



Security & Privacy Disclosure, Final Report

Character.AI (ai.character.app) v1.16.4, versionCode 146

A twelve-network advertising stack auto-initializes via ContentProviders before the age-verification screen can ever run, Chinese ad SDKs sit inside that same stack with no documented transfer safeguard, a session-replay engine can screen-record intimate conversations, and a hardcoded Firebase key leaves quota and account-enumeration vectors open, unresolved through 109 days and a 19-day-overdue embargo, with every one of nine messages answered only by an automated Help Center ticket, never a human

**PUBLIC DISCLOSURE, 17 OCTOBER 2026 — EMBARGO REACHED 28
SEPTEMBER 2026, SIX FINDINGS REMAIN UNADDRESSED**

Target	Character Technologies, Inc., Menlo Park, California, USA (no confirmed EU main establishment)
Package	ai.character.app
Version Audited	v1.16.4, versionCode 146 (APK SHA-256 d066de86...499385ce)
Platform Audited	Android, root-level static analysis of the production APK
Initial Disclosure (R1)	2026-06-30 (REF CAI-2026-R1-001)
Disclosure / Embargo Date	2026-09-28
Report Published	2026-10-17
Overall Severity	CRITICAL
Findings	6 (1 Critical, 4 High, 1 Medium)
Correspondence	9 messages sent by RFI-IRFOS across two Zendesk-fragmented threads, automated Help Center ticket replies received to every single one (4 distinct ticket numbers confirmed: 1415230, 1421971, ZPP6W5-9N5DW, 29LG2J-P47R9), 0 substantive replies ever received from privacy@character.ai or any named human
Authorities BCC'd / CC'd	Österreichische Datenschutzbehörde (DSB) · European Data Protection Supervisor (EDPS)

security@rfi-irfos.com • rfi-irfos.com

ZVR 1015608684 • GISA 39261441 • UID ATU83405245 • Steuernummer 68 696/8736 • Graz, Austria

Contents

1	Executive Summary	4
2	Subject & Operator Context	5
3	Engagement Summary	5
4	Findings	6
4.1	C1: A 12-network advertising stack auto-initializes via ContentProviders before the age-verification screen can ever run	6
4.2	H1: Chinese ad SDKs (Pangle/ByteDance and Mintegral) wired into the pre-consent init stack, an undocumented Chapter V transfer	7
4.3	H2: Amplitude Session Replay: on-screen recording capability on a platform built for intimate conversation	7
4.4	H3: Biometric and automated age-assurance of minors without a documented Art. 35 DPIA	8
4.5	H4: Hardcoded, unrestricted Firebase/Google API key: quota, enumeration, and password-reset-flood vectors	9
4.6	M1: No network security config or certificate pinning for any first-party endpoint	10
5	Impact Analysis	10
6	The Complete Chain: A Statutory Disclosure Answered Nine Times by the Same Account-Deletion Template	11
7	Data Protection – Article by Article	12
8	Regulatory & Legal Landscape	12
9	Indicators of Compromise	13
10	Disclosure Timeline & Outcome	13
11	Residual Risk & Recommendations	14

Executive Summary

Character Technologies, Inc. operates Character.AI, an AI-companion chat platform whose user base was substantially composed of minors until a policy change in November 2025. A root-level static analysis of the production Android binary (v1.16.4) found six findings, one CRITICAL, four HIGH, one MEDIUM.

The lead finding is structural rather than a single misconfiguration. The app bundles a twelve-network advertising and analytics stack (AdMob, AppLovin, Vungle, Meta Audience Network, Unity, ironSource, InMobi, Fyber, Moloco, CloudX, Amplitude, AppsFlyer, Braze) registered as Android ContentProviders, which the Android lifecycle runs during Application.onCreate, before any Activity, including the age-verification screen, can appear. No consent management platform or IAB TCF framework exists anywhere in the binary to gate this. The sequence itself, provable directly from the manifest, means the advertising stack was live and capable of accessing device and advertising identifiers before a user's age was ever established, during a period when most users were minors (C1). Two Chinese-origin networks, Pangle (ByteDance) and Mintegral (Mobvista), sit inside that same auto-init stack, raising a Chapter V international-transfer question that the PRC National Intelligence Law's state-cooperation obligation makes non-trivial (H1). Amplitude Session Replay, a screen-recording engine, is registered and capable of capturing on-screen content on a platform whose core function is intimate, sometimes romantic conversation, content that can reveal Art. 9 special-category data (H2). The app's own age-assurance system, while genuinely more sophisticated than commonly assumed, mechanically constitutes Art. 9 biometric processing and Art. 22 automated decision-making over minors without a documented DPIA (H3). A hardcoded, unrestricted Firebase/Google API key enables quota-exhaustion, user-enumeration, and password-reset-flood vectors against the live service (H4). No network security config or certificate pinning exists for any first-party endpoint (M1).

RFI-IRFOS disclosed these findings on 30 June 2026 to privacy@character.ai and support@character.ai, BCC the Österreichische Datenschutzbehörde and the European Data Protection Supervisor. Character.AI's Zendesk-based Help Center routed every single message into an automated ticketing system across two threads over the following 74 days; across nine messages sent, Character.AI's system returned an identical generic account-export/deletion boilerplate reply every time, under at least four distinct ticket numbers, never once engaging with a single finding or the three yes/no questions restated in every follow-up. The embargo, stated once as 2026-09-28 and never contested or revised by either side, is now 19 days overdue with zero substantive reply ever received.

ID	Severity	Title
C1	CRITICAL	A 12-network advertising stack auto-initializes via ContentProviders before the age-verification screen can ever run
H1	HIGH	Chinese ad SDKs (Pangle/ByteDance and Mintegral) wired into the pre-consent init stack, an undocumented Chapter V transfer
H2	HIGH	Amplitude Session Replay: on-screen recording capability on a platform built for intimate conversation
H3	HIGH	Biometric and automated age-assurance of minors without a documented Art. 35 DPIA

ID	Severity	Title
H4	HIGH	Hardcoded, unrestricted Firebase/Google API key: quota, enumeration, and password-reset-flood vectors
M1	MEDIUM	No network security config or certificate pinning for any first-party endpoint

Subject & Operator Context

Character Technologies, Inc. is a US company headquartered in Menlo Park, California, operating Character.AI, an AI-companion chat platform. The company has no confirmed EU main establishment, meaning there is no one-stop-shop supervisory authority and each concerned national authority is independently competent; RFI-IRFOS accordingly addressed this disclosure to the Österreichische Datenschutzbehörde with the European Data Protection Supervisor informed in parallel. Until a policy change in November 2025, the platform's user base was substantially composed of minors. This audit covers v1.16.4 (versionCode 146) of the production Android binary.

Engagement Summary

RFI-IRFOS sent the initial disclosure on 30 June 2026 to privacy@character.ai and support@character.ai, BCC the Österreichische Datenschutzbehörde and the European Data Protection Supervisor. Character.AI's Zendesk-based Help Center auto-replied twice within minutes with a generic account-export/account-deletion template (ticket ZPP6W5-9N5DW), addressing neither the findings nor the three yes/no questions posed in the disclosure.

A follow-up sent six days later, 6 July 2026, restating the three most severe findings, received the identical boilerplate again. From 17 July 2026 onward the correspondence continues across a second Zendesk thread, with further follow-ups on 27 July (two messages the same day, the second explicitly noting the same canned reply had now arrived three times, character for character), 15 August, 24 August, 31 August, and 12 September 2026, the supervisory authorities moved from BCC to CC from this point onward. Every single one of these nine messages received only an automated Help Center acknowledgment in return, confirmed across at least four distinct ticket numbers (1415230, 1421971, ZPP6W5-9N5DW, 29LG2J-P47R9), each reusing the identical account-export/deletion template verbatim. Not one human reply addressing any finding or question was ever received.

Findings

C1: A 12-network advertising stack auto-initializes via ContentProviders before the age-verification screen can ever run

CRITICAL (CVSS v4.0 AV:N/AC:L/AT:N/PR:N/UI:N/VC:H/VI:L/VA:N/SC:N/SI:N/SA:N, 8.5) — advertising and analytics SDKs capable of accessing device and advertising identifiers before any consent surface or age gate exists, on a platform whose user base was substantially minors until November 2025

```
<provider android:name="com.google.android.gms.ads.MobileAdsInitProvider" android:
  initOrder="100"/>
<provider android:name="com.applovin.sdk.AppLovinInitProvider" android:initOrder
  ="101"/>
<provider android:name="com.vungle.ads.VungleProvider" android:initOrder="102"/>
<provider android:name="com.facebook.ads.AudienceNetworkContentProvider"/>
<provider android:name="com.google.firebase.provider.FirebaseInitProvider"
  android:directBootAware="true" android:initOrder="100"/>
```

Consent management search of the Hermes bundle:

OneTrust [0] Didomi [0] Usercentrics [0] IABTCF [0] (single unstructured gdprConsent flag only)

Age-verification flow: runs in an Activity, after ContentProvider init

Impact: The manifest declares ContentProviders for AdMob, AppLovin, Vungle, and Meta Audience Network, plus a directBootAware Firebase provider, alongside classes for Unity, ironSource, InMobi, Fyber, Moloco, CloudX, Amplitude, AppsFlyer, and Braze, twelve networks in total. Android ContentProviders execute during Application.onCreate, structurally before any Activity, including the age-verification screen, can run. No consent management platform (OneTrust, Didomi, Usercentrics) or IAB TCF framework exists anywhere in the Hermes JavaScript bundle to gate this sequence, only a single unstructured gdprConsent flag. This means the twelve-network advertising stack was live and structurally capable of accessing device and advertising identifiers before a user's age was ever established, during a period when the platform's user base was substantially composed of minors. RFI-IRFOS blast-radius escalation: the lawful-basis gap alone would be HIGH, raised to CRITICAL because the population exposed during the relevant window was substantially minors and the processing is pre-consent by construction, not by a configuration error that could be toggled off case by case. Whether each individual network actually transmitted data before the gate at runtime is held for dynamic (R2) testing, which RFI-IRFOS does not perform; the sequence itself is structurally proven in the binary regardless.

H1: Chinese ad SDKs (Pangle/ByteDance and Mintegral) wired into the pre-consent init stack, an undocumented Chapter V transfer

HIGH (CVSS v4.0 AV:N/AC:L/AT:N/PR:N/UI:N/VC:H/VI:L/VA:N/SC:N/SI:N/SA:N, 7.5)
 — data routed to PRC-origin networks subject to the National Intelligence Law, inside the same stack already proven to run pre-consent

```
Lcom/bytedance/sdk/openadsdk/api/PangleAd
Lcom/bytedance/sdk/component/embedapplog/PangleEncryptManager
Lcom/google/ads/mediation/pangle/PangleConstants (Pangle wired into AdMob mediation)
Lcom/mbridge/msdk/activity/MBBaseActivity (Mintegral / Mobvista)
```

SCC / transfer impact assessment documentation found: none

Impact: The binary embeds Pangle (ByteDance, Beijing) with its own encrypted transport, wired in as an AdMob mediation network, and Mintegral (Mobvista), both inside the same C1 auto-init stack. Data routed to these PRC-origin networks engages GDPR Chapter V (Art. 44, Art. 46 Standard Contractual Clauses plus a transfer impact assessment), and the PRC National Intelligence Law Art. 7 compels assistance to state intelligence on demand, a factor any transfer impact assessment must address. No SCC or transfer impact assessment documentation was found. Whether these networks are active in EU mediation waterfalls at runtime is held for dynamic (R2) testing; capability and integration are demonstrated directly in the binary.

H2: Amplitude Session Replay: on-screen recording capability on a platform built for intimate conversation

HIGH (CVSS v4.0 AV:N/AC:L/AT:N/PR:N/UI:N/VC:H/VI:N/VA:N/SC:N/SI:N/SA:N, 6.9)
 — screen-recording infrastructure capable of capturing Art. 9 special-category content revealed through intimate chat

```
Lcom/amplitude/android/sessionreplay/config/ConfigManager
Lcom/amplitude/android/sessionreplay/config/MaskLevel
assets/amp-rrweb.js (233 KB, rrweb screen-recording engine)
```

Default masking behaviour for conversation content at runtime: unverified

Impact: The binary bundles Amplitude Session Replay, registered as a React Native plugin, together with a 233 KB rrweb DOM/screen-recording engine (assets/amp-rrweb.js). On a platform whose core function is intimate, emotional, and at times romantic conversation, screen recordings can capture content revealing health status, sexual orientation, religious belief, or political views, Art. 9 special-category data, and historically that of minors. A MaskLevel configuration capability exists in the code; whether conversation content is masked by default at runtime is unverified and a fair, direct question for the controller.

H3: Biometric and automated age-assurance of minors without a documented Art. 35 DPIA

HIGH (CVSS v4.0 AV:N/AC:L/AT:N/PR:N/UI:L/VC:L/VI:N/VA:N/SC:N/SI:N/SA:N, 5.8)
 — Art. 9 biometric processing and Art. 22 automated decision-making over a minor population with no documented impact assessment

```
useStartAgeVerificationSheet, isUser018, isUserCoppa, tCOPPAStatus,
  EntrywayBirthdateSheet
https://withpersona.com/verify?inquiry-id=%s
biometric_bucket, affects_liveness, does_not_affect_liveness
```

Documented Art. 35 DPIA covering age-assurance: not found

Impact: Age is determined through a three-stage system: self-declared date of birth, a behavioural age classifier, and escalation to Persona government-ID or biometric liveness verification to unlock the 18+ experience. RFI-IRFOS recognises this as genuine, engineered child-safety infrastructure, considerably more sophisticated than the commonly repeated claim that Character.AI has no age gate at all. Mechanically, however, this constitutes Art. 9 processing of biometric data for age estimation, requiring a specific Art. 9(2) exemption; Art. 22 automated decision-making restricting access to a service; and high-risk processing of minors' data that triggers a mandatory Art. 35 Data Protection Impact Assessment. No DPIA documentation was found. The self-declared primary gate is also trivially defeated by entering any birth year, a data-accuracy question independent of the DPIA gap.

H4: Hardcoded, unrestricted Firebase/Google API key: quota, enumeration, and password-reset-flood vectors

HIGH (CVSS v4.0 AV:N/AC:L/AT:N/PR:N/UI:N/VC:N/VI:N/VA:H/SC:N/SI:N/SA:N, 6.5)
— an unauthenticated key enabling availability and user-enumeration attacks against the live service

```
<string name="google_api_key">AIzaSyBYjIdjN5T49bIWDGX00qyr_WMLRRVeMMU</string>  
<string name="project_id">character-ai</string>  
<string name="google_storage_bucket">character-ai.firebaseio.com</string>  
<string name="google_app_id">1:458797720674:android:78bc41f75bd4767feb9d3b</string>
```

Authentication required to exploit any of the four vectors above: none

Impact: res/values/strings.xml contains a hardcoded Firebase/Google API key, project ID character-ai, storage bucket character-ai.firebaseio.com, and Google App ID, none requiring authentication to use. RFI-IRFOS confirms four concrete, no-auth-required vectors enabled by this key's exposure: (1) quota-exhaustion denial of service against the live Firebase API, degrading the service for every real user, (2) user-enumeration via the Identity Toolkit fetchSignInMethodsForEmail endpoint, revealing which email addresses are registered, (3) password-reset-flood, triggering reset emails for all registered users simultaneously, inbox spam and support-channel collapse, and (4) storage probing against the Firebase Storage bucket if its rules are misconfigured, an industry-common baseline failure. Google's own framing that this key class is "public by design" describes the architecture, not what becomes possible once every actor holds it; the burden of demonstrating none of these four vectors is exploitable rests on the controller, per the accountability principle of Art. 5(2).

M1: No network security config or certificate pinning for any first-party endpoint

MEDIUM (CVSS v4.0 AV:A/AC:H/AT:P/PR:N/UI:N/VC:L/VI:L/VA:N/SC:N/SI:N/SA:N, 3.8) — TLS channel open to interception by an attacker able to install a trusted CA on a managed or hostile-network device

```
android:networkSecurityConfig attribute: not declared
res/xml/ network security config file: not present
targetSdkVersion: 35 (platform-default cleartext denial applies)
Certificate pinning for character.ai / characterai.io / auth.character.ai: none
```

Impact: No android:networkSecurityConfig attribute and no network security config file exist in res/xml/. Because targetSdkVersion=35 denies plaintext HTTP by platform default, this is scored lower than it otherwise would be, a point credited to the app. However, the absence of any network security config means no certificate pinning exists for character.ai, characterai.io, or auth.character.ai, leaving the TLS channel open to interception by an attacker able to install a trusted CA on the device, a realistic scenario on managed devices or hostile networks. Art. 32(1)(a) calls for state-of-the-art measures proportionate to the sensitivity of the data in transit, which here includes intimate conversational content.

Impact Analysis

This report's CRITICAL finding is structural, not a single misconfigured flag. C1 demonstrates that the Android lifecycle itself, ContentProviders executing before any Activity, guarantees a twelve-network advertising stack is live before the age-verification screen the company itself built can ever run. The company's own child-safety engineering, the COPPA block, the Persona liveness escalation, the under-18 chat shutdown, all real and acknowledged in this report, sits at the front door, while the advertising and analytics stack initializes in the basement before that door is ever reached. H1 and H2 compound the exposure of the same pre-consent window, routing data to PRC-origin networks and enabling screen-capture of intimate conversation respectively. H3 and H4 are more conventional gaps, an undocumented DPIA for a child-facing biometric system, and a hardcoded key with concrete availability and enumeration consequences, but both sit inside a platform whose core function already concentrates unusually sensitive signals.

What ties all six findings together procedurally is the same failure already documented against other targets in this disclosure series with a ticketing-platform support backend: a statutory GDPR disclosure sent to a dedicated privacy address was never once engaged with by a human. Character.AI's Zendesk-based Help Center treated this exactly like an account-deletion request, repeatedly, across nine messages and 109 days.

The Complete Chain: A Statutory Disclosure Answered Nine Times by the Same Account-Deletion Template

Step 1 – Disclosure sent: RFI-IRFOS sends the full six-finding disclosure to privacy@character.ai and support@character.ai on 30 June 2026, BCC the Österreichische Datenschutzbehörde and the European Data Protection Supervisor. Character.AI's Zendesk Help Center auto-replies twice within minutes with a generic account-export/deletion template, ticket ZPP6W5-9N5DW.

Step 2 – Six days, first follow-up: On 6 July 2026, RFI-IRFOS follows up restating the three most severe findings. The identical boilerplate returns under a new ticket number, 29LG2J-P47R9.

Step 3 – Seventeen days, a direct request for routing: On 17 July 2026, RFI-IRFOS explicitly asks that the matter be routed to Character.AI's privacy or legal team directly, noting two closed tickets with zero engagement. The same automated reply follows.

Step 4 – Twenty-seven days, the pattern is named explicitly: On 27 July 2026, two messages are sent, the second one explicitly noting that the same canned reply has now arrived three times, character for character. The automated reply arrives a third time regardless.

Step 5 – Continued silence through August and into September: Further follow-ups are sent on 15 August, 24 August, and 31 August 2026, each restating the open findings and the days elapsed; the supervisory authorities move from BCC to CC from 31 August onward. Each receives the identical automated acknowledgment.

Step 6 – Final follow-up: On 12 September 2026, a ninth and final message is sent. The automated Help Center reply follows once more, now under at least a fourth distinct ticket reference across the correspondence.

Step 7 – Embargo reached and passed: The embargo, stated once as 28 September 2026 and never contested or revised by either side, passes with all six findings unaddressed.

Step 8 – Publication: This report publishes 19 days after the embargo date. Across 109 days and nine messages sent, every single response Character.AI's systems produced was an automated Help Center ticket acknowledgment reusing the same account-export/deletion template. It was never reviewed by a human, let alone resolved.

Data Protection – Article by Article

Art. 6(1), Art. 7(1), Art. 5(1)(a), Art. 5(2) GDPR; ePrivacy Art. 5(3) – Lawfulness, conditions for consent, accountability: C1's twelve-network advertising stack initializes with no consent management platform to gate it, before any age-verification screen can run.

Art. 44, Art. 46, Art. 35 GDPR – Transfers and impact assessment: H1's Pangle and Mintegral integrations have no documented SCC or transfer impact assessment addressing the PRC National Intelligence Law's state-cooperation obligation.

Art. 5(1)(c), Art. 25, Art. 9, Art. 35 GDPR – Data minimisation, privacy by design, special categories: H2's session-replay infrastructure can capture Art. 9 content revealed through intimate conversation with unverified default masking.

Art. 9, Art. 22, Art. 35, Art. 8 GDPR – Special categories, automated decision-making, impact assessment, children's data: H3's biometric age-assurance of a substantially minor population has no documented DPIA.

Art. 32(1)(b), Art. 25 GDPR – Security of processing, data protection by design: H4's hard-coded, unrestricted API key enables four distinct no-auth-required attack vectors against the live service.

Art. 32(1)(a), Art. 32(1)(b) GDPR – Security of processing: M1's absence of any network security config leaves the TLS channel open to interception via a trusted CA on a compromised device.

Regulatory & Legal Landscape

Character Technologies, Inc. is a US controller headquartered in Menlo Park, California, with no confirmed EU main establishment, meaning there is no one-stop-shop supervisory authority and each concerned national authority is independently competent. For this Austrian-initiated matter the competent authority is the Österreichische Datenschutzbehörde, with the European Data Protection Supervisor informed in parallel. A US controller targeting the EU market must also designate an Art. 27 EU representative; none was found in the binary or in the company's public privacy documentation.

Fine exposure under GDPR: Art. 83(5) applies to the Art. 6, 7, 9, 22, and 44 issues in C1, H1, and H3, up to EUR 20,000,000 or 4% of total worldwide annual turnover. Art. 83(4) applies to the Art. 32 and 25 issues in H2, H4, and M1, up to EUR 10,000,000 or 2% of total worldwide annual turnover. A platform whose user base was substantially composed of minors during the exact window in which a twelve-network advertising stack is structurally proven to have initialized pre-consent is the kind of fact pattern a supervisory authority is likely to treat as a priority matter.

Indicators of Compromise

The following identifiers, SDK packages, and manifest declarations were extracted from the production APK via static analysis only. No Character.AI servers, accounts, or infrastructure were probed; the hardcoded API key named in H4 was identified by inspection, not by use.

```
C1 -- AdMob, AppLovin, Vungle, Meta Audience Network, Unity, ironSource, InMobi,
      Flyer, MoLoco, CloudX, Amplitude, AppsFlyer, Braze (12 networks, ContentProvider
      auto-init)

H1 -- Pangle (com.bytedance.sdk.openadsdk), Mintegral (com.mbridge.msdk)

H2 -- Amplitude Session Replay (com.amplitude.android.sessionreplay), amp-rrweb.js

H3 -- Persona (withpersona.com/verify), biometric_bucket, isUserCoppa

H4 -- google_api_key: AIzaSyBYjIdjN5T49bIWDGX00qyr_WMLRRVeMMU,
      project_id: character-ai

M1 -- networkSecurityConfig: not declared
```

Disclosure Timeline & Outcome

2026-06-30 Initial disclosure sent to privacy@character.ai and support@character.ai, BCC the Österreichische Datenschutzbehörde and EDPS. Automated Help Center reply, ticket ZPP6W5-9N5DW.

2026-07-06 Follow-up sent restating the three most severe findings. Identical automated reply, ticket 29LG2J-P47R9.

2026-07-17 Follow-up explicitly requesting routing to the privacy/legal team. Automated reply follows.

2026-07-27 Two follow-ups sent; the second explicitly notes the canned reply has now arrived three times, character for character. Automated reply follows regardless.

2026-08-15 Follow-up sent. Automated reply follows.

2026-08-24 Follow-up sent. Automated reply follows.

2026-08-31 Follow-up sent; supervisory authorities move from BCC to CC from this point onward. Automated reply follows.

2026-09-12 Ninth and final follow-up sent. Automated reply follows.

2026-09-28 Embargo date reached, as stated once in the original disclosure and never revised. Correspondence record: nine RFI-IRFOS messages, at least four distinct automated Help Center ticket numbers, zero substantive replies.

2026-10-17 This report published, 19 days after the embargo date. All six findings confirmed unaddressed on the basis of the correspondence record.

Residual Risk & Recommendations

C1 – Immediate: gate all twelve advertising and analytics ContentProviders behind a consent management platform with IAB TCF support, verified to initialize strictly after the age-verification Activity, not merely after app launch.

H1 – Immediate: document a Standard Contractual Clause and transfer impact assessment for the Pangle and Mintegral data flows, or remove both networks from any EU-facing build until that documentation exists.

H2 – Near-term: confirm and document the default MaskLevel applied to session-replay recordings of conversational content, and disclose Amplitude Session Replay as a distinct processing purpose.

H3 – Near-term: commission and document an Art. 35 DPIA covering the biometric age-assurance pipeline, and close the birth-year self-declaration bypass.

H4 – Immediate: rotate the exposed Firebase/Google API key and apply App Check or equivalent request-origin verification to close the quota, enumeration, and password-reset-flood vectors.

M1 – Near-term: add a network security config with certificate pinning for character.ai, characterai.io, and auth.character.ai.

Across all findings: establish an intake path for statutory security and data-protection disclosures that does not route through the same Zendesk Help Center queue used for account-export and deletion requests. Across nine messages and 109 days, every single response was the identical automated template; the dedicated privacy mailbox itself never engaged with a single finding, a process gap worth correcting independently of the six technical findings in this report.

RFI-IRFOS security@rfi-irfos.com • rfi-irfos.com

ZVR 1015608684 • GISA 39261441 • UID ATU83405245 • Steuernummer 68 696/8736 • Elisabethinergasse 25/10, 8020 Graz, Austria All findings derive from root-level static analysis

of the published application using publicly available tooling. No Character.AI servers, accounts, or infrastructure were probed; the hardcoded API key named in this report was identified by inspection, not by use. Technical evidence is archived under chain of custody and will be provided to regulatory authorities on request. Research conducted under ISO/IEC 29147, the freedom of scientific research (Art. 17 StGG) and GDPR Art. 89. REF CAI-2026-R1-001.