



DATATHONX

3 E M E E D I T I O N

14 ET 15 FEVRIER

DSCC

INTRODUCTION

La compétition DatathonX est un événement d'envergure conçu pour réunir des passionnés de la science des données, du machine learning (ML), du deep learning (DL) et de l'intelligence artificielle (IA). Cette édition se concentre sur trois domaines clés : la santé, l'environnement et l'agriculture. L'objectif principal est de stimuler la créativité et la collaboration entre des profils variés (étudiants, chercheurs, professionnels) afin de développer des solutions innovantes répondant à des problématiques réelles.

DATATHONX

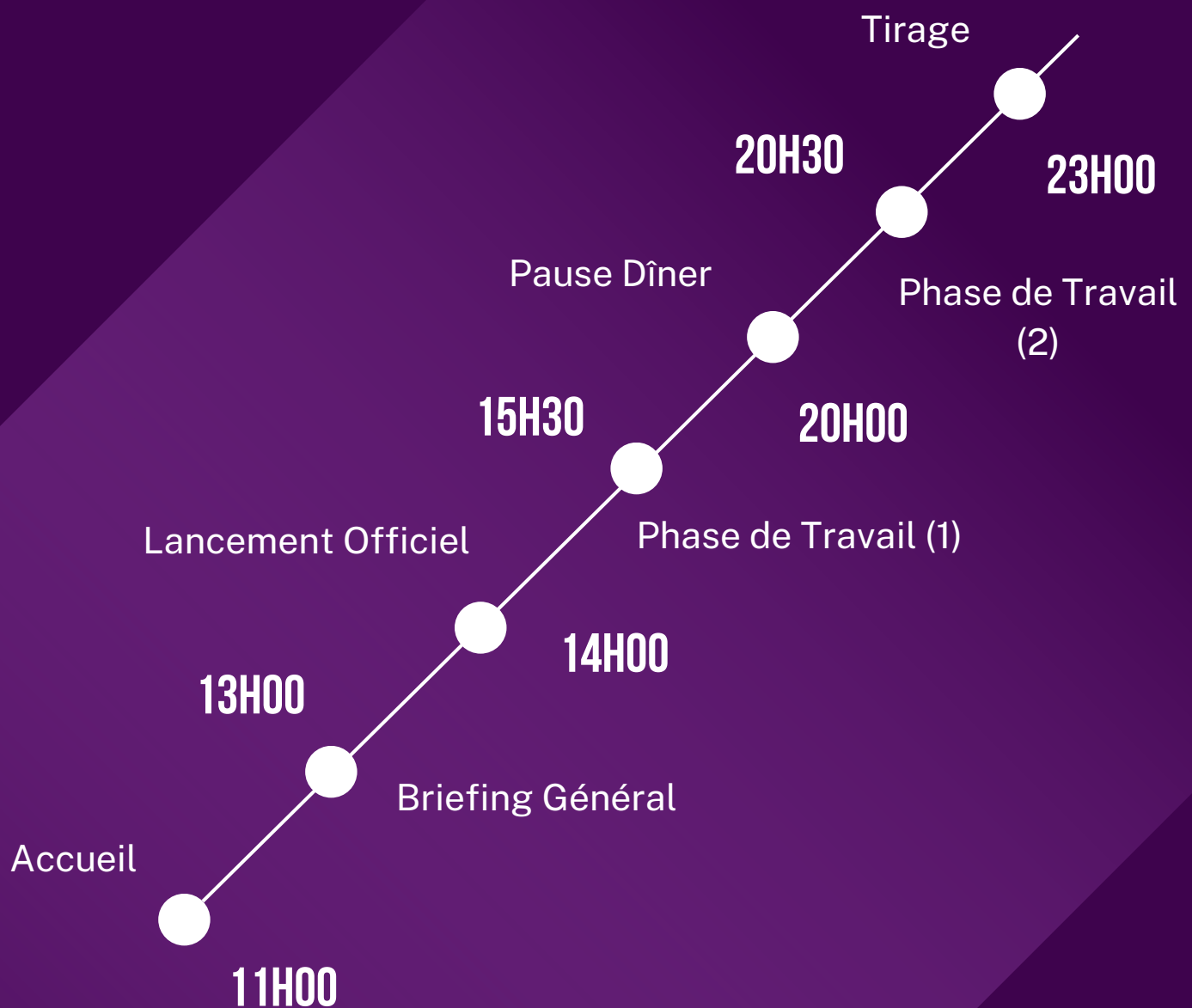
PUBLIC VISÉ

Cette compétition s'adresse aussi bien aux étudiants en cursus scientifiques, ingénieurs, développeurs, data scientists, qu'aux professionnels et chercheurs issus du secteur privé ou public. L'hétérogénéité des équipes permet la complémentarité des compétences et une plus grande diversité de points de vue.

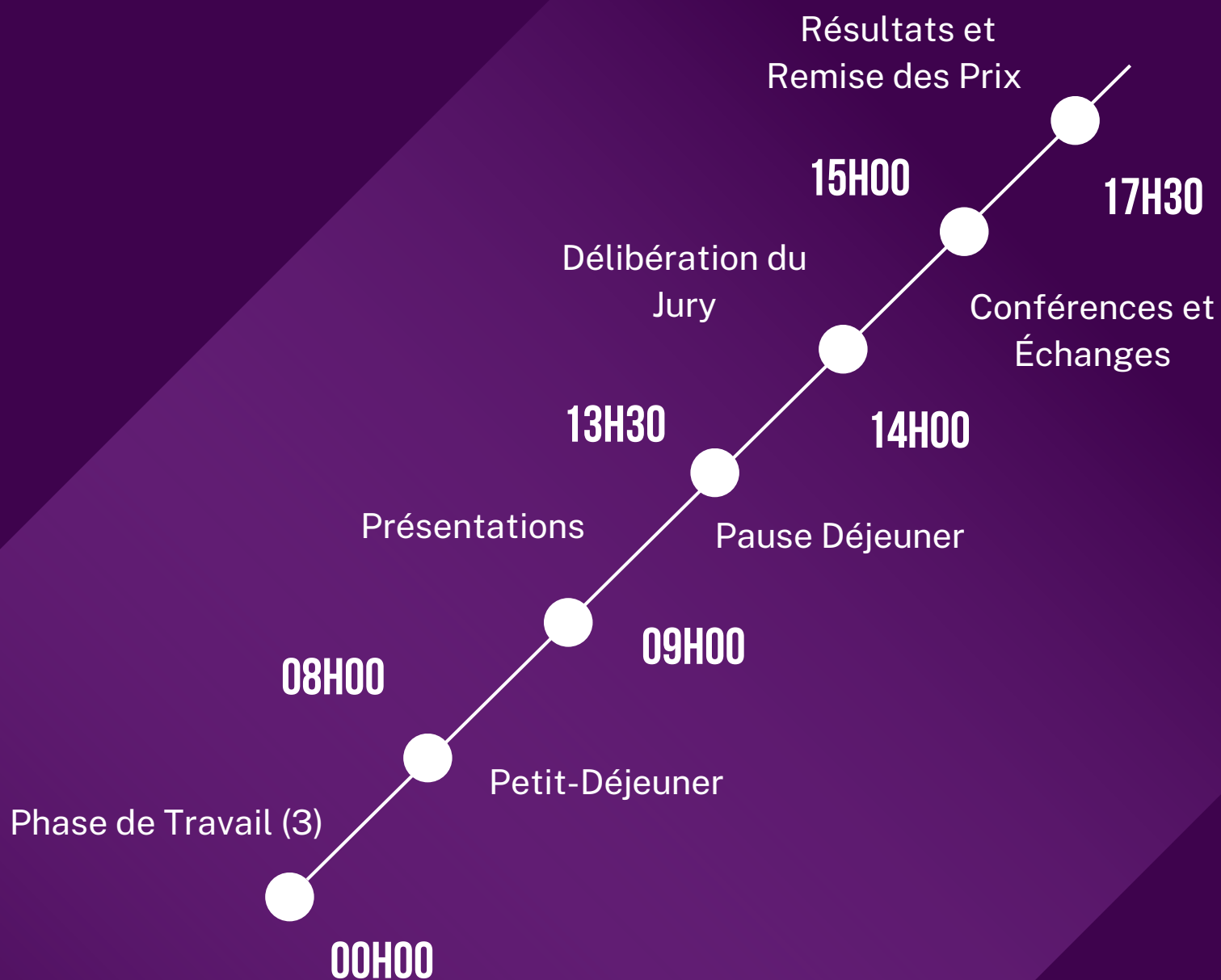
PRÉSENTATION DE CE DOCUMENT

Ce document constitue la feuille de route pour tous les acteurs du DatathonX. Il détaille les objectifs de la compétition, le cadre dans lequel s'inscrivent les projets, les techniques pouvant être mobilisées, les critères d'évaluation, ainsi que le déroulement planifié de l'événement.

PLANNING



PLANNING



OBJECTIFS

PROMOUVOIR

L'EXCELLENCE EN IA

Mettre en avant l'utilisation de techniques avancées (ML, DL, IA) pour relever des défis complexes et encourager les bonnes pratiques dans le domaine.

RÉSoudre DES PROBLÉMATIQUES RÉELLES

Les projets doivent s'inscrire dans une logique d'impact concret, que ce soit dans la santé (diagnostic, suivi, prédiction, etc), l'environnement (analyse de l'empreinte carbone, suivi de la biodiversité, etc) ou l'agriculture (optimisation des cultures, analyse de sols, etc.).

STIMULER LA COLLABORATION

Favoriser l'interdisciplinarité et la mise en commun des compétences. Les équipes sont composées de profils variés (techniques, scientifiques, management, design, etc.).

RENFORCER LA COMMUNAUTÉ DATA

Créer un réseau solide d'experts, de mentors, de sponsors et de partenaires autour de thématiques data-driven.

APPROCHES TECHNIQUES

- Choix d'algorithmes adaptés (ML, DL, IA) pour classification, régression, clustering, etc.
- Apprentissage supervisé et non supervisé pour une analyse optimisée.
- Évaluation et amélioration des performances des modèles.

PRÉPARATION DES DONNÉES

- Collecte, nettoyage et transformation des données.
- Gestion des valeurs manquantes, normalisation et réduction de dimension.
- Structures de stockage pour une organisation et accessibilité optimales.

VISUALISATION DES RÉSULTATS

- Formats visuels clairs et pertinents.
- Outils pour une meilleure interprétation et communication des analyses.
- Livrables synthétiques et faciles à comprendre.

STRUCTURE ATTENDUE

Pour être complet et pertinent, chaque projet doit inclure une **problématique** détaillant le contexte et les enjeux, des **objectifs** clairs avec des métriques de succès, une **méthodologie** expliquant les étapes et justifiant les choix techniques, des **résultats et analyses** illustrés et interprétés, une évaluation de l'**impact** potentiel, et une **documentation** complète pour garantir la traçabilité et la reproductibilité.

LIVRABLES

Les équipes doivent fournir :

- **Rapport Final** : Un document structuré en PDF (ou équivalent).
- **Code Source** : Déposé sur un dépôt Git (GitHub, GitLab) ou sous forme d'archive, accompagné d'une documentation pour l'exécution.
- **Présentation Orale** : Des slides synthétiques (PowerPoint, Keynote ou PDF) mettant en avant les points clés du projet.
- **Démonstration** : Si possible, une démo fonctionnelle (interface ou script exécutable).

CRITÈRES D'ÉVALUATION DES PROJETS

Critère	Poids (%)	Description Détaillée
Pertinence de la problématique	20	Pertinence du sujet (santé, environnement, agriculture) et clarté des besoins et enjeux
Qualité technique	30	Algorithmes choisis et optimisés avec justification, prétraitement et validation rigoureux (cross-validation, échantillon de test), et pipeline automatisé et structuré
Innovation	20	Approche originale ou combinaison inédite de méthodes, avec prise de risques et expérimentation.
Impact et applicabilité	20	Solution pertinente en contexte réel, avec intégration facilitée (technologies disponibles, coûts, maintenabilité)
Présentation	10	Pitch clair et concis, supports visuels et démonstrations pertinents, et réponses efficaces aux questions du jury

CONCLUSION

Le DatathonX est à la fois un challenge technique et une opportunité de réseautage et d'apprentissage. Les participants repartiront avec de nouvelles compétences, des contacts dans la communauté data, et pour les meilleurs, des prix récompensant leur travail. En suivant ces directives, chaque équipe disposera des clés nécessaires pour mener un projet ambitieux, structuré et conforme aux attentes. Nous comptons sur l'enthousiasme et la créativité de tous pour faire de ce DatathonX un véritable succès !



DSCC