ACPR
BANQUE DE FRANCE

PRÉSENTATION DE L'EXPÉRIMENTATION

Amélioration des modèles de détection en LCB-FT par mutualisation confidentielle des données et calculs collaboratifs

PÔLE FINTECH-INNOVATION

fintech INNOVATION

LE LAB Banque de France

30 MARS 2022

INTRODUCTION

-

Dominique Laboureix

Secrétaire Général de l'ACPR

Ordre du jour

1. Objectifs de l'expérimentation
2. Méthode proposée
3. Principales étapes de l'expérimentation
4. Récapitulatif : qui fait quoi, quand ?
5. Questions et réponses

1. Objectifs de l'expérimentation

CONTEXTE

Constat :

- La LCB-FT représente une charge administrative importante ;
- Elle est donc un domaine privilégié des algorithmes de détection, notamment d'IA ;
- L'efficacité de ces algorithmes pourrait être accrue s'ils étaient entraînés à une échelle plus large que celle de l'établissement ou du groupe.

Objectif : étudier l'amélioration de la performance prédictive des algorithmes de détection de transactions bancaires suspectes.

Contraintes : de confidentialité liées à la sensibilité des données et aux exigences réglementaires et juridiques.

Moyen : technologies de mutualisation sécurisée ou de calculs collaboratifs sur données sensibles.

Format : expérimentation scientifique, combinant ateliers métier et étude technologique approfondie.

ACPR
BANQUE DE FRANCE

POURQUOI UNE EXPÉRIMENTATION ?

- Éclairer un sujet technique complexe :
  - à droit constant ;
  - sans préjuger des solutions possibles.
- Donc une approche :
  - axée sur l'innovation ;
  - technologiquement neutre (évaluation objective et quantitative des résultats) ;
  - collaborative ;
  - déconnectée des objectifs de contrôle de l'ACPR ;
  - respectueuse des données sensibles ;
  - visant un objectif de bien public (efficacité accrue des dispositifs LCB-FT conduisant à une réduction des activités illicites et criminelles ciblées).

ACPR
BANQUE DE FRANCE

BÉNÉFICES ATTENDUS DE L'EXPÉRIMENTATION

- Évaluer **l'état de l'art technologique et scientifique** permettant de répondre :
  - à la mise en œuvre de méthodes collaboratives respectant la confidentialité ;
  - à l'objectif de performance des algorithmes de surveillance des transactions.

- **Rétroagir** sur les réflexions et les développements technologiques et scientifiques en cours : les enrichir par une vision concrète des **enjeux métiers**.

- Faire **collaborer acteurs bancaires et prestataires techniques** ou regtechs, sous l'égide de l'ACPR.

- Contribuer, *in fine*, à **l'amélioration de la LCB-FT**.

ACPR
BANQUE DE FRANCE

2. Méthode proposée

DEUX PRÉREQUIS À L'EXPÉRIMENTATION

CONFIDENTIALITÉ

- MCD : Mutualisation Confidentielle des Données (calculs collaboratifs inclus)
- Vise à garantir tout au long de l'expérimentation que seuls les individus et entités accrédités pour voir une variable x d'une transaction y le pourront
- Les technologies recouvrent essentiellement le domaine des PET (*Privacy Enhancing Technologies*)
- L'expérimentation inclura un "Tech Sprint MCD" pour sélectionner et évaluer l'état de l'art en la matière, objectif annexe de l'exercice.

MODÉLISATION

- Modèles de détection des transactions suspectes
- Méthode à la main des banques volontaires (IA ou procédural)
- L'expérimentation inclura une série d'ateliers modélisation (des données, de leur mutualisation, du processus de détection mutualisé).

ACPR
BANQUE DE FRANCE

PARTIES PRENANTES ET LEURS RÔLES : L'ACPR

L'ACPR conçoit, facilite et encadre l'expérimentation pour:

- veiller à la confidentialité des données
- garantir une collaboration efficace entre pairs et avec l'assistance de prestataires technologiques
- valider les protocoles expérimentaux et assurer le caractère probatoire des résultats
  - en favorisant l'innovation pluridisciplinaire (LCB / MCD / IA)
  - en restant dans un cadre juridique et réglementaire adapté.

Pour ce faire, elle pourra faire appel aux ressources techniques et aux experts de la Banque de France et de l'ACPR dédiés à l'innovation.

PARTIES PRENANTES ET LEURS RÔLES : LES BANQUES

Les banques participantes assemblées en équipes décident des détails du protocole expérimental, en accord avec l'ACPR :

- Coéquipier(s)
- Données (périmètre, profondeur, largeur)
- Processus de benchmark (sur données solo)
- Mécanisme de mutualisation (nature, complexité, intérêt métier)
- Type de modèle de détection (sur données mutualisées)
- Procédure d'annotation (des données mutualisées d'apprentissage)
- Procédure de revue manuelle (des données mutualisées d'évaluation)
- Prestataire MCD
- Ressources (humaines / matérielles / financières)

PARTIES PRENANTES ET LEURS RÔLES : LES PRESTATAIRES

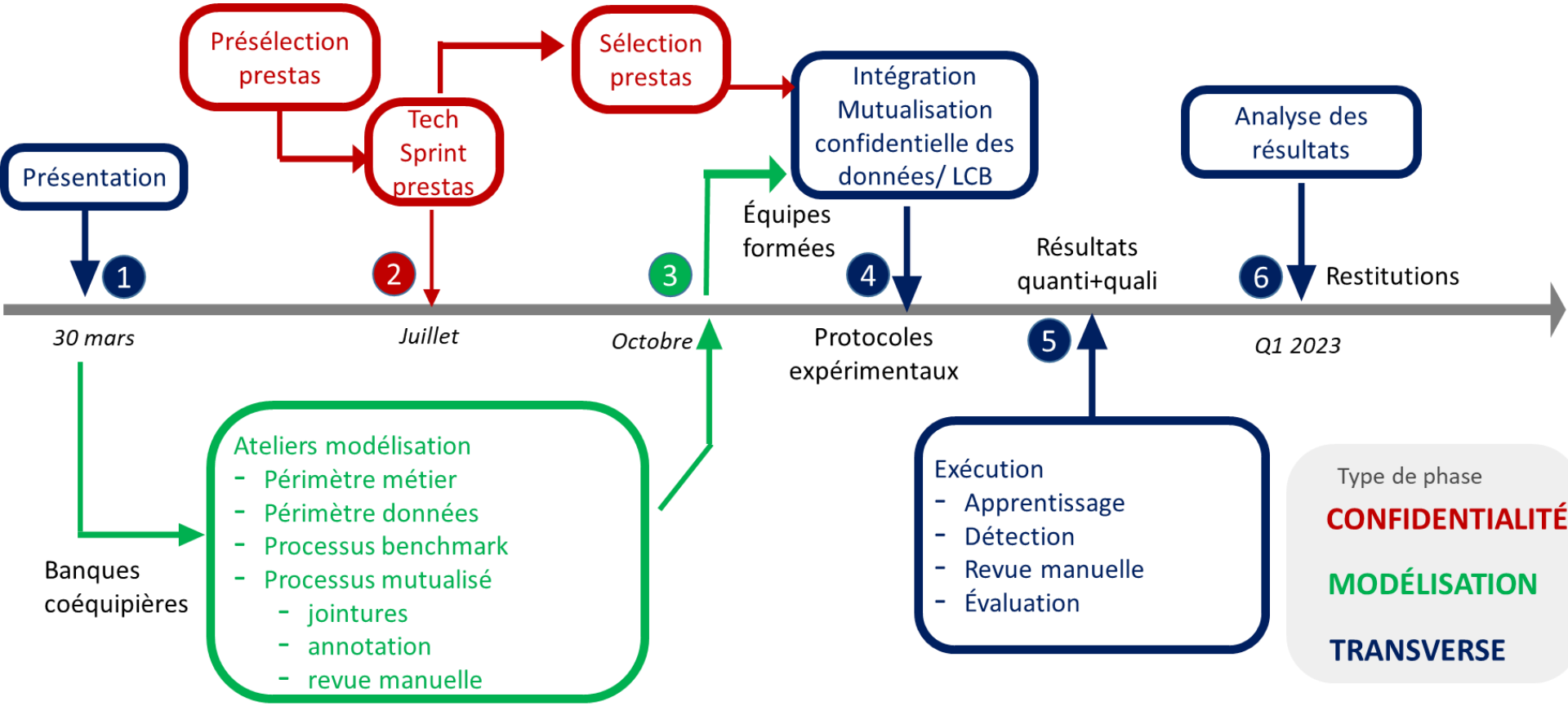
- Répondent à l'appel à candidatures s'ils peuvent proposer une solution (déjà commercialisée ou non)
  - pertinente aux plans fonctionnel et réglementaire ;
  - possible à intégrer dans le protocole expérimental en Q3 2022.
- Réalisent le PoC du Tech Sprint à titre gracieux.
- Positionnent leur prestation pour la réalisation du protocole
  - Solution générique ;
  - Solution adaptée au protocole d'une banque participante.

Avant d'entrer dans le détail des différentes étapes de l'expérimentation…

des questions sur les objectifs et la méthode proposée ?

3. Principales étapes de l’expérimentation

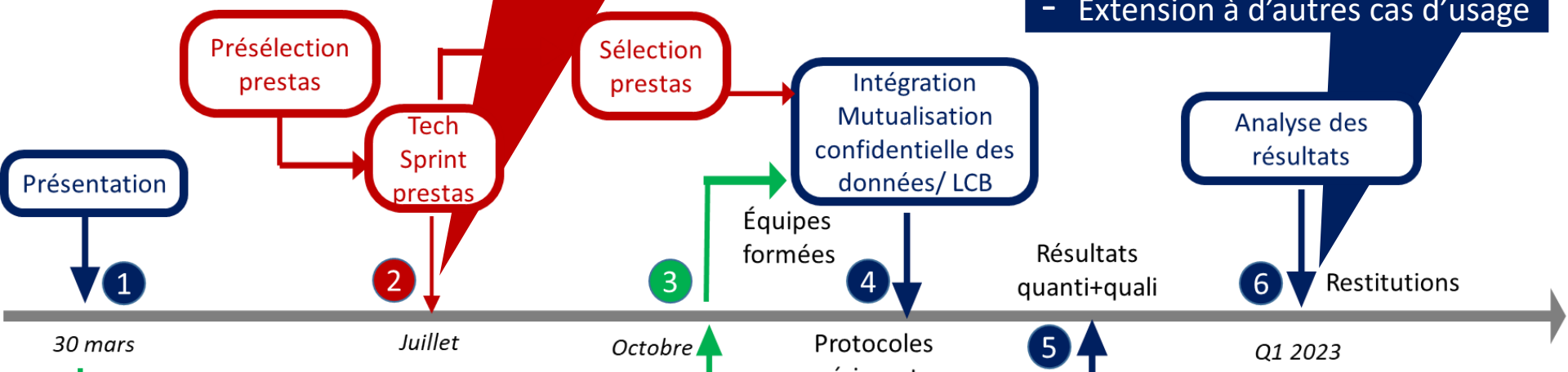
FEUILLE DE ROUTE : LES ÉTAPES



DES GAINS À CHAQUE ÉTAPE

- Étude de l'état de l'art
- Application aux données confidentielles du secteur bancaire

- Gain empirique associé à la mutualisation par rapport aux processus actuels
- Perspectives d'industrialisation pour le marché
- Extension à d'autres cas d'usage



Banques coéquipières

Ateliers modélisation
- Périmètre métier
- Périmètre données
- Processus benchmark
- Processus mutualisé
  - jointures
  - annotation
  - revue manuelle

Exécution
- Apprentissage
- Détection
- Revue manuelle
- Évaluation

- Étude entre pairs de cas métier et scénarios LCB
- Stratégie de progression par partage de données

Type de phase

CONFIDENTIALITÉ

MODÉLISATION

TRANSVERSE

ACPR
BANQUE DE FRANCE

LA MÉTHODE RÉSUMÉE EN UNE PHRASE

L'objectif étant d'étudier l'amélioration de la performance prédictive des algorithmes de détection de transactions bancaires suspectes,

en respectant **les contraintes de confidentialité** imposées par les exigences légales relatives aux données sensibles,

les ateliers modélisation entre banques participantes appariées définiront le protocole expérimental supportant cette étude,

le Tech Sprint MCD tenu en parallèle sélectionnera les solutions technologiques garantissant la confidentialité et la sécurité des données utilisées dans l'expérimentation,

et **la suite de l'expérimentation** réunira banques et prestataires en équipes pour réaliser chaque protocole et analyser les résultats expérimentaux.

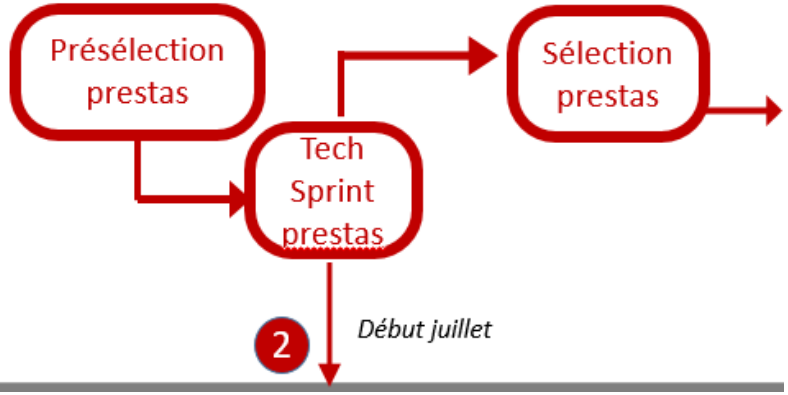
ACPR
BANQUE DE FRANCE

MUTUALISATION CONFIDENTIELLE DE DONNÉES

- Définition fonctionnelle et non conceptuelle
- PETs pressenties : F(HE), SMPC, TEEs, … (éventuellement DP)
- Autres approches sont moins pressenties mais pas exclues:
  - *Blockchain*
  - Apprentissage fédéré
- Ne pas négliger :
  - ni les approches hybrides
  - ni les approches « minimalistes » à la confidentialité
- Une maturité suffisante est requise pour :
  - adapter la solution au PoC en ≈ 2 mois
  - intégrer le protocole expérimental en ≈ 4 mois

ACPR
BANQUE DE FRANCE

LE TECH SPRINT MCD



- **Appel à candidatures** (avril) :
  les candidats décrivent leur solution
- **Présélection** des prestataires par l'ACPR sur cette base (mai)
- **Réalisation du PoC** défini par l'ACPR sur un scénario fictif et un protocole représentatif
- **Tech Sprint** (juillet) : présentation publique
  - de la solution proposée
  - des travaux du PoC
- **Sélection** de prestataires par les banques constituant leurs équipes, en présence de l'ACPR (avant octobre)

ACPR
BANQUE DE FRANCE

LE TECH SPRINT – MODALITÉS DU *PROOF OF CONCEPT*

L'ACPR fournira aux candidats :

- le scénario expérimental retenu pour le PoC
- les jeux de données synthétiques, documentés
- éventuellement une spécification de certains composants
  - mécanismes de jointure
  - modèles de détection solo
  - modèles de détection mutualisés.

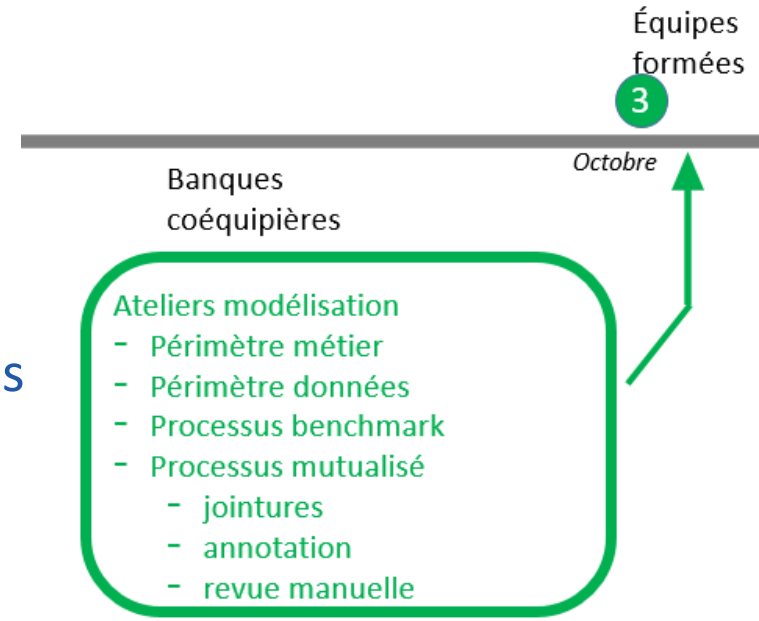
Chaque prestataire implémentera sa solution et préparera sa démonstration (preuve formelle algorithmique, argument architectural, ou toute autre garantie de confidentialité ou d'intégrité).

ACPR
BANQUE DE FRANCE

LES ATELIERS DE MODÉLISATION

Objectifs

- susciter la formation d'équipes, chacune définissant son protocole expérimental
- s'assurer d'un socle méthodologique minimal commun répondant aux attentes de l'ACPR
- identifier et valider les options les plus intéressantes (et réalisables)



Format

- Exercice sur table, restitution possible à l'issue

Angles d'étude

- Fonctionnel (LCB)
- Juridique (traitement des données, secret commercial)
- Technique (hors MCD : SGBDR, Data Engineering, ML)

Périmètre couvert : processus de benchmark et mutualisé

LES ATELIERS DE MODÉLISATION

Date	Thème de l'atelier
Mai	Préfiguration des équipes ; objectifs et contraintes
Juin	Données d'entrée (et filtrage initial)
Juillet/août	Données de sortie
Septembre	Modèles de détection
Octobre	Annotation et évaluation

Programme provisoire

(à ajuster en fonction des questions identifiées lors de l'atelier initial en mai)

ACPR
BANQUE DE FRANCE

LES ATELIERS : CONTENU PROVISOIRE

Atelier « données d'entrée »

- Périmètre de l'expérimentation
- Schémas de données d'entrée (potentiellement hétérogènes)
- Mécanisme de mutualisation
- Scénarios de filtrage initial

Atelier « données de sortie »

- Nature et type de données de sortie (classification des alertes)
- Potentiel d'enrichissement de ces données pour revue manuelle

Atelier « modèles de détection »

- Nature et caractéristiques des modèles utilisés pour l'expérimentation
- *Pour les modèles à base de règles :* adaptation des règles aux données mutualisées
- *Pour le ML, supervisé ou non :* ajustements (mineurs) des algorithmes.

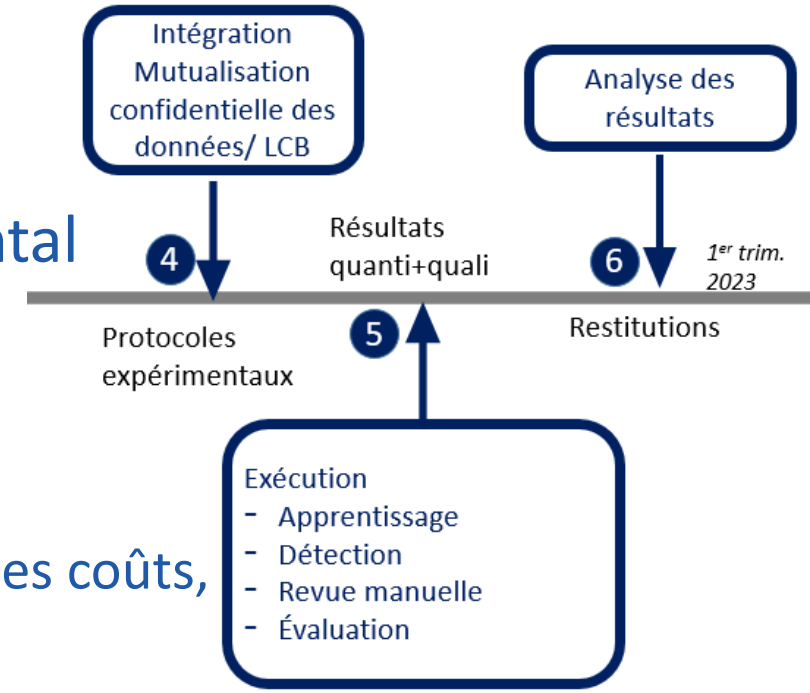
Atelier « annotation et évaluation »

- *Pour le ML, supervisé ou non :* labélisation des données d'apprentissage et de test
- Labélisation des données d'évaluation

ACPR
BANQUE DE FRANCE

ÉTAPE 4 : INTÉGRATION MCD/LCB

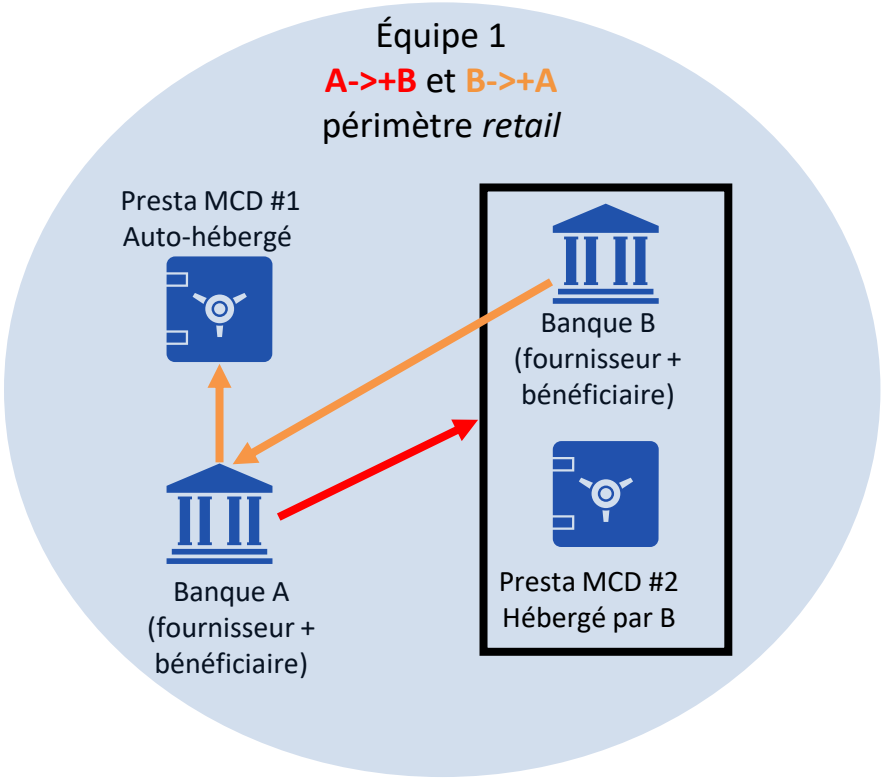
- Convergence entre les travaux MCD et LCB
- Finalisation du protocole expérimental
  - conditions matérielles d'exécution
  - calendrier
- Convention multipartite (précisant le traitement des données, la répartition des coûts, les responsabilités)
- Développement de la solution logicielle et/ou matérielle
  - mise en œuvre technique des livrables issus des ateliers métier
  - intégration de la solution du prestataire MCD
  - déploiement et opérations.

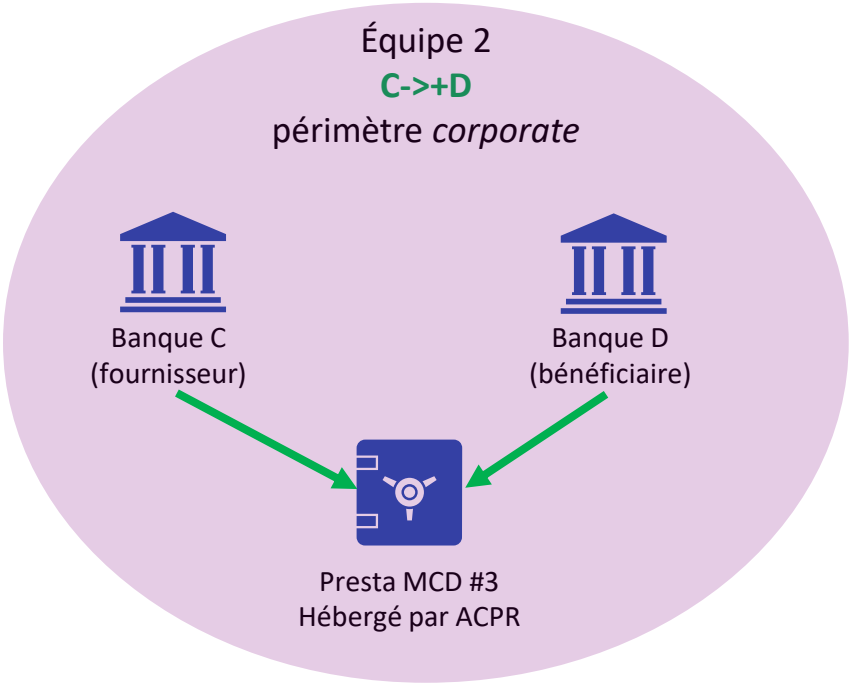


INTÉGRATION : MODES OPÉRATOIRES POSSIBLES

On note **A->+B** l'enrichissement confidentiel des données transactionnelles de la banque B par les données associées de la banque A

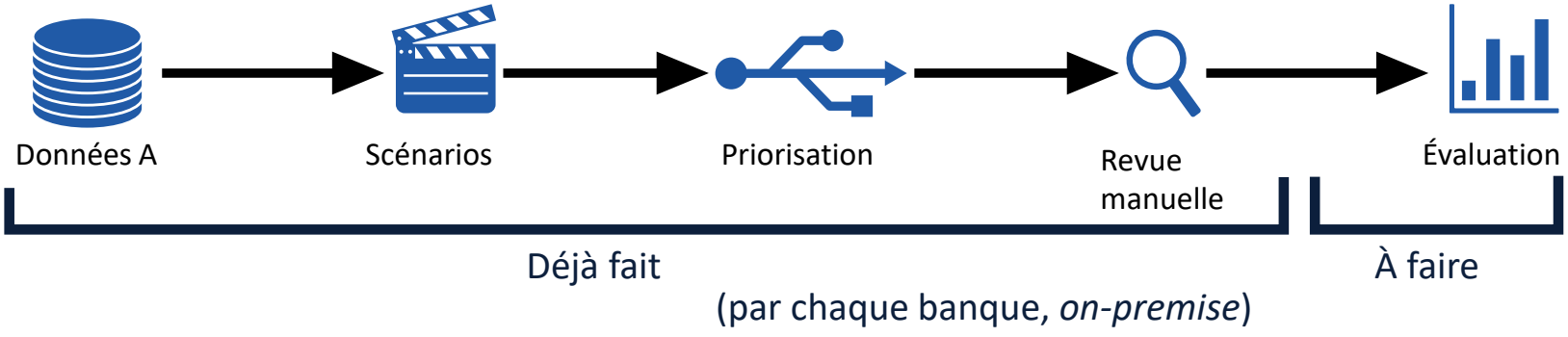
Deux exemples possibles d'organisation :



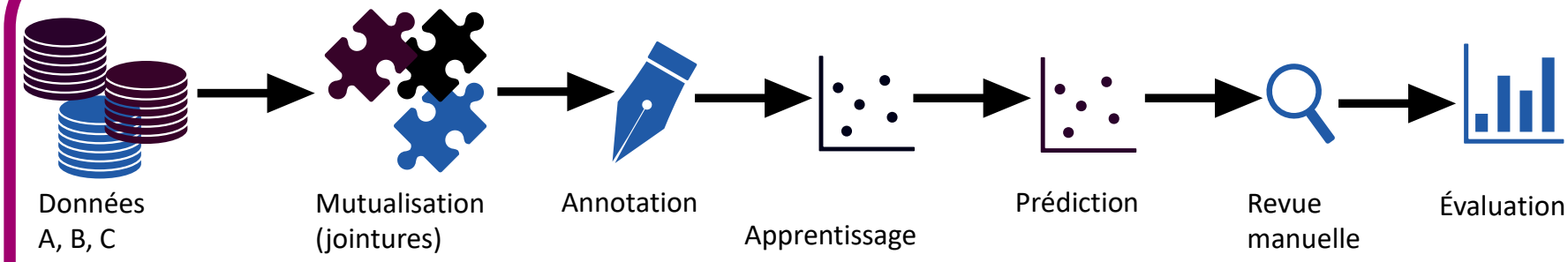


ÉTAPE 5 : PHASE D'EXÉCUTION

Processus benchmark



Processus mutualisé



ÉTAPE 6 : ANALYSE DES RÉSULTATS

- Réunion de chaque équipe avec l'ACPR
  - « *Post-mortem* » approfondi sur l'expérimentation
- Réunion de synthèse avec tous les participants
  - Présentation par chaque équipe de ses résultats
  - Partage d'éléments complémentaires avec les autres équipes (à discrétion des équipes)
- Restitution publique
  - Événement public
  - Rapport de synthèse.

ACPR
BANQUE DE FRANCE

Prochaines étapes : qui fait quoi, quand ?

À SUIVRE POUR LES BANQUES VOLONTAIRES

Qui :

1 contact par établissement

+ expertise thématique :

- métier : LCB-FT et TMS
- data : *data stewards / engineers / scientists*
- informatique : SGBD, IT
- juridique et réglementaire

Quoi, quand :

Avril
- S'inscrire à l'atelier de mai

Mai
- Atelier 0 : validation du programme d'ateliers, préfiguration des équipes, point sur le lancement du Tech Sprint

Juin
- Atelier 1 : données d'entrée et filtrage initial

Été
- Atelier 2 : données de sortie
- Tech Sprint : participation en tant que membres du jury des PoC

Sept.
- Atelier 3 : modèles de détection

Oct.
- Atelier 4 : annotation et évaluation
- Choix du prestataire MCD

ACPR
BANQUE DE FRANCE

À SUIVRE POUR LES PRESTATAIRES MCD

Qui :

- Engineering / Produit pour répondre à l'appel + réaliser le PoC
- Représentants pour pitch et démo au Tech Sprint

Quoi, quand :

Avril
- L'ACPR publie l'appel à candidatures, règlement inclus
- Chaque prestataire intéressé candidate et décrit sa solution

Mai
- L'ACPR sélectionne les participants au Tech Sprint
- L'ACPR fournit les éléments nécessaires au PoC

Mai-juin
- Chaque prestataire sélectionné réalise le PoC

Juillet
- Présentation (solution + PoC) au Tech Sprint MCD

Octobre
- Chaque banque participante (en consultation avec l'ACPR) a choisi son prestataire

ACPR
BANQUE DE FRANCE

5. Questions et réponses en direct

Et aussi :

Concernant l'expérimentation expe-lcb@acpr.banque-france.fr

Pour candidater au Tech Sprint techsprint2022@acpr.banque-france.fr