

Grille d'Évaluation — Demo Day

Adservio | Dr Olivier Vitrac

2026-02-03

Contents

1 Grille d'Évaluation — Demo Day	1
1.1 Principes d'Évaluation	1
1.2 Grille d'Évaluation	1
1.2.1 1. Clarté du Cadrage (0-4 points)	1
1.2.2 2. Qualité de la Trace & Auditabilité (0-4 points)	2
1.2.3 3. Gestion des Échecs (0-4 points)	2
1.2.4 4. Réalisme Ingénierie (0-4 points)	2
1.2.5 5. Transférabilité (0-4 points)	3
1.3 Barème Global	3
1.4 Questions Croisées Obligatoires	3
1.4.1 Pour le groupe Testing Agent :	3
1.4.2 Pour le groupe RAG Code/Docs :	3
1.4.3 Pour le groupe MCP Server :	3
1.5 Format de Présentation	4

1 Grille d'Évaluation — Demo Day

Adservio — Workshop Agents IA

1.1 Principes d'Évaluation

- On évalue des **systèmes**, pas des individus
 - L'échec documenté vaut mieux que le succès inexpliqué
 - La trace est aussi importante que le résultat
-

1.2 Grille d'Évaluation

1.2.1 1. Clarté du Cadrage (0-4 points)

Score	Critère
0	Problème mal défini, périmètre flou
1	Problème identifié mais mal délimité
2	Cadrage correct, quelques ambiguïtés
3	Cadrage précis, contraintes explicites

Score	Critère
4	Cadrage exemplaire, falsifiable

Questions à poser : - Le problème est-il clairement énoncé ? - Les contraintes sont-elles explicites ? - Le périmètre est-il réaliste ?

1.2.2 2. Qualité de la Trace & Auditabilité (0-4 points)

Score	Critère
0	Pas de trace ou trace inutilisable
1	Trace partielle, format incohérent
2	Trace présente mais lacunaire
3	Trace complète et structurée
4	Trace exemplaire, rejouable

Questions à poser : - Peut-on reconstruire le raisonnement de l'agent ? - Les échecs sont-ils tracés ? - Le format est-il exploitable ?

1.2.3 3. Gestion des Échecs (0-4 points)

Score	Critère
0	Échecs ignorés ou cachés
1	Échecs mentionnés sans analyse
2	Échecs analysés superficiellement
3	Échecs analysés avec causes identifiées
4	Échecs transformés en apprentissage

Questions à poser : - Qu'est-ce qui n'a pas marché ? - Pourquoi ? - Qu'auriez-vous fait différemment ?

1.2.4 4. Réalisme Ingénierie (0-4 points)

Score	Critère
0	Prototype jetable, non maintenable
1	Code fonctionnel mais fragile
2	Solution viable avec limitations connues
3	Solution déployable avec ajustements
4	Prêt pour intégration Adservio

Questions à poser : - Cela pourrait-il tourner en production ? - Quelles sont les limitations ? - Quel effort pour industrialiser ?

1.2.5 5. Transférabilité (0-4 points)

Score	Critère
0	Solution ad-hoc, non réutilisable
1	Quelques éléments réutilisables
2	Patterns identifiés et documentés
3	Solution adaptable à d'autres contextes
4	Contribution à la doctrine Adservio

Questions à poser : - Qu'est-ce qui est réutilisable ? - Comment cela s'articule avec les autres parcours ? - Quelle valeur pour Adservio au-delà du workshop ?

1.3 Barème Global

Total	Appréciation
0-7	Insuffisant — Objectifs non atteints
8-11	Passable — Bases acquises, lacunes
12-15	Satisfaisant — Objectifs atteints
16-18	Bien — Maîtrise démontrée
19-20	Excellent — Référence pour Adservio

1.4 Questions Croisées Obligatoires

Chaque groupe doit répondre à au moins une question sur les **autres parcours** :

1.4.1 Pour le groupe Testing Agent :

- Comment votre agent pourrait-il bénéficier d'un index RAG sur la documentation ?
- Quels outils exposeriez-vous via MCP pour votre agent ?

1.4.2 Pour le groupe RAG Code/Docs :

- Comment votre système de détection d'incohérences pourrait-il alimenter un agent de test ?
- Quels outils déterministes renforceraient la fiabilité de vos réponses ?

1.4.3 Pour le groupe MCP Server :

- Comment un index RAG améliorerait-il le choix d'outils par l'agent ?
- Quels tests automatisés valideraient le comportement de vos outils ?

1.5 Format de Présentation

- **Durée :** 15 minutes par groupe + 10 minutes questions
- **Structure suggérée :**
 1. Problème et cadrage (2 min)
 2. Démo live (5 min)
 3. Analyse des traces (3 min)
 4. Échecs et apprentissages (3 min)
 5. Questions croisées (2 min)

Grille à utiliser par les évaluateurs et à partager avec les participants avant le Demo Day.