



Étude de rentabilité

Résumé

Cette étude vise à évaluer la rentabilité de la modernisation d'un cas d'usage métier simple par l'intégration de solutions d'intelligence artificielle. Pour ce faire, nous analyserons tout d'abord le coût du processus métier dans son état actuel, en détaillant ses étapes et ses inefficacités. Ensuite, nous comparerons trois grandes catégories de technologies de compréhension du langage naturel : les approches basées sur l'apprentissage automatique, celles s'appuyant sur des modèles de langage pré-entraînés, et enfin celles exploitant des systèmes fondés sur la connaissance (knowledge-based). Cette comparaison prendra en compte des critères tels que les coûts de mise en œuvre, les performances, l'évolutivité et l'adaptabilité de chaque solution, afin d'identifier la plus pertinente pour optimiser le cas d'usage étudié.

Présentation du cas d'usage

Prenons un cas d'usage simple qui concerne de nombreuses entreprises, celui de l'attrition des talents. Bertrand est le PDG d'un groupe d'assurance sur Niort. Dans son entreprise, beaucoup d'employés sont mécontents et démissionnent. Au vu des chiffres et de la fréquence, cette situation l'inquiète et il décide d'en discuter avec Sophie sa Directrice des Ressources Humaines. Après avoir écouté Bertrand et partagé ce constat avec lui et afin de mieux comprendre les causes du départ des employés Sophie lui propose de faire une étude sur la satisfaction au travail dans son entreprise. Bertrand trouve que l'idée de Sophie est bonne et lui demande de revenir vers lui avec des axes d'amélioration au plus tard dans 3 mois.

Sophie se souvient d'avoir entendu qu'il existait une étude spécifique portant sur la satisfaction au travail des employés. Après quelques recherches, elle retrouve

des travaux menés par Frederick Irving Herzberg¹ qui est un psychologue clinicien américain. Il a notamment travaillé sur la motivation des personnes au travail². Elle décide de construire un questionnaire afin de collecter des informations portant sur l'accomplissement, la reconnaissance, le travail lui-même, la responsabilité et la progression sociale. Les réponses données par les employés à ces questions vont lui permettre de mieux comprendre quels facteurs se révèlent être en déséquilibre et une fois qu'ils auront été identifiés, elle pourra proposer des actions afin de corriger ces écarts et s'assurer que les employés de son entreprise sont satisfaits.

Ce processus RH qui concerne l'étude de la satisfaction au travail peut être découpé en 6 activités distinctes.



La première activité consiste pour Sophie et son équipe à formaliser les questions de l'enquête. Elles se mettent d'accord sur 8 questions et rencontrent leur premier problème. L'entreprise compte 9 000 personnes et il n'est pas envisageable d'effectuer des entretiens individuels en face à face. Ils décident donc de s'appuyer sur un produit logiciel disponible sur l'intranet de l'entreprise afin de diffuser un questionnaire auprès des employés.

La seconde activité consiste à demander à l'équipe informatique de préparer la diffusion de ce questionnaire auprès des employés.

La troisième activité concerne les employés qui vont devoir répondre aux questions qui sont posées. Bien évidemment les réponses seront traitées de manière anonyme.

La quatrième activité consiste à récupérer toutes les réponses des employés. Ces réponses seront enregistrées dans une base de données de l'entreprise hébergée sur un serveur en interne. Oui, c'est ici une précaution car ces données sont sensibles et peuvent contenir des informations à caractère personnel.

La cinquième activité consiste pour Sophie et son équipe à lire les réponses des collaborateurs afin d'évaluer leur satisfaction.

Enfin, la sixième et dernière étape consiste à résumer toutes ces informations issues de l'analyse et de générer un rapport qui sera ensuite présenté à Bertrand.

1 https://fr.wikipedia.org/wiki/Frederick_Herzberg

2 https://fr.wikipedia.org/wiki/Theories_des_besoins_et_des_motivations

Quand le travail est effectué par l'homme

Tout se met ainsi en place, le questionnaire est préparé, les employés vont répondre et l'équipe de Sophie va analyser ce qui a été dit. Mais voilà, les choses ne se passent pas comme prévues.

Afin d'analyser les critères de la satisfaction au travail, Sophie et son équipe vont s'appuyer sur l'étude d'Herzberg. Ce dernier a identifié 17 facteurs de satisfaction et d'insatisfaction au travail. Chacun de ces facteurs est porteur d'une sémantique particulière qu'il faut pouvoir repérer à la lecture des réponses des collaborateurs. En plus de ces facteurs directement liés à la satisfaction au travail, il faut aussi capter tous les mots qui sont en lien avec des états émotionnels³ et ceux qui sont en lien avec des états affectifs⁴ afin de pouvoir mieux évaluer ce que les employés apprécient et ce qui doit être amélioré. Cela va prendre du temps, beaucoup de temps, même si les réponses des employés sont plutôt courtes.

Faute de temps, Sophie et son équipe ne peuvent pas lire toutes les réponses fournies par les employés et doivent procéder par échantillon afin de respecter le délai donné par Bertrand. Elles lisent ainsi quelques réponses, évaluent les facteurs de satisfaction et d'insatisfaction au travail, les émotions et les états affectifs puis préparent une présentation pour Bertrand afin de lui rendre compte de ce qui va bien et de ce qu'il faudrait améliorer.

Ces résultats donnent en effet certaines pistes mais présentent un inconvénient. Faute de temps, toutes les réponses n'ont pas été analysées et il est possible que certains facteurs liés à la satisfaction au travail aient été oubliés et n'apparaissent pas. L'examen est donc imparfait.



Bertrand est toutefois content d'avoir ces premiers résultats et décide deux choses. La première c'est qu'il souhaite que ces études auprès des collaborateurs soient menées 2 fois par an et il demande aussi à Sophie de discuter avec le service informatique afin de trouver des solutions, notamment des solutions d'intelligence artificielle⁵ capables de comprendre le langage naturel⁶ afin de pouvoir analyser les réponses de tous les collaborateurs et d'accélérer le temps consacré à leurs analyse. Sophie va présenter son cas d'usage à Julien, le Directeur du Service Informatique.

3 <https://fr.wikipedia.org/wiki/%C3%89motion>

4 https://en.wikipedia.org/wiki/Richard_Lazarus

5 https://fr.wikipedia.org/wiki/Intelligence_artificielle

6 https://fr.wikipedia.org/wiki/Compr%C3%A9hension_du_langage_naturel

Quand le travail est effectué par l'homme et la machine

Évaluer la satisfaction au travail nécessite de pourvoir lire les réponses, d'identifier certains mots spécifiques et de les comprendre. Aujourd'hui ces processus nécessitant de l'intelligence humaine sont exécutés par des personnes, car il est difficile de les moderniser avec des technologies traditionnelles.

Mais il est possible que demain, les entreprises combinent les capacités des humains et des machines afin d'exécuter ensemble les mêmes processus plus efficacement. Pour rendre cela possible, il nous faut utiliser des solutions logicielles capables de reproduire les capacités humaines de traitement symbolique de l'information. Et c'est bien l'un des objectifs de l'intelligence artificielle.

Julien et son équipe identifient 3 types de solutions qui pourraient répondre au besoin de Sophie. Le premier type de solution utilise des technologies d'intelligence artificielle d'apprentissage comme par exemple celles qui utilisent l'apprentissage automatique⁷ (Machine Learning ou ML) telles que Microsoft Azur, IBM Watson ou bien encore AWS Comprehend. Le second type de solution, également apprenante, utilise des modèles de langage⁸ comme ChatGPT, Grok, Mistral, Gemini, ou CoPilot. Enfin, il existe un troisième type de solution qui est basé sur de la connaissance⁹ (knowledge based) comme le laboratoire d'intelligence linguistique Coeos™ d'Aeteos.

Toutes ces solutions traitent ou comprennent le langage naturel et offrent un large éventail de fonctions permettant de répondre aux besoins de Sophie. Mais quel est leur coût d'utilisation. Julien décide de les estimer. Il va tout d'abord estimer le coût de ce travail fait par l'équipe des ressources humaines.

Estimation du coût du travail

Sophie travaille dans le domaine de l'assurance. La taille moyenne d'une grande entreprise dans ce secteur est de 11 000 personnes mais la société de Sophie a uniquement 9 000 collaborateurs. Elle va donc envoyer son questionnaire à toutes ces personnes, mais uniquement 70 % des collaborateurs vont répondre à l'enquête. Elle va devoir, avec son équipe, traiter 6 300 réponses au total. Dans chaque questionnaire, chaque collaborateur a répondu à 8 questions. Pour chacune des réponses données, Sophie et son équipe vont devoir les lire et reconnaître toutes les sémantiques possibles liées à 2 états affectifs (négatifs et positifs), 8 émotions primaires et 17 facteurs de la satisfaction et de

7 https://fr.wikipedia.org/wiki/Apprentissage_automatique

8 https://fr.wikipedia.org/wiki/Mod%C3%A8le_de_langage

9 https://en.wikipedia.org/wiki/Knowledge-based_systems

l'insatisfaction au travail, soit un total de 27 champs sémantiques différents puis, sur détection de l'un de ces facteurs, rendre compte. Comme la détection de l'un de ces facteurs dans une réponse prend 10 secondes, l'équipe de Sophie aura une charge de travail équivalente à 2,56 ETP (Équivalent Temps Plein) par étude, soit 5,12 ETP au total, Bertrand ayant demandé à ce que cette enquête soit faite 2 fois par an.

En tant que directrice des Ressources Humaines, Sophie a beaucoup d'autres choses à faire, c'est pourquoi elle participe à ce projet seulement 1 jour par mois. Elle a confié ce travail d'analyse à des experts en ressources humaines juniors qui sont encadrés par un expert senior à mi-temps. La société étant sur Niort et non dans une grande ville comme Lyon, Marseille ou Paris, les locaux et la rémunération de ses équipes sont plus raisonnables. Le salaire chargé annuel tout compris (primes, mutuelle, repas, PC, logiciels, Smartphone, transport) d'un profil RH junior sur Niort a été estimé à 65 425€, celle du senior à 87 753€ et celui de Sophie à 125 307€. Au total, le coût de cette activité d'analyse a été estimé à 385 880€.

Estimation de la modernisation

Julien va ensuite comparer le coût du travail de l'équipe de Sophie à ce que lui coûterait l'usage de l'une de ces solutions logicielles permettant de comprendre le langage naturel.

Tout d'abord, il y a le coût de la licence.

Que ce soit l'usage de solutions de traitement du langage naturel comme celle d'IBM, d'AWS, de Google ou de Microsoft, ou bien que ce soit l'usage d'un modèle de langage comme par exemple ChatGPT, le coût mensuel par personne est environ de 30€. Dans certains cas et en fonction des éditeurs, certains composants logiciels ou services additionnels doivent être ajoutés (services de stockage, atelier logiciel de création de modèle, base de données, ...) et chaque utilisateur impliqué sur le projet doit avoir un accès. Nous estimons pour cette étude que ce coût représente 800€ par mois tout compris. Quant à Coeos™, son coût mensuel est de 1 460€, tout compris (nombre d'utilisateurs illimités, usage illimité).

Ensuite, il y a le coût à l'usage

Que ce soit des solutions de traitement du langage naturel ou bien des modèles de langage, ils fonctionnent tous sur le même principe, à l'usage. Plus il y a de caractères à traiter et plus on va devoir payer. Plus il y a d'utilisateurs qui utilisent

la solution et plus le coût de licence va augmenter. On paye pour les données que l'on va envoyer et on paye aussi la quantité des données que l'on va recevoir.

Dans ce cas précis, que traitons-nous exactement ? Nous traitons des réponses courtes à 8 questions. La taille d'une réponse est de quelques mots, un peu comme la taille d'un SMS ou bien d'un Tweet. Disons que la taille de la réponse est de 213 caractères. Donc on doit traiter 6 300 questionnaires, contenant 8 réponses, soit 50 400 fichiers d'environ 213 caractères (un total de 10,24 Mo de données) par étude, soit 100 800 fichiers par an (20,48 Mo de données).

Ces solutions logicielles ont souvent un tarif par paquet de caractères, soit 100, 1000 ou 10000 et chaque paquet coûte quelques centièmes ou millièmes d'Euros. Autant dire que vue la volumétrie à traiter ce coût sera négligeable, probablement quelques dizaines d'Euros. Exemple, si on utilisait ChatGPT, modèle GPT4(8K), le prix d'un paquet (token) serait de 0,03 euros et nous devons consommer un total de 1 154 paquets. Cela représente un coût de 34,62€. Un autre exemple serait l'usage de AWS Comprehend. Dans ce cas on devrait analyser 107 352 paquets et chaque paquet va coûter 0,0001€. Nous aurions dans ce cas un coût total de 10,73€. Là aussi, ces estimés concernent une seule étude et pour avoir le coût annuel, nous devons multiplier ces prix par deux. Comme ces solutions seront à la fois utilisées pendant la phase projet mais aussi pendant la phase d'exploitation ou de maintenance, il convient aussi d'ajouter à ces estimations le coût de cette consommation liée au projet ou bien à celle de maintenance du modèle. Tout compris, nous avons estimé que ce coût serait de 200€ par mois.

Quant à Coeos™, son coût n'est pas à l'usage. Il peut donc y avoir un nombre illimité d'utilisateurs qui traitent une quantité illimitée de données. Cela ne change rien au coût final.

Nous sommes également dans un cas d'usage métier bien spécifique où la solution logicielle doit être capable d'identifier correctement 2 états affectifs, 8 émotions primaires et 17 facteurs de satisfaction et d'insatisfaction au travail, soit un total de 3 cadres de référence contenant 27 entités parentes. La bonne nouvelle est que ces technologies sont apprenantes et il suffit donc de leur apprendre ces contextes sémantiques.

La première difficulté que l'équipe informatique de Julien va rencontrer est qu'ils ne connaissent pas tous les mots associés à ces cadres de référence et d'ailleurs Sophie ne le sait pas non plus. Cette liste de mots, correctement associée à chacun des concepts de ces 3 cadres de référence doit être fournie. Alors faisons ce premier travail.

En ce qui concerne les facteurs de satisfaction et d'insatisfaction au travail, un total de 1 045 mots ont été trouvés. En ce qui concerne les émotions, un total de 5 163 mots ont été trouvés et en ce qui concerne les états affectifs positifs et négatifs, un total de 4 190 mots ont été identifiés.

Comme ces technologies sont apprenantes, il va falloir fournir ces mots à la machine. Mais cela n'est pas aussi simple. Tout d'abord, lors de la collecte de ces mots, nous devons nous assurer que ces mots sont de bonne qualité, qu'ils sont correctement qualifiés. Ils ne doivent pas comporter d'erreur d'orthographe, ne doivent pas être discriminants ou malfaisants, doivent être conformes dans le contexte du cas d'usage à traiter. Il faut donc les sélectionner avec soin et cela va prendre du temps.

Une fois que cette première phase du projet est faite, il faut créer plusieurs fichiers pour chaque mot afin de le placer dans une phrase ayant une forme grammaticalement positive, puis négative (il existe en français pas moins de 7 formes négatives), puis interrogative, utiliser sa forme canonique (sans fautes), puis utiliser des variations de ce même mot comme par exemple sa forme au masculin ou au féminin, sa forme plurielle, ses formes composées, ses variantes sémantiques régionales, ses expressions argotiques ou familières et enfin le représenter avec des fautes d'orthographe courantes ou bien des fautes de frappe. Pour finir, si ce mot est un verbe, il conviendra de le conjuguer dans des formes courantes et de considérer certaines formes grammaticale spécifique comme par exemple son participe passé. L'ensemble de ces mots ainsi trouvés doivent être placés dans des phrases et des contextes variés, chaque phrase doit être représentée par un fichier (ou une entrée spécifique) et ce fichier doit être correctement étiqueté, labellisé. Cette approche permet d'obtenir un jeu de données de qualité à faire apprendre à la machine. Selon les estimations fournies et la qualité de l'analyse souhaitée, il peut y avoir entre 20 et 50 voir 100 fichiers (ou entrées) par mot dans le cas où ils seraient porteurs d'un double sens (polysémiques). Enfin, il faudra aussi créer des jeux de données spécifiques afin d'éviter pour certains mots des associations malheureuses qui feraient que la machine ne serait pas capable de correctement identifier le mot. Notons également qu'il est possible d'utiliser des outils informatiques ou bien des techniques afin de générer automatiquement certaines variantes, mais pas toutes, simplifiant à la fois le travail à effectuer et le temps passé à constituer ce référentiel.

Ce jeu de données doit être préparé, injecté dans la machine, le résultat doit être analysé, les fichiers doivent être corrigés et ainsi de suite. Chaque itération (cycle d'apprentissage) va améliorer la qualité de la compréhension jusqu'à l'obtention d'un résultat acceptable. Classiquement, il faut compter entre 10 à 20 itérations ou « époques ». Ces variations quant au nombre de cycles sont liées au pré-

entraînement du modèle. Moins il sera entraîné et plus le nombre de cycles sera important.

Dès que cela est terminé, il faut passer à la phase de test et tester plusieurs cas, si nécessaire ajuster et entraîner de nouveau avant de pouvoir déployer cette solution et l'utiliser pour aider Sophie et son équipe.

En un mot, c'est un véritable projet !

Comme Julien souhaite aller vite, au lieu de générer pour chaque mot 20, 50 ou même 100 fichiers, il ne va retenir que les variations les plus courantes et va générer pour chaque mot uniquement 9 fichiers (ou entrées spécifiques). Il a donc un total de 93 582 fichiers à générer et chaque génération de fichier, avec la labellisation et tous les contrôles devant être faits, est estimée entre 1 à 3 minutes par fichier en fonction de la difficulté de génération de la variation souhaitée, soit un total de 285 jours ou bien 1,35 ETP.

Dès que cette première phase du projet sera terminée, le modèle pourra être entraîné. Julien opte pour 12 cycles d'entraînement supervisé représentant un total de 238 jours ou bien 1,13 ETP.

Pour réduire les coûts de ses projets, Julien a pour habitude de s'appuyer sur un centre de Delivery near-shore se trouvant en Tunisie. Cette destination est idéale pour de nombreuses raisons. Tout d'abord, en Tunisie les professionnels ont un bon niveau, ils parlent français et le coût de la main d'œuvre est moins cher qu'en France. De plus, la Tunisie n'a qu'une seule heure de décalage horaire avec la France ce qui permet à ses équipes de travailler facilement avec ces prestataires tunisiens.

Le Data Scientist junior va lui coûter 320€, le senior qui va les encadrer quant à lui va coûter 540€ et le chef de projet du prestataire va lui coûter 640€. Comme ce projet est simple, ce chef de projet sera présent 1 jour par semaine. A Niort, un consultant va assurer le relais avec cette équipe et il sera accompagné d'un Directeur de projet qui va lui aussi être présent 1 jour par semaine.

Au total, 653 jours de prestations seront effectués near-shore, 129 seront effectués sur Niort et 31 jours de prestation seront confiés à des sachants en psychologie afin de constituer la liste des mots à prendre en compte pour ces 3 cadres de référence. Au total, le coût de ce projet sera de 378 k€ (à un TJM de 464€). Mais Julien sait aussi que si la solution cible utilisée est un modèle de langage, celui-ci est beaucoup plus opérationnel qu'une solution de machine learning classique car il est pré-entraîné, il va donc avoir moins de travail. Il estime qu'il aura simplement 40 % de ces coûts et au total ce projet lui coûtera 151 k€. Aujourd'hui, il existe aussi de nombreux modèles de langage plus simples,

moins pré-entraînés, qui peuvent être installés sur un simple serveur et qui offrent a priori une forme de sécurité un peu plus importante. L'usage de ces modèles augmenteraient de fait la charge de travail du côté projet et n'ont pas été considérés dans cette étude portant sur la rentabilité des solutions d'intelligence artificielle.

Dès que ce projet sera terminé, cette solution pourra être mise en œuvre et aider l'équipe de Sophie. Cependant, des nouveaux termes peuvent apparaître et il se peut également que la solution ait été mal entraînée sur certains termes, notamment ceux pouvant porter un double sens, sur certaines fautes non identifiées. Afin de sécuriser cela, Julien va devoir mobiliser des Data Scientists dans un centre de Delivery en Tunisie. Les 6 premiers mois, deux ingénieurs seront affectés à ce projet à temps plein puis l'année suivante un seul ingénieur et la dernière année un ingénieur à mi-temps. Ainsi, les éventuels ajustements pourront être pris en charge dans des délais raisonnables. Cette phase de maintenance ajoute un coût de 169 k€ au projet.

Pour information et afin d'être complet, des entreprises comme OpenAI, Google Meta, DeepSeek, Mistral, xAI, Baidu ou Anthropic s'appuient sur des centres de services, souvent situés dans des régions comme l'Afrique et l'Asie du Sud-Est, pour des tâches liées à l'annotation, la correction, et l'amélioration des modèles d'IA. Ces centres jouent un rôle clé dans l'optimisation des performances des modèles. Ces centres de services, parfois appelés "centres d'annotation de données", emploient des personnes pour effectuer des tâches essentielles au développement et à la maintenance des modèles. Ces tâches incluent l'étiquetage, la classification ou la structuration des données pour entraîner des modèles d'IA. Ces personnes examinent les sorties des modèles pour repérer les erreurs, biais ou contenus problématiques (comme des réponses inappropriées ou inexactes) et corrigent. Cela s'appelle « l'apprentissage par renforcement à partir de retours humains » (RLHF, Reinforcement Learning from Human Feedback). Ces entreprises se trouvent souvent dans des pays où la main-d'œuvre est moins coûteuse (certains employés étant payés 2 dollars de l'heure), comme le Kenya, l'Ouganda, les Philippines ou l'Inde. Outre l'exploitation de cette main d'œuvre, le traitement de ces données par ces prestataires pose aussi le problème de la souveraineté des données. En résumé, ce qui est appelé l'apprentissage par renforcement à partir de retours humains est le seul chemin possible pour créer un ensemble de données soigneusement annotées et corrigées permettant la création de modèles de qualité. Ce que ces entreprises font pour corriger leurs modèles doit aussi être fait par l'équipe de Julien et cette phase de maintenance est non seulement essentielle mais doit être maintenue dans le temps.

Qu'en est-il du côté de Coeos™ ? Et bien comme Coeos™ a une approche dite « basée sur de la connaissance », ce travail a déjà été effectué par l'éditeur et la

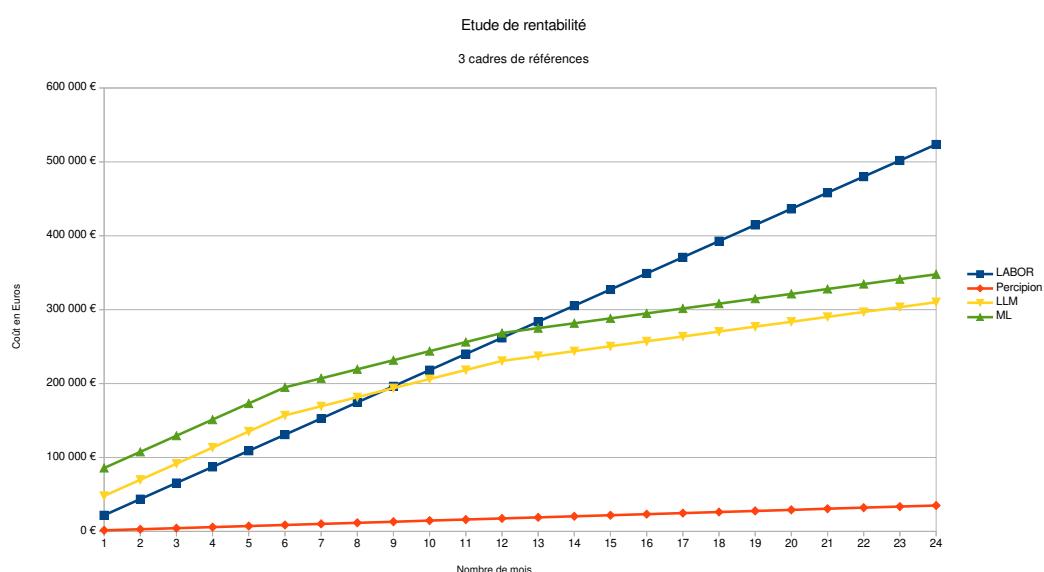
solution est pour ainsi dire prête à l'emploi. Le cadre de référence de la satisfaction au travail est compris dans l'abonnement. Aucun coût supplémentaire. En ce qui concerne les éventuelles mises à jour du cadre de référence, cette activité est prise en charge par l'éditeur sans surcoût.

Pour l'estimation de coûts liés à ce projet, nous n'avons pas chiffré les coûts du transport, de l'hébergement, la charge de travail liée à la gouvernance ainsi que la révision de ces estimés en fonction du niveau de risque pris (composante near-shore). En fonction de la solution qui sera utilisée, le résultat est fourni sous la forme de données qui doivent encore être exportées vers d'autres produits logiciels afin de générer un rapport pour l'utilisateur final. Cette dernière étape n'a pas été chiffrée.

Maintenant que tous les coûts sont connus, il est temps de les comparer et de voir en combien de temps il y aura un retour sur investissement (ROI).

Le retour sur l'investissement

Sur une durée de 2 ans, si cette activité est prise en charge par l'équipe de Sophie, le coût du travail est de 523 k€. Si cette activité est confiée à une intelligence artificielle apprenante, son coût sur 2 ans serait de 348 k€ pour une solution de machine learning et de 310 k€ pour une solution utilisant des modèles de langage et si cette solution utilisait une approche basée sur de la connaissance, comme Coeos™, ce coût serait de 35 k€.



La première conclusion à laquelle nous arrivons est que dans tous les cas, la modernisation de cette activité par l'usage d'une technologie d'intelligence artificielle permettra à cet assureur de réduire ses coûts de fonctionnement.

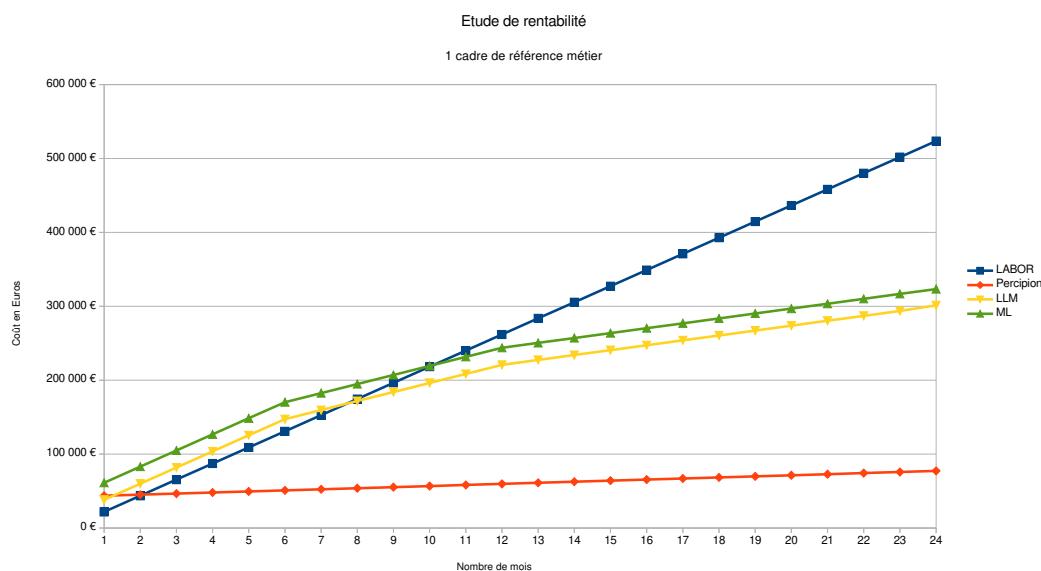
La seconde conclusion est que les technologies basées sur de la connaissance ont un coût d'exploitation moins élevé.

Toutes ces solutions présentent bel et bien des gains. Mais en combien de temps sont-elles rentables ?

En allant évaluer le retour sur investissement de ces solutions (ROI), les technologies de machine learning présentent des gains à partir du 12ème mois, les technologies utilisant des modèles de langage présentent des gains à partir du 9ème mois et celles étant basée sur de la connaissance offrent des gains immédiats.

Ce projet est complexe car il nécessite la création de 3 cadres de référence spécifiques, Julien s'est donc également posé la question de savoir ce qu'il en serait si il n'y avait qu'un seul cadre de référence spécifique à créer et que dans ce cas précis, ce nouveau cadre de référence n'était pas disponible du côté des technologies basées sur de la connaissance. Là aussi, il devrait y avoir une phase projet et ces technologies pourraient alors être mieux comparées.

La charge de travail la plus importante sur ce type de projet reste la constitution d'un référentiel de données correctement qualifié et présentant chaque mot de multiples manières dans des contextes variés ainsi que l'entraînement de la solution.



Du côté des technologies basées sur de la connaissance, notamment Coeos™, certains de ces travaux ne sont pas nécessaires car Coeos™ possède non seulement une structure de donnée innovante représentant les données dans la mémoire de la machine comme elles seraient représentées dans la mémoire à long-terme humaine mais aussi une unité de traitement permettant d'évaluer si deux mots peuvent être similaires ou non ce qui va réduire de manière importante l'identification correcte du nombre de variation que ce mot pourrait avoir par rapport à sa forme canonique. Pour cette évaluation, si l'on devait créer un cadre de référence dont la complexité est égale à celle du cadre de référence contenant les facteurs de la satisfaction et de l'insatisfaction au travail, la constitution de ce cadre de référence contenant 1 045 objets serait fait en un mois pour un coût projet de 20 391€. Ce cadre de référence étant alors spécifique à ce projet, le coût de son usage par Coeos™ serait alors nul.

Pour ce second cas et sur une durée de 2 ans, si cette activité est prise en charge par l'équipe de Sophie le coût du travail serait de 524 k€. Si cette activité est confiée à une intelligence artificielle apprenante, son coût sur 2 ans serait de 323 k€ pour une solution de machine learning et de 301 k€ pour une solution utilisant des modèles de langage et si cette solution utilisait une approche basée sur de la connaissance, comme Coeos™, ce coût serait de 77 k€.

En allant évaluer le retour sur investissement de ces solutions (ROI) dans ce second cas, les technologies de machine learning présentent un gains en 10 mois, les technologies utilisant des modèles de langage quant à eux présentent des gains à partir du 8ème mois et celles étant basée sur de la connaissance offrent des gains à partir du 2ème mois.

Il est très important de noter que la forte variabilité des cycles d'apprentissage en fonction des technologies apprenantes utilisées (machine learning versus modèle de langage) ainsi que la charge de travail liée à ces activités n'a pas été prise en compte dans cet article. Selon le Gartner¹⁰, « La production de valeur est lente (plus de deux ans) ». L'usage de telles solutions seraient rentables à partir de 24 mois. Et le Gartner ajoute que « d'ici 2028, plus de 50 % des entreprises qui ont élaboré des modèles d'IA de grande envergure en partant de zéro renonceront à leurs démarches en raison des coûts, de la complexité et du passif technique liés à leurs déploiements ». Rentable oui, mais pas tout de suite. Maintenable dans le temps, pas si sûr. Permettent de générer de la valeur ? A en croire une récente étude du MIT¹¹ (Juillet 2025), il semble bien que non : « malgré 30 à 40 milliards de dollars d'investissements des entreprises dans les IA génératives (GenAI), ce rapport révèle un résultat surprenant : 95 % des organisations n'obtiennent aucun retour sur investissement ».

10 <https://www.gartner.fr/fr/articles/adoptez-cette-approche-pour-evaluer-le-rendement-du-capital-investi-de-l-ia-generative>

11 https://mlq.ai/media/quarterly_decks/v0.1_State_of_AI_in_Business_2025_Report.pdf

Conclusion

En conclusion, la modernisation de cette activité par l'usage de technologies d'intelligence artificielle offre bien des gains à l'entreprise dans tous les cas. Les technologies comme Coeos™ qui sont basées sur de la connaissance sont plus simples à mettre en œuvre et permettent également de générer des gains de productivité ou des gains financiers plus rapidement voire immédiatement que des technologies qui seraient basées sur de l'apprentissage par la machine.

En plus de ces critères financiers, le choix d'une solution logicielle permettant de moderniser des activités comme celle de l'équipe de Sophie est également lié à d'autres contraintes.

Par exemple, ces solutions doivent respecter le cadre réglementaire de l'AI Act Européen et pouvoir être mise en production légalement. Un grand soin doit être pris quant aux données qui seront injectées dans les solutions apprenantes afin que ces solutions ne soient discriminantes ou traitent des connaissances qui seraient erronées. Ces technologies étant « ouvertes », une mauvaise injection de données peut rapidement faire varier ou rendre faux les résultats du traitement. Comment peut-on, avec ce type de technologie, garantir que les résultats d'une analyse effectuée la veille seront identiques à ceux obtenus aujourd'hui ? Le cadre de référence étant en constante évolution, les relations statistiques du modèle évoluant, il est impossible de garantir cette reproductibilité. Du côté de Coeos™, cela n'est pas possible car les cadres de référence sont hautement qualifiés et protégés de toute injection de données.

Ces solutions apprenantes sont des solutions logicielles produites par un éditeur qui pour construire son produit va créer son propre code source et s'appuyer soit sur d'autres composants ou librairies qu'il a créées ou bien sur des composants ou des librairies tierces qui ne lui appartiennent pas. Comme ces solutions sont des solutions d'intelligence artificielle, ces éditeurs doivent aussi avoir une maîtrise sur l'origine de ces composants tiers et être capable de tous les identifier. Lors de l'usage de langage de haut niveau, notamment de langage de script, l'usage d'une quantité importante de librairie est nécessaire et cette exigence devient vite très complexe. Du côté de Coeos™, cela n'est pas le cas car Coeos™ a été entièrement programmé en langage C et n'utilise que des librairies standard normées (C99 ANSI).

La grande majorité des solutions d'intelligence artificielle apprenantes sont des sortes de boîtes noires et dans la majorité des cas, l'éditeur est dans l'incapacité de produire un traitement qui soit à la fois transparent et qui permette d'expliquer les relations sémantiques entre les concepts du cadre de référence afin de fournir à l'utilisateur des résultats qu'il peut vérifier. Comment faire confiance à ce type

de solutions ? Coeos™ permet lors de l'exécution d'afficher toutes les étapes du traitement et offre à l'utilisateur un haut niveau de transparence.

Dans l'exemple donné dans cet article, nous nous sommes principalement appuyés sur les travaux qu'il était nécessaire de mener afin d'injecter dans une machine un nouveau cadre de référence en prenant en compte uniquement des variations simples de la forme canonique d'un mot. Mais cette approche, très conservatrice, ne permet pas de rendre compte de la richesse de l'expression en langage naturel ni de la substitution d'un mot par une représentation iconique (Émojis). L'usage d'une telle représentation imagée doit également être interprétée et correctement associée à son concept parent. De plus, certaines phrases utilisent des mots introduisant des valences. Ces dernières viennent modifier le sens du mot et doivent aussi être prises en compte. Cette prise en compte va venir augmenter de manière significative le nombre de données devant être injectée dans la machine et va donc venir augmenter de manière importante le nombre de variations possibles d'un même mot. Pour préciser ce point, il existe 38 mots dans la langue française qui sont porteurs d'une valence sémantique dite positive et 49 mots qui sont eux porteurs d'une valence sémantique dite négative. En plus de ces valences sémantiques, il existe aussi des valences psychologiques liées à nos états affectifs positifs et négatifs (pour un total de 836 mots) et celles liées à nos émotions primaires (pour un total de 729 mots). Ainsi, si lors de l'apprentissage, ces variations ne sont pas dans le corpus des données que la machine doit apprendre, la prise en compte de ces valences sera mise à mal et la machine ne pourra pas être capable de traiter ces cas et produira des analyses fausses. Coeos™ permet non seulement de transcrire les Émojis présents dans un verbatim dans le mot correspondant afin qu'il soit pris en compte par la solution de compréhension du langage mais il est aussi capable de traiter les valences sémantiques et psychologiques ce qui lui confère une compréhension beaucoup plus fine du langage naturel.

Posez-vous aussi la question sur l'hébergement de votre solution de compréhension du langage naturel, de l'usage du réseau et des éventuelles données qui sortiraient de vos serveurs vers des serveurs tiers. Vos données sont elles réellement sécurisées ? Sont-elles partagées avec d'autres solutions ? Sont-elles captées ? Comment faire confiance ? Dans notre exemple, Sophie doit analyser des données sensibles dont le traitement doit rester anonyme et sécurisé. Est-ce que ces solutions peuvent le garantir ? Du côté de Coeos™, la réponse est simple. Coeos™ peut être installé où vous voulez et cette solution logicielle n'utilise pas le réseau. L'activité réseau étant nulle, les traitements effectués par Coeos™ peuvent être effectués hors-ligne et en temps réel.

Avant de finir, abordons également l'impossibilité de nombreuses solutions de traitement du langage naturel de traiter des verbatims volumineux. Basés sur des API, ces solutions n'acceptent que des textes dont la taille est limitée. Mais dans

certains cas, comment faire lorsqu'il est nécessaire d'analyser l'ensemble d'un document ? Comment effectuer une analyse sémantique d'un texte dans son ensemble si à cause de contraintes techniques, il est nécessaire de découper ce texte en un ensemble de fichiers ? Dans ces cas, l'analyse devient plus complexe et nécessite en fin de traitement, de consolider les différents résultats de ces différents traitements. Cela ajoute une difficulté complémentaire. Coeos™ est capable de traiter des documents de toute taille en une seule passe.

Enfin, à l'instar de solutions logicielles d'intelligence artificielle apprenantes qui consomment une quantité d'énergie importante, Coeos™ ne consomme que 11,34Wh, soit par analogie 1 centilitre d'eau par cœur CPU. Là encore, Coeos™ est en avance sur sa concurrence et respecte bien les objectifs de l'Union Européenne.

Julien a identifié des technologies permettant d'aider Sophie dans son travail. L'activité de lecture et d'identification de sémantiques particulières liées à des états affectifs, des émotions ou bien celles liées à des facteurs de la satisfaction ou de l'insatisfaction au travail peuvent bel et bien être identifiées par une solution logicielle et l'usage de ces solutions offrent toutes des gains. L'approche basée sur de la connaissance, telle que celle proposée par Coeos™, est la moins coûteuse et est celle qui offre un retour sur investissement le plus rapide. De plus, de nombreux avantages fonctionnels, réglementaires et de sécurité font de Coeos™ un choix pertinent pour moderniser de nombreux cas d'usage dans l'entreprise.

D'abord l'Homme puis la machine

Après 8 années de recherche et de développement en neurosciences, psychologie cognitive et clinique, linguistique, philosophie et sciences informatiques, nous avons créé Coeos™, un laboratoire d'intelligence linguistique français, sûr, souverain et éco-responsable, capable d'aider l'homme dans son travail, permettant de transformer des données textuelles en visions stratégiques pour une prise de décision éclairée et rapide.

C'est un peu comme si dans ce laboratoire travaillaient des assistants capables de lire du texte, de le comprendre et d'en saisir les nuances, d'identifier des mots et de les associer avec des concepts, d'extraire le sens du texte par une analyse contextuelle précise, d'anticiper, de raisonner et d'en tirer des conclusions, de produire des rapports, de communiquer et d'agir.

Avec cette nouvelle technologie, Aeteos™ aide les entreprises à améliorer leur efficacité décisionnelle et opérationnelle en leur permettant de construire leur

propre laboratoire d'intelligence linguistique et de moderniser de nombreux cas d'usages.

Coeos™ permet une analyse quotidienne des données, qu'elles soient internes ou externes à votre organisation. Il peut aussi être utilisé pour monitorer l'espace numérique et contribuer à collecter des informations essentielles, identifier proactivement des risques et émettre des alertes. En fournissant une analyse contextuelle précise, Coeos™ aide à prendre des décisions éclairées, tout en garantissant aux utilisateurs un contrôle total.

Quel que soit le domaine d'activité, Coeos™ permettra d'accélérer l'analyse de l'information et facilitera la prise de décision dans plus de 47 cas d'usages.

Avec son propre jeu de données, il complétera le diagnostic des équipes. Tout reste sous leur contrôle.

Il apportera une expertise là où aucun expert n'est disponible en fournissant des informations sur des données sensibles, libérant ainsi vos équipes qui pourront se concentrer sur des activités à plus forte valeur ajoutée.

Si vous souhaitez en savoir plus, n'hésitez pas à revenir vers nous !