Maxence Foulon

Ressource gratuite · Checklist

Plan B IA : 8 protections pour qu'un projet IA survive à une coupure, à une hausse de prix ou à un krach

En 15 minutes, tu sais ce qui s'arrête dans ton entreprise si ton fournisseur d'IA coupe demain. En une journée, tu as un plan pour que ça ne t'arrête pas.

Lecture 10 min 16 sources vérifiées Mis à jour le 7 novembre 2026

L'essentiel en 30 secondes

- Anthropic a perdu 104 milliards de dollars sur neuf mois, surtout en écritures comptables, et des investisseurs doutent de son entrée en bourse en novembre. Le 5 novembre, le Nasdaq a eu sa pire séance depuis avril 2025.[1, 2, 3]
- Le risque concret pour une entreprise cliente, c'est la dépendance : en juin 2026, Claude Fable 5 a été coupé pour tous les clients pendant près de trois semaines, après une décision du gouvernement américain.[4, 5, 6]
- La parade tient en 8 protections. La première prend 15 minutes : l'inventaire. Pour le secours, Mistral a publié le 29 octobre les poids de Mistral Large 4 : loin derrière les meilleurs, mais téléchargeable.[7, 8]

Contexte

Ce qui se cache derrière le mot « bulle »

En octobre 2026, trois signaux se sont croisés. Les obligations de CoreWeave, le plus connu des loueurs de GPU (les « neoclouds »), ont vu leur rendement approcher 15 % sur le marché, et son

action a perdu la moitié de sa valeur entre le 1er octobre et le 5 novembre.[9, 10] Elon Musk a prédit la disparition ou le rachat de la plupart des neoclouds, plus de 200 selon lui, faute de crédit.[11] Et Anthropic a déposé sa demande d'introduction en bourse, avec une perte nette de 80,7 milliards de dollars sur six mois.[4]

Ce dernier chiffre fait peur, mais il se lit mal. 72,4 milliards viennent de la réévaluation de titres financiers, surtout des obligations convertibles en actions : plus l'entreprise prend de la valeur, plus ces titres pèsent dans les comptes, sans qu'un dollar sorte. La perte d'exploitation est de 693 millions, pour 16,3 milliards de chiffre d'affaires.[4]

Début novembre, la tension est montée d'un cran. La mise à jour de son dossier d'introduction en bourse (le S-1/A, déposé auprès de la SEC, le gendarme boursier américain) ajoute 23,3 milliards de perte nette au troisième trimestre, là encore surtout comptable (21,9 milliards de réévaluation), pour 17 milliards de chiffre d'affaires.[1] Le 5 novembre, deux jours après les élections de mi-mandat américaines, le Nasdaq a eu sa pire séance depuis avril 2025 (–5,5 %), et des investisseurs ont dit à Reuters douter que l'entrée en bourse d'Anthropic se fasse en novembre.[2, 3]

Le vrai poids

Les engagements de calcul. Anthropic a signé environ 895 milliards de dollars d'achats et de locations, presque tous pour du calcul et de l'infrastructure : 447,5 milliards avant fin 2030, le reste jusqu'en 2045.[1] C'est ce pari, et celui de toute la chaîne qui finance les GPU, qui se joue.

Pourquoi ça te concerne

Tu ne paies ni CoreWeave ni Nvidia. Mais ton fournisseur d'IA loue son calcul à cette chaîne, et il obéit au droit américain. En juin 2026, le Département du Commerce a imposé des restrictions d'export sur Fable 5 et Mythos 5. Faute de pouvoir trier ses clients par nationalité, Anthropic a coupé ces deux modèles pour tout le monde le 12 juin, deux jours après le lancement de Fable 5. Retour à partir du 1er juillet.[4, 5, 6] Plus tôt dans l'année, le 27 février, le président américain avait ordonné à l'administration fédérale d'arrêter d'utiliser les modèles d'Anthropic.[1] Son dossier d'introduction en bourse (le S-1) prévient que ces restrictions d'export peuvent se reproduire, durer plus longtemps ou toucher d'autres modèles.[4]

01

Fais l'inventaire (15 minutes)

Avant de choisir quoi que ce soit, sache de quoi tu dépends. Liste chaque usage de l'IA dans ton équipe, y compris ceux que personne n'a déclarés : l'extension de navigateur de la compta, l'assistant du commercial, le script d'un stagiaire.

- [ ] Pour chaque usage : qui s'en sert, pour quelle tâche, avec quel outil et quel modèle.
- [ ] Ce qui se passe s'il disparaît une semaine : rien, gêne ou blocage.
- [ ] Où partent les données : quel fournisseur, quel pays. Si tu ne sais pas, note « à demander ».

Fiche de dépendance IA (à copier dans un tableur)

```
Usage | Équipe | Outil | Modèle | Fournisseur | Pays des données |
Si coupé 1 semaine | Secours testé ? | Responsable
Exemple : réponses au support | Service client | Outil de tickets +
API | (modèle) | (fournisseur) | à demander | Blocage | Non |
(prénom)
```

02

Teste un modèle de secours, pas seulement un nom

Pendant la coupure de Fable 5, les autres modèles d'Anthropic, dont Opus 4.8, restaient disponibles.[5] Une équipe qui avait déjà testé ses tâches sur un autre modèle pouvait basculer en changeant un réglage. Une équipe qui ne l'avait pas fait découvrait les écarts en production. Fable 5 venait de sortir, mais le mécanisme vaut pour n'importe quel modèle.

- [ ] Choisis deux secours : un autre modèle du même fournisseur (coupure ciblée sur un modèle) et un modèle d'un autre fournisseur (coupure plus large, hausse de prix, décision d'un État).
- [ ] Passe ton jeu de tests (voir 04) sur les deux, et note les écarts.
- [ ] Pour les données sensibles, garde une option hébergée en Europe ou chez toi.

03

Une seule porte d'entrée dans ton code

Si le nom du modèle est écrit en dur dans quarante fichiers, changer de fournisseur prend des semaines. Si tous les appels passent par une seule fonction interne et que le modèle se règle dans la configuration, la bascule technique prend une heure ; reste à vérifier les résultats avec ton jeu de tests (voir 04).

- [ ] Tous les appels d'IA passent par un seul module interne.
- [ ] Modèle principal et modèles de secours sont dans la configuration, pas dans le code.
- [ ] Les prompts sont stockés à part, versionnés, avec le modèle pour lequel ils ont été testés.
- [ ] Un interrupteur permet de basculer sans redéployer.
- [ ] Le fournisseur de secours est validé pour tes données (contrat, pays, sécurité) avant qu'une bascule automatique lui envoie quoi que ce soit.

Gabarit de configuration (à adapter)

```
ia:
  principal:
    fournisseur: "à remplir"
    modele: "à remplir"
  secours_meme_fournisseur:
    modele: "à remplir"
  secours_autre_fournisseur:
    fournisseur: "à remplir"
    modele: "à remplir"
  bascule:
    automatique_si: "plus de 5 % d'erreurs sur 10 minutes"
    responsable: "prénom"
  budget:
    plafond_mensuel_eur: "à remplir"
    alerte_a: "80 %"
```

04

Ton propre jeu de tests

Les éditeurs publient les tests où ils brillent. Moonshot, l'éditeur chinois de Kimi, présente Kimi K3.1 Pro à 0,7 point de Claude Opus 5.5 sur Terminal-Bench 4.0 (65,7 contre 66,4), à moitié prix : 2 $ contre 4 $ par million de jetons en entrée, 10 $ contre 20 $ en sortie.[12, 13, 14] Ça ne dit rien de tes contrats, de tes tickets ni du ton de ta marque. Vingt cas réels le disent.

- [ ] Rassemble 20 demandes réelles, anonymisées, avec la réponse que tu attends.
- [ ] Ajoute 5 cas pièges : information manquante, demande hors périmètre, donnée personnelle à ne pas répéter.
- [ ] Note chaque modèle candidat avec la même grille, à chaque changement de modèle ou de prix.

Prompt : transformer tes cas réels en jeu de tests

```
Voici 20 demandes réelles de mon équipe, anonymisées, avec la
réponse que nous jugeons bonne.
Pour chacune :
1. Liste ce que la réponse doit absolument contenir (3 critères
vérifiables au plus).
2. Indique ce qui la rendrait inacceptable (erreur, ton, information
inventée).
3. Propose une note de 0 à 2 par critère.
Ajoute ensuite 5 cas pièges inspirés de ces demandes : information
manquante, demande hors périmètre, donnée personnelle à ne pas
répéter.
Rends le tout sous forme de tableau : Cas | Entrée | Critères |
Inacceptable si | Barème.

[Colle tes 20 demandes ici]
```

05

Les questions à poser avant de signer

Ton contrat est ta seule protection juridique contre une hausse de prix ou le retrait d'un modèle. Le dossier d'introduction en bourse d'Anthropic (S-1) évoque lui-même de possibles changements de tarification, au résultat ou à la tâche.[4] Je ne suis pas juriste : fais relire tes clauses par un juriste. Voici les questions à poser à ton fournisseur.

E-mail à envoyer à ton fournisseur d'IA

```
Bonjour,

Avant de signer ou de renouveler notre contrat, pouvez-vous nous
préciser par écrit :
1. Le préavis en cas de hausse de prix, et s'il existe un plafond.
2. Le préavis avant le retrait d'un modèle que nous utilisons, et la
durée de la période de migration.
3. Comment nous exportons nos données, nos prompts et nos journaux :
format et délai.
4. Dans quels pays nos données sont traitées et stockées, et par
quels sous-traitants.
5. Ce qui se passe pour nous si une restriction d'export ou une
décision d'un État vous oblige à couper un modèle : préavis, modèle
de remplacement, remboursement.
6. Vos engagements de disponibilité et les compensations prévues.

Merci,
(prénom)
```

06

Un budget qui encaisse une hausse

Les prix bougent. En septembre 2026, Anthropic a baissé d'environ 25 % le coût des tâches courantes avec Fable 5.1, et d'environ 40 % le coût par tâche avec Opus 5.5.[4] Mais ces baisses se font dans un marché où les fournisseurs s'engagent sur des centaines de milliards de calcul et où leurs loueurs empruntent cher.[1, 9] Rien ne garantit qu'elles dureront.

- [ ] Mesure le coût par tâche (par ticket traité, par document produit), pas seulement la facture du mois.
- [ ] Calcule ce que devient ton projet si le prix double. S'il ne tient plus, repère le modèle moins cher qui suffit.
- [ ] Mets un plafond et une alerte de dépense chez chaque fournisseur.

07

Une option ouverte, en lisant la licence

Un modèle aux poids ouverts, que tu héberges toi-même ou chez un hébergeur européen, ne dépend d'aucune décision de son éditeur une fois téléchargé. Mais « ouvert » ne veut pas dire « libre ». La licence de Kimi K3.1 Pro, par exemple, impose un accord séparé à qui le revend en service au-delà de 20 millions de dollars de chiffre d'affaires sur 12 mois ; l'usage interne en est exempté.[15]

Côté européen, Mistral a publié le 29 octobre les poids de Mistral Large 4, sous une licence MIT modifiée qui autorise l'usage commercial ; seule l'exploitation d'un service qui revend l'accès au modèle demande une licence séparée.[7, 16] Il est loin derrière les meilleurs : 26 sur l'indice d'intelligence d'Artificial Analysis, un classement indépendant, le 7 novembre, contre 46 pour Kimi K3.1 Pro et 57 à 60 pour les meilleurs modèles américains.[8] Avec environ 1 000 milliards de paramètres, il tourne en pratique chez un hébergeur, pas sur le serveur d'une PME.[16] Un secours n'a pas besoin d'être le meilleur : il doit tenir tes tâches critiques, et seul ton jeu de tests le dira.

- [ ] Repère un modèle ouvert capable de tenir tes tâches critiques, même moins bien.
- [ ] Lis sa licence : usage commercial, revente, mentions obligatoires.
- [ ] Passe-le dans ton jeu de tests et note l'écart avec ton modèle principal.

08

Un plan de coupure écrit, testé chaque trimestre

Le jour où ton modèle disparaît, personne ne doit chercher qui décide. Une page suffit.

Plan de coupure IA (une page)

```
Déclencheur : le modèle principal est indisponible, interdit, ou son prix augmente de plus de ___ %.
Décideur : ___ (remplaçant : ___).
Délai de décision : ___ heures.
Bascule : passer sur ___ (même fournisseur), puis sur ___ (autre fournisseur) si besoin.
Ce qui s'arrête pendant la bascule : ___.
Message aux équipes : « L'assistant ___ fonctionne avec un autre modèle jusqu'au ___. Signalez toute réponse anormale à ___. »
Test : une fois par trimestre, basculer une journée (ou une partie du trafic) sur le secours et comparer les résultats.
```

- [ ] Écris la page et partage-la avec les équipes concernées.
- [ ] Planifie le premier test de bascule dans le mois.

Sources

1. SEC EDGAR, Anthropic, PBC, Form S-1/A https://www.sec.gov/Archives/edgar/data/2130357/000162828026068676/anthropicpbc-sx1ax1.htm, 2 novembre 2026.
2. Reuters, via Investing.com, S&P 500 futures steady after midterm rally; Nasdaq lags on chip weakness https://www.investing.com/news/stock-market-news/sp-500-futures-steady-after-midterm-rally-nasdaq-lags-on-chip-weakness-5003237, 5 novembre 2026.
3. Reuters, via Investing.com, Exclusive: AI backlash could threaten Anthropic IPO timing, potential investors say https://www.investing.com/news/stock-market-news/exclusiveai-backlash-could-threaten-anthropic-ipo-timing-potential-investors-say-5006110, 5 novembre 2026.
4. SEC EDGAR, Anthropic, PBC, Form S-1 https://www.sec.gov/Archives/edgar/data/2130357/000162828026067168/anthropicpbc-sx1.htm, 26 octobre 2026.
5. Greenberg Traurig, AI Company Anthropic Suspends Access to Claude Fable 5, Claude Mythos 5 Following US Export Control Directive https://www.gtlaw.com/it/insights/2026/6/ai-company-anthropic-suspends-access-to-claude-fable-5-claude-mythos-5-following-us-export-control-directive, 17 juin 2026.

6. Engadget, US government allows Anthropic to redeploy its Mythos and Fable AI models https://www.engadget.com/2205599/anthropic-redeploy-mythos-fable-ai-models/, 30 juin 2026.
7. Mistral AI, Mistral Large 4 (Le Chonk) is now open weight https://mistral.ai/news/mistral-large-4-oss/, 29 octobre 2026.
8. Artificial Analysis, LLM Leaderboard, Intelligence Index v5.0.1 https://artificialanalysis.ai/leaderboards/models, consulté le 7 novembre 2026.
9. CNBC, CoreWeave shares tumble as bond yields near 15% on concern about AI borrowing binge https://www.cnbc.com/2026/10/14/coreweave-stock-bond-yields-ai-debt.html, 14 octobre 2026.
10. StockAnalysis, CoreWeave (CRWV) stock price history https://stockanalysis.com/stocks/crwv/history/, consulté le 6 novembre 2026.
11. CNBC, Elon Musk says more than 200 neoclouds face 'guppy-level death' https://www.cnbc.com/2026/10/26/neocloud-stocks-today-coreweave-nebius-iren-elon-musk.html, 26 octobre 2026.
12. Moonshot AI, Kimi K3.1 Pro https://www.kimi.ai/blog/kimi-k3-1-pro, 23 octobre 2026.
13. Artificial Analysis, Kimi K3.1 Pro https://artificialanalysis.ai/models/kimi-k3-1-pro, consulté le 6 novembre 2026.
14. Anthropic, Pricing (Claude API) https://platform.claude.com/docs/en/about-claude/pricing, consulté le 6 novembre 2026.
15. Moonshot AI, Kimi K3.1-Pro License https://huggingface.co/moonshotai/Kimi-K3.1-Pro/raw/main/LICENSE, Hugging Face, consulté le 6 novembre 2026.
16. Hugging Face, Mistral AI, Mistral-Large-4-1T-A52B https://huggingface.co/mistralai/Mistral-Large-4-1T-A52B, consulté le 7 novembre 2026.

Tu veux un regard extérieur sur ta fiche de dépendance ?

Je prends des projets en freelance : je conçois et livre des outils IA branchés sur les logiciels qu'une équipe utilise vraiment. Envoie-moi ta fiche, on regarde ensemble où ça casserait.

Parler de ton projet **Voir mes projets**

KakaoTalk, WhatsApp ou e-mail à copier : maxence.pro/fr/#contact

Les autres ressources gratuites : maxence.pro/fr/resources/

© 2026 Maxence Foulon · maxence.pro · Partage libre, en citant la source.