

LE PETIT LIVRE DE L'IA · CHAPITRE 4.5 · FICHE DE RÉVISION, STYLE EXAMEN

Hypertrucages, fraude et sécurité

CP 146 · LPD 8 L'IA sert aussi les fraudeurs : voix clonée du directeur, faux courriel parfait, fausse demande de changement de compte de salaire. Elle crée aussi de nouvelles failles : un agent qui lit un document piégé peut être détourné. La sécurité passe par des procédures simples et un double contrôle.

1 L'essentiel à savoir par cœur

- **Hypertrucages** (voix, vidéo) : l'appel urgent du « directeur » pour un virement doit être vérifié par un autre canal.
- Les courriels frauduleux sont désormais **sans fautes** : on vérifie la demande, pas l'orthographe.
- Changement d'**IBAN de salaire** : toujours une confirmation par un canal connu.
- Un **agent** d'IA connecté à la messagerie peut être manipulé par un contenu piégé : limiter ses droits.

2 signatures au-dessus d'un seuil	**1** canal connu pour vérifier	**LPD 24** annonce	**2.12.2026** nouvelle interdiction UE

2 La base légale, article par article

- L'escroquerie est punissable (CP 146) ; l'accès indu à un système informatique aussi (CP 143bis).
- L'entreprise prend des mesures de sécurité adaptées au risque (LPD 8) ; violation de la sécurité à risque élevé annoncée au PFPDT dans les meilleurs délais (LPD 24).
- Le travailleur suit les directives de sécurité de l'employeur (CO 321d) ; il répond du dommage causé intentionnellement ou par négligence, selon les circonstances (CO 321e).
- Règlement européen : les hypertrucages doivent être signalés comme tels par ceux qui les diffusent (RIA art. 50, par. 4).

3 Les règles, cas par cas

Situation	Règle	Article
Appel vocal urgent du directeur pour un virement	Rappeler sur un numéro connu, double signature	CP 146
Demande de changement d'IBAN de salaire	Confirmation par un canal connu, délai d'un jour	CO 321d
Agent d'IA avec accès à la messagerie	Droits limités, pas d'envoi sans validation	LPD 8
Fuite de données par un outil d'IA	Évaluer le risque, annoncer au PFPDT si élevé	LPD 24
Formation contre l'hameçonnage	Exemples récents générés par IA	Bonne pratique

4 La méthode, dans l'ordre

1. Vérifier par un autre canal.
2. Double signature.
3. Limiter les droits des agents.
4. Former avec des exemples récents.
5. Annoncer les violations graves.

> **« Urgence plus argent égale vérification. »**

6 Les pièges

Se fier à la voix ou à l'image : elles s'imitent désormais en quelques secondes d'enregistrement.

7 Le cas type du chapitre

La voix du patron. Un vendredi à 17 h, la comptable reçoit un appel : la voix du directeur, parfaitement imitée, lui demande un virement urgent de 48 000 francs pour un « rachat confidentiel ».

a. Quels signaux d'alerte ?
b. Que faire ?
c. Quelle procédure installer ?

Corrigé en fin de fiche.

8 Le coin de l'expert

- Les attaques par « injection de consigne » cachent des instructions dans un document ou une page web lus par l'IA : ne jamais donner à un agent le droit d'agir seul sur la base de contenus externes.
- Le règlement 2026/1744 ajoute dès le 2 décembre 2026 une interdiction des systèmes produisant des images intimes réalistes de personnes identifiables sans leur consentement (RIA art. 5, par. 1, let. b bis).
- Le harcèlement par hypertrucage (montage d'un collègue) relève de la protection de la personnalité (CO 328) et peut être pénal : réagir vite, conserver les preuves.

10 Questions d'examen

1. Vrai ou faux ? Un courriel frauduleux se reconnaît à ses fautes.
2. Vrai ou faux ? Une voix peut être imitée à partir de quelques secondes d'enregistrement.
3. Vrai ou faux ? Un agent d'IA peut être détourné par un document piégé.
4. Vrai ou faux ? Une violation de la sécurité à risque élevé s'annonce au PFPDT.
5. Vrai ou faux ? Un changement d'IBAN de salaire se fait sur simple courriel.
6. Vrai ou faux ? Les hypertrucages diffusés doivent être signalés dans l'UE.
7. Cas express : La voix du directeur demande un virement urgent : on paie ?
8. Résolvez le cas type de la section 7, « La voix du patron », questions a à c.
9. Citez les étapes de la méthode, dans l'ordre, et la base légale du chapitre.

11 Corrigés

1. Faux : L'IA écrit sans fautes.

2. Vrai : D'où la vérification par un autre canal.

3. Vrai : Injection de consigne.

4. Vrai : LPD 24.

5. Faux : Confirmation par un canal connu.

6. Vrai : RIA art. 50.

7. Non. Rappeler sur un numéro connu ; double signature.

8.

1. Urgence, confidentialité, montant inhabituel, fin de semaine, contournement de la procédure.

2. Raccrocher, rappeler le directeur sur son numéro connu, ne rien payer avant.

3. Double signature au-dessus d'un seuil, mot de passe convenu pour les demandes urgentes, formation ; voir la cybersécurité et la fraude à la paie.

9. 1) Vérifier par un autre canal. ; 2) Double signature. ; 3) Limiter les droits des agents. ; 4) Former avec des exemples récents. ; 5) Annoncer les violations graves. Base : CP 143bis, 146 ; LPD 8, 24 ; CO 321d, 321e, 328 ; RIA art. 5, 50.