Reputation in anonymous contexts

Clément Renaud, Gabriele de Seta & Dino Ge Zhang - 2023

Table of contents

1. Executive Summary

Abstract

This document discusses the relationship between anonymity and reputation in online communities, with case studies from 4chan and Reddit, as well as Japan's 2ch and China's Social Credit System. The author argues that design decisions about anonymity and reputation need to take into account users' elasticity and cultural contexts, and recommends mapping continuums instead of binaries, accounting for elasticity, questioning universality, providing flexible ways to enforce rules, and prioritizing interoperability. The document also provides references for further reading.

Breakdown

- The report explores the concept of reputation in an anonymous social environment online.
- The report is divided into three sections: definitions, case studies, and recommendations.
- The research is conducted by a group of three social scientists and the Unirep Protocol team.
- The definitions section lays the groundwork for the report by defining anonymity and reputation and discussing their relationship to trust.
- The case studies section uses a case study approach to illustrate how anonymity and reputation changed through history and in different global contexts.
- The recommendations section provides practical design knowledge derived from the history and discussion.
- The recommendations include: mapping continuums instead of binaries, accounting for elasticity, questioning universality, providing flexible ways to enforce rules, and prioritizing interoperability.
- The report suggests that reputation in an anonymous social environment online is a complex and nuanced concept that requires careful consideration and flexible design

2. Introduction

The digitalization of our social lives has led to two dynamics, apparently contradictory but also complementary. On the one hand, the new topology of exchanges online, etc gave birth to new forms of fame and referral, recreating new reputation systems within and across platforms. On the other hand, the ability to appear only under pseudonymous identities shaped conversations to become an integral part of the online experience. While these new forms of reputation and anonymity can be understood as the continuity of pre-existing ones, their new proximity through online communication systems came with new questions.

How does reputation exist in an anonymous context? Can reputation be measured while preserving anonymity ? What does it mean to be anonymous and well-known? Is reputation exchangeable? To support the ongoing reflection around these questions, the present report offers an overview of some of the most recent research in human & social sciences around the concept of reputation in an anonymous social environment online.

The report is divided into three sections. The first one is dedicated to definitions and explores the concepts of anonymity, reputation and trust and their possible relationships. The second part provides several relevant case studies from existing online platforms and shows how they managed to balance the creation of reputation systems with anonymous / pseudonymous identities. To provide a diversity of experiences, the case studies were selected to reflect a variety of scale (from niche websites to very large platforms), time frame (from the early 2000s to today), and geographical origins (US, Japan and China). The third and last part provides some recommendations specifically to support the work of the Unirep Protocol team.

3. Approach & Methodology

The report is the product of the encounter of a group of three social scientists (an anthropologist, a sociologist and a geographer) specialized in the study of digital technology and the Unirep Protocol team. The work started with a 1-hour brainstorming session to define the key concepts and questions to study. After that, the working group had regular meetings and exchanges in chat about the structure and content of the report.

The research content is mostly the product of desk research as well as exchanges between the researchers. The definitions rely on a literature review of the fields. The case studies were picked for their relevance and capacity to illustrate and inform on existing approaches to the design of reputation systems. The collection of recent and historical views from existing works to provide an overview of the landscape of ideas around the topic. The outcome and recommendations are the product of conversations with the working group.
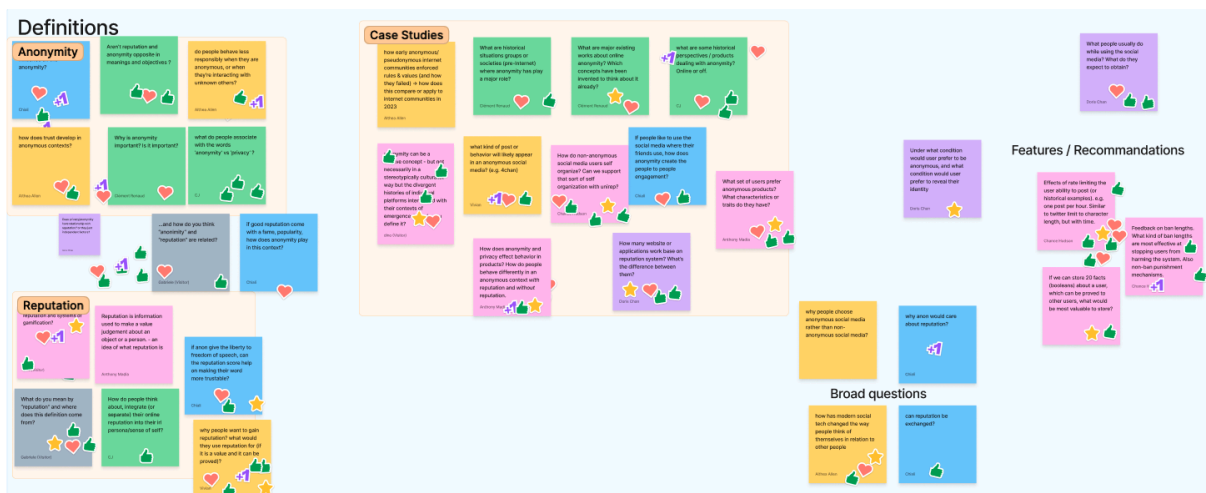

Fig. The board used in the original brainstorming (see link)

4. Definitions: Anonymity & reputation

In this section, we lay the groundwork for this research report by defining the main two concepts of *anonymity* and *reputation*, explain their relationship to *trust*, and discuss their *effects* throughout history.


Anonymous, *Men with cameras*. Detail from mass panoramic photograph of rally in Piazza Venezia. *Rivista Illustrata del Popolo d'Italia* (Oct. 1932)

4.1. What is anonymity?

Strictly speaking, anonymity means 'being unnamed', or 'having no name'. If we look at the literary history of the word, the adjective *anonymous* crossed over from Greek into English from at least the sixteenth century, and was used quite literally to indicate a piece of writing or an author with no name (Ferry, 2002). Being anonymous is impossible in non-mediated (face-to-face) communication: it requires some form of *mediation*, and for centuries this mediation was *writing*. Anonymous texts are not just nameless – there is always an author or a persona who penned them, but they are not identifiable, reachable, or trackable. In this sense, being anonymous is a *technique*.

It was only in the twentieth century that the adjective became a noun: *anonymity* (Ferry, 2002, p. 193). Anonymity was no longer only about written works with no attributed author or authors wishing to remain unnamed. With the advent of modernity – characterized by large-scale urbanization, industrialization, mass cultural production, and communication infrastructures – anonymity became a much more complex *technique of mediation*. The multiplication of authorities and systems capable of identifying individuals as subjects, citizens, employees, consumers (and much more) made anonymity something relevant for increasing numbers of people.
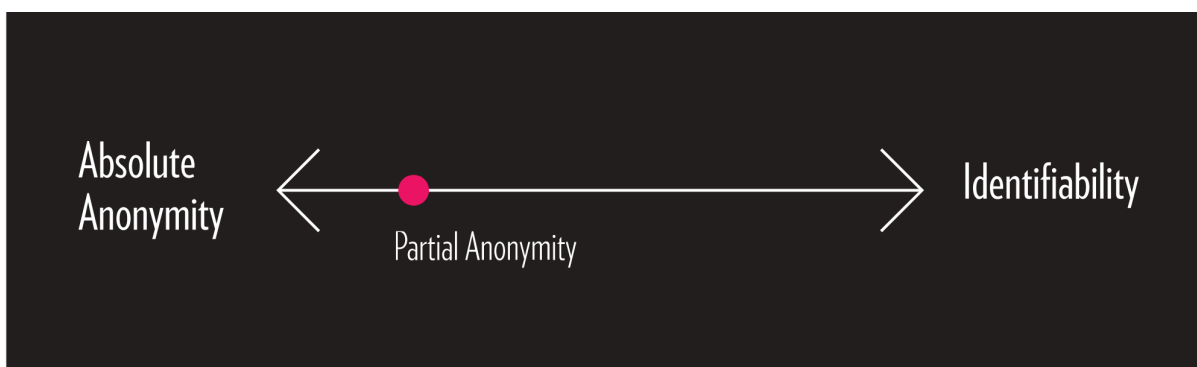
"This continuum should provide a basis to guide future research efforts to systematically examine, deconstruct, analyze and categorize social platforms along the **clear-gray-dark spectrum** in the context of the fluidity of two evolving dimensions in online social media: **technical anonymity** and **social anonymity**."

(Kenny, Lynn & Sinclair, 2022)

4.1.1. Why is anonymity important?

The importance of anonymity is *contextual.* This means that the relevance of being anonymous depends on the situation and constraints faced by individuals. In contemporary liberal democracies, we tend to take anonymity for granted in specific circumstances: we vote anonymously, we share anonymous feedback with stores or government bureaus, we make purchases with anonymous cash, we confess to a priest without sharing our name, we anonymous leak information to highlight misdemeanors, and so on. In many other circumstances, we expect not to be anonymous. But for most of human history, anonymity was the norm. Think about how many personal details we know about individuals from pre-Renaissance times: very few.

Anonymity has several implications that impact the sender and received of a message in different ways: it can grant personal safety to the bearer of a sensitive message, but also threaten the reputation of a receiver unable to control the source; it can allow secret information to be made public, but it can also put its validity in question. As a contextual technique of mediation, anonymity is also intertwined with other situated concepts such as *identity*, *privacy*, *trust*, *reputation*, and so on. Furthermore, anonymity is not necessarily a binary choice, but can be conceptualized as a continuum of information masking ranging from identifiability to non identifiability (Marx 1999, p. 100), and it includes compromises such as *pseudonymity* – the use of a unique identifier (a nickname, a credit card number, a private key) to establish a trusted relationship with another party.

4.1.2. Behavioral and social effects of anonymity

As Gary T. Marx observes, "Ironically, anonymity is fundamentally social" (1999, p. 100). An individual needs an audience of at least another person in order to be anonymous, and there is no point in being anonymous to oneself only. In this sense, anonymity is both *performative* and *relational* – it requires behavioral enactment in a social context (Scott & Orlikowski, 2014). This enactment results in *practices of anonymity*, which Bachmann et al. (2017) define as "constellations of partial unknowability, invisibility, and untraceability" which "emerge in complex intersections, entailing and combining amongst others, social practices, technologies and infrastructures, ethics and politics" (p. 243). To simplify and summarize, anonymity is a *contextual continuum of performative and relational techniques of mediation* that need to be practiced in order to have a social effect.

Given its relationship to *identity*, *privacy*, *trust* and *reputation*, anonymity has been theorized as having negative behavioral effects. One classical example is *deindividuation theory*, which posits that anonymity reduces accountability, removing the impact of sharing information on the source's reputation, thereby letting people

express themselves more freely but also resulting in a decline of trust and heightened dangers for the privacy and reputation of others. While this sort of theory has its appeal, more nuanced discussions in social psychology and communication studies have complicated the theorization of the behavioral effects of anonymity. The SIDE model (social identity model of deindividuation effects), for example, emphasizes how anonymity changes the broader balance between personal and social identity without necessarily resulting in deindividuation – its effects, again, have to be understood in context (Spears & Lea 1994).

4.1.3. Anonymity and ICTs

The *context* in which anonymity is practiced has been further complicated by the proliferation of information and communication technologies (ICTs). For centuries, the medium of writing had pretty clear and limited affordances that linked it to identity – signatures, calligraphy, etc. Mass media brought a broadcast model of communication in which a recognizable source disseminated information to an anonymous audience. ICTs afford new possibilities for identification and anonymity to both sources and receivers. Communication scholars have recognized the need for a more nuanced understanding of anonymity that differentiates between source and receiver, publicness and privacy, confidentiality and secrecy. In order to do this, two key elements are needed:

> First, anonymity must be viewed on a continuum from fully anonymous to fully identified. Thus, a source is not simply anonymous or identified, but may also be partially so. Second, anonymity is usefully viewed as a perception of the communicators involved. (Anonymous, 1998, p. 387).

In this sense, ICTs amplify the features of anonymity discussed above: it is a *continuum*, it is *contextual*, it is *relational*, and it is *performative*. The continuum of anonymity needs to be mapped in detail – not only distinguishing between anonymity and pseudonymity, but also identifying different expectations and self-perceptions from the point of view of source and receiver. Given the rapid changes in communication media, normative conclusions are not that useful, and traditional definitions of anonymity miss the point of what is at stake in the context of contemporary surveillance capitalism, where every individual on Earth is tracked according to "endless dimensions" of data (Nissenbaum, 1999, p. 141).

Summary: While the term literally means 'having no name', anonymity is a continuum of choices and options of identification. This continuum has a long history, and our current understanding of it is relatively modern. Anonymity is a social practice, and requires mediation. Anonymity is tightly connected to concepts like identity, privacy, trust and reputation, and has no determined behavioral effects. The advent of ICTs has further complicated the continuum of anonymity, requiring careful mapping of its uses.

4.2. What is reputation?

On a basic level, reputation can be defined as a refraction of the self, seen in the eyes of others. It is composed of the expressed and circulated opinions of others. Reputation is always an opinion about an opinion as it is not just our opinion of others or others' opinion of us but also what opinions we *should* have about other people – it is therefore meta and communicative. Anthropologically speaking, my own reputation can be comprehended as "what I think about what you (plural and collective) think of me". From this perspective, reputation is akin to a sense of image/impression management at the micro-level - think about everyday practices of switching identities or mannerism in different contexts from work to family (Goffman 1959). On a macro-level, across spatiotemporality, reputation can be stabilized and mundanised into specific objects in public memories, museums, and textbooks as national heroes, celebrities, historical figures and so on. Reputation in its definition already implies a hierarchical meta-system of judgements, filtering, guidance, and forth. As Origgi (2018) puts it, reputation "is an idea about ideas that guides us in selecting what we ought to think, or pretend to think, of other people".

Following this argument, there are many ways a reputation can be formulated from informal pathways such as rumors, gossips, deliberate misinformation, hearsay, to certain established system of authority or a certain impartial kind of objectivity such as classification – this would then impact the social weight given to these opinions or reputation. From informal pathways to formalized systems, reputation then goes from the ineffable to transparently measurable systems, which lies a spectrum of reputation metrics. The perceived solution to the interpersonal trust is various "objective" systems of reputations or reputational devices such as classification and ranking. Reputation, in both qualitative and quantitative sense, is the foundation of our collective mind.

4.2.1. How is it related to fame or popularity?

Reputation is a meta-concept that can cover or overlap with other conceptual frameworks such as fame and popularity. Fame, for example, has its own academic field in celebrity studies (Rockwell & Giles 2009). In this sense, seen in the realm of popular culture, fame, especially fame in contemporary settings, can be perceived as a particular expression of reputation where generating and maintaining fame or popularity itself is the profession - For example, Reality TV celebrities and Instagram/TikTok influencers. Fame in itself does not have to be justified by merits but a perpetual circulation of opinions and discourses - in simple words, being talked about a lot in public venues such as media, and fame reproduces itself with a limited life span. But any fame of a particular personality has its own limit in terms of reach depending on many factors such as education backgrounds, geographical locations, cultural histories and so forth. Reputation as a concept is more generic, therefore more unexacting and prone to distrust, which feeds into the following section.

4.2.2. Relationships with trust

Trust is inherently risky – reputation is therefore perceived as mean to avoid/minimize the risks; while at the time such perception then leads to risks where people's trust in reputation is manipulated. Trust therefore is a mix bag of emotions, values, rational assessments, personal histories, as well as bigger societal institutions such as sciences; for example, think why you trust your doctor's prescriptions. Today, the perceived solution is that interpersonal trust is various "objective" systems of reputations or reputational devices such as classification. For example, our society is obsessed with rankings, but ranking cannot be claimed as 'objective' just by outlining its various measures, mathematical formulas, and instruments; to begin with, the selection of what to measure is based on particular viewpoints, biases, and cultural institutions that reinforces the pre-existing prestige, for example, university ranking.

With various technologies such as search engines and social media, the idea of a measurable and therefore 'objective' system of reputation is deeply intertwined with our notion of trust. With information overload being the default condition of our epoch, information is always sorted, filtered, evaluated, and commented by other users, as well as algorithmic recommendations - whether it is the upvotes in reddit, pagerank, youtube algorithms, restaurant/trip reviews. In this sense, reputation, in both qualitative and quantitative sense, is the foundation of our collective mind. For Rogers (2013), search engine is an epistemological machine that in many ways determines what can and cannot be seen. There are also other collaborative filtering systems across the global internet (see our examples later). As Nick Seaver (2021)'s work demonstrates, "the dominant frame for making sense of listener variation is avidity: a level of enthusiasm for music, which manifests as a willingness to expend effort in finding listening material."

4.3. Summary

As a social system of beliefs and opinions, the term reputation already implies a hierarchical meta-system that is both qualitative (e.g. impressionistic and interpersonal) and quantitative (e.g. measurable and global). With digital technologies, these meta systems then become quantified and 'objectified' in various forms of rankings, in which biases are, quite often arbitrarily, reformatted into standards (transparent or not) that are considered trustworthy.
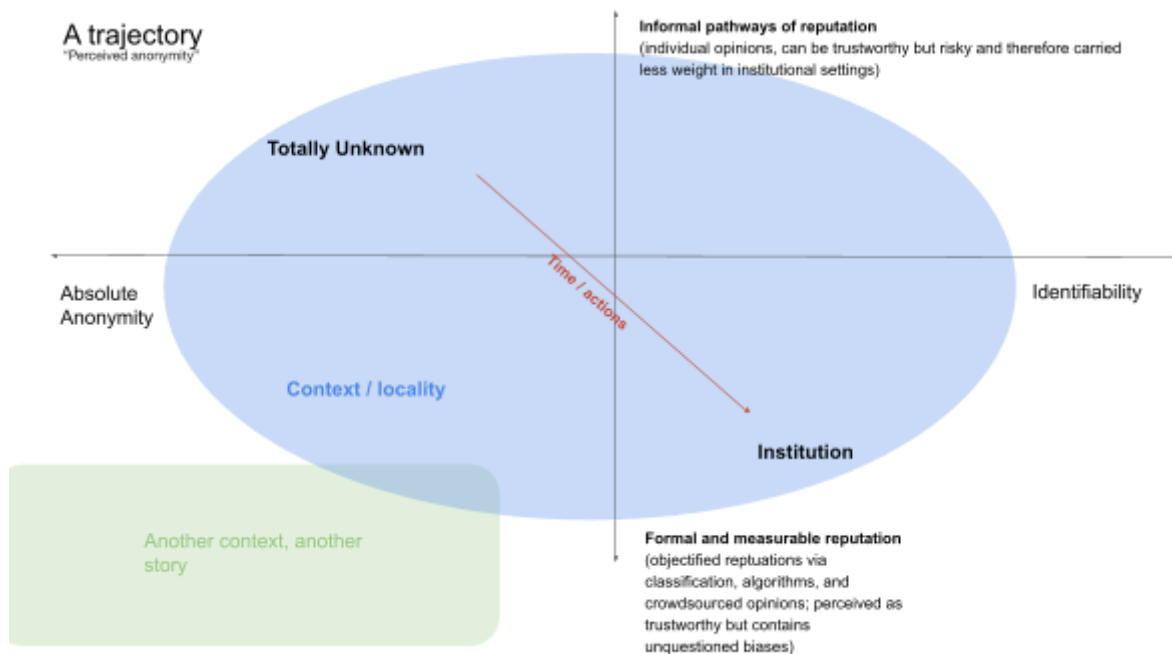
A trajectory
"Perceived anonymity"

Informal pathways of reputation
(individual opinions, can be trustworthy but risky and therefore carried less weight in institutional settings)

Totally Unknown

Time / actions

Absolute Anonymity

Identifiability

Context / locality

Institution

Another context, another story

Formal and measurable reputation
(objectified reputations via classification, algorithms, and crowdsourced opinions; perceived as trustworthy but contains unquestioned biases)

Diagram of the relationship between anonymity and reputation

4.4. Takeaways

In this section, we have defined the two main concepts of this report: *anonymity*, and *reputation*. After illustrating their relationship to other concepts like *trust*, we have discussed their *effects* throughout history, deriving some key takeaways:

1. Anonymity is a **contextual continuum** of **performative and relational techniques of mediation** that need to be practiced in order to have a **social effect**. This means that anonymity is not a fixed position but has various degrees, and always happens through interaction between humans and technological media
2. Reputation is a hierarchical meta-system that contains both qualitative and quantitative information, which should not invalidate each other and must be seen relationally. Objectification of reputation in the form of statistical rankings and algorithms should not be taken for granted as the human factor (or qualitative factor) cannot be entirely eliminated but merely presupposed and concealed

5. Case Studies: Anonymity & reputation in context

In this section, we use a *case study approach* to illustrate how anonymity and reputation changed through history and in different global contexts, demonstrating that specific configurations of anonymity and reputation lead to situated uses and behaviors.

5.1. Early online communities in the United States

As discussed above, ICTs offer both sender and receiver new possibilities for identifying themselves through a continuum including various forms of partial anonymity. From the early years of the internet, choosing the ideal position on this continuum was a key concern for users. The classic 1993 *The New Yorker* cartoon drawn by Peter Steiner, featuring a dog sitting at a computer and saying "On the internet, nobody knows you're a dog" is a key example of this concern. But the internet is not a singular medium, and has historically supported protocols

and methods with their specific forms of partial anonymity. When using e-mail, for example, a sender needs to self-identify themselves, but the stable e-mail address links them to a fixed identity, and the email service provider often collects and stores personal information at registration. On the bulletin board systems (BBSs) popular in the 1980s and 1990s, users responded to the possibility of highly anonymous communication by creating fixed identities and tying their reputation to them through both interactional rules and social roles (Myers, 1987). In Multi-User Dungeons (MUDs) and online communities, the power granted by anonymity led to the emergence of controversial practices like trolling and the occurrence of traumatic incidents. In one notorious example, the sexual harassment of user avatars in the LambdaMOO community led users to discuss the boundaries between reality and virtuality, and to adopt stricter moderation policies; strikingly, the culprit of the "cyber rape" was found to be not an individual, but a group of university students controlling a shared account (Dibbell, 1993). As these examples testify, design decisions about the degree of partial anonymity offered by a communication medium are not definitive, and often have unintended consequences that lead to the emergence of new identity practices or the negotiation of new rules of behavior.

5.2. 2000s forums (4chan reddit)

Since the 2000s, internet culture has quickly become mainstream, and being part of online communities, social networks and virtual worlds has become an experience shared by billions of people around the world. In this context, design decisions about specific configurations of anonymity and reputation are critical. To illustrate this, compare these two case studies of well-known communities: 4chan and Reddit.

5.2.1. 4chan

An American imageboard website founded in 2003. Users can create threads by sharing an image, and reply to existing threads with image or text comments. 4chan threads expire and are automatically pruned if they are not active enough. 4chan has no registration system, and users can post under any nickname, or not choose one and remain anonymous – in fact, "Anonymous" is the default nickname displayed when a user decides to not identify themselves (Knuttila 2011). Posts are identified with a "tripcode" – the numeric hash of a user-chosen password – that allows users to be identifiable across posts without sharing any personal details if they want. Anonymity is such a central feature of 4chan that it spawned a global, distributed movement: the Anonymous hacktivist collective (Coleman, 2015). 4chan's default anonymity allows users to post controversial, extreme or even harmful content that they would normally not do under their name or nickname – in this sense, it protects reputation by allowing for "undisciplined performativity" (Curlew, 2019). On the other hand, 4chan users want to cultivate reputation in some cases, and make use of different affordances: some, like tripcodes, are offered by the imageboard's design; others, like posting time-stamped photos or other forms of identity verification, are user-driven.

5.2.2. Reddit
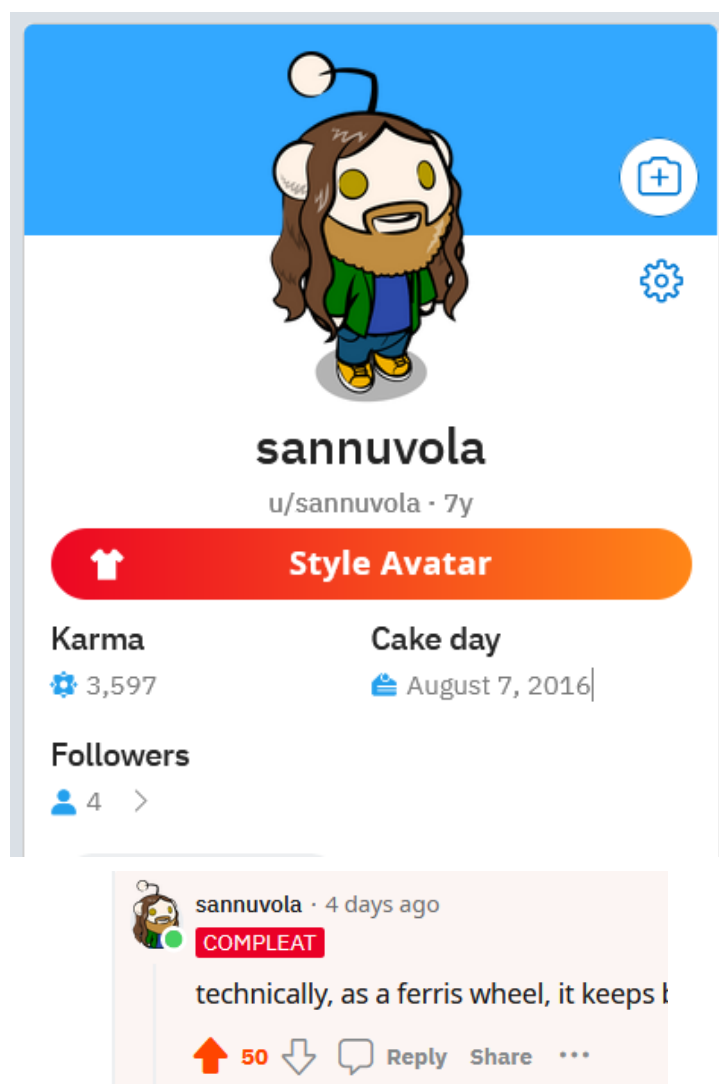
An American discussion board website founded in 2005. Users can create their own community (a 'subreddit'), start threads by sharing a link, a text, one or more images, or a video, and reply to existing threads with textual comments. In contrast to 4chan, content is not automatically deleted, and might be archived after months. Reddit requires user registration, which ties an account to an email account used to verify it. A user's profile aggregates not only basic information on their Reddit account, but also all the threads they opened, and the

comments they posted on other threads. User reputation is not only tied to the publicity of user profiles, but is also quantified through two systems: votes, and karma. Users can rate a given thread or comment post with an upvote or a downvote, influencing its visibility – most upvoted content tends to be displayed near the top of a page, and downvoted content moves to the bottom, and the most popular posts across Reddit become displayed on the website's frontpage. Users gain karma, a numerical score, by participating in different subreddits, contributing content that is upvoted by others, leaving comments; being downvoted reduces karma, and Reddit does not disclose other specifics of its karma algorithm. Reddit's partial anonymity means that, while users do not need to share personal information about themselves, their activities are tied to multiple reputational metrics – in this sense, public profiles and karma consolidate user reputation. On the other hand, users who seek a more anonymous engagement – for personal safety or other reasons – make use of untraceable functions like voting to claim a degree of agency in their community (Brown et al., 2018).

These two case studies of established American platforms that are quite well-known in the English-speaking world, highlight the elastic and open-ended relationship between the continuums of anonymity and reputation. On 4chan, full anonymity encourages fringe behaviors, but users make use of design affordances to establish their own reputation systems. Conversely, on Reddit, where multiple systems encourage the quantification of reputation, users negotiate ways of maintaining anonymity while still steering community activity. One conclusion that can be drawn from this: design decisions about the intersection of anonymity and reputation need to take into account *elasticity*, and expect users to push and pull a product or platform towards their needs.

5.3. Early internet communities in East Asia

Japan and China as a culturally divergent case of anonymity and collectivity

East Asia is often seen as alternative to the dominant notions of privacy/public sphere, individuality/collectivity, and therefore anonymity and reputation - these differences are often attributed to Confucianism without specifying the radical internal divergences of cultural practices and internet governance within Asia. However, each east Asian country has its own approach to privacy and anonymity: Japan never implemented a consistent system of real name policy for the internet and anonymity remains a core concern in public life while China has a rather comprehensive system of real-name registration for any online communication and anonymity is not an option anymore, at least seen from the perspective of governance.

5.3.1. 2Channel

The logo design of 2ch (left); Densha Otoko (Shosuke Murakami, 2005); Suicide Circle (Sino Sono, 2001)

Japan's 2ch, founded by Hiroyuki Nishimura in 1999, was a forerunner in the category of anonymous image board or BBS but it was preceded by Ayashii World in 1996 and Amezou World in 1998. 2channel was then followed by its backup site Futaba channel in 2001. These sites had always been precarious and close to shutdowns. The main features of these forerunners - boards, threads, "sage" (discontinue a thread), topic recurrency, tripcodes, and and other basic structures - are inherited by 4chan and other following imageboards in the East and West. The core feature of 2ch and Ayashii as image boards was its anonymity, as a safe haven for Japanese people to speak their minds. One of core concepts coming from the imageboard culture is 名無し, or "nanashi" (literally no name person or anonymous). There are two options of anonymity on 2chan,

> Pseudonym anonymity (this is very similar to modern social media such as Twitter but without a profile that aggregates your posts): your handle becomes your identity; consistently posting under the same pseudonym you can accumulate credibility and reputation. This is entirely voluntary though as you can have as many as handle as you like and it's possible for others to impersonate you - the sense of community also depends on a certain degree of mutual trust because every post was made in isolation (i.e. as a reader, there was no sure way of knowing who was posting multiple times). The reputation gained in this case can be limited to a single board.

> Nameless anonymity; 名無し: it is understood generally that this is predominantly the way people opt to post but this claim can be disputed by a study by Matsumura et al (2004). In different boards, people opted for different strategies of anonymity; in some cases, reputation is valued more than anonymity and the OP (original poster) would opt for pseudonym but the replies will be in nameless anonymity.

Seen in the light of its Western counterpart 4chan, Japan's 2ch is a peculiar case because it is seen as embedded in a particular Japanese need to speak without consequences; "Nanashi is more important in Japan than it is in the West" (Yotsuba Society). This results in both stories of solidarity and hatred/delusion; two movies captured the spirit of 2ch pretty well: *Densha Otoko* tells the story of collective efforts of the discussion boards coaching a male "otaku" gradually going out of comfort zone and starting a relationship with a woman for the first time; *Suicide Circle*, where online anonymous discussion boards organised mass suicide events.

5.3.2. Nimingban: a facade of anonymity

The mascot of nimingban (left) and the various kaomojis popular on nimingban (right); Source: link

China's nimingban (literally "anonymous board") was originally created by Acfun's staff in 2011 after 2ch's model; Acfun was modeled after Niconico Douga. When nimingban was created, its model was a complete copy of 2ch therefore anonymity was actually enforced until several events of "sensitive" political discussions that led to its temporary censorship. Unlike almost all the online forums in China, there were no pre-emptive censorship (such as censoring certain terms/words/characters) on the website; but due to its multiple cycles of closures, negotiations with the Chinese cyber police, in 2017 the website instigated a system dubbed "front-end anon back-end real-name" [sic] where a new user of the website has reply to specific thread to be assigned a "cookie" (which is verified by a one time password sent to a China-mainland phone number, which is again registered with a real name and Chinese ID). While nimingban's facial anonymity is in stark contrast to 2ch's voluntary or rather socially enforced anonymity, its subculture is remarkably similar to that of 2ch and the website also developed its own mascot, kaomojis and various subcultural practices over time.
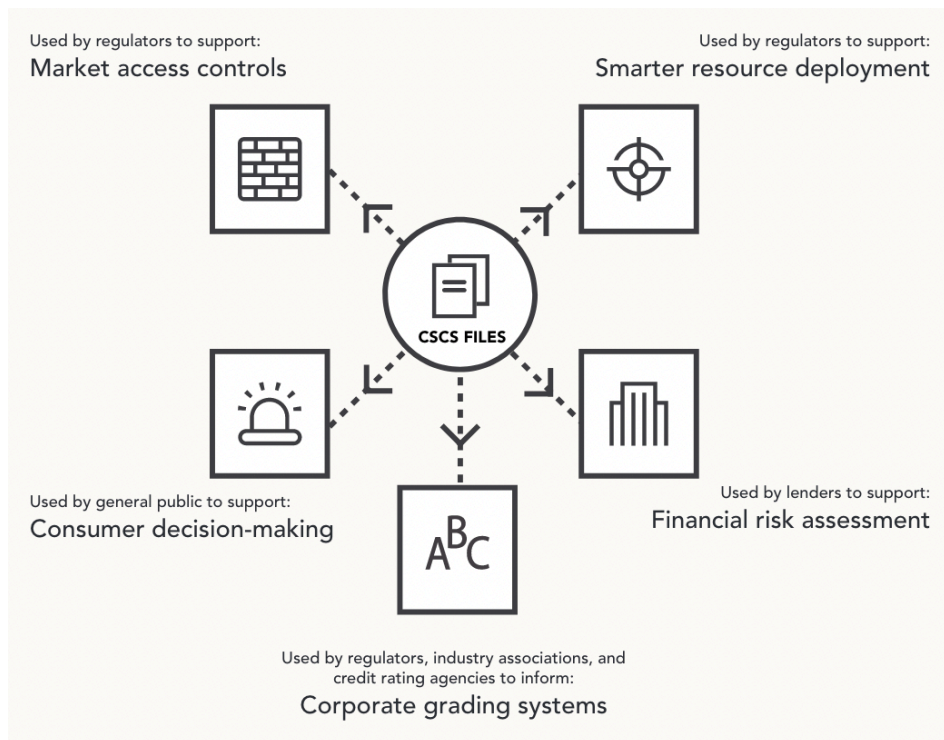
The two cases established culturally and regionally specific notions of anonymity and reputation even when the system is designed to encourage or enforce anonymity, users may still choose to opt out of it due to a need for reputation; and in other scenarios, when the state/regulator/administrator does not allow full anonymity it can still be pursued or at least performed in creative ways, socio-technically negotiated by users themselves..

5.4. China's Social Credit System

The Social Credit System (SCS) is a well-known example of a reputation system, developed in China over the past few decades. The widespread media coverage of the system as an Orwellian surveillance device for individuals is an overstatement to say the least[1]. For instance, there is no single unified score for individuals. The system has low automation and its implementation is still largely a work in progress. At the technical and organizational level, the SCS system has two main components: a record of regulatory compliance (CSCS files) and a system of penalties and market access restrictions.

---

[1] https://merics.org/en/report/chinas-social-credit-system-2021-fragmentation-towards-integration

The SCS is a general framework for market trust, inspired by Western credit rating approaches. The idea was first proposed in 1999 as a part of China's accession to the World Trade Organization (WTO), and it became a national law in 2019. The concept of a social credit system originated in the 2000s with the central government's desire to show local and foreign investors which companies and territories are (un)worthy of their trust. The framework law finally promulgated in 2014 concerns above all companies and points to specific practices such as tax evasion, non-payment of debts, pollution, accidents in factories, food safety and sanitation, academic fraud or counterfeiting. Evildoers identified through cooperation protocols between local governments and private companies are added to a database of "untrustworthy" (严重失信) parties to operate in a particular industry.

Used by regulators to support:
Market access controls

Used by regulators to support:
Smarter resource deployment

CSCS FILES

Used by general public to support:
Consumer decision-making

Used by lenders to support:
Financial risk assessment

Used by regulators, industry associations, and credit rating agencies to inform:
Corporate grading systems

From *Trivium China* (2020)

So far, pilot projects have been set up in about 30 cities. Cooperation between companies and local governments is proving to be laborious, resulting in fragmented and inefficient administrative protocols. Abuses such as the public humiliation of offenders raise the question of additional sanctions not provided for by the law. Finally, the polysemy of the term credit (信用) in Chinese sows confusion in legal debates. The implementation appears quite inconclusive and testifies to a very real problem: the difficulty of designing systems that requires cooperation between administrations, the private sector, software and machines.

5.4.1. The SCS Data Platform

Far from being a monolithic system, SCS relies on a general framework promoted as a national law that took 25 years to write. The implementation of these general guidelines is left to the discretion of various governmental bodies, companies and institutions - with each their own technical and organizational culture. For instance, each provincial government is in charge of gathering data from local institutions on its own database to be shared across National Credit Information Sharing Platform (全国信用信息共享平 台, NCISP). The approach (and available resources) varies greatly across provinces, with some governments deciding to have part of their data

available to the public and others keeping everything entirely private. The ethical questions about what to publicize and what to keep private (especially blacklists) are subject to settlements in court to ratify interpretations of the law in jurisprudence.
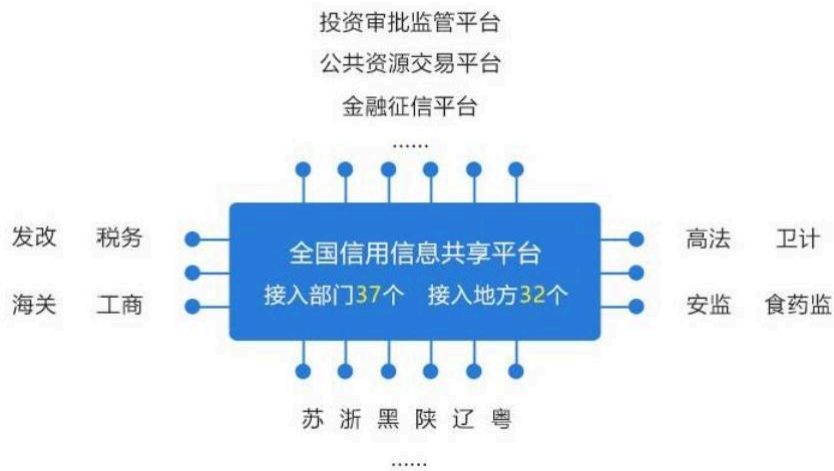

Fig. Technical schema of the data platform organization

In a sense, the NCISP wants to become an universal reputation system. The specifications were drafted by the social credit standards drafting committee SAC/TC470 (全国社会信用标准化技术委员会) which is comprised of the National Development and Reform Commission (which has supervised the development of the legal framework since 2015), the China Standardization Research Institute, and a number of other research institutions and private technology companies[2].

When it comes to data, there are two primary sources/types of data are considered

- Public Credit Information (公共信用信息) shared on NCISP
- Market Credit Information recorded by 3rd party players (often loosely defined)

The data is recorded by various government bodies using the Unified Social Credit Identifier (USCI). However, the data distributed by agencies shows large gaps in redundancy, accuracy, and completeness.

5.4.2. Mechanisms and Regulatory Systems

Beyond the data aspect, the second important aspect of the SCS is dedicated to the design of mechanisms and regulatory systems that can be used to enforce decisions and take actions according to specific scores on the SCS platform. This part is particularly controversial and is approached as a legal discussion that has been ongoing for years. The table below shows the main tools and mechanisms used in SCS cases.

Blacklist / red list with categories
ex. "generally untrustworthy," "seriously untrustworthy," and "particularly seriously untrustworthy."

---

[2] The published standard has been released https://webstore.ansi.org/standards/spc/gb394442020

<table>
<tr><td>
Unified Rewards and Punishments
decided by Memorandums of Understanding (MOUs) between agencies

Objections and credit repair
Moving but integral part of the framework

Market access controls in wider industry
Private companies relies on these lists to build their own platform (ex. e-commerce)
</td></tr>
<tr><td>
SCS Mechanisms and Regulatory Systems
</td></tr>
</table>

Few basic rules are used for legal cases. A written record of all decisions and actions taken has to be kept to ensure accountability and transparency. There can be no issuance of penalties or blacklisting without human intervention. However, the legal framework governing these actions does not include technical specifications. As a result, it evolves through jurisprudence for litigious claims, such as un-blacklisting or credit repair.

5.5. Takeaways

The case studies presented in this section illustrate how different configurations of anonymity and reputation differ throughout history and in different global contexts. We offer three takeaways from this:

1. Anonymity is *elastic*, and users will push and pull on it according to their expectations and needs. The contrasting examples of 4chan and Reddit are one example of how users balance anonymity and reputation against platform design.
2. Cultural contexts matter but not in determinist ways. The East Asian example shows that uses may vary according to local/cultural settings but can also largely differ within the same geographical space.
3. The general framework for SCS exists first as an overarching statement that is used to define technical specifications (the legal framework). To prevent abuses, enforcement of this system is better left with little to no automation (through human/ community action).
4. From SCS, we can also see that the real bottleneck and trickiest part is the integration of existing systems, both in terms of interoperability and change of habits for humans.

6. Recommendations:

In this section we look back at the concepts discussed in the report and at the case studies we studied to extrapolate five recommendations. These recommendations answer the question:

Which practical design knowledge can be derived from this history and discussion?

➔ **MAP CONTINUUMS, NOT BINARIES**
When thinking about concepts like anonymity, reputation, trust, and so on, avoid constructing arbitrary binaries (anonymous/not anonymous, with/without reputation, trust/trustless, etc.). Instead, draw a line between these extremes and map different possibilities on it. This will reveal actual expectations and needs, and help design a better product that can respond to them.

- ➔ **ACCOUNT FOR ELASTICITY**
  Expect users to push and pull against your design decisions about key concepts like anonymity and reputation. While users might be on board with your conceptual decisions, their practices might challenge them in edge cases, reconfiguring their relationship through emerging strategies. It is important to <u>leave room for elasticity</u> in your product.
- ➔ **QUESTION UNIVERSALITY**
  In a world where local history and context still play a key role in shaping how technologies are adopted and used, the term "universal" is a slippery one. Rather than aiming for universality, <u>think about how a product can achieve large-scale adoption while maintaining flexibility and elasticity</u>. The most successful general-purpose technologies are reconfigurable, privileging adaptability to emerging needs over imposed universal uses.
- ➔ **PROVIDE FLEXIBLE WAYS TO ENFORCE RULES**
  The enforcement of a reputation system can lead to unexpected outcomes and abuses that can sometimes be hard to manage with a community. Room should be left for designing systems to enforce decisions based on reputation (downgrading, banning, etc) - maybe a blueprint can be provided.
- ➔ **INTEROPERABILITY IS KEY**
  You can't just merely code anonymity or reputation systems into existence. Any reputation system is always a mix of pre-existing ones, augmented by the affordances of user needs and practices. The portability of reputation can only exist by cooperating with existing reputation systems (both online and offline) which requires a very different kind of work than designing a new system

7. References

Anonymous (1998). To reveal or not to reveal: A theoretical model of anonymous communication. *Communication Theory*, 8(4), 381-407

Bachmann G., Knecht M., Wittel A. (2017). The social productivity of anonymity. *Ephemera: Theory & Politics in Organization*, 17, 241-258.

Vincent Brussee. 2022. "China's Social Credit Score – Untangling Myth from Reality." February 11, 2022. https://merics.org/en/opinion/chinas-social-credit-score-untangling-myth-reality.

Creemers, Rogier. 2018. "China's Social Credit System: An Evolving Practice of Control." SSRN Scholarly Paper ID 3175792. Rochester, NY: Social Science Research Network. https://doi.org/10.2139/ssrn.3175792.

Ferry A. (2002). "Anonymity": The literary history of a word. *New Literary History*, 33(2), 193-214.

Goffman, E. (1959). The presentation of self in everyday life. University of Edinburgh Social Sciences Research Centre.

Kenny D., Lynn, T., Sinclair, G. (2022). From clear to dark: The social media platform anonymity continuum. *DIGITAL 2022, Advances on Societal Digital Transformation*, 13-17 Nov 2022, Valencia, Spain.

Marx, G. T. (1999). What's in a name? Some reflections on the sociology of anonymity. *The Information Society*, 15, 99-112.

Nissenbaum H. (1999). The meaning of anonymity in an information age. *The Information Society*, 15, 141-144.

Origgi, Gloria. 2018. Reputation: What It Is and Why It Matters. Princeton University Press.

Rogers, R. (2013). Digital Methods. MIT Press.

Schaefer, Kendra. 2020. "China's Corporate Social Credit System." Trivium China. https://www.uscc.gov/sites/default/files/2020-12/Chinas_Corporate_Social_Credit_System.pdf.

Scott, S. V., Orlikowski W. J. (2014). Entanglements in practice: Performing anonymity through social media. *MIS Quarterly*, 38(3), 873-893.

Seaver, N. (2021). Seeing like an infrastructure: Avidity and difference in algorithmic recommendation. Cultural Studies, 35(4–5), 771–791. https://doi.org/10.1080/09502386.2021.1895248

Spears R., Lea M. (1994). Panacea or panopticon? The hidden power of computer-mediated communication. *Communication Research*, 21, 327-459.