harmonic

Code Red: Analyzing China-Based App Use

www.harmonic.security

Executive Summary

Chinese-developed generative AI (GenAI) applications are emerging as a high-risk blind spot within enterprise environments. To better understand this trend, Harmonic Security conducted a 30-day analysis across approximately 14,000 US and UK-based end users.

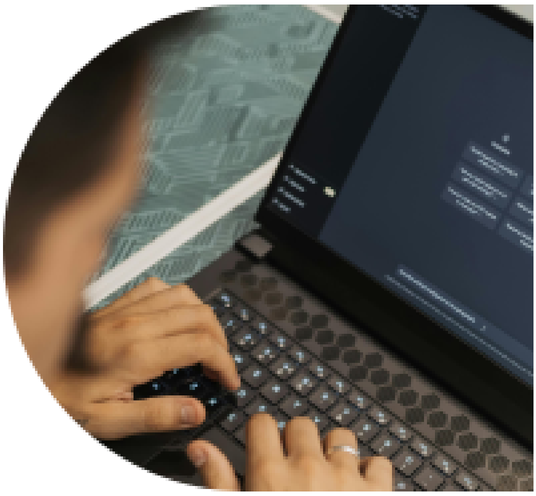
Harmonic's study identified widespread, unsanctioned use of Chinese GenAI tools such as DeepSeek, Moonshot Kimi, Manus, Baidu Chat, and Qwen. These tools offer free, powerful capabilities that are increasingly attractive to developers; but pose outsized risk to intellectual property and security.

Key findings include:

- 7.95% of employees in the average enterprise used at least one Chinese GenAI tool.
- 1,059 users uploaded over 17MB of content, much of it potentially sensitive.
- 535 incidents of sensitive data exposure were recorded across five platforms.
- The majority of exposed content was software-related:
  - 32.8% involved code, proprietary logic, or access credentials
  - Other sensitive categories included M&A documents (18.2%), PII (17.8%), and financial data (14.4%).

These platforms offer minimal transparency around data retention or model training practices. In many cases, their policies allow for submitted content to be stored or reused, creating unacceptable risk for organizations handling proprietary software, regulated data, or client information.

Engineering-heavy organizations are particularly exposed, as developers increasingly turn to GenAI for coding assistance. However, few realize the implications of submitting internal source code, API keys, or system architecture into foreign-hosted models.



Overview and Methodology

In response to the rapid growth in popularity of DeepSeek and similar Chinese LLMs, Harmonic conducted a 30-day behavioral analysis of approximately 14,000 end users, primarily based in the United States. The focus was on activity tied to SaaS-based GenAI applications and was collected via the Harmonic Security Browser Extension, before anonymized and sanitized for analysis.

Please note: this excluded self-hosted or open-source LLMs, which remain a separate but important risk vector.

Our data includes file upload volumes, usage frequency, and detections of sensitive content transmission. This dataset offers a focused lens into how these tools are being used within corporate environments and where the highest exposure risks lie.

Development of DeepSeek Alternatives

DeepSeek’s rapid rise has spurred the growth of several competing GenAI platforms, each gaining traction across enterprise environments. The four additional platforms we observed (Moonshot Kimi, Manus, Baidu Chat, and Alibaba’s Qwen) are often accessible via simple browser interfaces, making them easy for employees to use but difficult for security teams to monitor.

What unites these apps is a general lack of transparency regarding data use. Several platforms either do not specify retention policies or confirm they may retain and reuse inputs for training, which can create legal liabilities, particularly around trade secrets, intellectual property, and personal data.

Chinese GenAI App Usage Patterns

Usage of Chinese GenAI apps is already present across many enterprise environments. On average, 7.95% of employees in each organization used at least one Chinese GenAI chat tool over a 30-day period. These tools are easily accessible and often used without security oversight.

While user counts vary, data upload volumes suggest meaningful engagement. Each company uploaded an average of 1.2MB of data to these apps. enough to include code snippets, business content, or structured data.

The most used apps by total data uploaded were:

1. Moonshot Kimi
2. DeepSeek
3. Qwen
4. Baidu Chat
5. Manus

Code Exposure and High-Risk Data Leakage

Among the 1,059 users who engaged with Chinese GenAI tools, 535 incidents of sensitive data exposure were detected.

The majority of exposure occurred via DeepSeek, which accounted for roughly 85% of the total, followed by Moonshot Kimi .

In terms of content type, code and development artifacts represented the largest category, making up 32.8% of all exposures. This included proprietary code, access keys, and internal logic.

Other exposed data types included:

- Mergers & acquisitions data – 18.2%
- Personally identifiable information (PII) – 17.8%
- Financial information – 14.4%
- Customer data – 12.0%
- Legal documents – 4.9%

These findings highlight how technical users—especially developers and “vibe coders”—are unintentionally leaking valuable assets. This kind of exposure presents risks not just to security, but to intellectual property.

Key Recommendations

Shadow AI use is already common. Nearly 1 in 12 employees are using Chinese GenAI tools, often without oversight.

Blocking alone is rarely effective and often misaligned with business priorities. Even in companies willing to take a hardline stance, users frequently circumvent controls.

A more effective approach is to focus on:

- **Focus on Education.** Train employees on the risks of using unsanctioned GenAI tools, especially Chinese-hosted platforms.
- **Provide Alternatives.** Provide approved GenAI tools to meet developer and business needs.
- **Enforce Controls.** Enforce policies that prevent sensitive data, particularly source code, from being uploaded to unauthorized apps.

www.harmonic.security

North America
1390 Market Street
Suite 200
San Francisco
California 94102
USA

Europe
York House
221 Pentonville Rd
London
N1 9UZ
UK