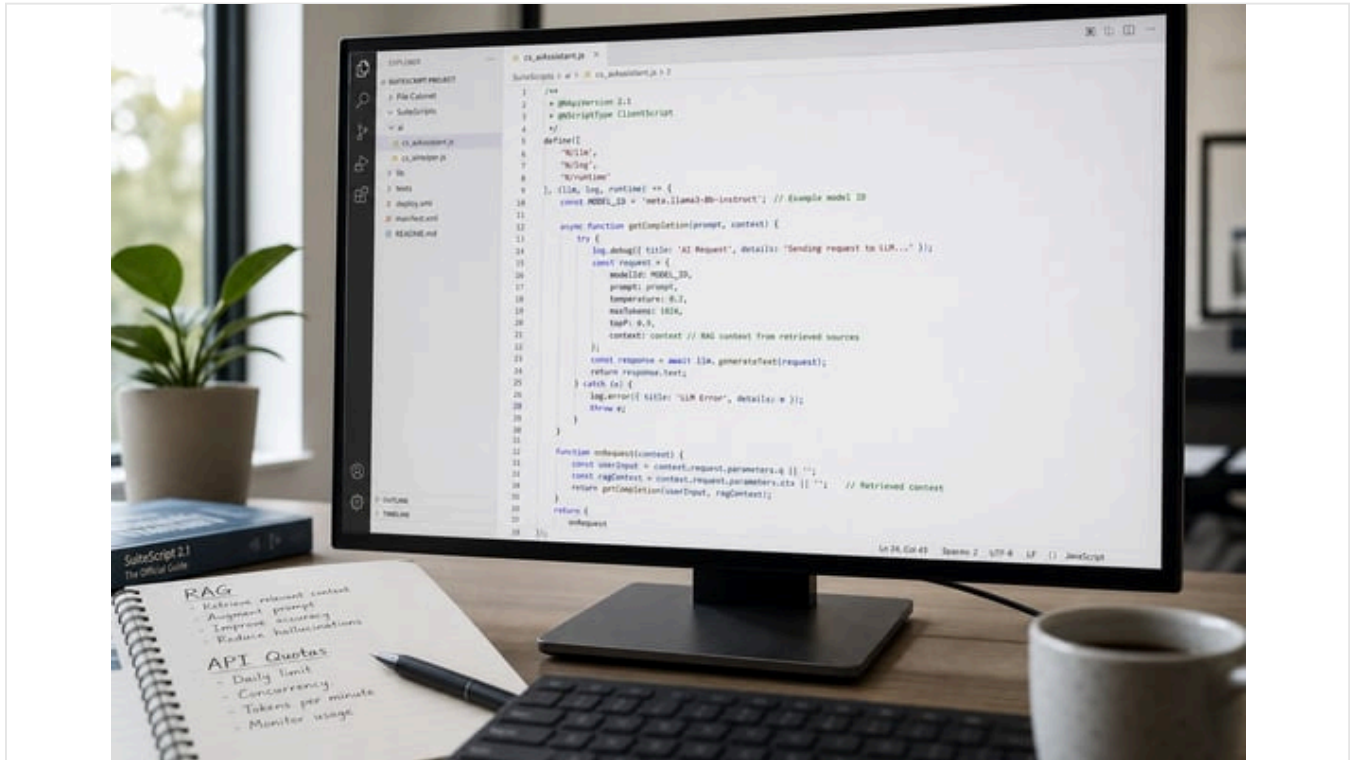


Module NetSuite N/LLM : Guide d'intégration de l'IA générative dans SuiteScript

Publié le 6 mai 2026 40 min de lecture



Résumé analytique

Le nouveau **module SuiteScript N/LLM** d'Oracle NetSuite (introduit en 2024) intègre **nativement** des capacités de grands modèles de langage (LLM) et d'IA générative (GenAI) dans la plateforme ERP NetSuite. Il permet aux développeurs SuiteScript d'appeler directement les services d'IA générative d'Oracle Cloud Infrastructure (OCI) depuis des scripts NetSuite, facilitant ainsi la génération de texte à la demande, les chatbots, la synthèse, la recherche intelligente, et bien plus encore, le tout basé sur les données propres au client. Les fonctionnalités clés incluent la génération de contenu (`llm.generateText`), l'évaluation de prompts (`llm.evaluatePrompt`), les plongements vectoriels (embeddings) (`llm.embed`) et la **génération augmentée par récupération (RAG)** en fournissant des documents NetSuite au LLM. Le module prend en charge les réponses en flux (`llm.generateTextStreamed`), l'intégration avec Prompt Studio (modèles de prompts partagés) et la **gouvernance de l'utilisation** (quotas mensuels gratuits et **limites de concurrence**). Par défaut, NetSuite fournit à chaque compte un quota mensuel gratuit de jetons LLM (quotas distincts pour la génération et les embeddings) (Source: www.houseblend.io). Les développeurs peuvent surveiller cela via des API (`llm.getRemainingFreeUsage()`, `llm.getRemainingFreeEmbedUsage()`) ou la page des préférences IA (Source: www.houseblend.io) (Source: docs.oracle.com), et peuvent également configurer leurs propres identifiants OCI pour une utilisation illimitée (facturée au client) (Source: www.houseblend.io) (Source: www.houseblend.io).

Ce rapport fournit un **guide d'intégration** complet du module SuiteScript N/LLM. Nous examinons la stratégie IA de NetSuite et le contexte de l'intégration de modèles génératifs dans les systèmes ERP ; nous détaillons l'architecture technique et l'API (méthodes, paramètres, limites et exemples de code) ; nous explorons des cas d'utilisation types (chatbots, **synthèse de rapports**, recherche sémantique, saisie intelligente de champs, etc.), incluant des études de cas de clients NetSuite ; et nous abordons la gouvernance, la confidentialité des données et les orientations futures. Nous citons la documentation officielle, des analyses tierces, des statistiques sectorielles et des exemples concrets. En particulier, nous montrons comment N/LLM permet le « mini-RAG » au sein de NetSuite : les scripts peuvent interroger des données transactionnelles (via SuiteQL ou des recherches), construire des « documents sources » à partir de ces données et les transmettre avec des prompts afin que les réponses du LLM soient ancrées dans les enregistrements du client (Source: blogs.oracle.com) (Source: houseblend.io). Cette approche ancrée atténue les hallucinations et renforce la pertinence (Source: blogs.oracle.com) (Source: houseblend.io).

Principales conclusions :

- Capacités** : Le module N/LLM introduit une riche [API SuiteScript](#) (Tableau 2) incluant des méthodes comme `llm.generateText(options)` pour la génération de texte libre, `llm.chat(options)` comme alias pour les prompts conversationnels, `llm.evaluatePrompt(options)` pour exécuter des modèles stockés dans Prompt Studio, `llm.embed(options)` pour obtenir des plongements vectoriels, et des utilitaires tels que `llm.createDocument()` (pour construire un contexte RAG) et `llm.getRemainingFreeUsage()` (pour vérifier les quotas) (Source: www.houseblend.io) (Source: docs.oracle.com). Les variantes en flux (`generateTextStreamed`, `chatStreamed`, `evaluatePromptStreamed`) permettent de traiter les jetons au fur et à mesure de leur arrivée, réduisant la latence (Source: gurussolutions.com). L'API inclut également des objets spécifiques au chat (`llm.ChatMessage`, `llm.ChatRole`) pour créer des sessions de conversation (Source: www.houseblend.io) (Source: docs.oracle.com).
- Gouvernance** : Oracle impose des **contrôles d'utilisation stricts**. Chaque compte (et chaque SuiteApp distincte) bénéficie d'un *pool mensuel gratuit* d'appels LLM pour la génération de texte et d'un pool distinct pour les appels d'embedding (Source: www.houseblend.io) (Source: docs.oracle.com). Par défaut, les appels sont déduits de ces quotas gratuits (qui se réinitialisent mensuellement), mais les clients peuvent lier un identifiant OCI Generative AI externe pour basculer les appels vers le quota (et la facturation) de ce compte pour une utilisation illimitée (Source: www.houseblend.io) (Source: www.houseblend.io). Une limite technique de **5 appels simultanés** s'applique par script pour les méthodes de génération et 5 pour les méthodes d'embedding (Source: www.houseblend.io) (Source: www.houseblend.io). Le dépassement des quotas ou de la concurrence génère des erreurs que les scripts doivent gérer. Les administrateurs peuvent suivre l'utilisation sur la page des préférences IA de NetSuite, qui affiche des ventilations mensuelles par type (Génération vs Embedding) (Source: www.houseblend.io) (Source: docs.oracle.com).
- Intégration et bonnes pratiques** : Le module est conçu pour la **confidentialité et la sécurité**. Il est important de noter que le LLM n'a *aucun accès direct* aux données d'un compte à moins qu'elles ne soient explicitement fournies via un script (Source: www.techfino.com) (Source: www.oracle.com). La conception de sécurité d'Oracle maintient les appels d'IA générative sur OCI sous un accès basé sur les rôles, et aucune donnée client n'est partagée avec des fournisseurs de modèles ou d'autres locataires (Source: www.oracle.com). Les meilleures pratiques (telles que documentées par Oracle et des consultants) incluent le fait de *ne transmettre que des données vérifiées et non sensibles* au LLM et de toujours valider l'exactitude des résultats (Source: www.techfino.com) (Source: gurussolutions.com). N/LLM complète d'autres modules IA de NetSuite (par exemple *N/documentCapture* ou *N/machineTranslation*, voir Tableau 1) et des outils comme **Prompt Studio**, qui permet aux administrateurs de définir des prompts et des actions réutilisables pouvant être invoqués via `llm.evaluatePrompt` (Source: gurussolutions.com) (Source: docs.oracle.com).

Les études de cas et les analyses sectorielles renforcent le potentiel de N/LLM. Par exemple, le magazine Houseblend note que des clients de NetSuite tels que BirdRock Home (un détaillant à fort volume) et Overture Promotions ont déjà utilisé l'IA/ML sur leurs données NetSuite pour obtenir des « améliorations stratégiques exploitables » (Source: www.houseblend.io). Le PDG de BirdRock rapporte traiter « des milliers de commandes quotidiennement » dans NetSuite, un cadre idéal pour les prévisions et les insights pilotés par l'IA (Source: houseblend.io) (Source: www.houseblend.io). Plus largement, les données d'enquête montrent une adoption explosive de la GenAI : Gartner a constaté que fin 2023, environ 29 % des organisations avaient des déploiements GenAI en direct (contre <5 % en 2022), avec une prévision de >80 % des entreprises utilisant des API ou des applications GenAI d'ici 2026 (Source: www.houseblend.io) (Source: www.techradar.com). McKinsey projette que l'IA générative pourrait ajouter **2,6 à 4,4 billions de dollars** par an à l'économie mondiale, doublant l'impact des générations d'IA précédentes (Source: www.houseblend.io) (Source: www.techradar.com). Cependant, les experts avertissent que sans une intégration minutieuse et une gouvernance des données solide, la plupart des projets pilotes d'IA générative ne parviennent pas à offrir un retour sur investissement mesurable (Source: houseblend.io) (Source: www.houseblend.io). Dans le cas de NetSuite, exploiter efficacement N/LLM nécessite une planification (par exemple, structurer les données dans des prompts, contrôler l'approvisionnement en champs sensibles, tester les résultats) et un alignement avec la stratégie IA et les normes de conformité de l'entreprise (Source: houseblend.io) (Source: www.houseblend.io).

En résumé, le **module N/LLM de NetSuite** représente une nouvelle frontière puissante dans la personnalisation ERP. Il intègre une technologie LLM de pointe directement dans l'écosystème SuiteCloud, permettant une nouvelle automatisation et une intelligence accrue tout en préservant l'intégrité de la gouvernance des données. Ce rapport approfondit ses fonctionnalités, illustre des exemples pratiques et examine les implications plus larges pour les entreprises utilisant NetSuite ERP.

Introduction et contexte

NetSuite et l'essor de l'IA générative

Oracle NetSuite est une plateforme **ERP cloud** multi-locataire de premier plan qui centralise les données d'une entreprise à travers la finance, le CRM, les stocks, le commerce, et plus encore (Source: houseblend.io) (Source: houseblend.io). Son architecture de données unifiée — souvent appelée « *Suiteness* » — signifie, comme le note le vice-président principal du développement d'Oracle, que « toutes les données commerciales et les flux de travail dans un seul système » fournissent une base idéale pour des insights pilotés par l'IA (Source: houseblend.io) (Source: houseblend.io). Plus de 40 000 organisations utilisent NetSuite dans le monde, traitant des volumes de transactions élevés et des données de référence riches (clients, articles, transactions, etc.) qui sont prêtes pour l'intelligence machine (Source: houseblend.io) (Source: houseblend.io).

Parallèlement, l'**IA générative** a rapidement transformé l'informatique et les logiciels d'entreprise. La percée publique des LLM comme ChatGPT (2022) et les offres concurrentes (Anthropic Claude, Google Gemini, etc.) a démontré que l'IA peut effectuer de nombreuses tâches cognitives — rédiger du texte, répondre à des questions, résumer des documents — avec une fluidité de niveau humain (Source: www.houseblend.io) (Source: www.techradar.com). Les analystes projettent des impacts spectaculaires : McKinsey (2023) estime que la GenAI pourrait contribuer à environ **2,6 à 4,4 billions de dollars** par an à l'échelle mondiale, doublant environ la valeur de l'IA précédente (Source: www.houseblend.io) (Source: www.techradar.com). Rien qu'entre 2023 et 2025, les éditeurs de logiciels et les utilisateurs professionnels ont adopté la GenAI à un rythme sans précédent. Gartner a rapporté que fin 2023, près de **30 % des organisations interrogées** avaient déjà déployé des solutions GenAI, contre <5 % en 2022 (Source: www.houseblend.io). D'autres études prévoient que plus de **80 % des entreprises** utiliseront des applications ou des API GenAI d'ici 2026 (Source: www.houseblend.io). Cependant, de nombreuses entreprises signalent également que le simple fait d'« utiliser des outils d'IA » ne s'est pas encore traduit par des gains financiers (Source: www.techradar.com) (Source: www.techradar.com), probablement parce qu'une intégration et une exécution efficaces restent difficiles.

Les entreprises ont réagi en intégrant l'intelligence artificielle au cœur de leurs applications métier. Les principales plateformes ERP et CRM ont annoncé des fonctionnalités basées sur l'IA : analyses prédictives, traitement automatisé des documents et agents conversationnels. Oracle NetSuite a « enrichi par l'IA » sa suite de manière agressive : les dirigeants soulignent que l'IA est devenue une condition sine qua non et insistent sur le fait que les données centralisées de NetSuite ouvrent des opportunités uniques. Par exemple, lors du SuiteWorld 2023, Oracle a annoncé **NetSuite Text*Enhance**, une fonctionnalité générative qui utilise les données de l'entreprise pour pré-remplir ou améliorer les champs de texte libre dans les domaines de la finance, des RH, de la chaîne logistique, et plus encore (Source: www.oracle.com) (Source: www.prnewswire.com). Lors du SuiteWorld 2024, des dizaines de nouvelles capacités basées sur l'IA ont été introduites (SuiteAnalytics Assistant, rapports narratifs, détection d'anomalies, etc.), avec l'insistance sur le fait qu'elles sont **intégrées sans coût de licence supplémentaire** (Source: www.houseblend.io) (Source: www.prnewswire.com). Evan Goldberg (EVP d'Oracle NetSuite) a déclaré : « *Nous avons intégré l'IA au cœur de la suite... et nous ajoutons maintenant davantage de capacités d'IA générative... sans coût supplémentaire* » (Source: www.prnewswire.com) (Source: www.oracle.com). En février 2025, la feuille de route de NetSuite incluait des agents avancés (pour le CPQ, SuiteAnswers), une API de gestion des prompts et une utilisation étendue des modèles génératifs pour rationaliser les flux de travail (Source: www.prnewswire.com) (Source: www.prnewswire.com).

Au sein de cette initiative IA plus large, le **module SuiteScript N/LLM** est la pièce maîtresse. Annoncé en 2024 et déployé en 2025, N/LLM offre aux développeurs un accès *programmatique* aux services LLM directement dans leur code JavaScript. En pratique, il permet à tout SuiteScript (User Event, Suitelet, RESTlet, etc.) de « **discuter avec un grand modèle de langage** » dans le cadre de la logique métier habituelle. La documentation d'Oracle le décrit comme un module SuiteScript 2.1 qui « apporte la puissance de l'IA générative dans votre environnement NetSuite SuiteScript » (Source: gurussolutions.com). Le module est construit sur la plateforme OCI Generative AI d'Oracle (exploitant actuellement les modèles Cohere) et s'intègre étroitement aux fonctionnalités de données et de gouvernance de NetSuite (Source: www.houseblend.io) (Source: www.houseblend.io).

Le timing et la motivation sont clairs : alors que les entreprises adoptent rapidement l'IA générative dans toutes leurs fonctions, l'activation d'appels LLM natifs dans les scripts ERP permet aux entreprises d'exploiter leurs données métier dans des flux de travail intelligents. Contrairement aux outils tiers, cette approche « verrouille » le LLM au sein de l'environnement sécurisé de NetSuite. Les données ne quittent jamais NetSuite, sauf via des prompts explicites, ce qui apaise les préoccupations liées à la gouvernance des données. Comme le note Oracle, aucune donnée client n'est partagée avec des fournisseurs de modèles externes ; les requêtes de chaque entreprise s'exécutent sur l'infrastructure multi-tenant d'OCI sous des contrôles d'accès stricts (Source: www.oracle.com). En bref, N/LLM intègre une IA de pointe là où elle dispose du plus grand contexte — avec les propres enregistrements de l'entreprise — tout en maintenant une sécurité de niveau entreprise.

Ce rapport examinera le module N/LLM en détail. Nous commençons par passer en revue la stratégie IA de NetSuite et le rôle de l'IA générative dans l'ERP (Section 2). Nous décrivons ensuite l'**architecture technique** de N/LLM (Section 3) et présentons ses méthodes d'API SuiteScript, ses paramètres et ses limites (Section 4). Nous fournissons des exemples de code illustrant comment appeler le LLM, utiliser des embeddings et effectuer du RAG. Ensuite, nous analysons des cas d'utilisation réels et des études de cas (Section 5), montrant comment les organisations peuvent appliquer N/LLM pour des tâches telles que l'interrogation en langage naturel des données, le reporting automatisé, les chatbots intelligents, etc. La Section 6

traite des considérations pratiques : quotas d'utilisation, confidentialité des données, conformité et meilleures pratiques. Enfin, nous abordons les orientations futures (Section 7) et concluons avec des recommandations. Tout au long du document, nous citons la documentation officielle de NetSuite, des analyses sectorielles et des exemples pour étayer toutes les affirmations.

Module NetSuite N/LLM : Aperçu technique

Modules IA dans NetSuite

NetSuite fournit plusieurs modules SuiteScript intégrés pour les tâches d'IA et d'automatisation, comme résumé dans le Tableau 1. Le nouveau **module N/LLM** est spécifiquement destiné à l'IA générative :

MODULE IA SUITESCRIPT	SERVICE OCI UTILISÉ	DESCRIPTION
N/llm	OCI Generative AI	Accède aux grands modèles de langage pour la génération de texte à la demande, les questions-réponses, la synthèse et d'autres tâches. Permet d'envoyer des prompts (avec des documents de contexte optionnels) aux LLM et renvoie des réponses générées par l'IA qui peuvent être réécrites dans les enregistrements NetSuite (Source: docs.oracle.com) (Source: houseblend.io).
N/documentCapture	OCI Document Understanding	Effectue l'OCR et l'extraction de données structurées à partir de documents (PDF, images, reçus) (Source: houseblend.io). Utile pour automatiser la saisie de données (ex: lecture des champs de factures) et alimenter les enregistrements NetSuite.
N/machineTranslation	OCI Language Translator Service	Fournit une traduction de texte entre les langues prises en charge (Source: houseblend.io). Permet aux scripts de localiser du contenu ou de créer des fonctionnalités multilingues (ex: traduction automatique de notes clients, rapports).

Tableau 1 : Modules IA de SuiteScript et leurs services OCI sous-jacents (source : documentation Oracle (Source: houseblend.io). Le module N/LLM est nouveau dans SuiteScript 2.1 (depuis 2024.1) et se distingue par son interface directe avec les LLM (actuellement les modèles Cohere sur OCI).

Chacun de ces modules s'exécute sur le backend cloud de NetSuite et est soumis à la sécurité et à la gouvernance de la plateforme. En particulier, les appels via N/LLM transmettent les données via le service d'IA générative OCI d'Oracle ; ils n'invoquent jamais directement des points de terminaison HTTP externes qui contourneraient NetSuite. Cela garantit que l'utilisation de l'API est suivie et que les autorisations NetSuite basées sur les rôles s'appliquent à toutes les données transmises (le LLM ne peut voir aucun champ d'enregistrement à moins qu'un script ne les lise explicitement et ne les inclue dans le prompt (Source: www.techfino.com) (Source: www.oracle.com). Il est important de noter qu'Oracle fournit un **pool d'utilisation gratuite limité** pour les appels IA (génératifs et embeddings) chaque mois (Source: www.houseblend.io) (Source: docs.oracle.com), renouvelé régulièrement. Si un compte a besoin de limites plus élevées, un administrateur peut configurer des **identifiants de compte OCI** dans Configuration > Société > Préférences IA. Une fois lié, les appels N/LLM puiseront dans le quota de ce compte OCI (avec une facturation selon les tarifs OCI) au lieu du pool gratuit (Source: www.houseblend.io) (Source: docs.oracle.com).

Modèles LLM pris en charge et RAG

N/LLM est actuellement construit sur la plateforme d'IA générative d'OCI, qui, dans les cas d'utilisation NetSuite, est alimentée par les modèles de Cohere. Le modèle par défaut « **Command-A 03-2025** » (un modèle de génération de texte Cohere) est utilisé si aucun autre modèle n'est spécifié (Source: docs.oracle.com) (Source: docs.oracle.com). Dans SuiteScript, on peut spécifier `options.modelFamily` dans `llm.generateText` pour choisir les modèles ; les valeurs prises en charge sont indiquées dans la documentation d'Oracle (Source: docs.oracle.com). Par exemple, `llm.ModelFamily.COHERE_COMMAND` correspond au modèle Command-A (avec prise en charge du RAG et du préambule), tandis que `llm.ModelFamily.GPT_OSS` fait référence à un grand modèle GPT open-source (avec prise en charge du préambule mais sans RAG) (Source: docs.oracle.com). Les clients JavaScript les voient comme des énumérations de chaînes simples. L'interface utilisateur de NetSuite utilise toujours par défaut le dernier modèle approuvé si aucun n'est défini.

Une capacité fondamentale est la **Génération Augmentée par la Récupération (RAG)**. N/LLM permet au script de fournir des « documents sources » pour étayer la réponse du LLM. Ceux-ci sont créés avec `llm.createDocument({ text: "...", title: "..." })` et transmis via le tableau `options.documents`. Lorsqu'ils sont fournis, le LLM utilise le texte de ces documents comme contexte supplémentaire pour la génération et renvoie des **citations** dans la réponse qui pointent vers les documents originaux (Source: houseblend.io) (Source: blogs.oracle.com). En pratique, un SuiteScript peut interroger les enregistrements NetSuite (par exemple en utilisant SuiteQL ou `N/query`) pour récupérer les données pertinentes, formater chaque enregistrement ou note en une courte chaîne de texte et l'ajouter à un tableau de documents. Ensuite, l'appel `llm.generateText({ prompt: userQuestion, documents: myDocuments, ... })` garantit que la réponse est « basée sur vos propres données NetSuite, et non sur des connaissances aléatoires d'Internet » (Source: blogs.oracle.com). Cette approche de « mini-RAG » — récupérer d'abord les données pertinentes de l'entreprise, puis générer des réponses — améliore considérablement la précision factuelle et la pertinence métier (Source: houseblend.io) (Source: blogs.oracle.com). (Sans RAG, le LLM ne s'appuierait que sur son corpus d'entraînement, qui peut ne pas refléter les spécificités de l'entreprise.)

Oracle prend entièrement en charge le RAG : en transmettant un tableau d'objets `llm.Document`, la réponse est accompagnée d'un tableau `.citations` d'objets `llm.Citation` qui identifient les documents utilisés comme sources (Source: houseblend.io) (Source: blogs.oracle.com). Cela permet au script de lier les réponses à des enregistrements spécifiques ou à des articles de la base de connaissances. Les premiers exemples d'Oracle illustrent l'utilisation du RAG pour des tâches telles que répondre à des questions sur les données de vente ou la documentation d'une entreprise (Source: houseblend.io) (Source: blogs.oracle.com).

Intégration SuiteScript

Le module N/LLM est disponible dans **SuiteScript 2.1**. En pratique, les développeurs l'incluent avec l'appel `require` ou `define`. Par exemple, dans un script ou un Suitelet, on pourrait écrire :

```
define(['N/llm'], function(llm) {
  // Utiliser llm.generateText, llm.embed, etc.
  const result = llm.generateText({ prompt: "Bonjour le monde !" });
  log.debug('Réponse IA', result.text);
});
```

(L'exemple de script d'Oracle « Envoyer un prompt au LLM » illustre cela, en envoyant le prompt *"Bonjour le monde !"* au LLM par défaut et en lisant `response.text` (Source: docs.oracle.com.) Dans le débogueur SuiteScript 2.x, le code utilise `require` (comme dans l'exemple), mais dans les scripts déployés, on utilise `define` (Source: docs.oracle.com). En dehors de la syntaxe, le module N/LLM fonctionne comme tout autre module NetSuite : il renvoie des objets et des valeurs au script et peut être appelé depuis des scripts User Event, des Suitelets, des RESTlets ou des scripts planifiés. Ses méthodes génèrent des erreurs spécifiques à NetSuite (exceptions SuiteScript) en cas d'échec (ex: dépassement des quotas).

Comme les appels génératifs impliquent un calcul externe, ils sont soumis aux limites de gouvernance de NetSuite. La documentation officielle des limites note que chaque type de méthode a une limite de concurrence (5 appels) (Source: www.houseblend.io) et que chaque invocation SuiteScript du module entraîne une consommation de jetons (tokens) sur le pool gratuit. La limite de gouvernance d'Oracle pour `llm.generateText` est effectivement de 100 unités (ceci est indiqué sous « Gouvernance » dans la référence de l'API (Source: docs.oracle.com), bien qu'en pratique l'utilisation soit mesurée par jetons). Les développeurs doivent prévoir des délais potentiels (les appels peuvent prendre quelques centaines de millisecondes ou plus) et inclure une gestion des erreurs pour les exceptions de quota/concurrence.

Méthodes et objets clés (voir Tableau 2) : Le module N/LLM expose à la fois des méthodes (fonctions) et des types d'objets. Le Tableau 2 résume les plus importants :

MÉTHODE / OBJET	DESCRIPTION
llm.generateText(options)	Envoie un prompt textuel au LLM et renvoie un objet <code>llm.Response</code> contenant le texte généré (<code>.text</code>), les éventuelles citations (<code>.citations</code>) et les statistiques d'utilisation. Alias : <code>llm.chat(options)</code> . Prend en charge <code>options.prompt</code> (chaîne), <code>options.documents</code> optionnels (<code>llm.Document[]</code>) pour le RAG, et des paramètres avancés (modèle, <code>maxTokens</code> , température, pénalités) (Source: docs.oracle.com) (Source: www.houseblend.io).
llm.generateTextStreamed(options)	Identique à <code>generateText</code> mais renvoie un <code>llm.StreamedResponse</code> qui permet d'itérer sur les jetons au fur et à mesure de leur arrivée. Utile pour créer des interfaces de chat en direct. Alias : <code>llm.chatStreamed(options)</code> .
llm.evaluatePrompt(options)	Exécute un prompt de Prompt Studio par ID dans un script. Les options incluent l' <code>id</code> du prompt enregistré et une carte de <code>promptValues</code> pour ses variables de modèle. Renvoie un <code>llm.Response</code> . Alias : <code>llm.executePrompt(options)</code> (Source: www.houseblend.io).
llm.evaluatePromptStreamed(options)	Version en flux de <code>evaluatePrompt</code> (alias <code>llm.executePromptStreamed</code>). Traite les jetons en temps réel.

| **llm.embed(options)** | Convertit le texte saisi en un vecteur d'embedding à l'aide du modèle d'embedding Cohere. Attend `options.input` (chaîne ou tableau de chaînes). Renvoie un objet `llm.EmbedResponse` contenant `.embeddings` (tableaux de nombres) et `usage`. Les appels d'embedding utilisent un quota distinct de celui de la génération de texte (Source: docs.oracle.com) (Source: gurussolutions.com). | | **llm.createDocument(options)** | Crée un objet `llm.Document` avec des propriétés telles que `.text` et `.reference` (un titre ou un lien source). Ces objets peuvent être rassemblés dans un tableau et transmis en tant que `options.documents` pour le RAG. (Par ex., après avoir interrogé des enregistrements, vous pourriez faire `docs.push(llm.createDocument({ text: recordSummary, reference: recordId }));`) | | **llm.getRemainingFreeUsage()** | Renvoie le nombre de jetons génératifs gratuits restants dans le pool mensuel du compte. Utile pour surveiller la consommation. Fonctionne de manière synchrone. (Existe aussi en version asynchrone `.promise()`) | | **llm.getRemainingFreeEmbedUsage()** | Renvoie les jetons gratuits restants pour les appels d'embedding. | | **llm.ChatMessage** | Constructeur pour les messages de chat. À utiliser pour créer des tableaux de conversation pour `chatHistory`. Exemple : `llm.createChatMessage({ role: llm.ChatRole.USER, content: "..."});` | | **llm.ChatRole** | Énumération pour les rôles de chat (`USER`, `ASSISTANT`). Utilisée lors de la création d'un `ChatMessage`. | | **llm.Citation** | Représente une citation dans une réponse d'IA générative renvoyant à un document. Contient `.content` (texte de l'extrait) et `.sourceIndex` (index dans votre liste `documents`). | | **llm.Response / llm.StreamedResponse** | Renvoyé par les appels de génération ou d'invite. `.Response` possède `.text` (chaîne) et `.citations` (tableau de `llm.Citation`). `.StreamedResponse` est itérable et produit des jetons ou des réponses partielles. |

Tableau 2 : Méthodes et objets sélectionnés du module N/LLM. Voir la documentation d'Oracle pour plus de détails (Source: docs.oracle.com) (Source: gurussolutions.com).

Cette interface API offre aux développeurs une boîte à outils puissante. Pour une utilisation classique, on pourrait appeler `llm.generateText({ prompt: "Résumé des ventes au T1", modelParameters: {...} })` puis utiliser `response.text` dans un champ d'enregistrement ou un e-mail. Pour les assistants conversationnels, on peut accumuler des objets `llm.ChatMessage` et utiliser l'option `chatHistory` pour conserver le contexte. Pour le RAG, on construit des documents via `llm.createDocument()`, puis on les transmet. Dans tous les cas, l'objet renvoyé fournit également des détails sur l'utilisation (jetons consommés), qui peuvent être journalisés ou stockés.

En arrière-plan, NetSuite gère l'authentification avec OCI ; le script n'a qu'à référencer le nom du module `N/llm`. Par défaut, Oracle gère les identifiants (modèle mutualisé sur OCI). Les utilisateurs avancés peuvent passer en mode « illimité » en remplissant les préférences IA avec leur propre locataire et clé OCI, auquel cas les appels acceptent les paramètres de configuration OCI (Source: docs.oracle.com) (Source: www.houseblend.io).

Exemple de code : Envoyer une invite (Hello World)

L'utilisation la plus simple de N/LLM est illustrée par l'exemple officiel de NetSuite (Source: docs.oracle.com). Dans un script SuiteScript 2.1, après avoir requis `'N/llm'`, on pourrait écrire :

```

define(['N/llm'], function(llm) {
  // Envoyer une invite "Hello World!" au modèle LLM par défaut.
  const response = llm.generateText({
    prompt: "Hello World!",
    modelParameters: {
      maxTokens: 100,
      temperature: 0.5,
      topK: 40
    }
  });
  // Extraire le texte de la réponse et le journaliser.
  const reply = response.text;
  log.debug('Réponse du LLM', reply);
  // Vérifier les jetons gratuits restants.
  const remaining = llm.getRemainingFreeUsage();
  log.debug('Jetons restants', remaining);
});

```

Cet exemple envoie une invite fixe et récupère `response.text` (Source: docs.oracle.com). Il appelle également `llm.getRemainingFreeUsage()` pour interroger le quota de jetons gratuits restants pour la génération de texte. Dans la console, un développeur verrait la réponse du LLM (par ex. « Hello World to you too! ») et l'utilisation mise à jour. Cela illustre le flux de travail synchrone de base. (Remarque : dans un script déployé sur NetSuite, utilisez `define` comme ci-dessus ; dans le débogueur SuiteScript, on peut utiliser `require` dans l'éditeur d'extraits comme indiqué dans la documentation (Source: docs.oracle.com).)

L'exemple montre d'autres paramètres : `temperature`, `topK`, etc., qui contrôlent la créativité. Par défaut (si omis), NetSuite utilise le modèle Command-A de Cohere avec une température de 0,2. Modifier ces paramètres affecte la variété de la sortie (température/aléatoire plus élevé).

Dans une application réelle, des champs comme `prompt` seraient probablement dynamiques (basés sur des données d'enregistrement ou des entrées utilisateur) plutôt qu'une chaîne codée en dur. Mais cet échantillon clarifie la structure de l'appel. Il souligne également que, par défaut, *tous les appels génératifs sont décomptés d'un pool d'utilisation par compte*. Dans notre exemple, chaque invocation de `generateText` consomme des jetons ; le solde gratuit du compte peut être vérifié par programmation (par ex. pour avertir l'administrateur ou passer en mode payant si le solde est faible).

Autres méthodes : Évaluation d'invite et Embeddings

Au-delà de la simple génération de texte, N/LLM fournit des fonctionnalités de plus haut niveau. Le **Prompt Studio** de NetSuite permet aux administrateurs de définir des invites réutilisables (avec des espaces réservés) et de les regrouper avec des actions Text Enhance. La méthode SuiteScript `llm.evaluatePrompt(options)` permet au script d'*invoquer une invite enregistrée* par ID, en transmettant des valeurs dynamiques. Par exemple, on pourrait définir une invite « Cher {{customer}}, votre solde créditeur est de {{balance}}. » dans Prompt Studio, puis dans SuiteScript appeler `llm.evaluatePrompt({ id: 123, promptValues: {customer: "Acme Corp", balance: "$500"} })`. Le LLM remplit le modèle avec les valeurs données, en utilisant les paramètres de modèle prédéfinis enregistrés avec l'invite. La documentation de l'API d'Oracle appelle cela « Évaluation d'invite » (Source: gurusolutions.com). L'alias `llm.executePrompt` est équivalent. Contrairement à une invite libre, `evaluatePrompt` réutilise des modèles d'invite entraînés et garantit la cohérence entre les utilisations.

Exemple (intégration Prompt Studio) :

```

const response = llm.evaluatePrompt({
  id: 456,
  promptValues: { "reportPeriod": "T1 2026", "totalRevenue": "$1,200,000" }
});
log.debug('Résumé du rapport généré :', response.text);

```

Cela exécuterait le modèle d'invite avec l'ID 456, en y substituant les valeurs.

Pour les embeddings, la méthode est `llm.embed(options)`. Cela convertit le texte en un vecteur numérique. NetSuite utilise actuellement le modèle `embed_v4.0` de Cohere pour cela. Les embeddings permettent la recherche sémantique et les comparaisons de similarité. Les scripts de la documentation montrent, par exemple, le calcul de la similarité cosinus entre les embeddings de descriptions d'articles (Source: docs.oracle.com). Le point important est que **l'utilisation des embeddings a son propre quota**. Les blogs d'Oracle et de partenaires notent que les embeddings consomment des jetons d'un pool mensuel gratuit distinct (Source: docs.oracle.com) (Source: gurussolutions.com). Les développeurs récupèrent les embeddings comme suit :

```
const embedResp = llm.embed({ input: ["Description article A", "Description article B"] });
const vectors = embedResp.embeddings; // Tableau de tableaux de nombres
```

On peut ensuite calculer la similarité (par ex. produit scalaire, cosinus) pour classer les enregistrements associés. Cette approche (« recherche sémantique ») peut trouver des enregistrements similaires ou regrouper des données.

Les méthodes de streaming (`generateTextStreamed`, `evaluatePromptStreamed`) sont identiques dans leur utilisation, sauf qu'elles renvoient un objet `StreamedResponse` spécial sur lequel vous pouvez boucler. Par exemple, dans un Suitelet, vous pourriez télécharger des jetons partiels vers le navigateur client au fur et à mesure de leur arrivée pour une interface de chat plus réactive (Source: gurussolutions.com). Les méthodes acceptent les mêmes `options` mais permettent une lecture en temps réel des valeurs `response.token`.

Exemple : Suitelet de questions-réponses avec RAG (Retrieval-Augmented)

Pour illustrer un cas d'utilisation plus complet, considérons un Suitelet qui répond à des questions en texte libre sur les données NetSuite (ventes, inventaire, etc.). Un récent blog de développeur Oracle explique la création d'un tel Suitelet « Sales Insights » (Source: blogs.oracle.com) (Source: blogs.oracle.com). Les étapes clés sont :

- Formulaire UI** : Créer un formulaire avec une zone de texte pour la question de l'utilisateur et un champ de sortie pour la réponse de l'IA. Inclure une section pour les « Détails avancés » afin d'afficher les informations de support (documents, citations et jetons restants) (Source: blogs.oracle.com).
- Interroger les données** : Lorsque l'utilisateur soumet, utiliser `N/query` ou SuiteQL pour récupérer les enregistrements pertinents. Dans l'exemple, les lignes de commande client de 2016–2017 sont interrogées (avec ventilation par article, quantité, revenu, emplacement) (Source: blogs.oracle.com) (Source: blogs.oracle.com).
- Créer des documents** : Pour chaque ligne/résultat, formater un court extrait de texte (par ex. "Article : Téléphone X, Quantité vendue : 250, Revenu : 50 000 \$") et appeler `llm.createDocument({ text: snippet, reference: recordId })`. Rassembler ceux-ci dans un tableau `documents[]`. Cela fournit un contexte structuré au LLM (Source: blogs.oracle.com).
- Générer la réponse** : Appeler `llm.generateText({ prompt: userQuestion, documents: documents, modelParameters: { maxTokens: 200, temperature: 0.3 } })`. Parce que `documents` est fourni, le LLM les utilise comme base factuelle. La méthode renvoie un `llm.Response` avec `.text` (la réponse) et `.citations` (indiquant quels documents insérés ont contribué à la réponse) (Source: blogs.oracle.com).
- Analyser et afficher** : Afficher `response.text` dans le champ de réponse. Dans la section avancée, lister les documents originaux et les points saillants des `response.citations`. Afficher également `llm.getRemainingFreeUsage()`.

En suivant ce modèle, le Suitelet garantit que les réponses à des questions telles que « *Quel a été le produit le plus vendu au T1 2017 ?* » sont basées sur les données réelles de NetSuite. Le LLM ne « hallucine » pas de faits aléatoires ; au lieu de cela, il fait référence aux documents fournis. Cet exemple démontre comment N/LLM permet **l'interrogation en langage naturel des données ERP** avec un minimum de codage.

Limites d'utilisation et gouvernance

NetSuite régit strictement l'utilisation de N/LLM pour équilibrer utilité et coût. Comme indiqué, chaque compte dispose d'un **pool d'utilisation gratuite limité** pour les appels IA, réinitialisé mensuellement (Source: www.houseblend.io). Plus précisément, les pools sont suivis séparément pour la *génération de texte* et les *embeddings* (Source: www.houseblend.io) (Source: docs.oracle.com). Par exemple, si vous exécutez un appel

`generateText` avec une invite de 100 jetons et une réponse LLM, cela consomme 100 jetons du pool de génération de texte. Si vous appelez ensuite `embed` sur des chaînes, ces jetons proviennent du pool d'embedding. Le nombre exact de jetons gratuits n'est pas documenté publiquement (il varie probablement selon le compte ou l'édition) (Source: www.houseblend.io), mais les administrateurs peuvent surveiller la consommation dans Configuration > Société > Préférences IA. L'API SuiteScript fournit également les méthodes `llm.getRemainingFreeUsage()` et `llm.getRemainingFreeEmbedUsage()` pour vérifier le solde restant pour le mois en cours (Source: docs.oracle.com) (Source: docs.oracle.com).

Chaque appel de méthode est évalué en unités de gouvernance NetSuite, mais celles-ci correspondent à l'utilisation des jetons. Si un script tente de dépasser le quota gratuit, il recevra une erreur. Pour continuer l'utilisation, il faut soit attendre la prochaine réinitialisation mensuelle, soit configurer la facturation OCI. En ajoutant une information d'identification OCI Generative AI externe dans les paramètres IA (pointant vers la location OCI du client), tous les appels N/LLM puiseront dans le quota de ce compte externe (sous réserve des limites OCI) (Source: www.houseblend.io) (Source: docs.oracle.com). Ce changement débloque une utilisation illimitée (au-delà du plafond gratuit de NetSuite) mais entraîne des frais OCI standard (par modèle et par million de jetons, etc.).

Outre les quotas de jetons, Oracle impose une **limite de concurrence de 5 appels simultanés**. Cela signifie qu'un script ou plusieurs scripts en parallèle peuvent avoir au maximum cinq appels `generateText` (ou `evaluatePrompt`) actifs à la fois, et de même cinq appels `embed`. La documentation indique que « jusqu'à cinq appels LLM simultanés sont autorisés pour les méthodes de génération et jusqu'à cinq pour l'embedding » (Source: www.houseblend.io). Tenter un 6ème appel simultané provoquera immédiatement une erreur. En pratique, les scripts serveur sont généralement séquentiels, mais cette limite est pertinente si, par exemple, un script planifié démarre plusieurs threads parallèles ou si un Suitelet génère plusieurs requêtes en arrière-plan.

NetSuite assure la transparence : la page des préférences IA (« SuiteScript AI Preferences ») affiche un tableau de l'utilisation par mois et par type (Source: docs.oracle.com) (Source: www.houseblend.io). Initialement, le tableau indiquera « Aucun enregistrement à afficher » jusqu'à ce que le premier appel IA réussi soit effectué (Source: docs.oracle.com). Une fois rempli, les administrateurs peuvent voir, pour chaque mois, le nombre total de jetons consommés pour « Générer » et « Embed » et la quantité de quota gratuit utilisée par rapport à celle restante. Cela aide à la budgétisation et au dépannage (par ex. identifier une augmentation soudaine de l'utilisation ou une information d'identification OCI expirée).

La gouvernance intégrée encourage l'efficacité. Par exemple, les API de streaming (`generateTextStreamed`) sont utiles pour réduire la *latence perçue* sur les grandes réponses (Source: www.houseblend.io), mais chaque jeton compte toujours de la même manière dans l'utilisation. De plus, les scripts doivent éviter les appels répétés inutiles. La mise en cache ou la réutilisation des réponses (le cas échéant) peut préserver le quota.

Exemples de cas d'utilisation et avantages

Le module N/LLM permet un large éventail de capacités pilotées par l'IA au sein de NetSuite. Quelques exemples illustratifs :

- Chatbots interactifs** : Fournissez une aide contextuelle ou un chat d'assistance au sein de NetSuite. Par exemple, un Suitelet de centre d'assistance peut utiliser N/LLM pour répondre à des questions telles que « Comment puis-je affecter une recherche enregistrée à un tableau de bord ? » ou « Quel est le code de taxe de vente approprié pour un fournisseur ? » en interrogeant un LLM avec la question de l'utilisateur (Source: www.techfino.com) (Source: www.houseblend.io). Étant donné que la documentation de support de NetSuite est intégrée à l'entraînement du LLM (via des partenariats Oracle) et/ou fournie sous forme de prompts, les réponses peuvent être très précises. Techfino (un partenaire NetSuite) note que l'utilisation la plus sûre consiste à poser des questions procédurales générales qui n'impliquent pas de données en temps réel, par exemple « Comment activer les filtres de recherche globale ? » (Source: www.techfino.com). Cela réduit le volume de tickets du centre d'assistance et guide les utilisateurs directement dans l'application (Source: www.techfino.com) (Source: www.techfino.com).
- Requête de données (recherche en langage naturel)** : Comme illustré dans l'exemple du Suitelet, les utilisateurs peuvent *poser des questions sur les données transactionnelles*. Au lieu d'écrire du SuiteQL, un analyste pourrait taper « Quel a été notre produit le plus vendu au dernier trimestre ? » et recevoir un résumé tout en voyant les enregistrements justificatifs listés. Ce scénario de « dialogue avec mes données » a même été mis en avant par Oracle : le titre de leur blog « Now You Can Talk to NetSuite » fait référence à l'utilisation de N/LLM combiné au RAG pour débloquent des requêtes personnalisées (Source: blogs.oracle.com) (Source: blogs.oracle.com). Houseblend et Oracle soulignent que les clients disposant de données riches (par exemple, commandes de détail, historiques d'inventaire) sont directement avantagés (Source: houseblend.io) (Source: www.houseblend.io). Par exemple, BirdRock Home (un détaillant d'articles de maison avec « des milliers de produits » et un volume de commandes élevé) peut utiliser N/LLM pour poser des questions sur la chaîne d'approvisionnement ou générer des rapports de tendance à partir de ses données NetSuite (Source: houseblend.io) (Source: www.houseblend.io).
- Génération automatique de contenu** : Des tâches telles que le remplissage de champs libres, la génération de résumés ou la création de documents texte peuvent être automatisées. La fonctionnalité **Text Enhance** d'Oracle (annoncée en février 2025) utilise N/LLM en arrière-plan. Par exemple, un utilisateur financier créant une écriture de journal pourrait recevoir des suggestions générées par l'IA pour le champ mémo ou un

e-mail de communication client en utilisant `llm.generateText` avec un prompt tel que « rédige une note pour la facture n°1234 basée sur ces détails... ». Le LLM peut intégrer des données ERP pertinentes (montant de la facture, conditions du fournisseur, etc.) pour rédiger un texte précis. De même, la copie de descriptions de produits, la traduction de documents de politique ou la rédaction de propositions peuvent être accélérées.

- Recherche sémantique et recommandation** : En intégrant du texte (noms d'articles, notes de fournisseurs, mémos de transaction) dans un espace vectoriel, les scripts NetSuite peuvent implémenter la recherche sémantique. Par exemple, l'exemple de Suitelet « Trouver des articles similaires » utilise `llm.embed` pour calculer des embeddings pour les descriptions d'articles, puis calcule la similarité cosinus (Source: docs.oracle.com). Le résultat est une fonctionnalité dynamique d'articles liés : cliquez sur un article et voyez une liste classée d'articles « similaires » par leur sens (et non seulement par mots-clés). Cela peut aider au merchandising ou à la détection de fraudes (par exemple, trouver des descriptions en double, des produits connexes, etc.). Les embeddings peuvent également servir de base à des moteurs de recommandation (suggestions de ventes croisées) ou au regroupement d'enregistrements.
- Reporting et insights améliorés** : Au-delà des questions-réponses ponctuelles, N/LLM peut générer des rapports narratifs. Un Suitelet ou un script planifié pourrait s'exécuter chaque nuit, résumant les indicateurs clés en prose. Par exemple : « La marge brute pour le T4 était de X \$ (en hausse de 5 % sur un an). La région la plus performante était la côte Ouest avec Y \$. La rotation des stocks s'est améliorée de 10 %. Actions clés : recommander 100 unités de la pièce Z. » Un tel récit peut être directement écrit dans des tableaux de bord ou envoyé par e-mail aux responsables. Ce type d'analyse en langage naturel a été un point fort des annonces de SuiteWorld. Des prompts prédéfinis (via Prompt Studio) peuvent standardiser ces rapports, et les scripts peuvent les remplir avec les valeurs de données actuelles.
- ChatOps et agents** : Oracle introduit déjà des « agents » IA dans divers modules (CPQ, SuiteAnswers). En pratique, un développeur NetSuite pourrait créer des flux de travail personnalisés pilotés par l'IA. Par exemple, un chatbot dans un Suitelet de service pourrait non seulement répondre aux FAQ, mais aussi créer des enregistrements sur commande (par exemple, « créer une demande de service pour le client X »), en combinant N/LLM avec des actions SuiteScript. Cela suit la tendance émergente de l'« IA agentique » où les modèles génératifs déclenchent des actions (Source: www.techradar.com). Bien que les scripts N/LLM actuels soient principalement des actions de « génération ou évaluation », le contrôle sous-jacent de SuiteScript signifie que l'on peut analyser les réponses du LLM et effectuer une automatisation plus poussée.

Ces cas d'utilisation soulignent un potentiel transformationnel : les tâches banales (comme la rédaction d'e-mails, la synthèse de données, la réponse à des questions routinières) peuvent être déléguées à l'IA, libérant ainsi les utilisateurs pour un travail à plus forte valeur ajoutée. L'intégration native de NetSuite signifie que les développeurs n'ont pas à assembler des outils disparates ; la logique de l'IA s'exécute là où résident les données, préservant la sécurité et réduisant les frais généraux d'intégration. Par exemple, Techfino note que N/LLM est « confidentiel par conception » – aucune donnée sensible ne fuit à moins d'être incluse intentionnellement (Source: www.techfino.com).

Études de cas et perspectives de l'industrie

Des preuves concrètes de l'impact de l'IA dans NetSuite commencent à émerger. Publiquement, Oracle a mis en avant des clients tirant parti de l'IA (même si ce n'est pas toujours de l'IA générative). Houseblend rapporte que BirdRock Home utilise l'entrepôt d'analyse et l'IA de NetSuite pour la prévision et l'aide à la décision (Source: www.houseblend.io). Le profil de BirdRock indique qu'il traite « des milliers de commandes par jour » au sein de NetSuite (Source: houseblend.io), ce qui implique des données riches pour l'IA. En février 2023, des études de marché ont montré que BirdRock avait sélectionné les plateformes d'IA générative d'Oracle NetSuite pour l'analyse et la BI (Source: www.appsruntheworld.com). Cela suggère que BirdRock utilise activement les capacités d'IA générative de NetSuite pour l'analyse. De même, Overture Promotions (une société de services marketing) est citée comme utilisant les analyses de NetSuite pour améliorer les décisions de la chaîne d'approvisionnement (Source: www.houseblend.io). Houseblend résume : ces clients ont appliqué l'IA/ML sur les données ERP (BirdRock pour la prédiction du taux de désabonnement, Overture pour la prévision de la chaîne d'approvisionnement) et ont obtenu des « améliorations stratégiques exploitables » (Source: www.houseblend.io). Bien que les études de cas de BirdRock et Overture décrivent principalement l'analyse/la prévision, elles illustrent le type d'enrichissement de données par l'IA que le module N/LLM peut exploiter.

Les analystes de l'industrie renforcent cette tendance. Des cabinets comme Gartner et TechTarget notent que l'IA est désormais une fonctionnalité intégrée des ERP modernes. Par exemple, itpro.com rapporte que les annonces de SuiteWorld 2023 étaient centrées sur l'IA intégrée à NetSuite (Owen Pettifor, TechTarget) (Source: www.pnnewswire.com) (Source: www.oracle.com). Les rapports d'Axios font écho à la position d'Oracle selon laquelle l'intégration de l'IA sans coût supplémentaire est une stratégie délibérée (Source: www.houseblend.io), contrairement aux concurrents qui vendent souvent des modules complémentaires d'IA. Une analyse de TechRadar avertit que de nombreux projets pilotes d'IA générative ont échoué en raison d'un manque d'intégration : « plus de 95 % des projets pilotes d'IA non structurés échouent sans une intégration ciblée » (Source:

[houseblend.io](https://www.houseblend.io)). Cela souligne que si N/LLM est puissant, le succès dépend d'une mise en œuvre disciplinée. Comme l'a plaisanté un expert, le déploiement de N/LLM « dépend d'une mise en œuvre structurée alignée sur la gouvernance des données d'entreprise » (Source: www.houseblend.io).

En résumé, les anecdotes des clients et les études de l'industrie indiquent que les entreprises disposant de **données NetSuite volumineuses et de cas d'utilisation clairs** sont prêtes à tirer profit de l'IA générative. BirdRock et Overture sont des exemples d'entreprises riches en données. La dynamique mondiale (McKinsey, Gartner) montre une courbe d'adoption de l'IA en croissance rapide (Source: www.houseblend.io) (Source: www.techradar.com), et le choix d'intégration de NetSuite le place dans une position forte pour offrir ces avantages.

Confidentialité des données, sécurité et bonnes pratiques

Bien que l'IA générative ouvre de nouvelles capacités, elle soulève également des préoccupations concernant **la confidentialité des données, la sécurité et la conformité**. La conception de NetSuite tente de résoudre ces problèmes intrinsèquement. Par défaut, **aucun enregistrement NetSuite ou donnée client n'est envoyé au LLM à moins qu'un script ne le lise explicitement et ne l'inclue dans le prompt** (Source: www.techfino.com). Cette « confidentialité par conception » est essentielle : comme le note Techfino, le LLM « *n'a pas accès à vos enregistrements NetSuite à moins que vous ne passiez explicitement ces données via un script* » (Source: www.techfino.com). En d'autres termes, les appels génératifs se produisent via NiS dans le cloud de NetSuite, et seuls le prompt et le texte du document (fournis par le script) sont envoyés à OCI. Oracle assure explicitement aux clients qu'« aucune donnée client n'est partagée avec les fournisseurs de LLM ou vue par d'autres clients ou des tiers » (Source: www.oracle.com). De plus, si une entreprise entraîne un modèle personnalisé sur ses données, seule cette entreprise peut l'utiliser. Toutes les actions d'IA respectent la sécurité basée sur les rôles de NetSuite : un script s'exécutant sous les autorisations d'un utilisateur donné ne peut lire que les champs que cet utilisateur est autorisé à voir, et la sortie ne peut pas exposer de données cachées.

Néanmoins, les administrateurs et les architectes doivent être prudents. Les bonnes pratiques (reprises par Oracle et ses partenaires) incluent :

- Limiter les données sensibles dans les prompts :** Ne transmettez que les données sûres et nécessaires. Les informations personnellement identifiables (PII) ou les données de carte de crédit, par exemple, ne devraient jamais être transmises à un modèle externe. Techfino conseille explicitement de « *ne transmettre que des données vérifiées que vous êtes à l'aise de partager* » (Source: www.techfino.com). Par exemple, lors de l'interrogation des soldes clients, le script pourrait inclure les totaux et le statut du compte, mais omettre les coordonnées du client.
- Valider toutes les sorties de l'IA :** Parce que les LLM produisent du texte « *créatif* », NetSuite avertit que les réponses « *doivent être vérifiées pour leur exactitude* » (Source: www.houseblend.io). Un modèle recommandé consiste à traiter les sorties du LLM comme des *brouillons* ou des suggestions. Par exemple, si vous générez un e-mail client, le script devrait exiger une révision ou une approbation humaine avant l'envoi. Houseblend souligne qu'environ 95 % des projets pilotes d'IA non ciblés échouent par manque de révision (Source: www.houseblend.io). Les processus automatisés devraient inclure des vérifications de cohérence. Quelques options : demander au script de vérifier certains faits (par exemple, que les montants correspondent aux montants dans les données) ; utiliser des scores de confiance lorsque disponibles ; ou acheminer les sorties via une intervention humaine.
- Limiter le débit et surveiller l'utilisation :** Gardez un œil sur les quotas et implémentez des vérifications afin qu'un bug ou une entrée malveillante n'épuise pas les jetons. Les méthodes d'API `getRemainingFreeUsage()` et `getRemainingFreeEmbedUsage()` peuvent être utilisées dans les scripts pour avertir ou arrêter les opérations si l'utilisation est faible (Source: www.houseblend.io). De même, gérez les erreurs générées en cas d'épuisement du quota avec élégance.
- Utiliser des alias et des méthodes asynchrones si nécessaire :** De nombreuses méthodes ont des versions asynchrones basées sur des promesses (par exemple, `llm.generateText.promise()`) et des variantes de streaming. L'utilisation de promesses peut éviter un comportement bloquant dans le code de style Node.js. Des alias comme `llm.chat()` (pour `generateText`) ou `llm.executePrompt()` (pour `evaluatePrompt`) peuvent rendre le code plus intuitif. GURUS Solutions recommande de tirer parti de ces commodités (Source: gurussolutions.com).
- Concevoir des pipelines de données :** Pour les scénarios RAG, l'étape de récupération est critique. Les données NetSuite doivent être préparées avec soin. Cela peut impliquer des scripts planifiés pour pré-construire et actualiser des corpus de « documents ». Par exemple, des scripts planifiés s'exécutant quotidiennement pourraient exporter des résumés clés dans un enregistrement personnalisé ou un magasin JSON, garantissant que le LLM voit toujours des informations à jour. Houseblend note que les données ERP nécessitent souvent une « planification minutieuse » pour être encodées en tant que prompts (Source: houseblend.io). L'utilisation intelligente de SuiteQL et des recherches enregistrées peut garantir que les données pertinentes sont mises en évidence (par exemple, l'exemple Sales Insights utilise SuiteQL pour récupérer efficacement les indicateurs de résumé (Source: blogs.oracle.com)).

- **Rester conscient de la région** : Les fonctionnalités d'IA générative sont déployées région par région. Vérifiez que le tenancy OCI GenAI souhaité est disponible dans votre région NetSuite. La documentation d'Oracle (par exemple, « Disponibilité des fonctionnalités d'IA générative ») doit être consultée. Si un script est déployé dans une région sans support GenAI, les appels échoueront.

En suivant ces directives, les organisations peuvent exploiter la puissance de N/LLM tout en minimisant les risques. Il est essentiel que les entreprises intègrent N/LLM dans un cadre de gouvernance de l'IA plus large — en alignant les politiques d'utilisation des données et les règles de conformité comme elles le feraient pour toute autre intégration sensible (Source: www.houseblend.io).

Analyse des données et considérations de performance

Bien que N/LLM soit principalement une intégration d'API, il existe certains aspects de performance et d'analyse que les développeurs devraient prendre en compte :

- **Taille du prompt et de la réponse** : Les modèles génératifs ont des limites de jetons (le modèle Cohere Command-A d'Oracle peut gérer une certaine taille de contexte maximale). Le paramètre `modelParameters.maxTokens` limite la longueur de la réponse. Les requêtes qui dépassent les limites du modèle généreront une erreur. En règle générale, gardez les prompts concis et envisagez de résumer ou de tronquer les documents sources. Les documents très longs (par exemple, des pages HTML complètes) doivent être pré-traités en points ou résumés plus courts avant d'être transmis.
- **Comptabilité des jetons** : Chaque mot ou symbole compte dans l'utilisation. Une bonne pratique consiste à mesurer la consommation de jetons pour planifier les coûts. La propriété `.usage` de l'API ou des méthodes distinctes (`getRemainingFreeUsage`) facilitent cela. Par exemple, si un prompt+réponse moyen consomme environ 500 jetons, le pool gratuit de 50 000 jetons d'une entreprise permet environ 100 appels de ce type par mois. Les tailles réelles des pools gratuits ne sont pas publiques, mais la surveillance peut les déterminer empiriquement. Connaître votre utilisation des jetons peut également vous indiquer quand provisionner la facturation OCI ou limiter l'utilisation.
- **Latence** : Les appels génératifs nécessitent des allers-retours vers OCI. En pratique, les scripts doivent supposer une latence de quelques centaines de millisecondes à quelques secondes par appel, selon la charge du modèle et la taille du prompt. L'exécution synchrone du script sera suspendue jusqu'au retour du résultat. Par conséquent, les scripts d'événement qui s'exécutent sur les formulaires utilisateur (déclencheurs de script client) doivent être utilisés avec précaution. Les Suitelets (côté serveur) ou le traitement asynchrone sont des cas d'utilisation plus typiques. Utilisez les méthodes de streaming ou un indicateur de chargement dans les interfaces utilisateur pour améliorer l'expérience utilisateur.
- **Traitement par lots et parallélisme** : Si un script a besoin de plusieurs éléments de données générés, envisagez de les traiter par lots en un seul appel (si cela a un sens sémantique) plutôt que par des appels séparés. Cependant, la limite de concurrence (5) permet quelques requêtes parallèles si elles sont effectuées correctement. Par exemple, un RESTlet recevant plusieurs prompts pourrait émettre jusqu'à 5 appels `generateText` en parallèle en utilisant des variantes de promesse, puis les attendre ensemble. Mais au-delà, les appels seront mis en file d'attente ou échoueront.
- **Analyses d'utilisation** : Les organisations devraient suivre les modèles d'utilisation. Par exemple, il est possible de créer une recherche enregistrée ou un enregistrement personnalisé pour journaliser le contexte et le résultat de chaque appel LLM (utilisateur, date, type d'appel, jetons utilisés). Avec le temps, l'analyse de ces données pourrait révéler les utilisateurs intensifs, les prompts réussis par rapport aux échecs, ou les opportunités d'optimisation des prompts. Cela peut également alimenter les modèles de CMDB ou de refacturation interne pour l'utilisation de l'IA.

Discussion, implications et orientations futures

L'intégration de l'IA générative par NetSuite via le module N/LLM est à la fois un **facteur limitant** et un puissant catalyseur pour l'automatisation axée sur les données. Les implications sont les suivantes :

- **Démocratisation de l'IA dans l'ERP** : En intégrant l'IA au niveau de la plateforme, NetSuite rend les analyses avancées et les capacités linguistiques accessibles aux clients qui pourraient manquer d'expertise interne en IA. Les utilisateurs n'ont plus besoin de configurer des services d'IA externes ou d'écrire des intégrations complexes. Cela abaisse la barrière à l'entrée pour les petites et moyennes entreprises souhaitant utiliser l'IA. Le fait qu'Oracle ait inclus cette fonctionnalité dans le produit de base (sans licence IA premium) suggère un changement dans l'industrie : l'IA passe du stade expérimental à celui d'attente standard (Source: www.prnewswire.com) (Source: www.houseblend.io).
- **Gains de productivité commerciale** : S'il est bien utilisé, le module N/LLM peut rationaliser les flux de travail. Le vice-président d'Oracle, Evan Goldberg, résume bien la situation : ces fonctionnalités d'IA « conseillent et assistent nos clients et les aident à stimuler leur productivité, à augmenter leurs profits et à optimiser leur activité » (Source: www.prnewswire.com). Par exemple, les équipes commerciales peuvent

automatiser la génération de propositions ; les agents de support peuvent obtenir des réponses instantanées de la base de connaissances ; la finance peut produire des ébauches d'analyses en quelques secondes. Cela se traduit par des économies de temps réelles et une réduction potentielle des coûts. Les études de McKinsey et Gartner suggèrent qu'un impact économique majeur est possible si les entreprises exploitent pleinement ces outils (Source: www.houseblend.io) (Source: www.techradar.com).

- Nouveaux paradigmes de développement** : Les développeurs SuiteScript pensent désormais en termes de *sorties de LLM*. Les déclencheurs de données traditionnels (par exemple, « au chargement de l'enregistrement, faire X ») peuvent être augmentés par « à la demande, demander à l'IA... ». Nous pourrions voir une croissance des SuiteApps axées sur l'IA (par exemple, assistants de chat, connecteurs de recherche intelligente). L'API de gestion des prompts (annoncée en 2025) et Prompt Studio sont étroitement couplés au module N/LLM : il est probable que les développeurs SuiteScript intègrent des modèles de prompts dans le code (par exemple, en invoquant des prompts d'entreprise standard pour plus de cohérence) et partagent les solutions d'IA plus largement. L'écosystème pourrait évoluer de la même manière que les outils de BI ont standardisé le reporting SQL ; nous pourrions bientôt avoir des « prompts standard » pour les résumés financiers, les ébauches juridiques, etc.
- Gouvernance des données et accent sur la qualité** : Comme de nombreux analystes le soulignent, le succès des fonctionnalités d'IA dépend de la qualité et de la gouvernance des données. Le taux d'échec élevé des projets pilotes d'IA sans objectif clair (Source: www.houseblend.io) suggère que les entreprises investiront dans le nettoyage et la structuration de leurs données ERP pour alimenter ces outils. Par exemple, s'assurer que les fiches clients contiennent des descriptions riches permet au LLM de générer de meilleurs profils clients. Les organisations définiront probablement des politiques sur les données pouvant être traitées (HIPAA, RGPD, etc.). La mise en œuvre du module N/LLM en production nécessitera souvent un effort de gestion du changement : formation des utilisateurs, mise en place de processus de révision et mesure du retour sur investissement.
- Dynamique concurrentielle** : La stratégie de NetSuite contraste avec celle de certains rivaux. Houseblend et les rapports d'actualité notent que NetSuite **inclut** l'IA dans sa plateforme de base, alors que les concurrents (par exemple, SAP avec « Joule », Microsoft avec Copilot) regroupent souvent l'IA en tant qu'extension premium ou exigent une licence séparée (Source: www.houseblend.io) (Source: netsuite.folio3.com). Cela pourrait constituer un avantage concurrentiel pour NetSuite, attirant les clients qui souhaitent utiliser l'IA sans frais élevés. Cependant, à mesure que davantage de fournisseurs déploient l'IA, les clients exigeront une innovation continue. NetSuite devra probablement continuer à étendre le module N/LLM (nouveaux modèles, vision, etc.) pour garder une longueur d'avance.
- Fonctionnalités futures** : À l'avenir, nous nous attendons à ce que le module N/LLM gagne en capacités. Les orientations possibles incluent la prise en charge de **familles de modèles supplémentaires** (par exemple, l'intégration avec des modèles d'OpenAI ou de Google, une fois approuvés pour OCI) et la gestion de dialogues multi-tours (suivi de session sur plusieurs appels). L'engagement d'Oracle envers OCI (avec l'infrastructure « Supercluster » GPU (Source: www.oracle.com) suggère qu'ils ajouteront des modèles plus puissants ou spécialisés au fil du temps. La mention formelle des rôles de chat et des messages dans l'API laisse entrevoir des flux de travail conversationnels. Nous pourrions également voir des *agents* plus profonds — par exemple, une réponse LLM qui planifie automatiquement des tâches ou met à jour des enregistrements via des hooks SuiteScript (une idée connue sous le nom d'IA agentique (Source: www.techradar.com)). L'intégration avec des outils d'analyse et de planification (comme un lien plus étroit entre N/LLM et l'Analytics Warehouse pour le RAG à grande échelle) est probable.
- Considérations éthiques et de conformité** : Plus d'IA signifie plus de surveillance. Les régulateurs en Europe et dans d'autres régions commencent à se pencher sur l'utilisation de l'IA. Les entreprises utilisant N/LLM devront peut-être assurer la conformité (par exemple, journaux de conservation des données, explicabilité des décisions générées par l'IA). NetSuite (Oracle) devra fournir des fonctionnalités de transparence, et les clients exigeront des audits (qui a posé quelle question, sur quelle base le modèle s'est appuyé, etc.). La fonctionnalité de citation intégrée aide à retracer les réponses jusqu'aux données sources, ce qui est un point positif.
- Impact sur l'écosystème plus large** : Enfin, la disponibilité du module N/LLM pourrait susciter des innovations au-delà de NetSuite. Par exemple, des développeurs indépendants pourraient créer des SuiteApps basées sur l'IA (marketplace), ou intégrer l'IA de NetSuite avec des flux de données/IoT. Le réseau SDN (SuiteCloud Dev Network) et les partenariats accéléreront les services d'intégration d'IA. Les supports de formation et les exemples communautaires (comme ce guide) vont proliférer, aidant les développeurs à élaborer des meilleures pratiques.

En résumé, l'intégration de N/LLM dans NetSuite signale que *l'IA générative en entreprise est arrivée*. Sa conception réfléchie (soutenue par OCI, sécurisée, avec gouvernance) répond à de nombreuses préoccupations qui ont condamné les premiers projets pilotes d'IA. La combinaison des données riches de NetSuite et de la plateforme d'IA générative d'Oracle offre un puissant catalyseur. Cependant, la création d'une valeur commerciale réelle nécessitera une adoption disciplinée : se concentrer sur des cas d'utilisation clairs, surveiller les résultats et intégrer les sorties dans les flux de travail. Les entreprises qui le font peuvent s'attendre à des gains de productivité et à de meilleures stratégies basées sur les données ; celles qui négligent la gouvernance pourraient se retrouver parmi les 95 % de projets pilotes ayant échoué (Source: www.houseblend.io).

Conclusion

Le module SuiteScript N/LLM de NetSuite est un ajout révolutionnaire au paysage ERP, transformant efficacement la plateforme en un environnement natif pour l'IA. En fournissant une interface transparente vers les LLM au sein du framework familier SuiteScript, il ouvre une myriade de possibilités : de la rédaction de descriptions de produits, à l'élaboration de résumés financiers et à la réponse aux requêtes des utilisateurs, jusqu'à l'activation d'analyses et de chatbots sophistiqués basés sur le RAG. Ce guide a exploré le contexte historique, les détails techniques et les exemples pratiques d'utilisation de N/LLM. Les points clés incluent : l'API riche du module (voir Tableau 2), sa dépendance au service d'IA générative d'OCI, l'importance de fournir des données d'entreprise pour étayer les réponses, et les mécanismes de gouvernance (quotas, concurrence, surveillance de l'utilisation) mis en œuvre par NetSuite.

Notre analyse indique un fort potentiel de valeur. Les données de l'industrie montrent une adoption et un enthousiasme rapides pour l'IA générative (Source: www.houseblend.io) (Source: www.techradar.com), et des clients NetSuite comme BirdRock et Overture exploitent déjà l'IA pour obtenir des informations (Source: www.houseblend.io) (Source: www.houseblend.io). Parce que NetSuite intègre l'IA générative sans coût de licence supplémentaire (Source: www.pnnewswire.com) (Source: www.houseblend.io), les clients ont une barrière à l'entrée faible pour expérimenter. Cependant, une expertise est nécessaire : les entreprises doivent traiter N/LLM comme une capacité stratégique, et non comme un gadget. Les meilleures pratiques (ne transmettre que des données sécurisées, réviser les résultats, s'aligner sur la gouvernance des données) sont essentielles pour éviter les pièges. Les études de cas et les rapports d'experts avertissent que les projets d'IA sans cible échouent couramment (Source: houseblend.io) (Source: www.houseblend.io). En revanche, une mise en œuvre qui source des données de haute qualité, conçoit des prompts réfléchis et intègre des contrôles peut débloquer une automatisation et des perspectives puissantes.

À l'avenir, nous prévoyons que le module N/LLM et son écosystème évolueront rapidement. Les futures améliorations d'Oracle (nouveaux agents, plus de fonctionnalités d'IA dans Suite, prise en charge de modèles divers, etc.) élargiront ce que les développeurs SuiteScript peuvent faire. Les organisations qui investissent dès maintenant dans la compréhension et l'exploitation de N/LLM seront bien positionnées pour récolter les premiers avantages. Plus précisément, nous recommandons aux lecteurs de :

- **Expérimenter avec le module** dans des environnements sandbox pour apprendre l'API et mesurer les modèles d'utilisation.
- **Commencer par des applications à faible risque**, telles que des chatbots de help-desk utilisant des prompts générés par script non sensibles (Source: www.techfino.com) (Source: www.techfino.com) ou la synthèse automatique de documents, pour se familiariser.
- **Impliquer les parties prenantes de la gouvernance des données** dès le début, en s'assurant que les données transmises au modèle sont conformes aux politiques.
- **Surveiller et itérer** : suivre l'utilisation des jetons, les journaux d'erreurs et les commentaires des utilisateurs pour affiner les prompts et les cas d'utilisation.
- **Rester informé** : surveiller les mises à jour d'Oracle (par exemple, API de gestion des prompts, modèles supplémentaires) et les actualités de l'industrie (versions des fonctionnalités OCI GenAI) pour améliorer continuellement les solutions.

En conclusion, N/LLM est une nouvelle capacité puissante qui, si elle est exploitée correctement, peut transformer NetSuite d'un référentiel de données passif en un système intelligent et interactif. Il incarne la tendance de « l'IA générative opérationnelle » dans les logiciels d'entreprise. Les preuves suggèrent que l'adoption généralisée de l'IA générative dans l'ERP est en marche (Source: www.houseblend.io) (Source: www.techradar.com), et l'approche de NetSuite a placé la barre très haut. Ce rapport visait à documenter minutieusement le fonctionnement de N/LLM et la manière dont il peut être appliqué ; les entreprises devraient maintenant réfléchir à la façon d'intégrer ces connaissances pour obtenir un avantage stratégique.

Références

- Aide Oracle NetSuite SuiteScript – Exemples de scripts du module N/llm (exemples de code) (Source: docs.oracle.com) (Source: docs.oracle.com).
- Aide Oracle NetSuite SuiteScript – llm.generateText(options) (référence API) (Source: docs.oracle.com) (Source: docs.oracle.com).
- Aide Oracle NetSuite SuiteScript – Limite d'utilisation de l'IA et utilisation (Source: docs.oracle.com) (Source: docs.oracle.com).
- Blogs Oracle : « Now You Can Talk to NetSuite: Unlock Your Data with N/LLM and Intelligent Querying » par Wilman Arambillette (29 avril 2025) (Source: blogs.oracle.com) (Source: blogs.oracle.com).
- Communiqué de presse Oracle : « NetSuite Extends AI Capabilities ... no additional cost » (SuiteConnect 2025) (Source: www.pnnewswire.com) (Source: www.pnnewswire.com).

- Communiqué de presse Oracle : « NetSuite Embeds Generative AI ... SuiteWorld 2023 » (17 oct. 2023) (Source: www.oracle.com) (Source: www.oracle.com).
- Blog partenaire NetSuite (GURUS Solutions) : « NetSuite N/llm Module » (aperçu des fonctionnalités) (Source: gurussolutions.com) (Source: gurussolutions.com).
- Blog Techfino (partenaire Oracle) : « NetSuite Chatbot Assistant: Using the N/llm Module » (juillet 2025) (Source: www.techfino.com) (Source: www.techfino.com).
- Houseblend : « NetSuite N/LLM Module: Methods, Limits & RAG » (février 2026) (Source: www.houseblend.io) (Source: www.houseblend.io).
- Houseblend : « Enriching NetSuite Data: A Guide to the N/LLM Module » (nov. 2025) (Source: houseblend.io) (Source: houseblend.io).
- Houseblend : « Building AI/ML Models on NetSuite » (nov. 2025) (Source: www.houseblend.io) (Source: www.houseblend.io).
- Gartner (via Houseblend) – Prévisions d'adoption de l'IA (2023) (Source: www.houseblend.io).
- Analyse McKinsey (2023) – impact économique de l'IA générative (Source: www.houseblend.io).
- TechRadar / Tom's Hardware – articles sur les tendances de l'IA en entreprise (2025) (Source: www.techradar.com) (Source: www.houseblend.io).
- Documentation Oracle NetSuite (module N/llm, Prompt Studio, etc.) via docs.oracle.com (Source: docs.oracle.com) (Source: docs.oracle.com).

Étiquettes: netsuite-nllm, api-suitescript, ia-generative, integration-api, implementation-rag, personnalisation-erp, oracle-cloud-infrastructure

AVERTISSEMENT

Ce document est fourni à titre informatif uniquement. Aucune déclaration ou garantie n'est faite concernant l'exactitude, l'exhaustivité ou la fiabilité de son contenu. Toute utilisation de ces informations est à vos propres risques. Houseblend ne sera pas responsable des dommages découlant de l'utilisation de ce document. Ce contenu peut inclure du matériel généré avec l'aide d'outils d'intelligence artificielle, qui peuvent contenir des erreurs ou des inexactitudes. Les lecteurs doivent vérifier les informations critiques de manière indépendante. Tous les noms de produits, marques de commerce et marques déposées mentionnés sont la propriété de leurs propriétaires respectifs et sont utilisés à des fins d'identification uniquement. L'utilisation de ces noms n'implique pas l'approbation. Ce document ne constitue pas un conseil professionnel ou juridique. Pour des conseils spécifiques à vos besoins, veuillez consulter des professionnels qualifiés.