new entrenched hierarchy. So the Trust Engine will be designed so that novices and intermediates from the outer rings can move into the experts circle if their ideas are excellent and their contributions to the community are significant. This will prevent a community of experts with a fixed view of the world from disregarding new, transformational ideas that a large body of novices or intermediaries acknowledge. Similarly, experts can not attain their position and then fail to contribute to the community. Over time, their past acknowledgements age and they can fall out of the experts circle if their contributions wane. These parameters for the Trust Engine can be modified to come extent within communities based on their desires. And if sub-groups within a community determine that they need to go their separate ways, they can fork the community and create their own community with their own standards.

Similarly, ideas can move from the novice ring into the experts circle through several means. A large number of novices and intermediaries could endorse the idea, promoting it for expert attention. Or an expert may notice a great idea from the novices, and because their opinion carries greater weight, they could promote it into the circle of experts. Similarly, half-baked ideas by experts could be critiqued by the public. If a large number of intermediates or novices provided



evidence why an idea wasn't feasible, it would fall out of the experts circle. In this way, experts can't promote ideas that have poor evidentiary weight.

## **CivSocial for government**

CivSocial is designed as a platform for enabling 21st Century government. CivSocial communities are designed to facilitate community participation in the business of government, and to enable government officials to interact with the public and efficiently source the best knowledge and ideas for government policies and programs.

There are three major obstacles to government crowdsourcing inputs from the public: 1) there is no platform specifically designed for government to collaborate with citizens; 2) traditional platforms are low-trust and government officials can't determine the identity and credibility of people they interact with; and 3) government officials don't have time to engage with a mass of users; they need to engage with a small group of expert advisers who can distill the gems of knowledge for their use.

In CivSocial communities, the expectation is that the experts circle will consist of a group of global experts on that topic. They are responsible for curating the knowledge and ideas in their community. If the community is one focused on assisting with the work of government, the experts circle forms a sort of advisory council. Currently, government officials use these advisory councils but they are often offline, membership is determined based on fame, knowing someone important, or other non-meritocratic factors, and they don't have easy access to a wide network of interested citizens looking to shape policy and programs.

CivSocial communities can provide government officials with an expert circle selected based on merit, who have access to the best knowledge and ideas curated by intermediaries and novices who could number in the tens of thousands.

How would this work? A U.S. senator might create a CivSocial community for their constituents to provide them policy advice and expertise. The senator could designate members of their staff and selected members of the local community to

serve as the initial experts circle. These experts would pose questions or make requests for inputs from the senator's constituents. Since user identities are verified, the senator could be reasonably sure that the community members are in fact constituents that live in the senator's district. As those constituents make contributions to the community, and other constituents acknowledge their expertise or ideas, a community of intermediaries develops and new experts join the circle from the public. This circle of experts sources inputs from across the breadth of the senator's constituents and distills the best ideas and evidence down into concise materials that are useful for the senator as they conduct Congressional business.

The senator could also use their constituent community for purposes other than sourcing and distilling policy inputs. They could hold virtual town hall meetings. Verified user identities would give them confidence that participants are actually constituents, and they would have a reasonable expectation of civility in the discussion due to the reputation score. The senator could engage a community to assist in information gathering related to their oversight of the federal government. The senator could use the community to seek feedback on the performance of government programs in their district. Or the senator could use the community to engage constituents in assisting programs in their district. There are numerous possibilities for enhancing the ability of elected officials to effectively represent their constituents.

This type of citizen-government engagement is critical for 21st Century governance. Problems have become far too complex for stovepiped, industrial age Departments and Agencies to tackle without the public's help. Studies have shown that the public is increasingly disillusioned with democracy. A significant part of this can be attributed to citizens believing that they have almost zero say in policy beyond their vote.

### **Community governance**

CivSocial's community focus is not just about human communication and connection, it also drives how the platform is governed. Communities make decisions about everything on the platform, from overall organizational strategy,

to the norms and standards for the most minor algorithm. The platform is truly *by the people, and for the people*.

### **Data Escrow Service and Data Marketplace**

CivSocial is committed to the principles that users own their data and own any revenues from their voluntary use of that data. The platform doesn't collect any data without explicit user consent in plain language, and it expressly prohibits third party transfers of data. Users can maintain their data on the platform and they can license the use of their data through a blockchain based marketplace. All revenues from the sale or use of their data belong to the users with a small percentage withheld by the platform to cover operating costs.

### **Secure Messaging**

CivSocial is committed to the fundamental principle that users should be able to communicate with each other through private, secure messaging channels that are protected with the highest levels of encryption. The platform will provide users this ability, including quantum-secure modes of communication. However, CivSocial also recognizes that different communities on the platform will have different ethos about the levels of security they desire in their communications. Some communities will opt for lower standards of encryption to facilitate properly warranted investigations by law enforcement personnel.

Users will have the option to select a level of security that matches their particular ethos, and this security level will be clearly communicated to both parties during communications. Three levels of secure messaging will be available to users:

- Ultimate encryption is the highest level attainable and requires an extremely high investment to crack.
- High encryption allows law enforcement access in cases of serious crimes, terrorism or other compelling national security grounds.
- Standard encryption allows law enforcement access for criminal investigations

These encryption methods will be based on open source protocols to the maximum extent possible.



# Appendix I. How we got into this mess: Modern Democracy, the Enlightenment and Disinformation

First we should ask ourselves why these issues matter. Why do we care about weaponized information and narratives, disinformation, and the possibility of a post truth world?. The answer provided by the Arizona State University Weaponized Narrative Initiative is important, because it puts these issues into a much larger civilizational context. The rampant disinformation enabled by social media platforms poses a direct threat to the Enlightenment principles that our civilization is based on: **we should always aspire to discover truth using reason.**

These Enlightenment principles underpin modern democracy. The Founding Fathers of the United States were Enlightenment thinkers, and the U.S. Constitution is one of the most important Enlightenment-era documents. The U.S. Constitution set the foundation for modern democracy worldwide. These democracies are based on rule of law systems where empirical thinking is essential to their functions. Facts and evidence are critical for everything from judicial processes to administrative due process. Disinformation and the concept of a post-truth world directly threatens democracy.

If we are're in a post-truth world, and evidence doesn't matter, then the truth becomes "whatever you can convince people of." This is a direct threat to the evidence based, rule of law system that modern democracy is based upon. If we concede we're in a post-truth world, then countries, organizations or even individuals with strong information operation capabilities and a casual relationship with the truth can hold inordinate amounts of power.

If democracy is the superior system based on evidence, but evidence doesn't matter, then other political systems can seem more attractive. Countries like Russia or China can convince their own populations, and the populations of other countries, that democracy is inferior to authoritarian rule. If modern democracy is to survive, we need to push back on this idea that we're in a post-truth world and facts don't matter. There is an objective reality, facts do matter, expertise matters, and evidence matters.

Some people speculate that the challenges we're facing with weaponized narratives and disinformation spell the end of the Enlightenment. We should challenge that idea. The Enlightenment was partially about mercantile elites empowering themselves with information - using logic and reasoning to push back on the dogma of the church and the nobility, which dominated the world in the 18th century.

The Enlightenment was partially about Elites using the scientific method to challenge dogma about how the world worked. The elites asked questions such as: "Who are you to tell us the sun revolves around the earth when we have evidence the earth revolves around the sun?" They were using a new technology - the scientific method - to push back on established power structures. But the Enlightenment was also about moral authority. The Elites asked questions such as: "Who are you to tell us how to live just because of your bloodline or title?"

A case can be made that elements of what we're experiencing now are a reaction against the Elites and their institutions. So where Elites were rebelling against the Church and the Nobility, now super-empowered individuals and groups are using new technologies - modern information and communication systems - to rebel against the elites. Some malicious actors are turning the scientific method on its head, using quasi-evidentiary approaches to convince information-overloaded citizens that the facts back their position. Others are simply asking legitimate questions such as "Who are you to tell us what to think or how to live just because of your title, your degree or wealth?"

In this way, the current period could be seen as a challenge to the Enlightenment, but also as a possible evolution of the Enlightenment away from elite power systems and toward a more dispersed, and democratic power system focused on individuals. The biggest problem is that we've empowered individuals with incredibly robust information and communication tools, but we don't yet have the institutions and frameworks to ensure accountability and responsibility. Democracy has had over two hundred years to refine the checks and balances in everything from government use of power to peer reviewed research. We've only had a few decades to create those structures for the Internet and social media, and they are still sorely lacking.

To illustrate this, let's look at the U.S. collective intelligence system - defined as the way we as a society determine truth from fiction. In the mid-20th century our system for determining truth consisted of several major national newspapers, a rich local media ecosystem, four national TV channels, government, academia and the church. But as information channels have grown geometrically, confidence in all of these institutions has waned. Some of this is a natural result of the diversification of information outlets and proliferation of communications technologies, and some is due to manufactured outrage by malicious actors. The net result is that our collective intelligence system for determining truth from fiction broke.

Malicious actors have stepped into that gap, and have used new technologies to create their own negative collective intelligence systems. They attract people into these systems with emotionally pleasing disinformation, they keep them on an emotional hook and they never let them go. Malicious actors use computational propaganda tools like social media, big data, autonomous agents (e.g. bots), and new discoveries from cognitive psychology to manipulate perceptions and influence behavior for their nefarious ends. And emerging artificial intelligence tools like chatbots, affective computing, audio and video manipulation (#deepfakes), dynamic content generation, and psychometric profiling will provide substantially greater capabilities to manipulate populations.

If we do not address these problems now, our societal collective intelligence systems will be irrevocably broken. And if our collective intelligence system remains broken, we will have failed to extend the Enlightenment into the age of hyper-empowered individuals and groups. The search for truth through reason will end, and we will enter the post-truth world where democracy is unable to function effectively. Scenarios for that world are not pretty - they range from a world of computational propaganda induced informational chaos where every communication is an information operation and no one knows what to believe, to authoritarian cognitive security state models where information is tightly controlled by the government in order to preserve social stability.

But there are ways to avoid these futures. We need to build a new digital institution of modern democracy for the 21st Century. This institution is a

collective intelligence system that will enable citizens to determine truth from fiction and control their digital lives.



## Appendix II: A Few of The Problems with Existing Media

- **Social platforms treat their users as their product.** The social platforms sell their users' data and their attention to third parties interested in influencing the users. This can't be emphasized enough: *for profit social platforms exist to facilitate the manipulation of their users by marketers, politicians, interest groups, foreign agents, or anyone else who wants to influence user behavior and is willing to pay for it.* That is literally how they make money and continue to exist. Their purpose is to facilitate influence and manipulation.
- **Social platforms optimize for engagement over truth.** The social platforms need users to spend time on the platform (or affiliated network) so they can be influenced. The social platforms use a variety of techniques grounded in cognitive psychology to encourage users to spend more time on the platform. They nudge users with notifications and they optimize feed algorithms for engagement. This addicts users to the platform and incentivizes content that maximizes engagement over content that maximizes truth, happiness, or other positive virtues. Unfortunately, content that maximizes engagement frequently consists of content that is inflammatory, fear-based, or outright disinformation.
- **Social platforms don't disclose what they are optimizing for.** The social platforms don't disclose to users that a particular piece of content was selected for a feed because it is more likely to keep the user engaged with the platform.
- **Social platform disclosures are unintelligible:** Social platform users must agree to a long and unintelligible list of terms of service that provide the platforms access to a wide range of data and usage rights. Where privacy protections are facilitated, it is often unclear exactly how privacy controls work or what data will or won't be disclosed.
- **Social platforms have little accountability.** Social platforms have a financial disincentive to take any action that undermines engagement. Frequently, the most inflammatory content, comments, and people generate the most engagement. Trolls often hide behind anonymous accounts or pseudonyms which allow them to behave in ways they never would in a face-to-face interaction. Bots pose as humans to facilitate social

engineering of public opinion. Foreign agents manipulate conversations in other countries by exploiting existing divisive narratives. And the platforms have little incentive to police these activities because they generate valuable engagement.

- **Social platforms are closed:** Social platforms strive to create walled gardens where the costs of switching are so high users will not leave. They disable or inhibit data portability that would enable you to take your rich contact network to another platform. They limit the ability for users to customize their experience. They are not open and auditable, so users have no real idea whether they are fulfilling their promises on a range of issues, from privacy to preventing computational propaganda.
- **Social platforms don't explicitly support democracy.** Social platforms are beholden to their investors, and as for-profit companies, their job is to provide profitable returns to their investors. They have no inherent incentives to support democratic values or democratic nations. So they are willing to modify their rules and standards in exchange for market access. They sacrifice free speech for profit.
- **Opinion pundits are indistinguishable from journalists and have no accountability:** The line between journalism and opinion is nearly non-existent in today's media. Pundits bias their news reporting to their audience and there is a low bar for "expertise". This problem is especially acute on cable news which has 24 hours of broadcast time to fill. Supposed "experts" present information and audiences have little ability to check the bona fides of the expert, their background, or their track record.