The Future of Trust will be Dystopian or Decentralized: Escaping the Metaverse

JP Vergne

"A thousand years ago your heroic forebears subjugated the whole of planet Earth to the power of OneState. [...] It is for you to place the beneficial yoke of reason around the necks of the unknown beings who inhabit other planets—still living, it may be, in the primitive state known as freedom. If they will not understand that we are bringing them a mathematically infallible happiness, we shall be obliged to force them to be happy"
— *Yevgeny Zamyatin* ("We")

"The trust of the innocent is the liar's most useful tool"
— *Stephen King* ("Needful Things")

Trust me, I'm famous

The Facebook corporation, operator of the *facebook.com* and *Instagram* platforms, epitomizes much of the dilemma contemporary societies are facing when it comes to trust. Many still vividly remember when, about a decade ago, the firm was touted as the enabler of pro-democracy Arab Spring movements. Back then, we confidently trusted Facebook. After all, it was connecting everyone without discrimination—not even economic, since it was free. And, by standing up to communication censorship in autocracies, Facebook was enabling freedom fighters to organize, congregate, protest, and overthrow corrupt governments.

A documentary entitled *How Facebook Changed The World* came out in 2011 to explain how, by "decentralizing" communications across millions of smartphones kept in our back pockets, Facebook prevented dictators from controlling the media and free speech with a master switch located in the basement of their barricaded palace.

But what does it mean to "trust"? It means to have positive expectations about someone or something even when these positive expectations may not be reciprocated; thus, to trust is to accept making oneself vulnerable in our social interactions. If I trust Facebook, for instance, I expect the firm to keep private the content I share over its digital platforms.

Facebook and its CEO have constructed that trust among their external audiences by repeatedly emphasizing their "privacy-focused vision for social networking" and by claiming to put user "privacy first" and above every other principle.[1] At the time of the uprisings in Tunisia, Egypt, or Libya, most of us trusted Facebook.

Now, fast forward a few years. Following series of revelations about corporate involvement in illicit mass surveillance sponsored by governments, former NSA engineer Edward Snowden clarified that "Facebook's internal purpose […] is to compile perfect records of private lives to the maximum extent of their capability, and then exploit that for their own corporate enrichment. And damn the consequences."[2]

And there's more. Facebook male engineers have used their privileged access to the firm's servers to stalk female users.[3] The company acknowledged granting special privileges to influencers on its platforms, despite prior claims that every user was treated fairly based on the same rulebook. At times, of the 60,000 employees in the firm, it is the CEO-founder himself who personally decides to remove or not an account or a specific post from the platform, among billions.[4]

What all these anecdotes have in common, beyond shaking our trust in the firm, is their pointing out to the intrinsic *centralization* that characterizes the design of the Facebook organization.

Since the heydays of the firm's reputation as democracy enabler, it has become clear that Facebook's claimed intentions and objectives, both in terms of privacy, equal treatment of users, and decentralization of communications, are vastly at odds with its actions and actual goals. Corporate documents obtained by the Wall Street Journal revealed, in fact, that Facebook executives unabashedly acknowledged that discrepancy: "We are not actually doing what we say we do publicly", they wrote in an internal review.[5]

---

[1] https://www.facebook.com/notes/mark-zuckerberg/a-privacy-focused-vision-for-social-networking/10156700570096634/

[2] https://www.vox.com/recode/2019/10/31/20940532/edward-snowden-facebook-nsa-whistleblower

[3] https://www.news.com.au/finance/work/at-work/facebook-reportedly-fired-52-employees-who-were-caught-spying-on-users/news-story/8d6c6ce230c20ea3299772b62245c762

[4] https://fortune.com/2020/05/29/facebook-mark-zuckerberg-trump-looting-shooting-post/

[5] https://www.wsj.com/articles/facebook-files-xcheck-zuckerberg-elite-rules-11631541353

Digital Dystopia

Dystopia is traditionally understood as a fictional world whose blatant discrepancies with the actual world characterize a hellish society that must be prevented from ever coming to life. Digital technology and social media have enabled the development of virtual (as opposed to fictional) worlds in which we are enticed to spend an ever-increasing share of our disposable time. A *digital dystopia* is a virtual world built intentionally on a foundational discrepancy between what is said about the virtual world by its designers and what the virtual world really is.[6]

The platform enables free speech, we are told? In fact, it *controls* free speech using its own corporate doctrine rather than its home country's legal doctrine. Pro-democracy? Not if, to roll out one of its virtual worlds, it has to strike a backstage deal with an autocrat, as illustrated by Facebook's presence in India.[7] Privacy-conscious? The business model of most social media corporations is to mine users' private data to extract behavioral predictions and to sell them to advertisers through online auctions. The digital dystopia's true goal, here, is to maximize users' behavioural compliance with the corporation's untold business agenda, as demonstrated by information scholar Shoshana Zuboff.

Paradoxically, by enabling us to broadcast a digital identity different from our real-world identity, social media lets us leverage a potential divide between the actual and the fictional in order to enhance others' trust in ourselves within the virtual worlds that we participate in. Without our cooperation and complacency as users, the model would cease to work. Our cooperation is guaranteed, for the most part, by the individual disadvantage each of us would incur from removing ourselves from the virtual worlds (e.g., Instagram, LinkedIn), namely, a missed opportunity to appear smarter, sexier, more empathetic, or more accomplished than the next guy.

---

[6] I use the term to mean something different from, albeit related to, Jean Tirole's arguments in an economic analysis of social rating systems entitled "Digital Dystopia." https://www.tse-fr.eu/sites/default/files/TSE/images/conference/Digital_conf_2020/tirole_digital_dystopia_080120.pdf

[7] https://www.theguardian.com/technology/2021/apr/28/facebook-blocked-resignmodi-hashtag-india-coronavirus

The mutual reinforcement between those two lies—the former, corporate and the latter, personal—can serve as basis to create a competitive advantage and make digital dystopias particularly sustainable. To further grow a business rooted in such dystopia, a natural next step is to add neighboring activities around the core, formed around predictive knowledge of users' identities, so as to increase the amount of disposable time spent by users inside the dystopia.

Upstream, this can be done by crafting neighboring virtual worlds to encompass additional aspects of users' lives beyond their social communications and media consumption. These all-encompassing virtual worlds, in which users are expected to dive deep using VR headsets, will jointly form a corporation-controlled connected environment called "the metaverse". Building the metaverse, in fact, has recently and explicitly become Facebook's strategic priority according to the firm's CEO.[8]

Downstream, the metaverse needs a mechanism enabling users to live full-time within its boundaries. To that end, users should be able to earn and consume value across the expanding metaverse, which implies a need for its designers to create a unified means of economic exchange. That's why corporate platforms will find it appealing to create their own money. Besides, by minting its own digital currency, a metaverse can better enforce its own corporate doctrine by incentivizing user compliance– just like national currencies, used by the state to denominate taxes and fines, compel citizens to abide by nation-state law. Facebook, as part of its projects Libra and Diem, has been working on the design of a new form of money since 2018, so the firm can eventually bring together all its virtual worlds within One Metaverse powered by its own digital currency. In venture capital jargon, this tactic is akin to "building a moat" around a business to protect its historical competitive advantage over the longer term.

Dystopia Needs Distributed Trust

In a virtual world created by a corporation, we as users must delegate our trust to the managers, designers, and engineers employed by the corporate platform. The actual rules underlying the platform are hidden from us as proprietary algorithms. The true goals of the corporation are also

---

[8] https://www.theverge.com/22588022/mark-zuckerberg-facebook-ceo-metaverse-interview

hidden (perhaps in plain sight), at least to maintain a semblance of plausible deniability in the face of regulatory agencies.

We as users must delegate our trust because the decisions made about the virtual world's underlying rules, which affect us, have themselves been delegated down a corporate hierarchy, following the well-understood principle of the division of labor. By analogy, in the domain of "distributed computing", reliance on so-called "master nodes" and "slave nodes" helps break down complex tasks to increase computational efficiency and speed.

In corporations, a similar principle is at work as tasks and associated decision-making authority are *distributed* to various middle managers down the hierarchy. Consequently, in a virtual world, our willingness to make ourselves vulnerable in our interactions with strangers is premised on *distributed trust*. I need to trust, for instance, that the security engineer competently oversaw the coding of the platform's private messaging functionality, that the chief technology officer recruited an honest security engineer, and that the overall objectives set by the CEO during closed doors meetings are consistent with my well-being as a user.

Utopia Needs Decentralized Trust

Decentralization, as political principle, surfaces amid the 1789 French Revolution. Since then, decentralization has found advocates across the entire political spectrum, from leftist anarchism (Joseph Proudhon) to classical liberals (Thomas Jefferson) and free market libertarians (Murray Rothbard).

Unlike distribution, which refers to the dispersion of tasks and associated decision-making, decentralization is about the dispersion of information. As such, free-market proponents like Friedrich Hayek consider that, if the information needed to produce wealth is dispersed across millions of individuals, it becomes intractable to employ these individuals as part of a single firm and tell them what to do in a top-down, distributed fashion. Instead, it may be convenient to let the market aggregate information in a transparent way, that is, to coordinate tasks using the price mechanism rather than managerial authority.

The problem, in decentralized worlds, is that everyone is potentially a stranger with no authority over others, and nothing about strangers can be used as collateral nor leverage. Everyone has an intuitive understanding of this problem in a setting known as "the tourist trap". A tourist trap is a business premised on the assumption of no repeat customers– for instance, an independently-owned restaurant in an area that tourists tend to visit only once and where they stay just for a day or two. In this setting, the business has no collateral at stake from the perspective of a tourist, such as a reputation to maintain. And the tourist has no leverage over the business, such as a ballot to be cast in the next local elections (tourists do not live locally). That's why you would never rent out your car to a stranger, from 9am to 5pm, while you're at work sitting in the office on a weekday. In a decentralized world, without trust, no such transaction can happen.

However, a digital platform could start charging a fee to match trusted drivers with trusted car owners. It would supply insurance to users and a reputation system, essentially acting as the centralized provider of distributed trust to participants. The digital platform, in essence, would be *creating a market* where previously there was none– by allowing users to delegate their trust to the managers, engineers, and algorithms of the matchmaking platform.

For some users, a potential issue with this is that the platform's rules and data may be proprietary and hidden from sight. Because of that, some might remain reluctant to delegate their trust.

An alternative platform could be built by freelance developers following a different design. To join, a user would agree to stake currency as collateral in an escrow account (maybe $1,000) before they can drive a stranger's car. The state of the car would be documented with timestamped photos recorded in a shared database that everyone can view. Standard penalties would apply for minor damages such as scratches, much like in today's car rental business. Rental payments would be automated using per hour and per kilometer rates that depend on the car make and model. Every rental transaction would be recorded publicly, alongside other relevant data on the state of the vehicle and owner reputations. In the event of a dispute following an accident, a protocol would randomly select regular users with high stakes and reputations to act as witnesses and adjudicate the dispute.

In this utopian scenario, all the software and data would be open-sourced and managed by a community (much like with Linux and Wikipedia). This would ensure transparency and eliminate the possibility of having a discrepancy between the stated and actual goals of the organization. Every user would be able to view all the rules, encoded as open-source software, and a full history of past transactions. Perhaps more strikingly, in this scenario, we would not need a corporation to act as intermediary, since every transaction would be taking place peer-to-peer using decentralized information. So, instead of having corporate engineers provide distributed trust to users, we'd have *decentralized trust* provided peer-to-peer without the need for a managerial hierarchy. The organization itself could be owned by its users and developers and structured as a co-op— a template that readers aware of Spain's *Mondragon Corporation* will be familiar with.

The Future of Trust: Decentralized Utopia or Distributed Dystopia?

Albert Einstein argued that "every kind of peaceful cooperation among men is primarily based on mutual trust". Yet if there was a formula to manufacture that mutual trust, we would be able to turn it into *calculated risk*— something that insurance companies do routinely when they price the odds of a burglary or fire devastating a home. Put differently, the issue of trust can never be addressed directly as such, but only indirectly by first turning it into risk. And even if all the trust needed could be turned into calculated risk, then we'd still have to trust the calculator. Hence the paradox— and our society's dilemma.

As digitalization makes it necessary for all of us to transact routinely with complete strangers, we need to find new ways to establish a minimum level of mutual trust to enable interaction. The dominant digital platforms that shape our virtual worlds rely on distributed trust, that is, they compel their users to trust the calculators, organized as for-profit corporations. Recent history, illustrated with a few anecdotes pertaining to Facebook, has showed us the limits of distributed trust and the dystopian potential it entails with respect to the upcoming metaverse— an all-encompassing, sovereign superset of virtual worlds powered by corporate digital currency.

There is a much rosier alternative out there, a utopia, which relies instead on decentralized trust as it lets everyone access all information to remove any discrepancy between the stated and actual goals of our virtual worlds. Software engineers and cryptographers are designing that

alternative as we speak from new building blocks, such as blockchain, graph, and zero-knowledge proof technologies. Instead of trying to subsume all the virtual worlds within One Metaverse, the crafters of the decentralized alternative envision nearby islands interconnected by chains of communication and exchange— a *paraverse* of sorts, instead of a metaverse.[9] To avoid having to trust the calculators, users of the paraverse have access to all the rules, algorithms, and data written and published as open-source software and database entries. Unlike the metaverse, which is distributed but not decentralized, the paraverse can operate without giving a happy few special access to privileged information.

In all likelihood, the utopia and dystopia will both fail. The former, because it overestimates our thirst for individual autonomy and the second, because it underestimates our power to resist organized oppression.

And there lies the future of trust: as tension between the metaverse and the paraverse, on the thin line between a dystopia that's already knocking on our front door and a utopia that may never find its way to the back door.

---

[9] By contrast with "meta" (*beyond*), which can evoke a form corporate transcendence, "para" (*beside, apart from*) indicates juxtaposition and is not premised on the juxtaposed elements being subject to the same unified rule or code ("para" can also mean *abnormal*, as in "paranormal").