

# Marketplace Briefing Bot PoC

Ce document détaille les attentes d'Altavia concernant la conception d'une application de projet de preuve de concept (PoC) qu'Altavia souhaiterait déléguer à un partenaire.

Il fournit des informations contextuelles, définit l'objectif et les histoires d'utilisateurs du PoC, couvre les données fournies, les livraisons attendues et le cadre de collaboration qu'Altavia souhaite mettre en place avec ce partenaire.

## Le contexte

Altavia développe **une solution de place de marché** pour permettre aux marques clientes et aux distributeurs d'acheter tous leurs supports marketing (imprimés, PLV, goodies, numériques, ...). Une première version d'une application web orientée client sera en ligne d'ici juin 2024, avec au moins un client embarqué sur la solution.

## Objectif

Altavia souhaite construire un bot conversationnel visant à aider l'utilisateur, par exemple un **responsable marketing de L'Oréal Paris, à spécifier ses besoins et à préparer une commande pour la production de supports marketing.**

L'objectif de ce bot est d'aider les clients d'Altavia **à affiner leurs besoins et de les conseiller et les rediriger vers le produit adéquat** (sorte de moteur de recherche conversationnel). Cela les aidera à trouver plus facilement les produits correspondant à leurs besoins, et épargnera les chefs de projet d'Altavia, tout en augmentant l'autonomie des clients.

Nous limitons le champ d'application à un ensemble de produits correspondant à ceux qui sont pertinents pour un client d'Altavia. Nous limitons le PoC aux langues française et anglaise en termes de conversation et d'ensembles de données.

Altavia attend du partenaire qu'il lui fasse prendre conscience que la prise en charge de deux langues pour ce PoC peut avoir un impact significatif sur la faisabilité, le délai et le coût.

En termes de volumes, Altavia a traité environ **20 000 dossiers au cours de l'année 2023** et s'attend à ce que la solution finale puisse s'adapter à ce nombre d'occurrences.

Avant de commencer à concevoir et à s'engager sur cette application, qui sera en production face au client, **Altavia veut prouver sa faisabilité et que le résultat de l'application correspond à ses attentes en termes de qualité.** C'est la raison pour laquelle Altavia a décidé de réaliser un PoC.

## User stories - Histoires d'utilisateurs

Nous envisageons 2 histoires d'utilisateurs distinctes, correspondant à des variations du même cas d'utilisation principal.

### Histoire 1 :

#### "Basé sur des documents

Ce premier cas d'utilisation "basé sur des documents" **propose à l'utilisateur de télécharger un document initial**, qui est censé être un mémoire : à partir de ce mémoire, **le robot est censé poser des questions à l'utilisateur afin d'obtenir autant d'informations que possible**, de sorte qu'il recueille suffisamment de données pour être en mesure d'émettre des spécifications détaillées en vue de l'exécution d'une commande.

#### En tant que client

1. Je suis connecté à l'application web de la place de marché et j'ai ouvert l'interface utilisateur du bot chat, ou je me connecte à un bot Slack / Teams / WhatsApp ;
2. la conversation commence avec l'émission par le bot d'un premier message du type "Bienvenue". Veuillez m'envoyer votre dossier, afin que je puisse vous aider à concevoir votre commande.1" ;
3. Je sou mets un document bureautique (un document PDF / PowerPoint / Markdown / HTML / texte brut) directement dans l'interface utilisateur du robot, qui contient la description d'un briefing pour la commande de produits marketing ;
4. le robot "analyse" ce document et poursuit la conversation avec l'utilisateur : il continue à poser des questions à l'utilisateur de manière conversationnelle, tant qu'il "détecte" qu'il lui manque des détails pour pouvoir résumer les exigences de cette commande au moyen de spécifications, qui sont suffisamment détaillées pour que la commande puisse être exécutée ;
5. une fois que le robot a recueilli suffisamment de détails pour que la commande puisse être exécutée, l'utilisateur peut demander au bot de l'aider ;
6. une fois que le robot a recueilli suffisamment de détails sur la commande, il conclut la conversation en écrivant une proposition à l'utilisateur avec les spécifications détaillées concernant la commande, à savoir le produit correspondant disponible sur la place de marché qui est concerné et recommandé, la quantité et les caractéristiques aussi détaillées que possible, de sorte qu'une commande puisse être totalement pré-remplie.

En prime, idéalement, le robot devrait également indiquer au début de la conversation que l'utilisateur peut l'avertir d'interrompre ses questions : à partir de ce moment, le robot devrait passer à l'étape 5 avec les détails déjà collectés.

Nous prévoyons ce cas d'utilisation, car le robot peut poser de nombreuses questions et l'utilisateur n'a aucune idée du nombre de questions en attente à un moment donné, et nous voudrions nous assurer que l'utilisateur peut indiquer qu'il en a assez de répondre aux questions.

## Histoire 2 :

“From scratch” - the prior use case we would like to focus on re the PoC.

"A partir de zéro" - le premier cas d'utilisation sur lequel nous aimerions nous concentrer dans le cadre du PoC.

Ce deuxième cas d'utilisation "à partir de zéro" est similaire en termes d'objectif, qui est de recueillir les exigences de l'utilisateur relatives à une commande et de résumer les spécifications sous-jacentes, mais sans que l'utilisateur n'engage la conversation par le biais d'un document : il incombe au robot d'engager la conversation et de poser progressivement toutes sortes de questions à l'utilisateur de manière à recueillir tous les détails.

En tant que client,

1 Pour être contesté, il existe probablement de meilleures phrases d'introduction.

1. Je suis connecté à l'application web de la place de marché et j'ai ouvert l'interface utilisateur du bot chat, ou je me connecte à un bot Slack / Teams / WhatsApp ;

2. la conversation commence avec l'émission par le bot d'un premier message du type "Bienvenue. Je vais vous aider à définir vos besoins. Veuillez me dire ce que vous aimeriez commander.2" ;

3. L'utilisateur répond à cette première question ;

4. le robot continue à poser d'autres questions à l'utilisateur dans le cadre d'une conversation, tant qu'il "détecte" qu'il lui manque des détails pour pouvoir résumer les exigences de cette commande au moyen de spécifications, qui sont suffisamment détaillées pour que la commande puisse être exécutée ;

5. une fois que le robot a recueilli suffisamment de détails sur la commande, il conclut la conversation en écrivant une proposition à l'utilisateur avec les spécifications détaillées concernant la commande, à savoir le produit correspondant disponible sur la place de marché qui est concerné et recommandé, la quantité et les caractéristiques aussi détaillées que possible, de sorte qu'une commande puisse être totalement pré-remplie.

Le bonus mentionné s'applique également.

## Data - Données

Afin de construire ce PoC, Altavia fournit 5 fichiers de données, contenant du texte en langue française.

1. "products.json" : un fichier JSON représente le référentiel des produits ;

2. "configurators.json" : chaque produit appartenant à une catégorie, ce fichier JSON contient tous les configurateurs des produits, regroupés par catégories de produits ;

3. "categories.json" : un fichier JSON qui détaille toutes les catégories de produits ;

4. "orders.csv" : un fichier CSV contenant des exemples de commandes sur une période de deux semaines ;

5. **"briefs.csv"** : un fichier CSV contenant des exemples de briefs de clients de la marque. La première colonne est le nom du responsable du projet / succès client d'Altavia, la deuxième colonne le contenu textuel du brief, la dernière colonne est la référence de la commande correspondante dans l'ERP d'Altavia.

Si le partenaire pense que ces données ne sont pas assez qualitatives, pas assez grandes, pas assez complètes, Altavia se réjouit de ce retour d'information afin de fournir des données plus valables et de bénéficier ainsi d'une application plus qualitative.

## Safety

---

### Sécurité

Ce document ne couvre pas certains sujets liés à la sécurité de l'application. Nous attendons de notre partenaire qu'il aborde les cas limites suivants en termes de modifications et d'efforts qui devront être déployés pour les traiter chaque fois que cela est applicable, que ce soit au stade du PoC ou lors de la transition du PoC à l'application prête pour la production.

2 Même remarque que pour la première histoire d'utilisateur.

<sup>2</sup> Same remark as for the first user story.

*les clients : sont des marques et distributeurs*

1. How is the bot supposed to react if the provided file is actually not a brief document? Likewise, what will happen if the user submits another document during the conversation?
2. How is the bot supposed to react when the user asks questions totally out of the scope or purpose of the application? Same concern if the user answers with out-of-scope or irrelevant content.
3. How is the bot going to prevent generating not-safe-for-work content?

If the partner thinks of additional facets to take into consideration in order to design an AI safe solution, it is welcome to raise Altavia's awareness. For instance, those safety perspectives could be tackled when the project transitions from the PoC to the production-ready application.

## Deliverables and deadline

The expected final delivery end date is 8 weeks, from the moment the collaboration starts: following this schedule:

1. to witness a first version after 4 weeks: that version will be demoed by the partner in their own environment, so that Altavia can already have a first glimpse of the bot behavior ;
2. to benefit a first delivery of the API service end-points after 6 weeks, along with some documentation on how to use it, so that Altavia can prepare the integration of the bot within their web application or any instant messaging application, and be able to perform the first integration tests ;
3. to receive a final version of the PoC application after 8 weeks with all the deliverables.

Ideally, Altavia would like to cover the 2 user stories, because they would match more use cases requirements and offer more business flexibility: it is up to the partner to decide whether it makes sense to respond to 1 or to the 2 user stories ("document-based" or "from scratch") depending on the involved effort and budget and to provide guidance and advice regarding the one which makes more sense in terms of quality outcome.

Here is the list of deliverables Altavia expects, for either the "document-based" or the "from scratch" user story or both, given the fact that the partner will have motivated the reasons of their choice in their initial response.

1. The API web service endpoints coordinates if the partner hosts the model runtime.
2. The application source-code, both at the level model and at the API level, and their documentation, their dependencies coordinates, their compilation process, so that Altavia may be autonomous at hosting it in the future.
3. The model tensor files, as well as the coordinates of the base model used if any for fine-tuning.
4. The explanations about the approach: what kind of training has been done, what device has been used, is it a LoRA (in case of model fine-tuning), does it use RAG... ? The training process documentation and training, validation, and test datasets.

5. Some consulting for explaining how the model works and to grow Altavia's expertise and autonomy in the long-term.

Altavia expects the partner to challenge the design of the bot or to ask any questions that will help at improving the PoC application quality.

Since this PoC is part of an innovation initiative, Altavia would also like to know whether the partner is eligible to the French Crédit d'Impôt Recherche (CIR) and Crédit d'Impôt Innovation (CII). In addition, Altavia would like to know whether the partner has been granted some certifications as a training organization (which could offer some benefits to Altavia).

## PoC assessment

Once the PoC application will be available, Altavia will proceed to its assessment. To that purpose, Altavia will be working with the usual project managers, so that they can evaluate the relevance and quality of the bot questions and answers on one side, then the detailed specifications outcome at the end of the conversation. For now, Altavia has not yet specified the way the assessment will be performed, other than asking a panel of users (project managers) to simulate conversations following the two use cases and then rating the outcome. If the partner has some guidance to offer in terms of assessment framework, they are welcome to propose one and to raise Altavia's awareness on that concern.

## Beyond the PoC

Altavia wants to design this PoC bot application in order to subsequently build a production-ready application with the same purpose, provided the PoC demonstrates its relevance.

The collaboration with the partner should span beyond the delivery of the PoC application. In particular, Altavia expects that the work of the PoC be reused for building a production-ready application. Hence, that PoC application should not be disposable, but re-usable instead and extended.

The PoC limits the scope of the handled products to a restricted list of categories: Altavia intends to extend the scope of supported product categories as much as possible to make it as universal as possible, and expects the partner to anticipate that ambition by indicating how much of effort is left in terms of model training and in terms of runtime cost and number of models provided, be it make sense.

In addition, Altavia ultimately needs the production application to tackle more Latin-alphabet-based languages, so that it supports French, English, Spanish, German, Italian (later on Polish, Dutch, Czech). Ultimately, we should support Russian, Ukrainian, but also Asian languages such as Chinese, Japanese, Korean. The partner will raise awareness of any

additional involved effort, feasibility and cost for covering those languages, in particular regarding the non Latin-alphabet-based languages, which may require a specific management.

## Collaboration framework

IA generation is a strategic topic for Altavia: this is the reason why, we are starting initiatives via Proof of Concept projects, and we are open to operate those projects through collaborations with companies with a strong expertise in that field, in order to accelerate and to take benefit of partners with a larger experience.

We also have the ambition to acquire knowledge in that field: this is why we prioritize collaborations with organizations able to make us grow in that field. Not only do we want to benefit from state-of-the-art and relevant PoC applications, we also want to learn from those experiments and to gain more and more autonomy. We hence privilege organizations able to provide guidance and knowledge transfer in the AI field.

We believe that AI generative partners will always be ahead of time in terms of expertise and experience compared to us, and this is why we are confident that, even if partners makes us grow, we will regularly request them, be it on PoC projects, for consulting or for harnessing applications so that they become production ready. We would like to build a trustful, transparent and rich relationship framework with partners with those considerations in mind. We insist on the discretion of our partners when it comes to the privacy of the data we share with them and on the nature of the project involved.

The collaboration framework that we propose highlights 4 main principles that we would like to respect as much as possible.

- **Reversibility:** the access to the source code of the model, along with its tensor files, its training procedure, and its reference models coordinates in case of fine-tuning should be granted. You are free to propose a hosting of the model via web services API, but Altavia wants to ensure that it can operate the hosting, if required.
- **Intellectual property:** the intellectual property of the model source code, the tensor files, and its training data are the full property of Altavia.
- **Data privacy:** the data used to train or fine-tune a model or which transits from or to the AI system is private: in particular, it should not be stored by any third-party other than the partner. The inference data should not be used to fine-tune any third-party solution. In addition, the way the data is stored or is transmitted should be secured.
- **Interoperability:** any model trained should be accessible via API web services endpoints. In case of a LLM-based model, it should be possible to initialize its conversational context via this API and the layout of the web services should be as close as possible as the arising standard such as ChatGPT, Gemini, Claude, for instance adopting the [“Ollama”](#) standard. Those endpoints should be secured and clearly documented.

# NDA

Because of the sensitive nature of the data and of the project, Altavia intends to submit a Non Disclosure Agreement (NDA) to the partner, which should be accepted before starting any collaboration.