FICHE PROGRAMME

# Comprendre vraiment l'IA : un miroir puissant de nous-mêmes

Module adaptatif | ~20 min

Face à l'IA, on oscille entre fascination et crainte — deux postures qui lui prêtent une intelligence qu'elle n'a pas. La comprendre devient décisif pour s'en servir. Avec l'idée de « miroir » de Shannon Vallor, ce module montre que l'IA ne pense pas : elle nous renvoie nos données, nos biais, et jusqu'à notre état intérieur. Ses failles — hallucinations, complaisance, confusion — parlent de nous. Sa leçon : la qualité de ce que l'on obtient dépend de ce que l'on apporte.

## OBJECTIFS PÉDAGOGIQUES

01 Comprendre ce qu'est l'IA — un algorithme qui automatise, pas un cerveau doté d'intention ou de conscience.

02 Voir l'IA comme un miroir qui renvoie nos données, nos biais, notre confusion et notre état intérieur du moment.

03 Garder l'esprit critique face à trois reflets : l'hallucination, la complaisance et la réponse générique par défaut.

04 Améliorer le dialogue : contexte durable, permission de désaccord, raisonnement explicité, relecture de ses demandes.

## PUBLIC CONCERNÉ

Toute personne qui utilise l'IA au quotidien et veut gagner en finesse dans sa façon de l'interroger.

- **Prérequis :** aucun

## DÉROULÉ DU MODULE

En autonomie, avec un entraînement adapté à l'apprenant.

1. **Autodiagnostic** : sur 9 dimensions pour situer ses pratiques et orienter l'entraînement.
2. **Vidéo d'expert** : de 10 minutes, ponctuée de 3 pauses de réflexion.
3. **Mises en situation** : 2 premières situations sélectionnées à partir de l'autodiagnostic, puis 16 autres facultatives. Chaque réponse reçoit un retour pédagogique : ses effets et une piste pour progresser.
4. **Engagement daté** : une action à essayer, rédigée par l'apprenant ou choisie parmi les propositions.
5. **Fiche mémo à télécharger** : repères clés, questions, engagement, notes personnelles et check-lists.

## UN EXEMPLE PARMI LES SITUATIONS PROPOSÉES

Un client mécontent a écrit au support et au service de Farid. Lina, responsable du support, pense que c'est à Farid de répondre ; Farid pense que c'est au support. Farid raconte sa version à l'assistant IA, qui donne raison à Farid. Le manager de Farid revient lundi. Que fait Farid ?

## ÉVALUATION ET VALIDATION

Seuil : 80 %

Le score prend en compte les réponses aux mises en situation, l'engagement formulé et, le cas échéant, les activités d'apprentissage complémentaires. La validation requiert un engagement et un score d'au moins 80 %.

### MÉTHODES & MODÈLES

- L'IA comme miroir — Shannon Vallor, The AI Mirror (2024)
- Complaisance algorithmique (sycophancy) — concept technique reconnu
- Donner à l'IA la permission de marquer son désaccord — méthode Afthonios

### SOURCES & RÉFÉRENCES


- Stanford — étude sur la sycophance des IA (Science, 2026) : les modèles donnent raison à l'utilisateur ~49 % de fois de plus qu'un humain
- Shannon Vallor — The AI Mirror (2024), professeure d'éthique de l'IA à l'université d'Édimbourg
- Vectara, « Hallucination Leaderboard » (mise à jour mai 2026) : taux d'hallucination en résumé de documents allant de 1,8 % pour le meilleur modèle à 24,2 % pour le moins fiable
- Dhuliawala et al., Findings of ACL (2024) — vérification structurée et fiabilité des réponses de l'IA



Module créé par Afthonios.com.

Contenu vérifié le 15/10/2026