

Architecture Générale

Plain Text

Événement/Loi/Alerte Externe

↓

LLM 1 - Pertinence Checker

"Est-ce que cela concerne notre entreprise?"

→ NON → Rapport rejeté
→ OUI ↓

LLM 2 - Risk Analyzer

"Quels sont les impacts et recommandations?"

↓

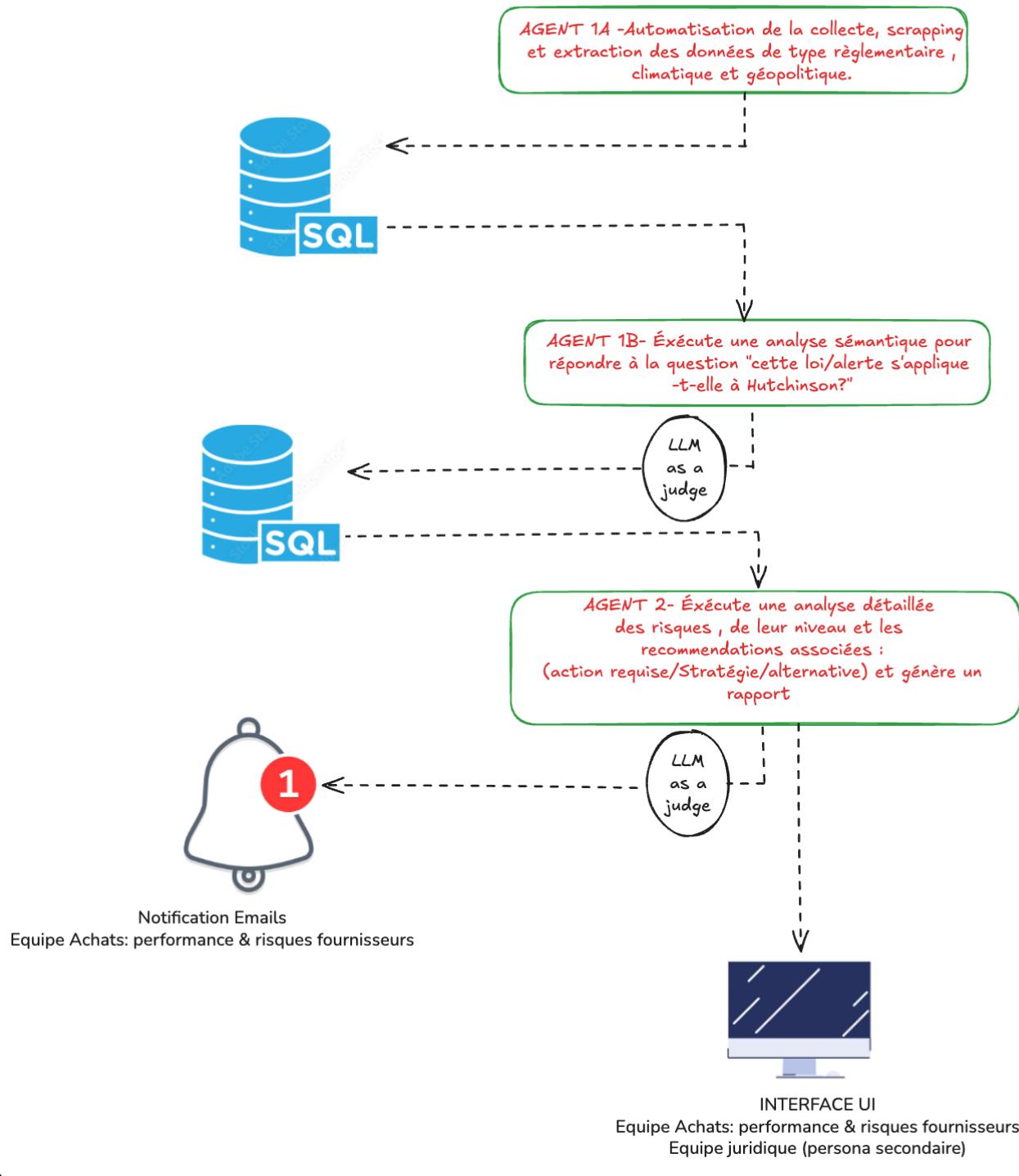
LLM Judge - Évaluation de Qualité

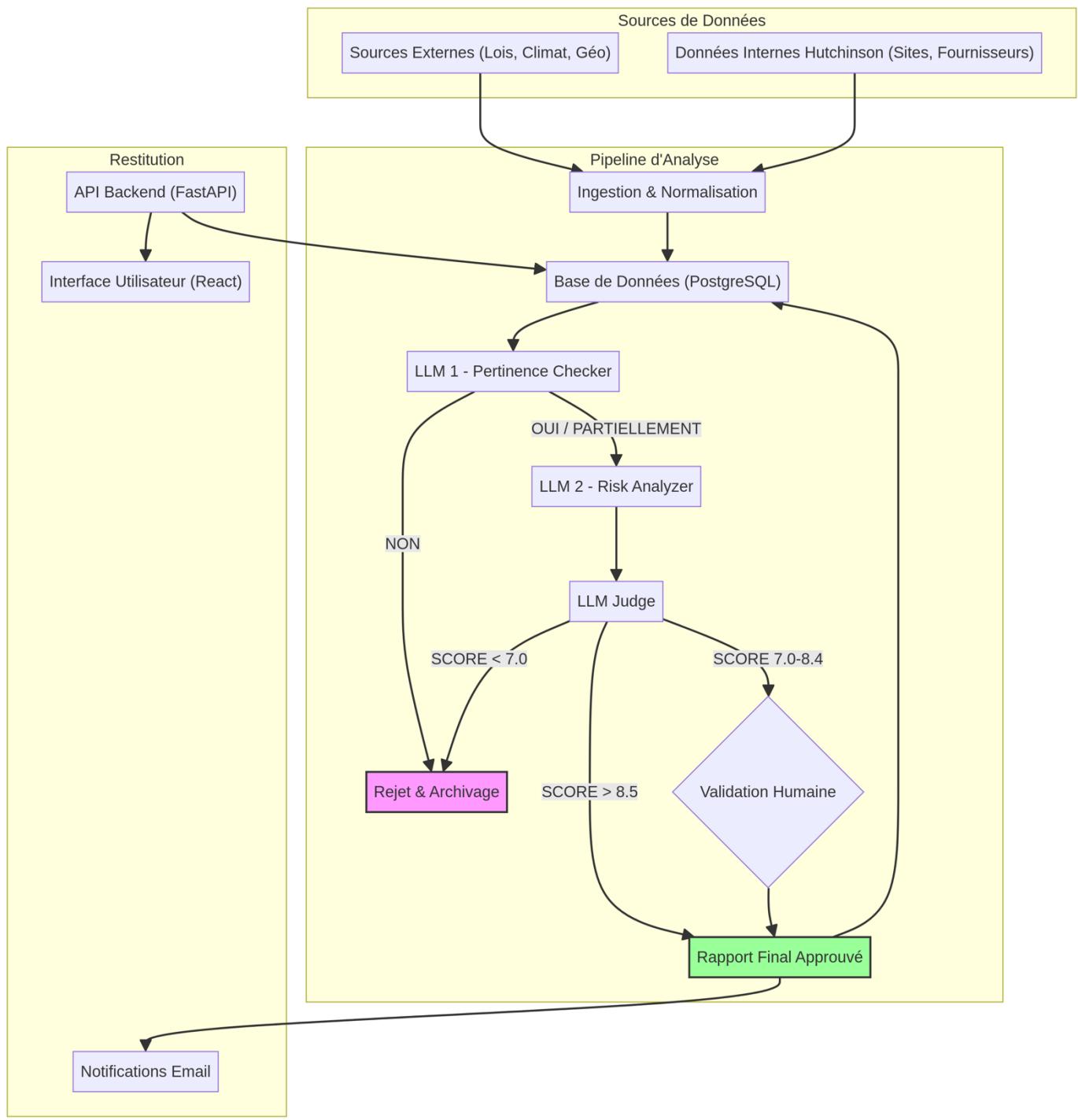
"Ce rapport est-il de bonne qualité?"

↓

Rapport Final

ARCHITECTURE DE LA SOLUTION





Agent 1A:

Objectif :

Automatisation de la collecte , du scrapping et l'extraction des données de type règlementations, climatiques , géopolitiques

Pilier 1 : LLM Pertinence Checker = Agent 1b

Objectif

Déterminer si un événement/loi/alerte externe s'applique réellement à l'entreprise Hutchinson, en croisant les données de la source avec les données internes de l'entreprise.

Entrées

1. **Données Externes** : Le texte de la loi, l'alerte climatique, l'actualité géopolitique, etc.
2. **Données Internes** : Sites Hutchinson, secteurs d'activité, produits, matières premières, pays d'approvisionnement, fournisseurs critiques.

Processus

Le LLM Pertinence Checker exécute une analyse sémantique pour répondre à la question : "**Cette loi/alerte s'applique-t-elle à Hutchinson ?**" Pour cela, il :

1. Extrait les critères d'application de la source (ex: "CBAM s'applique aux importations de caoutchouc en provenance de pays tiers").
2. Croise ces critères avec les données internes (ex: "Hutchinson importe du caoutchouc d'Indonésie").
3. Conclut avec un score de pertinence : **OUI, NON, ou PARTIELLEMENT**.

Structure de la Sortie

JSON

```
{  
  "event_id": "EVT-2024-001",  
  "event_type": "Réglementaire",  
  "event_title": "CBAM - Carbon Border Adjustment Mechanism",  
  "pertinence_check": {  
    "applies_to_company": true,  
    "confidence_score": 0.95,  
    "reasoning": "CBAM s'applique aux importations de caoutchouc. Hutchinson importe du caoutchouc naturel de l'Indonésie et le revende à ses clients.",  
  },  
  "traceability": {  
    "source_passages": [  
      {  
        "source": "EUR-Lex - CBAM Regulation, Article 2",  
        "passage": "The Carbon Border Adjustment Mechanism applies to imports of natural rubber and other products.",  
        "url": "https://eur-lex.europa.eu/..."  
      }  
    ],  
    "company_data_used": [  
      {  
        "data_type": "Matière Première",  
        "value": "Caoutchouc naturel",  
        "source_site": "HUT-FR-001"  
      },  
      {  
        "data_type": "Fournisseur",  
        "value": "Supplier_XYZ (Indonésie )",  
        "relevance": "Pays tiers concerné par CBAM"  
      }  
    ]  
  }  
}
```

Critères de Décision

Résultat	Condition	Action
OUI	Confidence Score > 0.80	Passer au LLM Risk Analyzer
PARTIELLEMENT	Confidence Score 0.50-0.80	Passer au LLM Risk Analyzer avec flag "À vérifier"
NON	Confidence Score < 0.50	Rejeter et archiver

Pilier 2 : LLM Risk Analyzer = Agent 2

Objectif

Analyser en détail les impacts réels de l'événement/loi sur l'entreprise et proposer des recommandations concrètes. Ce LLM ne s'exécute que si le LLM Pertinence Checker a conclu "OUI" ou "PARTIELLEMENT".

Entrées

1. **Données Externes** : Le texte complet de la loi/alerte.
2. **Données Internes** : Toutes les données de l'entreprise (sites, fournisseurs, chaîne d'approvisionnement).
3. **Résultat du LLM Pertinence Checker** : La conclusion précédente et la traçabilité.

Processus

Le LLM Risk Analyzer exécute une analyse détaillée pour répondre aux questions :

1. **Quels sont les impacts spécifiques ?** (géographiques, sectoriels, opérationnels)
2. **Quels sites sont impactés ?**
3. **Quels fournisseurs sont impactés ?**
4. **Quel est le niveau de risque ?** (Faible, Moyen, Fort, Critique)
5. **Quelles sont les recommandations concrètes et détaillées (actions requises, stratégies à mettre en place pour anticipation, alternatives?)**

Structure de la Sortie

JSON

```
{  
  "event_id": "EVT-2024-001",  
  "risk_analysis": {  
    "risk_type": "Réglementaire",  
    "risk_level": "Fort",  
    "description": "La loi CBAM impose une taxe carbone sur les importations de",  
    "impacts": {  
      "geographic": "Affecte les importations en provenance d'Indonésie, d'Inde",  
      "sectoral": "Affecte le secteur des élastomères et de la chimie",  
      "operational": "Augmentation des coûts d'approvisionnement, nécessité de",  
    },  
    "affected_sites": [  
      {  
        "site_id": "HUT-FR-001",  
        "site_name": "Hutchinson France - Clermont-Ferrand",  
        "impact_level": "Fort",  
        "reason": "Site qui utilise du caoutchouc importé"  
      }  

```

```

} ,
{
  "priority": "Haute",
  "action": "Évaluer les alternatives de fournisseurs en zones UE pour réduire les coûts",
  "timeline": "Court terme (3 mois)"
},
{
  "priority": "Moyenne",
  "action": "Réévaluer les contrats fournisseurs pour intégrer les coûts et les risques",
  "timeline": "Moyen terme (6 mois)"
}
]
},
"traceability": {
  "source_passages": [
    {
      "source": "EUR-Lex - CBAM Regulation, Article 5",
      "passage": "The carbon price shall be calculated on the basis of the weekly average price of allowances on the EU ETS market",
      "url": "https://eur-lex.europa.eu/..."
    }
  ],
  "company_data_used": [
    {
      "data_type": "Matière Première",
      "value": "Caoutchouc naturel",
      "source_site": "HUT-FR-001",
      "relevance": "Matière première importée affectée par CBAM"
    },
    {
      "data_type": "Fournisseur",
      "value": "Supplier_XYZ (Chine), Supplier_002 (Indonésie)",
      "relevance": "Fournisseurs critiques en pays tiers"
    }
  ]
}
}

```

Pilier 3 : LLM as a Judge - Évaluation de Qualité

Objectif

Évaluer la qualité des analyses produites par les deux LLM précédents (Pertinence Checker et Risk Analyzer).

Processus

Le LLM Judge évalue deux niveaux :

- Évaluation du Pertinence Checker** : La conclusion de pertinence est-elle correcte et bien justifiée ?
- Évaluation du Risk Analyzer** : Les impacts et recommandations sont-ils pertinents, complets et appropriés ?

Critères d'Évaluation

Critère	Description	Score
Source Relevance	La source citée est-elle fiable et pertinente ?	0-10
Company Data Alignment	Les données internes sont-elles correctement interprétées ?	0-10
Logical Coherence	La conclusion découle-t-elle logiquement des preuves ?	0-10
Completeness	L'analyse couvre-t-elle tous les aspects importants ?	0-10
Recommendation Appropriateness	Les recommandations sont-elles concrètes et applicables ?	0-10
Traceability	Chaque affirmation est-elle tracée à une source ou une donnée ?	0-10

Structure de la Sortie

JSON
<pre>{ "event_id": "EVT-2024-001", "judge_evaluation": { "pertinence_checker_evaluation": { "source_relevance": { "score": 9, "comment": "Source officielle et directe" }, "company_data_alignment": { "score": 8, "comment": "Données correctement interprétées" }, "logical_coherence": { "score": 9, "comment": "Conclusion bien justifiée" }, "traceability": { "score": 10, "comment": "Traçabilité complète" } }, "risk_analyzer_evaluation": { "impacts": { "score": 7, "comment": "Impact significatif et bien documenté" }, "recommendations": { "score": 8, "comment": "Recommandations concrètes et applicables" } } } }</pre>

```

    "source_relevance": { "score": 8, "comment": "Passages sources bien sélec
    "company_data_alignment": { "score": 9, "comment": "Impacts correctement i
    "logical_coherence": { "score": 8, "comment": "Recommandations cohérentes
    "completeness": { "score": 7, "comment": "Pourrait inclure une analyse fin
    "recommendation_appropriateness": { "score": 9, "comment": "Recommandation
    "traceability": { "score": 9, "comment": "Traçabilité complète" }

},
    "overall_quality_score": 8.5,
    "action_recommended": "APPROVE"
}
}
}

```

Seuils de Décision

Score	Action
≥ 8.5	APPROVE - Alerte immédiate
7.0 - 8.4	REVIEW - Validation humaine recommandée
< 7.0	REJECT - Rapport rejeté

Pilier 4 : Ground Truth - Validation Empirique

Objectif

Valider la performance globale du système en le testant par rapport à des analyses historiques fournies par le client.

Constitution de la Base de Vérité

Le client fournira un ensemble de scénarios réels (au minimum 10-15 cas) où une analyse de risque a été effectuée manuellement. Chaque cas doit inclure :

1. Le document source (loi, alerte climatique, actualité géopolitique).
2. La conclusion de pertinence (OUI/NON/PARTIELLEMENT).
3. Les impacts identifiés.
4. Les sites et fournisseurs impactés.
5. Les recommandations proposées.

Processus de Test

- Le système PING exécute une analyse complète sur chaque cas de la base de vérité.
- Les résultats du LLM Pertinence Checker sont comparés à la conclusion de pertinence du client.
- Les résultats du LLM Risk Analyzer sont comparés sémantiquement aux impacts et recommandations du client.
- Un score de similarité est calculé pour chaque cas.

Métriques de Performance

Métrique	Calcul	Seuil MVP
Pertinence Accuracy	% de cas où le Pertinence Checker conclut correctement (OUI/NON)	≥ 95%
Impact Similarity	Similarité sémantique moyenne entre les impacts du système et ceux du client	≥ 90%
Entity Accuracy	% d'entités (sites, fournisseurs) correctement identifiées	≥ 88%
Overall Reliability	Moyenne des trois métriques précédentes	≥ 90%

Résultat Attendu

JSON
<pre>{ "ground_truth_validation": { "total_cases_tested": 15, "metrics": { "pertinence_accuracy": 0.93, "impact_similarity": 0.91, "entity_accuracy": 0.89, "overall_reliability": 0.91 }, "conclusion": "Le système PING atteint une fiabilité de 91%, démontrant que" } }</pre>

```
}
```

Workflow Complet de Qualité

Plain Text

1. Événement Externe
2. LLM Pertinence Checker
 - └→ NON → Archiver
 - └→ OUI/PARTIELLEMENT ↓
3. LLM Risk Analyzer
4. LLM Judge
 - └→ APPROVE → Alerte Immédiate
 - └→ REVIEW → Validation Humaine
 - └→ REJECT → Archiver
5. Ground Truth Validation (Continu)
 - └→ Mesurer la fiabilité globale (Cible: $\geq 90\%$)

Résumé des Avantages

- ✓ **Séparation des Responsabilités** : Chaque LLM a une tâche clairement définie.
- ✓ **Réduction des Faux Positifs** : Le Pertinence Checker filtre les événements non pertinents avant le Risk Analyzer.
- ✓ **Meilleure Qualité** : Chaque LLM peut être optimisé pour sa tâche spécifique.
- ✓ **Traçabilité Complète** : Chaque étape est documentée et justifiée.
- ✓ **Évaluation Objective** : Le LLM Judge évalue chaque rapport de manière systématique.
- ✓ **Validation Empirique** : La Ground Truth garantit une fiabilité $\geq 90\%$.
- ✓ **Efficacité** : Pas d'analyse détaillée pour les événements non pertinents.