



U.S. DEPARTMENT
of STATE

A practical guide for U.S.
government agencies
developing and scaling
enterprise AI solutions

The Department of State **Generative AI Playbook**





TABLE *of* CONTENTS

<u>Letter from Leadership</u>	3
<u>Playbook Vision & Uses</u>	4
<u>Introduction</u>	5
<u>High-Level Execution Checklist</u>	6
<u>Prologue: Lay the Groundwork</u>	7
<u>Chapter 1 - Prove That It’s Possible: Alpha Development</u>	10
<u>Chapter 2 - Prove That It’s Useful: Beta Development</u>	15
<u>Chapter 3 - Launch: Going Live</u>	20
<u>Chapter 4 - Prove That It Can Scale: Driving Adoption</u>	23
<u>Chapter 5 - Pivot to Depth: Advanced AI Capabilities</u>	28
<u>Epilogue: Measure Impact</u>	31
<u>Appendix: Tools You Can Use</u>	36



GENERATIVE AI PLAYBOOK

U.S. DEPARTMENT OF STATE

A letter from the CIO and CDAO

The rapid rise of generative artificial intelligence (GenAI) presents a pivotal opportunity for U.S. federal government agencies to transform operations, enhance decision-making, and improve services for the American people. The U.S. Department of State adopted GenAI early to empower our workforce, streamline workflows, and improve productivity for mission impact.

With support from the General Services Administration's (GSA) Technology Modernization Fund (TMF), the Department of State developed and deployed StateChat, its first Sensitive But Unclassified (SBU)-enabled enterprise GenAI chatbot. TMF funding helped us build and deploy technology infrastructure that supports future secure AI applications and tools, like secure Application Programming Interfaces (APIs), to GenAI systems and controlled development environments.

This playbook chronicles the Department of State's journey to develop, test, and deploy StateChat as a case study. Its success is built on decades of work by State's IT professionals worldwide. As of June 2026, StateChat has over 62,000 users and adoption across 98% of U.S. diplomatic posts worldwide. StateChat positions the Department of State at the forefront of AI use within the federal government.

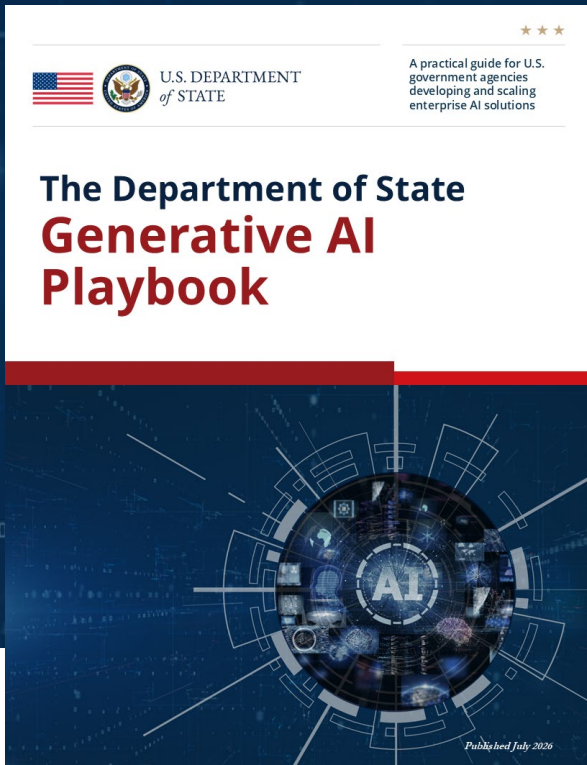
We hope this playbook can be a companion to you in your AI journey. Whether you're exploring AI's potential or scaling existing initiatives, the frameworks, tools, and insights included here reflect our lessons learned along the way and we hope they will help you anticipate challenges, seize opportunities, and deliver meaningful results for your agency and the American public we serve.

The Department of State is proud to support this collective effort, and we look forward to collaborating with other agencies to ensure the federal government remains a global leader in AI. Let us continue to innovate and build a future where AI drives American prosperity.

DR. KELLY FLETCHER
U.S. Department of State
Chief Information Officer

AMY RITUALO
U.S. Department of State
Acting Chief Data & Artificial
Intelligence Officer

[<< RETURN TO TABLE OF CONTENTS](#)



DOS Generative AI Playbook Vision & Uses

This playbook serves as a practical guide for U.S. government agencies developing and scaling enterprise AI solutions, with a focus on GenAI technology.

Based on U.S. Department of State experience, this playbook provides:

-  A repeatable framework
-  Real-world examples
-  Actionable steps to accelerate AI delivery
-  Sample deliverables (see appendix)

More than a technical manual, it balances strategic vision with operational detail to help agencies move from experimentation to implementation with confidence.

None of this work is one-size-fits-all; we hope agencies adapt the suggested processes and methodologies to fit their specific needs and uses, and should consult their own legal, security, and compliance offices before implementation.

Any references to commercial products or vendors do not constitute endorsement by the U.S. Department of State.

[<< RETURN TO TABLE OF CONTENTS](#)



GENERATIVE AI PLAYBOOK

INTRODUCTION

The release of ChatGPT in November 2022 marked a turning point in the practical use of AI for all, opening new possibilities for organizations worldwide. At the Department of State, its arrival sparked grassroots experimentation. Personnel around the world began using publicly available GenAI tools to streamline unclassified workflows like drafting public speeches and improving back-office management. These early experiments demonstrated the potential of applied AI at the Department of State, which extended from non-technical users to development teams building complex applications.

This playbook chronicles how the Department of State responded to the widespread demand for secure GenAI applications from 2022 through early 2026. Working in close partnership with offices across the agency, the Department of State leveraged some of America's leading commercial AI technology to establish a new GenAI platform for rapid prototyping and enterprise application development. StateChat was the first application built on that platform, a secure enterprise chatbot approved for sensitive data, whose creation story is documented thoroughly in these pages.

In parallel, the Department of State created the infrastructure to support safe, scaled GenAI development and adoption, including novel testing procedures, tailored training programs, and responsible use policies. Critical additions to the enterprise GenAI platform came next, including secure APIs to generative AI models and sandbox environments where technical experts and 'citizen developers' could prove new ideas. In every chapter of the story, users led the way. They signed up for user testing, provided feedback, voiced encouragement, and very directly helped shape a new generation of enterprise AI capabilities.

The Bureau of Diplomatic Technology's Center for Analytics (DT/CfA) led the design, development, and deployment of StateChat and the broader enterprise generative AI platform. The

achievements described in this playbook reflect DT/CfA's multi-year effort, made possible through close collaboration with partners across the Department of State, including DT Cyber Operations, DT Enterprise Infrastructure, DT Enterprise Chief Information Security Office, the Bureau of Diplomatic Security's Blue Team, the Records Office, and the Privacy Office.

This initiative represented unprecedented territory for every office involved. The Department of State had never before established an enterprise generative AI platform of this magnitude, and many of the requisite reviews, authorizations, and policies had never been applied to a generative AI system. This presented numerous novel challenges to include: authorizing a FISMA (Federal Information Security Management Act) High system for sensitive AI workloads, conducting a Privacy Impact Assessment for a generative AI platform, and establishing records management practices for AI-generated content. No established framework existed to guide these efforts.

The success of this work stemmed from the robust disciplines and processes already in place within each office. These existing capabilities, applied collectively, proved sufficient to address the demands posed by this emerging technology.

This playbook was written to support government agencies developing and scaling enterprise AI solutions, with a focus on GenAI technology. It provides a repeatable framework, real-world examples, and actionable steps to accelerate AI delivery. Each agency and organization should adapt the suggested processes and methodologies to fit their specific needs and uses and consult their own legal, security, and compliance offices before implementation. References to commercial products or vendors do not constitute endorsement by the U.S. Department of State.

Each chapter represents a critical milestone in this journey. It is not just about technology—it is about people, processes, and progress at the start of a new technological paradigm.

[<< RETURN TO TABLE OF CONTENTS](#)



GENERATIVE AI PLAYBOOK

HIGH-LEVEL EXECUTION CHECKLIST

This checklist outlines playbook chapters and sub-sections with each step of the GenAI delivery process.

Prologue: Lay the Groundwork	☐ Set the Vision: Develop a foundational enterprise data and AI strategy ☐ Put it into Action: Pilot AI solutions to demonstrate feasibility ☐ Ask for Help: Secure additional resource support, as needed
2-3 Years	
Chapter 1: Prove It's Possible	☐ Build the Team: Assemble multidisciplinary team ☐ Partner Up: Form strategic internal and external partnerships ☐ Experiment Early & Often: Rapidly prototype solutions to test ideas ☐ Ready, Set, Test: Confirm accessibility and basic functionality of MVP tool with a small group of trusted testers
4-6 Months	
Chapter 2: Prove It's Useful	☐ Take it for a Test Drive: Validate MVP tool functionality and collect feedback with a larger, multi-skilled group of testers ☐ Intentionally Support Users: Create user support resources and establish communication channels ☐ Listen, Prioritize, Repeat: Develop feedback collection and prioritization processes, and improve MVP iteratively
4-6 Months	
Chapter 3: Launch	☐ Prepare for Liftoff: Develop a launch plan ☐ Grant Access: Provide users with easy access in compliance with enterprise security policies ☐ Welcome Users: Provide targeted 'onboardings' to users to encourage use and adoption
4-6 Weeks	
Chapter 4: Prove It Can Scale	☐ Engage Users: Build trust through communications and training ☐ Foster Collaboration: Strengthen partnerships ☐ Refine Tools: Use feedback to improve and expand use cases ☐ Monitor & Improve: Track adoption and system reliability
6-8 Months	
Chapter 5: Pivot to Depth	☐ Complex Workflows: Integrate new tools into advanced workflows ☐ Deep Dive Trainings: Provide focused trainings to help users leverage more advanced functionality ☐ Orchestrating Agents: Develop interconnected, task-specific AI agents
4-6 Months	
Epilogue: Measure Impact	☐ Impact Survey: Conduct an impact assessment to evaluate AI tool effectiveness ☐ Lessons Learned: Analyze the collected feedback, reflect, and then document lessons learned for future experimentation
6-8 Weeks	

[<< RETURN TO TABLE OF CONTENTS](#)



GENERATIVE AI PLAYBOOK

PROLOGUE: LAY THE GROUNDWORK

AI initiatives need solid groundwork. Projects can derail before they start due to misaligned goals, skeptical stakeholders, and resource shortfalls. Before starting its AI journey, the Department of State focused on the below four steps.

Set the Vision: Establish Clear Data and AI Strategies

The Department of State established its Enterprise Data Strategy in 2021, an AI Strategy in 2023, and a combined Data & AI Strategy in 2025. Together, these strategies set a clear, Department of State-wide direction for data and AI in an environment where one-size-fits-all approaches don't work. They align staff on shared priorities, increase transparency about what has been accomplished and where the Department of State is headed, and create space for bureaus and missions to implement solutions tailored to their unique needs. They also publicly articulate the Department of State's vision for responsible and trustworthy AI.

Build the Base: Establish the Right Organizational Structures

In 2019, the Department of State launched the Center for Analytics (CfA) to embed data and analytics into decision-making, capture efficiencies, and help advance foreign policy. Led by the Chief Data and AI Officer (CDAO), CfA, housed in the Bureau of Diplomatic Technology (DT), developed an iterative, agile project implementation model called the "campaign model." The campaign model delivered data and machine learning solutions for critical Department of State use cases by integrating users, tech development, and

policymaking simultaneously. Matrixed, multi-disciplinary campaign teams learned through hands-on experimentation, identified process gaps and business needs early, and adapted governance in real time. The campaigns allowed for rapid innovation and iteration, enabling AI integration while keeping AI use safe and responsible.

Show Immediate Impact: Pilot Small-Scale Solutions

During CfA's early years, the Department of State leveraged the campaign model to run 12, six-month campaigns to streamline process, break down data silos, and deliver fast results, including with new technological, data and AI methodologies. Campaign teams worked within Department of State data governance structures, informed by the vision of executive leadership. Campaign teams facilitated requirements-gathering sessions with leaders and working-level staff to understand current processes, define solutions, and set clearly defined strategic goals.

The campaign model approach delivered results and built a repeatable process. Campaign teams documented lessons learned—both successes and failures—and shared them broadly to reduce mistakes and accelerate delivery. By starting small and moving quickly, teams tested continuously with small groups of targeted users, gathered real-time feedback, and refined solutions for maximum impact. This approach built best practices, drove a culture of innovation, informed use policies for safe experimentation, and generated momentum across the Department of State, laying the groundwork for scaling StateChat enterprise-wide.

Lesson Learned **– Small Pilots, Big Impacts**

One of the Department of State's first GenAI pilots was the AI for Reports Modernization (AIRMOD) campaign. AIRMOD used GenAI to make reporting faster and easier, improve report quality and free up staff time for mission-focused work. This effort led to the digitization, translation and extraction of key data points from over 115K documents, improved financial data accuracy by 28 percent, enabled tracking of over \$800M in lease liability, saved 76.5K hours, and avoided \$6M in costs while ensuring compliance with new accounting requirements. The Department of State later received a Best-in-Class award for this initiative.

[<< RETURN TO TABLE OF CONTENTS](#)



GENERATIVE AI PLAYBOOK

PROLOGUE: LAY THE GROUNDWORK

Identify Experts: Bringing Together Specialists with Varying Expertise

AI implementation requires new and different types of expertise. To support the development of StateChat, the Department of State methodically put together the StateChat delivery team, ensuring it combined product owners, application developers, platform and governance experts, testing specialists, customer success (including training and communications professionals), and innovation and research staff. This team worked with, and succeeded with the support of numerous other offices, including, but not limited to DT Cyber Operations, DT Enterprise Infrastructure, DT Enterprise Chief Information Security Office, the Records Office, and the Privacy Office.

See Chapter 1 for detailed guidance on building AI delivery teams and recommended skillsets.

Build the Foundation: Establishing an Enterprise AI Platform

From the beginning, the goal was to create a flexible GenAI platform that could adapt to the Department of State's evolving needs and support sensitive data. The solution had to be able to drive

data analytics, support application development, enable data engineering, and foster secure innovation in AI. In 2022 and 2023, while the technology landscape was rapidly changing, few platforms were able to bring extensive AI tooling while meeting the Department of State's strict security, compliance, and authorization standards at the pace required for mission use.

The platform was structured from the ground up as a managed enterprise service, not just as a single application system housing StateChat. The Department of State designed system security and contracts broadly, so that new applications or sandbox experiments could build and test with production data from day one. Business processes were built to provide enterprise visibility and ensure builders could access needed tools and data while following all rules of behavior. As a result, the system configuration work for StateChat paved the way for multiple major enterprise applications and over 120 builder projects, including crisis management tools responding to urgent life-threatening situations. While this playbook focuses on StateChat, the platform's broader impact and long-term value for the Department of State's technical future may be even greater over time.

PROLOGUE:

LAY THE GROUNDWORK

Summary & Insights



Establish clear foundational data and AI strategies first. Ensures alignment of priorities and principles, providing a shared direction that enables coordinated and safe adoption across a large Department of State without constraining mission-specific execution.

Build the right organizational structures. A dedicated innovation hub (like DT/CfA) and a cross-functional implementation model (like the “campaign” model) enable teams to develop technology and policy in parallel while learning through experimentation.

Start small with focused pilots. Testing AI solutions on a small-scale builds expertise, identifies compelling use cases, generates momentum, and creates a repeatable process for scaling enterprise-wide.

Build out multidisciplinary teams with emerging expertise. AI implementation requires new roles and skill sets — engineers, data scientists, security specialists, change management experts, user experience designers, and policy and governance professionals. Secure funding early, build hiring pipelines through fellowships and external partnerships, and design compelling roles to attract top talent. Equally important, invest in upskilling staff so institutional knowledge and emerging technical expertise grow together.

Leverage external funding mechanisms to scale. Programs like GSA's Technology Modernization Fund provide critical resources for strategic hiring, infrastructure investment, and moving from pilot to enterprise deployment.

Establishing a managed, enterprise AI platform as a shared service. This accelerates secure innovation, fosters collaboration, and provides a strong foundation for scaling solutions like StateChat and other mission-critical applications.



GENERATIVE AI PLAYBOOK

CHAPTER 1: PROVE THAT IT'S POSSIBLE ALPHA DEVELOPMENT

By late 2023, the Department of State had completed several successful AI campaigns, building expertise and demonstrating value through focused pilots. With that foundation in place, and TMF investment initiated in May 2024, the Department of State advanced StateChat into its “alpha” phase.

StateChat's alpha phase was short and targeted. If fundamental failures were to occur, the Department of State needed to discover them quickly. The alpha phase of development and testing sought to answer: could the Department of State offer a secure chatbot that avoided unacceptable risks, performed and improved over time, and worked globally?

This chapter describes how StateChat moved from concept to testable prototype during alpha by: (1) assembling the right team; (2) forming strategic partnerships; (3) building a minimum viable product; and (4) testing rigorously to validate core assumptions.

Build Strategic Partnerships Early

Building an enterprise AI solution is an organizational challenge as much as it is a technical one. Engaging cross-organizational stakeholders early supported successful adoption—not just after launch, but during development. This approach also allowed the team to identify blind spots and mitigate potential shortcomings from day one of development.

- **Internal Department of State Partnerships:** StateChat's success relied on strong internal partnerships across the Department of State. Before the core team was ever assembled, bureau staff from Diplomatic Technology and Diplomatic Security, networking specialists from DT Enterprise Infrastructure, accessibility experts from the Office of Accessibility and Accommodations, Accessibility Division, user experience designers, Independent Testing, Evaluation, Verification, and Validation (ITEVV) staff, and foreign policy advisors came together

during alpha development to navigate security protocols, governance, and ethical AI implementation. Cyber teams helped identify system requirements and supported the year-long process of securing an Authority to Operate (ATO). Without these partnerships, StateChat would not have made it to enterprise deployment.

The core StateChat delivery team structure can be found in Chapter 1, Section 4.

- **External Partnerships:**

The Department of State also explored partnerships with private sector vendors, research institutions, and other federal agencies to access expertise, pre-trained models, and technical capabilities. Those partners provided implementation expertise, access to lessons learned, best practices, and lynchpin technical capabilities, accelerating StateChat's progress.

Specifically, the Department of State benefited from working relationships with GSA's Technology Modernization Fund (TMF); three federal fellowship programs; two universities; four major AI labs; and several technology consultancies.

Alpha Development & Testing

Alpha development had clear objectives: prove that StateChat could meet cybersecurity standards, avoid unacceptable risks like toxicity, perform well enough to offer to beta users and improve over time. It also had to work across the Department of State's global footprint—from headquarters in Washington to embassies and consulates around the world. Alpha testing enabled the team to establish necessary safeguards, start experimentation (without getting stuck in planning mode), and stay close to users, all of which were essential to progress.



GENERATIVE AI PLAYBOOK

CHAPTER 1: PROVE THAT IT'S POSSIBLE ALPHA DEVELOPMENT

- **Experiment Early & Often: Building and Iterating the Alpha Prototype**

Rapid prototyping and iterative development were at the center of the alpha approach. Early versions of StateChat were intentionally spartan but served to test that the product could be useful to Department of State staff. Prototyping helped test ideas, assess feasibility, and uncover challenges before committing resources. Experimentation also addressed concerns like data security, compliance, and ethical considerations in a controlled manner.

The Department of State quickly tested new ideas, refined the tool based on user and tester feedback, and minimized risk through iteration. Experiments with an early, hand-selected user cohort put the spotlight on necessary capabilities that were lacking and allowed the team to develop them quickly. The goal was not perfection, it was iterating and learning fast.

Set up MVP (minimum viable product) requirements, testing, and security protocols

The Department of State drove alpha testing through a focused effort to identify foundational

safety and security issues. Alpha testing brought together 150 people with AI expertise, in cybersecurity and responsible AI roles, or willing early adopters for a targeted round of testing over a few weeks. This group included representatives from offices with security and governance expertise, offices that could validate the tool's core functionality, and known GenAI power users who could provide informed feedback on the prototype's viability.

Testers provided feedback on access, identity, authentication, and cyber threat resistance, enabling key adjustments and progress. Multiple cyber teams tested the early prototype to ensure it could support sensitive information for real-world users.

This disciplined testing approach delivered tangible results. The cyber teams' efforts helped secure a Conditional Authority to Operate (C-ATO)—allowing the project to continue development. The feedback loop between testers and developers enabled early and rapid iteration on core functionality, security controls, and user experience elements. Alpha testing outputs left the team confident and ready to move into the beta development (Chapter 2).

Lesson Learned from Alpha- Start Before You're Ready

Avoid getting stuck in the planning phase. The early alpha prototype was modest:

- One chat thread per user that disappeared after 24 hours
- Uploading documents took hours to process
- No connections to internal Department of State systems

Despite these constraints, launching early was the right decision. Experiments with real users immediately highlighted which capabilities mattered most; this included things like being able to upload working documents, and connections to Department of State datasets. This allowed the team to prioritize development efforts and build features users needed rather than features the team assumed they wanted.



GENERATIVE AI PLAYBOOK

CHAPTER 1: PROVE THAT IT'S POSSIBLE ALPHA DEVELOPMENT

A multi-layered testing approach was key for each prototype release during the alpha phase. The Department of State used a three-pronged approach for each major release:

- 1. Internal functional testing:** Validate technical infrastructure and functionality, including stress testing, security controls validation, Section 508 accessibility compliance, and regression testing.
- 2. AI performance testing:** Assess functionality, contextual accuracy, misinformation, cybersecurity, UI/UX, and reliable platform performance, aligning with NIST AI Risk Management Framework standards and government-wide mandates.
- 3. User acceptance testing (UAT):** Pull in a subset of users to test early iterations, ensuring the solution meets real-world needs and offers critical usability insights.

Prior to launch, to ensure that StateChat met the highest standards of trust and reliability, the Department of State leaned heavily on its Testing, Evaluation, Verification, and Validation (ITEVV) experts. ITEVV's approach was grounded in Executive Order 13960 (December 2020), which directs federal agencies to adopt AI in ways that protect civil rights, civil liberties, and privacy while building public trust. Drawing from the National Institute of Standards and Technology AI Risk Management Framework (NIST AI RMF), ITEVV evaluated StateChat against seven core characteristics: Safe, Secure and Resilient, Explainable and Interpretable, Privacy-Enhanced, Fair with Contextual Accuracy Maintained, Accountable and Transparent, and Valid and Reliable. ITEVV tested StateChat across 10 risk categories tailored to diplomatic AI systems, ensuring a thorough review of the platform's risk posture.

Lesson Learned: Security Comes First

Securing a FISMA High ATO unlocked StateChat's growth and value. Aiming for a "High" ATO presented initial complexity for the team but delivered long-term benefits.

The ATO enabled:

- Integration of sensitive data sources (cables, SharePoint, internal reports)
- Access to multiple Sensitive But Unclassified (SBU) data categories
- Real utility for users working with actual Department of State information
- Operation within a secure, closed environment
- Easy-to-understand expectations for users

The year-long ATO process required deep partnership with cybersecurity and governance teams from day one. Without the FISMA High ATO, StateChat would have been limited to unclassified, non-sensitive data severely limiting adoption and undermining the tool's core value proposition.



GENERATIVE AI PLAYBOOK

CHAPTER 1: PROVE THAT IT'S POSSIBLE ALPHA DEVELOPMENT

Assemble Your Team: Customer Success Starts at the Ground Floor

Weeks into alpha, the Department of State established barebones “customer success” operations – an office hour, an in-platform feedback button, a weekly newsletter, and a community of practice hosted on Microsoft Teams. While modest, these customer success efforts focused on surfacing user needs and were critical

as we moved towards beta testing (Chapter 2). StateChat's internal team structure was informal at the start; the team was small and streamlined, but as development and adoption grew, so did the team. By the time the Department of State reached more mature development phases, the staffing framework had evolved significantly. Table 1 below details the key roles and responsibilities for a mature GenAI delivery team.

Table 1: StateChat Delivery Team, Roles and Responsibilities

Role	Responsibilities
Product Owner	Business leader that sets the vision, interprets organizational needs into prioritized features, secures executive sponsorship to overcome bureaucratic hurdles
Application Development	Platform engineers and data scientists who build the product
Platform & Governance	Technologists and security analysts who authorize and securely administer an enterprise platform
Testing	AI testing specialists who test new functionality against expected performance requirements defined by the Product Owner
Customer Success	Communications, training, and change management specialists who drive enterprise adoption and fluency
Innovation & Research	AI and analytics specialists who explore emerging technologies, prototype new features, and inform product development and strategy

The StateChat delivery team had a dedicated product owner from day one. The product owner acted as a bridge between technical capabilities and business needs, ensuring the product solved

priority enterprise problems. Clear roles, responsibilities, and a solid product management framework helped the team work efficiently and collaboratively.

<< RETURN TO TABLE OF CONTENTS

CHAPTER 1:

PROVE THAT IT'S POSSIBLE: ALPHA DEVELOPMENT

Summary & Insights



Build a multidisciplinary team early and scale strategically. Start with product owner, developers, platform experts, and testers. Add customer success, governance and policy, and independent verification specialists as the product matures.

Forge internal and external partnerships early. Engage cybersecurity, legal, privacy, IT governance, and accessibility teams from day one. External partners—vendors, research institutions, other agencies—fill knowledge gaps and accelerate development.

Secure governance approvals as a foundational activity. Obtain a FISMA High Authority to Operate (ATO), or highest level for your agency, early by collaborating with cybersecurity teams. Budget a year for this process—it determines what data the tool can access.

Ship a minimum viable product focused on core questions. Build a working prototype that, while limited, is functional enough to test assumptions about security, risk, performance, and accessibility across the organization's environments.

Establish minimum requirements to move to production and test rigorously to validate feasibility and build momentum. Implement multi-layered, phased testing—internal functional, AI performance, and user acceptance testing. Independent verification builds trust and prepares the tool for enterprise use.



GENERATIVE AI PLAYBOOK

CHAPTER 2: PROVE THAT IT'S USEFUL

BETA DEVELOPMENT

The alpha phase of development and testing proved the art of the possible— the Department of State could launch a functional GenAI chatbot, safely and securely, to a select group of global users. The next step was to prove StateChat would be useful to the Department of State through beta (or, phase two) testing. Would users actually adopt the tool and find value in it?

Beta showcased emerging capabilities as the team built them, propagated successes surfaced by users, and constructed a backlog of promising, high-demand features.

During beta, the Department of State scaled StateChat from 150 to 8,000 users over five months by: (1) expanding access strategically; (2) building robust user support systems; and (3) establishing feedback mechanisms to guide development.

Expanding Access Strategically: Learning From Users as We Scale

Expanding the user base in a controlled manner and prioritizing user feedback allowed targeted

testing and established a contract with existing users; stick around, and we will build together.

Scaling fast— and in stages—was critical. By expanding access in phases, the Department of State was able to establish development and testing processes that could be repeated throughout the product's lifecycle, allowing the team to learn from early users, create introductory resources, and deliver them to new users in waves that would not risk team or user burnout.

- ***Keeping it Real: Proactively Manage Expectations***

Controlled testing set expectations with users and Department of State leaders on StateChat's limited, but growing, functionality. The team was upfront at every turn: "StateChat is in beta, and it shows" was a throughline in communications. Humility helped manage expectations.

Lesson Learned **-Feedback** **Will Flood In.** **Have a Plan!**

With targeted deployment, users will share feedback from day one—and that is exactly what you want. But without a system to collect, organize, and prioritize input, teams quickly drown in requests or miss what users actually need.

In StateChat's case, volunteers simply needed to request access through an automated form. As long as they could authenticate with a Department of State credential, they were in. Metrics on usage patterns, prompting success rates, and satisfaction complemented qualitative feedback, creating a complete picture of StateChat's performance before enterprise launch.



GENERATIVE AI PLAYBOOK

CHAPTER 2: PROVE THAT IT'S USEFUL

BETA DEVELOPMENT

Support Users at Every Turn: Driving Adoption and Engagement

Even the most intuitive technology needs thoughtful support to drive adoption—even more when introducing something groundbreaking. Many users did not have a foundational understanding of AI and many were skeptical of this new technology. Outreach and support were critical to GenAI adoption. The Department of State focused on an iterative change management strategy where each element—office hours, resources, training, and communications—worked together to move users toward adoption and build confidence in the tool.

- **Offer Regular “Office Hours” Across Time Zones**

The Department of State offered six “Office Hour” sessions per month, across three time zones allowing staff worldwide to learn during their working hours. Later, StateChat champions were established in distant time zones to host office hours for their regions, so the DC-based team did not have to run events in the early morning or late-night hours.

Office Hours serve as more than a support channel—they are a space for learning, experimentation, and community. From March 2024 to May 2026, the Department of State held 94 Office Hours with about 17,000 attendees.

Each session features live demonstrations where colleagues share how they apply StateChat in their work and open discussion. A skilled host facilitates, but the value comes from colleagues learning from one another. Hundreds of personnel attend each session, sustaining momentum over time.

- **Create a Centralized Resource Hub**

A core principle of the Department of State's approach to scaling adoption of StateChat and its generative AI tools was to “meet users where they are.” This meant the team intentionally designed tailored resources, training, and engagements to fit users' diverse skill levels, workflows, and contexts—rather than requiring users to adapt to a one-size-fits-all approach. An 80,000-person workforce distributed across nearly every time zone, working in dozens of languages and vastly different mission contexts, cannot be supported by a single channel, format, or schedule. This principle shaped every element of the support strategy.

The centerpiece of these resources was AI.State, a SharePoint site accessible to all Department of State employees that served as a centralized hub where users could easily find and access training materials, guides, and support resources.

Lesson Learned - Never Underestimate the Power of a Home Base

The Department of State initially underestimated the importance of a centralized resource as a valuable support tool. For StateChat, an early internal SharePoint site housed use cases from its global users, example prompts, demo videos, a self-paced learning module, and a dashboard with anonymized usage data by post and region.

This site became essential. When the team could not offer exactly what a colleague needed, they could at least direct users to a growing, authoritative repository of resources. The lesson: build this hub early, keep it current, and make it easily accessible.



GENERATIVE AI PLAYBOOK

CHAPTER 2: PROVE THAT IT'S USEFUL BETA DEVELOPMENT

- **Communicate Through Multiple Channels**

The Department of State used a multi-channel communication approach to reach users wherever they were. Weekly email updates, intranet announcements, and presentations at leadership meetings ensured broad visibility and gave users multiple entry points to learn about the tool. This strategy increased awareness, generated enthusiasm, and encouraged innovation by sharing use cases and success stories across the Department of State. More in Chapter 4.

Listen, Prioritize, Repeat: Feedback Mechanisms that Guide Product Development

As usage grows, users will share feedback and ask for additional functionality. Dedicated feedback channels, such as integrated feedback forms and customer satisfaction surveys, make it easy for users to report issues, suggest improvements and share success. Providing multiple avenues for feedback accommodates users with varying preferences and time constraints. With more feedback comes the need for a structured approach to collecting, analyzing, and prioritizing input, while staying agile enough to implement changes quickly.

- **Prioritize User Input Transparently**

The Department of State's goal was to demonstrate commitment to supporting users and improving the product based on real-world experience i.e. solicit feedback, acknowledge receipt, and provide updates on how input is being considered. When incorporating user feedback into new features or applications, clearly communicate how that input shaped the final product so users can recognize their influence.

Establish continuous improvement cycles in partnership with users through transparent communication and ongoing prioritization, not just one-time consultations.

Thunderdome: A Case Study in Collaborative Product Development

The Department of State developed "Thunderdome," a structured, prioritization approach to turn user feedback into actionable development priorities. Here is how it worked at the Department of State:

- **Step 1: Collect Input from Multiple Sources:**

The delivery team gathered feedback from diverse channels, including integrated feedback forms, customer satisfaction surveys, questions raised during office hours, and observations. This multi-source approach ensured the team hears from a broad audience, not just the loudest voices.

- **Step 2: Synthesize and Prepare:** Team reviewed all collected feedback, identified common themes, and surfaced top potential features or improvements.

- **Step 3: Conduct a Focused Prioritization Session:** Team facilitated a collaborative session to discuss collected feedback, assess technical feasibility, evaluate alignment with organizational priorities, and narrow options to a manageable set of high-impact features.

- **Step 4: Invite Users to Vote:** Users voted on the narrowed list of proposed enhancements.

- **Step 5: Create a Transparent Roadmap:** Team leveraged voting results to inform a high-level roadmap of upcoming enhancements, typically planned in six-month increments. The outcomes were shared with users, creating transparency and accountability.



GENERATIVE AI PLAYBOOK

CHAPTER 2: PROVE THAT IT'S USEFUL

BETA DEVELOPMENT

Lesson Learned: Thunderdome: Feedback to Features

StateChat's first Thunderdome analyzed and prioritized feedback from six channels, 800 survey respondents, and 22 interviews with StateChat power users, resulting in a six-month roadmap that reflected the top 10 most-requested features from users. Thunderdome transformed scattered input into a clear development plan that users could see themselves in, building trust and showing that their voices shaped the product.

- **Beta Testing is Not Endless**

Beta testing has a defined endpoint, after which, a product launches. However, launch is not the finish line—products should continue to evolve through ongoing improvements and maintenance after deployment. The Department of State prepared to move from beta testing and into enterprise deployment when key indicators aligned: core functionality performed reliably under load with critical bugs resolved, users consistently provided positive

feedback with sustained engagement, and support systems could handle scale. Importantly, the team distinguished between features requiring near-perfect reliability and those where acceptable imperfections enabled faster deployment and learning. This balanced approach—maintaining high standards for mission-critical functions while embracing controlled experimentation elsewhere—allowed the team to move confidently to launch, which Chapter 3 explores.

CHAPTER 2:

PROVE THAT IT'S USEFUL: BETA DEVELOPMENT

Summary & Insights



Expand access gradually to build scalable systems. Growing from 150 to 8,000 users over five months allowed the team to refine processes and prepare effective support resources.

Build a comprehensive change management strategy. Office hours, resource hubs, and multi-channel communication all work together to drive adoption and build trust. Understand your user personas, their pain points, and design resources around that.

Listen to users and prioritize transparently. Structured approaches like Thunderdome turn scattered feedback into actionable, user-informed roadmaps.

Communicate development decisions openly and often. Share what the team is building, why certain features were prioritized, and what is out of scope. Being upfront about beta limitations—"StateChat is in beta, and it shows"—keeps users as partners.

Know when beta is complete. Move to enterprise launch when technical stability, user validation, operational readiness, and organizational buy-in align.



GENERATIVE AI PLAYBOOK

CHAPTER 3: LAUNCH

GOING LIVE

Alpha proved StateChat was possible. Beta proved it was useful. Enterprise launch proved it could be released successfully to users across the Department of State. This chapter covers how the Department of State planned and executed a coordinated, phased rollout to thousands of users by: (1) creating a launch plan and (2) inviting enterprise users to the platform.

Prepare for Liftoff: Build Your Launch Plan

Launching an enterprise AI tool is a team sport. Everyone must row in the same direction—product owners, application developers, customer success specialists, platform and governance technologists, policy specialists, and testing experts—all synchronized through consistent cross-functional check-ins to knock down barriers as they pop up.

The Department of State developed a launch plan with several key components:

- user materials that explained the tool's capabilities;
- technical infrastructure that ensured smooth functionality;
- communications campaigns to build awareness;
- training programs to build user confidence; and
- targeted outreach to engage key stakeholders.

Within this framework, the team identified key sub-activities that needed to be completed prior to launch. These items informed a launch checklist that acted as a central, cross-workstream set of steps and milestones. This checklist synchronized developers, communications, trainers, and security

teams through daily stand-ups and weekly cross-functional meetings.

- **Treat Launch as a Process, Not an Event**

The Department of State treated launch as a three-week phased process. This allowed the team to test systems, refine processes, and build confidence incrementally. Before full Department of State launch, the team ran dress rehearsals for turning on access for all employees and piloted training with a single large bureau to stress test the platform.

- **Communicate Broadly and Address Concerns**

For StateChat, the launch included email campaigns, internal memos, intranet announcements, and leadership endorsements to ensure employees knew how to access the tool and understood its value. A key focus was addressing security, ethical, and privacy concerns upfront, demonstrating StateChat was a safe, approved tool for Department of State use. This comprehensive rollout-built momentum while establishing confidence in the tool.

- **Make Launch Day Visible and High-Profile**

Launch Day featured the Department of State's CIO, CDAO, and Dean of the School of Applied Information Technology at the National Foreign Affairs Training Center (NFATC) in a Department of State-wide webinar with live demonstrations by users around the world. The event centered on users and its message focused on awareness, relevance, and resources. That visible, high-profile endorsement, paired with immediate access, was effective.



GENERATIVE AI PLAYBOOK

CHAPTER 3: LAUNCH

GOING LIVE

Welcome Users to the Platform

Users should be able to access the tool without unnecessary obstacles. Clear instructions should be provided through multiple communication channels. However, granting access is not enough. The StateChat customer success team focused on two critical elements of building user confidence: leadership endorsement and hands-on training.

- **Shout it from the Mountaintops: Leadership Endorsement**

Department of State leadership played a pivotal role in adoption. Early user feedback revealed a common hesitation: uncertainty about whether they were "allowed" to use StateChat for their work. This is typical when introducing new technologies, especially AI.

To address this, senior Department of State leaders—Assistant Secretaries (A/S) or Under Secretaries (U/S) whenever possible—championed the tool directly with their teams.

These leaders showcased role-specific ways their staff could use StateChat in daily work and explicitly endorsed StateChat as an approved tool for handling SBU information. This top-down approach drove adoption across the organization, as subordinates carried the message forward and encouraged their teams to embrace the new technology.

- **Unlocking Barriers to Entry: Provide Immediate Training**

Leadership permission opened the door. Training helped users walk through it. Tailored training programs build user confidence and ensure successful adoption. Live training and direct engagement opportunities with the development team were offered in the weeks following launch, creating multiple pathways for users to develop comfort with the platform. These initiatives, discussed further in Chapter 4, complemented leadership endorsements.

Lesson Learned:
**Access + Training
= Adoption**

StateChat's first large-scale training event took place shortly before launch and included six sessions for a single Department of State bureau, reaching over 700 participants. In the weeks that followed, the team observed a 58 percent increase in StateChat users within that bureau. The lesson: do not wait to train users. Pair access with immediate opportunities to learn then track the impact.

[<< RETURN TO TABLE OF CONTENTS](#)

CHAPTER 3:

LAUNCH: GOING LIVE

Summary & Insights



Build a launch plan with a detailed checklist. Identify key components—user materials, technical infrastructure, communications, training, stakeholder outreach—and create a master checklist to keep teams aligned through daily stand-ups and weekly meetings.

Treat launch as a phased process, not a single event. A multi-week rollout allows time to test systems and build confidence. Run dress rehearsals and pilot training throughout.

Grant access simply and communicate broadly. Remove barriers to access with clear instructions across multiple channels. Address security concerns upfront. Make Launch Day visible with high-profile leadership and live demonstrations.

Secure leadership endorsement to drive adoption. Senior leaders championing the tool to addressing the common hesitation about whether users are "allowed" to use it. Top-down endorsement accelerates organization-wide adoption.

Pair access with immediate training. Tailored programs offered right after launch build confidence and adoption.



GENERATIVE AI PLAYBOOK

CHAPTER 4: PROVE THAT IT CAN SCALE DRIVING ADOPTION

The launch process secured leader buy-in and seeded confidence in this new set of tools. The next challenge was proving it could scale—that users would integrate the tool into their daily work and early adopters would become sustained advocates.

Granting access is not the same as driving adoption. Scaling requires strategy. The StateChat delivery team scaled adoption across the Department of State's 80,000-person global workforce through five interconnected pillars: (1) strategic communications; (2) targeted "onboarding" (or, training) programs; (3) global live training opportunities; (4) asynchronous learning resources; and (5) partnerships with StateChat/AI champions who multiplied the team's reach.

The team set up pillars during piloting and testing, then adjusted and expanded them to meet growing demand. It created others specifically for enterprise deployment to bolster what already existed. Together, these pillars transformed StateChat into an enterprise-wide solution embedded in the Department of State's workflows.

The 5 Pillars of Adoption

1. Strategic Communications: Transforming Technical Capabilities into Practical Applications

Communications played an important role in driving the use of StateChat. The team timed announcements to align with major functionality releases and delivered them through multiple channels. A Customer Relationship Management (CRM) tool tracked

engagement metrics, enabling the team to test different writing styles and content types. This approach allowed for A/B testing of email subject lines and content to identify which messages achieved the highest open rates, then adapting future communications based on what users gravitated toward.

The most effective approach focused on clear, relatable messaging that translated technical capabilities into practical applications. Communications emphasized not just how to use the tool, but why it mattered for specific roles or tasks; to make connections between technical capability and mission-critical work explicit. For example, instead of announcing "AI-powered document analysis," StateChat would explain *how* policy officers could use the feature to synthesize dozens of documents in minutes and identify emerging trends.

2. Targeted "Onboarding": Specific Bureaus, Regions, and Functions

The Department of State developed a targeted training framework branded "StateChat Onboarding" to bolster the launch strategy. These week-long, interactive and hands-on programs brought together employees from the same organization—whether by region, bureau, or function—ensuring training was relevant to specific workflows and accessible within appropriate time zones. For example, separate onboardings for each regional bureau allowed sessions to address the right languages and use case examples for posts in Africa versus those in Europe.

Lesson Learned - Community and Leadership Drive Adoption

The Department of State enlisted organizational leaders to sponsor onboarding. Senior officials sent direct invitations to their teams, which proved critical: a trusted leader's invitation carries far more weight than a generic announcement. Bureau-focused sessions also created community learning environments where colleagues explored the tool together. This combination—personalized invitations, relevant examples, and collaborative formats—accelerated adoption more effectively than individual training.



GENERATIVE AI PLAYBOOK

CHAPTER 4: PROVE THAT IT CAN SCALE DRIVING ADOPTION

The results validated the approach. Following a tailored onboarding for the Bureau of South and Central Asian Affairs (SCA), StateChat user growth increased by 31 percent in total users and 367 percent in average weekly users.

The Department of State continuously refined onboarding sessions based on participant feedback and usage data analytics, improving training materials, examples, and delivery methods.

3. Live Trainings: Access Across Time Zones

Live, virtual training sessions started shortly after enterprise deployment and gave users hands-on opportunities to explore StateChat's features and apply them to real-world tasks. Unlike "onboarding," live trainings were not sponsored by leaders or tailored to specific functions—they were open to all, and users signed up individually. Two types of live sessions were offered:

- **StateChat 101** – Open-enrollment training courses offered four to six times per month across various time zones, providing foundational training on StateChat's features and ensuring access for the global workforce.
- **Office Hours** – Real-time support sessions where users can come together with peers from around the world and view demos of applicable use cases, ask questions, troubleshoot

issues, share lessons learned, and receive updates on the tool and upcoming features.

To date, these training sessions have reached over 50,000 attendees.

4. Asynchronous Resources: Low Pressure, Self-paced Skill Development

During beta testing, the Department of State created basic documentation—FAQs, quick-start guides, and a SharePoint resource hub. As usage grew and platform capabilities expanded, these resources evolved into a comprehensive self-service ecosystem designed to support new and experienced users.

The team expanded the resource library with specialized content:

- A **Use Case Library** showcased practical applications across roles and job functions, with entries drawn from user-submitted examples. To keep content fresh, the team solicited creative use cases from end users, which were then refined and incorporated into the library.
- A **ProTip Library** provided scenario-based guidance grounded in Department of State workflows.

Lesson Learned— Pair Enterprise Deployment with Immediate, Accessible Training

StateChat's user base grew from 12,000 to 23,000 in the first quarter of FY 2025—its first quarter of live operation. As of July 2026, StateChat serves more than 62,000 users. This growth was driven by targeted onboarding campaigns and user-centric training strategies that made learning opportunities immediately available. The lesson: do not wait to train users. Offering multiple pathways to learn—onboardings, live training, office hours—immediately accelerate adoption by meeting users where they are.



GENERATIVE AI PLAYBOOK

CHAPTER 4: PROVE THAT IT CAN SCALE DRIVING ADOPTION

- **Troubleshooting Guides** addressed common technical issues.
- Self-paced **learning modules** allowed users to build skills at their own pace.
- A **Prompt Guide** provided 100-plus templates, to help craft more effective prompts.
- **Analytics dashboards** allowed users to view high-level StateChat data—such as total users, total prompts, or weekly active users—with results filterable by post and bureau.

The team also built StateChat to include its own user guide and key Department of State information. For example, if a user has trouble uploading a document, instead of searching through the user guide, they can simply ask StateChat, "How do I upload this document?" and receive step-by-step instructions right in the chat based on official resources.

5. Partnerships: Multiplying Impact

The Department of State's long-term training needs in AI far exceed the team's bandwidth.

To sustain and expand training capacity, the Department of State adopted a train-the-trainer approach—equipping internal partners with materials, support, and expertise to deliver StateChat training. Core stakeholders in this are:

- **National Foreign Affairs Training Center:** The Department of State's training institution, NFATC has helped institutionalize successful StateChat trainings within its long-term AI upskilling curricula. NFATC instructors now deliver, and develop, formal AI training to all Department of State personnel as part of their professional development.
- **AI Champions Program Pilot:** Rolled out in March 2026, the Department of State identified and trained over 50 enthusiastic early adopters who promote AI tools at their missions. These champions received StateChat training and ongoing support across a diversity of AI tools and best practices, enabling them to deliver localized training at their posts.

Lesson Learned - Mobile Access: Build for How Users Actually Work

State's 80,000 personnel are distributed globally and often work on the move. Recognizing this, the delivery team developed a prototype mobile application in the form of a "progressive web app" (PWA). This lightweight approach allowed users to access StateChat's interconnected tools from mobile devices without having to build a complex mobile app. They could learn, test, try the tool out on-the-go and on their own schedule, enabling access to technology anywhere. The lesson: interconnected tools must work wherever users are—at their desks, in transit, or at posts around the world. Build and provide training for the reality of how your workforce operates and learns, not just ideal conditions.

[<< RETURN TO TABLE OF CONTENTS](#)



GENERATIVE AI PLAYBOOK

CHAPTER 4: PROVE THAT IT CAN SCALE DRIVING ADOPTION

Lesson Learned- Train-the-Trainer Multiplies Reach

The train-the-trainer approach enabled the Department of State to reach 48,000+ attendees across more than 430 sessions without adding staff. By empowering partners—NFATC instructors, AI Champions, and Bureau Chief Data and AI Officers, among others—to deliver training within their networks, StateChat's impact multiplies, while building sustainable, decentralized training capacity across the Department of State.

- **Bureau Chief Data and AI Officers:** Bureau Chief Data and AI Officers across the Department of State's regional and functional bureaus were trained by the StateChat delivery team to promote StateChat adoption within their organizations. These partnerships leverage existing data governance networks to integrate AI tools into bureau-specific workflows.

Scaling AI: Measuring and Monitoring Adoption

Scaling AI demands benchmarking and ongoing measurement to understand what works, where gaps exist, and how to improve. As StateChat expanded, the Department of State developed systems to track adoption, analyze usage patterns while maintaining user privacy, and iterate based on data.

- **Tracking Key Metrics**
To effectively scale StateChat and ensure it met user needs, the team developed an analytics framework to monitor StateChat adoption and engagement. Key metrics included total prompts submitted, prompts per user per week, weekly average users, retention rate, and others. These metrics allowed the team to identify trends, measure the effectiveness of outreach and training efforts, and pinpoint areas where additional support was needed.
- **Privacy-First Data Analysis**

The team always prioritized balancing analytics with user privacy. Drawing inspiration from a leading commercial AI provider's Contextual Language Input Optimization (CLIO) methodology, the Department of State developed a privacy-protecting prompt analysis process. This approach enabled insights into usage patterns and common challenges while ensuring no sensitive or personally identifiable information was exposed. By prioritizing privacy alongside data-driven improvement, the team maintained user trust while upholding the Department of State's ethical standards.

- **Using Data to Drive Improvements**
Analytics drove continuous improvements across all five pillars. Low adoption rates in specific regions triggered targeted "onboarding" campaigns. Patterns in support requests guided updates to asynchronous resources and guides. Usage data helped identify which communications resonated most with users. The Department of State launched periodic innovation challenges, encouraging users to propose new use cases that the team refined and shared across all staff, fostering deeper engagement and ownership. This feedback loop ensured the adoption strategy remained responsive to emerging needs and challenges.

[<< RETURN TO TABLE OF CONTENTS](#)

CHAPTER 4:

PROVE THAT IT CAN SCALE: DRIVING ADOPTION

Summary & Insights



Use strategic communications to connect features to mission-critical work. Explain how features help users accomplish specific tasks—synthesizing key guidance, drafting memos, researching policy—rather than describing generic AI capabilities.

Deploy targeted onboarding that combines leadership endorsement with community learning. Direct invitations from trusted organizational leaders get users in the door. Bespoke onboarding sessions—to region or function—create collaborative spaces where colleagues learn together, reducing adoption barriers through positive peer influence.

Offer multiple pathways to learn through live training and asynchronous resources. Provide live sessions across time zones, self-paced modules, office hours, and use case libraries to accommodate diverse learning styles and schedules. Meet users where they are.

Scale training capacity through train-the-trainer partnerships. Equip partners—training institutions, AI champions and others—with materials and support multiplies reach without adding staff.

Measure adoption continuously and iterate based on data. Track metrics like weekly average users, retention rates, and user profiles—ensuring the tool reaches employees across different bureaus, roles, geographic locations, and career levels—to identify what works and where gaps exist. Use insights to trigger targeted interventions while maintaining user privacy and trust.



GENERATIVE AI PLAYBOOK

CHAPTER 5: PIVOT TO DEPTH

ADVANCED AI CAPABILITIES

Chapter 4 focused on breadth—scaling StateChat across the Department of State's 80,000-person workforce. Once StateChat achieved widespread adoption, the next challenge emerged: moving from breadth to depth.

Pivoting to depth involves thinking beyond basic chatbots and graduating from general tasks like drafting and summarizing to complex, mission-critical work such as negotiations and policy analysis. Depth also requires empowering users to build customized AI solutions and exploring how more autonomous AI tools can free personnel from routine work to focus on bespoke, high-value tasks. Achieving this depth involves connecting multiple information sources, automating multi-step processes, and enabling development teams and individual users to build specialized applications to meet specific mission needs.

This chapter covers how the Department of State pivoted to depth by (1) building interconnected tool suites within StateChat, (2) redefining workflows to include policy analysis and user-created solutions, and (3) laying the groundwork toward agentic AI systems.

Shifting to Interconnected, Integrated AI Tool Suites

Part of pivoting to depth also means evolving StateChat from a standalone chatbot into a suite of interconnected tools, a StateChat "ecosystem." The Department of State prioritized making it easy

for users to complete complex, multi-step tasks within the StateChat interface. To do so, the team focused on building additional AI functionalities, or tools, within StateChat, that supported those tasks most native to the Department of State's core mission. Today, users can move seamlessly between these AI tools without leaving the StateChat environment.

To accompany these expansions, the team developed "deep dive" trainings on specialized StateChat tools. These sessions helped users master advanced capabilities and get creative with customization, generating valuable feedback that informed future development.

- ***Expanding Capabilities: Specialized Policy Tools***
As StateChat matured, the Department of State expanded the definition of "workflow" beyond routine administrative tasks to include sophisticated policy analysis. This led to a focused effort on identifying and building specialized tools for foreign policy work—such as a pre-mortem analysis tool that helps users stress-test policy proposals.
- ***Expanding Capabilities: User-Created Solutions***
At the same time, the Department of State also leaned into pioneering self-service, user-created tools, enabling individuals to create customGPTs—bespoke AI assistants tailored to their specific needs. This transformed StateChat from a centrally managed tool into a platform for decentralized innovation.

FAMsearch - Integrating Policy Guidance into Workflows

FAMsearch, a functionality within the StateChat ecosystem, simplified how Department of State personnel access and use the 30,000-page Foreign Affairs Manual and Foreign Affairs Handbook (FAM/FAH). Users ask FAMsearch specific policy questions and receive synthesized answers drawn from relevant excerpts, complete with direct FAM/FAH citations. The tool represents one of StateChat's early steps toward agentic AI—over time, the team shifted from a FAMsearch tab to users being able to query the FAM/FAH directly in StateChat, with an agent retrieving the answer from FAMsearch autonomously.

[<< RETURN TO TABLE OF CONTENTS](#)



GENERATIVE AI PLAYBOOK

CHAPTER 5: PIVOT TO DEPTH

ADVANCED AI CAPABILITIES

Lesson Learned- customGPTs: Shifting from Consumer to Creator

Putting users in the driver's seat. The most important shift in advanced AI adoption is not technical. It's a change in mindset from "What can the AI team build for me?" to "What can I build to solve my problem?" Moving users from dependency to agency is as important as any technical capability the tools provide.

StateChat's customGPT feature makes this shift possible. Users can build their own specialized AI assistants directly within StateChat—whether helping create recurring tailored reports or answering mission-specific administrative questions. This democratization serves two purposes: first, it meets niche needs that would never rise to formal IT development but nonetheless save time and improve work quality. Second, the act of building their own tools demystifies AI for the personnel who use it; rather than treating AI as something delivered to them, users can craft it directly to serve their specific mission.

Agentic AI Systems: The Future Approaches

The Department of State envisions American diplomats commanding world-class AI capabilities to advance U.S. interests. This vision rests on two objectives: efficiency—saving time and eliminating redundant work—and capability advantage—ensuring American diplomats have the best tools to compete and collaborate on behalf of the American people.

Agent orchestration is how the Department of State will deliver on both. At the time of writing, Spring 2026, the Department of State is preparing for a global change toward agentic AI systems where multiple specialized agents work autonomously together to complete complex workflows.

- ***From Tools to Agents: The Evolution of FAMsearch***

FAMsearch demonstrates this evolution in action. Originally built as a standalone tool, FAMsearch allows users to query the 30,000-page Foreign Affairs Manual and Foreign Affairs Handbook and receive synthesized answers with direct citations. Over time, the team shifted from requiring users to open a separate FAMsearch tab to enabling StateChat to autonomously query the FAM/FAH in the background. As a result, when a user asks a

policy question, an agent now retrieves the answer from FAMsearch without the user switching applications—the orchestration happens invisibly. Similarly, new tools can "call" the FAMsearch agent for use in other developments. Build it once, the right way and reuse widely.

This represents the Department of State's broader direction: moving from tools users must explicitly activate to agents that work together autonomously. Imagine a foreign policy officer drafting a cable—an official report to and from U.S. embassies and posts abroad—that mentions visa policy. An agent could automatically flag relevant FAM/FAH guidance, while another agent checks recent cables on the same topic, and a third suggests standard formatting—all without the foreign policy officer requesting each action separately. The agents orchestrate the retrieval of relevant data based on context.

As these systems mature, diplomats will have more time and better AI tools to focus on high-impact work—policy analysis, negotiation, relationship-building—while AI systems handle routine processes. Agent orchestration represents the depth StateChat has been building toward: AI that doesn't just passively assist Department of State employees with State's work but actively accelerates and augments the employees.

[<< RETURN TO TABLE OF CONTENTS](#)

CHAPTER 5:

PIVOT TO DEPTH: ADVANCED AI CAPABILITIES

Summary & Insights



Evolve from standalone tools to interconnected ecosystems. Build multiple functionalities within a single interface—cables applications, drafting tools, formatting capabilities—so users complete complex, multi-step tasks without switching between systems.

Empower users to create their own AI solutions. customGPTs shift the paradigm from "What can IT build for me?" to "What can I build?" This democratization meets niche needs and transforms how users understand AI—from mysterious to accessible.

Provide specialized training as capabilities expand. Deep dive sessions on advanced features help users maximize the platform's potential while generating feedback that informs future development.

Transition from tools to autonomous agents. Move toward systems where specialized agents work together invisibly—automatically surfacing policy guidance, checking related cables, suggesting formatting—based on context rather than explicit user requests.



GENERATIVE AI PLAYBOOK

EPILOGUE: MEASURE IMPACT

After one year of enterprise deployment, the Department of State conducted a comprehensive survey-based impact assessment to measure how the tool affects Department of State personnel and mission outcomes. The voluntary survey ran from August 26 to September 26, 2025, and aimed to supplement existing usage data with detailed insights into how employees use StateChat, the benefits it provides, and the challenges they encounter. The results inform the development of new features, resolve common issues, and refine training strategies.

Impact Assessment Process and Methodology

The Department of State developed the survey in collaboration with experts from the Bureau of Intelligence and Research and the Bureau of Human Resources. The assessment was structured around four guiding research questions:

- Who uses StateChat?
- How is it being used?
- What is its impact?
- What facilitates or prevents adoption?

- ***Sampling and Data Collection***

Of the 3,700 employees who began the survey, 2,706 provided substantial responses for analysis, including 2,113 "regular users" (those who use StateChat at least monthly) and 593 "non-users" (those who use it less than monthly or not at all). Respondents included Civil Service, Foreign Service, and Locally Employed Staff across domestic and overseas locations, with 66 percent of StateChat users based overseas.

The survey used an opt-in approach rather than random sampling. As a result, statistical inference is not appropriate for this data; however, the substantial number of responses offers meaningful insight into user experiences. Branching logic tailored questions to regular users and non-users, allowing for targeted insights into usage patterns and barriers. Focus groups with non-users provided deeper understanding of their workflows and reasons for not adopting the tool.



GENERATIVE AI PLAYBOOK

EPILOGUE: MEASURE IMPACT

Impact Assessment Results

90%

of respondents report time savings

1.6 Hours

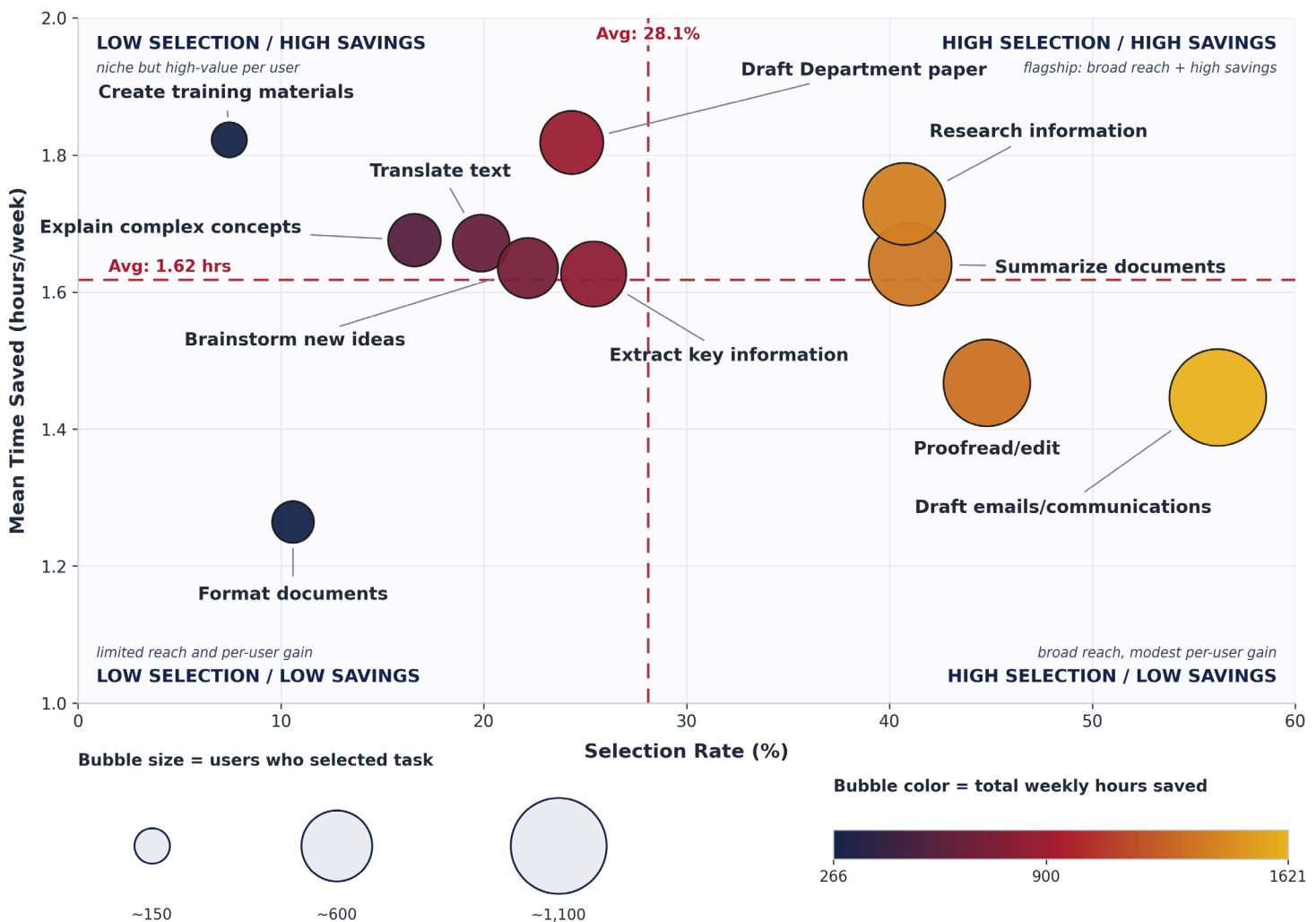
Reported average weekly time savings per task for StateChat users

"Improve the quality of my work"

Top reported use of time saved

Time Savings Impact Matrix

Selection Rate vs Mean Time Saved (bubble size = users who selected, color = total weekly hours saved)



<< RETURN TO TABLE OF CONTENTS



GENERATIVE AI PLAYBOOK

EPILOGUE: MEASURE IMPACT

Impact Assessment Results

A significant majority of regular users (90 percent) report saving time through StateChat, with 96 percent indicating they would recommend the tool to colleagues. On average, users save 1.6 hours per week per task—time they report using to improve work quality, respond to more requests, and pivot to higher-impact tasks.

StateChat users report using the tool most frequently for drafting communications, proofreading, and research. Notably, 51 percent of survey respondents report saving time drafting emails and communications, totaling over 1,600 hours per week among survey participants alone. Given that StateChat regularly has over 19,000 unique weekly users, data suggests users are saving well over 10,000 hours every week on drafting activities alone.

- **Powerful Use Cases**

The impact assessment revealed both common and "hidden gem" use cases with significant time savings:

- **Drafting communications:** This was the most common reported use case (selected by approximately 51 percent of users) with time savings averaging 1.6 hours per week.
- **Creating training materials:** While only 7 percent of respondents list this among their top StateChat uses, it shows the highest average weekly time savings at 1.82 hours, with 36 percent of users reporting saving 3+ hours per week on this task.
- **Summarizing documents:** A frequently cited use case with above-average time savings.

- **How StateChat Transforms Work: Real-World Examples**

The survey captured compelling examples of how StateChat is transforming workflows across the Department of State:

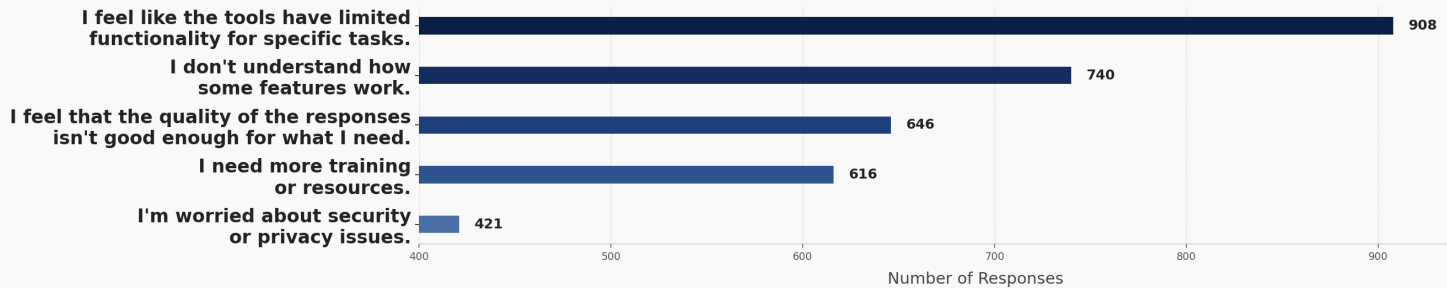
- **Automating HR Onboarding:** An HR coordinator reported how StateChat automated the onboarding workflow, reducing the process from 2–3 hours per new employee (with frequent errors) to a streamlined, error-reduced process. This time savings allowed them to launch a mentorship program and improve internal training resources—higher-value tasks that directly impact employee experience and retention.
- **Accelerating Document Drafting:** A user responsible for drafting a cable incorporating input from multiple posts used StateChat to synthesize information, reducing hours of manual revision. The time saved allowed them to immediately begin a new project on space technology outreach, which might have been delayed or missed otherwise.



GENERATIVE AI PLAYBOOK

EPILOGUE: MEASURE IMPACT

Top 5 reported challenges for users trying to use StateChat tools



• Challenges and Barriers to Adoption

The assessment also identified key barriers that prevent some users from realizing time savings with StateChat:

- Among regular users who report that StateChat does not save them time, the top blockers cited include insufficient quality of responses, more time spent verifying outputs than manually completing tasks, and poor integration with systems the user already uses.
- Focus groups revealed that early enthusiastic adopters sometimes disengage due to accumulated small barriers: promised features that don't materialize, rescheduled office hours, training mismatched to skill levels, and interface challenges compared to commercial alternatives.

Looking Forward: Recommendations for Future AI Implementations

Based on the impact assessment findings, future AI implementations should consider:

- **Diversifying Use Cases:** Develop and promote use cases for non-policy functions to ensure AI tools serve the entire workforce
- **Role-Based Training:** Tailor training to different roles and skill levels
- **Local Champions:** Continue fostering post-level

AI working groups with designated champions

- **Continuous Improvement:** Address small barriers quickly before they accumulate and cause user disengagement
- **Measuring Impact:** Measuring Impact: Continuously improve how impact is measured by refining quantitative metrics and qualitative feedback methods to better evaluate outcomes and guide development. The StateChat journey—from concept to enterprise-wide adoption—demonstrates the transformative potential of responsible AI in government. With over 60,000 users across 92 percent of global posts as of June 2026, StateChat has evolved from an experimental tool to an essential platform that saves tens of thousands of hours weekly, enabling diplomats to focus on what matters most: advancing America's interests through human connection and expert judgment. As other agencies embark on similar AI journeys, we hope the lessons from StateChat's implementation and impact assessment can provide a measure of guidance. The future of AI in government isn't just about technology—it's about empowering and augmenting public servants to work smarter, collaborate more effectively, and deliver better outcomes for the American people. By measuring impact continuously and adapting to user needs, agencies can ensure their AI investments deliver meaningful, mission-focused results that strengthen America's position as a global leader in responsible AI adoption.

[<< RETURN TO TABLE OF CONTENTS](#)

EPILOGUE:

MEASURE IMPACT

Chapter Summary & Insights



- Measuring impact is about more than usage stats—it's about understanding real-world benefits and challenges.
- User-centered assessments (surveys, focus groups) provide valuable insights.

Key Activities & Deliverables



- **Design and Run Impact Assessments:** Use surveys and focus groups to gather both quantitative and qualitative feedback from users and non-users.
- **Analyze and Share Results:** Turn data into actionable insights and communicate findings back to the user community.

APPENDIX:

AI Tools You Can Use at the Department of State, Beyond StateChat



GENERATIVE AI PLAYBOOK

APPENDIX: DOS AI TOOLS

Northstar: Northstar is a media analytics tool developed by the Department of State's Bureau of Global Public Affairs that uses artificial intelligence to supercharge how Department of State teams understand the global media environment. The platform gives Department of State teams centralized access to comprehensive insights, data, and analytics from 100,000+ global media sources in 100+ languages in near real-time, with one million pieces of news and social media content added to this repository every day, including from Department of State-owned social media accounts. By translating and summarizing these global conversations, Northstar frees time for Department of State teams to focus on higher-value strategic work.

SMART AI: The Department of State Messaging and Archive Retrieval Toolset (SMART) serves as the command-and-control platform for handling official correspondence. SMART leverages AI technologies to enhance knowledge management by analyzing, organizing, and generating personalized data of cables and record emails, tailored to users' interests and access privileges.

Proving Ground, Funhouse, and Prompter API Overview

The Department of State has developed a suite of complementary tools, connective tissue, and environments that extend beyond StateChat to support the broader AI ecosystem across the Department of State. These initiatives—Proving Ground, Funhouse, and Prompter API—represent the Department of State's strategic approach to democratizing AI capabilities while maintaining appropriate security controls and governance.

These platforms enable distributed innovation by providing secure, accessible environments where Department of State personnel with varying technical backgrounds can experiment, build, and deploy AI solutions tailored to specific diplomatic needs. Together, they form a comprehensive framework that supports the Department of State's vision for responsible AI adoption and enhances diplomatic capabilities while fostering a culture of innovation.

Proving Ground offers Department of State staff a secure space to build custom AI solutions using the same platform that powers StateChat. This free-tier environment allows technical staff to create workflow automations, specialized chatbots, and data visualizations without setting up independent development environments. By providing access to enterprise-grade tools and resources in a ready-to-use workspace, Proving Ground removes technical barriers for teams with ideas for improving diplomatic processes. This shared space enables innovation across the Department of State while maintaining security standards, serving as both a testing ground for potential enterprise solutions and a practical learning environment.



GENERATIVE AI PLAYBOOK

APPENDIX: DOS AI TOOLS

Funhouse provides a secure sandbox within Data.State where Department of State personnel can experiment with data science and AI techniques using Azure Databricks. Unlike Proving Ground's focus on application development, Funhouse specializes in data analysis, visualization, and machine learning experiments. This cloud-based environment allows users to test advanced analytical approaches to diplomatic challenges without establishing separate computing infrastructure. The sandbox encourages creative exploration of data-driven solutions in a secure setting that meets Department of State requirements, supporting the Department of State's commitment to innovation through accessible, properly governed data science tools at a low aggregate cost.

Prompter API provides Department of State developers with secure, pre-approved access to cutting-edge commercial AI capabilities from any Department of State system. With proper security to handle Sensitive But Unclassified information, this tool allows teams to add AI features to existing applications without navigating complex approval processes. Developers can incorporate capabilities like document drafting, translations, summarization, and coding assistance into their applications through simple, standardized connections.

[<< RETURN TO TABLE OF CONTENTS](#)



Published July 2026