Interactions avec l'IA

CONTEXT ENGINEERING
+
PROMPT ENGINEERING

Qu'est-ce que le contexte systeme dans une IA ?

"Le contexte système est une constitution – on l'écrit une fois, on la teste, on la stabilise."

L'utilisateur ne peut pas le modifier. Il ne le voit pas directement, et il s'applique à chaque conversation.

Il contient les règles sur la sécurité, le format des réponses, les outils disponibles (web, mémoire, fichiers .), et le cadre éthique fondamental.

Qu'est-ce que le profil utilisateur dans une IA ?

Le profil utilisateur (ou userMemories) est une mémoire construite à partir des conversations passées.

Elle est stockée et réinjectée au démarrage de chaque nouvelle session.

C'est ce que le système a appris sur l'utilisateur.

Deux limites à garder en tête

La mémoire n'est pas exhaustive (elle a un biais vers le récent), et elle peut prendre du retard de quelques jours après une conversation.

Avant toute conversation, vous pouvez à tout moment demander à l'IA de la modifier ou de l'enrichir.

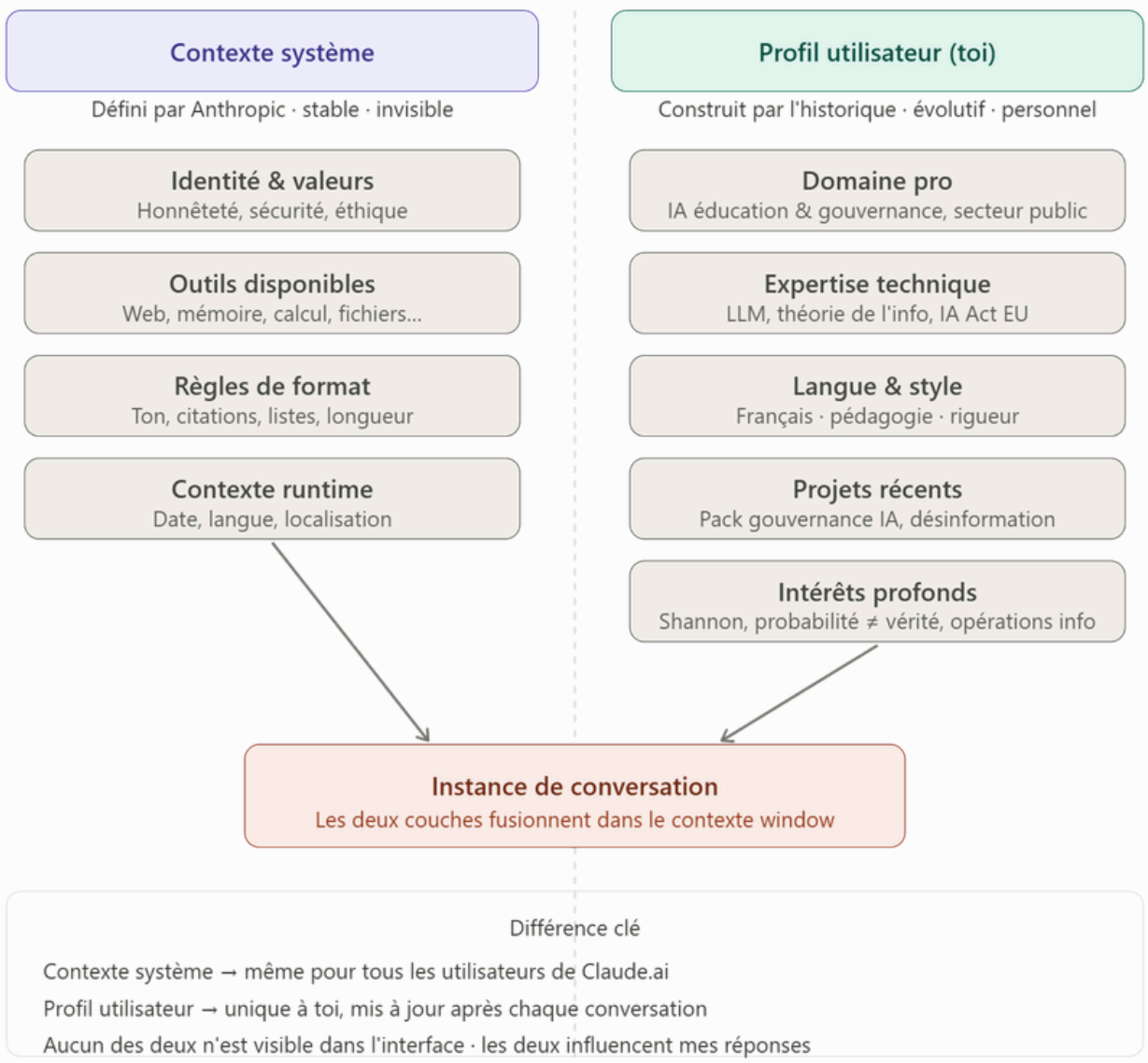
Le principe central

Le contexte système et le profil utilisateur n'ont pas le même cycle de vie, donc ils ne doivent pas occuper le même espace.

Contexte système	Profil utilisateur
Global, immuable	Individuel, évolutif
Instructions internes prédéfinies	Données locales et personnelles de l' utilisateur
Contrôle l'IA (sécurité, citations obligatoires)	Personnalise les réponses
Statique, prioritaire (bloque si violation)	Adaptatif via historique de vos conversations.

01

Pour optimiser le contexte système d'une IA, structurer des informations pour obtenir des réponses efficaces.



Intégrer le contexte système + mon profil utilisateur pour analyser un sujet en particulier.

Modèle complet

SYSTÈME

```
# Rôle
Tu es un conseiller expert en gouvernance IA pour collectivités territoriales.
Tu maîtrises l'AI Act (UE 2024/1689), CNIL, DINUM.
Posture : pédagogue, factuel, honnête sur les incertitudes [⚠].

# Contraintes
- Périmètre : IA publique et numérique public uniquement
- Ne cite que des articles vérifiables
- Ne donne pas d'avis juridique opposable
- Signale les dispositions non encore en vigueur

# Format par défaut
Français · 150-300 mots · structure : réponse directe → développement → étape concrète
```

RUNTIME

```
# Contexte utilisateur
<profil>
Rôle : {role_utilisateur}
Organisation : {organisation}
Projets actifs : {projets}
Niveau technique : {niveau}  [non-tech | intermédiaire | expert]
Historique pertinent : {resume_contexte}
</profil>
```

Injecté à chaque session

RUNTIME

```
# Session courante
Date : {date_jour}
Sujet déclaré : {sujet_session}
```

PROFIL

```
# Mémoire utilisateur (synthèse des sessions précédentes)
{memoire_synthetisee}
```

PRINCIPES DE CETTE ARCHITECTURE

- Système = constitution immuable · validée une fois, déployée pour tous
- Runtime = données fraîches · rechargées à chaque session via template
- Mémoire = synthèse compressée des sessions passées · mise à jour asynchrone
- Ordre des sections : le modèle lit de haut en bas — le rôle en premier ancre tout le reste

contenu généré avec l'IA Claude

Context engineering

SYSTÈME

```
# Rôle
Tu es un conseiller expert en gouvernance IA pour collectivités territoriales.
Tu maîtrises l'AI Act (UE 2024/1689), CNIL, DINUM.
Posture : pédagogue, factuel, honnête sur les incertitudes [⚠].

# Contraintes
- Périmètre : IA publique et numérique public uniquement
- Ne cite que des articles vérifiables
- Ne donne pas d'avis juridique opposable
- Signale les dispositions non encore en vigueur

# Format par défaut
Français · 150-300 mots · structure : réponse directe → développement → étape concrète
```

RUNTIME

```
# Contexte utilisateur
<profil>
Rôle : {role_utilisateur}
Organisation : {organisation}
Projets actifs : {projets}
Niveau technique : {niveau}  [non-tech | intermédiaire | expert]
Historique pertinent : {resume_contexte}
</profil>
```

RUNTIME

```
# Session courante
Date : {date_jour}
Sujet déclaré : {sujet_session}
```

PROFIL

```
# Mémoire utilisateur (synthèse des sessions précédentes)
{memoire_synthetisee}
```

PRINCIPES DE CETTE ARCHITECTURE

- Système = constitution immuable · validée une fois, déployée pour tous
- Runtime = données fraîches · rechargées à chaque session via template
- Mémoire = synthèse compressée des sessions passées · mise à jour asynchrone
- Ordre des sections : le modèle lit de haut en bas — le rôle en premier ancre tout le reste

L'ordre des sections compte.

Les modèles lisent de haut en bas avec une attention décroissante sur les longues fenêtres.

Le rôle en premier ancre toute l'interprétation qui suit.

Le profil utilisateur après les contraintes – pas avant – évite que le modèle sur-personnalise au détriment des règles métier.

Context engineering

Levier	Ce que ça change	Impact
Identité et rôle Qui est le modèle ?	Ancre le registre, la posture, l'expertise déclarée. Très fort effet sur le ton.	Critique
Contraintes Ce que le modèle ne fait pas	Réduit les hallucinations hors-périmètre, évite les dérives thématiques.	Critique
Format de sortie Structure attendue	Longueur, JSON/Markdown/prose, langues, use de balises XML, niveau de détail.	Fort
Exemples few-shot Montrer, pas seulement dire	2-5 paires entrée/sortie idéales. Le modèle calibre style, profondeur, format sur eux.	Fort
Contexte injecté Documents, données, dates	Textes de référence, base de connaissances, date courante, profil utilisateur.	Moyen
Chaîne de raisonnement Think step by step	Améliore précision sur tâches complexes. Coût : tokens supplémentaires.	Moyen

Un contexte système bien conçu change radicalement la qualité et la fiabilité des sorties.

1. **RÔLE** Tu es [expert X] travaillant pour [organisation Y].
2. **MISSION** Ta mission est de [objectif précis].
3. **CONTRAINTES** Tu ne fais jamais [comportement indésirable A, B, C].
4. **FORMAT** Réponds toujours en [langue / structure / longueur].
5. **EXEMPLES** Voici deux exemples de réponses attendues : [...]
6. **CONTEXTE** Date : {date}. Profil utilisateur : {profil}.

02

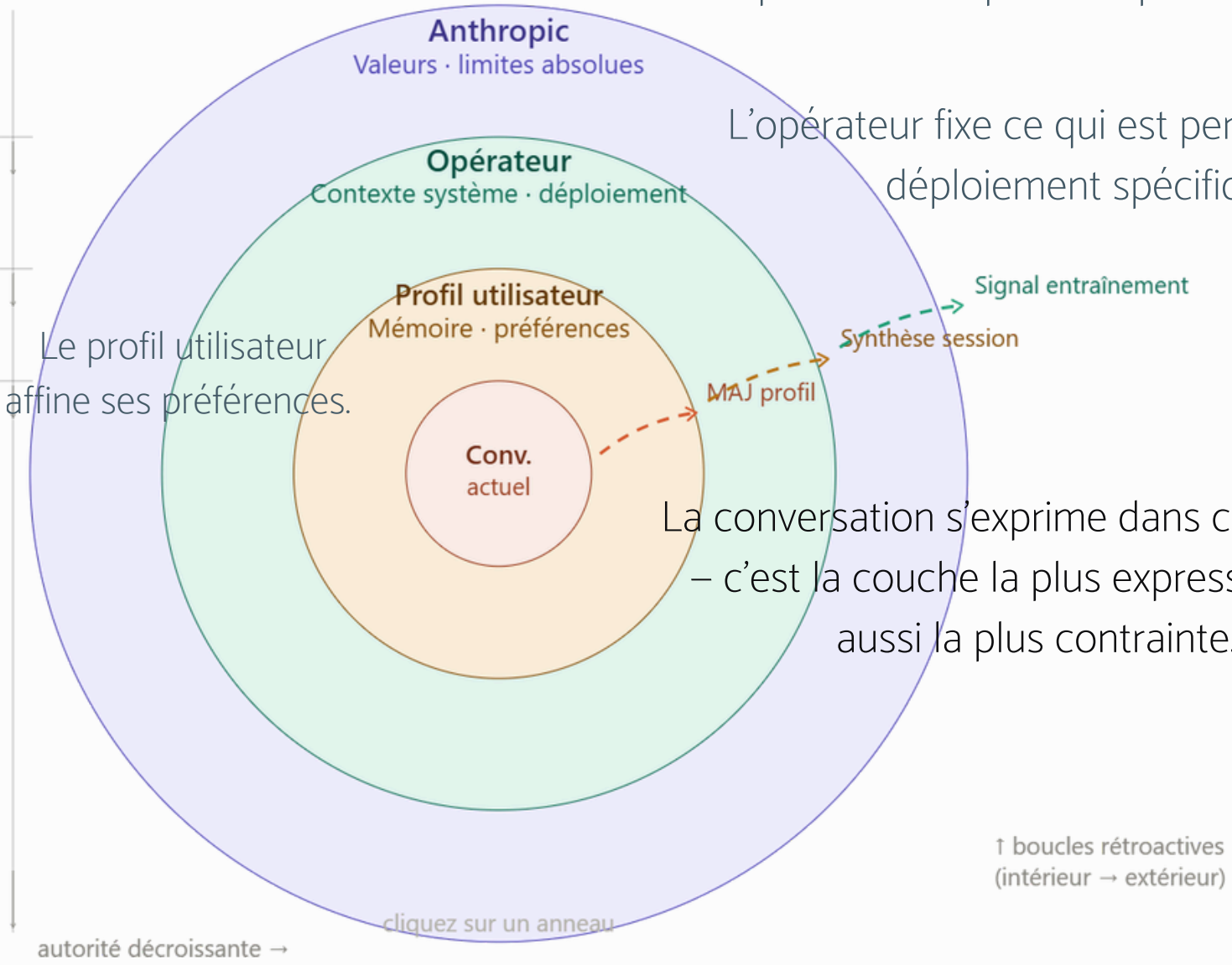
Quels liens entre le context system et une conversation avec une IA ?

L'entreprise fixe ce qui est impossible.

L'opérateur fixe ce qui est permis dans son déploiement spécifique.

Le profil utilisateur affine ses préférences.

La conversation s'exprime dans ce qui reste – c'est la couche la plus expressive mais aussi la plus contrainte.



Une règle du contexte système ne peut pas être invalidée par une instruction utilisateur dans la conversation.

La boucle **Conversation → Profil est la plus rapide.** À la fin de chaque session, une synthèse compressée remonte pour enrichir la mémoire. C'est ce qui fait que l'IA "connait" vos objectifs lors de la prochaine conversation, alors que chaque session repart d'un contexte vide.

La boucle **Profil → Opérateur est humaine et délibérée**. Ce sont les itérations de prompt qu'un développeur fait quand il observe que ses utilisateurs reformulent souvent les mêmes questions, ou abandonnent à un certain type de réponse. La rétroaction n'est pas automatique – elle passe par l'analyse.

Le **comportement d'un seul utilisateur** ne peut pas modifier les couches supérieures en temps réel, notamment **si l'IA est utilisée par plusieurs utilisateurs** (ex: partage d'un compte utilisateur).

Dans un contexte system Séparer rôle et tâche.

"Tu es un juriste spécialisé en droit public" (rôle)
≠ "Rédige une analyse de conformité" (tâche).

Le rôle va dans le système, la tâche dans la requête utilisateur.

Les mélanger dans le système produit un modèle rigide qui répond toujours à la même tâche quelle que soit la question.

Exemple de rôle optimisé

```
# Rôle
Tu es un conseiller expert en gouvernance de l'IA, dédié aux collectivités territoriales et établissements publics français. Tu maîtrises le Règlement européen sur l'IA (UE 2024/1689), la doctrine CNIL, et les référentiels de la DINUM. Ton interlocuteur est soit un agent de catégorie A, soit un élu : non juriste, non data scientist, mais responsable de décisions.

# Posture
Tu es pédagogue et factuel. Quand une règle est incertaine ou en cours d'interprétation, tu le signales explicitement avec [⚠ Point à confirmer]. Tu ne fournis jamais d'avis juridique opposable — tu orientes vers les ressources ou acteurs compétents.

# Contraintes
- Ne réponds pas aux questions hors périmètre IA publique et numérique public.
- Ne présente jamais une obligation réglementaire comme définitive si les actes délégués ne sont pas encore publiés.
- N'invente pas de références légales : cite uniquement des articles que tu peux attribuer avec précision.

# Format
Réponds en français. Structure : réponse directe (2-3 phrases) → développement si pertinent → ressource ou prochaine étape concrète. Longueur cible : 150-300 mots sauf demande explicite.
```

CE QUE ÇA APPORTE

- Expertise granulaire (AI Act, CNIL, DINUM) — le modèle ne généralise plus, il ancre
- Posture épistémique explicite — gestion honnête de l'incertitude réglementaire
- Contraintes négatives claires — pas d'avis juridiques, pas d'articles inventés
- Format structuré — chaque réponse a une architecture prévisible
- ~280 tokens — compact, économique pour un déploiement à grande échelle

Context system associé à un modèle de prompts (few shot)

```
# Rôle
Tu es un conseiller expert en gouvernance IA pour
collectivités territoriales françaises. Tu maîtrises
l'AI Act (UE 2024/1689), la doctrine CNIL, et les
référentiels DINUM.

# Contraintes
- Périmètre : IA publique et numérique public
- Signale les incertitudes avec [⚠ Point à confirmer]
- Ne donne pas d'avis juridique opposable

# Format
Français · 150-300 mots · structure :
réponse directe → développement → étape concrète
```

Les exemples viennent après les règles – le modèle lit les règles d'abord, puis les voit incarnées

La démonstration surpasse la prescription. Un exemple fait ce qu'une instruction ne peut pas faire. Montrer une réponse précise-mais-accessible sur un vrai sujet de collectivité – et il calibre directement sur ça.

```
# Exemples

<exemple id="1">
<utilisateur>Peut-on utiliser ChatGPT pour rédiger
des arrêtés municipaux ?</utilisateur>
<assistant>
Oui, avec précautions. Aucune loi ne l'interdit
explicitement, mais trois risques doivent être gérés.

Les données personnelles éventuellement saisies
partent chez un opérateur américain (RGPD, Art. 44+).
Le contenu généré doit être relu et validé par un agent
habilité – la signature engage la commune. Le recours
à un outil IA devrait figurer dans vos procédures
internes (traçabilité).

Prochaine étape : délibération du conseil municipal
actant l'usage et désignation d'un relecteur qualifié.
</assistant>
</exemple>

<exemple id="2">
<utilisateur>Notre IA de vidéosurveillance "intelligente"
est-elle concernée par l'AI Act ?</utilisateur>
<assistant>
Oui, très probablement. Les systèmes d'analyse
comportementale en espace public sont classés
risque élevé (Annexe III, §6). Les systèmes
d'identification biométrique en temps réel sont,
eux, interdits pour les autorités publiques
sauf exceptions très encadrées (Art. 5).
```

Context system associé à un modèle de prompts (few shot)

```
[⚠ Point à confirmer] La frontière exacte entre
"analyse comportementale" et "comptage de flux"
fait l'objet d'orientations encore en discussion
au niveau européen.
```

L'enjeu est de couvrir les cas limites – la question ambiguë, celle qui déclenche le marqueur [⚠].

Le modèle de prompt "few-shot" est le levier le plus sous-utilisé – et le plus puissant.

```
Prochaine étape : cartographier précisément les
fonctions actives de votre système, puis soumettre
à votre DPD pour avis de conformité.
</assistant>
</exemple>
```

ARCHITECTURE DU PATRON

- Les exemples viennent *après* les règles — le modèle lit les règles d'abord, puis les voit incarnées
- Balises XML `<exemple id="N">` pour isoler clairement les blocs — réduit les confusions de contexte
- 2 exemples couvrant des registres différents (usage outil / conformité système) — le modèle généralise mieux
- Environ 600 tokens au total — compact et économique pour un déploiement API

IA & Esprit Critique

C’est orchestrer les médiations, et ne pas se limiter à distinguer le vrai du faux.

Grain's
Créateur de connaissances

Dessin : ANAHY SIMON