

CONVERSATIONAL AI AS A COMPLIANCE LAYER FOR MANDATORY DIGITAL RESIDENT-COMMUNICATION REGULATION: EVIDENCE FROM A MULTI-REGION DEPLOYMENT IN URBAN HOUSING MANAGEMENT

ALEXANDER ROMASHCHENKO

VORONEZH, RUSSIA

EMAIL: ALEX.ROMASHCHENKO@GMAIL.COM

REESTRDOMA.RU

ORCID: 0009-0003-7570-9380

KEYWORDS	ABSTRACT
smart cities; e-government; regulatory compliance; conversational AI; housing management; GovTech; RegTech	<i>Cities increasingly regulate not only the content but the channel of communication between service providers and residents, backing such requirements with fines; manual compliance is costly for the small operators typically covered by such rules. We analyze telemetry from an AI assistant operating in residential building chats in multiple regions of Russia (1 January – 8 September 2026) around the effective date of Government Resolution No. 40, which made the same channel mandatory for housing organizations (fines up to RUB 300,000, ~€3,000 at early-2026 exchange rates). The assistant's response-generation feature was active for one month, of which the final eight days followed the mandate's effective date. Verification approval - 85.5% among residents who submitted documents - is computed on documents actually submitted, not on the much larger pre-provisioned base (Section 5.3). AI response volume, by contrast, is heavily concentrated: two-thirds of the 102 responses occurred in three internal test chats, leaving only 36 responses across 13 production chats - a pilot-scale production footprint, not yet a national or multi-region-scale one. We distinguish the assistant's dialogue layer from a separate transactional layer where formal service requests are actually created, and propose a minimal five-metric reporting set for evaluating AI-mediated compliance in comparable deployments.</i>

1. Introduction

A growing number of governments now legislate not only what information service providers must give residents, but the channel through which it must be given - increasingly a single, state-endorsed messaging platform, backed by administrative fines. Housing management is a useful setting in which to study this pattern: cities contain thousands of small, independent management organizations without dedicated digital-communications staff, creating a natural opening for AI-assisted compliance tooling that can read a mandated channel continuously, answer routine queries, guide identity verification, and point residents toward the appropriate formal channel.

This paper examines such a deployment: an AI assistant integrated with a housing-management platform, operating in multiple regions of Russia immediately before and after Government Resolution No. 40 (26 January 2026) made the same channel mandatory. Its contribution is fourfold: it documents this regulatory pattern; separates deployment- and verification-level metrics from far more concentrated AI response-volume figures, avoiding their conflation; distinguishes the assistant's dialogue-layer role from a separate transactional layer where legally significant requests are actually created; and proposes a minimal, reusable reporting set for cities evaluating similar AI-compliance layers.

2. Regulatory Context and Related Work

Government of the Russian Federation Resolution No. 40 (26 January 2026; published 2 February 2026; effective 1 September 2026) requires regulated housing organizations to interact with residents through a designated national messaging service, MAX. Non-compliance is treated as a violation of apartment-building management rules under Article 7.23.3 of the Code of Administrative Offences, with fines of up to RUB 50,000 for officials and RUB 300,000 for legal entities, and disqualification on repeat violations. This is a binding, platform-specific, penalty-backed channel mandate applied uniformly to a large and heterogeneous population of small regulated entities - conditions under which manual compliance is least tractable and AI-assisted tooling is most directly testable.

Conversational agents as a citizen-government communication mechanism were formalized by Androutsopoulou et al. (2019), and single-agency deployments have generally reported citizen-experience gains (Alsubayhay & Abdalla, 2025), though a review of chatbots in Spanish public administration found that rigorous evaluation of systems already in production, rather than prototypes, remains rare (Cortés-Cediel et al., 2023) - a gap this paper's production telemetry addresses. A separate literature on organizational adoption distinguishes the decision to adopt a chatbot from the determinants of its successful implementation, finding that resourcing and organizational conditions, not the technology itself, are typically the binding constraint on scale-up in single, discretionary agency deployments (Chen et al., 2024; Omonov & Ahn, 2025). A third relevant stream, regulatory technology (RegTech), has since Arner et al. (2017) framed technology-mediated compliance as a response to the rising cost of regulation for smaller supervised entities; this paper extends that logic, developed mainly around financial supervision, to a housing-sector mandate.

These three literatures can be connected through a simple compliance frame: the regulator acts as principal, the housing management organization as its agent, and AI as a monitoring and communication device the agent adopts to demonstrate performance of a delegated obligation. This framing separates the agent's delegable duties - availability and communication, which an AI layer can plausibly discharge - from duties that remain non-delegable, such as creating a legally binding record of a resident's request; that distinction motivates the dialogue-layer/transactional-layer separation introduced next.

3. System Overview

The regulated messaging channel examined here supports two functionally distinct layers, which this paper keeps separate throughout.

The dialogue layer is the chat interface, where an AI assistant operates on behalf of participating housing management organizations: it publishes and maintains the legally required contact and operating information, answers resident messages using building-specific data where it can do so with adequate confidence, and guides residents through identity verification - via a utility account number or a submitted document - before treating a message as coming from a verified resident. Chats typically also host other, unrelated bot accounts (≈ 2.5 per chat on average across the extract analyzed; Section 4), so the assistant is one of several automated participants, not the sole channel of engagement.

The transactional layer is a separate companion mini-application operated by the same housing-management platform and embedded within MAX, where residents submit formal service requests and access a paid-services marketplace. Requests are created exclusively in the mini-application: the dialogue layer can point a resident toward it - for example, by sharing a link - but does not itself generate a request or track its status. The deployment as a whole, both layers combined, is in continuous production use across multiple regions of Russia, placing it beyond a laboratory prototype (minimum TRL5).

4. Data and Methods

The analysis uses an operational data extract covering 1 January – 8 September 2026. The assistant's first chat connection occurred on 3 February 2026, but its response-generation feature was active only from 10 August to 8 September 2026 - a one-month window, of which the final eight days followed the regulation's effective date. Deployment- and verification-level figures in Section 5 are computed over the full period across all 252 chats; AI response-volume figures are computed over the one-month window only, and this distinction is preserved throughout the Results.

The extract includes chat-, message-, verification-, and service-request-level logs. Four categorical fields (chat type, chat status, message type, and service-request origin/status) are recorded in the source schema as unweighted indicator sets rather than true counts - each category is marked as present/absent rather than counted - and are excluded from quantitative reporting for that reason (see Section 7 for the related usage-telemetry gap). The bot-count figure cited in Section 3 (≈ 2.5 per chat) is derived from 633 bot membership-table records across the 252 chats in this extract ($633/252 \approx 2.5$); these are membership entries, which may include duplicate or historical records, not a count of distinct bots, so the true number of distinct non-focal bots per chat may be lower. Response counts in Section 5.2 are drawn from the message log itself (which records that the assistant sent a reply), a pipeline separate from the usage-metering table discussed below; this separation is why response counts remain available even though usage-metering data are not.

Ethics and data protection

All figures reported are aggregate counts computed by the platform operator from de-identified operational metadata; the author had no access to individual residents' documents or personal identifiers, and the analysis involved no direct engagement with human subjects. Verification-document handling within the platform is governed by the operator's own data-processing arrangements, outside the scope of this analysis.

5. Results

5.1 Deployment and retention

The chat table in this extract contains 252 rows in total - every chat record the platform has stored, including entries not linked to any building and internal test/staging chats - rather than a count of chats an administrator deliberately connected and linked to a real house. Of these, 89

rows are chats an administrator connected and linked to a specific house record; the relationship is chats-to-houses, not one-to-one - a single chat can serve more than one house, which is why linked houses (93) outnumber linked chats (89). This 89-chat, 93-house group is the more meaningful denominator for the deployment's actual reach, and the remaining 163 rows include unlinked and test entries.

This is a snapshot, not a retention rate over time: the extract records whether the bot was present in each row as of the extraction date, not continuous connection history, so it cannot show how long any given chat retained the bot. Bot-presence figures in this paper are computed across the full 252-row chat table, not restricted to the 89 administrator-connected, house-linked chats, because the presence/removal log itself is kept only at the level of the full table; a presence or retention figure specific to the narrower, 89-chat group is not available in this extract (Section 7). With that caveat, 209 of the 252 rows (82.9%) had the bot present as of the extraction date; the remaining 43 rows are chats from which the bot had been removed, or that had been closed, by that date. Aggregate participation figures over the same full table are 9,310 members (~38/chat) and 8,384 resident messages from 1,010 unique senders.

Metric	Value
Chat-table rows (total)	252
Rows with bot present as of extraction date (snapshot)	209 (82.9%)
Administrator-connected, house-linked chats	89 (93 buildings)

Table 1. Chat-table composition and bot presence, 1 Jan – 8 Sep 2026 (full period).

5.2. Response concentration

Over the one-month window in which the response feature was active, the assistant generated 102 responses, concentrated in 16 of 252 chats. Exactly 66 of the 102 responses (65%) occurred in three chats we classify as internal test/staging, based on chat title and known internal usage - a classification not independently verifiable from a structured field in the source schema (Section 7). The remaining 36 occurred across 13 chats we classify as production, spanning 15 buildings in 7 regions of Russia - a pilot-scale production footprint, not a national or multi-region-scale one in terms of AI activity specifically. Raw response counts therefore substantially overstate the assistant's live-production footprint, and any assessment of AI contribution to compliance should rely on this production subset rather than the aggregate.

5.3. Identity verification

Verification was enabled for 19 of 71 configured organizations across 21 waves, which generated 4,088 resident records provisioned in advance of any resident action - administrative placeholders, not a funnel of procedures residents had begun. Of these, 83 residents actively engaged by submitting a document; 71 (85.5%) were approved, 8 remained under review, and 4 were not verified. The compliance-relevant figure is the 85.5% approval rate among residents who submitted a document, not the 2.0% (83/4,088) that submitted relative to the pre-provisioned base - the latter measures outreach coverage, not verification reliability.

Verification stage	Count	Rate
Records pre-provisioned (no resident action yet)	4,088	-
Documents actively submitted by residents	83	-
Verified / approved (of submitted)	71	85.5%
Under review / not verified (of submitted)	8 / 4	9.6% / 4.8%

Table 2. Resident identity-verification funnel (rates computed on submitted documents, not on pre-provisioned records).

5.4. Service requests, appeals, and complaints

A service request cannot be created by sending a chat message: it is created exclusively through a form inside the companion mini-application (Section 3). Twenty-two service requests originated through that channel across 15 buildings ($\approx 2.2\%$ of the 1,010 unique residents who sent a chat message during the period), and a separate appeal-token mechanism issued 25 tokens to 7 residents across 8 chats, of which 6 tokens were opened (24%) and 2 used to completion (8%). That no request originated from an unstructured chat message reflects this architectural separation, not resident behavior or the assistant's filtering ability.

Three resident complaints about AI responses were logged during the period; given this small count, we report only the raw figure rather than a percentage-based error rate. A per-message conflict/tension indicator present in the source schema was not populated for this period, so it could not corroborate this figure.

6. Discussion

The dialogue layer examined in this paper helps satisfy the elements of the regulation that concern channel availability and communication: publishing required information, responding to residents, and guiding identity verification. Creating an official, legally significant service request happens in the separate transactional layer and does not depend on the dialogue layer succeeding. The assistant is therefore better described as a front-end of compliance - a layer that sustains channel availability and communication without itself creating a legally binding record - than as compliance's enforcement engine: its contribution should be evaluated as that of a communication and verification layer, not as a substitute for the structured request channel a regulator would actually audit.

The concentration of AI response volume in test rather than production chats (Section 5.2) suggests model capability is currently ahead of production adoption. This is consistent with, though not directly tested against, findings from single-agency, discretionary chatbot deployments in which organizational and resourcing conditions, not the technology, constrain scale-up (Chen et al., 2024; Omonov & Ahn, 2025). Whether that adoption-implementation gap is systematically wider under a non-discretionary, population-wide mandate is a hypothesis for future comparative work rather than a claim this single case can establish.

A minimal reporting set. The metrics below are proposed as a starting point for other cities and researchers evaluating similar AI-compliance layers, distinguishing what an AI layer can measurably contribute from what remains outside its scope. Data sources are, respectively, the chat connection log, the verification-workflow table, the response log combined with chat classification, and the complaint log.

Metric	Definition	Suggested use
Channel retention	Share of connected chats still active at measurement date	Track over time; report on a full-period basis
Verification approval rate	Share of submitted verification documents approved	Compute on submitted, not pre-provisioned, records
Production response share	Share of AI responses in genuine resident-facing chats vs. test/staging	Report separately from raw response counts
Complaint count	Resident complaints about AI responses	Report as raw n below ~ 30 responses; avoid % at low n

Table 3. A minimal reporting set for AI-mediated compliance with digital resident-communication mandates. A fifth metric, chat-to-application conversion, is included here for completeness of the proposed set, though not measured in this study; cities adopting comparable systems can log it from the outset (Section 7).

7. Limitations

- Single-country, single-deployment case study; AI response-volume figures cover a one-month window, of which only eight days followed the mandate's effective date.
- Chats are classified as production vs. test based on chat title and internal knowledge rather than a structured schema field, and this classification is not independently verifiable from the data alone.
- Bot-presence figures (Section 5.1) are a snapshot at the extraction date, not a time-series retention rate, and are computed across all 252 rows in the chat table - including chats never linked to a building and internal test entries - because the presence/removal log is kept only at that full-table level, not separately for the 89 administrator-connected, house-linked chats; building-level figures elsewhere in this paper are limited to that same 89-chat subset.
- A companion usage table shows zero logged language-model calls despite 102 recorded responses, indicating a gap between response-logging and usage-telemetry pipelines; no cost or token data are available.
- Chat-to-application conversion - whether the assistant's prompts lead residents to the transactional layer - is not logged in the current instrumentation and is a natural target for future data collection.

8. Conclusion

Regulations mandating a specific digital channel for resident communication, backed by fines, create an enforcement burden heaviest on the small management organizations least equipped to meet it manually. Production data collected around the effective date of one such mandate in Russia show that an AI dialogue layer can plausibly close part of this gap by acting as compliance's front-end, as the bot-presence figures (computed across the full chat table) and verification figures (computed on submitted documents) in Section 5 suggest, at a complaint level too small to characterize statistically. Creating the legally significant record a regulator would ultimately audit remains the task of a separate transactional layer, which the dialogue layer can direct residents toward but does not replace. The concentration of AI response activity in test rather than production chats (Section 5.2) shows that, at this stage, model capability runs ahead of production adoption. The binding constraint for cities is therefore less whether such systems can respond adequately than how quickly coverage can be extended and how reliably the dialogue layer converts engagement into transactional-layer action. The minimal reporting set proposed in Table 3 offers one concrete path toward making such mandates auditable at scale, without requiring every small operator to staff a dedicated response team.

9. Acknowledgements and AI Disclosure

AI assistance (Claude Sonnet 5, Anthropic) was used for language editing and drafting support, and for organizing the descriptive statistics reported in Section 5 from the underlying operational data extract. All interpretations, conclusions, and verification of reported figures against the source data are the sole responsibility of the human author, who takes full responsibility for the content, data, and conclusions of this manuscript.

References

- Alsubayhay, A., & Abdalla, M. (2025). Enhancing citizen engagement in e-government services through AI-driven chatbots. *Sebha University Conference Proceedings*, 3(3), 188–194. <https://doi.org/10.51984/sucp.v3i3.3840>
- Androutsopoulou, A., Karacapilidis, N., Loukis, E., & Charalabidis, Y. (2019). Transforming the communication between citizens and government through AI-guided chatbots. *Government Information Quarterly*, 36(2), 358–367. <https://doi.org/10.1016/j.giq.2018.10.001>
- Anthropic. (2026). Claude Sonnet 5 [Large language model]. Anthropic. <https://claude.ai>
- Arner, D. W., Barberis, J., & Buckley, R. P. (2017). FinTech, RegTech, and the reconceptualization of financial regulation. *Northwestern Journal of International Law & Business*, 37(3), 371–413.
- Chen, T., Gascó-Hernández, M., & Esteve, M. (2024). The adoption and implementation of artificial intelligence chatbots in public organizations: Evidence from U.S. state governments. *American Review of Public Administration*, 54(3), 255–270. <https://doi.org/10.1177/02750740231200522>
- Code of Administrative Offences of the Russian Federation, Article 7.23.3 (violation of rules governing the management of apartment buildings). Official Internet Portal of Legal Information, pravo.gov.ru.
- Cortés-Cediel, M. E., Segura-Tinoco, A., Cantador, I., & Rodríguez Bolívar, M. P. (2023). Trends and challenges of e-government chatbots: Advances in exploring open government data and citizen participation content. *Government Information Quarterly*, 40(4), 101877. <https://doi.org/10.1016/j.giq.2023.101877>
- Government of the Russian Federation. (2026). Resolution No. 40 of 26 January 2026, on the mandatory use of the MAX service in housing and communal services. Published 2 February 2026 on the Official Internet Portal of Legal Information, publication.pravo.gov.ru; effective 1 September 2026.
- Omonov, M. S., & Ahn, Y. (2025). Towards smart public administration: A TOE-based empirical study of AI chatbot adoption in a transitioning government context. *Administrative Sciences*, 15(8), 324. <https://doi.org/10.3390/admsci15080324>