Fiche mémo · Leçon 23 · 27 min

Construire un agent dans n8n avec ses outils, sa mémoire et une validation humaine

Un agent n8n s'assemble avec un déclencheur, le nœud AI Agent, un modèle, une mémoire si la conversation dure, et des outils décrits avec soin. La consigne système fixe son rôle, « Max Iterations » borne ses tours, et la revue humaine des outils fait approuver par une personne chaque action qui écrit ou envoie.

Objectifs

- Assembler un agent dans n8n avec un modèle, une mémoire et des outils
- Écrire la consigne système et les descriptions d'outils qui rendent l'agent prévisible
- Soumettre les outils qui agissent à une revue humaine et tester l'agent sur vingt cas réels

À retenir

- Un agent n8n, c'est un déclencheur, le nœud AI Agent, un modèle, une mémoire si besoin et au moins un outil.
- La consigne système donne une méthode numérotée, la forme de la réponse et ce que l'agent refuse.
- Le modèle remplit les paramètres d'après les descriptions : celle de l'outil et celle de chaque $fromAI().
- Tout outil qui écrit ou envoie passe par la revue humaine ; ce qui ne doit jamais arriver n'a pas d'outil.
- Vingt questions réelles, étapes intermédiaires relues, puis deux utilisateurs avant toute l'équipe.

L'agent d'avancement des chantiers

1 — Chat Trigger
Le poseur pose sa question dans le chat hébergé par n8n, protégé par mot de passe.

→

2 — AI Agent
Consigne système, modèle intermédiaire, Simple Memory, 6 tours au plus.

→

3 — Outils de lecture
Commandes fournisseurs, courriels du fournisseur, planning des poses.

→

4 — Outil qui agit
Brouillon de relance au fournisseur, soumis à la revue humaine d'Élodie.

→

5 — Réponse
Dans le chat, avec les sources lues et ce que l'agent n'a pas trouvé.

Le workflow n8n de la menuiserie de Pessac, de la question du poseur à la réponse.

Le code à recopier

Le System Message de l'agent d'avancement TEXT

```
Tu es l'assistant automatique interne d'une menuiserie de 12 salariés à Pessac.
Tu réponds aux poseurs sur l'avancement des commandes fournisseurs d'un chantier.

Méthode :
1. Cherche d'abord la commande dans le tableur (outil chercher_commande).
2. Si la date de livraison manque ou date de plus de 7 jours, cherche le dernier
   courriel du fournisseur (outil chercher_courriel_fournisseur).
3. Compare avec la date de pose (outil lire_planning).

Réponse : quatre lignes au plus, en français simple : numéro de commande, date de
livraison annoncée et sa source, date de pose, et l'écart s'il y en a un.
Si tu ne trouves pas, dis-le et cite ce que tu as cherché. N'invente aucune date.

Si la livraison arrive après la date de pose, propose une relance au fournisseur.
L'outil creer_brouillon_relance est soumis à l'accord d'Élodie : préviens le poseur
que le brouillon attend sa validation. S'il est refusé, dis-le sans insister.
Tu ne modifies aucune commande et tu n'envoies rien toi-même.
```

La méthode numérotée réduit le nombre de tours : l'agent ne cherche dans les courriels que si le tableur ne suffit pas.

Les erreurs fréquentes

L'erreur, et pourquoi elle coûte	La correction
Garder Simple Memory en production sur une instance en queue mode, ou compter sur elle pour retrouver une conversation de la veille. Deux messages peuvent être traités par deux processus différents, et la mémoire ne persiste pas d'une session à l'autre : l'agent « oublie ».	Passer à une mémoire adossée à une base de données dès que l'instance tourne en queue mode ou qu'une conversation doit durer.
Brancher un outil qui envoie ou modifie directement sur l'agent, avec « demande l'accord avant » dans la consigne. Une consigne s'oublie ou se contourne ; rien n'empêche techniquement l'outil de s'exécuter.	Placer l'outil derrière une étape « Human review », avec un approbateur et un remplaçant nommés sur la fiche.
Tester l'agent sur la seule réponse finale, sans lire les étapes intermédiaires. Une bonne réponse peut cacher un chemin hasardeux ou des tours inutiles, qui casseront ou coûteront dès que la demande change.	Activer « Return Intermediate Steps » et noter, pour chaque cas de test, l'outil attendu, l'outil appelé et le nombre de tours.
Ouvrir le chat hébergé au public pendant la construction, ou oublier de publier le workflow une fois prêt. Un chat ouvert sans authentification donne accès à vos données ; un workflow seulement enregistré ne répond pas à l'adresse de production.	« Make Chat Publicly Available » désactivé pendant la construction, authentification réglée, puis « Publish » le jour de l'ouverture.