



Cartofriches

Parcours Data Product Manager
Promotion Décembre – Mars 2024

Rédacteurs: [Omar CHOA](#) | [Samy SELLAM](#) | [Siham HOUCHE](#) | [Michael DEROUCHE](#)



Ordre du jour

1. Analyse du contexte
2. Discovery
3. Synthèse de la Discovery
4. Synthèse des améliorations
5. Benchmark
6. KPIs
7. Création du MVP
8. Lancement

1. Analyse du contexte


Business Model Canvas

Business Model Canvas				
Partenaires clés <ul style="list-style-type: none"> Ministère de la Transition écologique et de la Cohésion des territoires Cerema Collectivités territoriales Organismes et établissements publics, scientifiques et techniques (ADEME, EPCI...) Universités Associations d'aménageurs Bureaux d'études Entreprise Tech : fournisseurs de technologie SIG, IA et Cloud 	Activités Clés <ul style="list-style-type: none"> Collecte et mise à jour des données sur les friches Développement et maintenance de la plateforme cartographique Analyse prédictive et évaluation du potentiel de réhabilitation. Intégration et traitement de données SIG pour l'analyse spatiale. Animation de la communauté Cartofriches Organisation d'événements et de formations Communication et sensibilisation 	Propositions de valeur <ul style="list-style-type: none"> Accès à une plateforme cartographique interactive et gratuite répertoriant 10 952 sites friches en France (maj 11/03/24) Informations détaillées sur chaque site : localisation, surface, historique, potentialités, etc. Visualisation de l'évolution des friches dans le temps et de leur contexte géographique pour une planification efficace. Contextualisation autour des friches incluant l'existant autour. Outils d'aide à la décision pour les projets de reconversion des friches Mise en relation des acteurs du marché Sensibilisation du public à la problématique des friches Contribution au développement durable 	Relation Client <ul style="list-style-type: none"> Support technique Assistance à la recherche d'informations Mise en relation avec des experts Organisation d'ateliers et de formations Mise à disposition des guides d'utilisation Appui auprès des services déconcentrés de l'État Associations d'élus 	Segments de Clientèle <ul style="list-style-type: none"> Collectivités territoriales Aménageurs Promoteurs immobiliers Investisseurs Bureaux d'études Entreprises privées Grand public Associations environnementales Élus
Coûts <ul style="list-style-type: none"> Développement et maintenance de la plateforme Collecte et traitement des données Salaires des équipes Communication et marketing Frais de fonctionnement (hébergement, cloud, licences) Recherche et développement 		Revenus <ul style="list-style-type: none"> Subventions publiques Vente de données Analyses personnalisées Prestations de conseil Abonnements à la plateforme Sponsoring 		
Ressources clés <ul style="list-style-type: none"> Plateforme cartographique Base de données Équipe d'experts Financements Équipe développeurs en IA et SIG Serveurs et Infrastructures numériques Base de données urbaines 		Canaux de distribution <ul style="list-style-type: none"> Plateforme web Cartofriches Applications mobiles Webinaires, ateliers Documentation, rapports Réseaux sociaux Newsletters Presse 		

2. Discovery

Personas


Arthur Forestier | Chef de projets, Nantes Métropole

	Âge	Rattachement	Localisation
	32 ans	Direction aménagement urbanisme Nantes, Département urbanisme et habitat, Direction générale déléguée à la fabrique de la ville écologique et solidaire, Nantes Métropole	Nantes

Problématique

Arthur est en charge de l'accompagnement des porteurs de projets souhaitant reconverter des friches dans le métropole de Nantes. Il aide notamment les porteurs à trouver la friche qui correspond le mieux à leur projet. Il utilise régulièrement Cartofriches et aimerait voir de nouvelles fonctionnalités pour faciliter son travail de *matching*.

Jade Guillaume | Chargée d'études foncières, Cerema

	Âge	Rattachement	Localisation
	28 ans	Groupe Ingénierie de la donnée et innovation, Département Territoires Ville et Bâtiment, Direction territoriale Méditerranée, Cerema	Aix-en-Provence

Problématique

Membre de l'équipe Cartofriches, l'une des tâches principales de Jade consiste en la vérification visuelle, *via* photo aérienne ou Google Street View, des sites déclarés comme friches afin de mettre à jour les données BASOL et BASIAS qui sont intégrées dans Cartofriches. Elle aimerait automatiser cette tâche répétitive.

2. Discovery

Experience Map

Jade Guillaume | Chargée d'études foncières, Cerema

Action	1 Sélection d'une friche potentielle	2 Contrôle visuel numérique	3 Mise à jour de la fiche friche
Goal	Identifier une friche potentielle pour laquelle des données permettant de déterminer son statut réel sont disponibles	Déterminer le statut réel de la friche : sans projet, avec projet, reconvertie	Remplacer les données caduques de la fiche correspondante par des données à jour
Touchpoints & technologies	Portail administrateur de Cartofriches	Photos aériennes, Google Street View, Mapillary	Portail administrateur de Cartofriches
Emotions	 →  → 		
Pain points	Absence imprévue de données nécessaires pour aller au bout du processus de vérification	Données cartographiques non concluantes, caduques ou manquantes	Tâche répétitive et à faible valeur ajoutée

2. Discovery

Analyse Quantitative

Contrôle visuel numérique

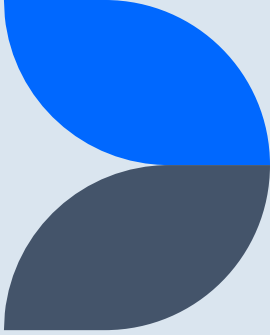
- **Question** : Quelle proportion de friches ont des données cartographiques non concluantes, caduques ou manquantes lors du contrôle visuel numérique ?
- **Données nécessaires** : Résultats des contrôles visuels, statut des données cartographiques (à jour, caduques, manquantes).
- **Analyse** : Evaluation de la proportion de friches avec des données cartographiques inadéquates lors du contrôle visuel.
- **Métrique** : Pourcentage de friches avec données cartographiques problématiques.
- **Calcul** : Proportion de friches avec données problématiques = $\frac{\text{Nombre de friches avec données cartographiques non concluantes, caduques ou manquantes}}{\text{Nombre total de friches soumises au contrôle visuel numérique}} \times 100$.

Mise à jour de la fiche friche

- **Question** : Combien de temps est consacré en moyenne à la mise à jour des fiches friches ?
- **Analyse** : Temps consacré à la mise à jour des informations d'une friche.
- **Données nécessaires** : Temps passé pour chaque mise à jour, nombre de mises à jour effectuées.
- **Métrique** : Durée moyenne pour la mise à jour d'une fiche.
- **Calcul** : Durée moyenne de mise à jour d'une friche = $\frac{\text{Temps total passé pour la mise à jour des friches}}{\text{Nombre total de friches mises à jour}}$

3. Synthèse de la Discovery

-



Problème 1

Automatisation de la
vérification des friches :
Processus manuel,
chronophage.

Problème 2

Prédiction de la **viabilité**
des projets sur les Fiches:
Difficulté à évaluer le
potentiel de réhabilitation.

Problème 3

Intégration de SIG et
Développement d'Analyses
Avancées: Manque d'outils
analytiques pour
l'optimisation des décisions.

4. Synthèse des améliorations



Amélioration 1

Automatiser via computer vision: Accélère la vérification des friches.

Amélioration 2

Outils Prédictifs: Evalue la viabilité des projets, réduit le taux d'échec.

Amélioration 3

Intégration SIG et Analyses: Fournit une analyse détaillée pour une meilleure planification.

5. Benchmark

-

Amélioration 1 | Automatiser la vérification des friches

<i>Des solutions métier mises en place par des entreprises concurrentes pour améliorer une situation similaire</i>	Détection à l'aide de drones des friches agricoles présentes dans le Loiret (lien , p. 19)
<i>Des produits digitaux existants basés ou non sur de l'IA</i>	SCO Friches Agricoles : pré-repérage automatisé des friches agricoles en Occitanie par imagerie satellitaire et IA (logiciel WaSaBI CNES + application mobile Vigifriche) (lien)

Amélioration 2 | Prédire la viabilité des projets

<i>Des solutions métier mises en place par des entreprises concurrentes pour améliorer une situation similaire</i>	Études et diagnostics sous maîtrise d'ouvrage direct de l'EPF Alsace (lien) Offre d'accompagnement de l'ANCT à destination des collectivités et porteurs de projets (lien)
<i>Des produits digitaux existants basés ou non sur de l'IA</i>	Bénéfriches (lien) Outil d'évaluation du potentiel de mutabilité des friches industrielles dans les Ardennes (lien)

Amélioration 3 | Intégrer un SIG et de l'analytique

<i>Des solutions métier mises en place par des entreprises concurrentes pour améliorer une situation similaire</i>	Outils Sogefi : Mon Territoire Carto (lien), HORIZON (lien)
<i>Des produits digitaux existants basés ou non sur de l'IA</i>	Decision Support Systems (DSS) : SADA , DESYRE , etc.

Amélioration retenue

Automatisation via Computer Vision

- Réduction du temps de vérification des friches de heures à minutes.
- Diminution significative des coûts opérationnels.
- Augmentation de la précision des données sur les friches.
- Facilitation de la réutilisation et de la réhabilitation des terrains en friche.

6. KPIs

-

Temps Moyen de Vérification d'une Friche

- **Description :** Ce KPI mesure la durée nécessaire pour vérifier le statut et les caractéristiques d'une friche, depuis l'identification initiale jusqu'à la classification finale dans la base de données. Une réduction significative de ce temps moyen indiquerait une efficacité accrue due à l'automatisation.
- **Données nécessaires :**
 - Date et heure de début de la vérification pour chaque friche.
 - Date et heure de fin de la vérification pour chaque friche.
- **Formule :** Temps moyen de vérification = Temps total passé sur la vérification de toutes les friches / Nombre total de friches vérifiées

Taux d'Erreur dans la Classification des Friches

- **Description :** Ce KPI mesure le pourcentage d'erreurs commises dans la classification des terrains comme friches. Il permet d'évaluer la précision des analyses réalisées par le système d'automatisation. Une diminution du taux d'erreur reflète l'amélioration de la précision et de la fiabilité des informations fournies par Cartofriches grâce à la vision par ordinateur.
- **Données nécessaires :**
 - Nombre de friches correctement classifiées (confirmation après vérification manuelle ou autre méthode de validation).
 - Nombre de friches incorrectement classifiées (déterminé après vérification manuelle ou autre méthode de validation).
- **Formule :** Taux d'erreur = (Nombre de friches mal classifiées / Nombre total de friches classifiées) x100



Type de Produit Data

Machine Learning, en utilisant des techniques de computer vision pour analyser les images des friches.



7. Création du MVP

a. Contenu MVP

Fonctionnalités à réaliser plus tard











- Classification avancée des types de friches

Fonctionnalités à écarter

- Analyse détaillée de la végétation environnante (exemple)
- Interface utilisateur avancée pour les utilisateurs finaux

7. Création du MVP

b. Machine Learning Canvas

Decisions  <p>How are predictions used to make decisions that provide the proposed value to the end-user?</p> <p>Les prédictions permettront d'accélérer l'étape de filtrage manuel effectuée par le Cerema avant de décider si un site non « vérifié » peut être basculé en site « vérifié ».</p> <p>Ainsi, les end users auront plus rapidement accès à un pool plus large de sites potentiellement utilisables pour la réalisation de leur projet.</p> <p>Cela contribuera à faire avancer plus vite l'objectif principal de Cartofriches, à savoir, faciliter la reconversion des friches.</p>	ML task  <p>Input, output to predict, type of problem.</p> <ul style="list-style-type: none"> - Entrée (X) : <ul style="list-style-type: none"> Données BDD Cartofriches Images Google Street View - Sortie (y) : friche / non friche - Type de problème : classification - Question spécifique que le modèle cherche à résoudre : prédire, à partir des caractéristiques répertoriées d'un site et des images correspondantes, si celui-ci est bien une friche ou non. 	Value Propositions  <p>What are we trying to do for the end-user(s) of the predictive system?</p> <p>Pour le Cerema :</p> <ul style="list-style-type: none"> - Fluidification du pipeline de traitement des données proposées - Gain de temps <p>Pour les end users :</p> <ul style="list-style-type: none"> - Augmentation sensible du nombre de sites dits « vérifiés » (actuellement surpassé par celui des sites non « vérifiés ») - Ce qui se traduit par des choix plus nombreux et plus fiables pour la réalisation de leur projet <p>What objectives are we serving?</p> <p>Les objectifs de Cartofriches au service de l'intérêt général, à savoir :</p> <ul style="list-style-type: none"> - Lutter contre l'artificialisation des sols - Sécuriser les sites sur les plans écologiques et sanitaires - Encourager l'ambition et le dynamisme économiques, sociaux et environnementaux des territoires 	Data Sources  <p>Which raw data sources can we use (internal and external)?</p> <p>BDD Cartofriches. Sa première couche est constituée de données provenant de BASOL et BASIAS, 2 BDD nationales. Les sites concernés sont classés en 2 catégories :</p> <ol style="list-style-type: none"> « Vérifiés » : ont fait l'objet d'un contrôle visuel numérique Non « vérifiés » : données non contrôlées et potentiellement obsolètes <p>Google Street View. Les images sont accessibles via le service Street View fourni par l'API Maps JavaScript.</p>	Collecting Data  <p>How do we get new data to learn from (inputs and outputs)?</p> <p>Le dataset de base sera constitué des sites « vérifiés ». Pour la première itération du modèle (= train + test), selon la quantité et la qualité des données, il peut être judicieux de n'en utiliser qu'une partie (p. ex., 50%), et d'ajouter 10% à chaque nouvelle itération en guise de nouveau jeu de données.</p> <p>Par la suite, une partie des sites non « vérifiés » pourra être utilisée sous réserve de la disponibilité de ressources (humaines) pour les labelliser.</p>
Making Predictions  <p>When do we make predictions on new inputs?</p> <p>La BDD Cartofriches est mise à jour à fréquence trimestrielle. Les prédictions pourront être effectuées à ce même rythme sur les nouvelles données ajoutées au cours du trimestre écoulé.</p> <p>Comme toutes les autres opérations ML du produit, les prédictions seront effectuées par Google Vertex AI sur Google Cloud, affranchissant ainsi le Cerema des contraintes matérielles associées en échange d'une tarification élastique.</p> <p>How long do we have to featurize a new input and make a prediction?</p> <p>Le feature engineering sera effectué pendant les 3 premiers mois de chaque trimestre, le dernier mois étant dédié à la réalisation des prédictions.</p>	Offline Evaluation  <p>Methods and metrics to evaluate the system before deployment.</p> <ul style="list-style-type: none"> - Temps de prédiction - Seuil de score - Précision vs rappel : préférer une précision élevée (les sites prédits comme étant des friches le sont réellement, quitte à classifier d'autres friches réelles en tant que non friches) - Précision moyenne : calculé sur l'ensemble des seuils de score 	Live Evaluation and Monitoring  <p>Methods and metrics to evaluate the system after deployment, and to quantify value creation.</p> <ul style="list-style-type: none"> - Temps de prédiction - Seuil de score - Précision - Statistiques de trafic - Taux de reconversion des friches - Taux de satisfaction utilisateur 	Features  <p>Input representations extracted from raw data sources.</p> <p>Les features de base seront celles qui figurent dans le fichier friches-standard.csv, disponible sur data.gouv.fr. Comme elles sont très nombreuses (50), il faudra identifier et exclure celles qui ont un faible impact sur / nuisent aux prédictions.</p> <p>Ces caractéristiques seront enrichies par les images de Google Street View qui, une fois transformées en matrices et traitées, deviennent des features complémentaires.</p>	Building Models  <p>When do we create/update models with new training data?</p> <p>Par souci de simplicité, nous choisissons la solution AutoML de Vertex AI, une plateforme Google. Cet outil user-friendly et prêt-à-l'emploi prend en charge le lifecycle ML de bout en bout, facilitant ainsi la création du MVP, puis la prise en main par le Cerema. D'autres avantages potentiels incluent une intégration simplifiée avec Street View et des tarifs plus avantageux.</p> <p>Le modèle pourra être ré-entraîné à fréquence trimestrielle, en accord avec le rythme de mise à jour de la BDD Cartofriches, et en fonction des métriques de monitoring.</p> <p>How long do we have to featurize training inputs and create a model?</p> <p>Nous prévoyons une période de 4 mois (1 trimestre) pour développer la première itération du modèle.</p>

7. Création du MVP

c. Challenger son MVP

Risques liés aux données

- Qualité des données
- Complétude des données

Réglementations

- RGPD

Réglementations

- Biais algorithmique
- Transparence et interprétabilité
 - Non-discrimination
- Conséquences sociales
- Confidentialité et protection des données

8. Lancement

-

Dans le but de déployer efficacement notre solution de data utilisant la vision par ordinateur, nous intégrons l'API statique Street View de Google ainsi que Google Earth, et employons les outils Vertex AI et AutoML pour automatiser la vérification des friches. Cette stratégie de lancement est conçue pour répondre aux besoins spécifiques des professionnels travaillant chez Cerema ou au sein d'entités partenaires :

1. Préparation Technique et Conformité

2. Développement de Supports

3. Stratégie de Communication

4. Formation et Sensibilisation

5. Support et Feedback

Merci