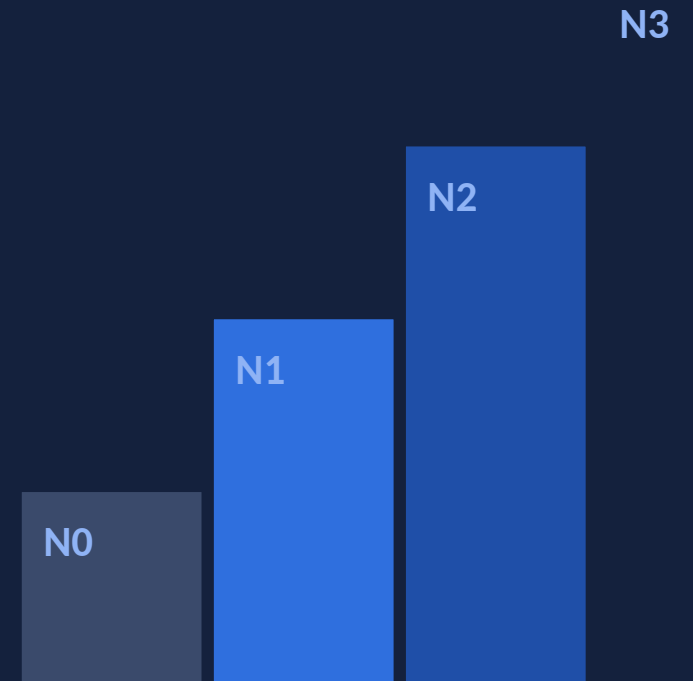


PRÉPARÉ POUR LES ÉQUIPES ACME • SÉCURITÉ • ARCHITECTURE • DATA

Gemini Enterprise

Feuille de route sécurité

De la découverte dans la console Google Cloud à une production industrialisée en Terraform, niveau par niveau.



Sécuriser par niveaux, comme on monte un escalier



Découverte

Où

Console Google Cloud

Qui

10 à 20 testeurs

Données

Données non sensibles

Durée indicative

1 à 2 semaines



Socle

Où

Terraform

Qui

Pilote de 50 à 200

Données

Données internes

Durée indicative

3 à 4 semaines



Production

Où

Terraform + CI/CD

Qui

Ouverture par groupes

Données

Données internes

Durée indicative

4 à 6 semaines



Avancé

Où

Terraform + organisation

Qui

Tous

Données

Sources étendues

Durée indicative

Selon besoins

Chaque marche apporte une protection mesurable, élargit le périmètre d'usage, et ne commence que lorsque la précédente est validée. Les mesures sont cumulatives : le niveau 2 contient tout le niveau 1.

Aller vite, sans prendre de risque



Démarrer demain

Test en console en quelques jours, sur des données sans enjeu, pendant que le socle de production se construit.



Une sécurité démontrable

Chaque mesure répond à un risque nommé, se teste en recette et laisse une trace d'audit.



Impossible à affaiblir par erreur

Plancher Model Armor, séparation des rôles, alerte dès qu'un garde-fou change.



RGPD dès la conception

Données en UE, sources interrogées sans copie, conservation minimisée, journaux réservés.



Industrialisé

La même configuration Terraform, relue et versionnée, en recette et en production.

*« La bonne question n'est pas : l'IA est-elle sûre ?
C'est : que peut-elle voir, que lui envoie-t-on, et comment le sait-on ? »*

Le reste de la présentation répond à ces trois questions, niveau par niveau.

Quatre décisions à prendre à froid

1

Fournisseur d'identité

Google ou Entra ID

En changer plus tard efface l'historique des conversations

Quand : Avant le niveau 1

2

Localisation

UE recommandée

Immuable après création ; conditionne le chiffrement CMEK

Quand : Avant le niveau 1

3

Sources du pilote

Drive, SharePoint...

Définit connecteurs, secrets et autorisations à préparer

Quand : Pendant le niveau 0

4

Conservation et journaux

DPO, CSE si besoin

Données personnelles de salariés

Quand : Avant le niveau 2

Six principes qui guident tous les choix



Tester en console, industrialiser en code

La console pour comprendre ; Terraform pour garantir, relire et reproduire.



Mesurer avant de bloquer

Observer les faux positifs, régler, puis appliquer. Identité et données : bloquants dès le jour 1.



Les droits de la source font foi

Mode fédéré par défaut : chacun ne voit que ce qu'il voit déjà, sans copie.



Défense en profondeur

Identité, connecteurs, filtrage et détection se complètent ; aucune mesure ne suffit seule.



Personne ne désactive sa propre sécurité

Admins de l'app ≠ sécurité. Plancher technique et alerte sur toute modification.



Décisions tracées et réversibles

Variables documentées, revue de code, validation par la bonne personne.

L'escalier de sécurité

Pourquoi des niveaux plutôt qu'un « big bang » ? Trois décisions sont irréversibles : localisation, identité, CMEK. On les prend après avoir appris.

N0 • Découverte

Console

- App de test en UE
- SSO Workspace
- Model Armor natif
- Recherche web coupée

N1 • Socle

Terraform

- Groupes, séparation des rôles
- Model Armor + DLP France
- Connecteurs fédérés
- Secrets en coffre
- Journaux, alertes

N2 • Production

Terraform + CI/CD

- Clés Acme (CMEK)
- Preuves verrouillées
- Politiques d'organisation UE
- Revue de code obligatoire

N3 • Avancé

Terraform + organisation

- VPC Service Controls
- Agent Gateway
- Agents maison filtrés
- Databricks via MCP

Six couches de protection, qui montent en puissance

Couche	Question à laquelle elle répond	N0	N1	N2	N3
1 • Identité et accès	Qui peut entrer, et qui peut changer les règles ?	SSO	Complet	Revue d'accès	Entra ID si besoin
2 • Protection des données	Qui contrôle le chiffrement et les secrets ?	Google	Secrets	CMEK	Politiques d'org.
3 • Connecteurs	Que peut voir l'assistant, avec quels droits ?	1 source test	Fédéré	IP fixes	MCP, Databricks
4 • Filtrage IA	Qu'est-ce qui entre et sort du modèle ?	Natif	DLP France	Red team	Agents maison
5 • Exposition	Où vont les données, que partage-t-on ?	UE, web coupé	Partages fermés	Ouverture gouvernée	VPC-SC
6 • Audit et détection	Comment le sait-on ?	Audit	Alertes	Logs verrouillés	SIEM



Tester directement dans la console Google Cloud

Où console Google Cloud Qui 10 à 20 testeurs Données non sensibles Durée 1 à 2 semaines

Ce que l'on active	Pourquoi	Où dans la console
Projet de découverte dédié	Isoler les essais ; supprimé ensuite	Resource Manager → Nouveau projet
App Gemini Enterprise en UE	Résidence européenne dès le départ	Gemini Enterprise → Créer une app → EU
Identité Google Workspace (SSO)	Aucun compte local, 2FA de l'annuaire	Gemini Enterprise → Paramètres → Identité
Licences d'essai, groupe testeurs	Périmètre maîtrisé et nominatif	Abonnements ; IAM → groupe
Google Drive fédéré (dossier test)	Vérifier que les droits Drive sont respectés	App → Data stores → Google Drive
Model Armor sur l'assistant	Voir les blocages : injection, données sensibles	Security → Model Armor ; App → Configurations
Recherche web publique coupée	Aucune question vers un moteur public	App → Configurations → Assistant
Journaux d'accès aux données	Premières traces d'audit	IAM → Journaux d'audit

Ce que l'on teste ensemble, et la règle d'or

- ✓ Un testeur ne retrouve pas un document Drive qui ne lui est pas partagé
- ✓ « Ignore tes instructions... », une carte de test ou un lien malveillant sont bloqués
- ✓ Les réponses citent leurs sources ; qualité jugée sur des cas d'usage réels
- ✓ Faux positifs de Model Armor observés sur le vocabulaire Acme



Règle d'or du niveau 0

Aucune donnée personnelle ou confidentielle.
L'app de découverte est jetable : elle n'a pas le chiffrement CMEK, qui doit être activé avant la création de l'app de production.

Critères de passage au niveau suivant

- ☐ Les 4 décisions structurantes sont prises
- ☐ Tests passés et faux positifs documentés
- ☐ Inventaire des réglages console établi, pour reprise en Terraform

Les six couches en code, sur un pilote métier

Où Terraform (projet 06)

Qui pilote de 50 à 200

Données internes

Durée 3 à 4 semaines



1 • Identité et accès

4 groupes ; séparation des rôles : les admins de l'app ne touchent ni filtres, ni clés, ni journaux



2 • Données

Secrets des connecteurs en coffre ; conversations conservées 30 jours au lieu de 60



3 • Connecteurs

Modèle générique fédéré : droits natifs, aucune copie ; IP sortantes fixes



4 • Filtrage IA

Model Armor entrée et sortie, FAIL_CLOSED, détecteurs France (NIR, IBAN...), plancher



5 • Exposition

UE, recherche web coupée, partages d'agents et de conversations fermés



6 • Détection

Journaux dédiés en UE lisibles par la sécurité seule ; alertes garde-fou et pic de blocages

Critères de passage au niveau suivant

- ☐ Recette de sécurité passée (13 tests) et points « à valider » levés
- ☐ DPO : conservation et journalisation validées ; pilote informé
- ☐ Valeur métier confirmée par le pilote

Maîtrise du chiffrement, preuves et exploitation

Où Terraform + CI/CD Qui ouverture par groupes Données internes Durée 4 à 6 semaines

Mesure	Pourquoi
CMEK : clé Acme (KMS, Europe)	Clé désactivée = données illisibles en 15 min
Journaux de sécurité verrouillés	Rétention et preuves non modifiables
Politiques d'organisation	UE et CMEK imposés à tout nouveau projet
Chaîne CI/CD Terraform	Plus aucune modification manuelle en production
Access Transparency	Accès éventuels du personnel Google journalisés
Exploitation	Revue d'accès, rotation des secrets, red team

Pourquoi CMEK maintenant ?
La clé doit être enregistrée avant la création de l'app qu'elle protège, et elle engage l'exploitation : rotation, révocation, astreinte.

Critères de passage au niveau suivant

- ☐ Test de révocation de clé réalisé en recette
- ☐ CI/CD en place : plus de droit d'écriture humain en production
- ☐ Tableaux de bord d'usage et de sécurité suivis

Étendre la valeur sans étendre la surface d'attaque

Où Terraform + organisation

Qui tous

Données sources étendues

Durée selon besoins

Mesure	Pourquoi	Quand
Agent Gateway	Pare-feu des agents : refus par défaut, règles par outil	Premier agent maison ou MCP
Model Armor par agent	L'app ne filtre pas les agents maison	Pour chaque agent
Agent Identity	Une identité par agent, sans clé, révocable	Avec Agent Gateway
Databricks via MCP	Pas de connecteur natif ; droits Unity Catalog	Selon les cas d'usage
VPC Service Controls	Bloque l'exfiltration, même avec des identifiants volés	Observation, puis actif
SIEM / SOC	Intégrer l'IA à la supervision existante	Selon l'outillage



Point clé : le filtrage Model Armor de Gemini Enterprise ne s'applique pas aux agents maison.

Agent Gateway est prêt dans le socle : une variable suffit, le jour où le premier agent qui agit arrive.

Console pour apprendre, Terraform pour garantir

Élément	Console (N0)	Terraform (N1 à N3)	Reste manuel
App, assistant, recherche web, langue	Oui	Oui	—
Identité, rôles, groupes	Oui	Oui	Appartenance aux groupes
Model Armor, plancher, DLP France	Oui	Oui	—
Connecteurs (Workspace, Microsoft, GitHub, Jira...)	Oui	Oui, modèle générique	OAuth initial, IP côté SaaS
Secrets des connecteurs	—	Référencés, jamais leur valeur	Création par la sécurité
CMEK	Possible	Oui, protégé contre la suppression	Révocation d'urgence
Conservation, observabilité	Oui	Par l'API, depuis Terraform	—
Journaux dédiés, exclusions, alertes	Oui	Oui	—
Licences	Oui	Possible (facturé)	Attribution en console
Agent Gateway, VPC-SC	Possible	Oui (options)	Règles d'entrée VPC-SC

Model Armor : le pare-feu des échanges avec l'IA



Filtre	Ce qu'il détecte	Réglage
Injection et jailbreak	« Ignore tes instructions... », consignes cachées	Seuil moyen
Liens malveillants	Phishing et malware connus	Activé
IA responsable	Haine, harcèlement, sexuel, dangereux	Seuil moyen
Données sensibles	NIR, CNI, passeport, IBAN, carte, secrets	DLP France

- FAIL_CLOSED**

Model Armor indisponible : l'assistant refuse
- Plancher**

Aucun filtre moins strict que le minimum
- Séparation + alerte**

Seule la sécurité modifie ; tout est signalé

Choix assumé : seuil moyen. Un outil qui bloque à tort est contourné ; plancher et détection couvrent le reste.

Journalisation : tout tracer, sans tout exposer

Journal	Contenu	Stockage UE
Audit d'administration	Qui a modifié filtres, clés, droits, connecteurs	Sécurité · 365 j
Audit d'accès aux données	Appels aux API, lectures de secrets	Sécurité · 365 j
Usage Gemini Enterprise	Qui a demandé quoi, sources, blocages	Usage · 90 j
Texte intégral des échanges	Questions et réponses complètes	Désactivé (DPO)
Contenu filtré Model Armor	Texte bloqué, filtre déclenché	Désactivé (Google)
Statistiques d'usage	Actifs, requêtes, retours	Console Analytics



Détecter sans surveiller

Questions exclues du journal par défaut, lisibles par la sécurité seule. Les blocages sont comptés et alertés sans exposer le contenu.

RGPD : finalité écrite, information des salariés, registre, CSE si contrôle d'activité.

Onze risques, et le niveau qui les traite

#	Risque	Exemple	Mesures	Niveau
R1	Accès non autorisé	Ex-collaborateur, compte partagé	SSO, groupes	0 → 1
R2	Divulgence de données	Un stagiaire retrouve des salaires	Connecteurs fédérés, DLP	1
R3	Injection de prompt	Consigne cachée dans un document	Model Armor	0 → 1
R4	Données personnelles	Un IBAN client collé dans une question	DLP France	1
R5	Contenus malveillants	Lien de phishing relayé	Model Armor	0 → 1
R6	Sécurité affaiblie	Un filtre coupé « pour une démo »	Rôles, plancher, alerte	1
R7	Perte de maîtrise	Données hors UE, non révocables	UE, CMEK	0 → 2
R8	Secrets exposés	Mot de passe dans un dépôt Git	Secret Manager	1
R9	Exfiltration	Export massif, identifiants volés	IP fixes, VPC-SC	1 → 3
R10	Abus non détecté	200 jailbreaks ignorés	Journaux, alertes	1 → 2
R11	Agent incontrôlé	Un agent appelle une API sensible	Agent Gateway	3

Qui fait quoi

Activité	Sécurité	Admin Gemini	DevOps	DPO	Métier
Filtres Model Armor, plancher, clés	R / A	I	C	I	—
Connecteurs et sources	C	R	C	C	A
Code Terraform, CI/CD	C (revue)	C	R / A	—	—
Conservation, journalisation du contenu	C	I	I	A	C
Ouverture à de nouveaux groupes	C	R	I	I	A
Enquête après alerte	R / A	C	C	I	—

R = réalise • A = approuve • C = consulté • I = informé

Le point clé : la sécurité approuve les garde-fous, le métier approuve l'ouverture, le DPO approuve la conservation. Personne n'est juge et partie.

Douze semaines jusqu'à la production



Durées indicatives : le chemin critique est rarement technique. Ce sont les décisions (identité, DPO) et les autorisations des sources (Microsoft, GitHub, Atlassian) qu'il faut lancer dès le niveau 0.

Ce qui reste à valider sur le projet réel

Point à confirmer en recette	Repli si besoin
Localisation du modèle DLP associé à Model Armor en UE	Autre localisation DLP, ou filtre de base temporaire
Paramètres exacts de certains connecteurs (Teams, GitHub)	Création une fois en console, puis reprise en code
Motif de blocage visible dans les journaux d'usage	Journalisation Model Armor temporaire en recette
Opérations surveillées par l'alerte « garde-fou modifié »	Ajustement du filtre après le premier changement
Exclusion effective des questions du journal par défaut	Vérification dans l'explorateur de journaux

Le socle Terraform est validé syntaxiquement et construit sur la documentation Google d'octobre 2026. Il sera confirmé lors de la recette du niveau 1 : chaque point a sa solution de repli.

Pourquoi cette approche est la bonne

- 1 **Démarrer demain** en console, sans exposer une seule donnée sensible
- 2 **Décider au bon moment** les choix irréversibles après avoir appris
- 3 **Sécurité vérifiable** un risque, une mesure, un test, une trace
- 4 **Protéger l'adoption** des réglages mesurés, que personne ne contourne
- 5 **Préparer la suite** agents, Databricks et périmètre sans refaire le socle

Prochaine étape

Atelier de deux heures avec sécurité, architecture et DPO pour acter les quatre décisions, puis lancement du niveau 0 dans un projet de découverte dédié.