

An abstract graphic on the left side of the slide, featuring several overlapping, wavy, translucent blue shapes that create a sense of depth and movement. The colors range from a deep navy blue to a lighter, almost white blue at the edges.

Projet M2 – Infrastructure as Code & IA

Déploiement d'une plateforme
d'extraction de texte d'images sur Azure
avec Terraform

Objectif pédagogique

Concevoir, déployer et sécuriser **une plateforme cloud complète d'extraction automatique de texte à partir d'images**, en utilisant :

Terraform

Infrastructure as Code

Microsoft Azure

Plateforme cloud cible

Azure AI Vision

Computer Vision / OCR

DevOps

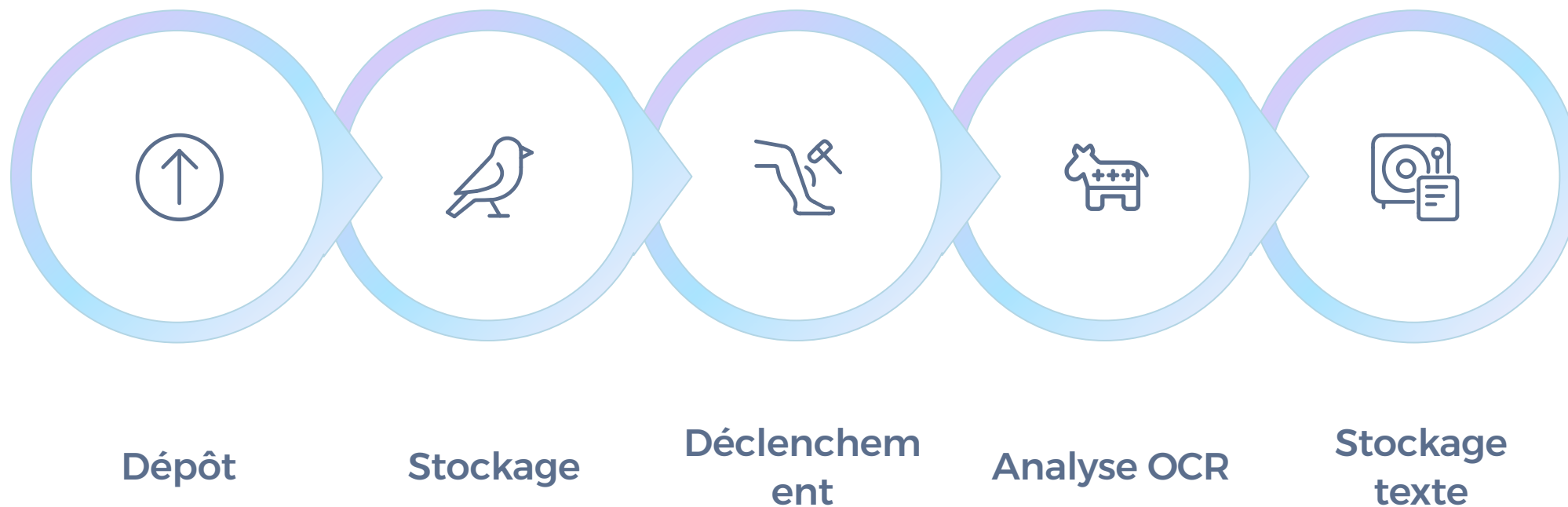
Sécurité, modularité, CI/CD, monitoring

Le projet met l'accent sur :

- **Architecture cloud**
- **Industrialisation avec Terraform**
- **Sécurisation d'API IA**
- **Observabilité**
- **Automatisation**

Description fonctionnelle

L'application devra :



Ce flux fonctionnel couvre l'ensemble du cycle de vie d'une image, depuis son dépôt jusqu'à la restitution du texte extrait via une API exposée.



Services Azure à déployer

Les étudiants devront provisionner via Terraform :



Microsoft Azure

Plateforme cloud principale pour l'ensemble de l'infrastructure



Azure Blob Storage

Stockage des images et des résultats d'extraction



Azure Key Vault

Gestion sécurisée des secrets et des clés d'accès



Azure AI Vision

Service d'intelligence artificielle pour l'extraction OCR



Azure Functions

Traitement serverless déclenché à l'upload d'image



Azure Application Insights

Observabilité, logs et monitoring de la plateforme



Travail demandé



Partie 1 – Infrastructure as Code

Les étudiants doivent :

1. Structurer le projet Terraform en modules

- network
- storage
- cognitive_service
- compute
- monitoring

2. Implémenter

- Resource Group
- Virtual Network + Subnet
- Storage Account
- Container Blob
- Cognitive Service (Vision)
- Function App
- Key Vault

3. Gérer les variables et outputs

4. Utiliser

- Backend distant (Azure Storage)
- State locking

5. Implémenter

- Naming convention
- Tags obligatoires
- Environnements (dev / prod)



Partie 2 – Intégration OCR

La Function devra :

01

Déclenchement à l'upload Blob

La Function est automatiquement activée lors du dépôt d'une image dans le container Blob.

02

Appel Azure AI Vision

La Function interroge le service Azure AI Vision pour lancer l'analyse OCR de l'image.

03

Extraction du texte

Le texte contenu dans l'image est extrait et structuré par le service cognitif.

04

Sauvegarde du résultat

Le résultat est sauvegardé dans : soit un autre container, soit un fichier JSON, soit une table storage.





Partie 3 – Sécurité

Les étudiants doivent :

Interdire les accès publics inutiles

Limiter l'exposition des ressources au strict nécessaire pour réduire la surface d'attaque.

Utiliser Managed Identity

Authentification sans secret explicite entre les services Azure via des identités managées.

Stocker les secrets dans Key Vault

Centraliser et sécuriser tous les secrets, clés et certificats dans Azure Key Vault.

Restreindre l'accès réseau

Mettre en place des règles réseau strictes pour isoler les ressources sensibles.

Implémenter RBAC minimal

Appliquer le principe du moindre privilège via le contrôle d'accès basé sur les rôles.



Partie 4 – Observabilité



Intégration Application Insights

Connexion de la plateforme au service de monitoring Azure Application Insights.



Logs d'analyse OCR

Enregistrement des événements et résultats de chaque analyse OCR effectuée.



Monitoring des erreurs

Détection et remontée automatique des erreurs survenant dans le pipeline de traitement.




Dashboard de supervision

Tableau de bord centralisé pour visualiser l'état et les performances de la plateforme.



Partie 5 – Industrialisation

 Bonus DevOps



Pipeline CI/CD

GitHub Actions ou Azure DevOps



terraform validate

Validation automatique du code Terraform



terraform plan automatique

Génération automatique du plan d'exécution Terraform



Déploiement conditionnel

Déclenchement du déploiement selon des conditions définies



Tests automatisés

Exécution de tests automatisés dans le pipeline CI/CD



Livrables attendus

1

Repository Git propre

Un dépôt Git organisé, versionné et documenté.

2

Code Terraform modulaire

Infrastructure as Code structurée en modules réutilisables.

3

Script de déploiement

Script permettant de déployer l'ensemble de la plateforme de manière automatisée.

4

README technique

Documentation incluant : Architecture, Diagramme, Sécurité, Coût estimé.

5

Démonstration fonctionnelle

Présentation en conditions réelles du fonctionnement complet de la plateforme.