CyberArmor.AI 客户门户 — 用户指南

最后核验时间：2026-10-29。全文已对照 `main` 分支提交 1f4497db6 的门户代码逐节核对。截图拍摄于 2026-10-24，环境为新加坡生产环境门户（am-app.cyberarmor.ai）、版本 1.0.13-rc.12，演示组织为"CyberArmor.AI Asia"；此后部分界面已有变化（详情窗口、身份 / SSO、设置、报告）。如图片与正文不一致，以正文描述为准。

本指南将逐屏演示如何使用 CyberArmor.AI 客户门户。每张图片都标有红色方框和箭头，精确指示应点击的位置。图中红色圆圈内的编号与图片下方带编号的说明一一对应。

我应使用哪个访问地址?

区域	客户门户地址
亚洲及中东地区（含巴基斯坦）	**https://am-app.cyberarmor.ai**
美国	**https://app.cyberarmor.ai**

两个门户的界面与功能完全相同。本指南以亚洲及中东地区门户为例。如果贵组织运行自己的 CyberArmor.AI 服务器，请使用管理员或合作伙伴提供的地址。

想看视频? 六集培训视频《客户门户培训》第 1 至 6 章逐一演示了本门户的每个视图。视频位于文档站点 **https://docs.cyberarmor.ai/zh/training/**、**https://cyberarmor.ai/training**，以及门户内的**支持 → 帮助与常见问题 → 培训视频**（第 15 节）。视频目前为英语（巴基斯坦）版。

目录

1. 登录
2. 熟悉界面布局（含"每一行都能打开详情"）
3. 仪表盘（任务控制台）逐项说明
4. 下载软件并获取安装令牌
5. 从合规策略包添加策略（示例：中国 PBOC / NFRA 与 MLPS 2.0）
6. 修改策略（含"优先级决定哪一条策略生效"）
7. 创建自己的策略（含银行卡 PIN 码示例）
8. 启用或停用策略
9. 查看您的设备（终端）
10. 查看运行状况（策略事件、遥测、安全事件、影子AI）
11. 报告与审计日志
12. 添加同事（用户）
13. 登录安全（身份 / SSO）
14. 适用于所有设备的设置（自动响应、私有网络地址、本地页面检查、本地模型）
15. 获取帮助与培训视频
16. 快速参考：全部菜单项
17. 逐一认识其他视图

---

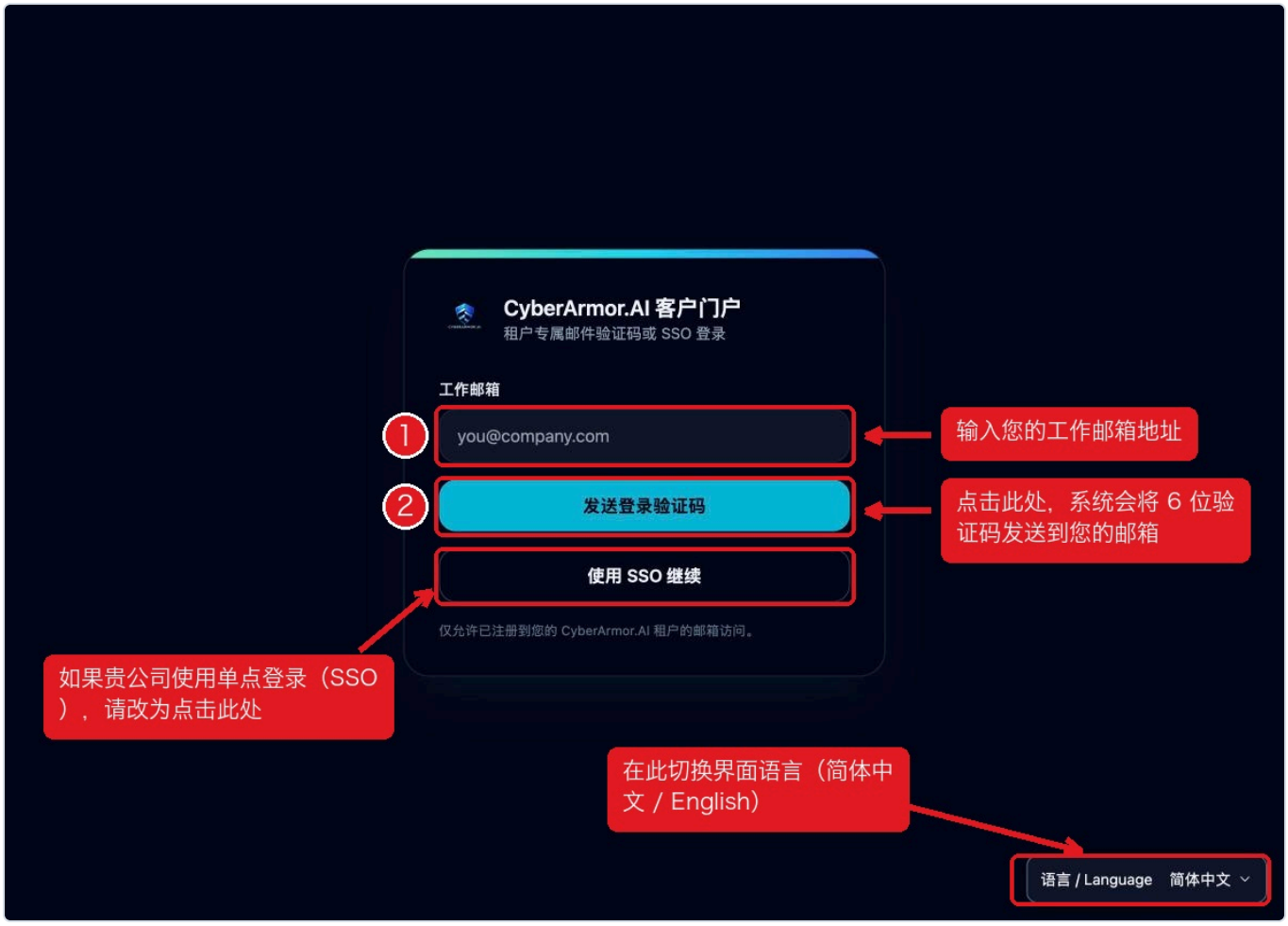
1. 登录

步骤 1. 在 Chrome、Edge 或 Safari 中打开 **https://am-app.cyberarmor.ai**。

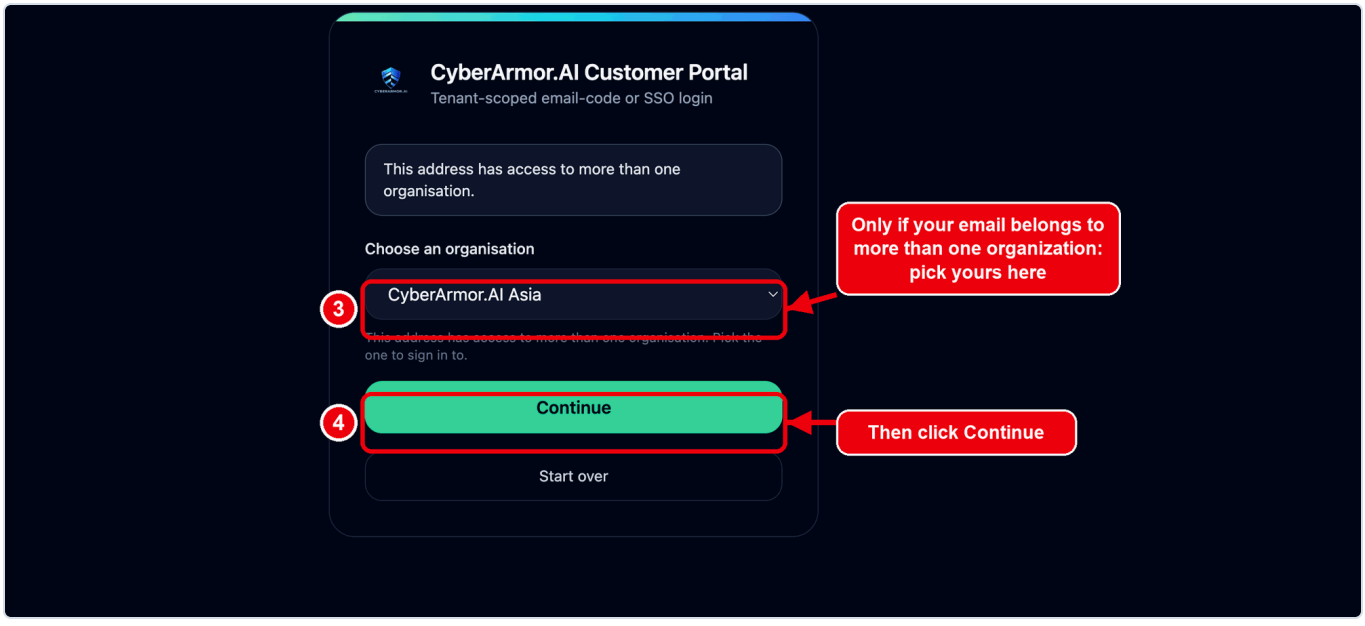
步骤 2. 输入您的工作邮箱（1），然后点击**发送登录验证码**（2）。一分钟内您会收到一封包含 6 位验证码的邮件。在下一个界面输入验证码，然后点击**验证并进入**。（**重新发送验证码**会再发一封；**使用其他邮箱**可重新开始。）

您无需使用密码。如果贵公司采用单点登录（Microsoft Entra ID、Okta 等），请改为点击**使用 SSO 继续**。

如果您已设置身份验证器应用（第 13 节），门户随后会要求输入**身份验证器验证码**：输入应用中显示的 6 位验证码或一个备用码，然后点击**验证并进入**。



步骤 3（仅部分用户需要）。 如果您的邮箱地址隶属于多个组织，门户会请您**选择组织**。请从列表中选择（3），然后点击**继续**（4）。



步骤 4. 首次登录时会自动启动一段简短的引导游览。点击**下一步**跟随游览（约两分钟，共 11 步），或点击**跳过游览**。您日后可随时从**帮助与常见问题**或 **?** → **开始引导游览**中重新播放。



2. 熟悉界面布局

门户的每个页面都采用相同的布局：

- **左侧菜单。** 包含门户的全部页面，按职能分组：*防护*、*检测与响应*、*资产清单*、*证据*、*设置与管理* 以及 *支持*。点击分组标题可将该组折叠。
- **筛选视图（菜单顶部）。** 输入几个字母，例如 `pol`，即可快速定位页面。
- **页面标题（顶部）。** 显示您当前所在的页面，下方附有一行简要说明。
- **限定租户范围（右上角）。** 提示您只能看到本组织的数据。其他组织的人员无法查看贵组织的数据。
- **?（右上角）。** 打开一个菜单：**开始引导游览**、**帮助与常见问题**、**文档**（即本文档站点）和**支持门户**。
- **您的组织、邮箱与角色（菜单底部）**，以及**退出登录**按钮。
- **语言（右下角）。** 在 English 与简体中文之间切换门户语言。

如果由合作伙伴管理贵组织并为您打开了门户，页面顶部会显示**合作伙伴会话**横幅，并提供 ← **返回合作伙伴控制台**。



每一行都能打开详情

点击任意列表中的任意一行——事件、安全事件、设备、策略、用户、制品、API 密钥、SIEM 连接器——都会打开一个详情窗口，用通俗的文字列出该记录的每个字段。点击该行中除其自身按钮之外的任意位置即可；使用键盘时，移动到该行并按回车。

- 窗口底部的 ▸ **查看原始 JSON** 会显示服务器发送的原始记录，供技术团队使用。显示前，凡是保存凭据的字段都会从这份副本中移除。
- **策略。** 由策略作出决定的记录会显示该策略的名称、处置动作与优先级，并提供**打开策略**，可直接打开该策略自身的详情。如果该策略已被删除或归档，或名称无法核实，该行会如实说明，而不会留空。
- **内容扫描。** 说明内容是否经过扫描以及扫描结果，例如"已扫描：未发现敏感数据"，或"未扫描——被目标地址策略拦截"。空的发现项列表绝不代表"未扫描"。
- **关联。** 对来自浏览器扩展或无代理传感器的记录，说明它是否已与同一台计算机上的终端代理关联：**已关联**（以及关联方式、计算机名称与操作系统账户）、**未关联**（以及原因）或**未检查**（以及原因）。

有两种详情窗口没有**查看原始 JSON**：策略详情窗口（完整显示其存储的条件，整条策略可在**编辑** → **JSON** 中查看），以及审计记录详情窗口（完整打印原始签名记录）。

---

3. 仪表盘（任务控制台）逐项说明

概览页面即**任务控制台**，也是您登录后看到的首个页面。它一眼即可回答三个问题：*是否已完成配置？数据是否正在接入？是否发生过什么情况？*



序号	名称	含义
1	**策略**	贵组织拥有多少条安全规则，其中有多少条已归档。
2	**终端**	有多少台计算机、手机和浏览器已接入 CyberArmor.AI。
3	**遥测**	CyberArmor.AI 已接收的活动记录总数、最近 24 小时的记录数，以及来源设备数量。
4	**安全事件**	有多少次操作被拦截、脱敏，或被标记为需要关注的发现项。
5	**AI 服务提供商**	您已接入并交由 CyberArmor.AI 管理的 AI 服务数量（例如 OpenAI 或 Anthropic）。
6	**API 访问日志**	您的门户与设备同 CyberArmor.AI 服务器通信的次数。
7	**BOM**（物料清单）	在您的设备上发现的 AI 工具、模型和程序库数量，以及今日新增的数量。
8	**租户就绪状态**	依据一份五步检查清单，反映贵组织配置完成度的评分。每个勾号代表一个已完成的步骤；点击每一步可打开完成该步骤的页面。
9	**建议的下一步操作**	指向最有价值的后续操作的快捷入口。**引导式入门**将引导您完成配置；**导出证据**会打开**报告**页面，证据导出位于该页面。

序号	名称	含义
10	**遥测——最近 24 小时**	过去一天内每小时一根柱状条，以及最常见的活动类型。柱状条越高表示该小时越繁忙。
11	**威胁态势**	分为两部分。**强制执行决策**（Enforcement decisions）按与**安全事件**完全相同的口径，统计最近 24 小时内的每一次拦截、脱敏与警告，注明每次拦截来自哪个数据存储，并提供 **Open Incidents** → 链接。下方是同一天的遥测记录按处置方式的分布：**Block**（被策略拦截）、**Stopped, unchecked**（因无法检查而被阻止，未涉及任何策略）、**Sent, unchecked**（未经检查即已发出，需要重点关注）、**Redact**（敏感数据在外发前已被移除）、**Warn**（已向用户发出警告）、**Detect**（记为发现项）、**Monitor**（仅记录）、**Allow**（经策略允许）。占比很小时显示为"<0.1%"，绝不显示为"0.0%"。

门户的每个视图对"拦截"的统计口径相同，且默认都显示最近 24 小时，因此任务控制台、遥测、安全事件、策略事件、代理程序目录、操作图谱与 AI 风险仪表盘上的数字彼此一致，并各自注明数字来自哪个数据存储。



序号	名称	含义
12	**近期活动**	最新的 10 条事件。点击任意行可查看全部详情（第 2 节）。**查看全部遥测**可打开完整列表。
13	**快捷操作**	八个一键快捷入口：策略、策略构建器、终端、代理程序目录、遥测、用户、API 密钥和报告。
14	**租户范围**	显示贵组织的名称与 ID、当前登录人员，并确认服务器已强制实施组织间的数据隔离。
15	**您的账户**（菜单底部）	显示贵组织 ID、您的邮箱、您的角色，以及**退出登录**。

4. 下载软件并获取安装令牌

所有需要安装的内容都集中在一个页面：**设置与管理** → **接入与下载**。



1. **控制平面地址。** 您的设备所连接的门户地址。安装过程中需要选择或输入该地址。亚洲及中东地区为 `https://am-app.cyberarmor.ai`。
2. **租户信号。** 绿色圆点表示您的设备正在实时发送数据。
3. **下载 ZIP。** 获取 **CyberArmor.AI Protect**，即适用于 Chrome、Edge、Brave 和 Opera 的浏览器扩展，这是最快的入门方式（约 3 分钟）。
4. **签发引导令牌。** 生成扩展程序接入贵组织所需的一次性令牌。
5. 从此处**复制**控制平面地址。

向下滚动至**终端代理**（Endpoint Agent）分组，可获取 Mac 与 Windows 安装程序：**CyberArmor.AI Endpoint Agent——macOS 安装程序（PKG，已签名并经过公证）**与 **CyberArmor.AI Endpoint Agent——Windows 安装程序（MSI，测试版）**。**浏览器、IDE 和 Office** 分组中是浏览器扩展 **CyberArmor.AI Protect**，**SDK 与 RASP** 分组中是开发者使用的软件包。在各张卡片上点击下载按钮（1），然后点击**签发引导令牌**（2）。每台计算机需单独签发一个令牌。



点击**签发引导令牌**后，令牌会显示在页面顶部。点击**复制**。该令牌：

- 仅可使用一次，且只能用于一台计算机；
- 将在 **30 分钟**后过期，因此请在即将安装前再签发；
- 属于机密信息。请勿通过邮件或聊天工具发送。



页面底部的**本地模型包**会在开启本地模型后（第 14 节）列出本组织的设备端模型包，供手动安装和离线（物理隔离）环境安装使用。

Windows、Mac 与浏览器扩展的详细安装步骤，请参阅配套的**安装指南：Windows、Mac 与 Chrome**。

---

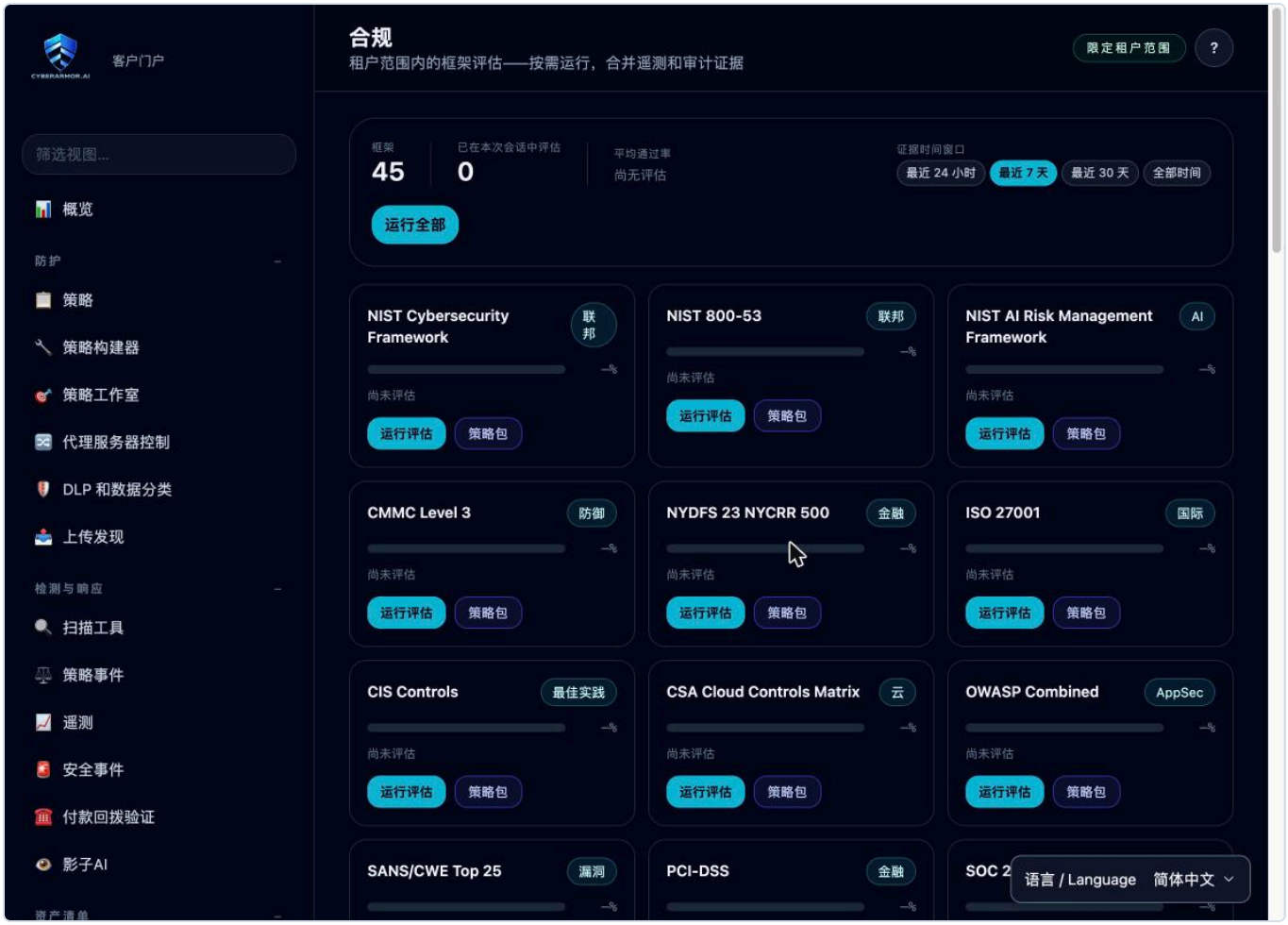
5. 从合规策略包添加策略（示例：中国 PBOC / NFRA 与 MLPS 2.0）

策略就是一条规则，例如"记录与 AI 服务的每一次对话"或"在数据发送至 AI 服务前移除其中的身份证号码"。您无需自行编写：**合规**页面中的大多数法规框架都配有现成的**策略包**。

合规页面列出 45 项法规框架，其中 44 项配有策略包（SANS/CWE Top 25 有意不提供策略包：其控制项针对的是软件缺陷，而非 AI 流量规则）。面向中国市场的客户，最直接相关的是 **PBOC / NFRA with MLPS 2.0 (China)**——它将人民银行与金融监管总局对技术外包的监管要求，以及网络安全等级保护 2.0 的审计留存要求，落实为可执行的技术控制项。

步骤 1. 在菜单中点击**证据** → **合规**。找到 **PBOC / NFRA with MLPS 2.0 (China)** 卡片，点击其上的**策略包**。

（同一张卡片上的**运行评估**功能，会以贵组织的真实活动数据为证据，评估当前对该框架的符合程度并保存结果；打开页面时会载入最近一次保存的结果，**查看详情**可逐项查看各控制项。）



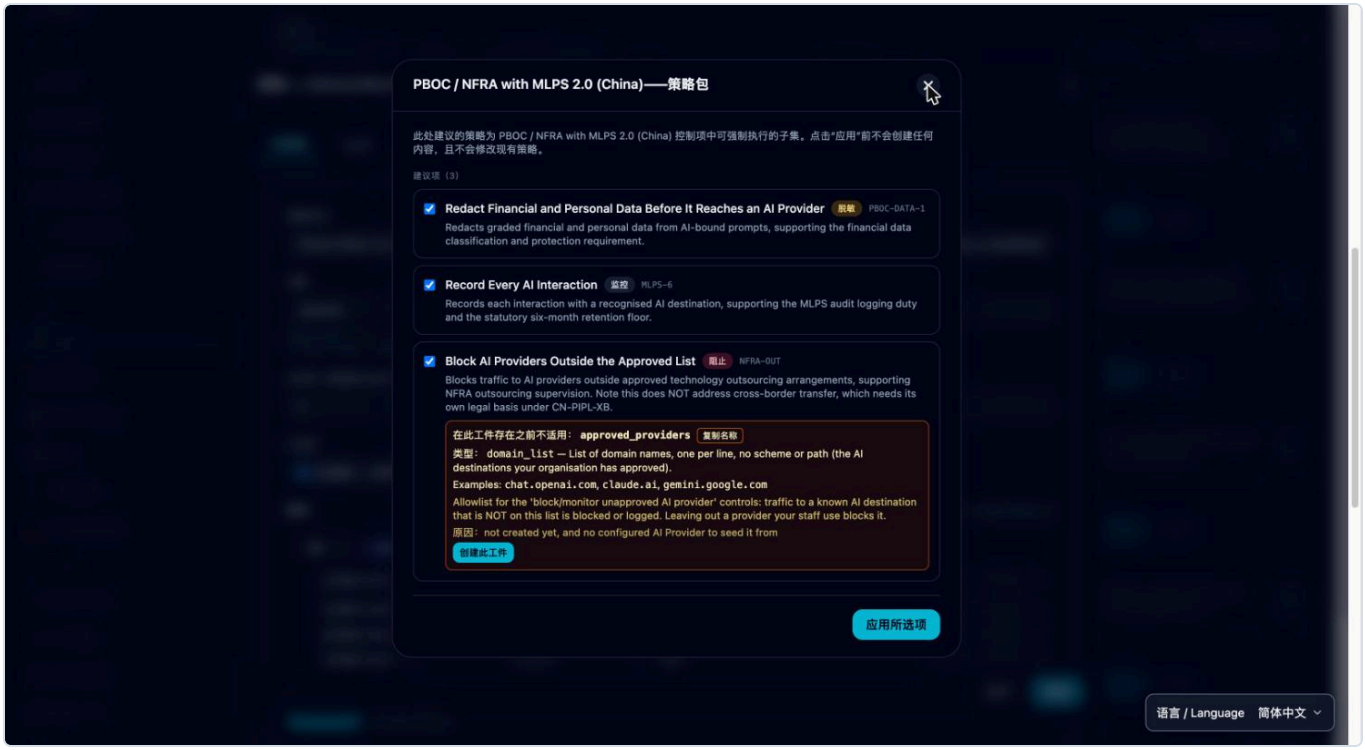
步骤 2. 策略包会列出其建议的各项策略，每项都带有彩色标签标明其作用（脱敏、监控、阻止）并附有通俗说明。本策略包包含三项：

- **Redact Financial and Personal Data Before It Reaches an AI Provider**（脱敏）——在发往 AI 服务的内容中移除受分级保护的金融与个人数据。
- **Record Every AI Interaction**（监控）——记录与已识别 AI 服务的每一次交互，对应等级保护的审计留存要求。
- **Block AI Providers Outside the Approved List**（阻止）——阻止流向已批准清单之外的 AI 服务提供商。

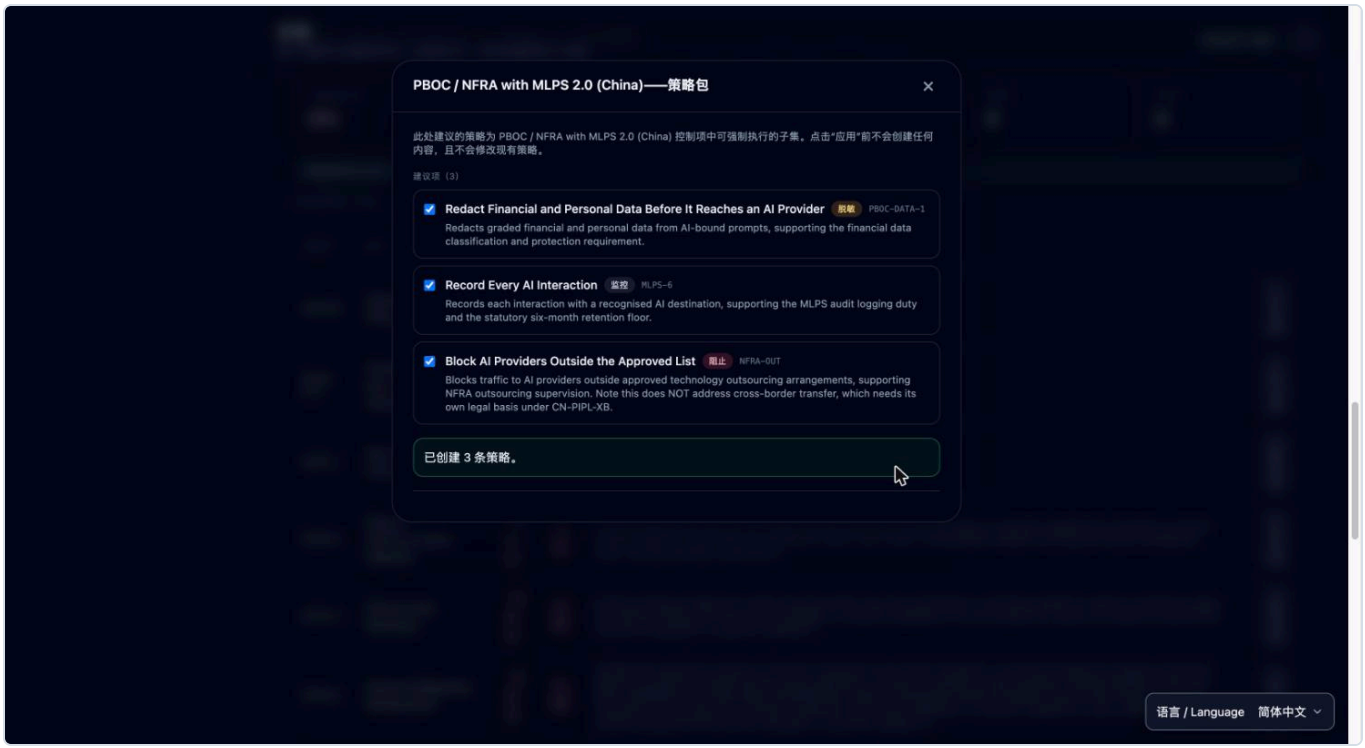
每一项建议的策略默认均已勾选。 请取消勾选您不需要的策略。若您希望先观察再拦截，可仅保留 **Record Every AI Interaction**：它只记录使用情况，不会中断任何人的工作，是一项稳妥的入门策略。然后点击**应用所选项**。

- 已从本策略包应用过的策略列在**已应用**下方，处于锁定状态。
- 若某项控制引用了尚未创建的清单（例如 `approved_providers` 已批准服务提供商清单），该项下方会出现一个提示框，其中的**创建此工件**按钮会打开**制品**页面并预先填好名称。在清单创建之前，该策略无法生效。
- 在您点击**应用所选项**之前不会创建任何内容，且您现有的策略绝不会被修改。

如果您保留了多条策略，请接着检查它们的优先级。 策略包创建的策略优先级一律为 **100**。详见第 6 节"优先级决定哪一条策略生效"。



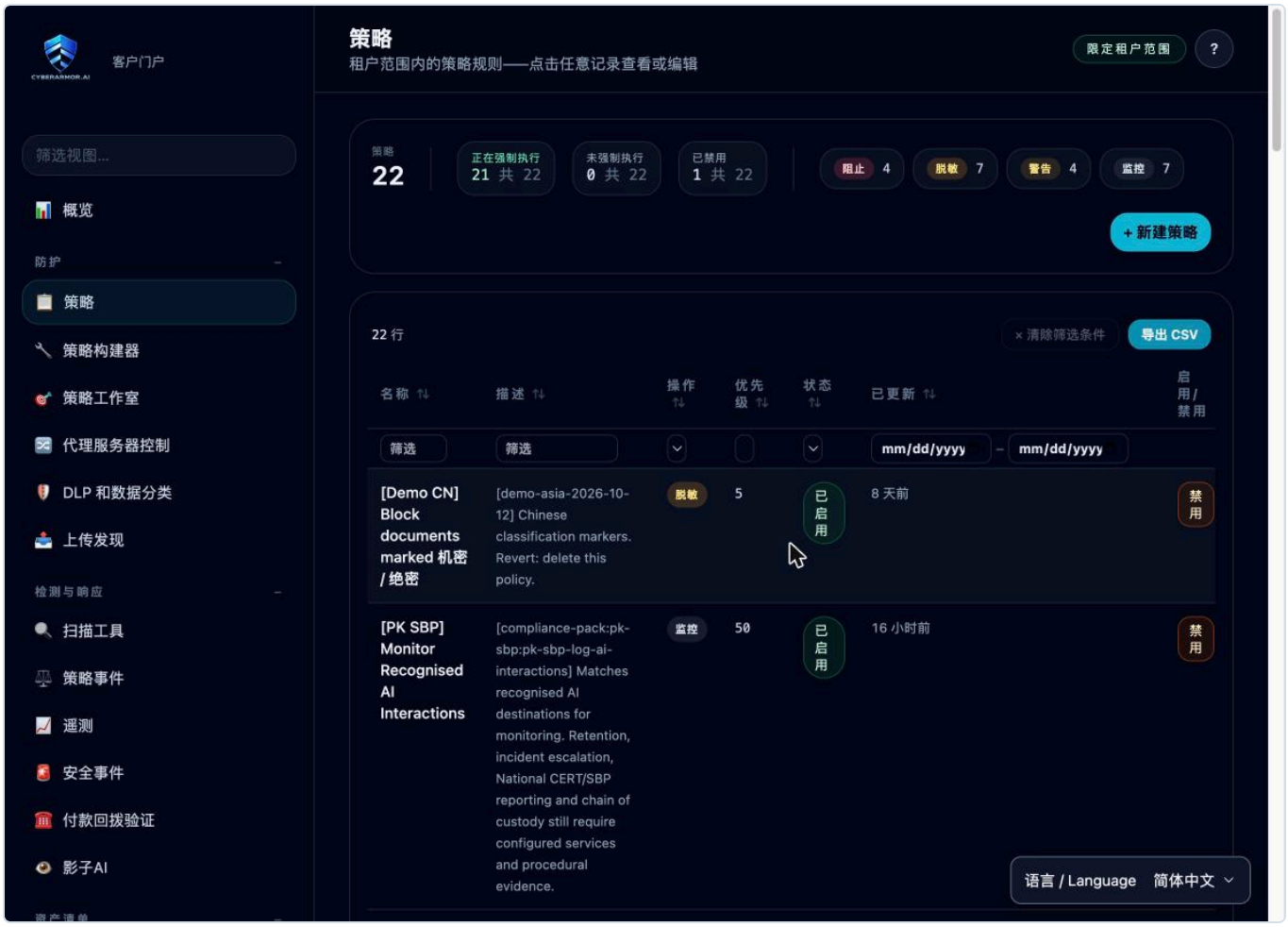
步骤 3. 门户会提示已创建的策略数量——本例中若保留全部三项，则为"已创建 3 条策略"——并重新运行该框架的评估，评分会在同一界面上更新。点击 **X** 关闭窗口。新建的策略会立即出现在**策略**页面中。



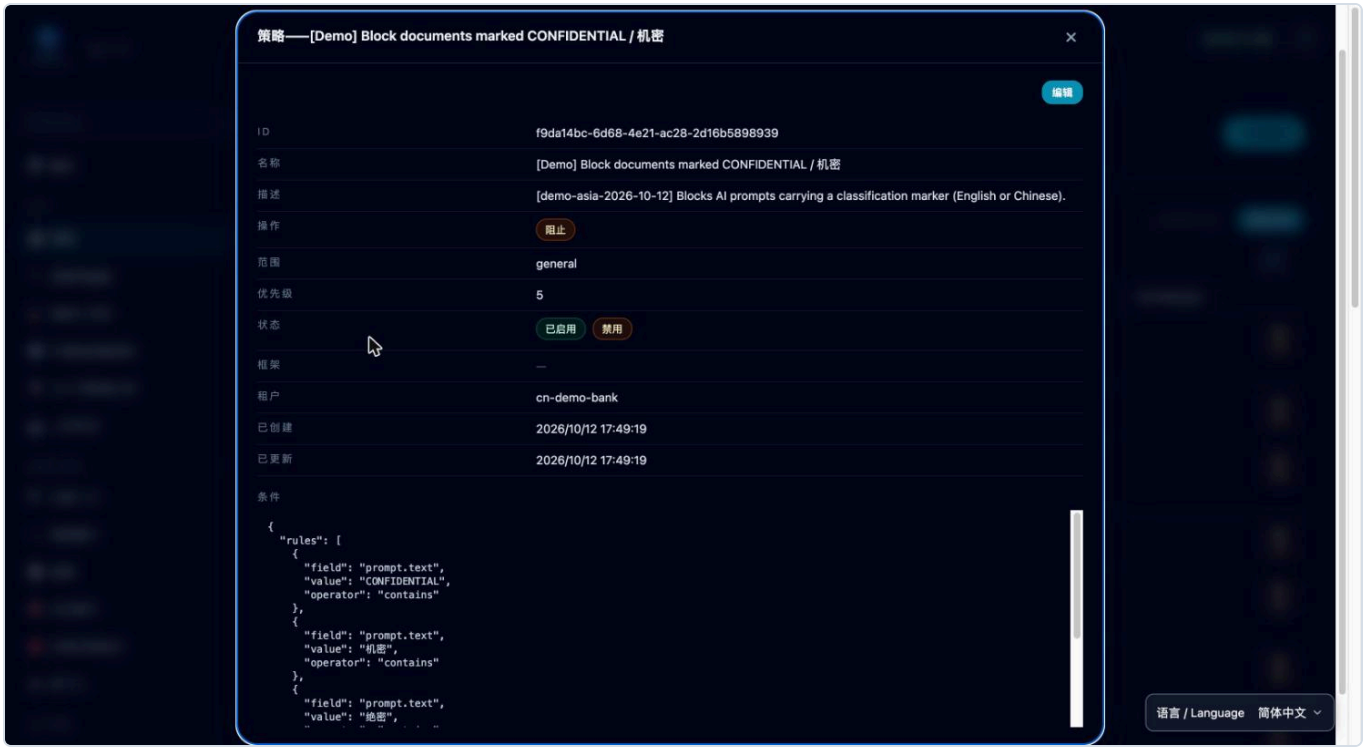
日后如需移除策略包，请再次打开**策略包**并点击**停用策略包**：该策略包创建的策略会被归档，而不是删除。

6. 修改策略

步骤 1. 在菜单中点击**防护** → **策略**。新建的策略已出现在列表中，名称带有其所属策略包的前缀：**[CHINA PBOC NFRA MLPS] Record Every AI Interaction**。点击该策略名称。



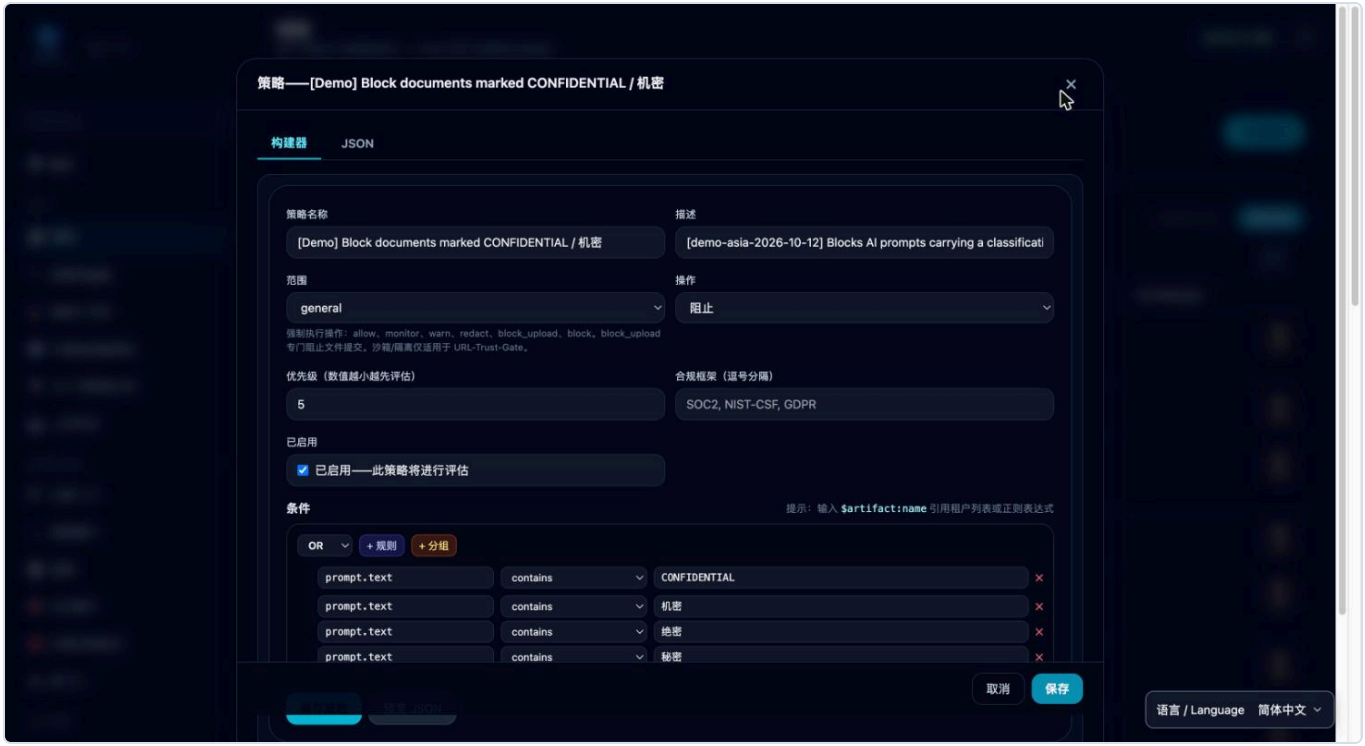
步骤 2. 策略详情随即打开，显示其全部设置：名称、描述、处置动作、范围、优先级、状态、合规框架及其条件。点击**编辑**。编辑模式有两个选项卡：**构建器**（表单）与 **JSON**（整条策略的文本形式）。



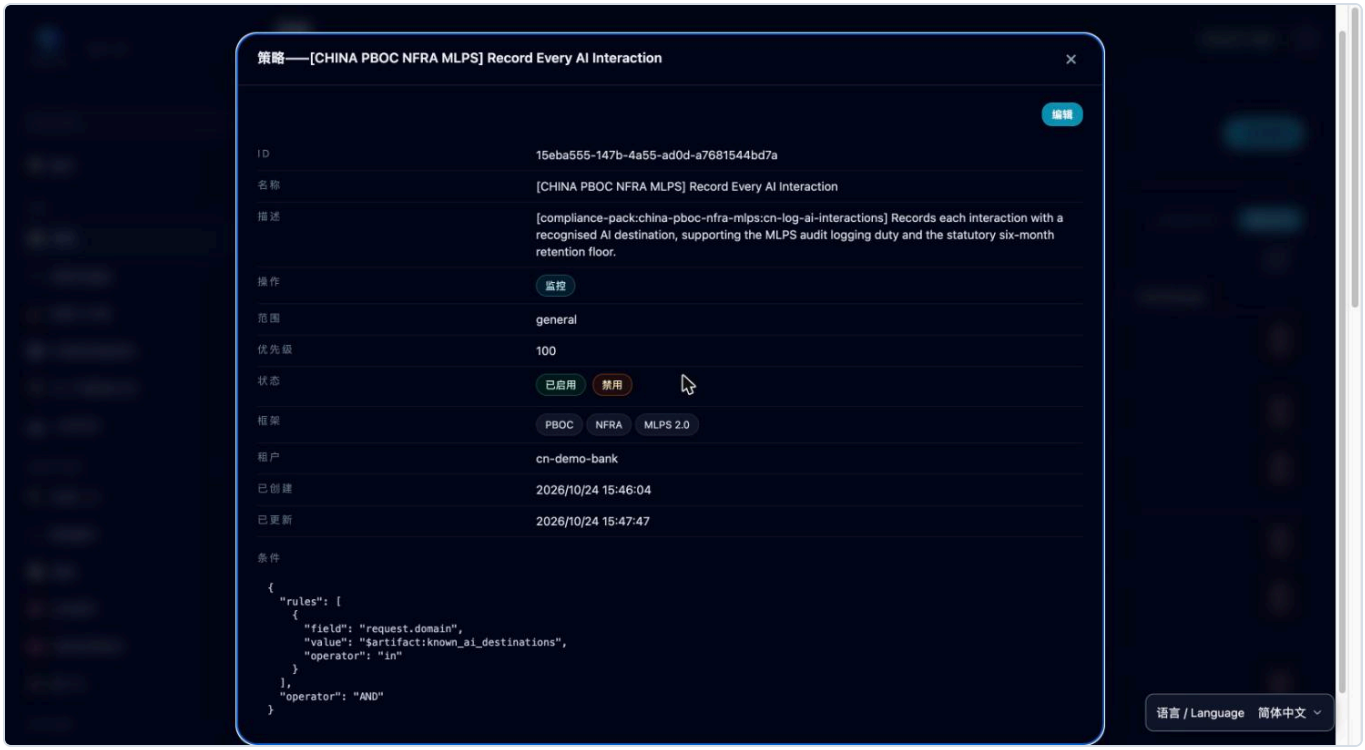
步骤 3. 修改您需要调整的内容。在本示例中我们修改了两项：

- **优先级**由 100 改为 **50**。数值越小的策略越先作出决定——详见下一小节。
- **合规框架**：我们填入了 **PBOC**、**NFRA** 与 **MLPS 2.0**，这样报告中就会显示该策略所支撑的法规。

点击**保存**。



步骤 4. 列表中现已显示优先级 **50**。设备会在下次与门户通信时自动获取该变更，您无需对设备进行任何操作。



优先级决定哪一条策略生效（请务必阅读）

优先级不只是"检查顺序"，它决定**哪一条策略真正生效**。

对于每一次请求，CyberArmor.AI 服务与终端代理会找出所有命中的策略。**命中的策略中，优先级数值最小的那些作出决定。**如果有多条命中的策略具有相同的最小数值，**则其中处置力度最强的动作生效**：block（及 block_upload）强于 warn，warn 强于 redact，redact 强于 monitor（及 audit-only），monitor 强于 allow。决策记录中会注明被搁置的那条策略，使这种并列情况清晰可见。

浏览器扩展对同一次请求的判定更为严格：无论优先级如何，它都会采用自己在该页面上能够执行的、处置力度最强的命中动作，并从优先级数值最小的命中脱敏策略中取得要脱敏的数据类别。

由此带来的后果是：**在服务端与终端代理上，数值更小的弱策略会压过数值更大的强策略**，但在浏览器中不会。若要在所有环节得到相同的结果，请为每一条策略设定各自的优先级数值，并把处置力度最强的动作放在最小的数值上。

策略包创建的策略，优先级一律为 100。 如果您从一个策略包中保留了多条策略（例如同时保留"记录每一次 AI 交互"与"阻止已批准清单之外的 AI 服务提供商"），二者并列时现在由阻止策略胜出；但仍建议在应用之后为它们重新排序，让顺序由您明确设定并有据可查：把处置力度**最强**的那一条调到**最小**的数值。同时请检查您自己的策略：您在优先级 50 设定的一条 *monitor* 策略，会先于策略包在 100 设定的 *block* 策略作出决定。

建议的排序方式（数值越小越先生效）：

优先级	处置动作	说明
8	`block`、`block_upload`	完全阻止、阻止上传——力度最强，应最先生效
85	`warn`	向用户显示警告
90	`redact`	发送前移除敏感内容
100	`monitor`、`audit-only`	仅记录——作为兜底，应最后生效

这些数值之间留有间隔是有意的：日后需要在两者之间插入新策略时，无需重排已有策略。您为特定业务单独创建的策略（例如优先级 5、6、10 的策略）仍然会在策略包的策略之前生效，这正是预期行为。

优先级为整数，默认值为 100。数值 `0` 虽然可以使用并且会排在 `1` 之前，但请不要使用——请为将来可能需要更高优先级的策略留出余量。

7. 创建自己的策略

步骤 1. 在**策略**页面点击 **+ 新建策略**。**策略构建器**随即打开。

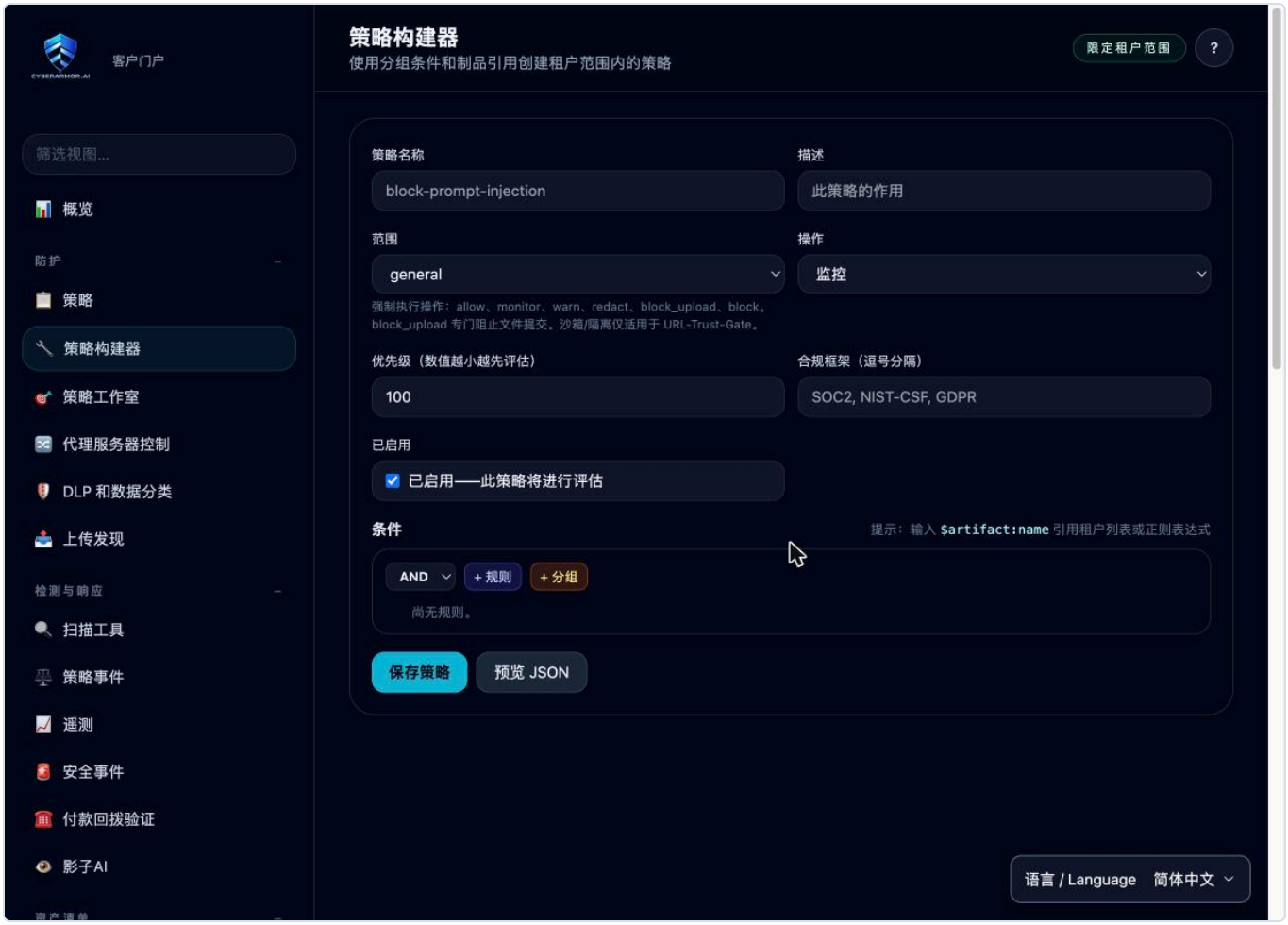
步骤 2. 填写以下内容：

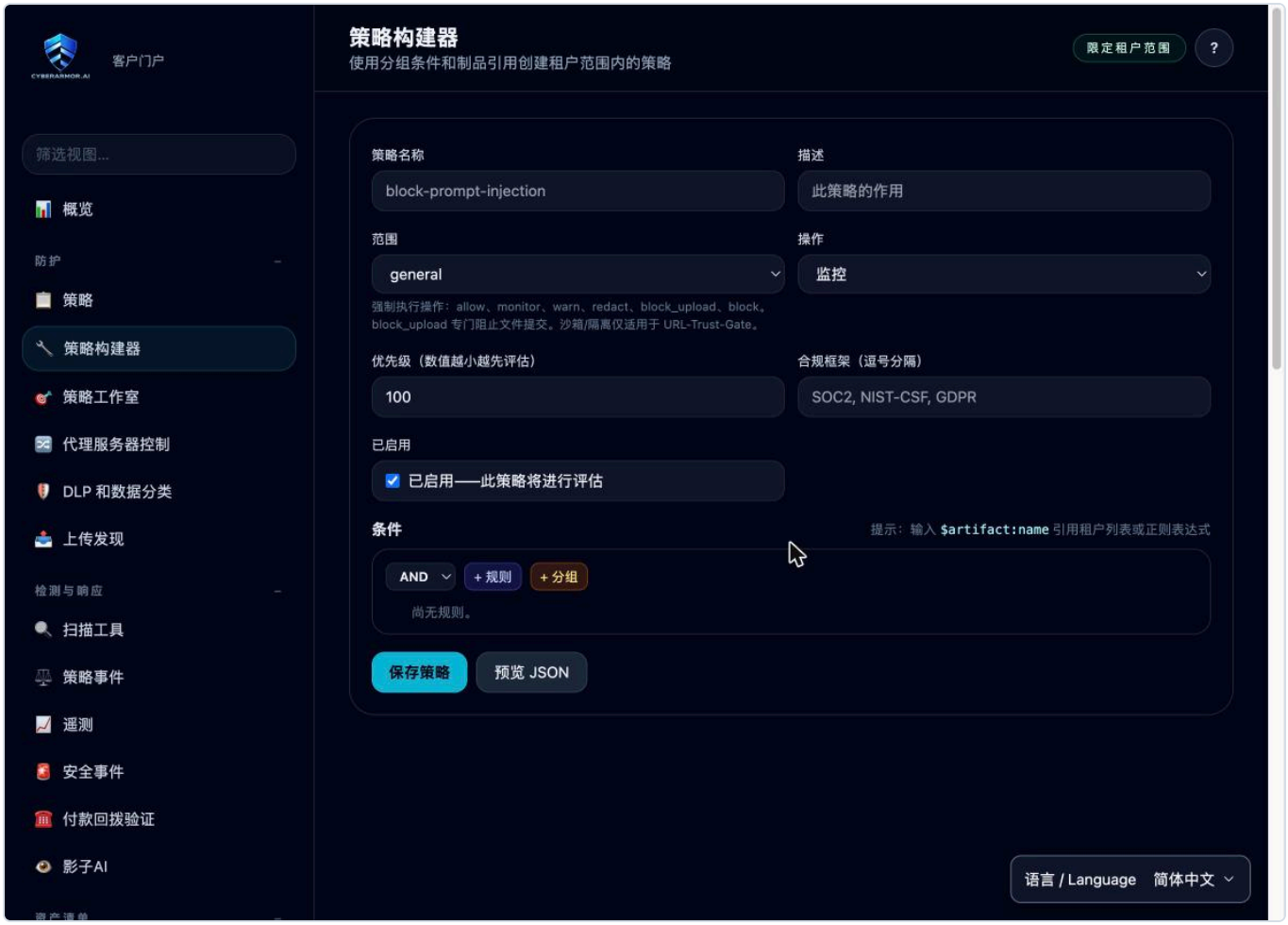
1. **策略名称。** 取一个所有人都能理解的名称。示例：*Demo – Monitor AI prompts containing a PRC resident ID number*（演示 – 监控包含居民身份证号码的 AI 提示词）。
2. **范围。** 策略的适用范围：`general`、`proxy`、`endpoint`、`identity` 或 `url-trust-gate`。除非技术团队另有要求，请保留 `general`；输入框下方的提示会说明每种范围执行哪些处置动作。
3. **操作。** 策略命中时执行的动作： - **monitor**：仅记录； - **warn**：向用户显示警告； - **redact**：在发送前移除敏感部分（在**需脱敏的 DLP 类别**中选择要脱敏的数据类别）； - **block_upload**：仅阻止文件上传； - **block**：完全阻止； - **allow**：放行。

建议先从 **monitor** 开始，在**策略事件**中查看结果后，再逐步加强处置力度。 4. **优先级。** 默认值为 100。若这条策略需要在策略包创建的那些策略之前生效，请填入更小的数值，并参照第 6 节"优先级决定哪一条策略生效"中的排序建议。 5. **+ 规则** 用于添加条件，**+ 分组** 用于添加以 AND、OR 或 NOT 组合的一组条件。我们的示例匹配包含中国居民身份证号码的 AI 提示词：字段 `prompt.pii_classes`，条件为 **contains**，值为 `pii.cn_resident_id`。在各输入框中开始输入时，门户会自动给出候选项（其他示例：`pii.cn_mobile` 对应中国手机号，`pii.credit_card` 对应银行卡号）。值也可以引用**制品**中的清单，写法为 `$artifact:` 加清单名称。

身份证号码只有在**校验位合法**时才会被判定为 `pii.cn_resident_id`。如果您想自行测试这条策略，请使用一个校验位正确的合成号码；随手编造的号码不会被检出。 6. 点击**保存策略**。

如果您的技术团队有需要，**预览 JSON** 可以文本形式显示完成后的策略内容。





示例：银行卡 PIN 码只能留在银行自己的网站上

借记卡 PIN 码只应在刷卡终端、ATM 上输入，至多是在银行自己的网站上。银行卡 PIN 码策略会在其他任何地方，对写在银行卡号旁边的 PIN 码进行脱敏或拦截。它由三部分组成：

1. **银行清单。** 打开**设置与管理 → 制品**。页面顶部的**列表库**提供现成的银行网站清单：**美国银行**、**美国信用合作社**、**中国的银行**和**巴基斯坦银行**，另有**银行卡组织**和**支付机构**。清单中的每个域名都已在该机构自己的网站或官方监管名单上核实，**查看来源**会显示出处。请对贵国的银行清单点击**添加到本组织**。添加清单绝不会删除您自己添加的条目；清单在被策略引用之前不起任何作用。清单有新域名时，按钮会显示为**添加新域名**。
2. **例外只限银行。** 只把银行清单设为 PIN 码可以去往的地方。切勿把**支付机构**或**银行卡组织**设为例外：支付机构、钱包或商户平台绝不会在网页上要求输入银行卡 PIN 码，而且任何人都可以在这类网站上发布付款页面，因此在那里输入 PIN 码本身就是一个危险信号。
3. **策略本身。** 处置动作为 **redact**（要脱敏的类别中包含 `pii.card_pin` 与 `pii.credit_card`）或 **block**，优先级低于您的其他脱敏策略，并以 AND 连接两条规则：文本中银行卡号旁边出现了银行卡 PIN 码，且 `request.domain` **not_in** 您的银行清单（例如 `$artifact:china-banks`）。CyberArmor.AI 通过一个脚本为组织写入这条策略并核验结果；请联系您的 CyberArmor.AI 联系人。

策略生效后，浏览器扩展会在 PIN 码被输入 AI 聊天时将其移除（脱敏）或阻止该消息（拦截），并暂扣银行卡号旁边带有 PIN 码的网页表单：拦截时不发送任何内容并显示策略名称；脱敏时清空 PIN 码、遮盖卡号并暂扣表单，让使用者看到原因。在银行自己的网站上，一切照常。

---

8. 启用或停用策略

策略页面的每一行都有一个**禁用**按钮（如果该策略已关闭，则显示为**启用**）。点击一次即可切换策略状态，且不会删除该策略。策略详情的**状态**一行中也有同样的开关。在第 6 节步骤 4 的图片中，我们的演示策略处于**已停用**状态，因此显示**启用**按钮。

9. 查看您的设备（终端）

资产清单 → **终端**列出了接入贵组织的每一台计算机、浏览器和手机。

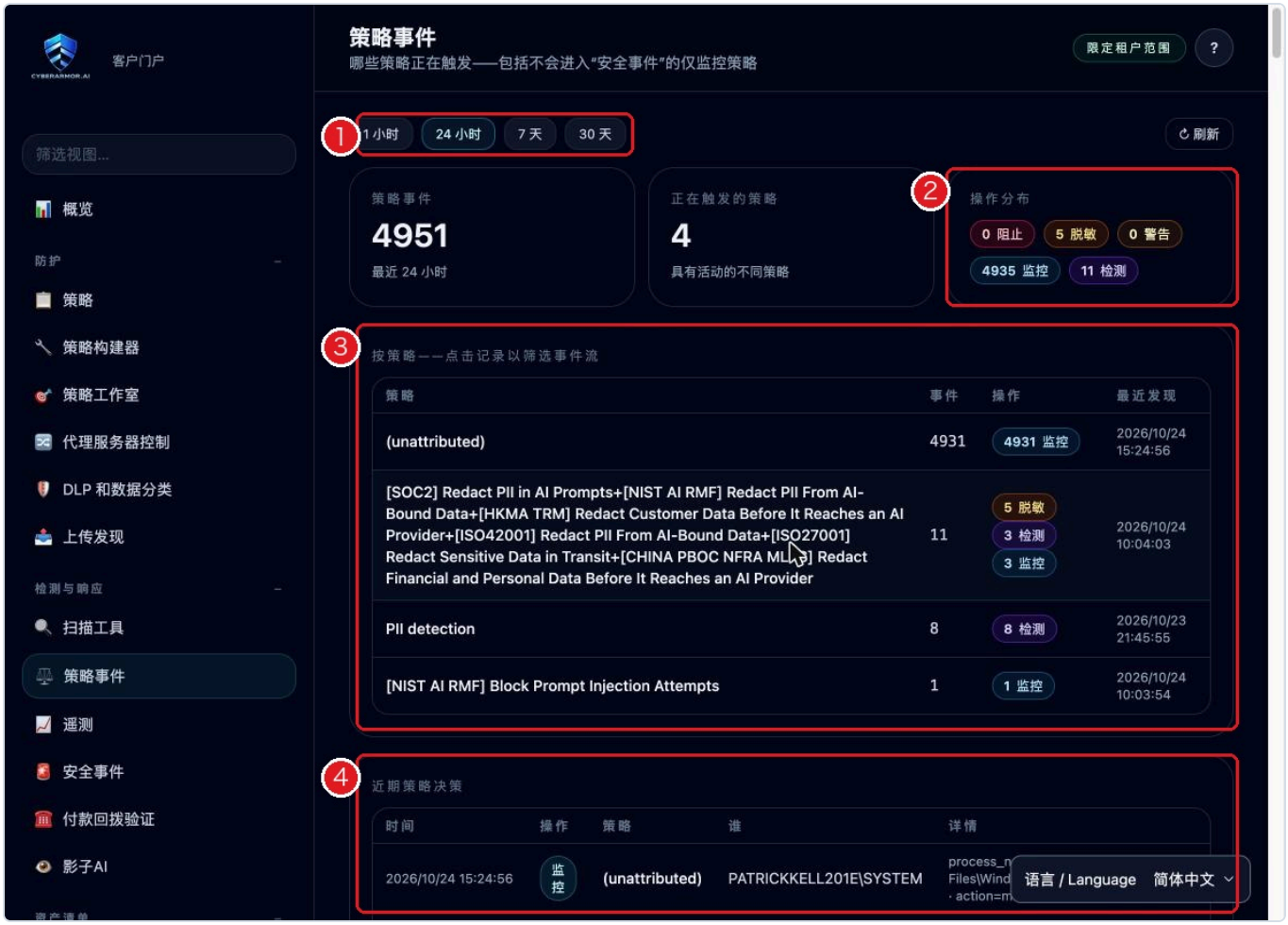


1. **汇总信息。** 显示有多少设备处于**健康**、**警告**、**陈旧**（近期未上报）或**离线**状态，有多少设备已强制启用 AI 流量检查，按平台统计的总数，以及需要关注的提示（例如某台设备的 AI 流量过滤器是为另一个组织设置的，或设备数量上限）。
2. **每台设备一行：** 显示设备名称、健康状况、监控状态、**AI 流量**检查是否已强制启用、其**过滤器所属组织**、最后在线时间、平台、用户、版本、最近 24 小时的事件数以及处于活动状态的监控项。
3. **健康**（绿色）表示该设备已受保护且上报正常。
4. **导出 CSV** 可将列表下载为电子表格。

点击某一行可查看设备详情：健康状况、凭据、AI 流量、最后心跳、操作系统、代理版本、事件、活动监控项、注册时间以及原始记录。租户管理员可在此**重命名**、**禁用**、**启用**、**停用**（Retire）或**归档**设备。重新安装的计算机留下的旧注册会单独列为**已被取代**，已归档的设备可恢复到列表中。

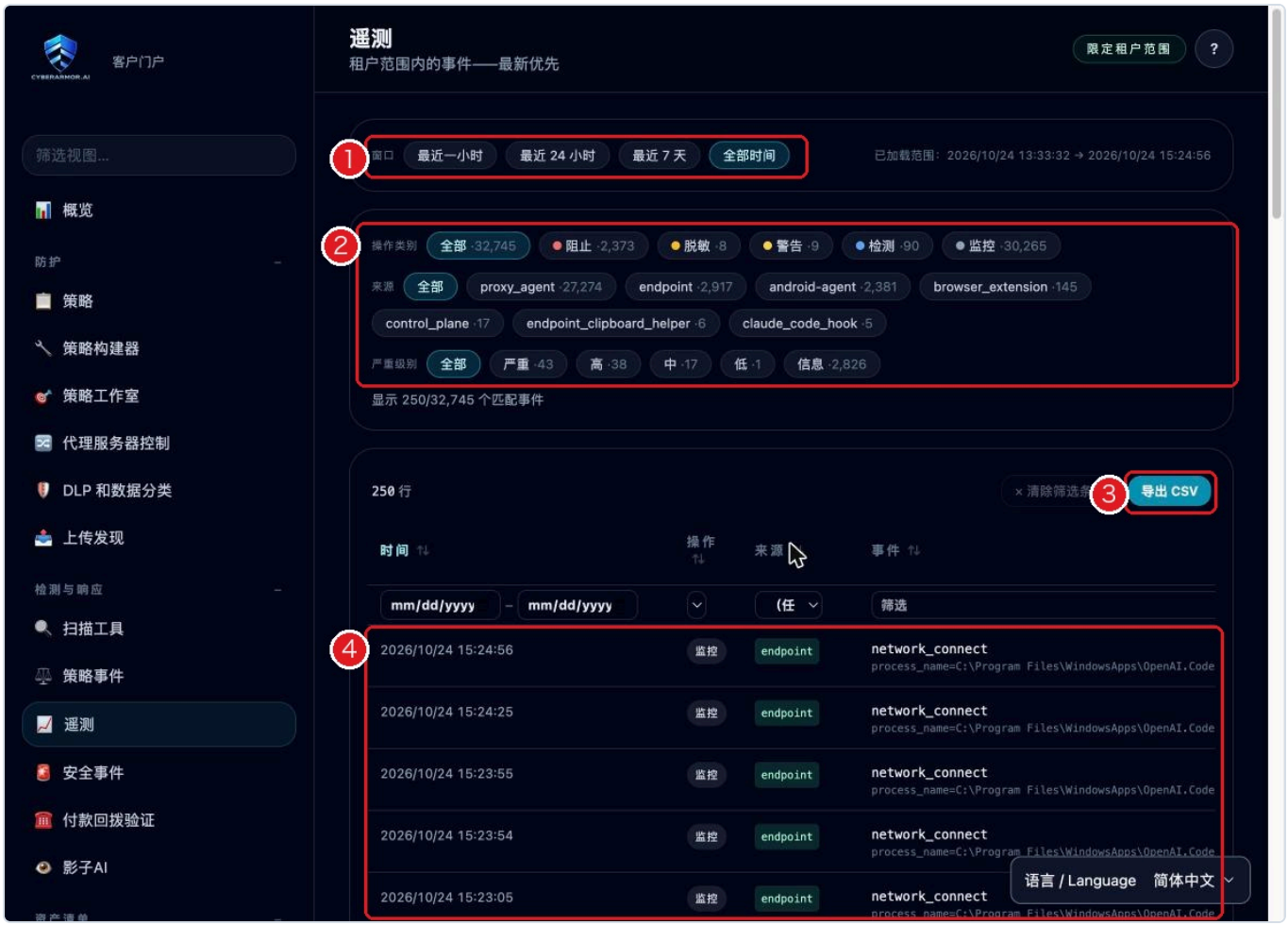
10. 查看运行状况

策略事件：哪些策略正在触发



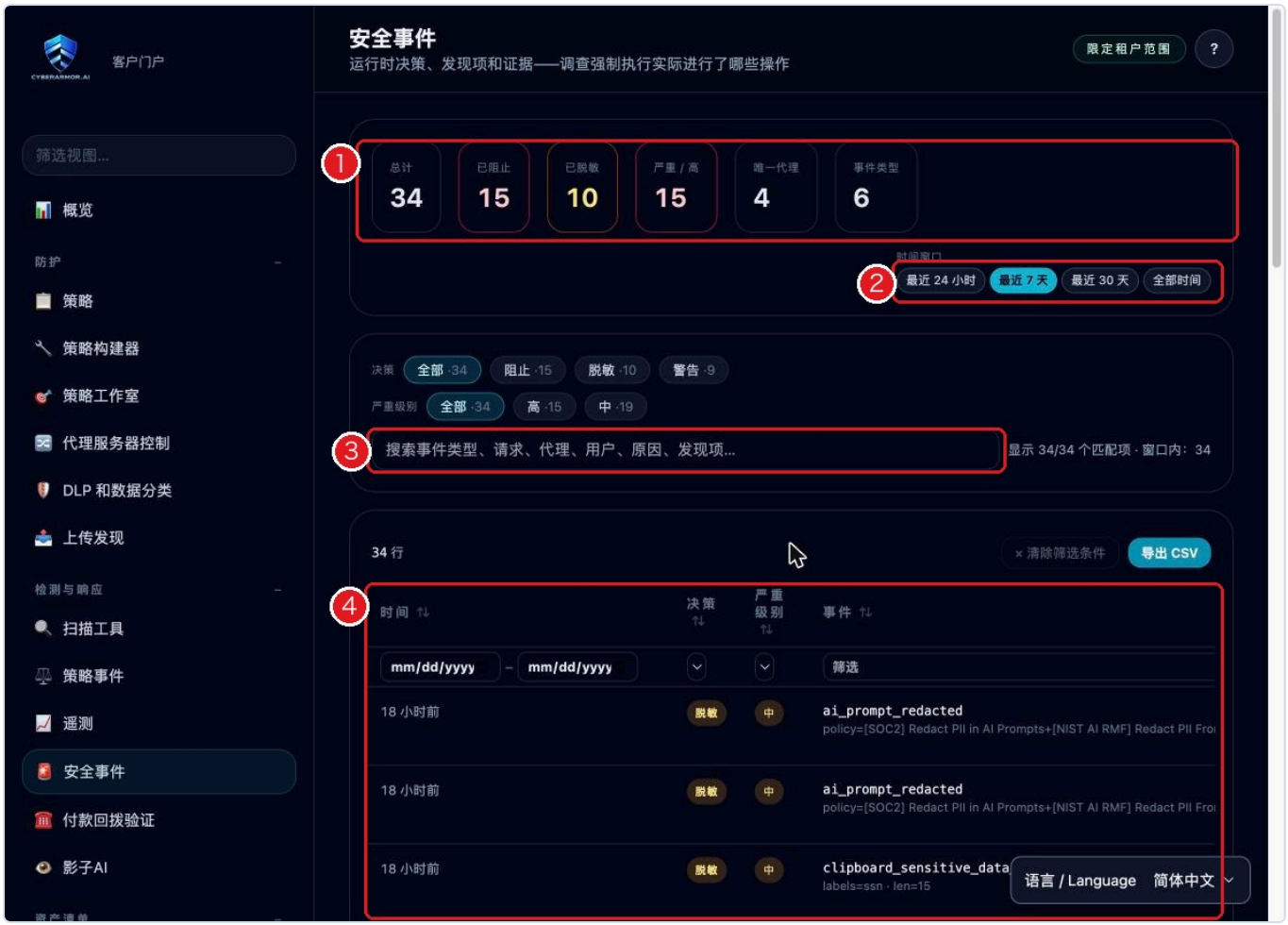
1. **时间范围：** 最近 1 小时、24 小时、7 天或 30 天，以及 **↻ 刷新**。
2. **策略事件、正在触发的策略、操作分布：** 显示有多少事件被拦截、脱敏、警告、检出或仅监控。
3. **按策略统计：** 显示各策略的触发频次。点击某一行可仅查看该策略的事件。
4. **近期策略决策：** 列出每一次决策，包含时间、操作、策略、用户和详细信息。点击某条决策可查看详情：策略名称及**打开策略**、内容扫描结果，以及该决策的上下文。

遥测：全部事件



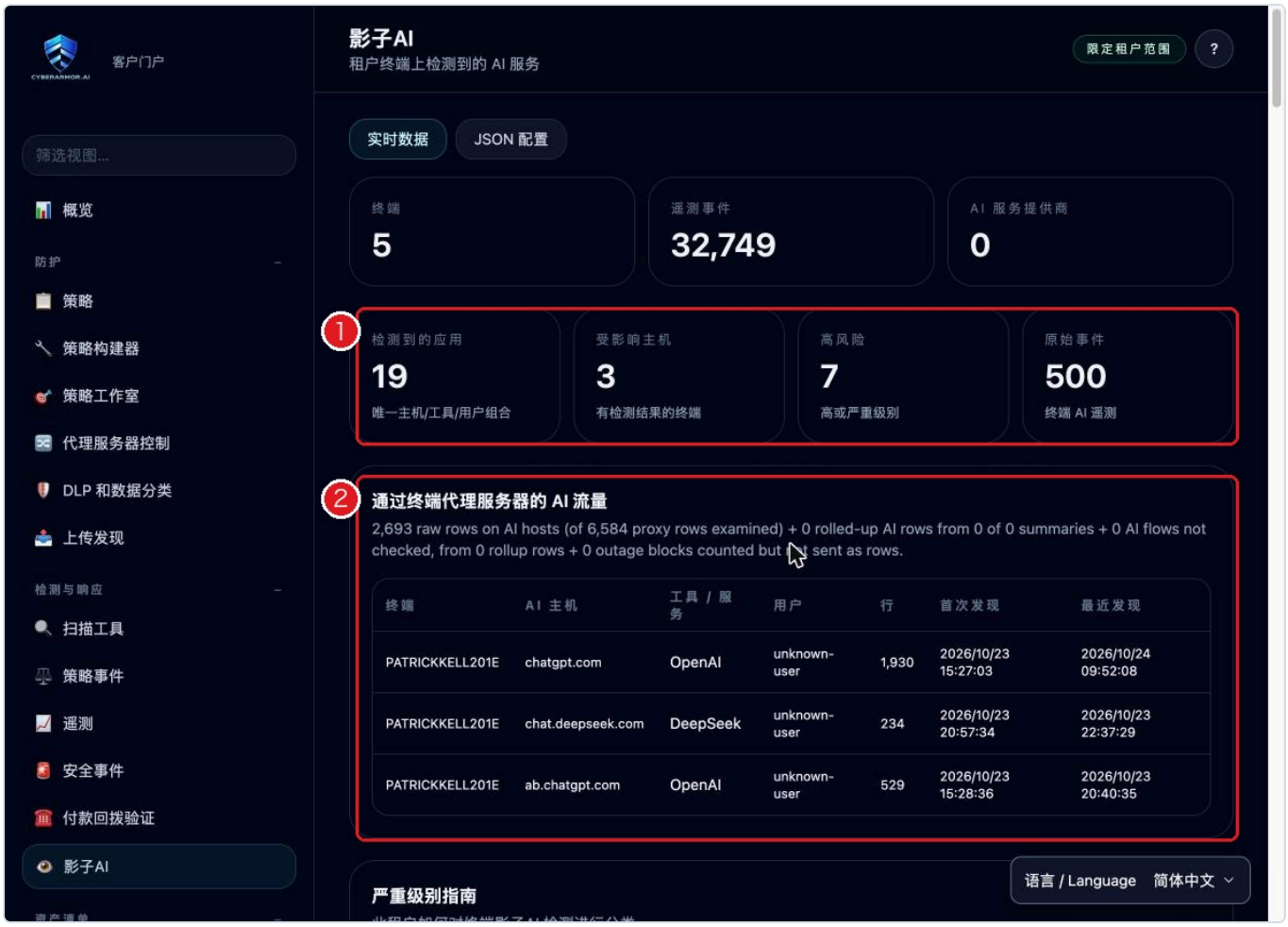
1. **时间窗口：** 最近一小时、24 小时、7 天，或全部时间。
2. **筛选器：** 点击可仅显示某一种操作（Block、Redact 等）、某一个来源（终端、浏览器扩展等）或某一种严重级别。
3. **导出 CSV** 可下载当前显示的内容。
4. **事件列表，按时间倒序排列。** 点击任意行可查看其完整详情，包括策略、内容扫描与关联信息（第 2 节）。

安全事件：哪些操作被阻止或被修改



1. **汇总数据：** 总数、已拦截、已脱敏、严重或高危、涉及的代理数量以及事件类型。
2. **时间范围**，以及按决定与严重级别的筛选。
3. **搜索**：可按事件类型、用户、设备或原因检索。
4. **事件列表。** 例如 `ai_prompt_redacted` 表示某条 AI 提示词中的个人数据在发送前已被移除。点击某条可查看详情：决定、原因、策略及**打开策略**、内容扫描、发现项、用户、设备与关联信息。

影子AI：正在使用的 AI 工具



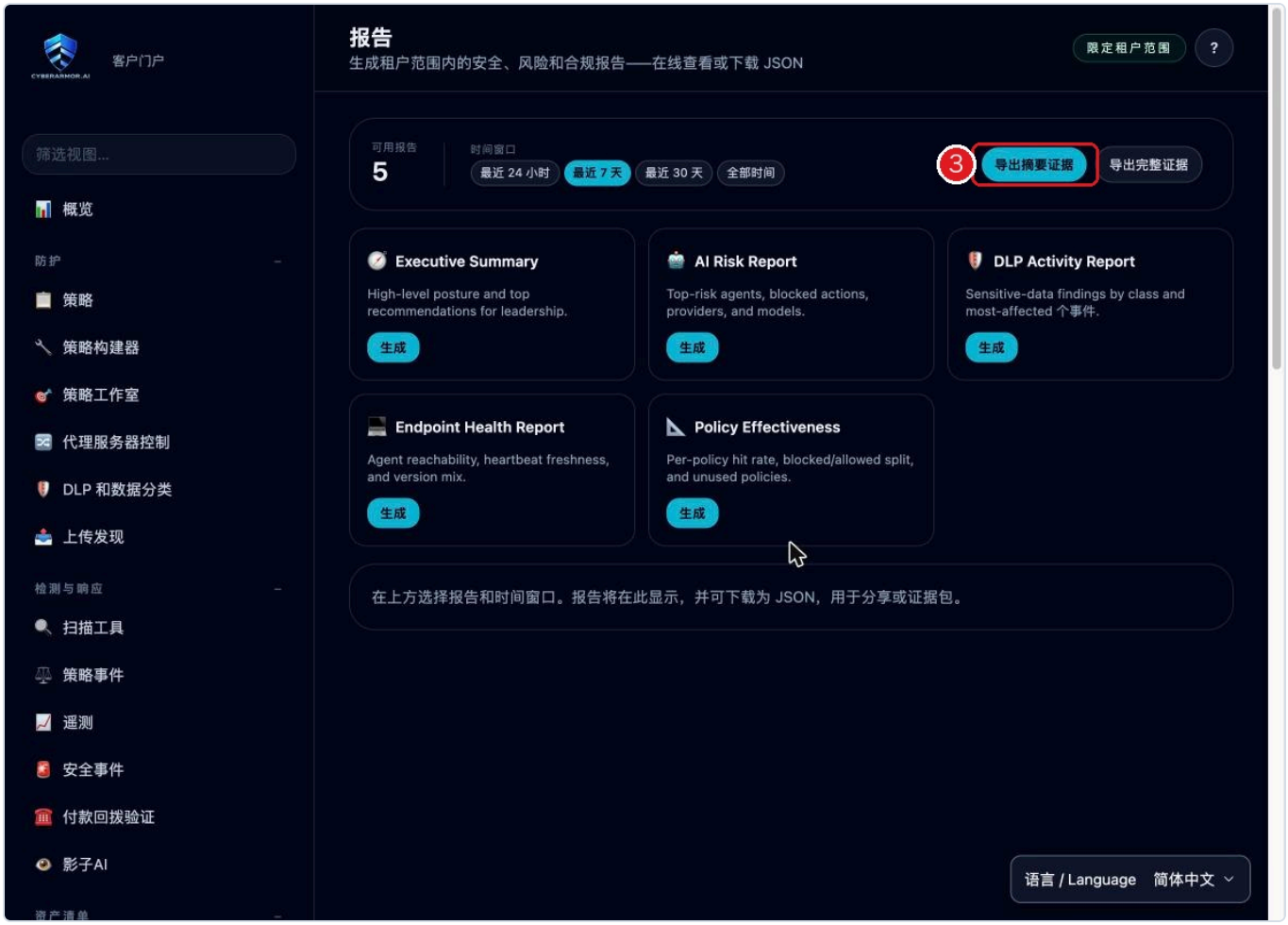
1. **检出的应用、受影响的主机、高风险发现项：** 对您设备上发现的 AI 服务与工具的汇总。
2. **通过终端代理的 AI 流量：** 显示哪台计算机使用了哪项 AI 服务（例如 ChatGPT 或 DeepSeek）、使用量，以及首次和最近一次出现的时间。

点击某条流量记录或某个检出项，可查看其详情及**处置操作**。窗口会显示该 AI 服务当前的设置（是否在已批准清单中、对应的策略、是否已审阅），租户管理员可直接在此处置：**批准**（加入已批准的 AI 清单；不写入任何策略）、**阻止**、**监控**、**警告**、**脱敏**（各自为该服务写入一条策略；监控、警告与脱敏策略会排在所有拦截策略之后，因此今天被拦截的内容不会因此被放行），或**标记为已审阅**（记录您的姓名与时间）。每项操作在更改之前都会先征求确认。**打开此主机的事件**与**打开策略事件**可跳转到相应记录。

---

11. 报告与审计日志

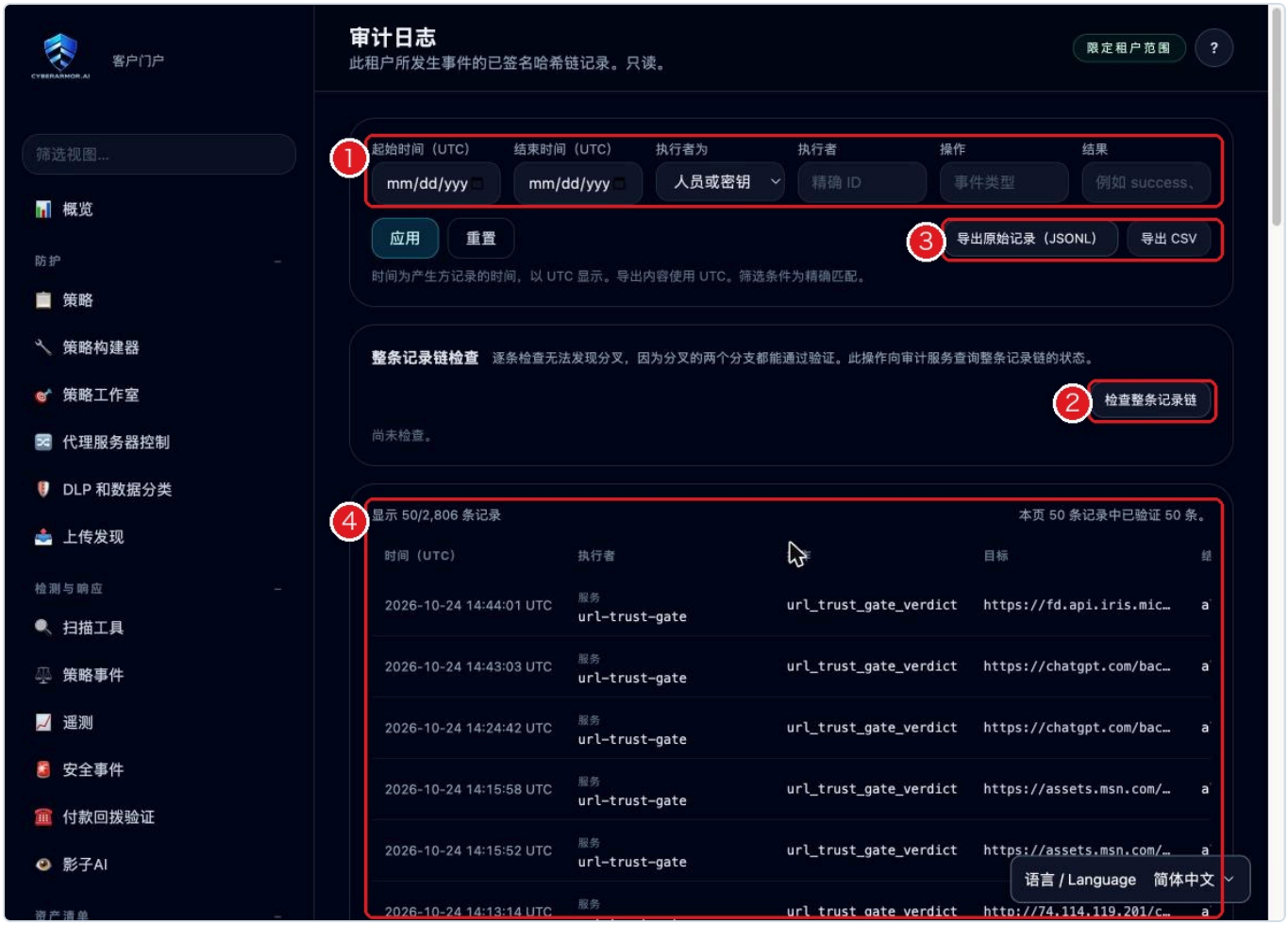
报告



1. 报告的**时间范围**：最近 24 小时、7 天、30 天，或全部时间。
2. 九种现成报告：**执行摘要**、**合规态势**、**安全事件与响应**、**AI 使用情况**、**AI 风险报告**、**DLP 活动报告**、**付款回拨验证**、**终端健康报告**和**策略有效性**。点击**生成**，报告会显示在页面上。**下载 PDF** 会打开浏览器的打印对话框：请选择**另存为 PDF**。**下载 JSON** 保存数据本身。
3. **导出摘要证据** / **导出完整证据** 可下载供审计机构与监管机构使用的证据包。

审计日志

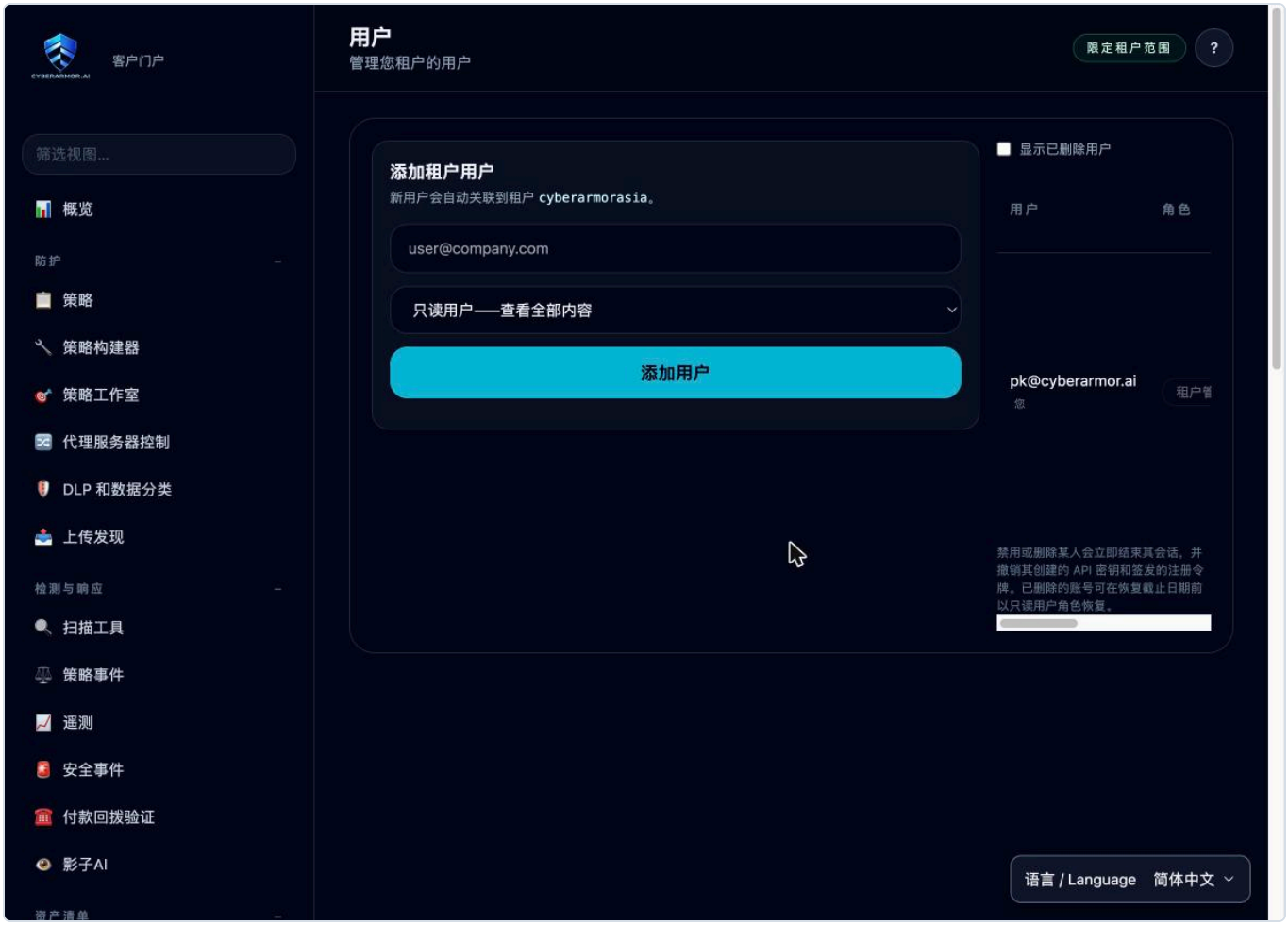
审计日志是贵组织内所发生操作的签名记录：谁登录过、哪些策略被修改、运行了哪些命令。每条记录都以 HMAC-SHA256 签名，并以哈希链与前一条记录相连；整条链会在检查点处以 ML-DSA-87 签名封存——这是一种后量子签名，任何持有公钥的人都可以验证。该日志为只读。



1. **筛选器：** 日期范围、操作者、操作类型和结果。点击**应用**。
2. **检查整条记录链** 会请审计服务验证每一条链及其签名检查点，并报告任何分叉。
3. **导出原始记录（JSONL）** 会下载审计服务所保存的原始签名记录，这就是证据文件。**导出 CSV** 生成电子表格。
4. **记录列表。** "本页 50 条记录中已校验 50 条"表示显示的每一条记录均通过了签名校验。点击某条记录可查看其全文及校验结果。

12. 添加同事（用户）

设置与管理 → 用户。 只有租户管理员可以添加用户。



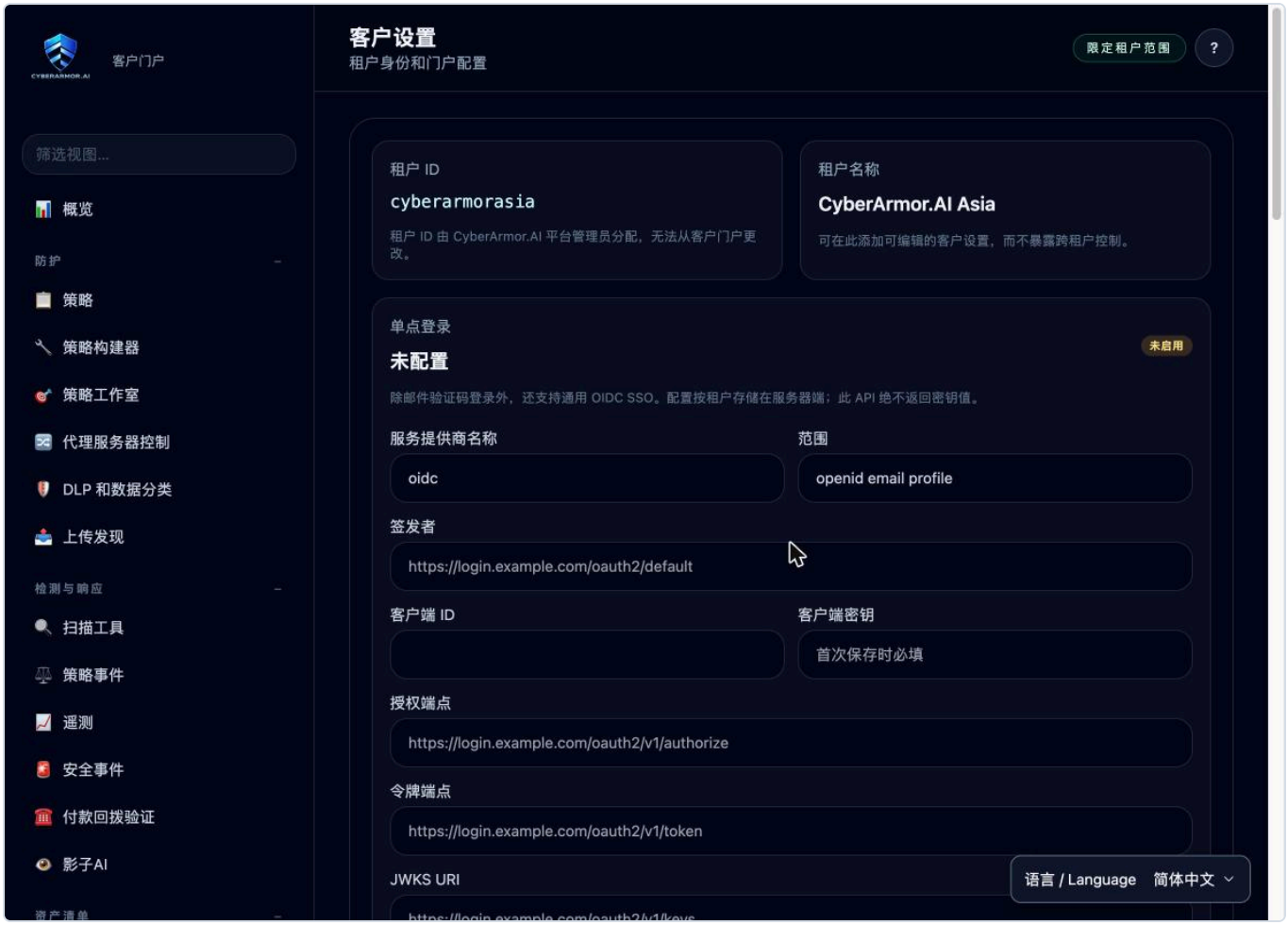
1. 输入该同事的工作邮箱。
2. 选择角色： - **Viewer**（查看者）： 可查看全部内容，但不能进行任何修改； - **Deployer**（部署员）： 在查看权限基础上，还可安装代理程序； - **Analyst**（分析员）： 在上述权限基础上，还可审批操作； - **Tenant admin**（租户管理员）： 拥有完整控制权限。
3. 点击**添加用户**。对方将收到邮件，并使用邮件验证码登录。

每一行都显示该人员的角色（租户管理员可直接修改，修改后立即保存），并提供**重命名**、**禁用**、**退出登录**、**重新发送邀请**、**删除**和**更改邮箱**。点击某一行可查看该人员的详情，包括是否已设置身份验证器应用（MFA）以及最近一次登录时间。

---

13. 登录安全（身份 / SSO）

设置与管理 → **身份 / SSO** 控制人员的登录方式。



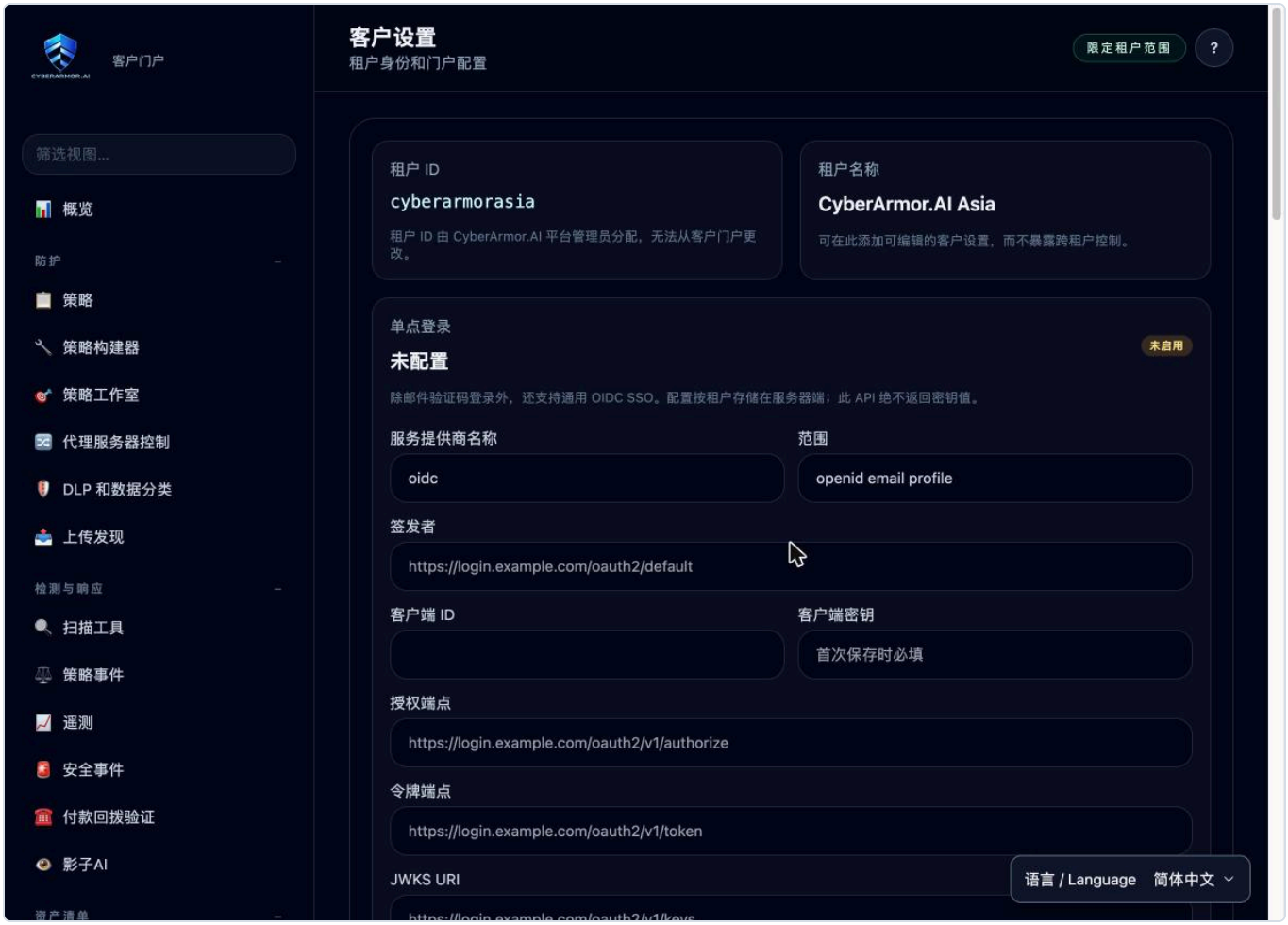
- **单点登录。** 租户管理员填写贵组织的 OIDC 设置（例如 Microsoft Entra ID 或 Okta），开启**为此租户启用 SSO**，并可选择让人员在首次单点登录时自动创建为查看者账户（**JIT 预配**；默认关闭）。点击**保存 SSO 配置**。邮件验证码登录方式仍然可用。
- **多因素认证。** 租户管理员勾选**为此租户用户提供 MFA**。这一设置只是让 MFA 可用，并不会为任何人开启。随后每位人员需在本页面自行设置身份验证器应用：点击**配置身份验证器应用**，用 Microsoft Authenticator 或 Google Authenticator 等应用扫描二维码，输入 6 位验证码，并保存仅显示一次的备用码。此后该人员登录时都会被要求输入应用中的验证码。关闭该设置会停止对所有人的验证要求，但会保留各人的注册信息，重新开启后即恢复。
- **在所有位置退出登录** 会结束您本人账户在本组织中的所有会话，包括当前会话——例如在使用过共享或丢失的计算机之后。

该勾选框在更改时立即保存。

---

14. 适用于所有设备的设置

设置与管理 → **设置**显示贵组织的 ID 与名称（由 CyberArmor.AI 设定，无法在此修改），以及适用于所有设备的设置。只有租户管理员可以更改，其他人可以查看。登录相关设置位于**身份 / SSO**（第 13 节）。



自动终端响应

允许终端代理在检测到破坏性行为时不等人工介入而直接处置。**新建组织默认关闭此功能。**关闭时，检测结果仍会上报，但不会采取任何处置。

需要做出两个相互独立的选择，二者都设定后才会有处置发生：

1. **允许此租户中的代理自动响应**（总开关）。
2. 为每种检测选择处置动作：**疑似勒索软件**与**批量删除**。二者初始均为**仅报告**，另一个选项是**终止进程**。

点击**保存响应设置**。保存为**仅报告**以外的处置动作时，系统会请您确认，并列出将被启用的内容。

终止进程只会作用于计算机自身传感器观察到执行了该变更的进程（macOS 上为 Endpoint Security，Windows 上为 ETW），且仅在它仍是同一个进程时执行。Linux 计算机，以及传感器无法归因的检测，只会上报。勒索软件终止进程功能在能力状态页面（https://cyberarmor.ai/status）上处于**可配置**（Configurable）级别：它已在一台 Windows 11 测试计算机上运行过，在 Mac 上已构建但尚未在硬件上验证，且尚未在客户计算机上运行。仅靠检测并不能阻止加密，已被修改的文件仍保持修改后的状态。目前不提供将计算机与网络隔离的选项：因为还无法远程解除一台被隔离的计算机。

私有网络地址与本地页面检查

- **私有网络地址（浏览器扩展）。** 当有人通过私有 IP 地址打开路由器、打印机、NAS 或内网应用时，浏览器扩展的处置方式：**允许（默认）**、**警告**（用户可以继续）或**阻止**（用户无法继续）。无论选择哪一项，扩展都绝不会把这些地址发送给 CyberArmor.AI。选择后立即保存；浏览器会在下次同步策略时应用。
- **先在设备上检查网页（本地 Trust Gate）。** 默认开启。在安装了 CyberArmor.AI 代理及其模型包的计算机上，页面会先在本机检查，只有有风险或无法读取的页面才会发送到服务器。关闭时，所有页面都在服务器上检查。更改时立即保存。

本地模型

在每台计算机上运行 CyberArmor.AI 的检测模型——使用为每台设备单独加密的签名模型包——而不是把内容发送到服务器进行检查。支持搭载 Apple 芯片的 Mac（macOS 14 或更高版本）和 Windows；其他设备仍像以前一样由远程判定。本地模型在能力状态页面上处于**可配置**（Configurable）级别。

新建组织的本地模型默认关闭——状态显示为 **off (not yet rolled out)**（关闭，尚未推送）——除非 CyberArmor.AI 的分阶段推送已覆盖到该组织。**要开启本地模型**，请由租户管理员选择**开启**（On）并点击**保存**。可选项包括：

- **跟随 CyberArmor.AI 发布**（会显示该发布当前为开启还是关闭）、**开启**或**关闭**。管理员选择的开启或关闭会一直保持，直到管理员再次更改。
- **包含 PHI 模型（医疗）**。
- **模型集合： 标准**（提示词注入、PII、有害内容、输出安全）或**完整**（全部五个模型），并显示各自的下载大小。
- **软件包版本： 最新发布版本**，或固定某个版本。
- **更新计划**（按各设备的本地时间）： **每日非工作时间**——时间窗口初始为 **02:00 至 05:00**；**租户计划**（仅在勾选的日期、在时间窗口内）；或**仅手动**。设备每天检查一次，只下载有变化的部分；在时间窗口内一直合着盖的笔记本电脑会在 36 小时内补上更新。
- **立即推送更新（所有设备）**会要求所有设备立即更新；设备所在行的**推送**按钮只针对该设备。

标题行显示"M 台符合条件的设备中有 N 台使用模型包 vX"。尚未上报的设备会显示为**未知——未报告**，绝不会显示为最新；不计入统计的设备（手机、平板、旧的重装注册）列在**不计入 M** 之下。管理员关闭本地模型后，每台设备会在下一次每日检查时删除其模型包，检查工作改回由服务器完成，与以前相同。

---

15. 获取帮助与培训视频

支持 → 帮助与常见问题提供：

- **参加 2 分钟游览（▶ 运行引导游览）**；
- **资源： 文档**、**支持门户**和**接入与下载**；
- **培训视频：** 面向客户的八个视频——《安装：从登录到受保护的计算机》《防护实战》，以及《客户门户培训》第 1 至 6 章。**观看**会在新标签页中播放视频；**章节**会打开该视频在文档站点上的页面及章节列表。**全部视频及其章节**会打开完整列表；
- **常见问题**，例如*代理程序与浏览器扩展是如何接入我的租户的？*以及"*脱敏*"*具体会做什么？*

同样的视频也位于 **https://docs.cyberarmor.ai/zh/training/** 和 **https://cyberarmor.ai/training**。视频为英语（巴基斯坦）版，在演示设备上录制。本指南及其配套指南位于文档站点的**用户指南**栏目，提供网页版以及 PDF 和 Word 下载。



16. 快速参考：全部菜单项

菜单分组	页面	用途
—	**概览**	任务控制台仪表盘（第 3 节）。
防护	**策略**	查看列表、打开、编辑、启用和停用策略（第 6、8 节）。
防护	**策略构建器**	通过表单创建策略（第 7 节）；优先级字段请参见第 6 节。
防护	**策略工作室**	以 AI 决策的形式查看策略，每条策略均可**测试**（第 17 节）。
防护	**代理服务器控制**	故障模式、绝不解密的站点，以及 AI 流量检查所使用的策略（第 17 节）。
防护	**DLP 和数据分类**	数据分类标签、自定义规则、检测模式以及实时扫描（第 17 节）。
防护	**上传发现**	浏览器扩展发现的、尚未收录到目录中的上传类网站；将其添加后，上传拦截即可覆盖这些站点。
检测与响应	**扫描工具**	粘贴任意文本，查看 CyberArmor.AI 能从中检测出哪些内容。
检测与响应	**策略事件**	查看哪些策略被触发以及触发频次（第 10 节）。
检测与响应	**遥测**	来自所有设备的全部事件（第 10 节）。
检测与响应	**安全事件**	哪些操作被拦截、脱敏或标记（第 10 节）。
检测与响应	**付款回拨验证**	付款须经人工回拨登记号码确认后方可放行（第 17 节）。
检测与响应	**影子AI**	正在使用的 AI 工具与服务，以及对每一项的处置（第 10 节）。

菜单分组	页面	用途
资产清单	**终端**	您已接入的设备及其健康状况（第 9 节）。
资产清单	**补丁管理**	保持设备上的软件为最新版本，支持审批流程与维护窗口（付费附加功能）。
资产清单	**物料清单**	在设备上发现的 AI 模型、工具与程序库，含变更、漏洞与 IOC 扫描。
资产清单	**代理程序目录**	在您的审计记录与活动中出现过的 AI 代理身份，以及您的设备。
资产清单	**AI 服务提供商**	贵组织自有应用所使用的 AI 服务（OpenAI、Anthropic 等）的凭证。
资产清单	**委托管理**	查看每位人员向哪个 AI 代理授予了哪些委托权限：可签发、限制和撤销。
证据	**合规**	框架评估与策略包（第 5 节）。
证据	**审计日志**	经签名的只读记录（第 11 节）。
证据	**API 访问日志**	对 CyberArmor.AI 服务器发起的每一次调用（未签名）。
证据	**报告**	现成报告（PDF 与 JSON）与证据导出（第 11 节）。
证据	**操作图谱**	将每个 AI 代理与其使用过的 AI 服务、模型和工具相关联的链路图。
证据	**AI 风险仪表盘**	整体风险评分、风险最高的代理及改进建议。
设置与管理	**接入与下载**	安装程序与安装令牌（第 4 节）。
设置与管理	**API 密钥**	用于对接贵组织自有系统的密钥。
设置与管理	**SIEM 配置**	将事件发送至您的安全监控系统（Splunk、Sentinel 等）。
设置与管理	**身份 / SSO**	单点登录、多因素认证、在所有位置退出登录（第 13 节）。
设置与管理	**制品**	策略规则中可复用的清单与模式，以及银行与支付清单的列表库（第 7 节）。
设置与管理	**用户**	添加和管理同事（第 12 节）。
设置与管理	**设置**	自动终端响应、私有网络地址、本地页面检查与本地模型（第 14 节）。
支持	**帮助与常见问题**	引导游览、培训视频与常见问题（第 15 节）。

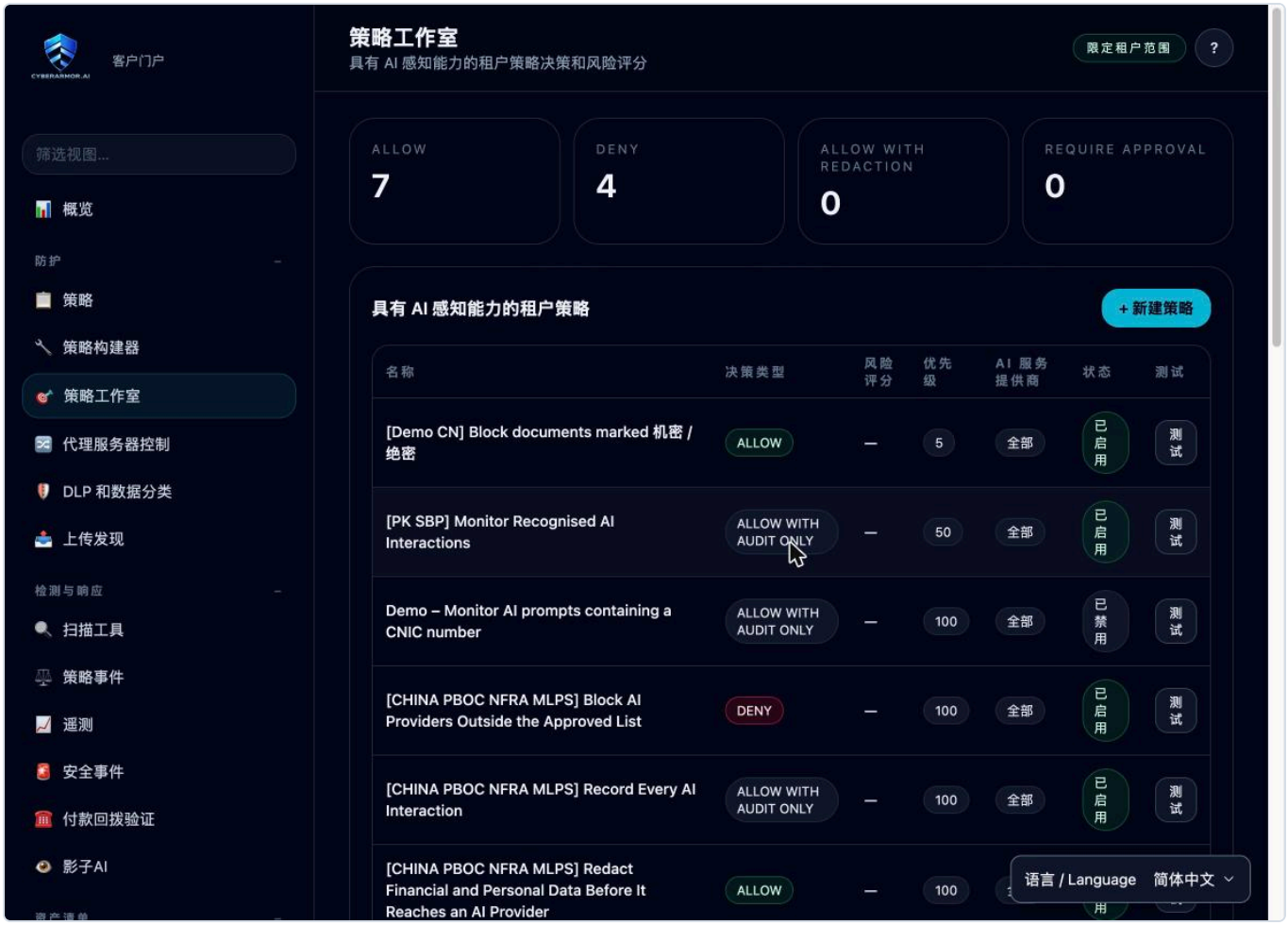
---

17. 逐一认识其他视图

本节介绍第 1–16 节未逐步讲解的页面。每个条目先说明**该视图的用途**，再给出界面截图。这些页面上的每个列表同样点击即可打开详情（第 2 节）。

防护

策略工作室（Policy Studio）——以 AI 决策的形式查看全部策略：**Allow**、**Deny**、**Allow with redaction** 与 **Require user decision**，并以表格列出面向 AI 的策略。点击某行的**测试**会打开**评估策略**面板：选择服务提供商和模型，输入一段测试提示词并点击**评估**。它会显示决定、原因、作出决定的策略、检查过的策略以及任何问题，且不会更改任何内容。



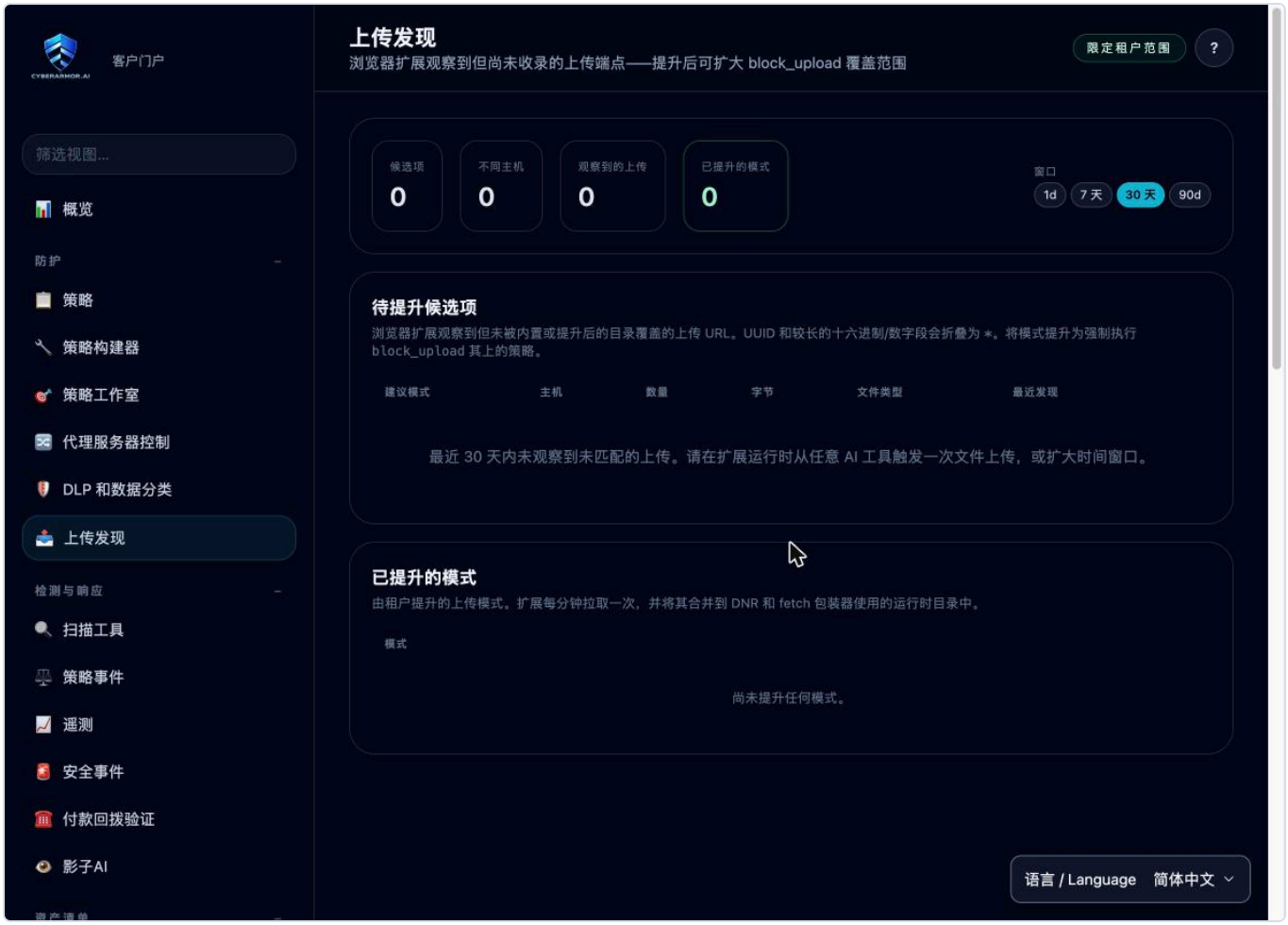
代理服务器控制（Proxy Controls）——设定当 CyberArmor.AI 无法检查时计算机如何处置 AI 流量（**故障时关闭**或**故障时开放**；选择后立即保存），维护加密流量绝不解密的主机清单（TLS 拦截绕过），以及 AI 流量检查所使用的策略，可勾选**显示已归档项**。



DLP 和数据分类（DLP & Data Classification）——四个数据分类标签（公开、内部、机密、受限），您的**自定义分类规则**（第一条命中的规则生效），**DLP 检测模式**及其中被脱敏策略引用的那些，以及**扫描并分类内容**：粘贴文本并点击**分类**即可查看标签与发现项。



上传发现（Upload Discovery）——浏览器扩展观察到、但尚未纳入内置目录的上传地址，以及上传次数、文件类型和使用者。租户管理员点击**提升**即可将其纳入，使上传拦截覆盖该地址。



检测与响应

扫描工具（Scan Tools）——对粘贴的文本运行四种扫描：**提示词注入扫描**、**敏感数据扫描**、**输出安全扫描**与**完整检测流水线**。每种扫描都会显示发现项和 ▸ **原始 JSON** 展开项。适合在制定策略前查看策略将会"看到"什么。



付款回拨验证（Wire Verification）——超过贵公司所设阈值的付款会被暂停，直到有人拨打贵公司自己的**登记目录**中的号码回拨确认对方（绝不使用随付款指令提供的号码），并由贵公司要求数量的人员确认——发起付款的人绝不能作为确认人。队列显示每笔被暂停的付款、谁可以放行以及无人处理时的后果；**我正在拨打电话**、**已在通话中验证**和**已拒绝**用于记录回拨结果。CyberArmor.AI 不持有任何资金，也不监控贵公司的付款系统：持有付款的系统必须把每笔付款发送给 CyberArmor.AI，在完成这一项对接之前，此视图不会显示任何内容。付款回拨验证在能力状态页面上处于**试点就绪**（Pilot-ready）级别。

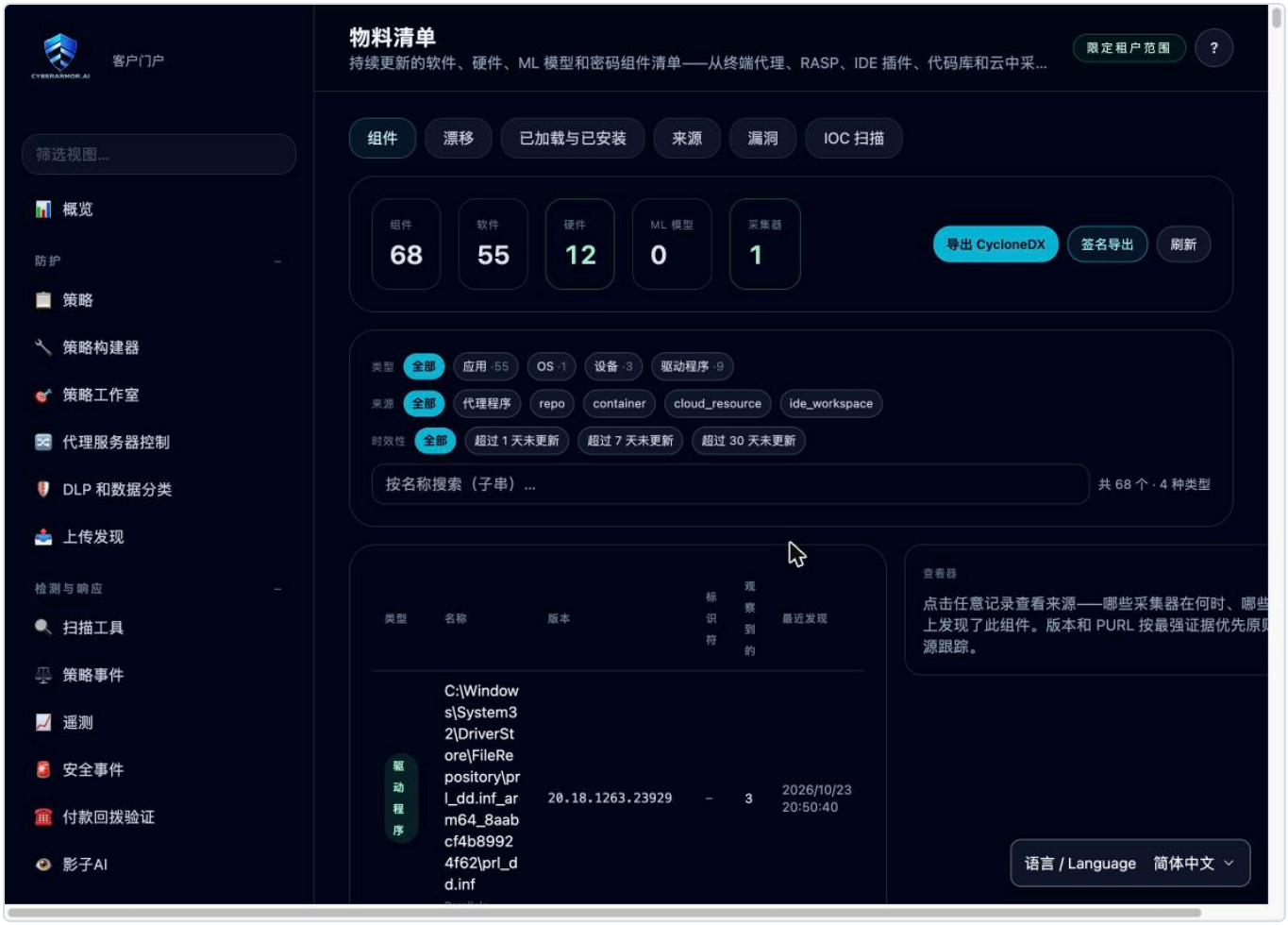


资产清单

补丁管理（Patch Management）——通过审批、维护窗口和按应用自动批准，保持设备上的软件为最新版本（winget、Homebrew、apt、yum/dnf）。此为付费附加功能：未开通时，页面会说明尚未启用。



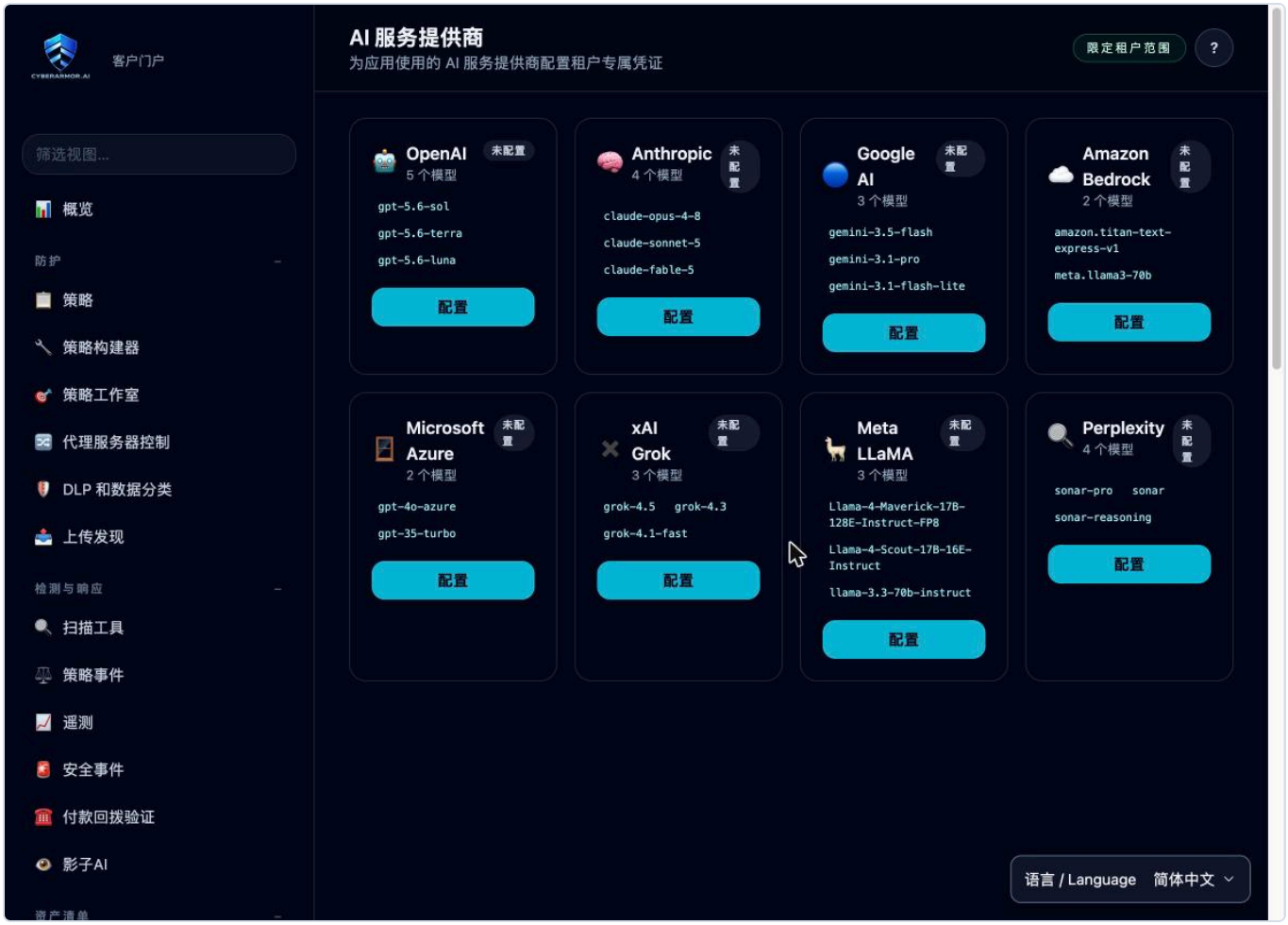
物料清单（Bill of Materials）——持续更新的软件、硬件、ML 模型与采集器清单，包含以下选项卡：**组件**、**变更**（新增、移除、版本变化）、**已加载与已安装**、**来源**（代码库、制品仓库、云资产清单）、**漏洞**（OSV、CISA KEV 与 EPSS，以及 VEX 状态）和 **IOC 扫描**（说明匹配项出现在哪里以及应如何处理）。**导出 CycloneDX** 与**签名导出**可生成证据文件。



代理程序目录（Agent Directory）——在您的活动中出现过的 AI 代理（含事件、拦截、脱敏、服务提供商、模型、工具、用户与主机），第二个选项卡中是您的终端代理。



AI 服务提供商（AI Providers）——贵组织自有应用可通过 CyberArmor.AI 使用的每个 AI 服务提供商各有一张卡片，显示其模型与状态。**配置**用于保存该服务提供商的 API 密钥与设置；**测试**用于检查配置。

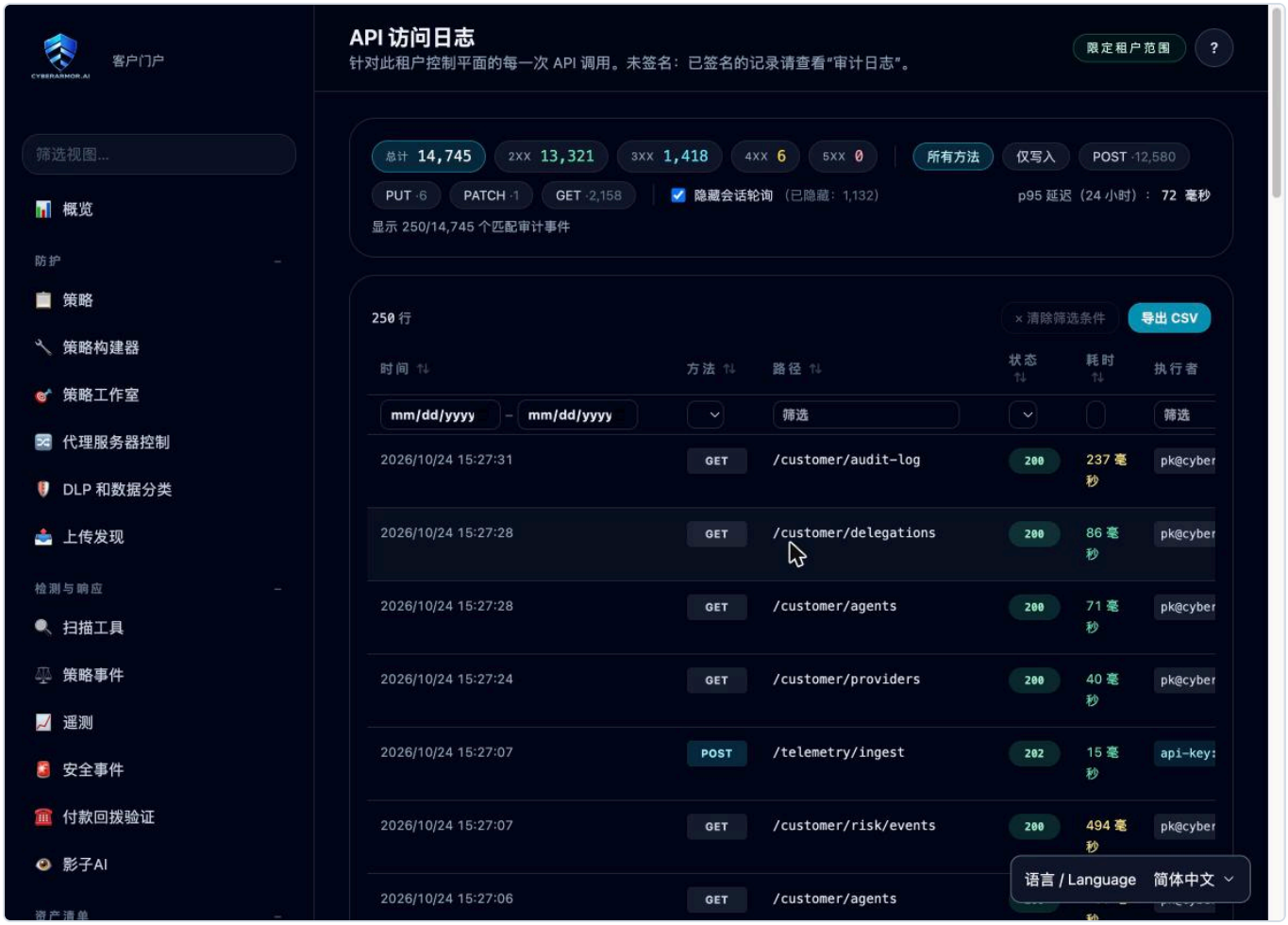


委托管理（Delegation Manager）——记录哪位人员向哪个 AI 代理委托了什么：签发带有范围与到期时间的委托，或**撤销**委托（撤销无法撤回）。



证据

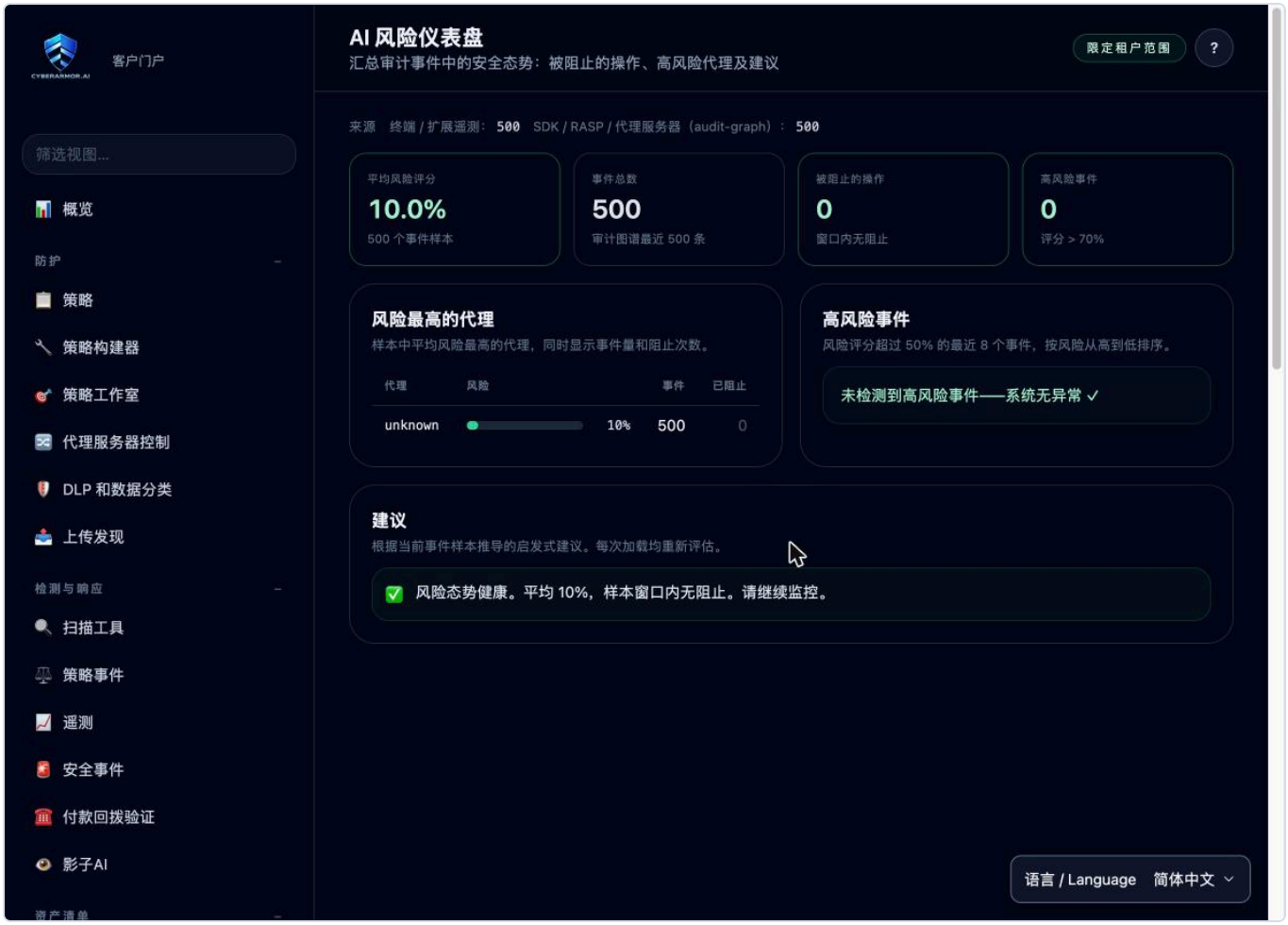
API 访问日志（API access log）——针对 CyberArmor.AI 服务器中本组织部分的每一次调用，包含方法、路径、状态码、耗时与调用者。此日志未签名；签名记录请查看**审计日志**（第 11 节）。



操作图谱（Action Graph）——以图形方式呈现哪些代理和人员使用了哪些 AI 服务、模型与工具，并标出被拦截的连线。点击某个节点可查看其详情并逐步走到相邻节点；点击某项操作可查看其详情。

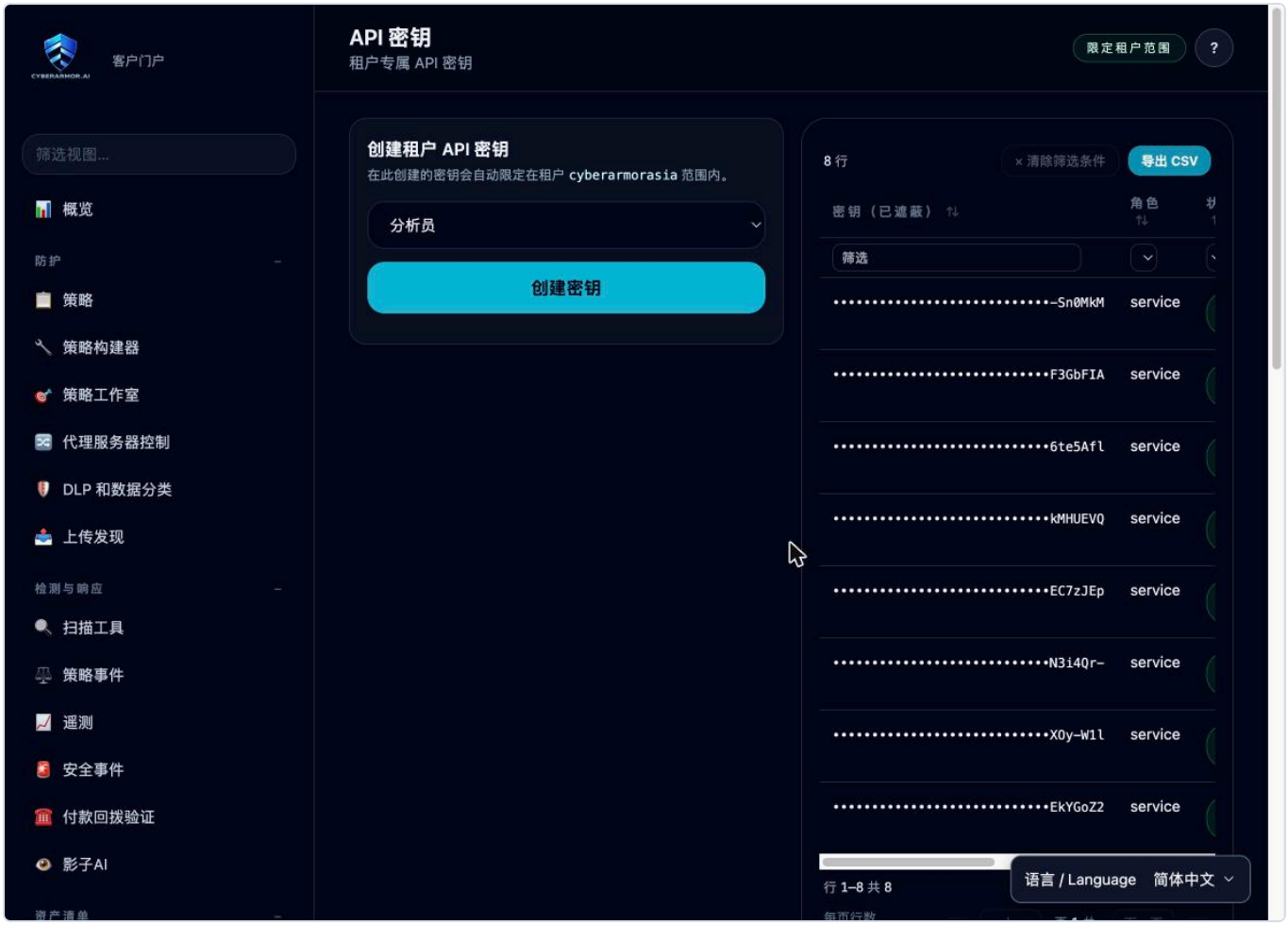


AI 风险仪表盘（AI Risk Dashboard）——在您选择的时间窗口内，显示平均风险评分、风险最高的代理、高风险事件以及改进建议。

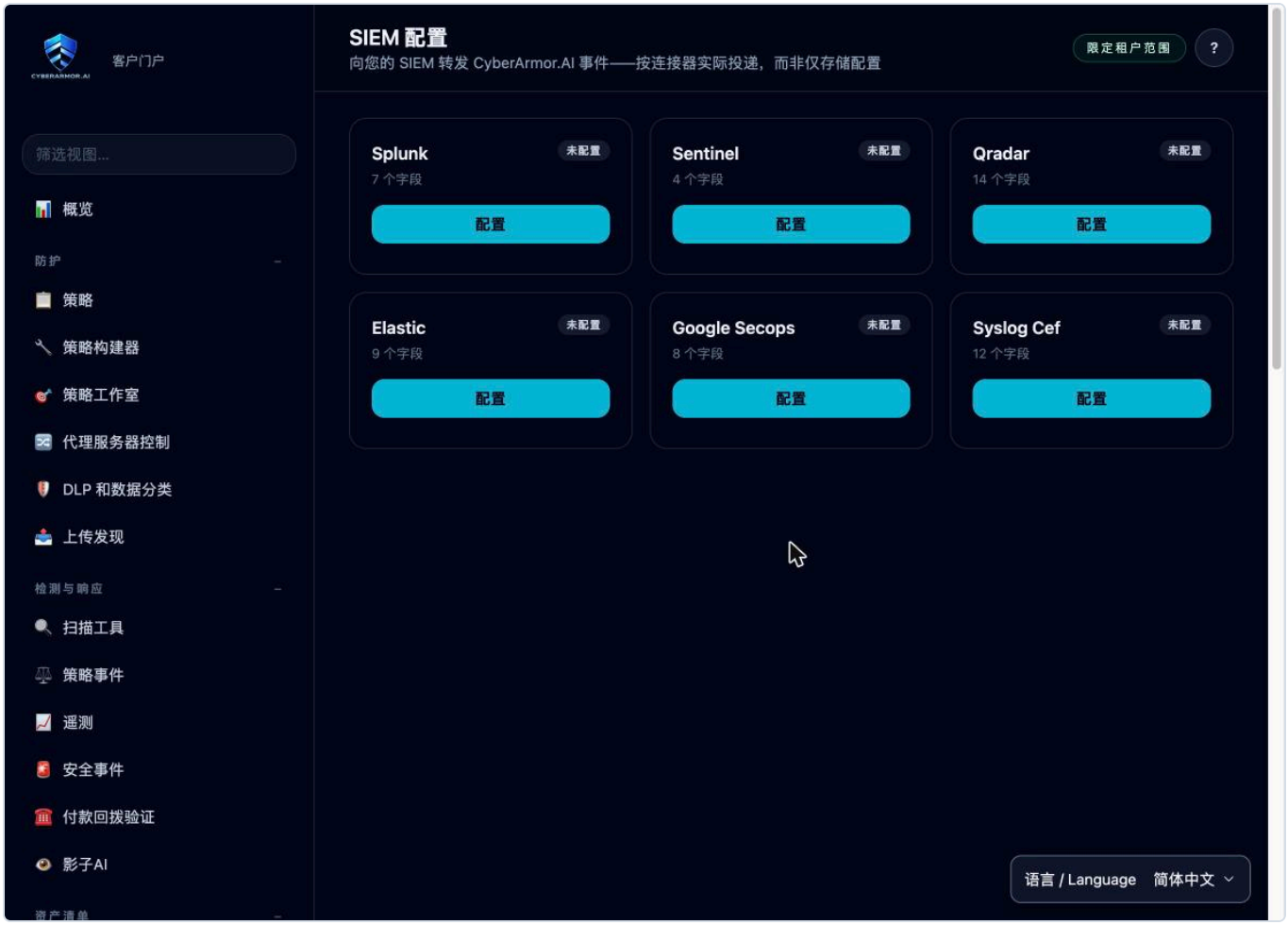


设置与管理

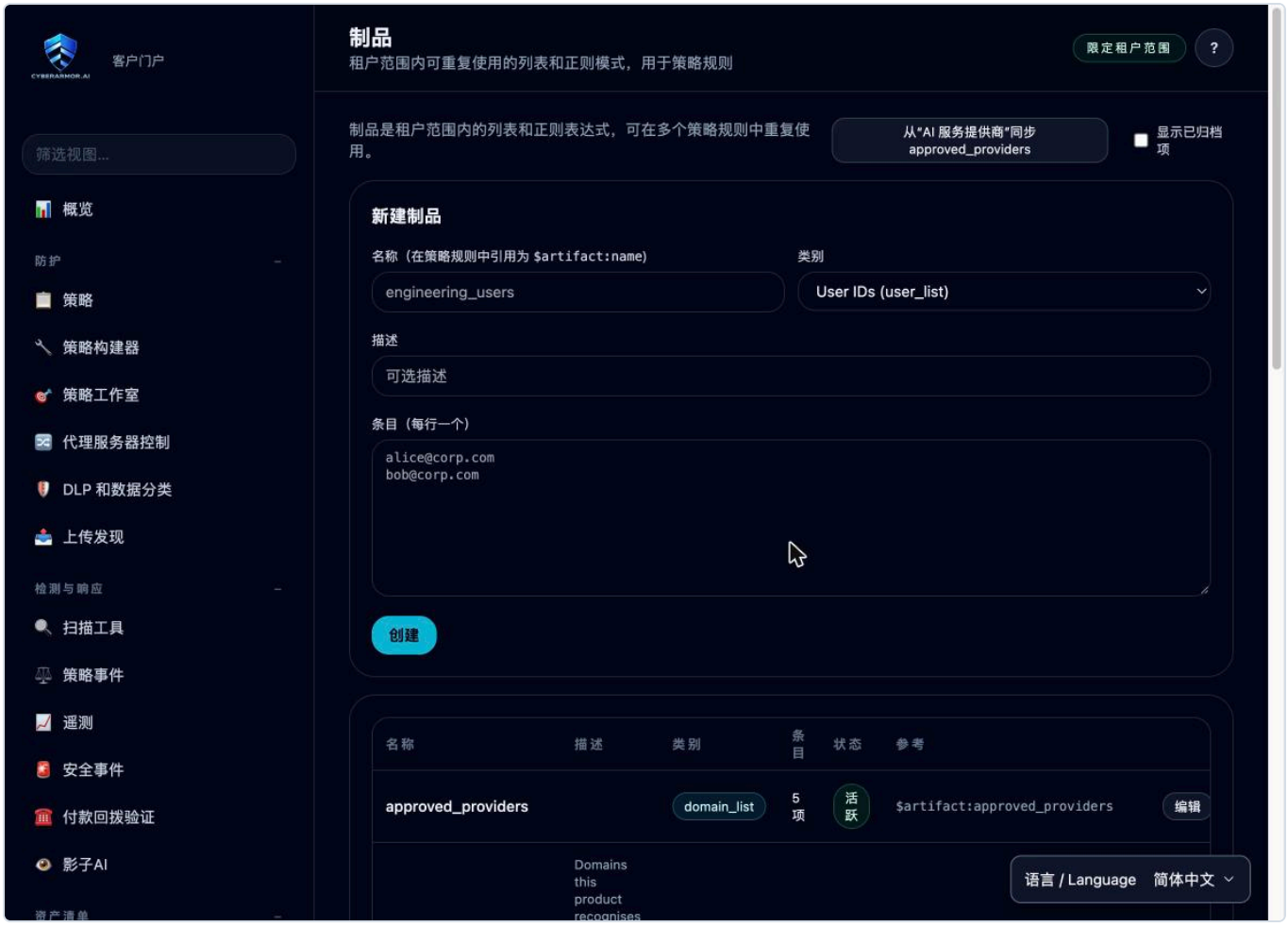
API 密钥（API Keys）——用于对接贵组织自有系统的密钥，每个密钥都对应一个角色。新密钥只在创建时显示一次。**禁用**可关闭某个密钥。



SIEM 配置（SIEM Config）——Splunk、Microsoft Sentinel、QRadar、Elastic、Google SecOps 与 Syslog CEF 各有一个连接器。租户管理员点击**配置**并保存连接器，然后用**测试**发送一条测试事件。密钥类字段保存后不会再次显示。



制品（Artifacts）——策略以 `$artifact:` 加清单名称的形式引用的可复用清单与模式（例如已批准的 AI 服务提供商）。银行与支付清单的**列表库**位于页面顶部（第 7 节）。每个制品都可以编辑、禁用、归档和恢复。


客户门户
筛选视图...
概览
防护
策略
策略构建器
策略工作室
代理服务器控制
DLP 和数据分类
上传发现
检测与响应
扫描工具
策略事件
遥测
安全事件
付款回拨验证
影子AI
资产清单
制品
租户范围内可重复使用的列表和正则模式，用于策略规则
限定租户范围
?
制品是租户范围内的列表和正则表达式，可在多个策略规则中重复使用。
从"AI 服务提供商"同步 approved_providers
显示已归档项
新建制品
名称（在策略规则中引用为 $artifact:name）
engineering_users
类别
User IDs (user_list)
描述
可选描述
条目（每行一个）
alice@corp.com
bob@corp.com
创建
名称
描述
类别
条目
状态
参考
approved_providers
domain_list
5 项
活跃
$artifact:approved_providers
编辑
Domains this product recognises
语言 / Language
简体中文