Privacy-First Architecture for Identity-Protected Digital Communication v1.0

Karthikeyan Sukumaran

karthikeyansukumaran@naachichat.com

Legal Disclaimer This White Paper is published by Naachi Research and Innovation Foundation solely to present its research findings and proposed technical architecture. Nothing in this document constitutes an offer to sell or solicitation of an offer to buy any tokens, securities, or financial instruments. The Foundation is a social venture focused on research and development in emerging digital technologies. The technical architecture described herein represents current plans which may change. Success depends on many factors outside the Foundation's control, including market conditions, regulatory developments, and technology adoption. Any forward-looking statements are based solely on the Foundation's analysis and may prove incorrect.

Abstract

This paper proposes a messaging architecture that decouples communication identity from personal phone numbers. In current systems, the phone number serves simultaneously as a user identifier, authentication credential, and contact address creating a single point of failure that exposes users to enumeration attacks, harassment, and unsolicited communication at scale. We present a four-layer architecture: (1) a username-based identity layer that masks phone numbers by default, (2) a consent-gated connection protocol that requires mutual approval before messaging, (3) an institutional verification framework using cryptographic attestations, and (4) an invitation-chain traceability system for platform accountability. The architecture is analysed against India's 1.18 billion subscriber ecosystem where phone numbers are linked to Aadhaar, bank accounts, and UPI, making exposure uniquely consequential. We present evidence from NCRB, CERT-

In, and TRAI data showing 86,420 cybercrime cases (2023), 41.68 billion spam calls (2025), and a 3.5 billion account enumeration vulnerability in WhatsApp [1]. The proposed system addresses these vulnerabilities while maintaining compliance with the Digital Personal Data Protection Act, 2023 and the IT (Intermediary Guidelines) Rules, 2021.

1 Introduction

Digital messaging systems have become the primary communication infrastructure for over 800 million Indians. Yet these systems share a fundamental architectural flaw: they use personal phone numbers as the primary user identifier. This means that initiating a conversation with someone, joining a group, or contacting a business requires exposing a piece of personal data that, in the Indian context, is linked to biometric identity (Aadhaar), financial accounts (UPI, bank), voter registration, and government services.

The consequences are measurable. The National Crime Records Bureau recorded 86,420 cybercrime cases in 2023, a 31.2% increase from the previous year, with a conviction rate below 3% [2]. CERT-In handled 1,592,917 cyber incidents in 2023 [3]. Truecaller's 2025 India report identified 41.68 billion spam calls and 129.03 billion spam SMS messages, with India accounting for 67% of global spam call volume [4][5].

The architecture proposed in this paper treats these not as enforcement failures but as design failures. When a system requires phone number exposure as a precondition for communication, no amount of downstream regulation can fully mitigate the resulting privacy loss. The solution must be architectural.

This paper suggests, a messaging system designed from first principles around identity protection. The remainder of this paper is organised as follows. Section 2 describes the problem architecture. Section 3 presents the system design. Sections 4 through 7 describe each architectural layer in depth. Section 8 analyses the regulatory alignment. Section 9 covers threat analysis. Section 10 discusses implementation. Section 11 presents the use case architecture. Section 12 concludes.

2 Problem Architecture

2.1 The Phone Number as Single Point of Failure

In existing messaging platforms, the phone number serves four simultaneous functions: registration credential, user identifier, contact address, and authentication factor. This conflation creates a cascading vulnerability. A study presented at NDSS 2026 by Gegenhuber et al. demonstrated that WhatsApp's contact discovery mechanism allowed enumeration of 3.5 billion accounts at 100 million phone numbers per hour, exposing phone numbers, profile photographs, and encryption keys without rate limiting [1]. Of the enumerated accounts, 57% had public profile photographs and 29% had text in profile fields.

Research by McDonald et al. (2021) at the ACM CHI Conference established that phone numbers as identifiers create pervasive privacy, security, and access risks, with participants reporting harassment from phone number recycling where reassigned numbers carry the digital history of previous owners [6].

2.2 The Indian Context: Compounding Vulnerability

In India, the phone number carries uniquely heightened sensitivity because it functions as a nexus connecting multiple identity systems. Dev et al. (2018), surveying 213 Indian WhatsApp users, found that phone numbers in India are typically linked to voter identification and financial data, creating compounding vulnerability [7]. A follow-up study of 674 users found Indians showed significantly higher concerns about social contact from professional colleagues compared to users in other countries [8].

Table 1 summarises the identity systems connected through a single Indian mobile number.

Table 1: Identity Systems Linked to a Single Indian Phone Number

System	Linkage Mechanism
Aadhaar	Mobile number mandatory for biometric authentication
Bank Accounts	KYC-linked mobile, OTP-based transaction auth
UPI	Phone number is the payment address prefix
Voter ID	Linked via Election Commission registration
Government Services	DigiLocker, CoWIN, Ayushman Bharat all phone-linked
Messaging (WhatsApp)	Phone number = user identity, visible to all contacts

2.3 Scale of Exposure

The scale of India's digital ecosystem makes these architectural vulnerabilities consequential at a level unmatched by most countries. Table 2 presents the key metrics.

Table 2: India's Digital Communication Scale (2024-2025)

Metric	Value
Telephone subscribers	1.18 billion
Internet subscribers	~900 million
WhatsApp MAU (India)	500-600 million
Cybercrime cases (2023)	86,420
Spam calls identified (2025)	41.68 billion
Spam SMS identified (2025)	129.03 billion
CERT-In incidents handled (2023)	1,592,917
Cybercrime conviction rate	<3%

2.4 Group Communication: Amplified Exposure

Group messaging amplifies the phone number problem by order of magnitude. WhatsApp Communities expose phone numbers to all subgroup members through the Announcement Group. Guild (2020) documented that WhatsApp group member phone numbers became publicly visible on Google search results [13]. Rosenberg and Asterhan (2018) found teachers' phone numbers necessarily shared in classroom groups, creating boundary violations [14]. Tamil Selvan and Kalaiyarasan (2024) identified privacy and security concerns as a key drawback of WhatsApp use in education [15].

In an Indian university context, a single student joining 10 course groups, 3 club groups, and a hostel group exposes their phone number to potentially 500+ contacts most of whom are strangers. Each of these contacts can then initiate unsolicited communication at any time, indefinitely.

3 System Design

The platform's architecture addresses the vulnerabilities described in Section 2 through a four-layer design. Each layer operates independently but is designed to compose with the others. The layers are:

Table 3: Architecture Layers

Layer	Component	Function
1	Identity Masking	Username-based communication; phone number used only for verification
2	Consent Gateway	Connection request protocol; no messaging without mutual approval
3	Institutional Verification	Cryptographic attestation of professional/organisational identity
4	Invitation-Chain Traceability	Referral graph for accountability without mass surveillance

The design philosophy follows Cavoukian's seven Privacy-by-Design principles [16]: proactive not reactive, privacy as the default, privacy embedded into design, full functionality (positive-sum, not zero-sum), end-to-end security, visibility and transparency, and respect for user privacy. The system is designed for India's regulatory environment under the DPDPA 2023 [17] and IT Rules 2021 [18].

4 Layer 1: Identity Masking

4.1 Description

The identity masking layer decouples the communication identifier from the phone number. Users register with a phone number (for OTP-based verification) but communicate exclusively through a self-selected username. The phone number is stored in the system's authentication database but is never transmitted to other users, displayed in chat interfaces, or exposed through contact discovery APIs.

This follows the model pioneered by Signal in February 2024, where phone numbers were hidden by default from non-contacts and usernames were designed to not be stored in plaintext [19]. The Freedom of the Press Foundation subsequently recommended this as a critical security practice for journalists [20].

4.2 Architecture

The identity layer maintains two separate data stores:

Authentication Store: Phone number → Hashed credential mapping. Used exclusively for OTP verification during registration and account recovery. Phone numbers are stored as salted hashes. This store is not queryable by other users or through any public API.

Identity Store: Username → Public profile mapping. Contains the user's chosen display name, optional bio, verification status, and institutional affiliation. This is the only identity information visible to other users.

The separation ensures that compromise of the Identity Store does not expose phone numbers, and that the Authentication Store cannot be used for social graph enumeration. This directly addresses the Gegenhuber et al. vulnerability [1], where WhatsApp's unified contact discovery

API allowed mass enumeration because the same phone number served as both authentication credential and contact identifier.

4.3 Contact Discovery

Unlike phone-number-based systems where contact discovery automatically scans the user's address book, the platform uses an explicit search model. Users find other users by searching for their exact username. There is no address book upload, no phone number matching, and no suggested contacts based on phone proximity or shared contacts. This eliminates the primary vector for enumeration attacks and unsolicited contact.

The trade-off is reduced convenience in initial connection. This is intentional. The friction of sharing a username through an out-of-band channel (face-to-face, email, business card, QR code) serves as the first layer of consent.

4.4 Defence Against Enumeration

Username search is rate-limited to prevent automated enumeration. Failed search queries do not reveal whether a username exists the system returns identical responses for nonexistent usernames and users who have disabled discoverability. Additionally, usernames can be changed at any time, breaking any previously established linkage.

5 Layer 2: Consent Gateway

5.1 Description

The consent gateway implements a connection request protocol that requires explicit mutual approval before any messaging can occur. This is architecturally distinct from existing messaging platforms where possession of a phone number grants immediate messaging access.

5.2 Protocol

The connection protocol operates as follows:

Step 1 Request: User A locates User B through username search and sends a connection request. The request contains A's username, verification status, institutional affiliation (if any), and an optional introductory message limited to 200 characters.

Step 2 Review: User B receives the request in a separate inbox (not the main chat list). B can view A's public profile, verification status, and institutional affiliation before making a decision.

Step 3 Decision: User B accepts, declines, or ignores the request. Declined requests cannot be re-sent for a configurable cooldown period (default: 30 days). Ignored requests expire after 7 days.

Step 4 Connection: If accepted, a bidirectional messaging channel opens. Either party can revoke the connection at any time, immediately terminating messaging access.

This protocol eliminates the class of attacks where acquiring someone's phone number (through a data breach, business card exchange, group membership, or social engineering) automatically grants messaging access. In this architecture, obtaining someone's username grants only the ability to send a request not to communicate.

6 Layer 3: Institutional Verification

6.1 Description

The institutional verification layer allows organisations to cryptographically attest that a user is their authorised representative. When a bank relationship manager, university faculty member, or healthcare professional communicates through this platform, their profile displays the verified institutional affiliation. This converts identity claims from unverifiable assertions into provable attestations.

6.2 Verification Architecture

The verification process follows the three-party model defined in the W3C Verifiable Credentials 2.0 standard [21]:

Issuer: The organisation (bank, university, hospital) registers as a verified institution on the platform. Institutional verification requires legal documentation (GST registration, AISHE code, RBI licence number, or equivalent). The institution receives a signing key pair.

Holder: The individual employee or representative receives a credential signed by their institution. This credential contains: the user's username, the institution's verified identity, the role/department, and an expiry date.

Verifier: Any user of the platform viewing the holder's profile can see the verified institutional badge, the institution name, and the role. They can verify the credential's validity by checking the institution's public key without needing to contact the institution directly.

Credentials expire and must be renewed, ensuring that employees who leave an organisation cannot continue to represent it. Institutions can also remotely revoke credentials.

6.3 Trust Hierarchy

The system supports three levels of verification:

Level 0 Unverified: Default state. Username only. No institutional claims.

Level 1 Peer Verified: User has been verified through QR code scan in physical proximity by an existing verified user, or through mutual friend attestation.

Level 2 Institution Verified: User holds a valid, non-expired credential from a registered institution. Profile displays the institution name and verified badge.

This graduated trust model allows the system to function for casual users (Level 0) while providing strong identity guarantees for professional contexts (Level 2).

7 Layer 4: Invitation-Chain Traceability

7.1 Description

The traceability layer maintains a directed acyclic graph (DAG) of how each user entered the platform. Every account records which existing user invited or verified them. This creates an accountability chain without requiring mass surveillance of message content.

7.2 Design

When User A invites User B to the platform, the system records the edge A → B in the invitation graph. This graph is visible only to platform administrators and is not accessible to other users. If User B subsequently invites User C, the chain A → B → C is recorded.

If User C engages in platform abuse (harassment, fraud, impersonation), administrators can trace the invitation chain to identify not only the abuser but the pathway through which they entered the ecosystem. This creates social accountability: users have an incentive to invite only people they trust, because their own reputation is indirectly linked to the behaviour of their invitees.

The invitation graph also serves as an anomaly detection layer. A single account that generates an unusually large number of invitations, or whose invitees have high abuse rates, can be flagged for review.

7.3 Privacy Constraints

The invitation graph stores only the minimum data required: inviter username, invitee username, timestamp, and verification method used. It does not store message content, contact lists, or behavioural data. Access to the graph is restricted to designated administrators and requires documented justification. This design complies with the data minimisation principle under DPDPA Section 4(2) [17].

8 Regulatory Alignment

8.1 DPDPA 2023 Compliance

The Digital Personal Data Protection Act, 2023 [17], with implementing DPDP Rules 2025 notified on November 14, 2025 [22], establishes India's data protection framework. The platform's architecture is designed for native compliance:

Consent (Section 6): The connection request protocol implements explicit, informed consent before any data exchange occurs. Users actively choose each communication partner.

Purpose Limitation (Section 4): Phone numbers are collected solely for verification and are not used for contact discovery, advertising, or third-party sharing.

Data Minimisation (Section 4(2)): The identity masking layer ensures only the minimum necessary data (username) is exposed to other users.

Right to Erasure (Section 12): Account deletion removes the user's identity from all active connections. Phone number hashes are purged from the authentication store.

Penalties: Non-compliance under DPDPA carries penalties up to ₹250 crore. A privacy-by-design architecture provides structural compliance rather than bolted-on controls.

8.2 IT Intermediary Guidelines 2021

The IT (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021 [18] require platforms exceeding 5 million registered users to appoint India-based compliance officers, publish monthly transparency reports, and maintain a grievance redressal mechanism. The invitation-chain traceability layer directly supports the traceability requirements under Rule 4(2), which mandates identification of the first originator of information upon court order. The platform architecture stores originator information (the invitation chain and message metadata) without requiring backdoors in encryption.

9 Threat Analysis

9.1 Enumeration Attacks

The Gegenhuber et al. attack [1] exploited WhatsApp's contact discovery API to enumerate 3.5 billion accounts. This platform eliminates this vector entirely: there is no contact discovery API, no phone-number-to-username mapping available to users, and no address book upload. Username search is rate-limited and returns uniform responses regardless of whether a username exists.

9.2 Unsolicited Communication (Spam)

The consent gateway makes bulk unsolicited messaging structurally impossible. A spammer would need to: (a) discover target usernames (no enumeration API), (b) send individual connection requests (rate-limited), and (c) have each request individually accepted. At any step, the attack fails. This contrasts with phone-number-based systems where a single data breach can expose millions of contactable numbers.

9.3 Identity Misrepresentation

In current systems, anyone can claim to represent an organisation in their profile bio with no verification. The platform’s institutional verification layer requires a cryptographic credential issued by a registered institution. Misrepresentation would require compromising the institution's signing key a significantly higher bar than editing a text field.

9.4 Harassment via Platform

The connection revocation mechanism allows any user to immediately terminate a connection, blocking all further messaging. The invitation-chain traceability allows administrators to identify repeat offenders and their referral networks. Combined, these create both immediate protection (revocation) and long-term deterrence (accountability).

9.5 Data Breach Impact

In a conventional messaging platform, a database breach exposes phone numbers, which are permanent identifiers. Here, a breach of the Identity Store exposes only usernames and public

profile data no phone numbers, no payment credentials, no government IDs. Usernames can be changed post-breach, restoring privacy. Phone numbers cannot.

Table 4: Breach Impact Comparison

Data Exposed	Phone-Number System	Suggested Platform
Phone numbers	Yes (permanent)	No
Linked financial accounts	Indirectly (via number)	No
Linked government ID	Indirectly (via number)	No
User identity	Permanent (number)	Changeable (username)
Recovery difficulty	Cannot change number easily	Change username instantly

10 Implementation Status

The platform is currently implemented(kindly check projects menu in website) with the following components operational:

Completed: Username-based messaging, connection request protocol, personal/professional contact separation, institutional verification display, invitation-chain recording, end-to-end encrypted messaging, Android and iOS applications deployed on Google Play Store and Apple App Store.

In Development: Professional business accounts with service listings and appointment scheduling; curated content feed with vertical scrolling across controlled categories (education, news, entertainment); UPI payment integration via TPAP partnership; Indian language support for regional accessibility; community-driven feature polling for democratic development.

Planned: Mini-app ecosystem enabling third-party developers to build privacy-respecting applications within the platform; advanced analytics for business accounts; open developer SDK.

11. Use Case Architecture

The four-layer architecture maps to specific Indian use cases where phone number exposure creates documented harm.

Table 5: Use Case Mapping to Architecture Layers

Sector	Problem	Layer(s) Applied
Banking	RMs expose personal numbers to thousands of clients	L1 + L3
Education	Students and teachers share numbers in class groups	L1 + L2 + L3
Healthcare	Doctors receive after-hours contact from any patient	L1 + L2 + L3
Real Estate	Brokers face persistent spam post-transaction	L1 + L2
Freelancers	One-time service contacts become permanent	L1 + L2
SMB Services	Business-personal boundary collapse	L1 + L3
Communities	Group membership leaks all members' numbers	L1 + L4

12.Conclusion

The evidence presented in this paper 86,420 cybercrime cases growing at 31.2% annually, 41.68 billion spam calls in a single year, 3.5 billion WhatsApp accounts enumerable through phone numbers, and a conviction rate below 3% establishes that phone-number-dependent communication architecture has created a systemic privacy crisis in India.

The platform's architecture addresses this crisis at the design level through four composable layers: identity masking, consent-gated connections, institutional verification, and invitation-chain traceability. Each layer is grounded in established research and aligns with India's regulatory framework under the DPDPA 2023 and IT Rules 2021.

The architecture does not require users to sacrifice functionality for privacy. The consent gateway adds a single step (connection request) in exchange for eliminating an entire class of attacks. The institutional verification layer adds trust signals that do not exist in current systems. The identity masking layer removes a permanent vulnerability (phone number exposure) and replaces it with a changeable, non-sensitive identifier (username).

The technical building blocks username-based identity (Signal, 2024), verifiable credentials (W3C, 2025), privacy-by-design (Cavoukian, 2011), and consent-based communication (EDPB, 2020) are all established and validated. What remains is implementation at India's billion-user scale.

References

[1] Gegenhuber, G. K., Frenzel, P. E., Gunther, M., Ullrich, J., & Judmayer, A., Hey there! You are using WhatsApp: Enumerating three billion accounts for security and privacy. Network and Distributed System Security Symposium (NDSS), 2026. https://publications.sba-research.org/publications/

[2] National Crime Records Bureau, Crime in India 2023, Ministry of Home Affairs, Government of India, 2025. https://www.ncrb.gov.in/

[3] Ministry of Electronics and Information Technology, Year-wise number of cyber security incidents (CERT-In), Open Government Data Platform India, 2024. https://www.data.gov.in/

[4] Truecaller, Truecaller India Insights Report 2025, February 2026. https://www.truecaller.com/

[5] Comparitech, 35+ phone spam statistics and facts for 2017-2024, 2024. https://www.comparitech.com/blog/information-security/phone-spam-statistics/

[6] McDonald, A., Sugatan, C., Guberek, T., & Schaub, F., The annoying, the disturbing, and the weird: Challenges with phone numbers as identifiers. Proceedings of CHI 2021, ACM. https://doi.org/10.1145/3411764.3445085

[7] Dev, J., Das, S., & Camp, L. J., Privacy practices, preferences, and compunctions: WhatsApp users in India. HAISA 2018. https://www.researchgate.net/publication/327867176

[8] Dev, J., Rashidi, Y., Das, S., & Camp, L. J., Lessons learnt from comparing WhatsApp privacy concerns across Saudi and Indian populations. SOUPS 2020, USENIX. https://www.researchgate.net/publication/348936324

[9] Telecom Regulatory Authority of India, Telecom subscription data as on 31st December 2024. https://www.trai.gov.in/release-publication/reports/telecom-subscriptions-reports

[10] DataReportal, Digital 2025: India, 2025. https://datareportal.com/reports/digital-2025-india

[11] Ministry of Micro, Small and Medium Enterprises, MSME Annual Report 2024-25, Government of India. https://msme.gov.in/msme-annual-report-english-2024-25

[12] Ministry of Education, All India Survey on Higher Education (AISHE) 2021-22, Government of India. https://aishe.gov.in/aishe-final-report/

[13] Guild, Is WhatsApp suitable for schools?, 2020. https://guild.co/blog/whatsapp-risk-use-in-schools/

[14] Rosenberg, H. & Asterhan, C. S. C., WhatsApp, Teacher? Student perspectives on teacher-student interactions. JITE: Research, 17, 205-226, 2018. https://www.informingscience.org/Publications/4081

[15] Tamil Selvan, P. & Kalaiyarasan, G., Systematic review on utilisation of WhatsApp in education. E-Learning and Digital Media, 2024. https://doi.org/10.1177/20427530241239424

[16] Cavoukian, A., Privacy by Design: The 7 Foundational Principles. Information and Privacy Commissioner of Ontario, 2011. https://www.sfu.ca/~palys/Cavoukian-2011-PrivacyByDesign-7FoundationalPrinciples.pdf

[17] The Digital Personal Data Protection Act, 2023, No. 22, Acts of Parliament, India. https://www.meity.gov.in/static/uploads/2024/06/2bf1f0e9f04e6fb4f8fef35e82c42aa5.pdf

[18] Ministry of Electronics and Information Technology, IT (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, as amended. https://www.meity.gov.in/static/uploads/2024/02/Information-Technology-Intermediary-Guidelines-and-Digital-Media-Ethics-Code-Rules-2021-updated-06.04.2023-.pdf

[19] Signal, Keep your phone number private with Signal usernames, Signal Blog, February 2024. https://signal.org/blog/phone-number-privacy-usernames/

[20] Freedom of the Press Foundation, Why journalists should enable Signal usernames, 2024. https://freedom.press/digisec/blog/enable-signal-usernames/

[21] World Wide Web Consortium, Verifiable Credentials overview, W3C Group Note, 2025. https://www.w3.org/TR/vc-overview/

[22] Press Information Bureau, Government notifies DPDP Rules to empower citizens and protect privacy, November 14, 2025. https://www.pib.gov.in/PressReleasePage.aspx?PRID=2190014

[23] Reserve Bank of India, Master Direction on Digital Payment Security Controls, RBI/2020-21/74, 2021. https://www.rbi.org.in/Scripts/NotificationUser.aspx?Id=12032&Mode=0

[24] European Data Protection Board, Guidelines 05/2020 on consent under Regulation 2016/679, Version 1.1, 2020. https://www.edpb.europa.eu/sites/default/files/files/file1/edpb_guidelines_202005_consent_en.pdf

[25] Data Security Council of India & SEQRITE, India Cyber Threat Report 2023, 2024. https://www.dsci.in/resource/content/india-cyber-threat-report-2023

[26] Mueller, R., Schrittwieser, S., et al., Security and privacy of smartphone messaging applications. IJPCC, 11(2), 132-145, 2015. https://doi.org/10.1108/IJPCC-07-2014-0044

[27] Press Information Bureau, Steps to curb cyber crime, Government of India, December 2024. https://www.pib.gov.in/PressReleasePage.aspx?PRID=2112244

[28] Indian Computer Emergency Response Team, Directions under section 70B of IT Act 2000, No. 20(3)/2022-CERT-In, April 2022. https://www.cert-in.org.in/PDF/CERT-In_Directions_70B_28.04.2022.pdf