



# L'IA souveraine en production

Pas une promesse. Une app qui tourne.

# **SOUVERAINETÉ**

**le mot est partout,  
la réalité nulle part.**

*« Chaque prompt envoyé à un SaaS IA transite par des serveurs que vous ne contrôlez pas. »*

# **Tout le monde fait de l'IA.** Personne ne la maîtrise.

**8**

**abonnements  
SaaS IA actifs**

8 road maps  
que vous ne contrôlez pas.

**0**

**données  
sous votre contrôle**

Où vont vos prompts ?  
Personne ne le sait vraiment.

**∞**

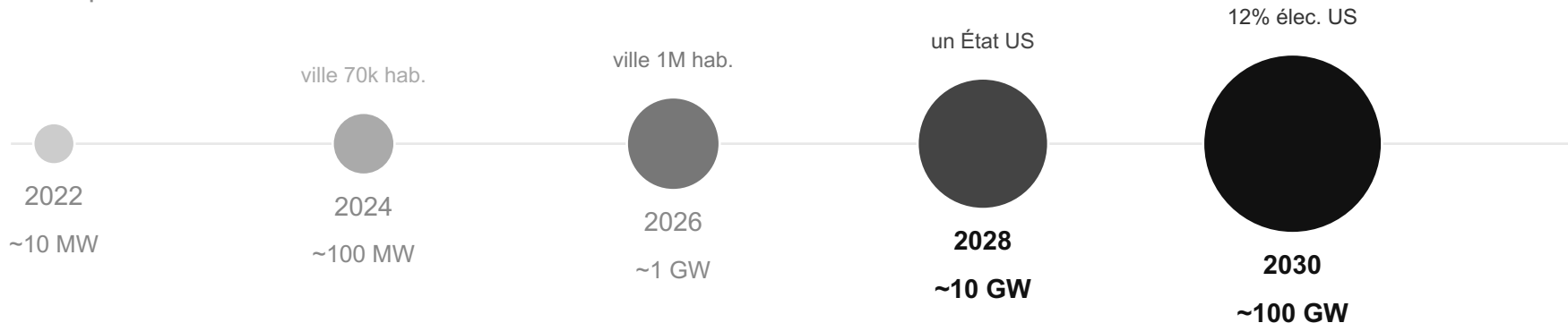
**dépendance  
fournisseur**

Changement de pricing,  
d'API, de politique — vous subissez.

# Mais à quel prix ?

## ×10 000

la capacité d'entraînement entre 2022 et 2030.



## Quel modèle choisir pour quelle tâche ?

*La réponse change tout — en coût, en énergie, en souveraineté.*

# Le paradoxe du sur-modèle

Tâche	Modèle utilisé	
Classifier un email spam	<b>SoloTiny — 4B</b>	● ○ ○ ○ ○
Analyser une offre fournisseur	<b>SoloSmall — 30B</b>	● ● ○ ○ ○
Rédiger un email de consultation	<b>SoloMedium — 80B</b>	● ● ● ○ ○
Raisonnement complexe multi-source	<b>SoloLarge — 150B</b>	● ● ● ● ●

« Pourquoi appeler GPT-5 pour classer un email ? On choisit son outil selon la tâche. »

**Solomund : notre  
vision de l'IA utile**

# Une stack ouverte, parce que l'IA n'arrête pas d'innover

WordPress

MariaDB

Nextcloud

Activepieces

LiteLLM

Qdrant

MinerU

OpenWebUI

*Chaque brique est remplaçable, extensible, intégrable.*

*« Construire et maintenir cette stack, c'est un job à plein temps. C'est le nôtre. »*

# Un process métier complet pour moins de 25 K

Approche classique

150 K+

- Intégrateur × consulting
- Vendor lock-in 5 ans
- Maintenance propriétaire
- Dépendance totale

Solomund

< 25 K

- ✓ Stack souverain ouvert
- ✓ Données chez vous
- ✓ Évolutif sans dépendance
- ✓ Livré en production

# **Présentation de cas**

Cas client

# Entreprise BTP & Construction, Suisse.

---

**200+**

projets  
gérés / an

---

**100+**

fournisseurs  
par projet

---

**4 j.**

de saisie  
manuelle / dossier

**App CORE**

# 200 pages de PDF.

## Des centaines d'articles à identifier.

CDC — CAN SIA

5.1 Terrassements généraux

5.1.1 Fouilles en masse

métré = 450 m<sup>3</sup>

5.1.2 Remblais compactés

5.2 Béton armé fondations

5.2.1 Semelles filantes

métré = 120 ml

5.2.2 Dalles de sol

**Le maître d'ouvrage surligne.**

L'acheteur recopie.


Article par article.

Dossier après dossier.

Depuis toujours.

**Jusqu'à maintenant.**

# Le workflow d'extraction.

 Vidéo — Upload PDF surligné → Pipeline → Liste d'articles structurés

# Sous le capot : 7 briques.

 Vidéo — Flow Activepieces s'exécutant brique par brique

**1**

Webhook  
réception

**2**

Nextcloud  
stockage

**3**

pdf-highlight  
extracteur

**4**

MinerU  
OCR

**5**

LLM  
SoloSmall

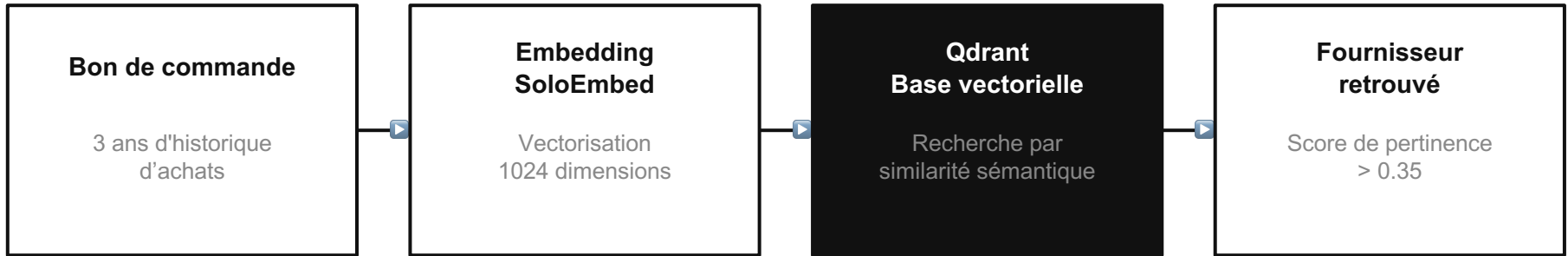
**6**

Structuration  
NocoDB

**7**

Notification  
acheteur

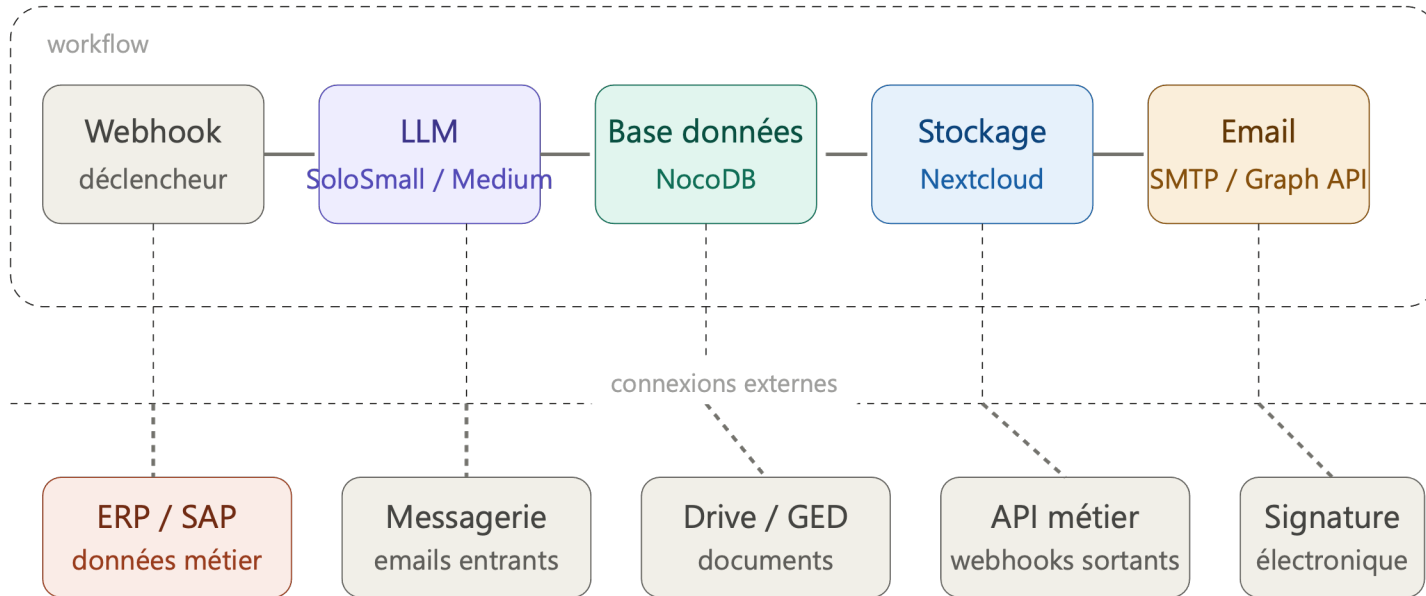
# 3 ans de mémoire fournisseurs par recherche sémantique.



**Par le sens, pas par les mots-clés.**

« Semelles filantes béton armé » → retrouve automatiquement les entreprises de gros œuvre qui ont livré ce type de prestation — même si le libellé exact n'est jamais le même.

# Des workflows connectés à vos outils d'entreprise.



« L'application s'intègre dans votre écosystème. Pas l'inverse. »

# De 4 jours à 4 heures.

## Avant

- Lecture PDF manuelle
- Recopie article par article
- Recherche fournisseurs par e-mail
- Consolidation manuelle des offres

## Avec Solomund

- ✓ Upload PDF surligné
- ✓ Extraction automatique — 8 à 60 min
- ✓ Matching sémantique Qdrant
- ✓ Analyse IA des offres reçues

**MERCI**



Souveraineté | Pertinence | Efficience