

Vers un actuariat agentique ? Graphes de connaissance et agents intelligents en pratique

Le 24/11/2025 à Paris

Swan BROUTARD, Actuaire consultante chez Exiom Partners

Aurélien COULOUMY, Actuaire ML Engineer chez Dylogy, Enseignant chercheur à l'ISFA

Idir HAREB, ML engineer chez Dylogy

EXIOM

Dylogy

ISFA
Institut de Science Financière et d'Assurances
GRANDE ÉCOLE D'ACTUARIAT & GESTION DES RISQUES



Université Claude Bernard Lyon 1



1. Agents et actuariat : définition, bénéfices, organisation

*Définir, borner et architecturer l'usage
des agents actuariels*

1.1 Comprendre le rôle des agents pour mieux assister l'expertise actuarielle

Définitions

Un agent intelligent est un système autonome, doté de la capacité de percevoir, raisonner, et agir sur son environnement pour atteindre un objectif donné.

Observer : il collecte des informations (prompts, API, bases de données, graphes, documents...).



Agir : il interagit avec des outils, exécute du code, requête d'autres données, rédige des rapports, etc.



Raisonner : il planifie, déduit, hiérarchise ses actions (via des modèles LLM).



Apprendre : il évalue et améliore son comportement selon des retours.



L'agent agit comme un assistant cognitif augmenté, capable de reproduire des raisonnements actuariels, d'appliquer des méthodes, et d'automatiser les séquences de travail, tout en restant traçable et contrôlable.

Tarification :

"Analyse le portefeuille Auto, identifie les segments sur- ou sous-tarifés, propose des ajustements tarifaires, calcule l'impact sur la perte attendue."
→ L'agent prépare les données, calcule, compare, recommande.

Réserves / clôture :

"Calcule les triangles, sélectionne les méthodes pertinentes, fournit une synthèse des provisions par branche avec justification méthodologique."
→ L'agent calcule, compare les méthodes, trace les choix.

Réassurance :

"À partir du traité XoL, extrait les clauses clés, simule la distribution des sinistres agrégés et estime le risque de dépassement de la couche."
→ L'agent fait du RAG sur documents, simule, rédige une note synthétique.

Préparation de scénarios :

"Génère une série de scénarios de mortalité, choc pandémie modéré, propose les paramètres (multiplicateurs, lapses), cite les études"
→ L'agent donne les scénarios, vérifie la cohérence, documente les sources.

1.2 Identifier les avantages et borner l'utilisation de ces méthodes en actuariat

Bénéfices clés



Automatisation cognitive

Réduction du temps sur les tâches à faible valeur (data prep, documentation, reporting).



Raisonnement explicable

Les chaînes d'action sont tracées : auditabilité, conformité, justification.



Augmentation du périmètre d'analyse

Possibilité de tester plus de scénarios, de combiner des sources multiples.



Co-pilotage intelligent

Interaction naturelle : l'actuaire exprime une intention, l'agent exécute en dialogue.



Capitalisation du savoir

Les graphes de connaissance servent de mémoire institutionnelle : savoirs, formules, méthodologies, retours d'expérience.

Objectif global : transformer les workflows actuariels en chaînes d'agents orchestrés - fiables, auditées, reproductibles.

Limites à maîtriser



Automatisation sans garde-fou

L'automatisation peut amplifier les erreurs → validations, revues et seuils d'alerte indispensables.



Transparence & explicabilité limitées

Certaines décisions restent difficiles à retracer → importance de chaînes d'action pleinement auditables.



Qualité & biais des données

Les agents héritent des biais des données sources → contrôle qualité renforcé et gouvernance des données.



Variabilité & dépendance à la configuration

Résultats sensibles aux prompts et paramètres → standardisation des workflows et règles d'usage.

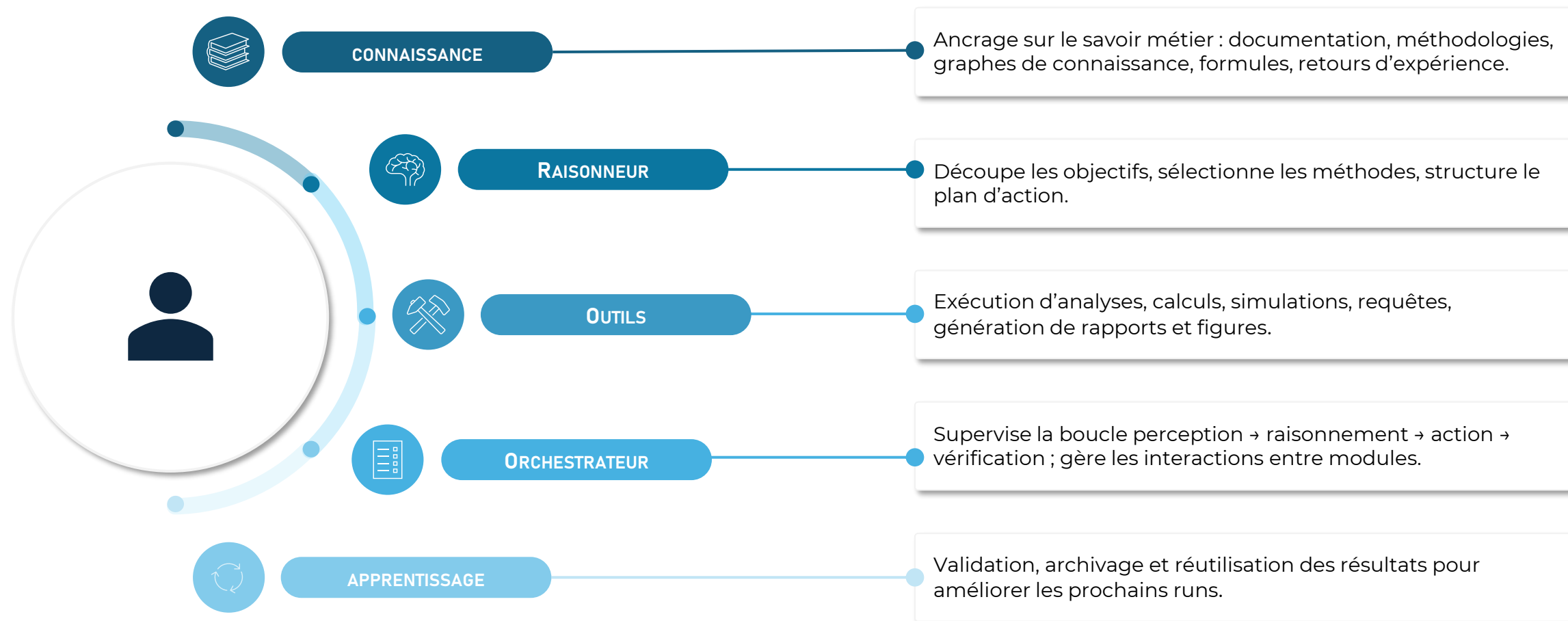


Sécurité & responsabilité

Accès à des données critiques → protection renforcée et rôle clair de l'actuaire dans la validation finale.

Point de vigilance : garantir un usage maîtrisé des agents - contrôlé, explicable, sécurisé et validé - afin d'éviter biais, dérives et erreurs automatisées.

1.3 Définir une architecture robuste pour développer un agent (actuariel)



2. Spécifications d'agents actuariels

Paramétrer et détailler les composants clés associés aux agents actuariels

2.1 Architecture détaillée d'agents (actuariels)

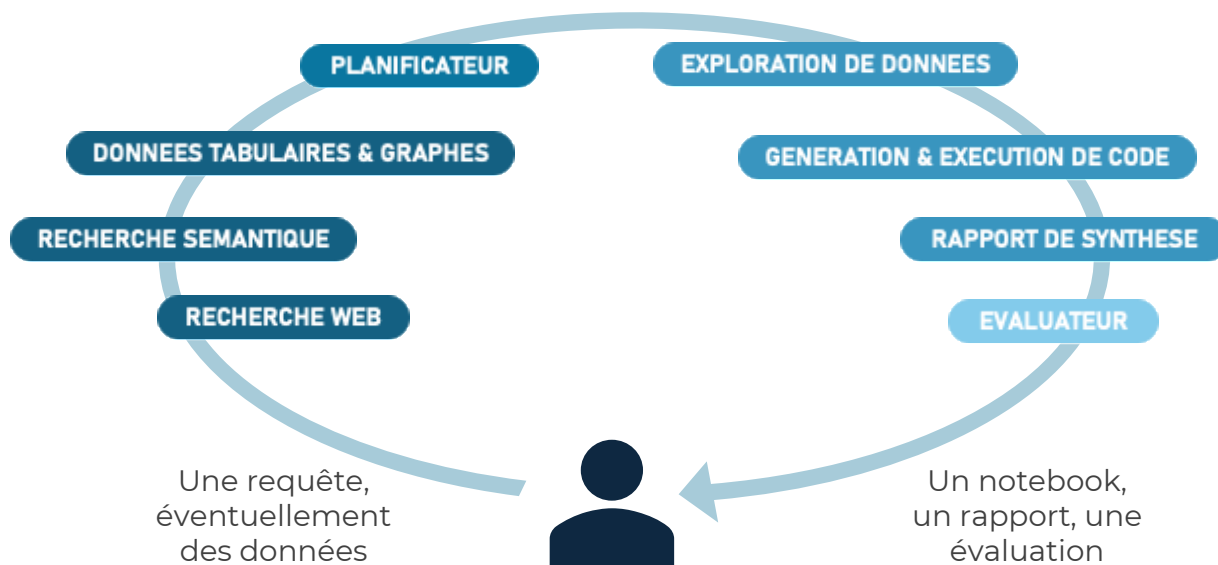
VUE D'ENSEMBLE

Quoi ?

- Créer un système actuarielle autonome qui traite des requêtes en langage naturel, produit des analyses, du code, des visualisations et des rapports.

Comment ?

- Basé sur LangGraph pour créer des applications multi-agents avec état et flux de contrôle explicite
- Implémente un flux de travail avec nœuds conditionnels, chemin d'exécution, mécanismes de réessaie
- Chaque nouveau nœud reçoit l'information cumulée des précédents et met à jour des critères clés
- Déployé via Chainlit pour l'interface web et partage en temps réel de la progression de l'analyse
- Une suite de LLM (séries Nova et Claude 3.7/4/4.5) opérés via AI gateway qui orchestre, limite, garantie l'usage de des modèles en fonction des usages ou contraintes



2.2 Les composants d'un agent - Connaissance



RECHERCHE WEB

Quoi ?


- Rechercher sur le web les méthodologies actuarielles pertinentes

Comment ?

- Récupère la demande
- Transforme en requête pour le service de recherche
- Modèle LLM : Amazon Nova Pro (eu.amazon.nova-pro-v1:0)
- Retourne 5–10 sources (titre, snippet, url) pour une latence : 3–5 secondes et un coût de 0,0075–0,008 USD.

Pourquoi ?

- **Collecter des éléments de contextes à jour (réglementation, news, code, blog)**


Web Research Findings

The following actuarial knowledge and methodologies were gathered to inform the analysis:

The following actuarial knowledge and methodologies were gathered to inform the analysis:

A Guide to Actuarial Price Modeling: Claim Frequency and Severity ...

<https://medium.com/@1998ameya/a-guide-to-actuarial-price-modeling-claim-frequency-and-severity-modeling-with-glms-f8d33231bc00>

In this guide, I will explore a popular actuarial model — the Generalized Linear Model (GLM) — and walk through an example with real-world data

Tweedie regression on insurance claims - Scikit-learn

https://scikit-learn.org/stable/auto_examples/linear_model/plot_tweedie_regression_insurance_claims.html

```

predict(X_train, y_label = "Claim Frequency", title = "train data", ax = ax[0, 0],) plot_obs_pred(df = df_test, feature = "DriveAge", weight = "Exposure",
Frequency", title = "test data", ax = ax[0, 1], fill_legend = True,) plot_obs_pred(df = df_test, feature = "VehAge", weight = "Exposure", observed =
"test data", ax = ax[1, 0], fill_legend = True,) plot_obs_pred(df = df_test, feature = "BonusMalus", weight = "Exposure", observed = "Frequency",
y_pred_product),("Compound Poisson Gamma", y_pred_total),): cum_exposure, cum_claims = lorenz_curve(df_test["PurePremium"], y_pred, df_
{:.3f})".

```

Modelling Insurance Claims Data Using the Tweedie Approach

<https://medium.com/@wardarahim25/modelling-insurance-claims-data-using-the-tweedie-approach-94db8b14bf55>

```

Import the required packagesfrom lightgbm.sklearn import LGBMRegressorfrom sklearn.metrics import mean_tweedie_deviance, make_scorerfrom
GridSearchCVneg_mtd_scorer = make_scorer(mean_tweedie_deviance, greater_is_better=False, power=1.5)# Instantiate the modelmodel = LGBMRegressor(
random_state=1, n_estimators=10000, learning_rate=0.001)# Define the hyperparameter grid (the following three are a good starting point for a
'min_data_in_leaf': np.arange(25, 125, 25), max_depth': np.arange(2, 8, 1),)# Hyperparameter tuning with early stoppingfit_params={"early_stop
KFold(n_splits=5)lgbm_cv = GridSearchCV(lgbm_model, lgbm_params, cv=cv.split(X_train), return_train_score=False, verbose=0, n_jobs=-1, sco
**fit_params)

```

[PDF] Frequency and Severity Models | Cambridge Core

https://resolve.cambridge.org/core/services/aop-cambridge-core/content/view/4A11A9A658AE9C795A5B674C7FE794B/9781139342674c6_p13

Chapter Preview. Many insurance datasets feature information about frequency, how often claims arise, in addition to severity, the claim size.

Figure : Exemple de résultats après recherche web

2.2 Les composants d'un agent - Connaissance



RECHERCHE SEMANTIQUE

Quoi ?

- Effectuer des recherches sur la documentation interne et récupérer des extraits pertinents de mémoires d'actuaire via RAG (Retrieval-Augmented Generation)

Comment ?

- +1000 mémoires / 100p en moyenne indexées, chunkées par 400 tokens
- Vectorise via Embedding : multilingual-e5-large-instruct-endpoint
- Base exploitée au moment d'une nouvelle demande utilisateur pour retourner 5-20 sources avec une latence de 1-2 secondes et un coût négligeable.
- Exemples de demandes: « DQ checks with a fix log », « EDA for frequency/severity », « GLM freq+sev and GBM challenger », etc

Pourquoi ?

- **Intégrer des méthodes ou processus internes**

Internal Knowledge Base - Top 5 Relevant Chunks

The following document chunks were retrieved from the internal knowledge base based on semantic similarity to your query:

- #### 1. Modélisation de la Prime Pure en Assurance Multirisque Habitation 0.pdf

Pages: 10

Content:

1. Preparation and data cleansing 2. Pure premium modelling in Emblem² 3. Model validation 4. Impact study The model validation helps us to accept the estimate and validate this new production later. Key words: - Non-life insurance pricing - Direct insurance - Multi-risk Home Insurance (MRH) - Generalized Linear Models
- #### 2. Travel pricing sophistication with GLM and Machine Learning approaches 0.pdf

Pages: 9, 10

Content:

pure premium in use of GLM "Claim Frequency X Claim Severity" and GLM Tweedie (compound Poisson-gamma) to model the pure premium for the GLM approach. To be clearer, we used two different methods to model the pure premium of the selected portfolios, which are the GLM "Claim Frequency X Claim Severity" for the GLM approaches. Then in a second step, in order to model the pure premium of the selected portfolios, we also applied the machine learning methods: tree and GLM with Poisson distribution and logarithm link function is applied on the number of claims of one portfolio to model its claim frequency. As a result, we obtained the frequency estimated by the model for certain one risk profile where its relativity is equal to 1) and relativity referring to different risk profile. So the frequency of each corresponding risk profile relativity. The frequency relativity table of B2B
- #### 3. Sophistication de la tarification du risque catastrophe naturelle 0.pdf

Pages: 9

Content:

appreciation of the potential explanatory variables re-activities. In order to eliminate any bias inherent in the Mexican portfolio's exposure data, a notional portfolio all geographic areas is constructed. The second step is to calibrate frequency and severity using parametric models such as GLMs and other machine learning methods. This step involves applying discount curves to reflect financial conditions. This approximation can be quite rough, when considering a limited number of discount curves (multi-layering, reinsurance facultative, min / max deductibles).
- #### 4. Tarification d'un produit Multirisque Professionnelle 0.pdf

Pages: 69

Content:

of databases (contracts, clients, accidents, external) for many years and including many AXA France regions in the scope. One of the main hurdles will then be to overcome this constraint, we chose to create a "R Markdown" code. The key advantage of this format is the associated structural constraint the user needs to acknowledge: The code contains the description of the results. We then have the option to re-use this R-Markdown code for other guarantees models by only having to change formulas, it's a semi-automated modelling of this gross premium, which is being possible thanks to the standardized guarantees database format. We are modelling the gross premium following the multiplication between the estimated frequency and the average cost planned in case of an accident. Variables are selected by GBM using the relative importance. Among these selected variables, we need to avoid redundancy of information

Figure : Exemple de résultats après recherche sémantique

2.2 Les composants d'un agent - Connaissance



DONNEES TABULAIRES & GRAPHES

Quoi ?

- Charger des données nécessaires aux calculs actuariels

Comment ?

- Pour la partie tabulaire : simple récupération d'un fichier csv ou xls
- Pour les graphes: récupération de pdf (type rapports sinistres) puis transformation en graphes de connaissance via modèles Claude Sonnet 4, ontologie et référentiels dédiés

Pourquoi ?

- Disposer d'informations structurées pour opérer les calculs**

policy_id	age	gender	vehicle_age	vehicle_power	region	annual_mileage	use	bonus_malus	prior_claims_count	exposure	premium_v
1	23	M	16	Low	IDF	13039	Commute	1.04	1	0.66	565.410000
2	69	F	14	Low	IDF	16100	Private	0.97	0	0.75	527.89
3	61	F	11	Low	IDF	11344	Private	0.92	0	0.57	404.94
4	47	F	17	Medium	Sud-Est	9676	Commute	0.96	0	0.41	393.07
5	47	M	2	Medium	IDF	3000	Commercial	0.77	1	0.9	667.01
6	75	M	18	Low	Sud-Ouest	14564	Private	1.03	0	0.51	370.99
7	23	F	17	Low	Sud-Est	11547	Private	0.73	0	0.94	515.63
8	64	M	10	High	Sud-Ouest	7949	Private	1.03	0	0.41	360.37
9	31	M	14	Low	Est	12506	Private	1.08	0	0.67	417.5
10	24	M	10	Medium	Centre	9665	Commute	0.93	0	0.65	379.19
11	53	F	5	Medium	IDF	9886	Private	0.87	0	0.87	472.52
12	83	F	15	Low	Ouest	14425	Commute	0.96	0	0.35	327.74
13	67	F	4	Low	Nord	6646	Commute	0.5	0	0.93	399.05

Figure : Exemple de jeu de données de tarification

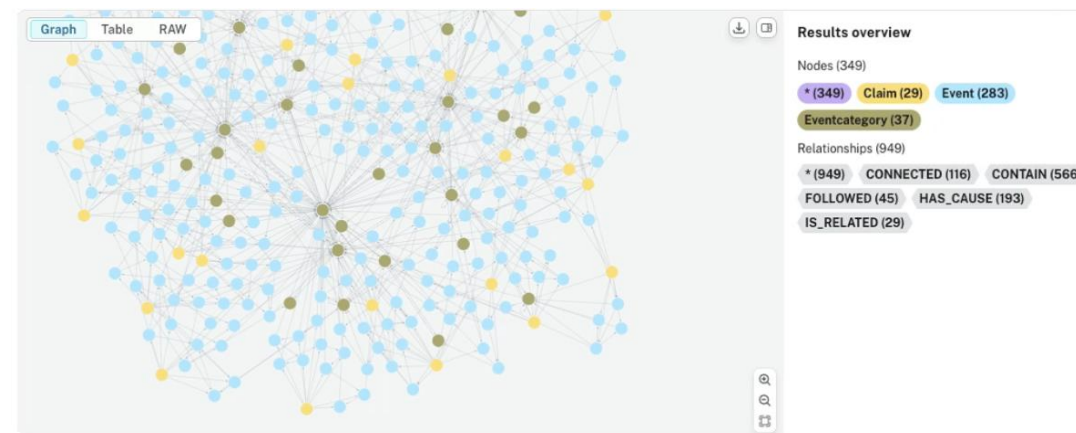


Figure : Exemple de graphe de connaissance sinistre

2.3 Les composants d'un agent - Raisonneur



PLANIFICATEUR

Quoi ?

- Créer un plan d'analyse actuarielle étape par étape

Comment ?

- Prend tout le contexte recueilli (recherches web, documents internes, détails sur les données, etc.) et génère un plan détaillé
- Couvre les aspects de préparation des données, sélection de modèles, sélection de variables, validation, l'interprétation et la visualisation
- Définit également la compétence nécessaire à l'évaluation
- Le tout opéré par un LLM: Claude Sonnet 4 avec une Latence de 15-30 secondes et un Coût de 0,02-0,04 USD.

Pourquoi ?

- **Créer un plan transparent avant toute programmation et garantir des compétences adaptées pour l'évaluation des résultats**

Plan: Data Quality → EDA → Methodology → Features → Modeling → Interpretation → Visualization + selected skill file path

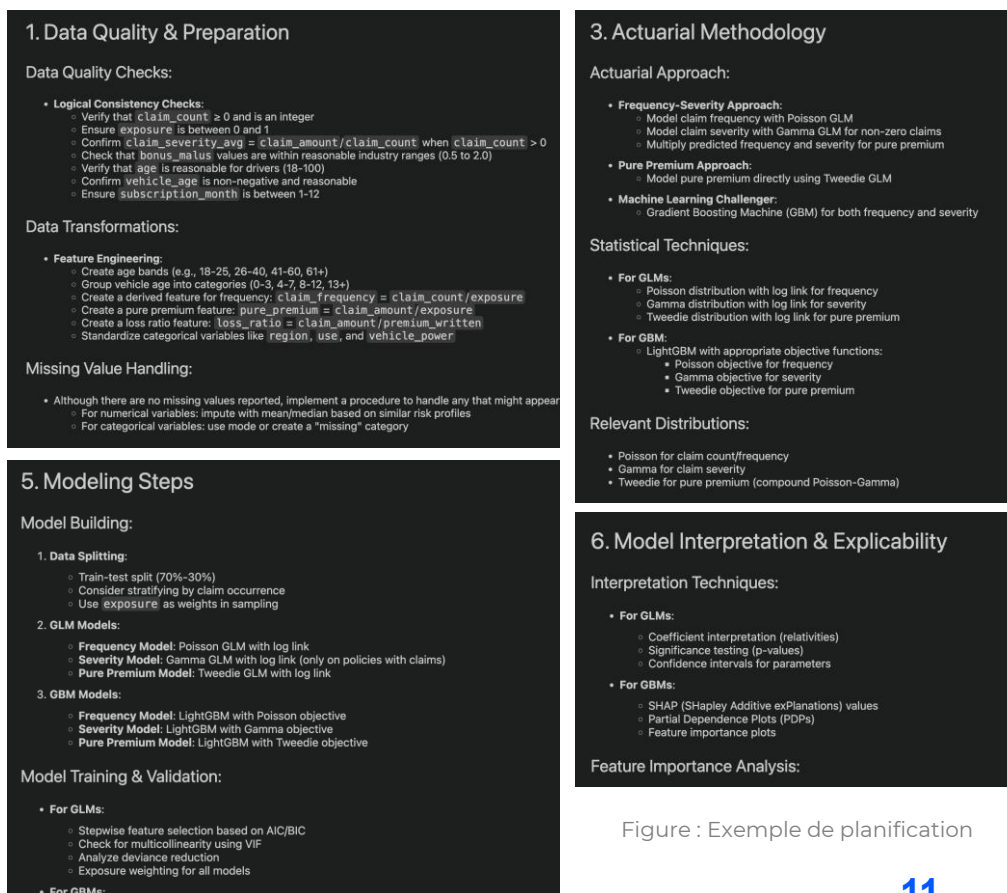


Figure : Exemple de planification

2.4 Les composants d'un agent - Outils



EXPLORATION DE DONNEES

Quoi ?

- Inspecter l'ensemble des données chargées (en csv / xlsx ou JSON pour les graphes)

Comment ?

- Analyse purement Python, sans IA
- Pour le tabulaire: retourne des statistiques descriptives (moyenne, médiane, quartiles, distributions) et dtypes en dictionnaire
- Pour les graphes: retourne également les nombres de nœuds, d'arrêtes, les distributions d'attributs de nœuds et arrêtes, etc.
- Les résultats sont intégrés à la phase de planification afin de garantir que le code correspond à vos données

Pourquoi ?

- Comprendre et résumer les caractéristiques des jeux de données

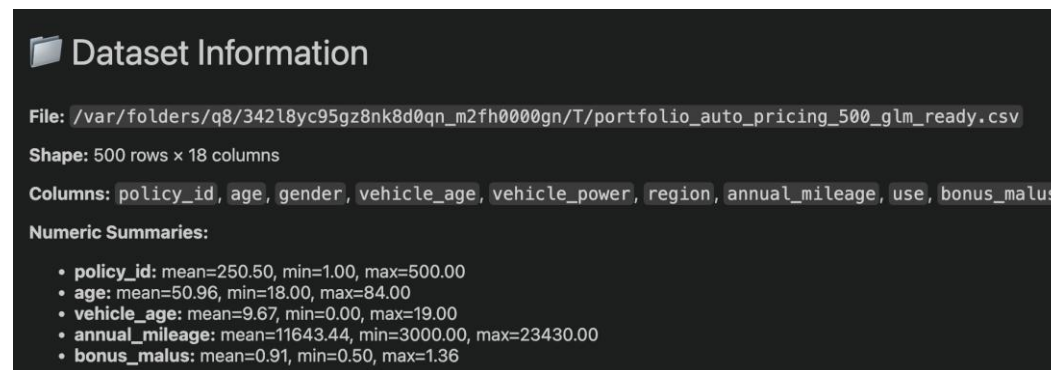


Figure : Analyse de structure de données tabulaires

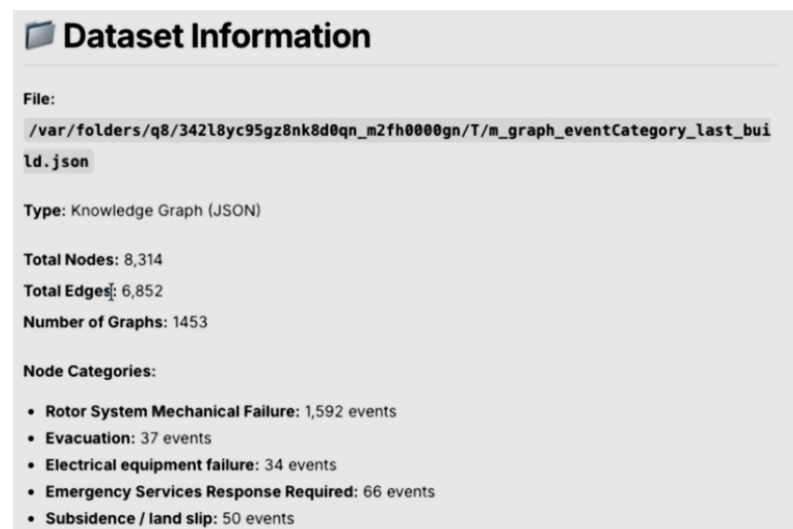


Figure : Analyse de structure de données en graphe

2.4 Les composants d'un agent - Outils



GENERATION & EXECUTION DE CODE

Quoi ?

- Traduire le plan en projet python complet
- Exécuter le code généré dans un environnement sécurisé

Comment ?

- Code tout ce qui est indiqué dans le plan, exécute et si cela échoue analyse l'erreur pour générer un code corrigé (3 fois)
- Utilise Claude Sonnet 4.5 pour un script de 16k tokens max (500–800 lignes)
- Opéré dans un environnement isolé bac à sable, avec runtime séparé, espace mémoire et log dédiées
- Après run, récupère les sorties, les résultats et les messages d'erreur, ainsi que les fichiers générées (visualisations en png par exemple)
- Contraintes : packages limités, pas d'accès réseau / internet, timeout à 4min.

Pourquoi ?

- Réaliser en toute sécurité les opérations mathématiques du plan

```
# =====
# GLM FREQUENCY MODEL
# =====
print("\n--- GLM FREQUENCY MODEL (Poisson) ---")

# Prepare data for GLM
df_train_freq = X_train.copy()
df_train_freq['claim_count'] = y_freq_train.values
df_train_freq['exposure'] = exp_train.values
df_train_freq['log_exposure'] = np.log(df_train_freq['exposure'])

# Build formula
formula_freq = 'claim_count ~ ' + ' + '.join(feature_cols_encoded)

# Fit Poisson GLM
try:
    glm_freq = smf.glm(formula=formula_freq,
                       data=df_train_freq,
                       family=sm.families.Poisson(),
                       offset=df_train_freq['log_exposure']).fit()
```

Figure : Extrait de codes générées

```
3. PREDICTIVE MODELING
=====
--- FEATURE SELECTION ---
All columns: ['policy_id', 'age', 'gender', 'vehicle_age', 'vehicle',
'claim_count', 'claim_severity_avg', 'claim_amount', 'vehicle_color',
'mileage_band']
Selected features for modeling: ['age', 'gender', 'vehicle_age', 've',
'subscription_month']
Excluded: policy_id (ID), dealer_postal_code (ID), claim_severity_avg

Training set: 400 policies
Test set: 100 policies

--- GLM FREQUENCY MODEL (Poisson) ---
✓ GLM Frequency Model Summary:
AIC: 386.88
BIC: -2134.35
Deviance: 190.34

Significant coefficients (p < 0.05):
vehicle_power_encoded: coef=-0.4176, p=0.0337, relativity=0.6586
```

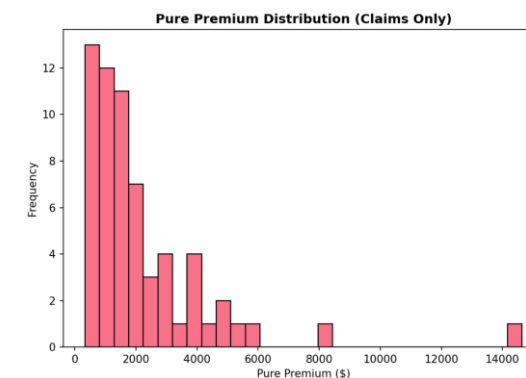


Figure : Extrait d'ouputs après exécution du code

2.4 Les composants d'un agent - Outils



RAPPORT DE SYNTHESE

Quoi ?

- Rédiger un rapport actuariel qui résume les opérations et résultats

Comment ?

- Récupère tous les éléments associés aux travaux (sources, données, plan, code et output, graphiques) et présente des explications suivant une structure établie et une tonalité claire (méthodes, découvertes, interprétations, recommandations)
- Cite tous les éléments exploités dans le rapport
- Utilise Claude Sonnet 4 pour obtenir un rapport en 20 à 40 secondes pour 0,03–0,06 USD.

Pourquoi ?

- **Produire un rapport synthétique et sourcé, associé à la requête initiale**

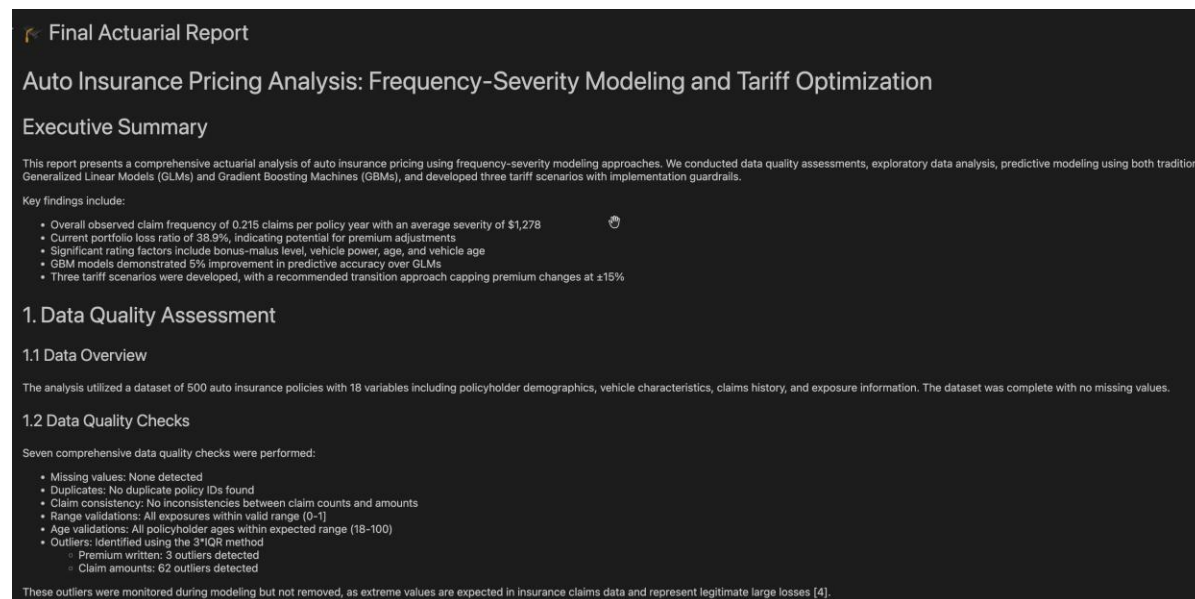


Figure : Exemple de rapport de synthèse pour une tâche de tarification actuarielle

2.4 Les composants d'un agent – Apprentissage et évaluation



EVALUATEUR

Quoi ?

- Evaluer automatiquement la qualité des travaux réalisés

Comment ?

- A partir du répertoire de contenant plusieurs fichiers .md d'évaluateurs (tarificateur, équipe provisionnement, spécialiste en gestion sinistres, data analyste, etc.), évalue le rapport, le code, les résultats.
- Retourne suivant une structure et méthode de scoring fixée des feedbacks experts (sur les modèles, la conformité, la qualité de code, la documentation)
- Processus long opéré par un LLM Claude Sonnet 4 qui prend 30–60 s. pour un Coût de 0,03–0,05 USD permettant d'établir un rapport de contre expertise

Pourquoi ?

- Assurer la qualité, la cohérence, l'amélioration continue des travaux produits par l'agent au regard des pratiques expertes internes

Actuarial Pricing Skills Assessment – skills.md

This file defines a **ready-to-paste prompt** to assess actuarial pricing work (reports, notebooks, slide decks, code, documentation, etc.).

Use it whenever you want to evaluate how well someone has carried out a non-life (or life) actuarial pricing study, especially when machine learning or advanced models are involved.

1. How to use this file

1. **Upload all relevant files:** pricing report(s), technical appendix, slide decks, data dictionaries, notebooks, and validation docs.
2. Call your LLM using the **Assessment Prompt** below.
3. Adapt bracketed parts as needed.
4. Optionally use the **Scoring Rubric** and **Checklists**.

Focus on **methods, assumptions, results, limitations & ethics** — not code style.

2. Assessment Prompt (copy-paste into your LLM)

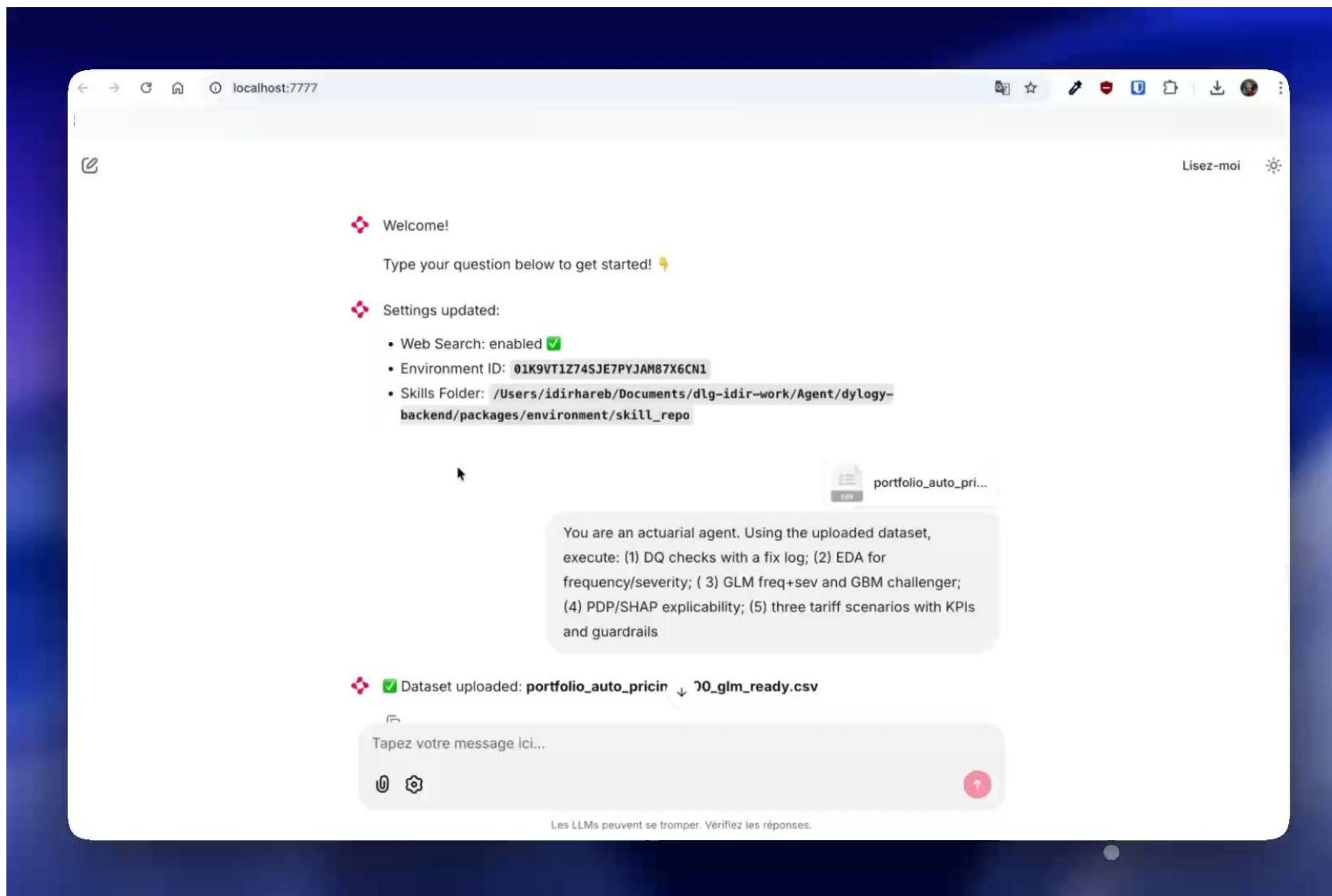
SYSTEM / ROLE

Figure : Exemple de « fiche expert » pour une évaluation d'une étude de tarification

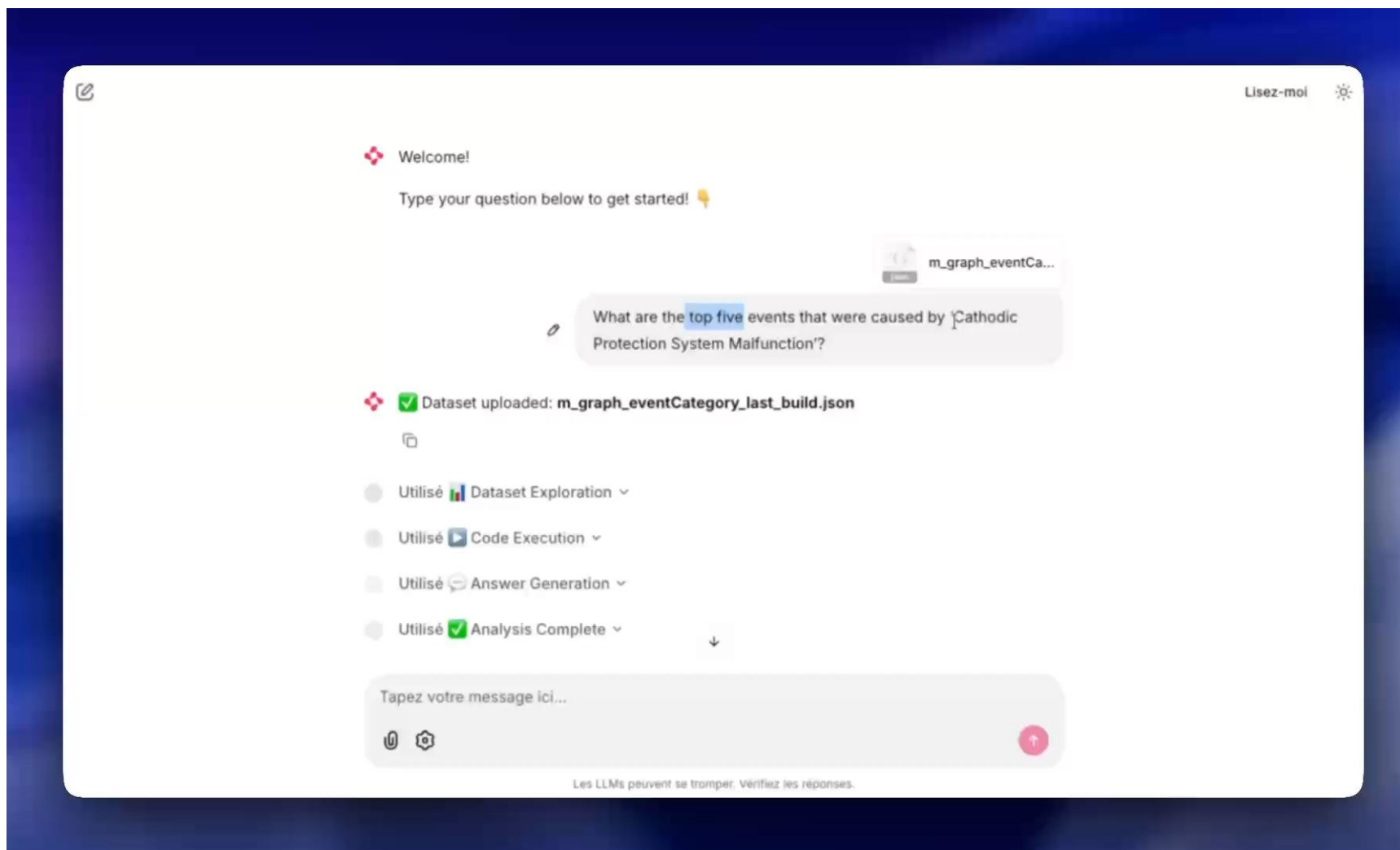
3. Expérimentations et exemples en pratique

Observer et comprendre les résultats associés aux agents actuariels

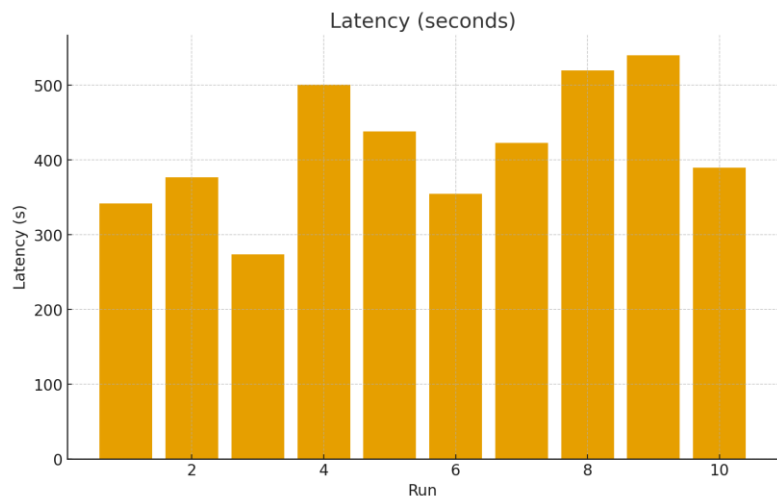
3.1 Démonstration sur un cas de tarification non-vie



3.2 Démonstration sur un cas de gestion sinistres

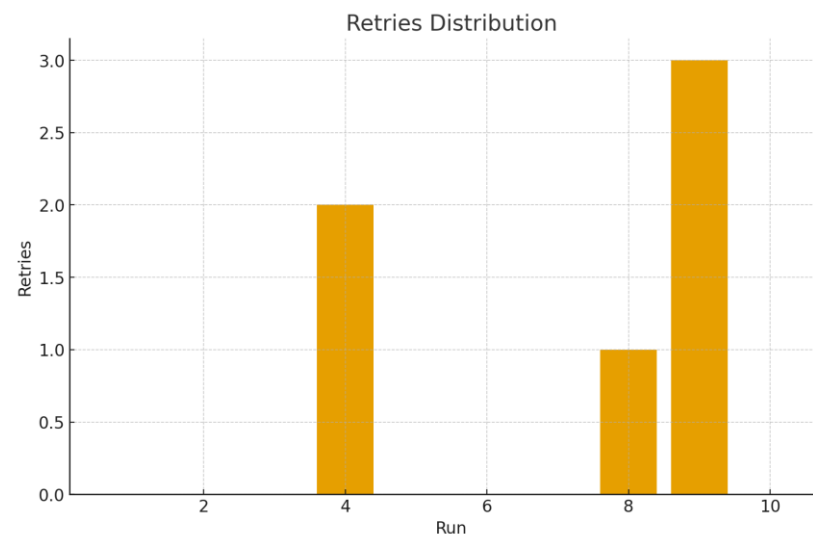
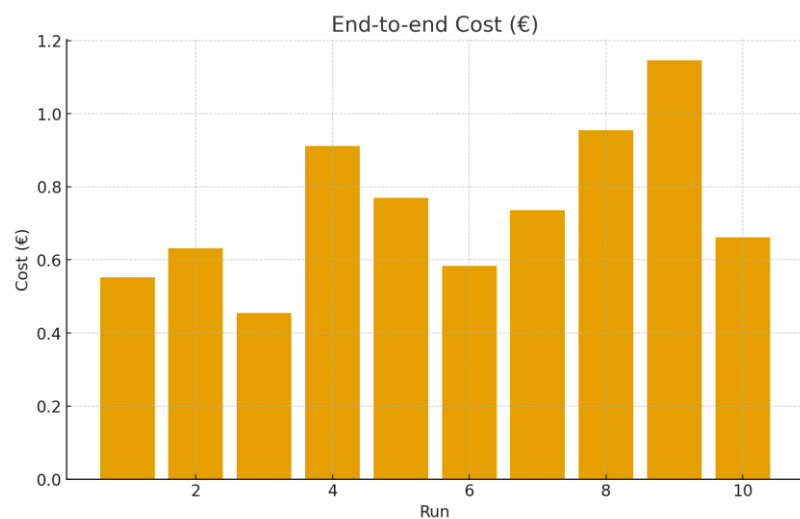


3.3 Résultats complémentaires et sensibilités



Autres éléments :

- Coût total : 0,40–1,10 USD (avec 1 retry et sans considération des préparations documentaires).
- Temps total : de 1–3 minutes (cas simples) à 6–8 minutes (cas complexes)



Résultats concernés	Incertitude
Execution Results	37%
Visualizations	28%
Web Research Findings	8%
Internal Knowledge Base	8%
Code Implementation	8%
Actuarial Plan	7%
Final Actuarial Report	6%
Skills Evaluation Report	5%
Dataset Information	1%

Figures : Tests de sensibilité sur le temps, les couts et les retry (pour 10 expérimentations)

3.3 Résultats complémentaires et sensibilités

Skill Evaluation Radar — Unified Style

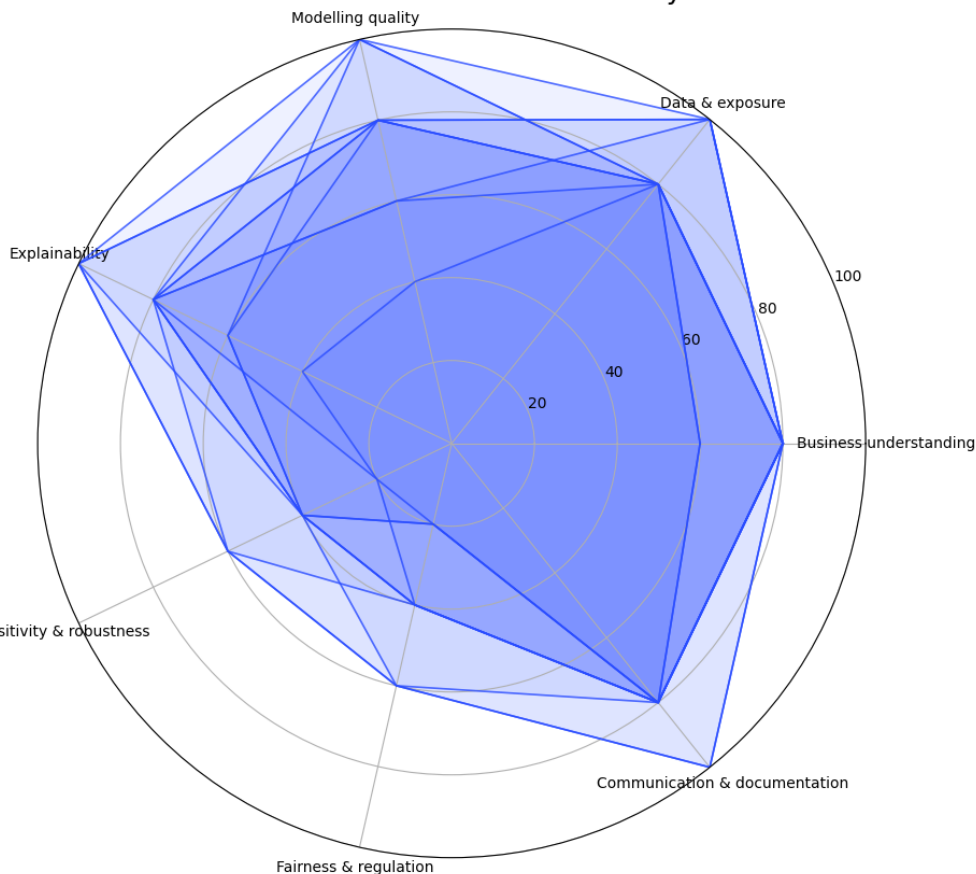
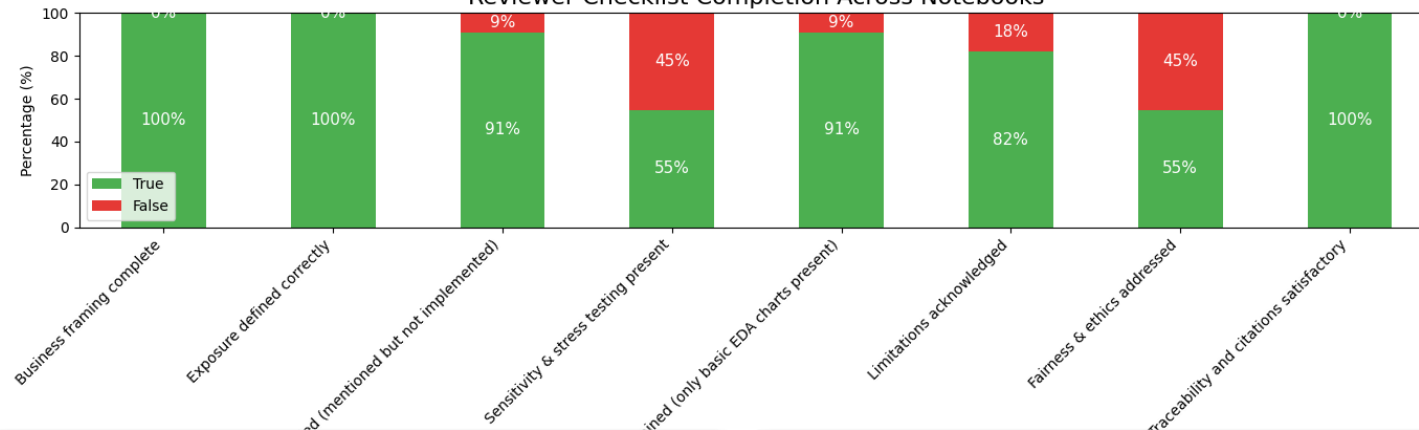


Figure : Radar d'évaluation unifié pour toutes les expérimentations

Reviewer Checklist Completion Across Notebooks



=== Fairness & regulation ===

- *Strengths: Basic segmentation and guardrails are in place, with some attention to premium impacts and risk-based pricing fairness.*
- *Weaknesses/recurring issues: Lacks formal fairness metrics, proxy discrimination testing, and explicit regulatory compliance assessment; limited consideration across protected characteristics appears repeatedly across reviews.*
- *Overall maturity: early-stage—foundational steps exist, but a structured fairness framework and clear regulatory alignment are absent.*

=== Sensitivity & robustness ===

- *Strengths: basic tariff/pricing scenario checks exist.*
- *Weaknesses: no formal sensitivity analysis or stress testing; no scenarios for inflation, mix shifts, or environmental factors; no assessment of model stability across segments.*
- *Overall maturity: low/early-stage; needs a structured sensitivity/stress-testing framework with broader scenario coverage and stability analysis.*

4. Conclusions et perspectives

Résumer et compléter les conditions pour déployer des agents sûrs, explicables et utiles

4.1 Ce qu'il faut retenir pour exploiter et contrôler ces agents

+ APPORTS DÉMONTRÉS

Exécution complète, contrôlée, auditée des workflows actuariels

Data prep → modèle → diagnostics → explicabilité → sensibilités → livrables.

Alignement strict aux domaines métier

cohérence, stabilité, calibrations, documentation, opposabilité.

Livrables professionnels et rejouables

notebook exportable, note technique PDF, Excel métier, PPT

Traçabilité intégrale

données, prompts, code, modèles, décisions → reproductibilité multi-run

⚠ RISQUES IDENTIFIÉS

Dérive actuarielle : segmentation inappropriée, variables mal spécifiées, fairness

Incohérence multi-exercice : instabilité des résultats d'un jour à l'autre

Interprétation erronée des sensibilités

Résultats conformes techniquement ≠ pertinents métier

Perte de savoir-faire

SÉCURITÉ & GOUVERNANCE : LES EXIGENCES D'UN ENVIRONNEMENT ACTUARIEL

SÉCURISATION TECHNIQUE

Whitelists strictes (packages, versions, dépendances gelées)

Sandbox d'exécution → aucun accès SI/web non autorisé

Validation avant/après tool-calls : cohérence inputs/outputs, timeout, stabilité technique

Détection d'effet de bord: dérive & biais divers

Versioning complet : données, scripts, modèles, évaluations et tout autre livrables

SÉCURISATION ACTUARIELLE








Hypothèses explicites & contrôlées : segmentation, exclusions, transformations

Guardrails métier : bornes tarifaires, règles IFRS17/S2, limites techniques

Contrôles de stabilité, raisonnabilité, cas de bord, scénarios

Opposabilité réglementaire : chaînage d'action transparent, explicabilité normative

4.2 Pourquoi ne pas se contenter de Copilot / ChatGPT

	ChatGPT	Agent actuariel
 Connaissance	Limitée (via fichiers plats)	Accès aux bases métier, graphes de connaissance
 Planification	Limitée et générique	Raisonnement + plan d'action spécialisé
 Outils	Etendus mais déliés du métier	Pré contexte pour code, requêtes, calculs, vérifications
 Sécurité	Standard	Renforcée: Sandbox + whitelists + guardrails
 Traçabilité	Limitée (Conversation)	Logs détaillée, explications, versioning
 Livrables	Texte, fichiers plats à la demande	Etendus et maîtrisé : Notebook / pdf / excel / ppt
 Eval & Audit	(Quasi) Impossible	LLM as a judge, Reproductible et mesurable

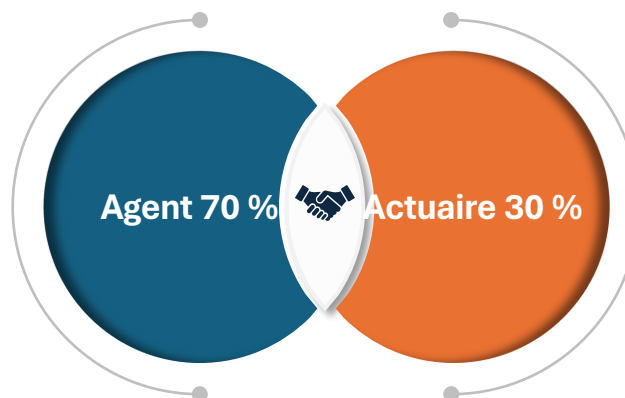
ChatGPT/CoPilot “répond”. L’agent “travaille” : il prépare, exécute, vérifie, documente et respecte les contraintes d’entreprise.

4.3 Le rôle de l'expert dans tout ça ?

OPERATEUR

Ce que l'agent réalise

- Traitement des données (qualité, outliers, dérive)
- Construction automatique des modèles
- Diagnostics complets et explicabilité
- Stabilité & robustesse
- Sensibilités techniques
- Reporting et documentation
- Traçabilité, vérifiabilité, reproductibilité



DECIDEUR

Ce que l'actuaire conserve

- Cadre : périmètres, segmentation, exclusions, objectifs
- Analyse avancée : compréhension des effets, cohérence business
- Arbitrages économiques, prudentiels, réglementaires
- Gestion des cas atypiques : extrêmes, ruptures, signaux faibles
- Validation interne & conformité (IFRS17/S2, gouvernance modèle)
- Décision finale : go/no-go tarifaire, limites, stratégie

L'agent exécute. L'actuaire supervise, interprète, décide et porte la responsabilité.

Evaluez cet atelier

