

# Xavier BLANC

JOURNÉE  
FRANÇAISE  
DES TESTS  
LOGICIELS

RAG Na Rock  
« Le RAG N'est pas Mort »

BEFFROI DE MONTROUGE



université  
de BORDEAUX



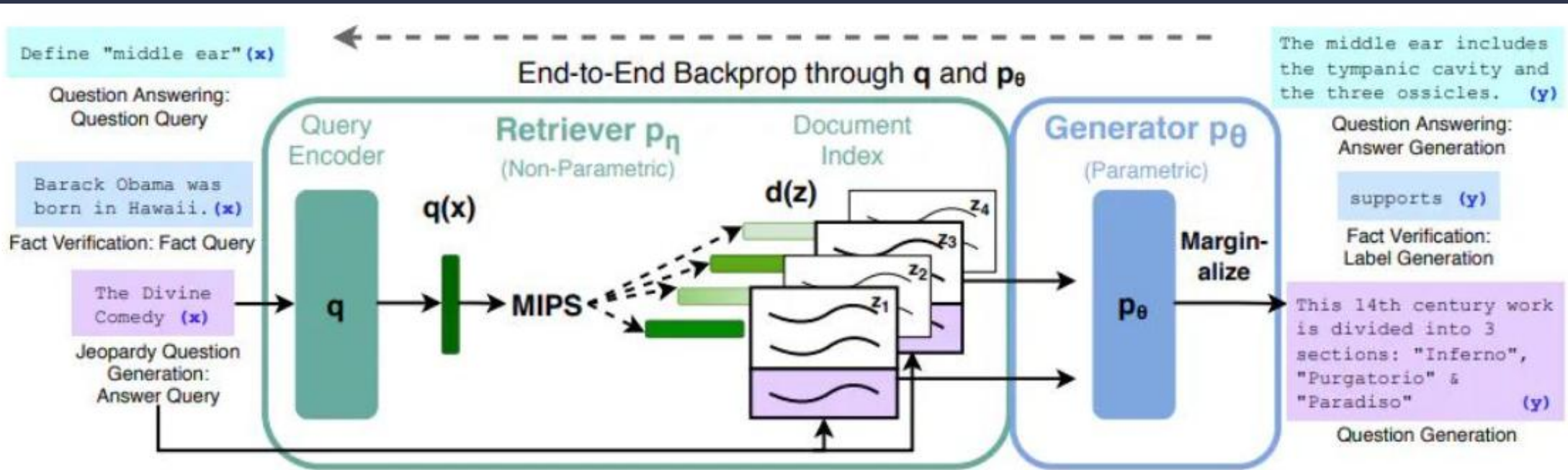
# Plan

- 1 - Qu'est ce que le RAG ?
- 2 - Cas d'étude : le rapprochement des incidents
- 3 - RAG Architecture
- 4 - Test et comparaison des technologies
- 5 - Mise en production
- 6 - Conclusion

# RAG = 1 Retrieval – 2 Augmented Generation

- Principe clé : retrouver de l'information pertinente relativement à une question, la donner à un LLM pour qu'il génère une réponse optimale
- Hypothèse sous-jacente : Le LLM n'a pas la connaissance suffisante pour répondre à la question. Il n'a pas été entraîné avec les données permettant d'y répondre de manière optimale.

# RAG : Architecture de tous les RAG



*Schéma de l'architecture du système RAG de Lewis et al*

# Cas d'étude : Le rapprochement des incidents

- Toutes les exécutions automatiques de test en erreur génèrent un incident.
  - L'incident contient la trace du run qui pointe vers le scénario de test qui a été exécuté
  - Il n'est pas rare que des incidents aient une même cause

=> Comment savoir si un nouvel incident est « similaire » à un incident passé ?

Rapprochement d'un incident : Retrouver des incidents similaires

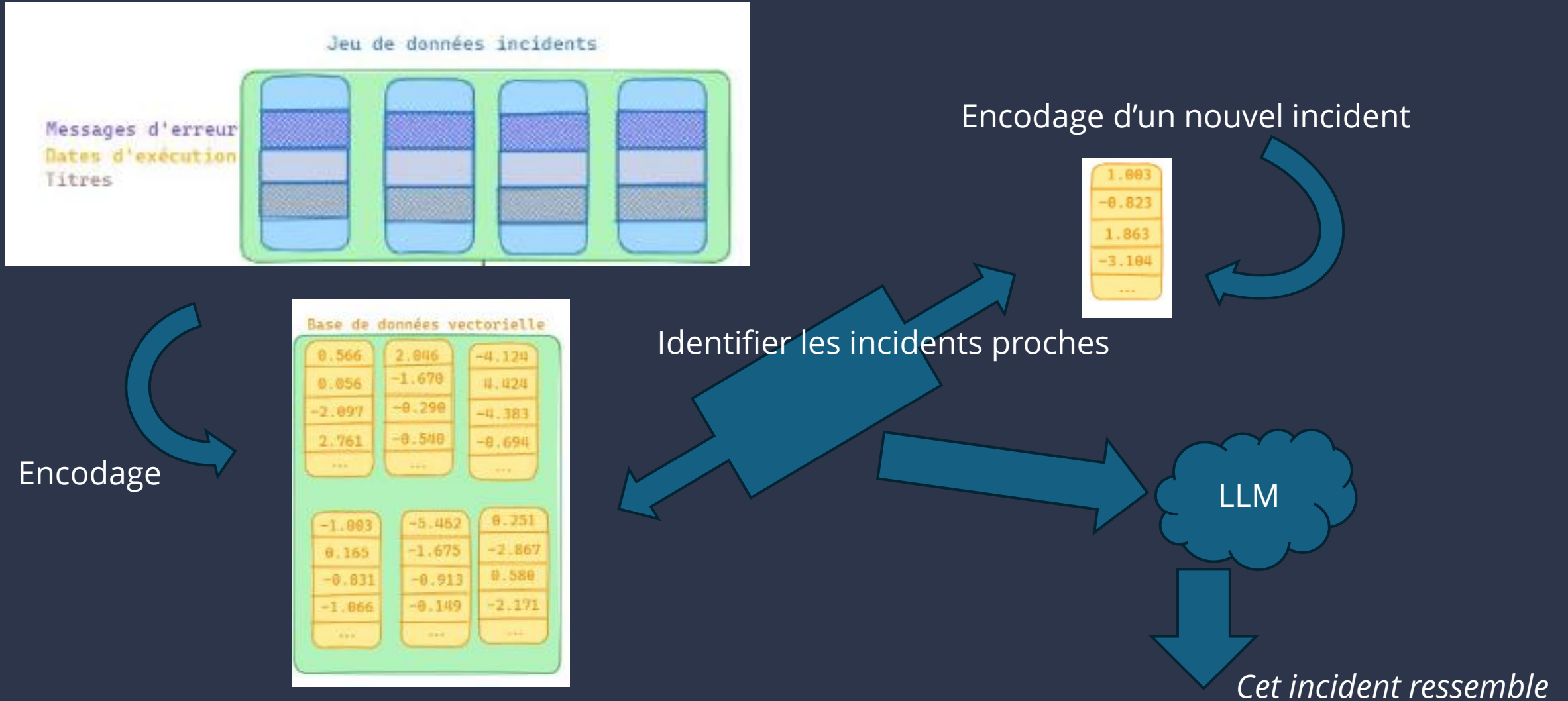
# Problématique du rapprochement des incidents

- On considère l'arrivée d'un nouvel incident
- Il faut identifier parmi tous les incidents passés lesquels sont proches du nouvel incident

=> Présenter une analyse avec les incidents proches et les raisons

=> Ce cas est typique du RAG car les incidents ne sont pas connus par le LLM lors de son entraînement

# Le RAG Narock : Architecture



# Le RAG Narock : Encodage des incidents

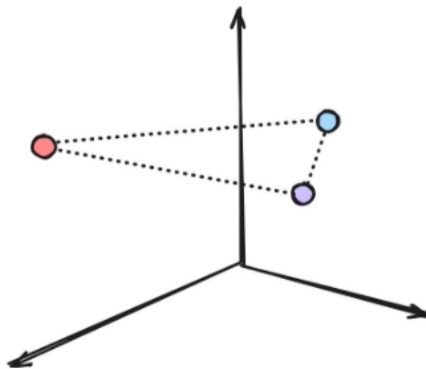
## Incidents Passés

Echec lors de l'attente de la visibilité de l'élément : #button : menu (#1) n'est pas visible

Echec lors de l'attente de la visibilité de l'élément : #B7B55208 (#1) n'est pas visible

Echec du survol de la souris : #main-tab (#1)

Messages d'erreur dans l'espace :



Embeddings messages d'erreur :

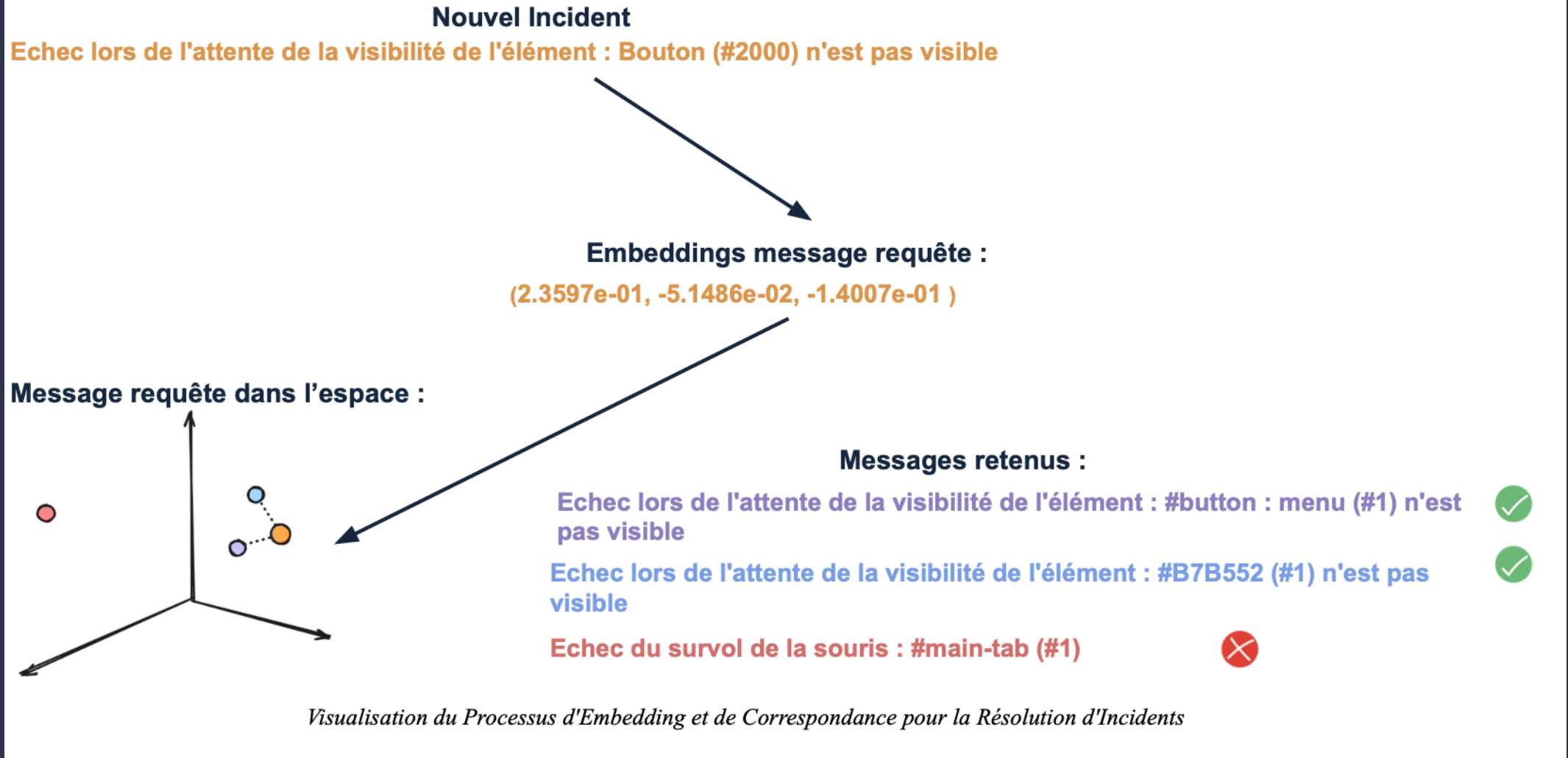
( 2.4000e-01, -1.6975e-01, -1.6522e-01 )

( 2.4125e-01, -5.7165e-02, -1.7241e-01 )

( 4.2197e-01, -4.9142e-02, -1.6197e-01 )

*Visualisation des embeddings de messages d'erreur dans un espace vectorielle*

# Le RAG Narock : Encodage des incidents



# Le RAG Narock : Génération d'analyse

## A quel incident je rattache :

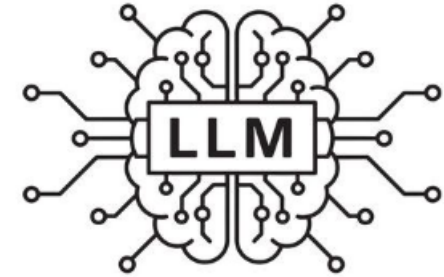
```
message : "Echec lors de l'attente de la visibilité de l'élément : Buton (#2000) n'est pas visible",  
scenario : "11327", executionDate : "2024-01-18 15:09:25", lastMessage : "Click on #edit-search-submit (#1)",  
title : "TECHNICAL ERROR - Error 500 with error message \"Sorry, we are temporarily unavailable\"",  
Etape : "3-5-1", Contexte de l'étape générale : "Submit the destination form IN form_validation (blocking)",  
Etape précédente : "None", Etape suivante : "Click on #edit-search-submit (#1)"
```

## Parmi ces incidents

```
message : "Message : \"Echec lors de l'attente de la visibilité de l'élément : #button : menu (#1) n'est pas visible",  
scenario : "11327", executionDate : "2023-12-30 05:25:13", lastMessage : "Click on #edit-search-submit (#1)",  
title : "Error message \"Your email address or password has changed. Please login.\"",  
Etape : "3-5-2", Contexte de l'étape générale : "Submit the destination form IN form_validation (blocking)",  
Etape précédente : "None", Etape suivante : "Click on #edit-search-submit (#1)"
```

```
message : "Echec lors de l'attente de la visibilité de l'élément : #B7B55208 (#1) n'est pas visible",  
scenario : "10023", executionDate : "2022-09-14 13:10:32", lastMessage : "Tap on \".fixed-panel [id] (#1)\"",  
title : "ANOMALY - \"Standard\" Room Offer is not Displayed - [WEB]",  
Etape : "6-1-2", Contexte de l'étape générale : "Submit the contact form IN form_validation (blocking)",  
Etape précédente : "Click on .sticky-block [id] (#1)", Etape suivante : "None"
```

Prompt qui contient l'incident actuel, les incidents candidats et des consignes pour le LLM



*Il serait judicieux de le rattacher au premier incident. La problématique est la même. Il porte sur le même scénario. Ce n'est pas exactement la même étape mais c'est très proche.*

....

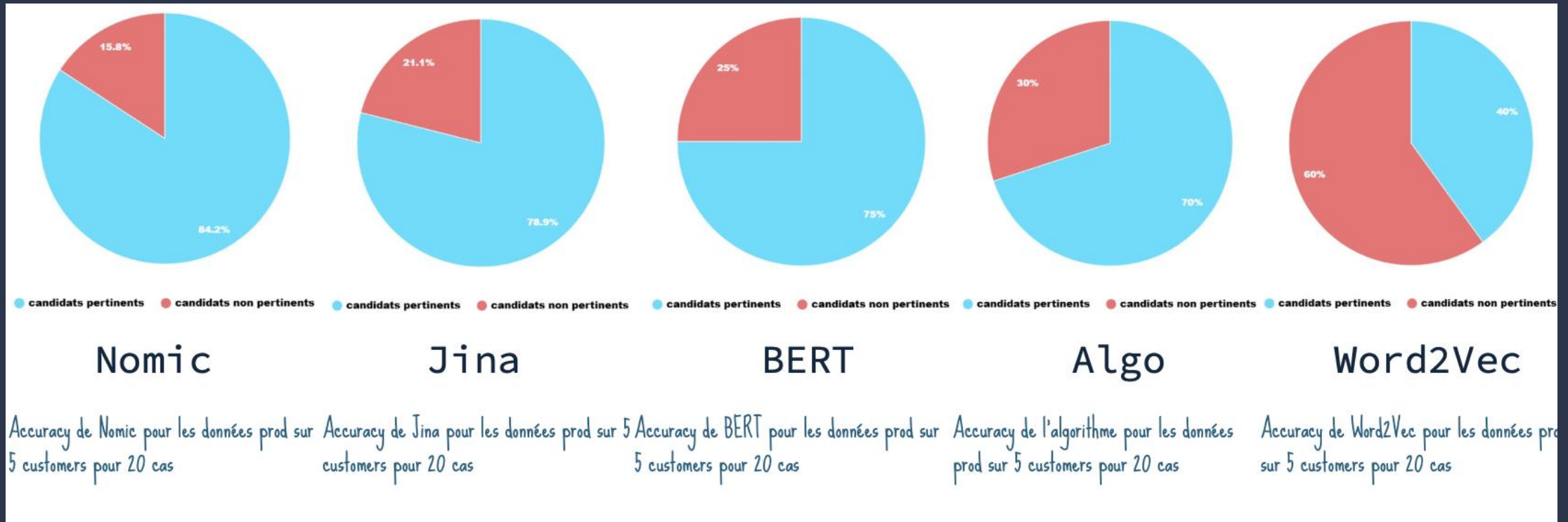
# Test de validation

Modèle	Index	top 1		top 3	
		Précision	MRR	Précision	MRR
word2vec-google-news-300	L2	97.2	97.2	92.9	97.4
	MIPS	97.2	97.2	92.7	97.4
glove-twitter-25	L2	96.7	96.7	92.7	97.1
	MIPS	96.7	96.7	92.7	97.1
en_core_web_md	L2	96.9	96.9	92.5	97.7
	MIPS	96.2	96.2	92.6	97.8
<u>paraphrase-MiniLM-L6-v2</u>	L2	89.5	89.5	88.1	96.9
	MIPS	88.8	88.8	88.0	96.5
<u>bert-base-uncased</u>	L2	87.2	87.2	90.6	97.3
	<u>MIPS</u>	87.0	87.0	90.8	97.6
xlnet-base-cased	L2	75.6	75.6	74.5	84.3
	MIPS	77.4	77.4	75.0	84.7
<u>gpt2</u>	L2	57.6	57.6	61.5	64.5
	<u>MIPS</u>	60.0	60.0	63.2	66.8
roberta-base	L2	57.4	57.4	84.9	91.6
	MIPS	57.5	57.5	86.3	93.8
t5-small	L2	54.2	54.2	67.8	75.1
	MIPS	53.9	53.9	68.0	75.3

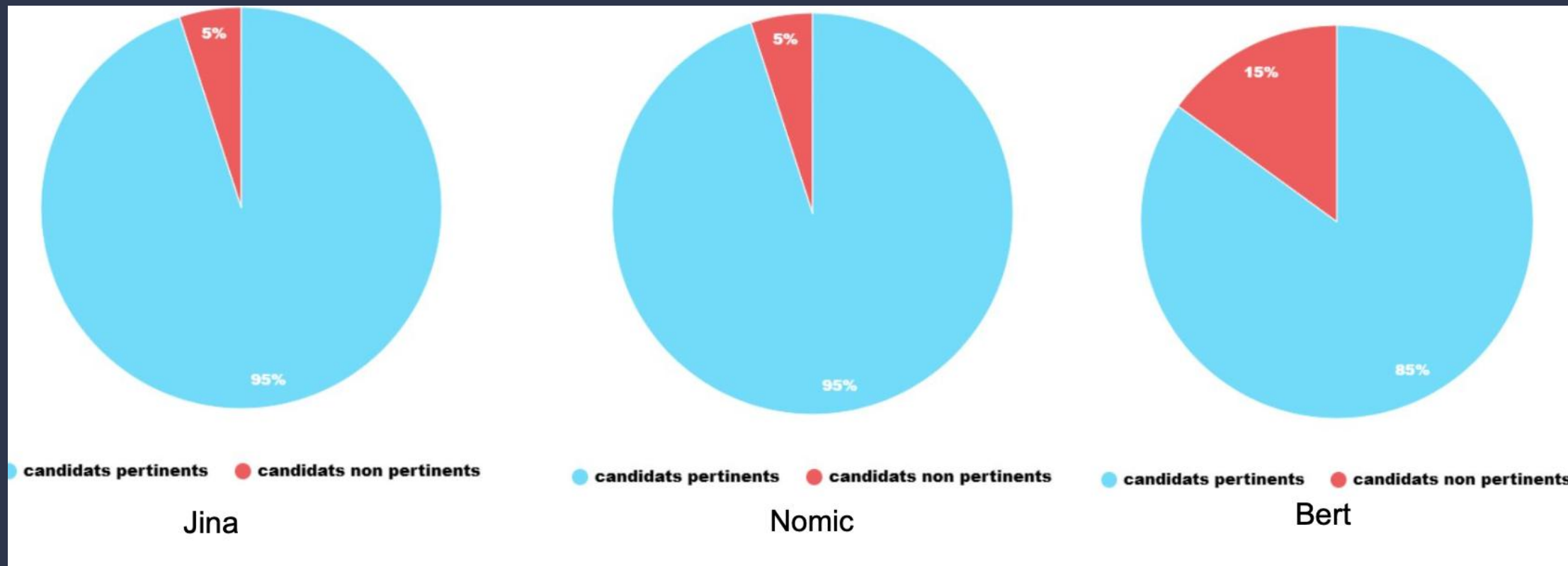
- 68 scénarios du même environnement client
- 226 runs en échec
- 6 639 messages test

Résultats de la recherche par index pour le top 1 et top 3

# Test de validation



# Test de validation



# Mise en production

The screenshot displays the Mr.Suricate software interface. A central dialog box titled "Gestion du rattachement de l'incident #114946" is open. The dialog contains the following elements:

- Date début:** 20/01/2026
- Date fin:** 22/01/2026
- Scénarios:** A list with "0 Exécutions disponibles" on the left and "1 Exécutions sélectionnées" on the right. The selected execution is "16024 - 78851007".
- Buttons:** "Analyser par IA" (with a star icon) and "Exécution #78851007" are located below the scenarios list.
- Form fields:** "Rattacher l'exécution à l'incident" (with a dropdown arrow), "Visibilité" (with a dropdown arrow), and "Classification" (with a dropdown arrow).
- Action:** A prominent green "RATTACHER" button is at the bottom right of the dialog.

The background interface shows a sidebar with navigation icons, a search bar at the top, and a main content area with a "TECHNICAL ERROR" header and a "Liste des Périodes" section. The top right corner of the interface includes the text "B&B - MOBILE - PREPROD" and a user profile for "Alix DELAGE".

# Peut on faire sans RAG ?

- Le LLM doit pouvoir obtenir la liste des anciens incidents (les plus probables)
- On peut faire un serveur MCP avec un Agent
  - Mais l'agent devrait faire un call au serveur MCP avant même de poser la question au LLM

=> Il faut exploiter un RAG quand on a beaucoup d'information, non disponible à l'entraînement, et qui évolue dans le temps.

# Conclusion

- RAG is not dead
- Nécessaire quand on a des données dynamiques
- Nécessaire quand les données ne sont pas disponibles lors de l'apprentissage
- Nécessite la définition d'une fonction de distance (comparaison des incidents)
- Des performances importantes

# Perspectives

- Efficacité de l'approche
- Maintenance de la base des embeddings
- Intégrer les retours des utilisateurs