Consumer Lending

The Consumer-Lending Phone Line

What roughly 1.8 million analyzed voice calls reveal about why consumers call, how those calls end, and where conversational AI can create the most value.

Anonymized industry insights | Three consumer-lending deployments | Trailing 12 months | July 2026

In Brief

1. Consumer-lending phone demand is dominated by **routine servicing**. Payments and account questions are roughly **half of all calls**, the most automatable work lenders run, yet fewer than half of calls reach a clear resolution.

2. The experience leaks value. About **two in five calls resolve** and **one in four is abandoned**; transfer-to-human rates vary roughly **40-fold** across lenders doing the same work, signaling wide, unrealized automation headroom.

3. **Originations is the next frontier.** Sales calls are longer, more positive, and rich in rate, quote, and document intents, the qualify-and-quote work AI can carry to free human loan officers for the deals that need them.

What these numbers measure

A diagnosis of the channel as it runs today, not Replicant's automation results

The figures in Sections 1–4 describe the **current phone channel** at these lenders (existing agents and IVR), measured by Replicant's Conversation Intelligence. "Resolved / partially resolved / abandoned / transferred" is the **caller's final disposition on that call**, whole-channel, not AI-only containment. Abandonment means the caller hung up before a resolution. This section sizes the opportunity; **Replicant's own deployment results are in Section 5**, where "containment" (calls handled without a live agent) is reported separately from channel resolution.

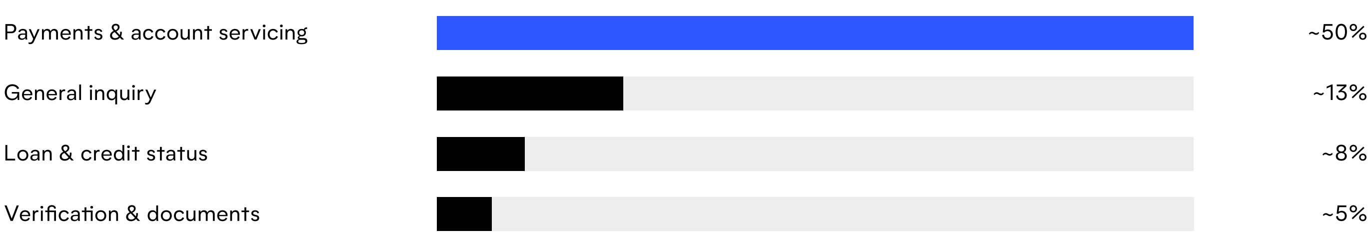
Section 1

Half of consumer-lending call volume is routine servicing

Conversation Intelligence tags a reason on every call. Grouping payments and account questions into a single servicing category, servicing is the dominant driver at every lender in the sample; the remainder is loan and credit status, general inquiry, and identity or document verification. The majority of live-agent minutes are spent on structured, data-backed requests that an AI agent can resolve end to end when connected to the servicing system of record.

Exhibit 1

Servicing accounts for roughly half of all lending calls



Share of tagged calls, equal-weighted across three deployments; deployment range for servicing 38–59%. Source: Replicant Conversation Intelligence, trailing 12 months (~1.8M calls).

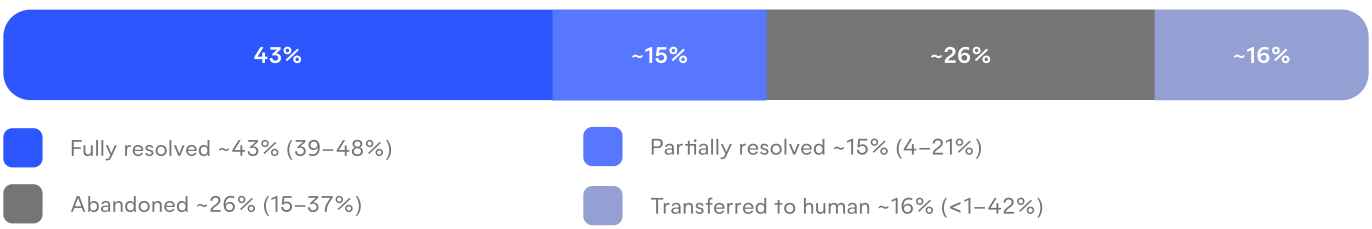
Section 2

Fewer than half of calls resolve, and one in four is abandoned

Despite the routine demand, only about 43 percent of calls reach a clear resolution. Roughly a quarter of callers abandon before their issue is resolved, the clearest quantification of friction in the channel. The dispersion is itself a finding: transfer-to-human rates span under 1 percent to 42 percent across lenders performing the same category of work, indicating that automation maturity is highly uneven and the ceiling sits well above where most operate today.

Exhibit 2

How consumer-lending calls end



Share of calls by final disposition, equal-weighted across three deployments; ranges show deployment minimum to maximum. Source: Replicant Conversation Intelligence, trailing 12 months.

Section 3

Origination calls are a distinct, higher-value opportunity

Servicing dominates volume, but origination (sales) calls are a categorically different conversation, and the one lenders most often ask where else AI can help. They are markedly longer and more positive, and their intents cluster around rate and loan inquiries, quotes, and document collection, motivated buyers at the top of the funnel.

~17 min

Typical origination call length, versus ~3.5 min for servicing

7.1/10

Origination caller sentiment, versus ~5.9 for servicing

~40%

Of origination calls are rate or loan inquiries

Exhibit 3

Two different conversations: servicing versus originations

Dimension	Servicing calls	Origination / Sales calls
Share of volume	~half of all calls	a distinct, smaller stream
Typical length	~3.5 minutes	~17-18 minutes
Caller sentiment	~5.9 / 10	~7.1 / 10
Top intents	Payments, balances, account, status	Rate & loan inquiry, quotes, documents, income
AI opportunity	Resolve routine requests, 24/7	Quote rates, pre-qualify, capture documents, book callbacks

Servicing figures equal-weighted across three lending deployments. Origination figures from a single mortgage-origination call set (~3,000 recent calls), reported separately and not blended into the market rates given the smaller, single-program sample. Source: Replicant Conversation Intelligence.

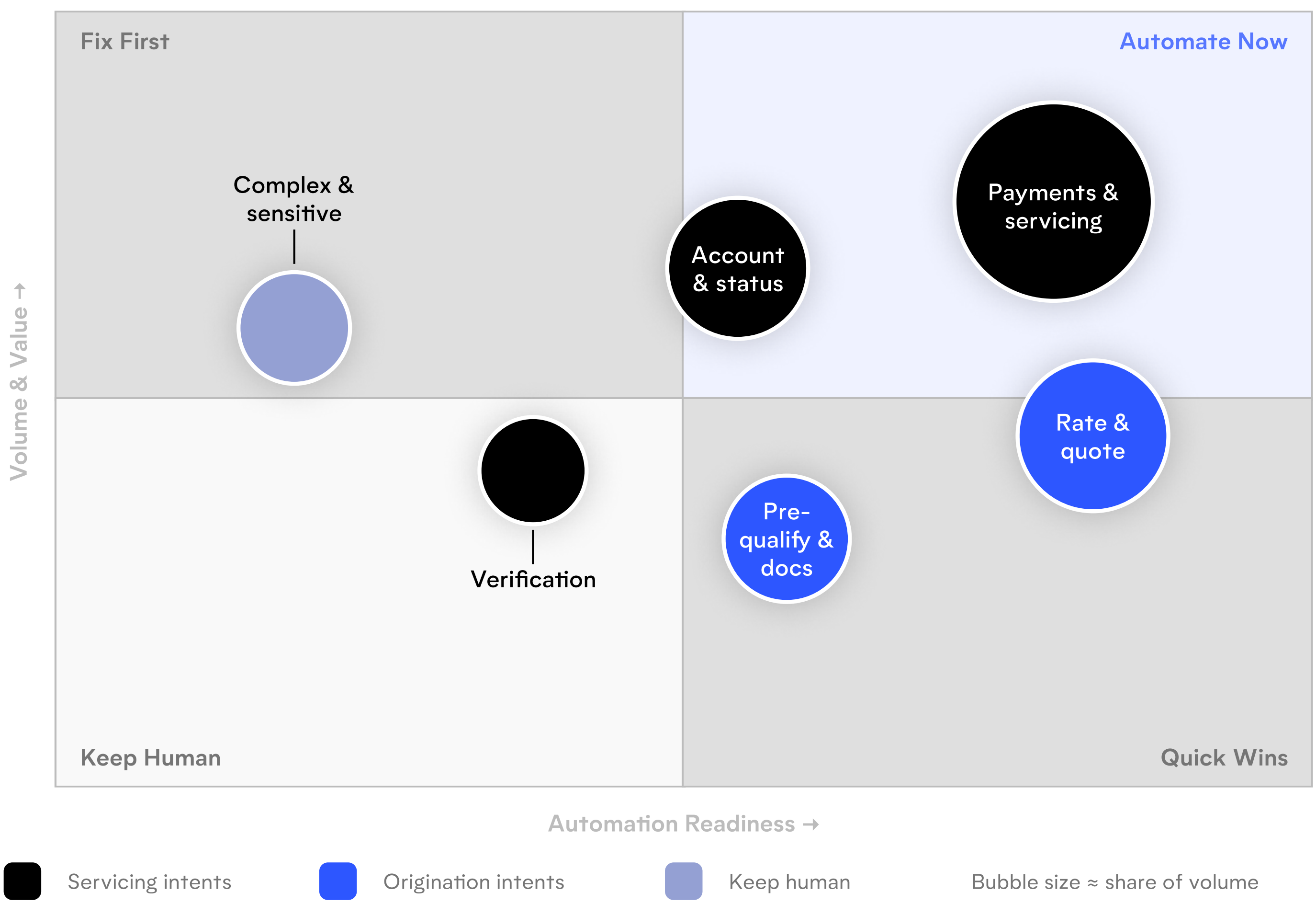
Section 4

Where to focus first: a priority map

Positioning the major call types by their readiness for automation against their volume and value points to a clear sequence. Servicing intents sit top-right, high volume and ready to automate now. Origination intents bring high value at solid readiness. Complex, sensitive conversations remain with people, augmented rather than replaced.

Exhibit 4

Automation priority map: readiness versus volume and value



Illustrative positioning by Replicant based on intent mix, call outcomes, and data availability across the sample; axes are directional, not to scale. Source: Replicant Conversation Intelligence.

Section 5

From diagnosis to outcome: Replicant in financial services

The sections above size the leak. The question a buyer asks next is what happens to it. Across Replicant's production financial-services deployments, automation converts that unresolved volume into measured containment, lower handle time, hard cost savings, and high customer satisfaction. "Containment" here means calls fully handled without a live agent, a deployment outcome distinct from the whole-channel resolution measured in the diagnostic.

65-75%

AI containment across deployments (one moved 65% to 75%)

4.8/5

CSAT, sustained (range 4.5 to 4.86)

-62%

Average handle time, with $1M+ annual savings

Exhibit 5

Measured results from Replicant financial-services deployments

Deployment	Measured result
Debt-resolution provider	AI containment **65% → 75%;** AHT **-62%;** CSAT **4.86/5**
Debt-relief provider	Containment **65%;** CSAT **4.75/5;** call time **-1.5 min;** runs on Conversation Intelligence
Consumer debt-relief provider	Resolution **74%;** CSAT **4.8/5; 99%** focus-group approval
Loan servicer	**$1.5M** annual savings; CSAT **4.5/5;** 35k calls/month automated

Results from Replicant financial-services deployments (debt resolution, debt relief, loan servicing), shown as anonymized descriptors and **distinct from the diagnostic cohort** in Sections 1–4. Containment (calls handled without a live agent) is a deployment metric, distinct from the channel-wide resolution measured in the diagnostic.

Section 6

Built for regulated financial services

Voice AI in lending touches consent, disclosures, PII, and money movement. Replicant is built to operate inside that framework, with enterprise security and data-privacy controls certified by independent audit, and conversation-level controls configured to each program's regulatory requirements.

Security & Data privacy	Controls for regulated conversations
Certified. SOC 2 Type 2 (independent annual audit), HIPAA, PCI, and GDPR compliant.	**Consent & recording.** Configurable consent capture and call-recording disclosures.
Encrypted. AES-256 at rest (GCP KMS); TLS 1.2 in transit.	**Consistent disclosures.** Every caller hears the same approved script, verbatim, supporting fair-lending consistency.
Hosted in the US. Google Cloud, continental US; redundant, high-availability infrastructure.	**PII redaction.** Sensitive data redacted from transcripts and recordings.
Access-controlled. Role-based access, IP restrictions, network segmentation, need-to-know.	**Human in the loop.** Guaranteed escalation for exceptions and adverse-action handling.
Tested. Independent penetration testing, bug bounty, IPS/IDS monitoring, incident-response plan.	**Your framework.** Retention, disclosures, and licensing configured per program (TCPA, Reg Z, adverse action, state licensing).

Platform security and data privacy per Replicant's Security & Data documentation. Conversation-level controls are configured to each customer's compliance program.

Section 7

Two agendas, one playbook

The opportunities the analysis surfaces are consistent across deployments and fall into two agendas. In servicing, the prize is deflecting and resolving routine, always-available demand. In originations, it is qualifying and quoting at the top of the funnel so that human loan officers spend their time on the deals that convert.

Servicing agenda	Originations agenda
Self-service payments. Payments, plans, and auto-pay resolved in-conversation.	**Rate & quote inquiries.** Answered and quoted instantly, at any hour.
Proactive notifications. Due-date and status alerts that deflect calls before they occur.	**Pre-qualification.** Income and basics captured to qualify leads.
Streamlined verification. Fewer repeated identity questions; pre-authentication in IVR.	**Document & income intake.** Collected and verified over voice and SMS.
Intent-led routing. Callers state the request directly, bypassing menu trees.	**Callback scheduling.** Hot leads routed to a human at the right moment.

The next step

This diagnosis is directional, three deployments, not your book. The definitive read is your own. Replicant can run this same Conversation Intelligence analysis on your calls, quantify where your channel leaks, and model the containment and resolution lift from automating it, before any commitment.

Methodology & scope

Market figures are drawn from Replicant Conversation Intelligence over the trailing 12 months and are equal-weighted across three production consumer-lending deployments spanning consumer lending, auto finance, and mortgage (~1.8 million analyzed calls); each deployment contributes equally and the deployment minimum-to-maximum range is shown, so no single program drives a headline. A metric is reported only where all three deployments contribute. Payments and account reasons are grouped into a single servicing category. Numbers are rounded and banded. The priority map is an illustrative synthesis of intent mix, outcomes, and data availability; its axes are directional, not to scale. The originations spotlight is drawn from a single mortgage-origination call set (~3,000 recent calls) and is reported separately, not blended into the market rates. Automated quality scores use per-program rubrics and are not compared across deployments. Debt settlement is excluded from this lending-focused view. No deployment, client, or individual caller is identified, and no underlying data left the analytics environment. The Section 5 results are Replicant financial-services deployments (debt resolution, debt relief, and loan servicing) shown as anonymized descriptors, separate from the diagnostic cohort; containment is a deployment metric, distinct from the whole-channel resolution measured in the diagnostic. No customer is named anywhere in this report.