Reputation Markets

Reality, Dangers, and Possibility

Since reputation networks already exist, and are critical to solving game theoretic coordination problems, the more important question isn't 'do they exist' or 'should they exist' — it is 'how can they be made better and more egalitarian?'

C 4
S S

Emmi Bevensee

Trust is the backbone of not just economies, but also the social fabric more broadly. With it, new worlds are possible, and without it, collapse is inevitable.

Trust is built on dynamic and complexly layered reputations. We each have our own methods of building trust, and different communities have their own systems for developing reputation. Mana in Maori culture maps and flows differently than social capital in western activist scenes, but they all point to the meaty question of how we can engage more deeply with each other. A healthy reputation system allows for these diverse and dynamic ecosystems to consensually co-mingle.

When someone unknowingly engages with a person who has harmed people under similar circumstances in the past, this is a failure of information sharing around reputation. A whisper through informal networks that someone has an abusive past works similarly to a price signal in markets. Whisper networks are inefficient and sparsely connected reputation markets. Whisper networks transmit local knowledge to impact the contextually dependent preferences of other actors. If the harm someone did is relatively minor it only shifts your trust preferences slightly. If it's severe it prevents you from wanting to engage at all. This process is analogous to the way price changes based on factors downstream in the supply-chain. Markets are a form of information network. If environmental and labor ethics crises

make the mining of coltan excessively expensive, iPhone consumers will likely change their preferences to something cheaper. The price reflecting market conditions shares a stripped form of information across geographic distance.

This market-based information sharing holds as long as Apple can't rely on corporate subsidies and colonial trade-agreements to externalize the costs of production and maintain an artificially low price. Similarly, without the artificial subsidies and accumulation, illegitimate forms of social capital, like coercive charisma and high school popularity, whither on the vine. In either case, faster and more informative signals diminish illegitimate hierarchies while simultaneously boosting the resilience of actors on the margins such as a new competitive firm or a new entry to a social network.

The existence of and problem-space functioning of reputation markets are the same under any political or economic model, even if the details of how they play out are different. A charismatic (anarcho-)communist committee member still has informal power deriving from their attractiveness, charisma, fluency with ingroup signals or manipulation, and their control over privileged information. This process functions similarly to when Elon Musk smokes weed with Joe Rogan and makes illegal business comments on Twitter and Tesla stocks tank. Healthier reputation systems would help to map the hit of a #metoo allegation more effectively against the trustworthiness of politicians. Since reputation networks already exist, and are critical to solving game theoretic coordination problems, the more important question isn't 'do they exist' or 'should they exist' — it is 'how can they be made better and more egalitarian?'

1. Sybils are possibly less of a problem in the types of systems I hope for because anyone with an across the board empty reputation in adulthood would be looked on with suspicion. People would know that they had spun up a new identity, and it would be harder (but possible) to rebuild reputation. This would prevent the need to centralize power in order to stop people from creating new identities.
2. In an assurance game there is an additional Nash Equilibria strategy, mutual cooperation, whereas in a prisoner's dilemma it's just mutual defection.
3. I've never read this book only explanations of it so apologies if I misunderstand it but multiple people pressed me to include it in this piece and it offers interesting fodder nonetheless.
4. Doctorow says, "A few people do very badly, and get downranked and eventually punted off the system – something that a normal complaints tipline would handle just as well." but that's not how that works at scale at all. If it was, then Facebook, Twitter, and Youtube would be in much better shape than they are with regard to nazis and there wouldn't be a huge cottage industry of machine learning auto-moderation springing up.

Universal Basic Income is developed that supports people to live peaceful and meaningful lives in relative safety and abundance regardless of personal competitive advantages. The concomitant technological advantages self-compound into worlds more fantastic than that we can imagine.

How Can We Know Each Other?

The most important part of reputation networks, technological or in our minds, is that they be accurate. They should not be exaggerated positively or negatively, and should as close as possible hold the intricate and beautiful complexity of a human. From that, we should have effective feedback mechanisms and cultivate the ability to learn from those results with bravery and right-sized humility. This could be technologically, or just through the words of trusted friends.

The more reputation systems we have interacting with each other, the better we can model each other. The ability to accurately model each other is the bedrock of trust-building and everything that entails. Through overlapping horizontal networks of reputation, we can break away from the dangerous myth of high-trust bordered societies, and into the evolutionarily agile realm of networked co-development. However high the potential of healthy reputation networks, the dangers of authoritarian and hyper-capitalist drift are not just dire, but already happening. As such we need to be innovative in our adaptation and resistance.

Teach me *how* to trust you and I'll try to do the same.

A Brief Primer on Trust and Reputation

There's a glut of scholarly research on reputation largely from within different schools of economics. Trust is generally seen as the expectation that someone will or won't do something. One mechanism of reputations is the bootstrap method which is just repeated interactions (or games). Most of this literature revolves around the concept of moral hazard (asymmetries of information, knowledge, and its relationship to risk-taking) and punishing or disincentivizing defection. The second trust mechanism is the Bayesian update which is just about changing your beliefs based on new information.

In a more technical sphere, reputation is generally thought of as a tool for facilitating interactions. In describing digital reputation system design, Dellarocas says, "Reputation is a summary of one's past actions within … a specific community, presented in a manner that can help other community members to make decisions… whether and how to relate to the individual." Reputation is closer to the Bayesian update method and has to do with expectations of what you expect someone to be, rather than to do. Reputation and trust are generally modeled game-theoretically, though modern approaches can model much more complex incentive ecosystems than that of a traditional prisoner's dilemma, including things like repeated games, complex social signaling, and community standing.

It's a Feature Not a Bug

Efficiency means less accumulation, less coercive hierarchy, and fewer obstacles to entry. All these factors lead to people being less able to artificially protect their reputation. For example, a very

popular person can often squash stories about their abusiveness. The more efficient reputation markets are overall, the less this is possible. In addition, people who are less able to natively grasp social dynamics to accumulate informal power around themselves are given much more numerous and meaningful opportunities to build the healthy forms of social capital such as resilient networks of friendship solidarity. So the intricate beauty of a shy and socially awkward person can more easily grow and network but a scene patriarch will have a hard time building an excess of power that makes them untouchable.

As laid out by Gillis and others, reputation is a constant site of contestation where equilibria are reached only through competing signals and claims. A reputation structure not prone to monopolistic accumulation allows for a range of other implications such as the legitimization or delegitimization of property titles. A robber-baron prone to exploitation loses favor in the community, which increases the costs of guarding their wealth as community favor begins to shift in favor of those who would decentralize their wealth. A massive polluter living in the hills will face exponential difficulty in hoarding their empire if the externalities are placed on the shoulders of those living at the base of the hill. Without state subsidized policing and banks, their entire business model could flounder under the weight of both labor and consumer resistance. This is made more likely if the true nature of their crimes is able to spread through people's awareness which requires efficient and transparent reputation systems.

Similarly, individuals are better able to learn and grow with access to better information. As in non-technological reputation systems, people aren't just represented by a single totalizing score

there is "no objective measure of merit" and that's why a unitary reputation system is so preposterous. Monopoly is dangerous and anything that fuels it is inhumane and leads to dystopias.

In this uninsulated network of overlapping markets, the geniuses born as "Syrian refugees" or the "one brilliant scientist with ALS" are more able to enter the fray despite the circumstances they were born into. No doubt the system envisioned by Doctorow is dystopic but it's quite different to the idealistic versions I advocate.

Doctorow's commentary on the danger of trolling reviews is well-played. Gamergate et al will use anything for a weapon, but so will anti-fascists and the like. Though culture-wars would obviously play out on reputation networks, it would be possible to test and amplify systems of abuse protection such as banning users found to be abusing the system. Additionally, one's own reputation would show affiliation with a white-nationalist or gamergate milieu, so review readers could have filters against those with those views. But the point isn't whether I can propose a patch in this essay, but more so that patches compete in the real world to achieve ethical efficiency.

ECONOMIC ABUNDANCE AND AUTOMATION

After a period of difficult transitions into the post-total automation era, society has reached a state of relatively maximal personal and interdependent freedom through high trust. This era of peak coordination is facilitated by a combination of incentive engineering and highly efficient reputation networks. As a result of the economic abundance created by competitive, overlapping reputation systems, a profound social safety net and borderless

monopoly, it's pretty clear that reputation markets alongside traditional markets have facilitated trade amongst strangers.

However failing whuffie may be, it's important to acknowledge that there is some nebulous exchange value between various forms of currency. Reputation does function like a regular currency, just with different strengths and weaknesses. That's because social capital has market dynamics and reputation is a form of money. You can trade popularity for money and money for popularity. That's precisely why capitalism is so dangerous, and why meritocracy is a dog whistle for "born rich and dominated my enemies," along the lines of Hinduvatna caste supremacy and Elon Musks across the world.

Any system that protects artificial concentrations of wealth around any kind of capital, will facilitate inequality. This is why the state protects big banks and tech giants from competition through IP laws, beck-and-call policing, subsidies, bailouts, and other forms of many-monopoly creating corporate welfare, and in turn, those corporations protect the state. This is similar to the ways that people like abusers can use the threat of violence to artificially protect their reputations and why a credible reputation market-style wikileaks for exposing rapists is so important even if there is a risk for system abuse at the margins. Removing insulation around accumulation prevents the type of 'meritocracy' that "ends up pooling up around sociopathic jerks who know how to flatter, cajole, or terrorize their way to the top."

Anyone who can hold a solid reputation without these forms of market manipulation could be said to have a form of healthy social capital. The kind that is just built through being a decent, trustworthy, person who works to do right by people. Of course,

but are instead comprised of richly textured webs of reputational information forming a more accurate, changing, and holistic interpretation of a person. Through this, people are also able to better understand and learn from the incongruencies between how they view themselves and are seen by others. Because the market is flattened, there are much wider pathways for building new trust. Additionally, because there are such diverse types of overlapping reputation systems, a bad score in one area is more easily seen in context with other information, having an additional flattening effect on the network topology.

In a future-tech system, this is made possible through the types of data that are made visible to the person, though users can also post private reviews about others. Then each user also has their own reputation metrics which can be reviewed when looking at the ways they've reviewed others publicly. Different systems of both anonymity and transparency will compete for ethical utility.

Aside from just the increased overall trust and concomitant economic coordination and abundance provided by reputation, there is also increased transparency around the fungibility of things like social capital through which murky exchange rates often protect those most manipulative. The tensions of legibility and illegibility should wrestle in a timeless battle for efficacy and egalitarian ethics.

No, Actually It's a Bug

For many, this may read as an obvious dodging of the sheer terror of ubiquitous reputation, rankings, scores, and mutually suspicious judgment. At the most extreme end is the *Black Mirror* episode where social scores become the metric for class

positionality and overall limits to personal agency favoring neurotypicality and productivity above all else. People could hack reputation systems to artificially inflate the reputations of paying clients. Reputation systems can pin people other against one another in artificial scarcity and vicious competition despite the non-rivalrousness of trustworthiness. Ableism could become further structurally entrenched, valuing those most neuronormatively able to recreate the status-quo. Reputation systems could contribute to the quantitative oversimplification of people. After all, the beauty of people is their sheer complexity and in a sense, describing any one feature of someone is minimizing the wholeness of them.

These and other dangers are real. But they're also already real. These systems exist already — both in informal relational ways, and in massive state and corporate databases. Each individual is scored and judged by everyone and everything in constant myriad ways. This doesn't displace the dangers, but it does put them in perspective. There is no Eden we can return to of perfect non-judgement where reputation ceases to exist. Gillis notes, "There's a reason credit preceded currency — as Graeber had to remind a number of economists — trust and goodwill are simply the foundation of the world we move in." We need avenues for those of us whose skills are not commodifiable by traditional markets to assert ourselves and our right to exist with maximized agency.

It's impolite to acknowledge that we judge each other. But when I share my car keys with someone I'm taking a rough heuristic of many variables surrounding their trustworthiness and reputation — and that makes sense. If they're aren't good at driving safely I probably shouldn't lend them my car, or they should have to find a different more net-beneficial and low-risk way of getting

While his critiques are generally insightful, they point only to a very specific type of reputation economy in which the reputation system is not transparent, is centralized, and is the primary currency.

With regard to transparency he writes, "Unlike other virtual currencies like Bitcoin, Whuffie isn't something you buy and sell: it's a score that a never-explained set of network services calculate by directly polling the minds of the people who know about you and your works, reducing their private views to a number." First off, if brain-to-computer interfaces become seamless, then we have already achieved rapid, strong-AI takeoff and have bigger issues to contend with. Additionally, that "never explained" bit is pretty important because it belies that there is no transparency and that people don't really have a stake in which reputation services themselves have reputations. If the systems are not state-run or monopolistic and are themselves subject to boycott and reputations they will be incentivized to address problems of monopolistic reputation systems as well as inscrutable systems of calculation. If the majority of users are screwed over, they will stop supporting these reputation services in favor of a system that better handles complexity, transparency, and equity. The absence of transparency and competition are the keys to black-boxing a system into the hands of those most powerful.

Doctorow is right to acknowledge that "reputation is a terrible currency." Currency must be fungible and a store of value to be usable as a basis for an economy, but reputation can facilitate transactions with more traditional currencies. As much as he may dislike the anomalies where Ebay's store-keeper reputation model fails, it downplays the fact that Ebay is currently worth around $33-billion. While one should be appalled at their near

Commons Quality Assurance App-ified

As in the guild societies of late-medieval times studied by Kropotkin and others, stateless quality assurance commons can arise through voluntary, bottom-up, competitive reputation networks (Carson, *Desktop Regulatory State,* p 221-240). These networks can transform the prisoner's dilemma into an assurance game where people compete to build their reputation, but obstacles to market entry are still low. They serve the additional function of being illegible to centralized state entities while being highly legible to individuals embedded in local networks. The technological era transforms both the dangers and potentials of these systems into new scales.

Many apps could compete to be the centralized source of various reputation and quality assurance organizations, allowing users to quickly choose the aspects of a given product that are most important to them, and engage based on their values and needs. Similarly to the assurance game played by firms, the quality assurance organizations themselves are in an assurance game and building trustworthy reputations. This Russian doll of reputation provides a more resilient, customizable, and modular approach to ensuring that quality assurance cartels cannot be sustained to the detriment of market entry, while still maintaining the benefits of various forms of safety and ethical consumption consideration tools.

Whuffie

Whuffie is a fictional reputation currency introduced in *Down and Out in the Magic Kingdom* by Cory Doctorow. Cory has since written that "inequality is even worse in reputation economies."

around. The advantage of relative abundance and low barriers to entry is that these alternative strategies become increasingly accessible and can be coupled with profound social safety nets.

We shouldn't just think of each other as numbers, but we also need to get a sense of how to trust each other. Everyone deserves love and access, but not everyone deserves my car keys. So these forces of complex de-quantified love and solidarity, coupled with accurate understandings of each others strengths and limits should be in tension. Systems that help us navigate these tensions should be analyzed, critiqued, and improved.

Existing Reputation Networks

To understand the functioning of possible reputation systems — both good and bad — it is important to analyze existing systems.

Dark Market Yelp

Dark markets that traffic in illegal drugs are able to simultaneously remain anonymous and build trust through reputation markets. This means that they are horizontally legible to buyers while (ideally) vertically illegible to the state. When someone seeks to buy a drug locally they are forced to accept the quality of whoever they personally know and the amount of trust is only relative to a small number of local buyers. Dark-web markets radically cut down on low-quality or dirty drugs through massive reputation markets, using tools like chemical purity testing that would make less sense at smaller scales. This also minimizes violence between buyer and seller by making robbery more difficult. In a sense, this is an agorist form of Yelp and provides the same benefits that Yelp does.

Even if users may skew Yelp reviews through exaggerated positive or negative reviews, it still provides a forum for economic resistance against malicious or harmful firms. If a shop employs neo-nazis or massively pollutes, people can write that in reviews, and most reasonable consumers will choose a different shop and cost them economically. Reputation markets are the backbone of things like boycott as well as quality assurance.

COMMON POOL RESOURCE MANAGEMENT

Elinor Ostrom's groundbreaking work on Common Pool Resource (CPRs) management shows how local reputation, coupled with things like graduated sanctions can effectively deter defection and prevent a tragedy of the commons. If you see local actors regularly, you have an incentive to maintain trust with them. Things like social pressure and social capital intersect to help people effectively manage rivalrous goods in common property. Unfortunately, this method is largely scale-dependent and so works differently on global CPRs such as the ozone or oceans.

SECURE-SCUTTLEBUTT AND THE DECENTRALIZED WEB

Problems of identity, like Sybil attacks, and reputation are being widely mapped and explored in the decentralized web space. Reputation is generally broken into two categories. The first is trust graph which has to do with trusting someone through networks of trust as in liquid democracy or page rank. The second is behavioral reputation, which is the accumulation of one's past actions, and must be Sybil attack resistant. It's often easy to spin-up a new identity on things like the blockchain so reputation

with the person you were matched with originally. The two of you have a lovely evening of engaging conversation, and joyous dancing, before deciding to leave together. Knowing in advance each other's preferences and needs you are able to trust each other more deeply — making the consent process, both yesses and no's, so much more fluid. As a whole society, consent is more effectively incentivized and facilitated leading to an almost complete eradication of sexual coercion.

TOTALIZING FASCISM

As a response to the rapid adaptations forced by climate change, a more powerful and technologically-advanced hyper-authoritarian era emerges. The promise of order is delivered through an iron-fist of repression and the social cleansing of undesirable elements. With the ubiquitousness of 3d-printing and decentralized internet, the world has become too complex to be administered with traditional methods of dictatorship, and so a new techno-dictatorship is developed to control the illegible populace. It is run through state-controlled reputation markets that become so efficient at surveillance that all persons are rendered completely legible to centralized enforcement. Every single action is monitored by state intelligence agencies down to the smallest dissenting gesture. No one whispers against the state or glances too long at its horrors for fear of being disappeared. Resistance is forced into the most extreme margins and reputation is enforced with sheer brutality. Those who fail to maintain adequate scores are re-educated through torture or just liquidated in the margins of society as their very existence compromises the utopian fascist myth of revitalization through ultra-violence (techno-authoritarian palingenesis).

inefficient choices by so-called "rational actors." For example, something like Solar Coin can incentivize the collectivized sharing of surplus solar, in a way that traditional markets wouldn't. Similarly, the Nori carbon-offsetting marketplace is designed to incentivize carbon reduction.

Thinking in this domain could help solve the incentive problems of environmental common pool resources at a scale that is inadequately solved by Ostromian localized solutions. If everyone had a carbon output reputation that both prevented their ability to collect various forms of capital if they were offenders, and also artificially incentivized carbon-negative behavior, we could break the cycle of negative environmental externalities. Corporations and their leaders could be more effectively held accountable, steered, or abolished. These scores and qualitative stories would likely reveal the inefficiency of massive hierarchical multinationals, and as such could help create more sustainable resource creation organizations.

AUGMENTED REALITY DATING

Imagine entering a club and hoping to meet someone. You flip on a projected screen in front of your eyes that scans the room for shared preferences and interests and suggests some people for you to chat with and some possible topics for you to discuss. As you approach someone you speedread their various reputation details. They share your kinks and hopes for the evening, but they have one critical review from an apparent ex amidst a bunch of mostly positive stories. You read further into the critical experience and something about it seems strange — so you check out the reviews of the person who wrote the review. All of their reputation notes are about them making up stories, so you decide to take a chance

becomes meaningless.[1] This is a strange feature of so-called "trustless" economies. Similarly, if you try to impose scarcity on trust systems with something like voting with one's own cryptocurrency, you just structurally enforce plutocracy in which the wealthiest secure their future power — similar to the property titles of landed elites stolen through colonialism and slavery. For these reasons, they tend to rely more heavily on the trust graph than behavioral reputation.

Secure-scuttlebutt (SSB) is different from the foundation. SSB is a peer-to-peer protocol for devices to communicate beyond the constraints of traditional internet infrastructure. It can run without the centralized infrastructure of modern internet and was designed to fulfill the requirements of an interstellar communication system for a galactic council. Currently, many different platforms and tools are built on top of it — including social media platforms. Scuttlebutt is different than traditional p2p technologies, like blockchain, in that it replicates through trust rather than being designed for trustlessness. Your reputation allows your messages to spread farther because each user replicates the messages of people close to them in the network. All blocks are visible so if, for example, I see a bunch of femmes I trust have blocked a guy, I can assume he is probably creepy and refuse to replicate his messages. This increases the health and resiliency of the network as a whole, by limiting the ability of malicious actors to harm others and spread harmful messages while simultaneously helping people discover people who are good-faith actors in a similar way to how meatspace networking works. Gossip is the name that scuttlebutt uses for blobs of information people can pass on from each other. Your choice on with whom and how to gossip is built on trust which is built on reputation. In these ways,

SSB is more tied to behavioral reputation in that while you can easily create a new identity, you have to build trust, reputation, and friendships again from scratch.

CHINA'S SOCIAL CREDIT (社会信用) SCORES

There are seemingly endless waves of frenzy around China's social credit system. Dystopian horror stories abound of people's scores being used to justify denial from flights and train trips — or even ending the college enrollment of family members of someone else with a low score. Forms of ingenuity and resistance to the centralized system have already developed, such as phone cradles that mimic the user taking steps while they relax to increase their score.

Other reports suggest that the system is less coherent and more complex than it is usually painted in often inaccurate or biased western media. Regardless of whether it is as much of a totalitarian surveillance system as those in the U.S. see it as, it nonetheless makes clear the dangers of the priorities of states determining the trustworthiness of individuals and then stigmatizing or fast-tracking users based on these characteristics. However terrifying the details of the more experimental features of the social credit system are, credit scores in the U.S. are also ubiquitous and malicious in their arbitrary and expansive power. The difficulty that marginalized people face in building their credit scores turns into a vicious structural cycle that expands intergenerationally and maintains the impact of illegitimate property titles descended from things like slavery and indigenous dispossession.

SEX WORK CLIENT BLACKLISTS

Some of the most important reputation networks in sex work are client reviews and blacklists in the form of websites like preferred 411. P411 is a reputation network that allows sex workers to review clients, especially those that are dangerous, as a way of deterring violence and police sting operations. Unfortunately, as a result of infrastructure centralization, FOSTA-SESTA was able to repress the website p411, who stopped offering new accounts. This puts sex workers, especially those most marginalized, at increased risk of violence while also preventing good clients from sharing their positive reputations. Sex work positive platforms like Switter and Tryst provide not only a venue for networking and advertising, but also to show trust in and boost other workers while posting the details of violent clients.

POSSIBLE REPUTATION MARKETS

Although there are pros and cons to the ways that reputation markets are presently employed, they open difficult and essential conversations for the existential and intimate challenges facing humanity. Imagination can help to chart a positive direction for this delicate needle threading. If there were grassroots, transparent, and accountable reputation systems that were also competitive with each other, we could have all kinds of niche systems designed to solve different types of problems.

CLIMATE CHANGE

One of the most powerful aspects of the digital currency space is the ability to incentivize non-zero sum behaviors that benefit the collective, but would normally be outranked by selfish but