

Traductions automatisées à grande échelle : remplacer un SaaS propriétaire par un LLM open-source

Mateo Fernandez Martinez

90 000€ par an

~~90 000€ par an~~

-87%



☰ Toutes les catégories

👕 Mode

🚗 Véhicules

🏠 Ménage & Maison

🏈 Sport

🔧 Bricolage & Jardin

💻 Ordinateurs & Réseaux

Back to Uni : t'es prêt ?

Que ce soit un PC, sac à dos ou une gourde, trouve tout ce dont tu as besoin pour débiter ton semestre bien préparé.

[C'EST PARTI](#)

Actuellement sur Ricardo

Dans ta région

Dernière minute

Populaires

Nouveau

Dès CHF 1.-

Page d'accueil de la plateforme Ricardo (www.ricardo.ch)

Ricardo + Angebot erstellen Favoriten Anmelden DE

Suche nach Artikel, Verkäufer oder Artikelnummer Q

Alle Kategorien
Kleidung & Accessoires
Fahrzeuge
Haushalt & Wohnen
Sport
Handwerk & Garten
Computer

Ricardo / Handy, Festnetz, Funk / Handys / Smartphones / iPhone XR 256GB defekt Bildschirm



Navigation arrows: < >

Thumbnail images:     

iPhone XR 256GB defekt Bildschirm

🕒 27. Aug. 2025, 22:05 Uhr

Nächstes Gebot
32.00

🛡️ MoneyGuard verfügbar

BIETEN

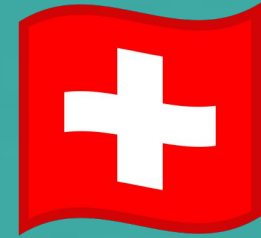
♥️ ZU FAVORITEN HINZUFÜGEN

🛡️ MoneyGuard: Optionale Absicherung ^

Profitiere von einer Geld-Zurück-Garantie bei Käufen bis zu CHF 5'000.- im Fall von Nichtlieferung oder wenn das Produkt von der Beschreibung abweicht.




[Mehr erfahren](#)



4 langues
nationales

Product Detail Page (ou page produit en français)

Ricardo

+ Mettre en vente

Favoris

Se connecter

FR

Recherche des produits, des vendeurs ou un numéro de produit

Économisé

50 kg de CO₂

Capacité de stockage

256 GB

Couleur

Bleu

Marque

Apple

Modèle

iPhone XR

État

D'occasion

TRADUIRE EN FRANÇAIS

iPhone XR 256GB defekt mit gesprungenem Bidschirm. Akku 80%. Funktion allg. defekt ... muss repariert werden.

Genaue Konfiguration im letzten Bild ersichtlich.

iPhone XR 256GB defekt

Bildschirm

27 août 2025 à 22:06

Prochaine offre

32.00

MoneyGuard

disponible

ENCHÉRIR

AJOUTER AUX FAVORIS

MoneyGuard: l'option de protection

Bénéficie de la garantie de remboursement pour tout achat inférieur à CHF 5000.- en cas de non-livraison ou si le produit ne correspond pas à la description.

TWINT

VISA

En savoir plus

Livraison

Retrait par l'acheteur à 8047 Zürich. Gratuit

Product Detail Page (ou page produit en français)

6

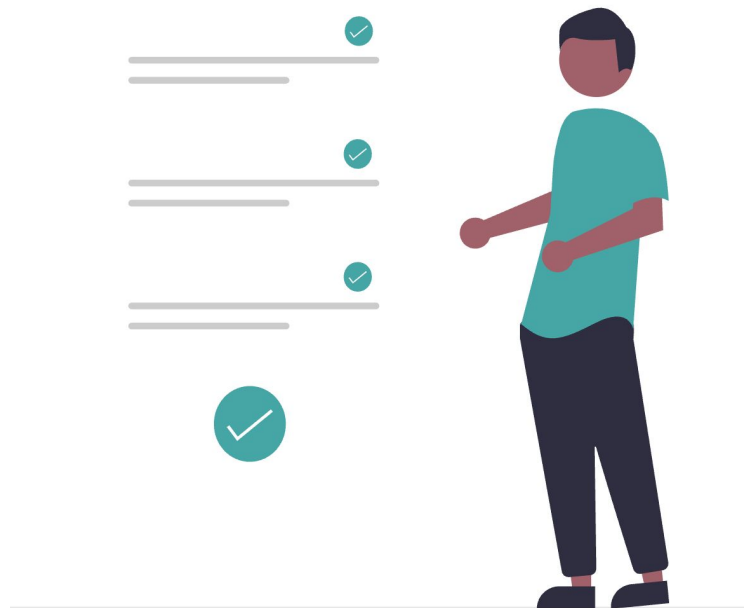
DeepL

7 500€/mois

smg

Sommaire

1. Contexte et enjeux
2. Remplacer DeepL
3. Construire l'alternative
4. Validation
5. Apprentissages



1. Contexte et enjeux :

Comprendre la mission

Swiss Marketplace Group (SMG)



Fondée en 2021

Swiss Marketplace Group (SMG)

Real Estate

 homegate.ch

 Achefer-Louer.ch

 Immo Scout24

 IAZI CIFI

 ALLE IMMOBILIEN .CH

 ImmoStreet.ch

 CASASOFT

 home.ch

 PUBLIMMO

 Flatfox

Automotive

 Auto Scout24

 Moto Scout24

General Marketplaces

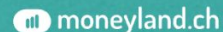
 tutti.ch

 anibis.ch

 Ricardo

Finance & Insurance

 Finance Scout24

 moneyland.ch

Ricardo

1999

Date de création

5M

Utilisateurs

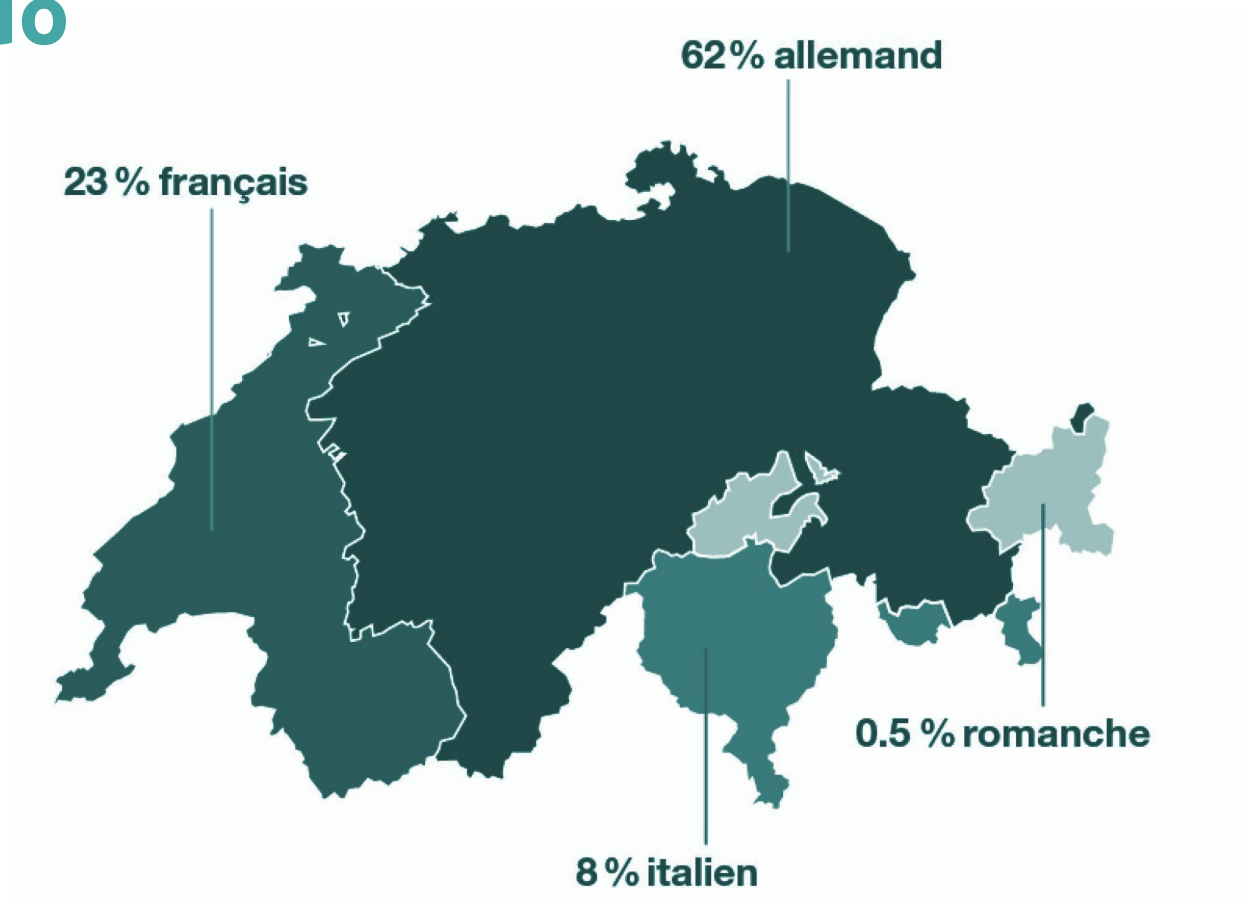
20k

Ventes
journalières



Économie
circulaire

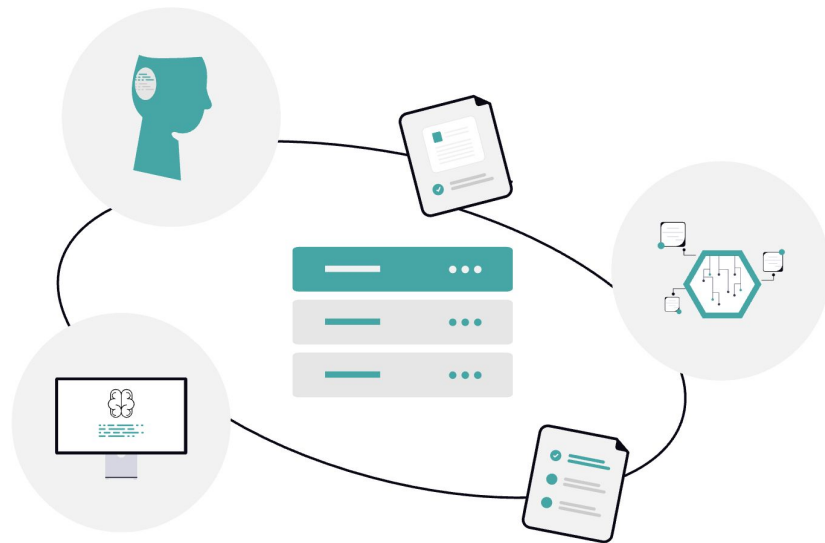
Ricardo



DeepL

3 problèmes majeurs :

- 7 500€/mois soit 90 000€/an
- Dépendance à un service externe
- Partage de nos données utilisateurs



Objectifs

- Réduction des coûts de traduction d'au moins **70%**
- Temps de réponse inférieurs à **3s** dans **95%** des cas
- Qualité de traduction similaire à DeepL (**BLEU > 30**)
- Traiter jusqu'à **20 requêtes/s**
- UX identique : **même taux de conversion**

2. Remplacer DeepL : Trouver une alternative

Services de traduction



Google Translate



Amazon Translate



Microsoft Translator

Solutions de traduction open-source



LibreTranslate



Apertium



OpenNMT

Sélection de modèles d'IA open-source

- Modèles NLLB de Meta
- Modèle MBart de Meta
- Modèles T5 de Google
- Modèles OPUS-MT de l'université d'Helsinki



Hugging Face

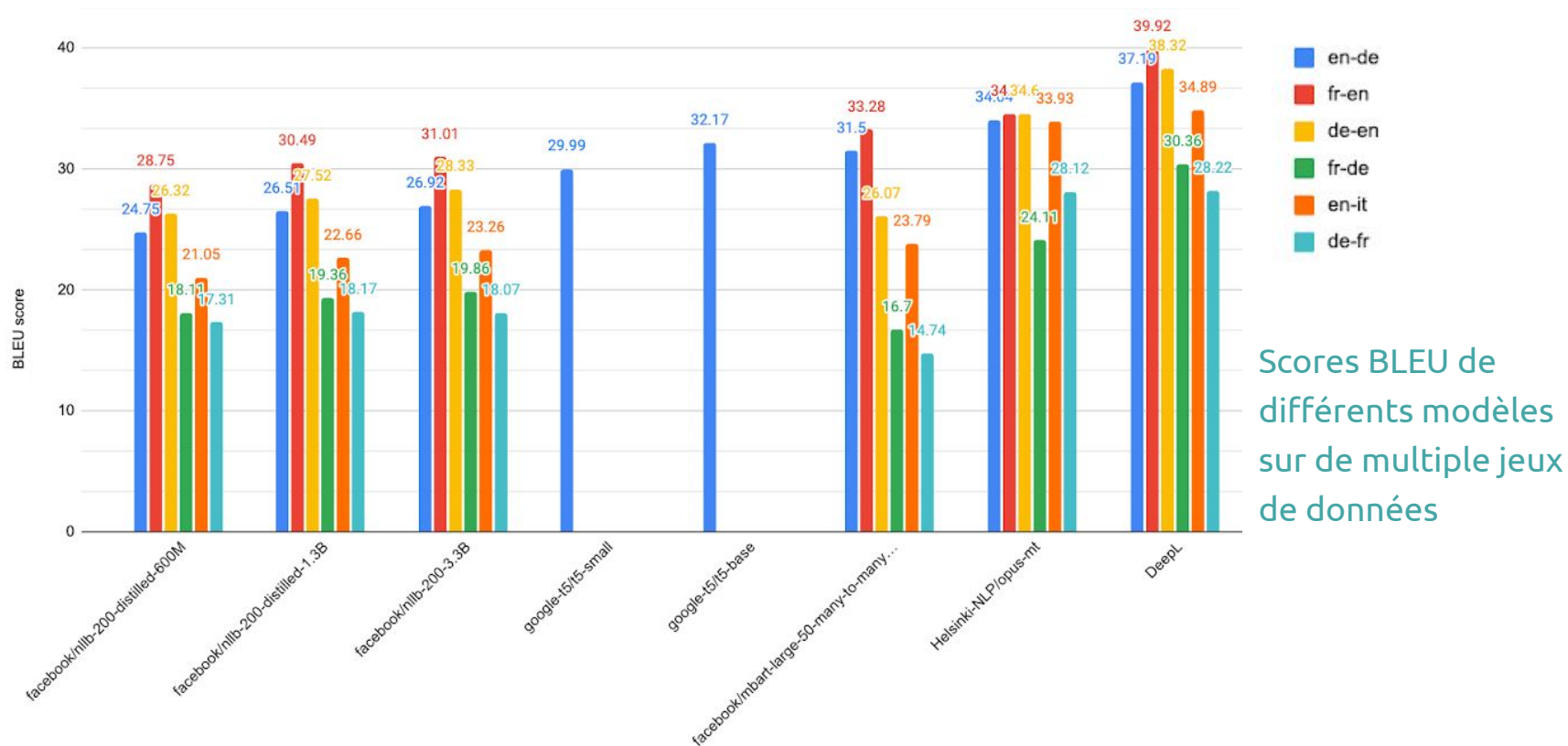
Sélection de modèles d'IA open-source

- Score BLEU calculé sur des corpus multilingues

Score BLEU	Interprétation
< 10	Traductions presque inutiles
10 à 19	L'idée générale est difficilement compréhensible
20 à 29	L'idée générale apparaît clairement, mais le texte comporte de nombreuses erreurs grammaticales
30 à 40	Résultats compréhensibles à traductions correctes
40 à 50	Traductions de haute qualité
50 à 60	Traductions de très haute qualité, adéquates et fluides
> 60	Qualité souvent meilleure que celle d'une traduction humaine

- Tests manuels sur des contenus provenant de Ricardo

Sélection de modèles d'IA open-source



Sélection de modèles d'IA open-source

Texte d'origine :

“Je vends mon iPhone 12 (64 Go) en très bon état 👍.”

Traduction OPUS-MT :

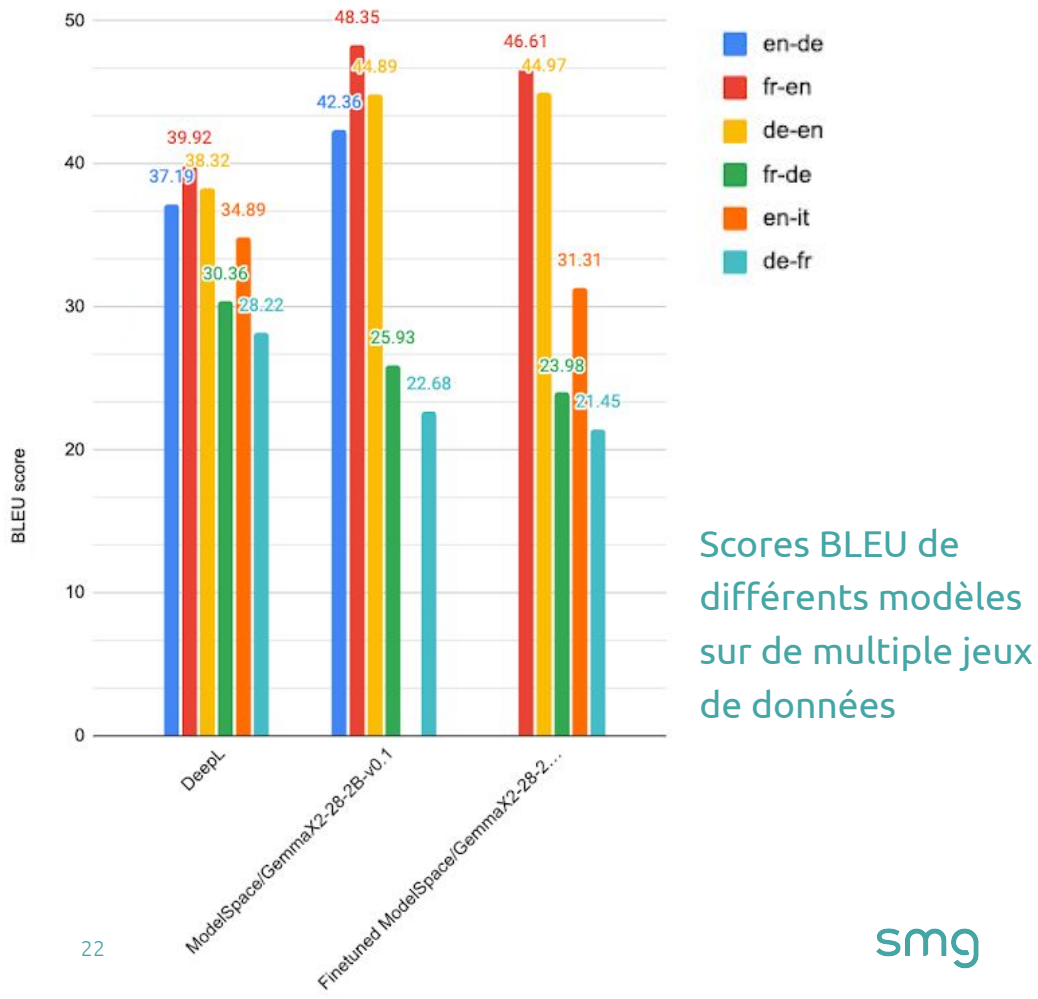
“I sell my phone 12 (64GB) in very good condition.”

Traduction DeepL :

“I’m selling my iPhone 12 (64 GB) in very good condition 👍.”

GemmaX2-28-2B

- Qualité de traduction remarquable
- Bonne robustesse face aux cas limites
- Taille réduite (2B de paramètres)



GemmaX2-28-2B



Texte d'origine :

“Je vends mon iPhone 12 (64 Go) en très bon état 👍.”

Traduction GemmaX2 :

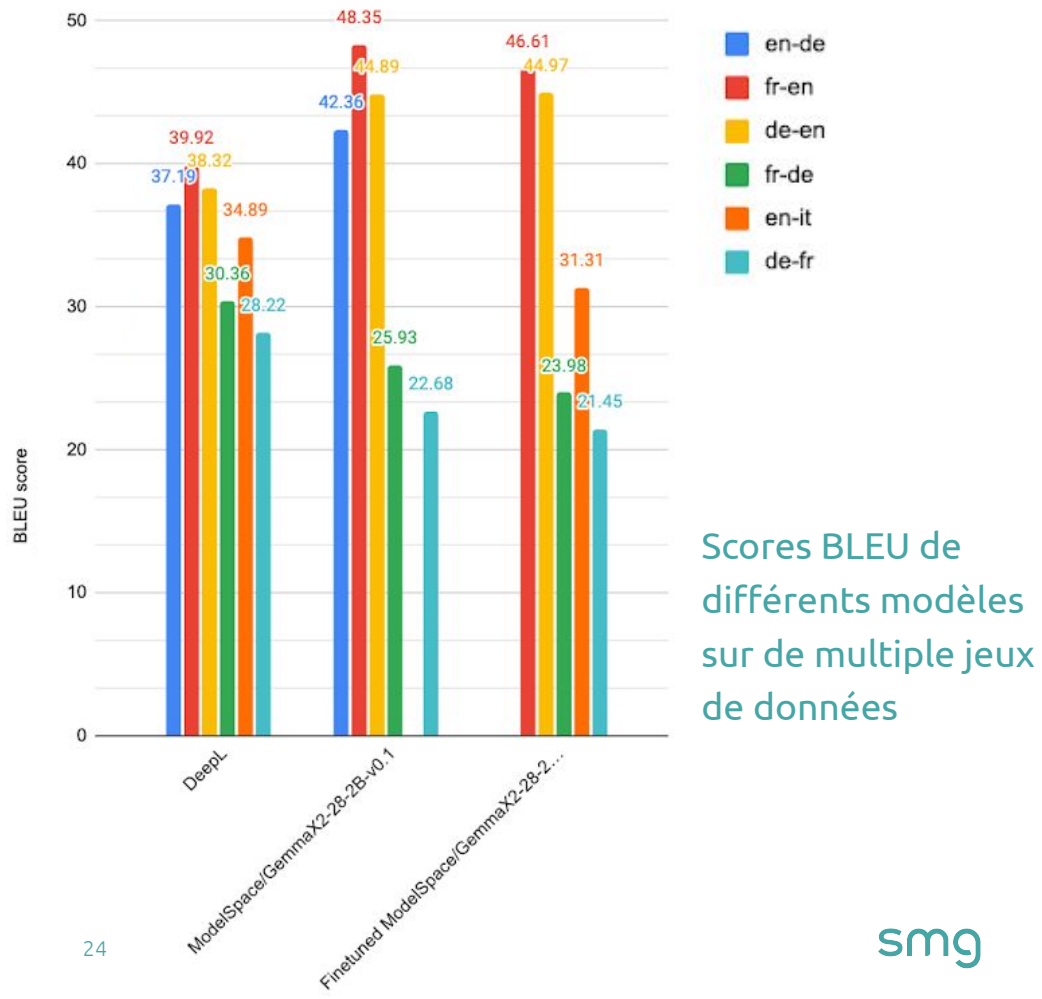
“I sell my iPhone 12 (64 GB) in very good condition 👍.”

Traduction DeepL :

“I’m selling my iPhone 12 (64 GB) in very good condition 👍.”

GemmaX2-28-2B

- Qualité de traduction remarquable
- Bonne robustesse face aux cas limites
- Taille réduite (2B de paramètres)



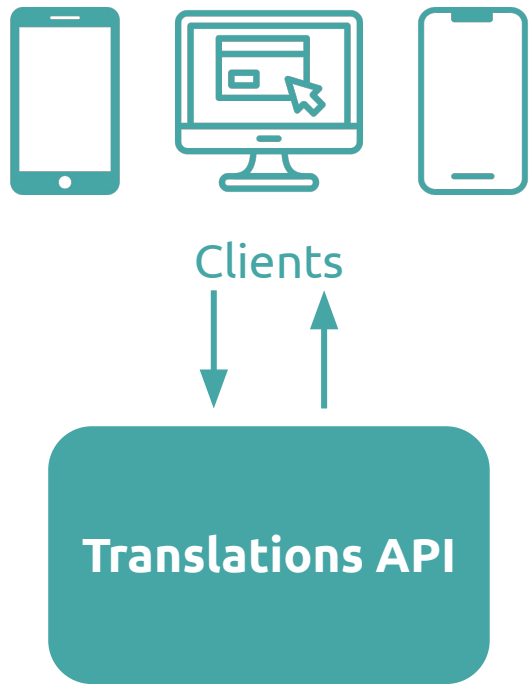
3. Construire l'alternative : Concevoir la solution technique

Architecture de la solution

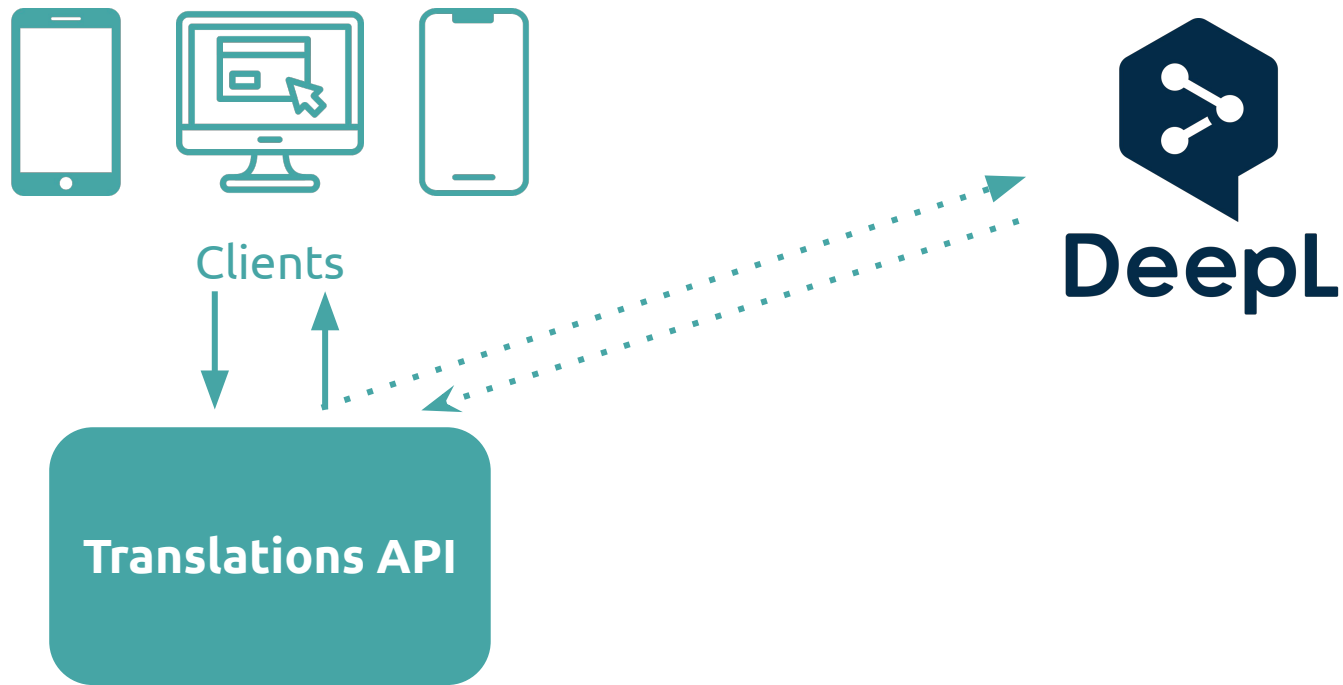


Translations API

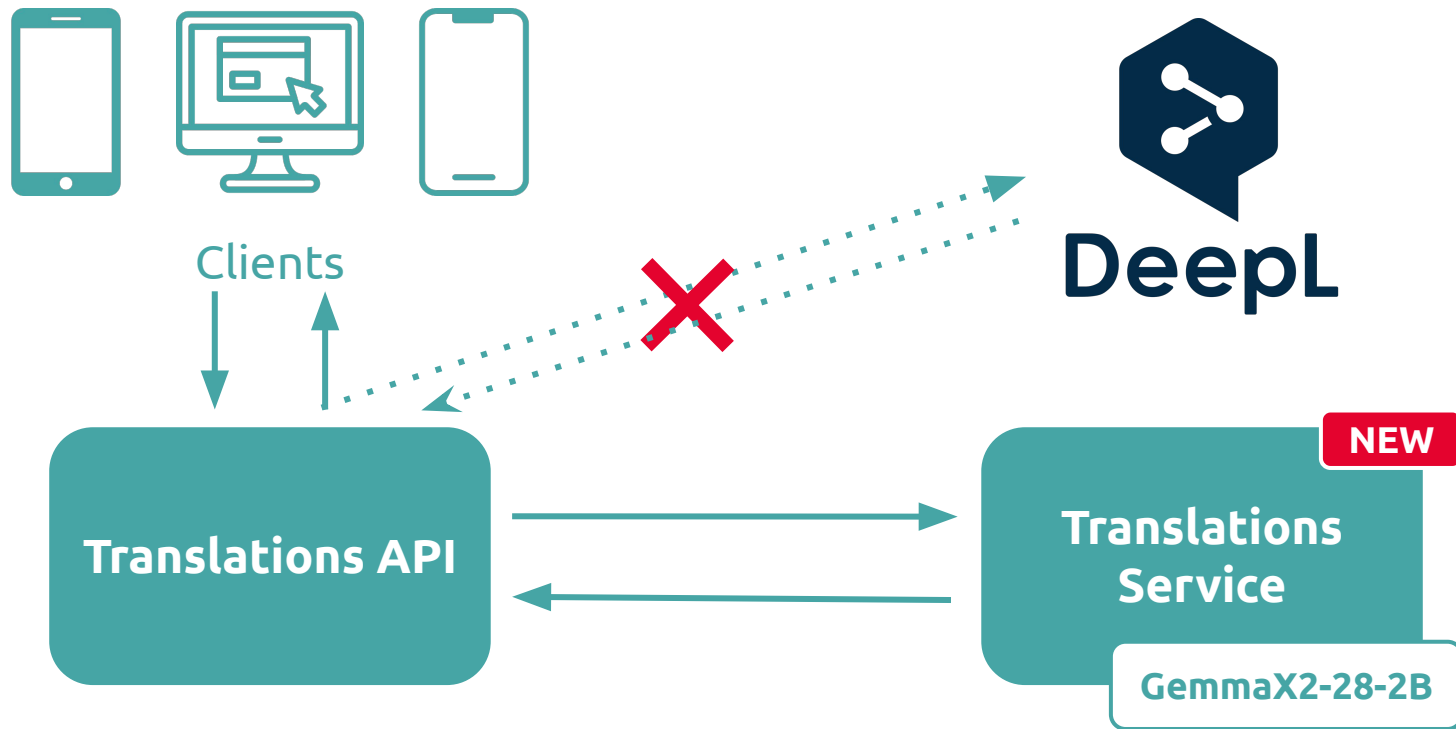
Architecture de la solution



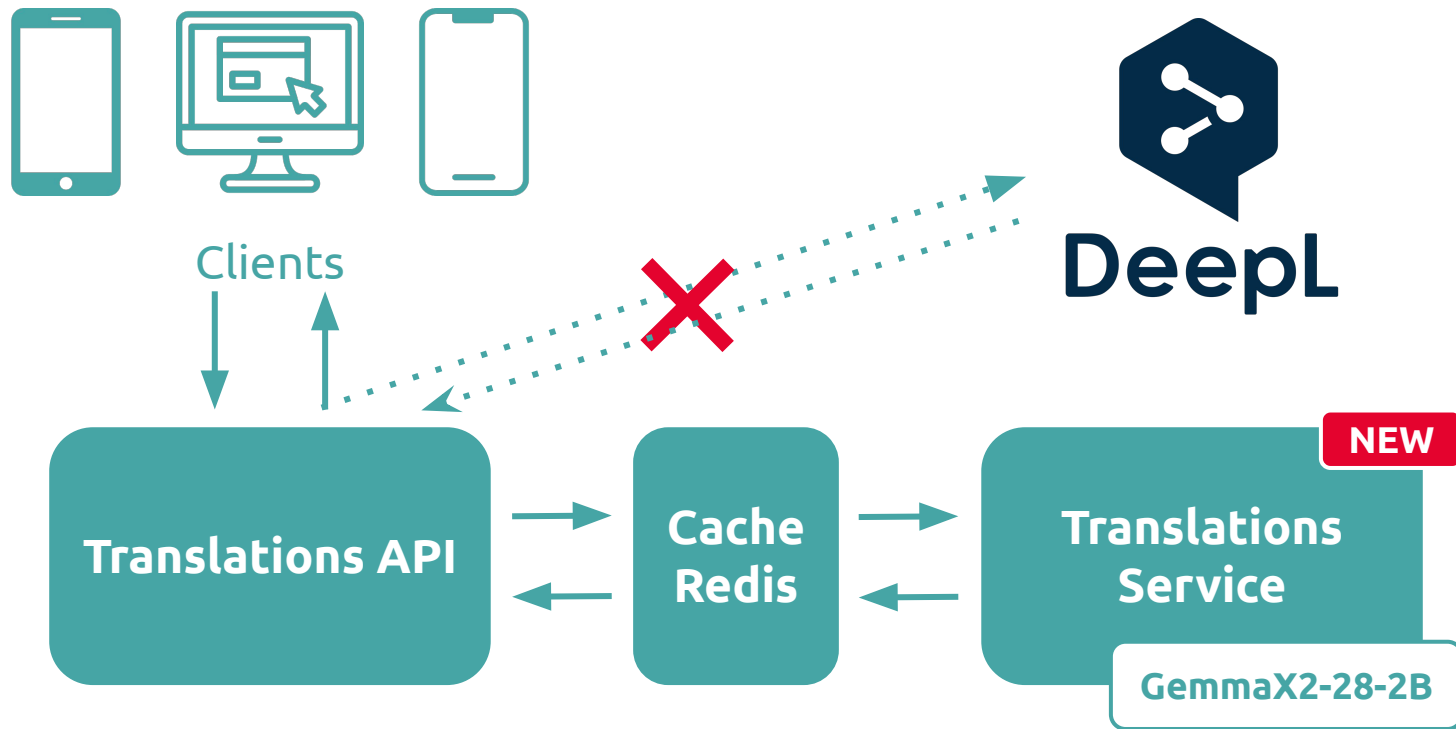
Architecture de la solution



Architecture de la solution



Architecture de la solution

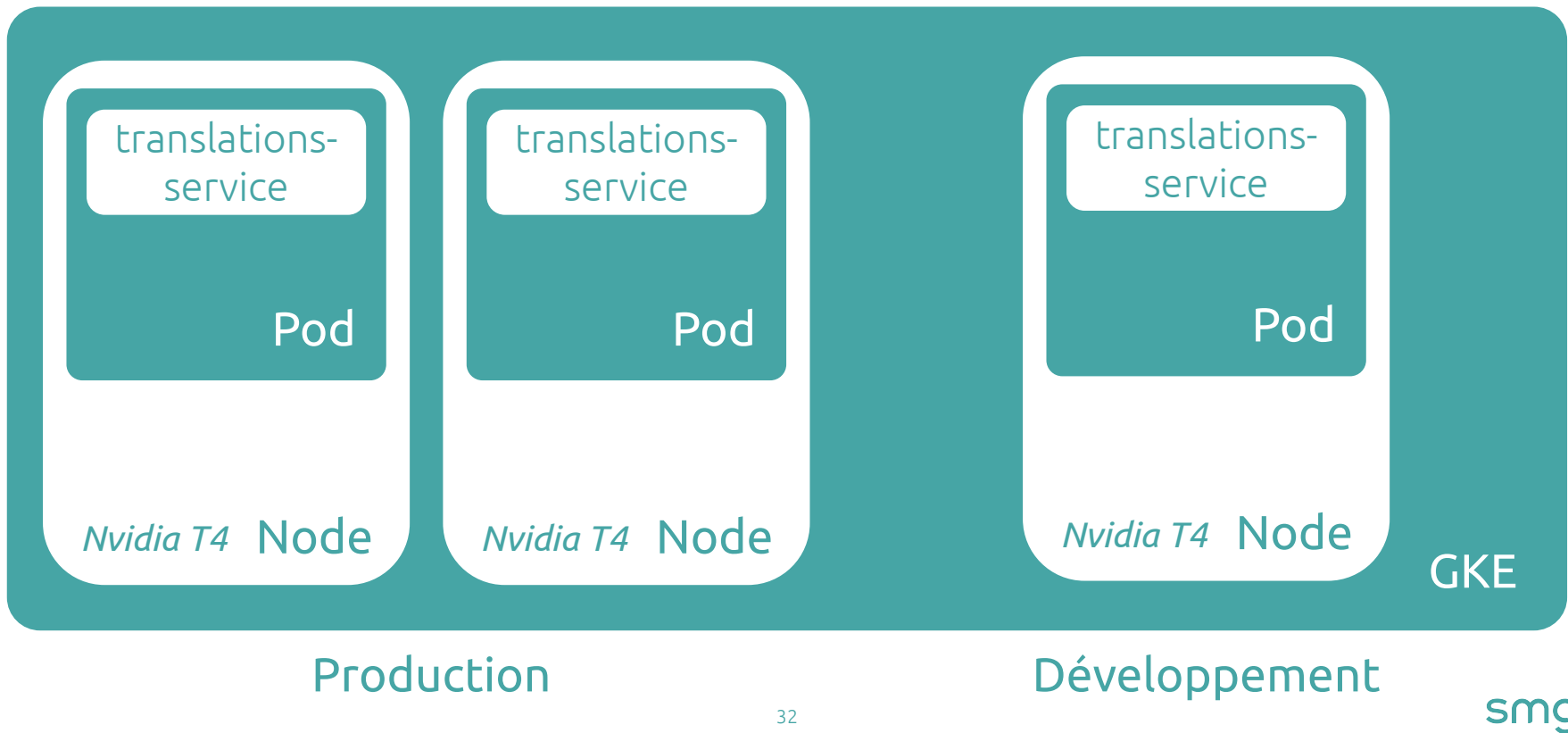


Optimisations

- Quantification FP16
- Utilisation de vLLM comme moteur d'inférence
- Parallélisation de la traduction

Résultat : passage d'un temps de réponse > 20s à moins de 3s

Infrastructure



Fonctionnement de la solution

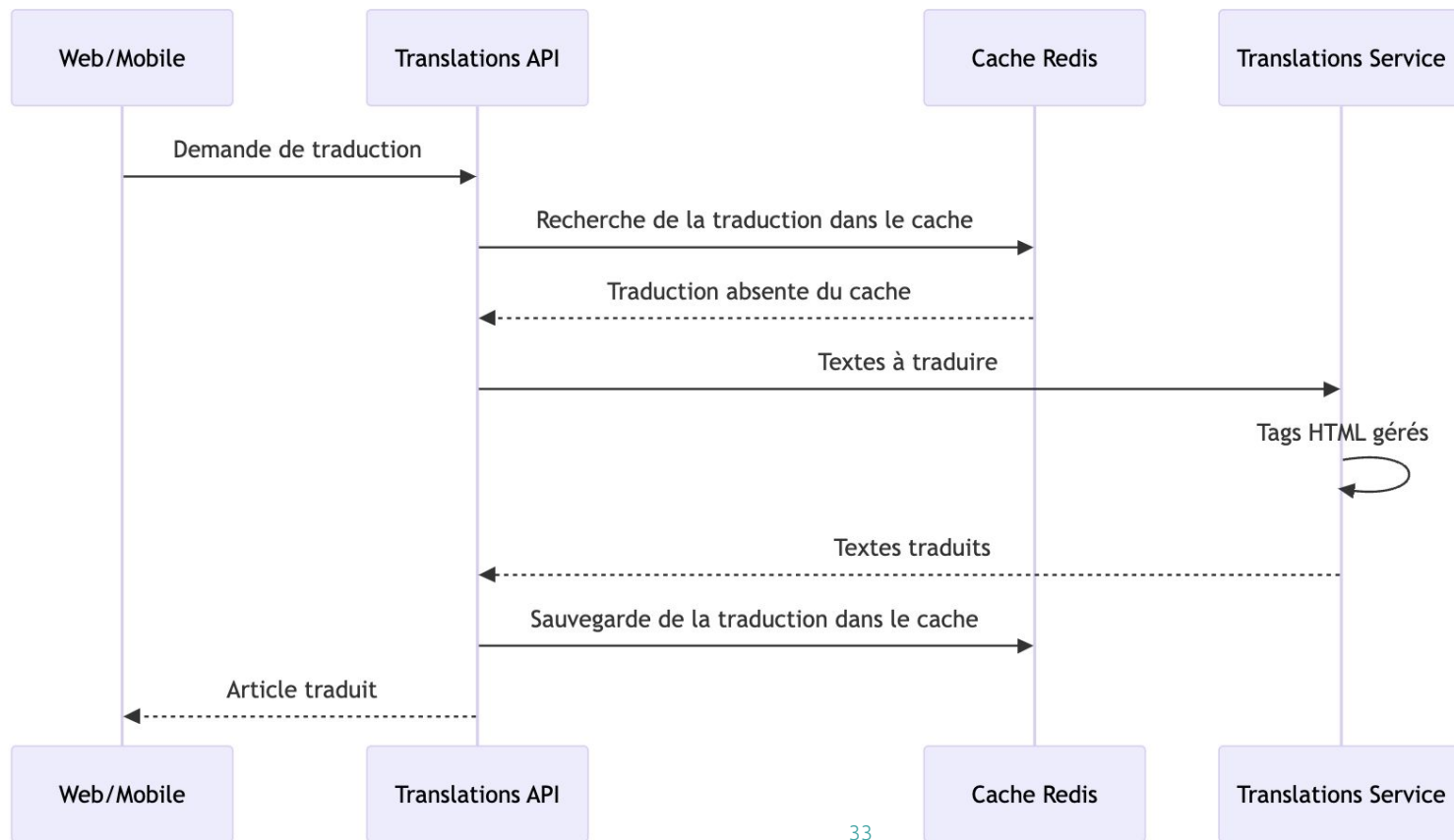


Diagramme de séquence montrant le nouveau système de traduction sur Ricardo

4. Validation :

Tester et confirmer les résultats

Validation de la solution

Qualité

Score BLEU > 30

Moyenne à 33,66

Test de charge

Load testing

Pic à 20 requêtes/s

0 redémarrage

Robustesse

2M traductions vs 1,8M

8k erreurs vs 50k

0,4% d'erreurs vs 2,83%

Performance

Latence moyenne de 2,7s
pour 95% des requêtes

A/B test

Même taux de conversion

Finance

87% d'économie

Validation de la solution

A/B test sur la plateforme web

Score BLEU > 30
Moyenne à 33,66



50% des utilisateurs web

Traduction via **DeepL**

Latence moyenne de 2,73 pour
95% des requêtes

Load testing avec Grafana k6
Pic à 20 requêtes/s
0 redémarrage

A/B test

Même taux de conversion

Robustesse

2M traductions vs 1,8M
8k erreurs vs 50k
0 erreurs vs 2,83%



50% des utilisateurs web

Traduction via **GemmaX2-28**

Validation de la solution

A/B test sur la plateforme web : 1M de sessions testées

Score BLEU > 30

Moyenne à 33,66



50% des utilisateurs web

Traduction via **DeepL**

Latence moyenne de 2,73 pour 95% des requêtes
Taux de conversion : **13,41%**

Load testing avec Grafana k6

Pic à 20 requêtes/s

0 redémarrage

A/B test

Même taux de conversion

2M traductions vs 1,8M

8k erreurs vs 50k

0 erreurs vs 2,83%



50% des utilisateurs web

Traduction via **GemmaX2-28**

Taux de conversion : **13,19%**

Validation de la solution

Qualité

Score BLEU > 30

Moyenne à 33,66

Test de charge

Load testing

Pic à 20 requêtes/s

0 redémarrage

Robustesse

2M traductions vs 1,8M

8k erreurs vs 50k

0,4% d'erreurs vs 2,83%

Performance

Latence moyenne de 2,7s
pour 95% des requêtes

A/B test

Même taux de conversion

Finance

87% d'économie

Validation de la solution

Bilan financier

Test de charge

Robustesse

Service	DeepL (CHF)	GemmaX2 (CHF)
API de traduction	7 000 CHF	3 instances GPU T4 ≈ 768 CHF
Cache Redis (identique)	188,34 CHF	188,34 CHF
Total mensuel	environ 7 188 CHF	environ 956 CHF

environ 7 708€

environ 1 025€

5. Apprentissages : Enseignements et perspectives

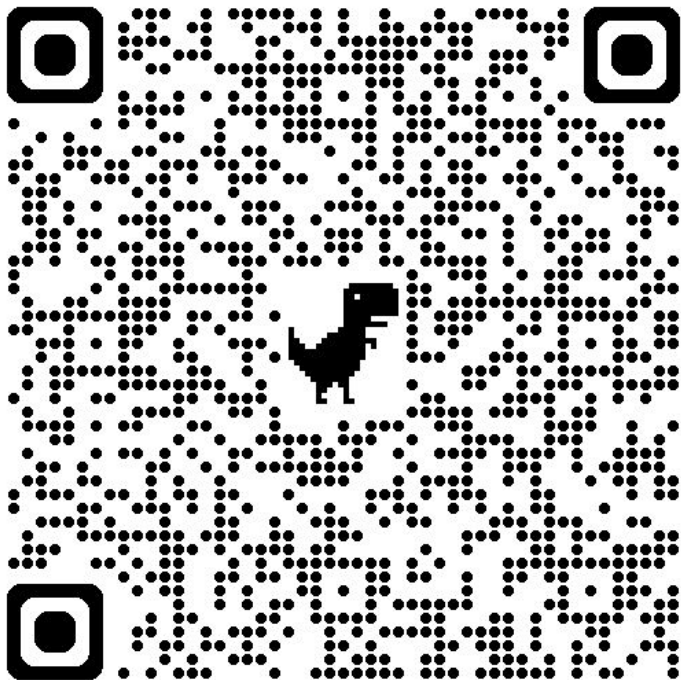
Apprentissages

- Les LLMs open-source ne sont plus des Proof Of Concept
- Une solution évolutive
- L'IA peut être durable



Conclusion

Pour aller plus loin...



<https://medium.com/smg-real-estate/build-a-low-cost-low-latency-low-energy-ai-translator-a-deepl-alternative-th-at-works-19b903d00845>