

Documentation Technique : Analyse Approfondie via IA dans l'Extension Anti-Phishing Gmail

O1 Objectif

Ce document décrit la structure technique de l'analyse approfondie d'un email suspect dans une extension Chrome Gmail anti-phishing, en utilisant plusieurs intelligences artificielles et APIs externes. Il est destiné aux développeurs et architectes de l'application.



Vue d'ensemble de l'architecture d'analyse approfondie

L'analyse approfondie se déclenche à la demande de l'utilisateur, lorsque l'analyse locale (DOM) détecte un score de risque élevé.



Etapes principales :

1. Lecture du contenu du mail via le DOM (analyse locale)
2. Nettoyage et filtrage intelligent des données extraites
3. Envoi du contenu filtré au backend/API centrale
4. Dispatch des données vers les IA externes pour analyse (GPT, VirusTotal, etc.)
5. Affichage du rapport final à l'utilisateur



IA/API utilisées pour l'analyse approfondie

IA/API	Rôle principal
OpenAI GPT	Analyse sémantique du corps (ton, urgence, manipulation)
VirusTotal	Scan de fichiers ou d'URLs
PhishTank	Vérification de URLs contre une base phishing connue
AbuseIPDB / WhoisXML	Vérification de la réputation IP/domaine expéditeur
Perspective API (opt)	Détection de langage toxique ou incitatif (optionnel)



Données extraites du mail via le DOM

Donnée	Exemple	Récupérable via DOM
Sujet	"Suspension de votre compte"	<input checked="" type="checkbox"/> Oui (h2 . hP)
Corps du mail (texte)	"Cliquez ici pour éviter le blocage..."	<input checked="" type="checkbox"/> Oui (div . a3s)
Liens (href)	http://phishing.link	<input checked="" type="checkbox"/> Oui

Donnée	Exemple	Récupérable via DOM
Nom de l'expéditeur	"Banque Nationale"	<input checked="" type="checkbox"/> Oui
Email / domaine expéditeur	<input type="text" value="support@banque-fake.com"/>	<input checked="" type="checkbox"/> Partiel
Pièces jointes (nom)	<input type="text" value="facture.pdf"/>	<input checked="" type="checkbox"/> Oui (nom uniquement)

🚫 Filtrage intelligent avant l'envoi aux IA

Donnée	Traitement recommandé	Objectif
Sujet	En clair	Contexte et ton du mail
Corps texte	Nettoyé (voir fonction ci-dessous)	Anonymisation + réduction de bruit
Liens	Transmis en clair (un par un)	Scan ciblé
Email expéditeur	Seulement le domaine (ex: <input)<="" td="" type="text" value="banque.com"/> <td>Protection RGPD</td>	Protection RGPD
Nom de l'expéditeur	Remplacé par <input type="text" value=" [nom]"/>	Anonymisation
Pièces jointes	Nom uniquement (pas de fichier)	Envoi uniquement si opt-in
HTML complet	Non envoyé	Trop lourd et peu pertinent

✓ Fonction JS de nettoyage :

```
function nettoyerTexteEmail(texte) {
    return texte
        .replace(/https?:\/\/[^/]+[\s]/g, '[URL]')
        .replace(/\b[A-Z0-9._%+-]+@[A-Z0-9.-]+\.[A-Z]{2,}\b/gi, '[email]')
        .replace(/Monsieur\s+[^,\n]+/gi, 'Monsieur [nom]')
        .replace(/\d{2,4}[-.\s]?\d{2,4}[-.\s]?\d{2,4}/g, '[numéro]')
        .trim();
}
```

🔄 Mapping : IA → Données → Filtrage

IA/API	Données requises	Filtrage appliqué avant envoi
OpenAI GPT	Sujet + Corps nettoyé + domaine	Anonymisation des noms/emails, remplacement URL
VirusTotal	URLs et pièces jointes (opt.)	Envoi de chaque URL et fichier hash (opt-in)

IA/API	Données requises	Filtrage appliqué avant envoi
PhishTank	URLs	Aucune anonymisation
AbuseIPDB	Domaine de l'expéditeur	Extraction simple, aucun contenu
Perspective API	Corps nettoyé (optionnel)	Idem OpenAI



Exemple de payload à envoyer au backend

```
{
  "subject": "Suspension de votre compte",
  "body_cleaned": "Bonjour,\n\nVeuillez cliquer sur [URL] pour éviter la
suspension.",
  "links": [
    "http://login-urgence.tk",
    "http://retrait.confirm.tk"
  ],
  "sender_domain": "banque-fake.com",
  "attachments": [
    "facture_urgence.pdf"
  ]
}
```



RGPD et gestion des résultats IA

Toutes les données analysées sont :

- **Pré-filtrées** côté navigateur, sans capturer d'identifiants personnels (sauf consentement explicite).
- **Transmises uniquement à la demande** de l'utilisateur, via clic sur "Analyse approfondie".
- **Non stockées durablement**, sauf sous forme anonymisée ou agrégée à des fins statistiques ou pour le tableau de bord des comptes premium.

Le système respecte les principes de minimisation, transparence, sécurité, et limitation de conservation du RGPD. Les utilisateurs professionnels auront un accès sécurisé à un tableau de bord permettant de consulter les historiques et alertes anonymisés.



Interface d'affichage des résultats

Utilisateurs particuliers (mode extension simple)

- Affichage immédiat via une **bannière contextuelle** dans l'interface Gmail
- Pas d'identifiant requis, pas de compte, pas d'historique conservé
- Bouton "Analyser en profondeur" déclenche une popup avec le **score final + verdict résumé**

Utilisateurs PME/TPE (mode SaaS)

- Interface de **tableau de bord web sécurisé** (accès par login/email)
 - Données consolidées par utilisateur, domaine ou collaborateur
 - Alertes, statistiques, et export PDF/CSV possibles
 - API token pour automatisation ou intégration SIEM future
-

Conclusion

Cette structure garantit une **analyse IA avancée**, efficace, **respectueuse de la vie privée**, et conforme aux pratiques RGPD. Elle permet de tirer parti des meilleures IA actuelles tout en assurant une intégration simple dans ton extension anti-phishing Gmail.