零信 AI 邮探秘之二：天生智能

构建智能邮件的安全架构

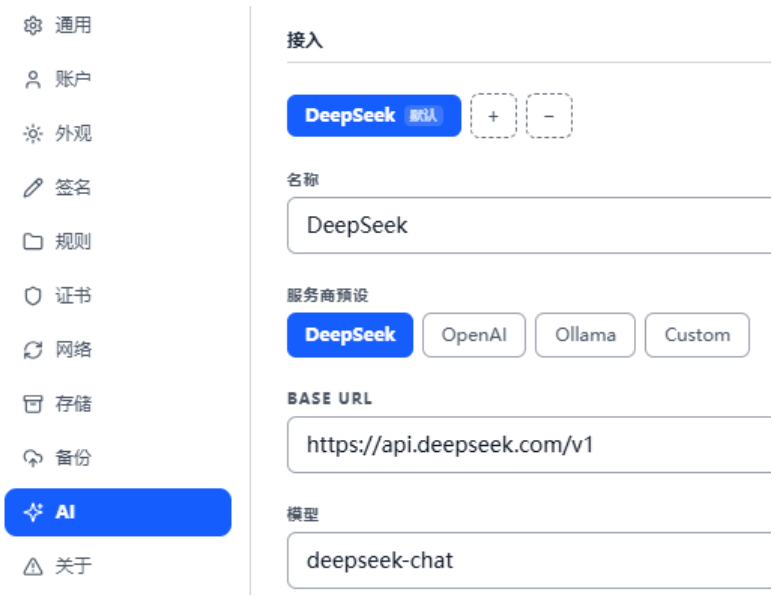
2026 年 9 月 16 日

生成式 AI 正在成为现代邮件服务的常见功能。AI 助手可以起草和修改邮件、总结对话、翻译内容、提取信息、分类来信，并帮助用户处理日益复杂的沟通。这些能力能够提升生产力，但也带来了传统邮件架构尚未充分解决的问题：谁控制与邮箱连接的 AI 服务？哪些服务商被允许处理通信数据？哪些信息可以被提交，以及应当在什么条件下提交？当 AI 生成的邮件越来越逼真时，收件人又该如何判断一封邮件是否确实来自其声称的身份？

最后一个问题尤其重要，因为生成式 AI 正在改变商业邮件入侵（BEC）的实施方式。攻击者可以利用 AI 生成措辞流畅、高度个性化的邮件，模仿高管、供应商、客户或同事的语言风格、语气、业务背景和沟通习惯。由此产生的欺诈邮件，可能不再包含传统安全检查和人工判断所依赖的明显警示信号。挑战已经不只是判断一封邮件看起来是否可疑，而是判断其背后的身份是否可信。

一、用户自带 AI 服务 API Key 能够带来什么

零信 AI 邮不绑定 AI 服务，由用户自带 AI 服务 API Key ，AI 就不再是产品内置的固定依赖，而是一种可配置的服务。用户或组织可以根据自身需求选择 AI 服务商，而不必被迫依赖某一个模型、供应商或数据处理环境。这一区别非常重要，因为不同用户对于模型质量、语言支持、成本、隐私、数据处理地域、监管要求以及提交数据的留存方式，可能有不同的期待。



用户自有的 API Key 可以支持多种实际的邮件功能。AI 服务可以起草或修改邮件、总结长篇对话、翻译内容、提取结构化信息、分类收到的邮件，或帮助用户识别讨论中的主要决定事项和后续行动。用户还可以针对不同任务接入不同的服务商，让用户根据信息的敏感程度和使用目的选择合适的服务。

对于组织而言，自带 API Key 也提供了更清晰的控制点。管理员可以确定哪些服务商获得批准、哪些用户或部门可以使用相关服务、应授予哪些权限，以及哪些信息可以被提交。例如，组织可以允许 AI 辅助处理普通业务邮件，同时对机密文件、客户记录、财务信息或知识产权设置更严格的限制。API Key 并不能解决所有数据治理问题，但它能够让邮件平台与 AI 服务之间的关系更加明确，也更容易管理。

这种模式还保留了一个重要的职责边界：AI 服务 API Key 授权的是对智能服务的访问，并不意味着该服务拥有邮箱、邮件账户或用户的加密身份。AI 可以被授予执行特定任务的权限，而不必成为可读取通信内容的、不受限制的永久保管者。

二、AI 辅助仍然需要密码学信任

对 AI 服务的自主选择，并不能自动建立对邮件本身的信任。AI 可以帮助生成清晰、流畅且符合语境的内容，却无法独立证明发件人就是其声称的那个人。生成语言的质量并不是身份凭证。一封写得非常好的邮件，仍然可能是欺诈邮件。

基于 AI 的检测可以帮助评估一封邮件是否表现异常，但概率并不等于证明。合法邮件也可能具有异常特征，而经过精心准备的欺诈邮件则可能看起来完全正常。当攻击者和防御者都开始使用 AI 时，仅仅依赖更强的内容检查，可能会陷入一场不断升级的对抗：欺骗越来越逼真，模式识别也越来越复杂。

因此，真正需要回答的问题是：收件人能否验证邮件背后的身份？这正是数字签名和密码学身份发挥作用的地方。在 S/MIME 邮件中，数字签名由与证书关联的发件人私钥创建。当证书链有效且受到信任时，收件人便可以验证邮件与其声称身份之间的关系，并检测签名后的内容是否发生过改变。

数字签名本身并不能使发件人成为可信对象，也不能消除所有形式的欺诈。但它提供了一种无法仅凭熟悉的显示名称、写作风格、标识或视觉相似性建立的证据。当 AI 生成的邮件使基于外观的判断变得不可靠时，这种证据尤其有价值。

要让这种证据真正发挥作用，邮件客户端的用户界面必须能够清楚地呈现它。如果证书信息始终隐藏在技术菜单中，收件人仍可能主要依赖姓名、地址和写作风格。零信 AI 邮 App 就是在用户决定是否信任一封邮件或是否执行其中请求的关键位置，展示经过验证的身份信息、

签名状态、身份认证级别以及相关组织信息。

对于企业而言，同样的身份信息还可以支持自动化策略。零信 S/MIME 自动化网关可以将证书有效性、签名状态、组织身份或信任域信息作为额外信号，用于邮件路由、放行、隔离或审批决策。密码学身份并不是要取代现有的内容与行为分析，而是为其提供可靠的补充。

三、两个密钥，构建更强大的邮件安全架构

AI 服务 API 密钥与 S/MIME 证书密钥解决的是不同问题。API Key 决定用户或组织选择哪一种智能服务，以及该服务如何在邮件流程中获得授权；密码学密钥则支持数字签名、加密、身份验证以及对受保护通信的访问。前者控制对智能能力的访问，后者支撑信任与保护。

“两个密钥”模型的价值，在于让两类功能相互连接，却不彼此混淆。在邮件撰写阶段，经过授权的 AI 服务可以帮助起草、修改、总结、翻译或分类内容。用户审核结果，并决定是否发送。获得批准后，密码学机制再应用最终的数字签名和加密控制，使邮件在传输、存储以及后续访问过程中，能够保有可验证的身份和机密性。

这种分工也维护了 AI 应有的角色。AI 助手应当帮助人们处理信息，但不应成为判断发件人是否可信的权威。信任决策需要建立在独立于生成文本说服力的证据之上。密码学提供了权威的可信基础，而 AI 服务仍然是一个受到控制的生产力工具。

因此，“两个密钥”模型的价值高于单独使用其中任何一个密钥。没有密码学信任的 AI，可以让邮件更快、更方便，却无法解决身份问题；没有智能辅助的密码学，可以提供强有力的保护，却无法提升邮件处理效率。两者结合，才能让邮件同时变得更高效、更可信。

四、自动化是实现普及 S/MIME 关键一层

“两个密钥”架构只有在自动化的支持下，才具有实际可行性。S/MIME 长期以来已经提供了成熟的签名和加密机制，但证书签发、续期、密钥处理、收件人配置、恢复以及策略管理，对于普通用户和大型组织而言仍然可能十分复杂。自动化可以消除这些运维负担，同时不能削弱用户或组织的控制权。

AI 集成也遵循同样的原则。零信 AI 邮简化 AI 服务商、API Key、权限以及特定任务的配置，同时让数字签名、加密、身份认证和密钥管理成为正常工作流程中可靠且易于理解的组成部分。安全性不应依赖用户每次收发邮件时都记住一系列技术步骤。

不同环境可能需要不同的实现方式。个人用户和小型组织可以使用零信 AI 邮 App 将 AI 辅助与自动化 S/MIME 保护结合起来的应用；大型企业则可以采用基于零信 S/MIME 自动化网关的基础设施，实现集中管理、策略执行、与现有邮件系统集成，以及组织范围内的证书和加密管理。具体交付模式可以不同，但核心原则保持不变：智能能力应当运行在可信的通信框架之内。

五、让智能运行在可信框架内

智能邮件的未来不会只由 AI 决定。AI 可以改善用户创建、理解和管理通信的方式，但它无法独立建立一封邮件背后的身份。随着生成式 AI 让 BEC 攻击变得更加逼真，如果缺少独立的身份保证来源，内容和行为分析的可靠性就会下降。

用户自带 AI 服务 API 密钥回答的是“谁控制智能能力”这一问题；密码学密钥回答的则是身份、完整性、机密性和受保护访问等问题。两个密钥不能相互替代，但彼此互补：前者让 AI 提升邮件生产力，后者提供 AI 无法自行创造的信任与保护。

当这些能力协同工作时，AI 就不再只是一个孤立的功能，而可以运行在一种新的邮件架构之中：用户选择自己的智能服务，控制访问权限，收件人能够评估可验证的身份，而邮件在发送之后仍然受到加密保护。这正是构建更强大、更安全的智能邮件体系的基础。

王高华

2026 年 9 月 16 日于深圳

欢迎关注零信技术公众号，实时推送每篇精彩 CEO 博客文章。
已累计发表中文 287 篇(共 85 万 1 千多字)和英文 129 篇(17 万 9 千多单词)。