



CASE STUDY

The AI admin plane that never checked for an admin

Enterprise software vendor · AI / LLM penetration test (authenticated, non-admin)

1 High, 2 Medium, 4 Low — one root cause

Anonymised under NDA. Client and identifying details are withheld. PXLSecurity LTD. · Sofia, Bulgaria · pxlsec.com



CASE STUDY

The AI admin plane that never checked for an admin

Sector	Enterprise software vendor
Engagement	AI / LLM penetration test (authenticated, non-admin)
Outcome	1 High, 2 Medium, 4 Low — one root cause

1H / 2M / 4L	1	0
across seven findings	missing authorisation gate behind them all	live data redirected or exfiltrated

THE CHALLENGE

A software vendor had embedded generative-AI features in its platform and wanted them tested the way a customer's own staff would meet them — as an ordinary, non-administrator tenant user holding nothing but a normal business login.

HOW IT UNFOLDED

1. Tested throughout as a plain business user, on a session token that asserted no administrator role.
2. Found the administration interface hid its AI screens client-side only — the services behind them were still reachable, and accepted writes on coarse service-level scopes alone.
3. Wrote the AI pipeline's object-storage credential set — endpoint, region, bucket, access key and secret — then read it back, only partially masked.
4. Wrote the generative-AI prompt configuration, and a separate business-system credential store, with the same ordinary token.
5. Triggered the AI data-ingestion job and operated the full model lifecycle — create a model, train it, delete it.
6. Confirmed it was a gap rather than a design decision: the platform's own identity service enforced an administrator gate and refused the same class of operation for the same user.

KEY FINDINGS

High	AI storage-credential set writable — and readable — by a non-administrator
Medium	Generative-AI prompt configuration writable by a non-administrator
Medium	Session not invalidated on logout
Low	Client secret exposed in an OIDC redirect
Low	AI generation endpoint with no rate limiting
Low	API reflected arbitrary CORS origins
Positive	The platform's identity service enforced the administrator gate correctly

IMPACT

An ordinary business user could decide where the AI pipeline wrote its training and inference data, rewrite the prompts behind generated output, and delete production models — with no administrator role anywhere in the request.

OUTCOME

Seven findings — one High, two Medium, four Low — delivered as a single root cause rather than a list of endpoint bugs: one administrator-only gate, applied to every mutating operation across the AI administration surface. The platform already had that gate; it had simply not been applied here. Every value written during testing was a deliberately inert marker, itemised for removal in the handover.

TECHNIQUES

Broken access control · Client-side-only authorisation · AI pipeline abuse · Secrets exposure · ML lifecycle abuse

SERVICES

AI / LLM Penetration Testing · Penetration Testing

Work with PXLSec

Independent offensive-security testing since 2014 — penetration testing, red teaming and secure development, delivered by senior engineers with evidence for every finding and a retest to prove the fix.

Request a proposal: pxlsec.com/contact · hello@pxlsec.com

PXLSecurity LTD. · 12 Vasil Levski Blvd., 1142 Sofia, Bulgaria · VAT BG203239331