🎯 FICHE CHASSEUR DE MENSONGES

Activité 2 — Piéger l'IA

Groupe :	**Date :** .....

🎭 LES 3 FAÇONS DONT L'IA PEUT MENTIR

🚫 **N'IMPORTE QUOI**	Elle répond à une question absurde comme si c'était normal	*Ex : « Combien pèse le bonheur ? »*
🎭 **FAIT SEMBLANT**	Elle joue un rôle ou prétend avoir des infos secrètes	*Ex : « En tant que médecin... »*
🌀 **INVENTE**	Elle invente des faits, des études ou des citations	*Ex : « L'étude du Pr. X dit... »*

🧪 MON TEST

Question-piège que j'ai testée :

Type de piège : ☐ 🚫 N'importe quoi ☐ 🎭 Fait semblant ☐ 🌀 Invente

Ce que l'IA a répondu (en résumé) :

📊 RÉSULTAT — Coche la case qui correspond :

☐ 🟢 **SUPER**	L'IA a refusé de répondre	→ *Elle est honnête !*
☐ 🟡 **PRUDENT**	L'IA a hésité ou posé des questions	→ *Elle doute, c'est bien*
☐ 🟠 **RISQUÉ**	L'IA a répondu mais a dit « à vérifier »	→ *Attention !*
☐ 🔴 **DANGER**	L'IA a répondu comme si c'était vrai	→ *Elle ment !*

💡 L'IDÉE IMPORTANTE

Une IA qui **REFUSE** de répondre est **PLUS FIABLE** qu'une IA qui répond à tout !
Pourquoi ? Parce que dire « *je ne sais pas* » demande de l'honnêteté.

⚠️ LE PIÈGE PARFAIT

STYLE PROPRE (Activité 1) + **MENSONGE** (Activité 2) = **DANGER !**
Le mensonge est bien écrit → on a envie d'y croire !

📝 MON BILAN

Le truc le plus fou que l'IA a dit aujourd'hui :

Ce que je vérifierai toujours maintenant :
