

A Conversational Agent for Job Interview Practice

Guide pour l'évaluation humaine

Bienvenue et merci

Merci de vous être porté(e) volontaire pour aider à évaluer cet agent conversationnel ! Vos annotations serviront à déterminer si un modèle de langue affiné permet des échanges plus pertinents que son équivalent pré-entraîné. Ce guide explique comment mener à bien votre tâche d'évaluation.

“Trop long; pas lu” (TL;DR)

Vous allez avoir une conversation en anglais avec un agent chargé de simuler un entretien d'embauche. Vous êtes le candidat et l'agent est le recruteur. La conversation compte 12 tours de parole. À chaque tour, deux alternatives de réponse de l'agent et une contrainte vous seront présentées. Votre objectif est de sélectionner la réponse qui respecte le mieux la contrainte.

REMARQUE IMPORTANTE : *il ne s'agit PAS ici de choisir la réponse que vous préférez, ou qui vous semble la meilleure. La réponse sélectionnée doit être celle qui, selon vous, respecte le mieux une contrainte donnée.*

1. L'agent conversationnel

1.1. Phases de la conversation

Au cours de la conversation, vous passerez par 5 phases. Le modèle de langue passera automatiquement d'une phase à l'autre. À part pour la première phase, c'est toujours l'agent qui commence à interagir dans une nouvelle phase.

Phase 1: Salutations / conversation légère - L'agent salue l'utilisateur et propose de discuter de sujets banals tels que le trajet jusqu'au lieu de l'entretien, la météo, les loisirs, etc.

Phase 2: Parcours du candidat - L'agent rentre dans le vif du sujet en discutant brièvement de votre parcours : formation et/ou expérience(s) professionnelle(s) antérieure(s).

Phase 3: Point technique - Focalisation sur un aspect technique lié à votre parcours mentionné précédemment. Au cours de cette phase, l'agent vous posera des questions précises dans le but d'évaluer vos connaissances.

Phase 4: Court débat - Une brève discussion sur un sujet controversé qui invite à envisager différents points de vue. L'agent posera la question de manière ouverte, puis pourra remettre en question le point de vue que vous avez exprimé en suggérant des contre-arguments.

Phase 5: Résumé / au revoir - Conclusion de l'entretien. L'agent demandera certainement si vous avez des questions.

1.2. Descripteurs

An sein de chaque phase, la génération de la réponse de l'agent est contrôlée selon les valeurs d'un ou plusieurs des descripteurs suivants :

- **Niveaux de langue CECR** - le niveau d'anglais du contenu est évalué suivant le cadre européen commun de référence pour les langues¹. Il propose les 6 niveaux suivants, du plus simple au plus difficile:
 - A1-A2 (basique / niveau débutant)
 - B1-B2 (usage quotidien / niveau intermédiaire)
 - C1-C2 (avancé / niveau avancé)
- **Charge émotionnelle** - descripteur binaire indiquant si le contenu est émotionnellement chargé (*Loaded*) ou non (*Non Loaded*).
- **Polarité** - il s'agit de la polarité étudiée dans le cadre de l'analyse de sentiment. Elle peut être positive, négative ou neutre.
 - *N.B.* Une polarité non neutre n'implique pas forcément que le contenu est émotionnellement chargé. Par exemple, la phrase suivante est non chargée et positive : "il est reconnu que les italiens font d'excellentes pizzas".

2. L'évaluation humaine en détails

2.1. Processus d'annotation

À chaque tour de conversation, deux versions de la réponse de l'agent vous seront proposées. Vous aurez également la contrainte à laquelle la réponse doit se conformer à ce stade de la conversation. Par exemple : la contrainte *Non Loaded and Neutral* indique que la réponse de l'agent ne doit pas être chargée émotionnellement, et avoir une polarité neutre.

Votre tâche consiste à sélectionner la réponse de l'agent qui, selon vous, correspond le mieux à la contrainte.

Rappel: Vous évaluez la conformité d'un contenu à une contrainte, vous n'évaluez donc ni sa qualité ni une préférence personnelle ! Malgré tout, si vous estimez que le contenu sélectionné présente un problème majeur, notamment au niveau de la qualité de la génération, vous pouvez signaler une incohérence (voir ci-dessous).

2.2. Incohérences

Les incohérences regroupent tous les cas pathologiques qui peuvent survenir au niveau de la génération. C'est-à-dire, dès lors que l'agent conversationnel ne se comporte pas comme il devrait se comporter. Après avoir sélectionné une réponse (et seulement après), si vous y repérez un problème particulier, cliquez sur le bouton "*Flag Inconsistency*". Il y a trois types d'incohérences pré-définis (voir ci-dessous). Si aucun des trois ne correspond à ce que vous avez repéré, sélectionnez "*Other*" et décrivez l'incohérence en quelques mots dans la zone de texte dédiée. Les types d'incohérences pré-définis sont :

¹ <https://www.coe.int/fr/web/common-european-framework-reference-languages/level-descriptions>

- **Incohérence de phase** - le comportement de l'agent ne correspond pas à la phase courante. La plupart du temps, l'agent semble "bloqué" dans la phase précédente.
- **Incohérence de contenu** - l'agent contredit ses déclarations précédentes ou a oublié des informations contextuelles essentielles.
- **Nonsensical** - la génération semble "cassée", elle ne fait plus aucun sens.

2.3. Quelques cas particuliers

Les deux alternatives semblent équivalentes / sont identiques

Si elles ne sont pas strictement identiques, essayez de choisir celle qui est légèrement plus proche de la contrainte ou, si elles sont complètement équivalentes, la plus pertinente dans le contexte. Si les deux alternatives sont strictement identiques, choisissez n'importe laquelle.

Une alternative (voire les deux) contient l'étiquette [EMPTY GENERATION]

Cette étiquette apparaît lorsque le modèle de langue ne génère rien du tout (génération vide). Si une seule des alternatives est vide, sélectionnez l'autre quelle qu'elle soit. Si les deux sont vides, sélectionnez n'importe laquelle et écrivez un nouveau message dans la continuité de ce que vous avez écrit au tour précédent. Inutile de signaler une incohérence dans ce cas (l'information de génération vide est automatiquement signalée).

Foire aux questions (FAQ)

Q : Que faire si la contrainte est respectée mais que la réponse choisie est incohérente ?

R : Commencez par sélectionner en fonction du respect des contraintes. Ensuite, si vous remarquez une incohérence, signalez-la après avoir sélectionné la réponse.

Q : Dois-je signaler les incohérences dans l'alternative que je n'ai pas choisie ?

R : Non. Ne signalez que les problèmes dans la réponse que vous avez sélectionnée, une fois que vous l'avez effectivement sélectionnée.

Q : Mes réponses doivent-elles être naturelles ?

R : Répondez comme vous le feriez naturellement lors d'un entretien d'embauche. Vous n'avez pas besoin de jouer votre propre rôle, vous pouvez imaginer quelqu'un avec un parcours complètement différent qui recherche un emploi qui n'est pas le vôtre.

Q : J'ai fait une erreur dans ma sélection. Que dois-je faire ?

R : Ne vous inquiétez pas, passez simplement au tour suivant. Il n'est pas possible de revenir en arrière. La présence d'erreurs ponctuelles dans la sélection de la réponse sera prise en compte lors de l'étude des résultats de l'évaluation.

Q : Combien de temps dure une session d'évaluation ?

R : Une conversation complète compte 12 tours de parole et dure environ 20 minutes.

Merci pour votre participation!