Décembre 2022

Tech Sprint ACPR sur la mutualisation confidentielle de données

Rapport de synthèse

Auteur : Laurent Dupont, Pôle Fintech-innovation, ACPR

1. Résumé

Le Tech Sprint en Mutualisation Confidentielle des Données organisé par l'ACPR en 2022 visait à éclairer les différentes méthodes de maintien de la confidentialité de données dans le cadre d'analyses collaboratives. Il s'inscrivait dans le cadre de l'initiative d'expérimentation en LCB-FT (Lutte Contre le Blanchiment et le Financement du Terrorisme) lancée par l'ACPR en mars 2022 et se poursuivant en 2023, dont il constituait un prérequis technologique.

L'objectif du Tech Sprint consistait à **identifier des solutions technologiques permettant de garantir que plusieurs acteurs peuvent collaborer en partageant en toute sécurité leurs jeux de données dont la confidentialité est à préserver**. Dans le cadre plus spécifique de l'expérimentation ACPR en LCB-FT, les solutions devaient permettre de **concevoir et mettre en œuvre des modèles de détection (notamment de transactions suspectes) opérant sur des données mutualisées entre pairs, tout en garantissant la conformité du traitement de ces données** vis-à-vis de toute contrainte réglementaire, législative ou propre à chaque établissement.

Une caractéristique essentielle du défi proposé par l'ACPR était **l'absence de « classement » des solutions évaluées** : plutôt que de juger les mérites absolus de chaque solution proposée, il s'agissait d'apprécier ses caractéristiques techniques et de permettre à chaque établissement financier participant à l'expérimentation ultérieure en mutualisation de données, d'évaluer l'adéquation des solutions à ses besoins et contraintes propres.

Douze solutions ont été conçues, mises en œuvre et présentées lors du Tech Sprint, certaines étant l'œuvre d'une startup seule et d'autres d'un consortium d'entreprises établies. Elles **couvraient une grande variété technico-fonctionnelle** : génériques ou spécialisées en LCB-FT ; reposant sur un intermédiaire ou totalement décentralisées ; couvrant la confidentialité des données en entrée mais aussi en sortie (*output privacy*) ; fournissant enfin pour la plupart des fonctionnalités de surveillance de transactions elles-mêmes très diverses.

À travers cette diversité, les solutions du Tech Sprint ACPR reflétaient collectivement **l'état de l'art mondial des technologies de maintien de la confidentialité** et de la cryptographie moderne. Ainsi le chiffrement homomorphe, les calculs multipartites sécurisés, l'apprentissage fédéré, les enclaves matérielles sécurisées, ou encore la confidentialité différentielle étaient représentés.

Les principaux enseignements que l'on peut tirer à l'issue de cet événement sont les suivants :

- **L'ensemble des approches proposées étaient applicables au défi du Tech Sprint**, tout en présentant leurs propres avantages et limitations. Par exemple, même les approches considérées *a priori* comme les plus consommatrices de ressources (temps de calcul ou coût de communication), telles que celles complètement décentralisées ou basées sur du matériel sécurisé, ont permis de réaliser la « preuve de concept » spécifiée par l'ACPR.
- Le type de mécanisme de mutualisation était très variable selon les solutions : croisement simple ou plus avancé entre données de transactions, voire construction de réseau, ou simples interrogations unitaires. Or **le mécanisme de mutualisation est apparu comme le facteur le plus limitatif parmi les solutions du Tech Sprint** en raison des contraintes qu'il impose sur le type de modèle de détection pouvant être intégré (et partant sur le gain potentiel attendu de la mutualisation de données).
- Outre les enjeux déjà identifiés d'ordre métier (maintien par les équipes LCB-FT du contrôle des modèles de détection opérant sur données mutualisées) et d'ordre technique (contraintes d'intégration au système d'information des établissements financiers), un facteur d'évaluation mis en avant par les acteurs financiers conviés au Tech Sprint était la complexité technologique des solutions proposées. Qu'elle soit avérée ou simplement perçue, cette complexité constitue à ce jour un frein à l'adoption des méthodes les plus sophistiquées de maintien de la confidentialité. À terme, toutefois, ces solutions présentent un indéniable intérêt en matière de garanties de sécurité, ainsi que l'ont brillamment démontré certains participants du Tech Sprint.

2. Introduction

2.1 Contexte et motivation du Tech Sprint

L'ACPR a conçu et lancé en mars 2022 un projet collaboratif sous forme d'expérimentation dont l'objectif principal consiste à prouver – ou infirmer – l'hypothèse selon laquelle la mutualisation de données permet d'améliorer la performance des systèmes de surveillance de transactions.

Les questions soulevées par cet objectif ont conduit l'ACPR à organiser, comme étape préalable de l'expérimentation, un défi appelé Tech Sprint en Mutualisation Confidentielle des Données (MCD). Ce Tech Sprint, dont le calendrier s'étendait de mai 2022 à la publication du présent rapport, comportait lui-même deux objectifs clefs :

- il visait tout d'abord à évaluer les différentes méthodes et techniques de maintien de la confidentialité et de l'intégrité de données mutualisées ;
- un objectif secondaire du Tech Sprint visait à susciter des partenariats entre établissements financiers participant à l'expérimentation en LCB-FT (appelés Établissements Participants par la suite), ayant constitué des équipes de pairs prêts à collaborer à cette fin, et fournisseurs de technologie MCD.

2.2 Calendrier

L'ACPR a publié le 16 mai 2022 un appel à candidatures, puis sur la base des dossiers soumis a sélectionné 12 participants au Tech Sprint. Chaque participant était composé soit d'une seule société soit de plusieurs sociétés faisant équipe à cette occasion.

Il était demandé à chaque participant au Tech Sprint de réaliser une « preuve de concept » (PoC) à partir de spécifications détaillées dans un document fourni le 13 juin par l'ACPR. La date limite pour achever la réalisation du PoC était fixée au 13 septembre, jour de présentation des travaux réalisés par les participants au Tech Sprint. En outre, une documentation de ces travaux devait être livrée à l'ACPR au plus tard le 3 septembre, et une période de questions-réponses approfondies s'étendait entre la journée de restitution et le 23 septembre.

2.3 Objet du présent rapport

Ce rapport de synthèse décrit la qualité et la variété des solutions présentées au Tech Sprint ; il résume les principaux enseignements de cette initiative ; enfin il présente chacune des 12 solutions de façon factuelle et concise.

Afin de décrire fidèlement les principales caractéristiques, similitudes et facteurs de différenciation des solutions MCD présentées, le contenu et le vocabulaire de ce rapport sont relativement techniques. On se reportera au glossaire en section 6 (ou à toute autre documentation concernant les technologies de maintien de la confidentialité ou de cryptographie) pour une définition des termes techniques les plus importants.

3. Description du Tech Sprint

3.1 Mutualisation Confidentielle de Données et technologies concernées

Lors de la conception du Tech Sprint, l'ACPR n'a pas souhaité limiter *a priori* le champ des technologies testées. Aussi a-t-elle choisi d'employer un terme *ad hoc*, celui de « Mutualisation Confidentielle de Données » (MCD). Sa définition est avant tout fonctionnelle et non conceptuelle : il désigne les technologies qui satisfont aux objectifs de l'expérimentation à venir, c'est-à-dire qui garantissent que plusieurs acteurs peuvent collaborer en partageant leurs jeux de données respectifs tout en en garantissant la sécurité. Plus précisément, le terme MCD désigne l'ensemble des technologies permettant de fournir des garanties de confidentialité (et idéalement d'intégrité) vis-à-vis des données utilisées dans l'expérimentation ACPR, depuis l'étape de mutualisation des données, en passant par la création et l'optimisation de modèles de détection, jusqu'à l'évaluation du gain de performance prédictive.

Le champ des technologies testées est donc proche de celui communément désigné par *Privacy-Enhancing Technology* (PET), tout en étant *a priori* plus large.

3.2 Modalités

L'ACPR a fourni aux participants un scénario de PoC, décrit dans le cahier des charges et comprenant :

- un cas d'usage hyper-simplifié (basé sur les protocoles expérimentaux anticipés et où les fonctionnalités spécifiques au domaine LCB-FT étaient à peine esquissés) ;
- le lien vers des jeux de données fictives ;
- une liste de tâches obligatoires ou optionnelles à réaliser.

Chaque participant devait à la fois mettre en œuvre sa solution de PoC et préparer une démonstration argumentée des avantages associés (sous forme de preuve algorithmique, d'éléments d'architecture ou de conception, ou toute autre forme de garantie de sécurité). Le cadre de l'exercice visait à répondre au scénario du Tech Sprint, mais aussi à anticiper l'application de la solution proposée à l'expérimentation ultérieure sur données réelles, visant à tester l'amélioration de la performance prédictive des modèles de détection LCB-FT.

3.3 Principes généraux

Neutralité technologique

Le Tech Sprint de l'ACPR fut tout d'abord conçu en respectant un principe de neutralité technologique, avec deux corollaires importants :

- d'une part, comme les autres initiatives de l'ACPR concernant l'innovation dans le secteur financier, l'événement est complètement indépendant du rôle de supervision de l'Autorité et décorrélé de ses missions de contrôle ;
- d'autre part, l'ACPR s'abstient de promouvoir ou de critiquer en quelque manière que ce soit les solutions techniques présentées à cette occasion, et *a fortiori* les fournisseurs à l'origine de ces solutions. En particulier, elle n'encourage ni ne décourage leur adoption par les entités supervisées du secteur financier.

En conséquence, tant le cahier des charges du PoC que les conclusions tirées du Tech Sprint sont présentés de façon aussi objective que possible, et en tout état de cause, sans aucun caractère prescriptif vis-à-vis des choix technologiques des acteurs financiers. Comme expliqué dans ce qui suit,

le Tech Sprint visait à étudier le spectre le plus large possible de méthodes et d'algorithmes. La section 4.1 démontre que cet objectif a été atteint.

Définition du défi

L'objectif de n'exclure *a priori* aucune méthode de confidentialité ni aucun cas d'usage peut se révéler en tension avec la recherche de résultats expérimentaux concrets, tangibles et quantifiables pour l'initiative de l'ACPR. Aussi le sujet du Tech Sprint a-t-il tenté de trouver un équilibre entre :

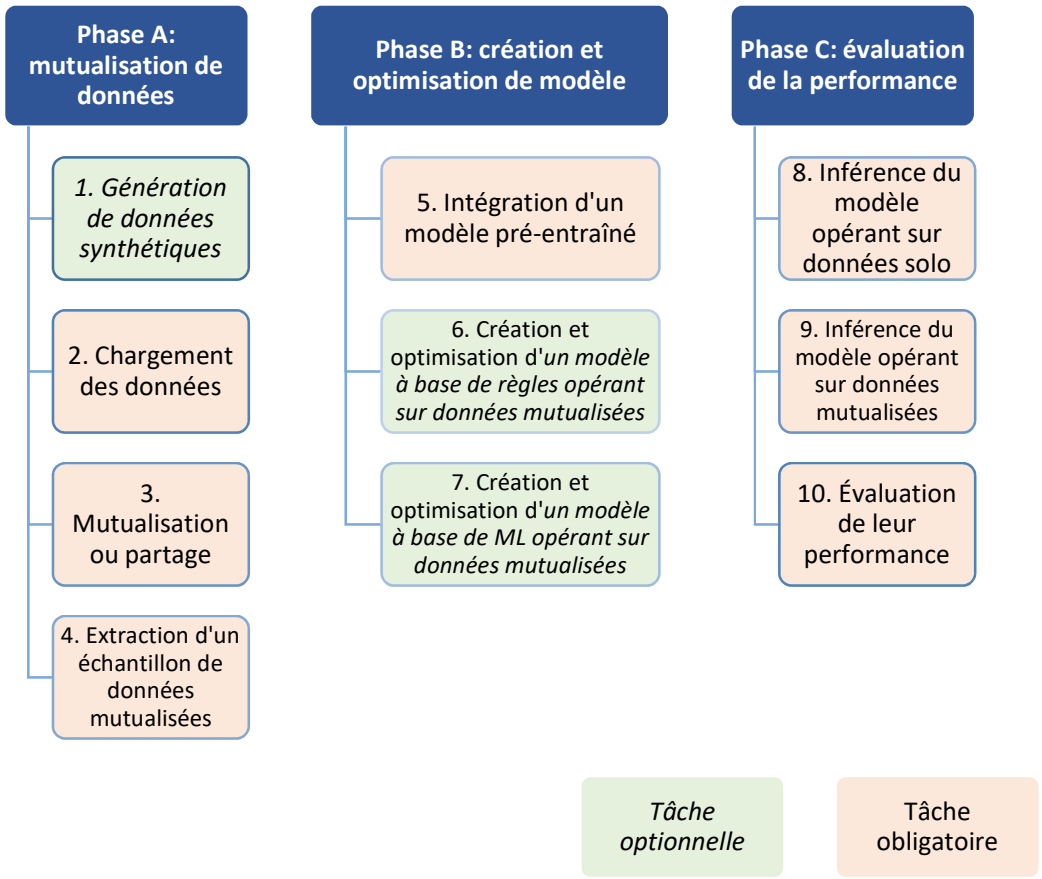
- d'une part, la généralité des exigences techniques ;
- d'autre part, un niveau de spécification et d'accompagnement permettant de maximiser les chances de réussite des étapes ultérieures de l'expérimentation, à savoir la conception puis l'exécution de chaque protocole expérimental sur données réelles par les Établissements Participants.

Dans la pratique, le cahier des charges du PoC définissait une liste de fonctionnalités de deux types :

- fonctionnalités obligatoires (ou « *must have* ») exprimées en termes génériques ;
- fonctionnalités optionnelles (ou « *nice to have* ») telles que le support de modèles de Machine Learning (ML), l'intégration de modèles de détection *ad hoc*, ou la confidentialité différentielle (*Differential Privacy* ou DP).

3.4 Tâches assignées au PoC

Le diagramme suivant résume la liste de tâches assignées dans le cahier des charges du PoC aux participants au Tech Sprint. La section 4 décrira m comment les 12 solutions ont traité chacune de ces tâches.



3.5 Exigences de documentation

Les participants au Tech Sprint devaient en outre produire une documentation de leur solution, incluant les éléments suivants pour chacune des tâches du PoC :

- le niveau de confidentialité maintenu pour chaque type de variable, et la façon de l'atteindre ;
- le modèle de menace considéré ;
- les garanties d'intégrité fournies, et la méthode utilisée pour les satisfaire ;
- la performance (ressources consommées et temps d'exécution) sur chaque environnement-cible considéré ;
- l'impact éventuel sur le temps de calcul de certains facteurs (par exemple l'utilisation de données réelles plutôt que les données fictives du PoC).

3.6 Évaluation des solutions du Tech Sprint

En pratique, ces travaux ne pouvaient être évalués comme un défi d'informatique typique ou un hackathon de data science classique.

En effet, d'une part, le défi proposé ne se prêtait pas à une solution unique : certaines solutions sont plus appropriées à de faibles volumes de données dont la confidentialité est à préserver et passent difficilement à l'échelle de données massives, tandis que d'autres proposent des garanties de confidentialité moins fortes mais une meilleure tenue en charge. D'autre part, chaque Établissement Participant a ses propres objectifs et contraintes, ne permettant pas de définir des critères simples de « classement » des solutions qui coïncideraient avec les choix de partenariat effectués, *in fine*, par les différents établissements participant à l'expérimentation. Au demeurant, les établissements étant seuls responsables de ce choix, le Tech Sprint n'avait pas pour vocation d'en préjuger.

Ainsi, l'évaluation des travaux réalisés et des livrables produits avait pour objectif de faciliter l'analyse technique de chaque Établissement Participant en fournissant l'information la plus complète possible et en mettant en commun les capacités d'analyse par la constitution d'un panel d'experts techniques sur la base du volontariat. Ce panel était composé d'1 à 3 membres représentant chacun des établissements participant à l'expérimentation LCB-FT, de 2 représentants de l'ACPR (1 ingénieur informatique et 1 data scientist) et de 2 représentants de la Banque de France (1 data scientist et 1 expert IT). Chaque membre du panel était sélectionné pour son expertise en sécurité des données, c'est-à-dire en cryptographie (si possible incluant les *PET*) et/ou en analyse de données confidentielles, avec idéalement une expertise complémentaire en cybersécurité.

Matériel utilisé pour l'évaluation

L'information utilisée pour évaluer les solutions du Tech Sprint devait inclure tous les types de matériel fournis au panel d'experts techniques, à savoir :

- les réponses à l'appel à candidatures reçues de chaque fournisseur candidat (entre le 16 mai et le 13 juin) ;
- la documentation de chaque solution (envoyée à l'ACPR avant le 3 septembre) ;
- la présentation de chaque solution lors de la journée de restitution (13 septembre) ;
- les réponses aux questions d'approfondissement suite à chaque présentation (jusqu'au 23 septembre).

Fiche d'évaluation et critères sur-mesure

Avant la journée de restitution des travaux, le panel s'était réuni pour discuter des principes et critères d'évaluation, afin de converger sur une fiche d'évaluation, dont le modèle figure en Section 7.

Cette fiche d'évaluation (et les critères qu'elle contient) était uniquement fournie aux Établissements Participants comme grille de lecture et non comme liste prescriptive de principes d'évaluation. Aussi, certains établissements ont opté pour d'autres critères, ou pour une pondération différente des mêmes critères.

Quelques critères sur-sur-mesure mis en avant par les Établissements Participants étaient les suivants.

Contrôle des modèles

Un établissement a souligné son besoin de garder le contrôle des modèles de détection résultant de l'expérimentation, un attendu logique mais qui doit être explicitement garanti pour toute solution déployée sur un scénario réel.

Complexité technique

Le niveau de complexité (ou de sophistication) inhérent aux garanties de sécurité de chaque solution a aussi été mentionné par certains Établissements Participants comme un critère important : en supposant que certaines des solutions (voire toutes) offrent un niveau de confidentialité équivalent tout au long du protocole expérimental, ce niveau de complexité deviendrait un facteur de sélection décisif (à minimiser bien sûr).

Questions d'intégration

Certains Établissements Participants ayant des contraintes d'intégration bien précises, certains se sont notamment éloignés des solutions matérielles telles que les TEE (enclaves matérielles sécurisées) en raison des coûts induits et de leur politique de gestion des systèmes d'information.

La suite de ce rapport n'a pas pour objectif de rendre compte des choix individuels effectués par les Établissements Participants mais de donner une description synthétique, sur la base des analyses menées avec l'aide du panel technique, des différentes caractéristiques des solutions proposées lors du Tech Sprint.

4. Vue d'ensemble des solutions proposées au Tech Sprint

4.1 Diversité des travaux

Les solutions conçues, mises en œuvre et présentées lors du Tech Sprint couvraient une grande variété de techniques et d'algorithmes, tout en représentant l'état de l'art des technologies de maintien de la confidentialité et de la cryptographie moderne. Leurs principales caractéristiques peuvent être, dans toute leur diversité, résumées comme suit.

Technologies

Elles englobaient l'ensemble du domaine des techniques de maintien de la confidentialité (voir page suivante).

Consortiums et équipes « solo »

Certaines solutions furent proposées, construites et présentées par un groupe de sociétés (dans certains cas ce partenariat préexistait au Tech Sprint, dans d'autres il avait été noué spécifiquement pour l'événement. D'autres étaient des solutions « tout-en-un », en général déjà construites par une société (souvent une startup) et ajustées au défi proposé.

Solutions génériques ou spécialisées

Certaines solutions sont des plateformes d'analyse collaborative adaptées à un cas d'usage générique, tandis que d'autres sont des produits spécialisés en LCB-FT ou en lutte contre la fraude.

Hébergement et déploiement

Les solutions du Tech Sprint adoptaient différentes approches d'hébergement et de déploiement : déploiement sur site (par chaque propriétaire de données), hébergement sur un *cloud* public (accompagné le plus souvent de fortes garanties de sécurité), ou dans certains cas ne nécessitant aucun déploiement applicatif.

Solutions décentralisées ou avec intermédiaire

Certaines solutions requièrent la désignation et l'implication d'un tiers de confiance ; à l'inverse d'autres sont entièrement décentralisées.

Fonctionnalités de surveillance de transactions

Les solutions proposées fournissent des fonctionnalités de surveillance de transactions également très diverses : modèles de Machine Learning à portée générique (supervisés ou non), analyse de réseaux, ou encore fonctionnalités spécifiques à un cas d'usage précis.

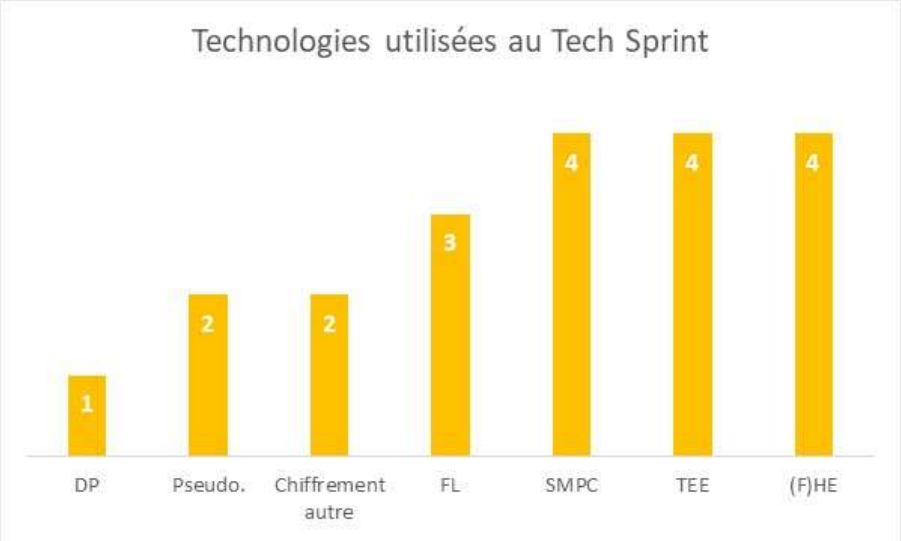
Flexibilité et facilité d'utilisation

De nombreuses solutions offrent une architecture modulaire et extensible, la plupart optant même pour une approche « orientée développeurs » en exposant une API[1], un SDK[2] ou un DSL[3].

[1] API (Application Programming Interface) désigne un accès à du code externe dont l'utilisation se fait à travers une interface simplifiée et permettant d'étendre les fonctionnalités d'un logiciel.
[2] SDK (Software Development Kit) désigne un kit complet facilitant l'accès à une ou plusieurs APIs, et fournissant un ensemble de bibliothèques logicielles, de documentation, d'exemples, etc. dans un ou plusieurs langages de programmation afin de l'intégrer au code utilisant la ou les APIs.
[3] DSL (Domain-Specific Language) désigne un langage de programmation, plus ou moins riche et complexe, créé spécifiquement pour communiquer avec un logiciel ou en contrôler les fonctionnalités, afin d'obtenir du code à la fois plus concis et plus expressif que par l'utilisation de langages de programmation standard (comme le fait une API ou un SDK).

4.2 Technologies de maintien de la confidentialité

Le spectre des solutions techniques, méthodes et algorithmes proposés par les participants peut être décomposé selon le type de technologie de maintien de la confidentialité utilisé. Le graphique suivant indique le nombre de solutions au Tech Sprint reposant sur chacun des principaux types de technologies de maintien de la confidentialité, tant pour la confidentialité des données d'entrée (*input privacy*) et dans certains cas leur intégrité, que pour la confidentialité des données en sortie (*output privacy*).



Abréviations utilisées

DP: confidentialité différentielle

FL: apprentissage fédéré

(F)HE: chiffrement (totalement) homomorphe

Pseudo. : pseudonymisation

SMPC: calculs multipartites sécurisés

TEE: enclave matérielle sécurisée

Il convient de souligner que ces technologies sont souvent utilisées en conjonction l'une avec l'autre – d'où le nombre total d'occurrences (20) supérieur au nombre de solutions (12) dans ce graphique. En particulier, les TEE (enclaves matérielles) sont souvent déployées en lien avec des méthodes de chiffrement complet ou avec un mécanisme permettant de pseudonymiser les attributs confidentiels.

Les garanties de confidentialité (et le cas échéant d'intégrité) associées à chaque type de technologie utilisée dans les solutions du Tech Sprint sont décrites dans ce qui suit.

Enclaves sécurisées (TEE)

Les solutions à base de TEE font généralement l'hypothèse qu'un attaquant distant a obtenu l'accès complet à la plateforme MCD, y compris à son infrastructure et aux autres participants au sein du Périmètre de Confidentialité et de l'enclave elle-même. Les propriétés de sécurité de la technologie TEE garantissent :

- que tous les calculs (mutualisation de données ; création, apprentissage et inférence de modèle de détection ; évaluation des résultats) sont exécutés sur une enclave matérielle sécurisée (via un processus dit « d'attestation à distance » et en certifiant que cette attestation est signée par une enclave et non par un simulateur) ;

- qu'aucune des parties prenantes, administrateur et opérateur de l'enclave inclus, n'a modifié le code des calculs ni leurs paramètres ;
- que les données et le code sont lisibles uniquement par l'enclave sécurisée et non par son administrateur ni son opérateur ;
- que la liste des identités de chaque participant aux calculs est celle attendue ;
- que l'intégralité de la pile logicielle utilisée pour les calculs est correctement identifiée et vérifiable.

Chiffrement homomorphe (HE, FHE)

Les solutions à base de chiffrement homomorphe font l'hypothèse d'un adversaire dit « honnête mais curieux » (voir glossaire). Le type de chiffrement utilisé, de nature probabiliste, garantit qu'à chaque chiffrement de la donnée, sa représentation chiffrée apparaîtra à un adversaire potentiel différente des opérations de chiffrement précédentes. Cette propriété empêche les attaques de type « *man in the middle* » ou MITM (c'est-à-dire dans le scénario du Tech Sprint, empêche un acteur malveillant de corrompre le résultat reçu par un analyste LCB-FT).

L'agrégation basée sur du chiffrement homomorphe est ainsi robuste à de telles attaques, néanmoins les problèmes de confidentialité de modèle subsistent en sortie de processus ; en outre l'utilisation de chiffrement homomorphe présuppose l'absence de collusion entre les nœuds participant aux calculs.

Calculs multipartites sécurisés (SMPC)

La plupart des solutions à base de calculs multipartites sécurisés font aussi l'hypothèse d'un adversaire honnête mais curieux. Une distinction au sein de cette catégorie se fait entre :

- le partage de secret non vérifiable (par exemple la méthode « *Shamir secret sharing* »), qui ne permet pas de vérifier l'exactitude de chaque part lors de la phase de réassemblage ;
- le partage de secret vérifiable, permettant en principe de vérifier l'honnêteté de chaque participant, au sens où ils ne fournissent pas de parts de secret inexactes.

La confidentialité différentielle (DP)

Les solutions à base de confidentialité différentielle fournissent des garanties de confidentialité en sortie (*output privacy*) : dans ce cas, le modèle de menace porte sur les utilisateurs ou les machines exécutant des requêtes via une API, notamment les data scientists en phase exploratoire ou de construction de modèle, les programmeurs travaillant à l'intégration de systèmes, et les équipes IT travaillant sur les modèles déployés. Dans ces 3 situations, les solutions de confidentialité en sortie réduisent le risque associé à une éventuelle fuite de données dans la mesure où toute information sortante est rendue conforme à la politique de confidentialité du propriétaire des données d'origine.

Dans l'ensemble, les solutions n'incluant pas de confidentialité en sortie sont moins adaptées que les autres au scénario proposé. La confidentialité différentielle – ou une méthode similaire de réduction du risque de ré-identification – a un double bénéfice :

- son principal avantage est de limiter l'impact d'une fuite de données potentielle ;
- un atout complémentaire, spécifique à l'expérimentation ACPR en LCB-FT, est que la DP facilite la spécification et la mise en œuvre d'un processus de revue manuelle d'alertes produites par l'inférence d'un modèle de détection opérant sur données mutualisées.

La DP pouvant être mise en œuvre par-dessus toute solution de maintien de la confidentialité en entrée, les deux bénéfices ci-dessus sont cumulables à d'autres catégorie de *PET*.

Apprentissage fédéré (FL)

Bien que 3 des solutions du Tech Sprint fournissent une capacité d'apprentissage fédéré sur données confidentielles, une seule d'entre elles a *de facto* appliqué le FL au scénario du PoC. Cette observation corrobore une hypothèse préalable au Tech Sprint : bien que l'apprentissage fédéré soit de prime abord attrayant pour ce type de scénario car c'est (avec certaines variantes de SMPC) l'une des rares méthodes ne nécessitant pas l'envoi des données par leur propriétaire et fournissant ainsi les garanties de confidentialité des plus fortes, il n'est pas réellement adapté à un scénario de LCB-FT car il impose trop de contraintes au processus de modélisation. Son principal défaut est que le processus d'agrégation du FL doit traiter les mises à jour de modèle indifféremment de la part de chaque participant, et ne permet donc que l'approche par empilement (et pas les approches par jointure standard ou *ad hoc*).

Pseudonymisation

Le risque de ré-identification peut aussi être réduit par des moyens plus traditionnels que la confidentialité différentielle, par exemple la pseudonymisation qui élimine en principe le risque de ré-identification accidentelle ou délibérée dans les cas où des données auraient fuité de la plateforme.

Autres méthodes cryptographiques

Outre les approches précédentes basées sur des technologies innovantes et souvent relativement sophistiquées, les autres approches se fondaient sur une méthode de chiffrement de bout en bout, plus traditionnelle mais combinée au recours à un tiers de confiance.

Certaines solutions reposaient sur un *cloud* public opérant aussi son propre KMS, si bien que le niveau de confiance envers l'opérateur et l'administrateur du *cloud* doit suffire à garantir qu'ils n'accéderont pas aux données en clair ni ne modifieront le code exécuté lors des calculs.

Sécurisation des données en transit

Outre les différents modèles de menace et garanties de sécurité associées décrites dans ce qui précède pour les données en cours d'utilisation (*« data in use »*), une solution du Tech Sprint, propriétaire et particulièrement innovante, fournit un modèle de menace particulièrement robuste pour les données en transit. Si ce type de garantie de sécurité sortait du périmètre fonctionnel requis par le Tech Sprint, le caractère novateur et éprouvé de cette solution justifie de la décrire succinctement :

- cette solution construit une « surcouche réseau sécurisée » garantissant le secret parfait au niveau de chaque nœud du réseau (l'adresse IP et le chemin utilisé sont masqués des nœuds participants et de chaque serveur) ;
- son architecture réseau « en anneau » fournit des garanties de sécurité et d'intégrité ainsi qu'une preuve d'origine (c'est-à-dire que la donnée provient d'un membre du réseau) ;
- elle est enfin résistante aux attaques de type « *man in the middle* » grâce à la distribution de ses nœuds de sortie, ne requérant ainsi aucun tiers de confiance au niveau du réseau.

4.3 Caractéristiques générales

Cette section présente les caractéristiques générales des solutions du Tech Sprint. Ces solutions ne pouvant être classées de façon adéquate sur la base de leur qualité intrinsèque ou absolue, l'exposé de leurs similitudes et de leurs différences devrait permettre d'appréhender leurs forces et faiblesses respectives, ainsi que les contextes et scénarios particuliers auxquels chacune est adaptée.

Tenue en charge et temps d'exécution

La difficulté de comparer les temps d'exécution

Le cahier des charges du Tech Sprint exigeait de fournir des mesures de temps d'exécution afin de permettre une évaluation comparative de la vitesse et de la tenue en charge des solutions proposées. Cette comparaison s'est néanmoins avérée difficile à effectuer car les valeurs mesurées dépendent largement des caractéristiques de l'environnement matériel d'exécution et de la définition précise des tâches réalisées, lesquelles ne sont pas forcément comparables entre solutions. Par exemple, avec certaines solutions basées sur du SMPC les données restent à leur emplacement d'origine : la tâche de chargement des données est alors virtuelle et son temps d'exécution ne peut être mesuré, contrairement aux solutions réalisant une étape de mutualisation physique et bornée dans le temps.

Plus précisément, la performance en temps d'exécution dépend de nombreux facteurs incluant le volume et le nombre de jeux de données, le type de mutualisation réalisée, les caractéristiques des machines utilisées pour les calculs, ou encore le niveau de sécurité requis des mécanismes de chiffrement (symétrique, asymétrique ou homomorphe). En outre, la logique de mutualisation et de croisement des données est définie *in fine* par les utilisateurs eux-mêmes (établissements financiers dans le cas de l'expérimentation ACPR), et – comme pour les calculs standards – l'efficience de la mise en œuvre de ces mécanismes ainsi que la puissance de calcul assignée au traitement sont les principaux facteurs dictant le temps d'exécution.

Par ailleurs, seule une minorité des participants au Tech Sprint ont fourni les mesures de temps de calcul pour la réalisation par leur solution de chaque tâche de PoC.

Pour ces deux raisons, la documentation fournie par les participants n'a pas suffi à réaliser une analyse comparée exhaustive de la tenue en charge et des temps d'exécution de chaque solution.

Observations générales

Certaines tendances générales concernant la tenue en charge ressortent néanmoins de l'analyse des résultats du Tech Sprint. En particulier, deux catégories de *PET* (à savoir SMPC et TEE) sont souvent considérées comme prohibitives en termes de ressources allouées d'une part à la communication entre nœuds participant à une analyse collaborative, d'autre part aux calculs réalisés sur chaque nœud :

- SMPC induit typiquement un surcoût de communication significatif, ce qui dans le cas du Tech Sprint pourrait conduire à de fortes restrictions concernant les classes de modèles de détection pouvant être construits et optimisés dans le respect des garanties de confidentialité. Les algorithmes de SMPC présentent généralement les caractéristiques suivantes : la connectivité nécessaire nuit aux temps de calcul, et la performance tend à croître lorsque de nouveaux participants rejoignent l'analyse collaborative (car il s'agit d'un *pool* virtuel de données, *a contrario* des calculs effectués sur données physiquement mutualisées entre participants).

- L'approche TEE induit typiquement un surcoût de calcul significatif, si bien que l'apprentissage d'un modèle au sein d'une enclave est souvent relativement lent[4] (entre autres raisons parce que le processus en question est séquentiel et non parallélisé). Une estimation grossière du surcoût associé au chiffrement et au déchiffrement réalisés en mémoire est que la plupart des scripts nécessitent 5 à 15% de temps d'exécution nécessaire par rapport au même script utilisant les mêmes ressources dans un environnement standard, avec une borne supérieure de 50 à 100% pour cette dégradation dans le cas de données et de requêtes réelles.

Il convient de noter que même les approches considérées a priori comme les plus consommatrices de ressources, telles que SMPC et TEE, se sont avérées applicables au scénario du PoC. Ainsi, leurs temps d'exécution sur des serveurs ou VM (machines virtuelles) standards étaient – lorsque ces informations ont été fournies par les participants – comparables aux autres solutions ou, au pire, d'un ou deux ordres de grandeur plus élevés (ce qui se justifie par les garanties de sécurité plus fortes associées à ces approches).

Traçabilité et auditabilité

Le cahier des charges du PoC présentait la traçabilité et l'auditabilité comme des prérequis intrinsèques aux solutions souhaitées. En particulier, les tâches obligatoires 4 et 10a consistaient à fournir une capacité d'extraction par échantillonnage des résultats respectifs de la mutualisation de données et de l'inférence du PDDM (*Pooled Data Detection Model* ou modèle de détection opérant sur données mutualisée).

De nombreuses solutions proposées pour le PoC ont soit négligé ces exigences de traçabilité et d'auditabilité, soit les ont réduites à leur plus simple expression, à savoir l'accessibilité de *logs* applicatifs sans prise en compte de la sensibilité de l'information sous-jacente. Même cette seconde option pose problème : si les *logs* contiennent des données d'entrée à maille fine ou des données issues de calculs intermédiaires cela peut conduire à une faille de confidentialité ; à l'inverse s'ils omettent de l'information détaillée afin de préserver la confidentialité des données leur intérêt sera en pratique limité (notamment pour le débogage).

Par contraste, l'une des solutions basées sur TEE s'est distinguée en offrant :

- des *logs* d'audit complets générés au sein de l'environnement de calcul confidentiel. Ces *logs* permettent une totale auditabilité et une transparence vis-à-vis des actions réalisées dans le Périmètre de Confidentialité (indiquant par exemple quand un calcul donné a été exécuté au sein de l'enclave sécurisée et par quel utilisateur ou quel programme) ;
- des messages d'erreur remontés directement à l'utilisateur lors de l'exécution. Ces messages contiennent le même niveau d'information que la trace complète sur données en clair, tout en préservant la confidentialité grâce à la suppression des informations confidentielles.

D'autres solutions offraient une fonctionnalité intermédiaire entre ces deux extrêmes, produisant une trace de toutes les actions réalisées par les utilisateurs mais sans la sophistication de la solution précédente (laquelle offrait un véritable environnement de data science opérant sur données confidentielles).

[4] Dans un futur proche, la performance des TEE devrait néanmoins grandement s'améliorer, notamment avec une nouvelle génération de moteurs de calcul confidentiel opérant sur GPU (*Graphical Processing Unit*) qui offrira un apprentissage de modèle bien plus rapide (par exemple le processeur Nvidia H100).

Gestion des clefs cryptographiques

La plupart des solutions du Tech Sprint reposent sur un KMS (*Key Management System* ou système de gestion de clefs de chiffrement) pour stocker et donner accès aux clefs de chiffrement.

Pour les solutions basées sur un TEE, les clefs secrètes sont automatiquement scellées dans l'enclave sécurisée, et ne sont donc visibles de personne, même des participants à l'analyse collaborative.

Certaines solutions adoptent une approche BYOK *(Bring Your Own Key)*, où les participants peuvent fournir leur propre clef, préexistant à l'analyse, plutôt que de faire générer une nouvelle clef ou paire de clefs par le système.

Plusieurs solutions sont également compatibles avec les systèmes tiers de gestion de clefs (par exemple Microsoft Azure Key Vault et son service outillé de *Hardware Security Module*).

Enfin, d'autres solutions ont choisi de ne pas traiter le commissionnement des clefs de chiffrement : tous les participants à l'analyse collaborative sont alors supposés partager un *key ring* (porte-clefs) via un protocole sécurisé de distribution de clefs, et gérer la création de clefs chacun de son côté.

Traitement des données confidentielles

Le cahier des charges du PoC classait les variables d'entrée en trois catégories :

- variables sans caractère confidentiel (donc visibles en revue manuelle) ;
- variables prédictives (avec un pouvoir prédictif pour les modèles de détection, donc à ne pas montrer en revue manuelle) ;
- variables de jointure (utiles pour le croisement des jeux de données, mais soit dénuées de pouvoir prédictifs soit délibérément exclues de l'ensemble des variables prédictives).

De nombreuses solutions proposées par les participants prenaient en compte ces trois catégories comme escompté, néanmoins quelques solutions ont négligé la deuxième catégorie, à savoir les variables prédictives, dans la mesure où elles supprimaient tous les attributs confidentiels suite à la phase de mutualisation des données. Il s'agit d'une limite importante à ces solutions en termes de fonctionnalités génériques, qui dans le cadre de l'expérimentation ACPR empêche l'apprentissage de modèles prédictifs sur des attributs transactionnels confidentiels ou de l'information client elle-même confidentielle (c'est-à-dire des attributs ayant un pouvoir prédictif mais ne devant pas être révélés à un établissement pair).

Capacité de modélisation confidentielle sur données mutualisées

Une caractéristique fondamentale des solutions du Tech Sprint est leur capacité à permettre la création et l'optimisation de modèles de détection opérant sur données mutualisées, tout en garantissant bien sûr la confidentialité des données traitées durant ces phases de création et d'optimisation. En d'autres termes, l'application d'une méthode de MCD à un scénario de modélisation en LCB-FT se traduit précisément par cette capacité. Par exemple dans le cas de modèles de détection à base de Machine Learning, une solution adéquate doit permettre aux data scientists de configurer et entraîner un modèle sur des données mutualisées contenant des attributs confidentiels exactement (ou quasiment) comme s'ils travaillaient sur des données non confidentielles[5].

[5] Une fonctionnalité similaire était attendue dans le cas des modèles à base de règles, avec un niveau de difficulté moindre car les règles peuvent être prédéfinies sans considérer les données d'apprentissage. Le niveau de confidentialité des données d'entrée a donc un effet bien plus limité sur le processus de modélisation que dans le cas du Machine Learning.

Une observation clef issue du Tech Sprint est que la capacité de modélisation confidentielle sur données mutualisées découle directement du mécanisme de mutualisation mis en œuvre par la solution. En pratique, le type de mécanisme de mutualisation est apparu comme le choix de conception le plus limitatif parmi les solutions du Tech Sprint en raison des contraintes qu'il impose sur le type de modèle de détection pouvant être intégré – et partant, sur le gain potentiel attendu de la mutualisation de données.

Les tableaux suivants décrivent les principaux types de capacité de modélisation confidentielle sur données mutualisées proposés au Tech Sprint, ainsi que leurs bénéfices et limitations respectifs, en explicitant le lien entre MCD et modélisation confidentielle. Le nombre de solutions du Tech Sprint correspondant à chaque type de capacité de modélisation confidentielle y est aussi figuré[6].

1. Requêtes à la demande	Proposées par **2** solutions
Définition	
Une capacité de requêtes à la demande signifie que la solution permet des requêtes « à la volée » sur des attributs spécifiques (dans le cas du PoC, typiquement une requête récupérant les attributs transactionnels d'un client ou d'un ensemble de clients).	
Bénéfices et limitations	
Cette capacité est bien plus restrictive que celles permettant la création et l'optimisation de modèles de détection : ceux-ci rendent notamment possible l'ajustement d'un modèle (basé sur des règles ou du Machine Learning) sur données historiques. Il convient toutefois de noter que certaines solutions du Tech Sprint permettaient des requêtes à la demande *en conjonction* avec une capacité de création et d'optimisation de modèles sur données mutualisées.	
Modèles de détection possibles	
Aucun, seules des requêtes *ad hoc* sont possibles.	

2. Modélisation sur données empilées	Proposée par **5** solutions
Définition	
Ce type de capacité de modélisation découle d'un mécanisme de mutualisation : - nécessitant que tous les jeux de données d'entrée partagent le même schéma (ou au moins un sous-ensemble commun de champs). - opérant par empilement de ces données (donc dans le cas de base de données relationnelles, par opérations UNION).	
Bénéfices et limitations	
L'empilement de données fut largement adopté par les participants car il s'accompagne hypothèses simplificatrices (en particulier, les données mutualisées peuvent être traitées de façon identique aux données solo). Toutefois, le type de modèles de détection envisageables est très limité en comparaison de la mutualisation par jointure, car l'empilement traite chaque Établissement Participant indifféremment. Ainsi, pour un champ X présent dans les jeux de données de A et de B, un tel modèle prendra en entrée une simple variable prédictive X définie sur l'empilement des jeux de données.	
Modèles de détection possibles	
La modélisation sur données empilées permet de créer et optimiser des PDDM : - utilisant comme variables d'entrée soit le schéma commun soit les champs communs - pouvant être entraînés sur l'union des jeux de données d'entrée (appelé apprentissage fédéré horizontal dans le cas du FL).	

[6] Le nombre total d'occurrences est supérieur à 12 car certaines solutions combinaient plusieurs capacités, par exemple requêtes à la demande et modélisation sur données jointes.

3. Modélisation sur données jointes	Proposée par **5** solutions
Définition	
Ce type de capacité de modélisation dérive d'un mécanisme de mutualisation : -ne nécessitant pas des schémas de données identiques -opérant par croisement de jeux de données (donc dans le cas de base de données relationnelles, par opérations JOIN). Le mécanisme de mutualisation peut employer une méthode de croisement quelconque (par exemple dans le cas du PoC, un *matching* exact ou approché sur les deux contreparties d'une transaction financière). La requête de jointure elle-même peut être une requête SQL standard, ou plus avancée comme le second exemple de requête donné dans le cahier des charges du PoC (transactions ayant un compte en commun et distantes d'au plus 2 mois).	
Bénéfices et limitations	
La modélisation sur données jointes est bien plus puissante que la modélisation sur données empilées car elle permet un ensemble d'attributs (variables prédictives) plus riche et un univers plus grand (relations entre transactions au lieu des simples transactions). Ainsi, pour un champ X présent dans les jeux de données de A et de B, un tel modèle sera compatible avec des modèles de détection prenant en entrée simultanément les variables X_A et X_B sur les transactions jointes.	
Modèles de détection possibles	
La modélisation sur données jointes permet de créer et optimiser des PDDM : - utilisant comme variables d'entrée l'union des schémas de données - pouvant être entraînés sur le résultat d'une requête SQL de jointure sur les jeux d'apprentissage respectifs (souvent appelé apprentissage fédéré vertical dans le cas du FL).	

4. Modélisation sur le graphe de clients	Proposée par **2** solutions
Définition	
Ce type de capacité de modélisation suppose que l'ensemble du réseau de transactions a été construit. La structure la plus naturelle représente typiquement chaque compte bancaire par un nœud et agrège les transactions au niveau des arêtes.	
Bénéfices et limitations	
La modélisation sur le graphe de clients permet de construire des modèles de détection, tant supervisés que non-supervisés, ayant une vue holistique de l'ensemble du réseau de transactions (tout en restant compatibles avec les modèles tabulaires plus classiques).	
Modèles de détection possibles	
La modélisation sur le graphe de clients correspond aux PDDM utilisant comme données d'entrée le résultat de l'unification des réseaux de chaque Établissement Participant.	

4.4 Traitement du PoC par les solutions

Cette section présente plus en détail la façon dont les 12 solutions du Tech Sprint ont, dans leur ensemble, su traiter et (partiellement ou intégralement) résoudre chacune des tâches demandées.

Étape A (mutualisation des données)

Tâche 1 : génération de données synthétiques

Le cahier des charges du PoC décrivait une tâche optionnelle de génération de données synthétiques : le but était de produire un jeu de données supplémentaire représentant les transactions d'une banque fictive C. Ce nouveau jeu de données pouvait être produit de façon synthétique *stricto sensu* (c'est-à-dire en émulant certaines propriétés des données existantes, donc vraisemblablement des transactions de A et B) ou de façon pseudo-aléatoire. Une contrainte était d'introduire de nouvelles connexions avec A et B, fournissant des motifs intéressants (tels que de nouveaux schémas d'activité suspecte) et non déductibles des deux jeux de données d'origine.

La moitié des solutions présentées ont choisi de mettre en œuvre cette tâche optionnelle et ont donc inclus une capacité de génération de données synthétique. Les fonctionnalités proposées reposaient sur différents types de mise en œuvre :

- certaines ont choisi la génération aléatoire en utilisant une approche *ad hoc* basée sur un modèle statistique ;
- d'autres ont appliqué la DP aux données d'entrée ;
- plusieurs ont mis en œuvre une approche plus robuste basée sur des GAN (*Generative Adversarial Networks* ou réseaux antagonistes génératifs) ;
- d'autres encore ont employé une combinaison de DP et de GAN, notamment en faisant appel à la méthode PATE-GAN qui fournit des garanties fortes de confidentialité différentielle sur un modèle entraîné par GAN[7].

Tâches 2, 3 et 4 : chargement et mutualisation des données

En pratique, les tâches de chargement et de mutualisation de données ont été réalisées sous forme de séquence atomique voire même simultanément.

On notera par ailleurs que certaines solutions ne contiennent pas d'étape de mutualisation des données à proprement parler :

- Les solutions basées sur du SMPC opèrent typiquement en créant un *pool* virtuel de données au sens où les jeux de données ne sont pas transférés en un même emplacement, mais où des morceaux (dans certains cas à la maille fine de l'octet) sont soumis à perturbation aléatoire et chiffrés par un protocole de partage de secret ;
- Le FL consiste à partager des modèles de détection, et non pas les données d'apprentissage alimentant ces modèles ;
- Comme détaillé dans la section suivante, quelques solutions exécutent des requêtes *ad hoc* sur des éléments de données individuels. Ce procédé sort du cadre du PoC du Tech Sprint et ne requiert pas de mutualisation préalable, néanmoins la plupart des solutions (toutes sauf une) proposent cette fonctionnalité en supplément de celle requise pour le PoC, à savoir la création et l'apprentissage de modèles opérant sur données mutualisées.

[7] Référence bibliographique : Jinsung Yoon and James Jordon and Mihaela van der Schaar. PATE-GAN: Generating Synthetic Data with Differential Privacy Guarantees, International Conference on Learning Representations (2019).

Le cahier des charges laissait toute liberté aux participants pour choisir les filtres à appliquer lors de la mutualisation. Cependant, le choix de filtres inadéquats (notamment des filtres trop larges) pouvant conduire à une explosion combinatoire (produit cartésien des jeux de données d'entrée) pour certains types de structures tels que les bases de données physiques, l'ACPR fournissait des filtres sous la forme de quelques règles basiques pouvant être exprimées trivialement en requêtes SQL :

- joindre les deux pattes d'une transaction, c'est-à-dire la vue d'une transaction donnée par les établissements A et B (voire C) ;
- joindre toutes les transactions impliquant un compte commun et éloignées de moins de 2 mois.

Un enseignement intéressant issu de l'ensemble des travaux du PoC est qu'un croisement sophistiqué de sources de données s'est avéré très difficile à réaliser en garantissant la confidentialité des opérations. Plus le mécanisme de mutualisation était simple, plus il a été adopté largement par les participants. En outre, le type de mécanisme de mutualisation fourni par une solution affecte directement le type de PDDM qui peut être créé et optimisé par cette solution.

Quant aux structures à base de graphes, un principe général était de charger les jeux de données d'entrée aussi complètement que possible dans le graphe de données mutualisées, dans la mesure où les exigences en stockage et requête à la volée sur de telles structures sont généralement proportionnelles à la somme (et non au produit) des tailles des jeux de données d'entrée. Les seuls participants ayant adopté cette approche ont de fait choisi de construire et d'analyser le graphe complet des transactions ou des comptes (cette seconde option fournissant déjà un niveau d'agrégation et donc de tenue en charge intrinsèque).

Étape B (création et optimisation de modèle)

Dans une seconde étape (taches 5 à 7), les solutions proposées pour le PoC devaient permettre à l'établissement A de charger son modèle de détection actuel, appelé pour les besoins de l'expérimentation un SDDM (*Solo Data Detection Model* ou modèle de détection sur données solo), dans le Périmètre de Confidentialité. Le scénario du PoC consistait en effet à exécuter l'inférence et l'évaluation des modèles comme une seule et même tâche, qu'il s'agisse des données solo ou des données mutualisées, afin d'obtenir un protocole expérimental plus simple et cohérent.

La solution devait ensuite permettre à l'établissement A de construire un modèle de détection sur données mutualisées (PDDM). Les participants étaient libres de choisir de mettre en œuvre soit un modèle à base de règles, soit un modèle à base de Machine Learning (voire les deux) ; certains ont en outre choisi de mettre en œuvre un modèle à base d'analyse de graphe. La phase d'inférence du PDDM, suivie de l'évaluation de sa performance prédictive, devait ensuite être effectuée dans l'étape C sur le(s) modèle(s) choisi(s).

Dans le cas d'un PDDM basé sur du Machine Learning, les solutions devaient permettre l'apprentissage d'un modèle sur un sous-ensemble des données d'entrée ; dans le cas d'un modèle basé sur des règles, elles devaient permettre la définition d'un ensemble de règles en adéquation avec le PoC.

Entraînement des modèles et variables cibles

Les jeux de données de test incluaient aussi des variables cibles hypothétiques sous la forme d'une annotation *is_alert_solo* (respectivement *is_alert_pooled*) associée à chaque transaction, indiquant si une transaction vue par l'établissement A avait été marquée comme suspecte sur la base des informations connues de A uniquement (respectivement sur la base des informations connues de A ou

de B). Ainsi le but de la tâche 5 était d'optimiser le SDDM en fonction de la variable *is_alert_solo*, celui de la tâche 7 d'optimiser le PDDM en fonction de la variable *is_alert_pooled*.

Ces deux variables cibles n'avaient pas été générées aléatoirement, sans pour autant représenter des « sources de vérité » (pour employer le jargon du Machine Learning) issues d'un scénario réel sur la base de transactions réelles. L'objectif du PoC n'était pas de mesurer la performance prédictive d'un modèle, mais de démontrer qu'un modèle de détection (par exemple détection d'activités suspectes) pouvait être intégré et entraîné au sein du Périmètre de Confidentialité[8].

Tâche 5 : intégration, création et optimisation du SDDM

Les participants au Tech Sprint devaient fournir à la fois :

- la création d'un SDDM, qui peut être un modèle de Machine Learning très simple dans la mesure où l'optimisation de la performance des modèles sortait du contexte du PoC ;
- le chargement et l'intégration d'un modèle pré-entraîné dans le Périmètre de Confidentialité.

Ces tâches étaient obligatoires pour le PoC car l'intégration, la création et l'optimisation de modèle sont supposées plus faciles à mettre en œuvre pour un SDDM que pour un PDDM. Elles visaient à fournir une solution cohérente à l'inférence des deux types de modèles (SDDM et PDDM) et à la comparaison de leurs performances[9].

Tâches 6 et 7: création et optimisation de PDDM

Le cahier des charges du PoC suggérait explicitement de considérer comme PDDM à la fois des modèles à base de règles et à base de Machine Learning ; outre ces deux catégories, certains participants ont choisi de mettre en œuvre des modèles d'inférence à base d'analyse de graphe.

La création et l'optimisation de modèles de détection sortaient du cadre du Tech Sprint, dont le but premier consistait à fournir une plateforme de calculs confidentiels, c'est-à-dire un service « clef en main » permettant de réaliser une procédure d'analyse collaborative avec des garanties de confidentialité et d'intégrité. Ainsi, concernant les phases de cette procédure liées aux modèles de détection, l'exigence était de permettre la création et l'optimisation de modèles, mais pas de fournir des fonctionnalités de ML ou d'AutoML[10]. De nombreux participants au Tech Sprint ont toutefois inclus dans leur proposition des fonctionnalités de ML, simples ou plus avancées, soit parce que leur solution fournie « sur étagère » comportait déjà de telles fonctionnalités, soit parce qu'ils considéraient que la valeur ajoutée dans le cadre du PoC le nécessitait.

[8] La motivation de ce choix est qu'il existe seulement deux façons de construire une solution de modélisation prédictive garantissant la confidentialité et opérant sur des données réelles dans un contexte réaliste : soit commencer par des données réelles et montrer d'un seul coup la possibilité d'entraîner un modèle avec une bonne performance prédictive et en toute confidentialité ; ou alors commencer par traiter la question de la mutualisation sécurisée au moyen d'une plateforme de MCD, et seulement ensuite alimenter un PDDM au moyen de données réelles afin de comparer sa performance à celle d'un SDDM. Le Tech Sprint ACPR a clairement opté pour cette seconde approche, après avoir constaté l'échec d'autres initiatives en LCB-FT et lutte contre la fraude ayant tenté la première approche.

[9] En principe, non seulement la création et l'optimisation de modèle mais aussi l'inférence sur données solo pourraient être réalisées hors du Périmètre de Confidentialité. La raison principale pour exécuter l'inférence du SDDM au sein de ce périmètre est d'assurer une comparaison significative entre les performances prédictives, ainsi que de fournir une expérience homogène aux programmeurs utilisant la plateforme MCD.

[10] L'AutoML désigne un ensemble de méthodes automatisant certaines tâches consistant à appliquer le ML à des cas d'usage pratiques.

La consigne générale pour les tâches 6 et 7 était d'imposer aussi peu de divergence que possible entre le modèle de référence sur données solo (SDDM) et celui sur données mutualisées (PDDM). Cela n'implique pas l'identité de ces modèles – ils différent au minimum en raison des variables plus nombreuses prises en compte par le PDDM. En outre, l'importance des variables (dans le cas d'un modèle à base de ML) ou les règles de décision (dans le cas d'un modèle à base de règles) doivent être ajustées en fonction de cette information supplémentaire.

Toutefois, le but reste de ne pas fausser la mesure du gain de performance associé à la mutualisation de données en changeant radicalement de modèle, par exemple en passant d'un simple arbre de décision pour le modèle « solo » (SDDM) à un réseau de neurones pour le modèle sur données mutualisées (PDDM).

Les fonctionnalités de ML représentées par les solutions couvraient un champ extrêmement large :

- en termes de classes de modèles, elles comprenaient du ML non supervisé, du ML supervisé, des moteurs de règles, ou de façon plus originale des modèles causaux[11].
- Certaines solutions n'autorisaient qu'un petit nombre de classes de modèles de ML : en particulier, celles basées sur du HE (*Homomorphic Encryption* ou chiffrement homomorphe) ne permettent généralement pas les modèles complexes, tout au plus les variations sur des modèles linéaires. Néanmoins, certaines solutions à base de HE permettent l'entraînement et l'inférence non seulement de régressions, mais aussi de GLM (*Generalized Linear Models*) voire de GBT (*Gradient-Boosted Trees*) et de perceptrons multicouches.
- Concernant la structure des modèles, certaines solutions opèrent sur données tabulaires (c'est surtout le cas des modèles de ML supervisé), d'autres sur des représentations en graphe (ML non supervisé surtout), quelques solutions offrant une API assez flexible pour accepter les deux types de modèles, de données sources et de variables prédictives.

Comme expliqué en Section 4.3, la capacité de modélisation confidentielle sur données mutualisées découlait directement du type de mécanisme de mutualisation des données.

Étape C (performance prédictive)

Tâches 8, 9 et 10: évaluation de la performance prédictive des modèles

Le cahier des charges du PoC demandait aux participants de calculer des métriques de performance à la fois sur le modèle de référence opérant sur les données solo (SDDM) et sur le nouveau modèle opérant sur les données mutualisées (PDDM). Outre les métriques standards (précision et rappel), les participants étaient invités à inclure toute autre métrique pertinente.

L'une des caractéristiques les plus atypiques (et contre-intuitives même pour les participants) du Tech Sprint MCD est que la performance prédictive n'entrait pas en ligne de compte dans le succès des travaux. Le Tech Sprint visait essentiellement à sélectionner les solutions les plus adaptées pour *permettre* la mutualisation de données et l'optimisation de modèles *avec des garanties de confidentialité* ; à l'inverse la procédure d'optimisation de modèles elle-même devrait être une brique totalement indépendante de la solution de LCB-FT envisagée – et partant, hors du cadre du PoC. Par

[11] Un modèle causal est un modèle conceptuel décrivant les mécanismes de causalité d'un système. Un tel modèle peut améliorer la conception d'analyses en fournissant des règles claires permettant de décider quelles variables doivent être inclues ou contrôlées. Les modèles causaux sont appliqués entre autres au Machine Learning. Dans les solutions présentées lors du Tech Sprint, il s'agissait plus précisément de réseaux bayésiens causaux, combinant les avantages des données historiques et d'une base de connaissance adaptée au domaine considéré. Ils permettent de calculer un risque à partir de valeurs d'incertitude et présentent l'avantage supplémentaire d'être convertibles en modèles à base de règles.

ailleurs, seule une signification limitée pouvait être attribuée à des valeurs de performance prédictive calculées sur les données fictives du PoC.

Certains participants ont néanmoins choisi de démontrer leur capacité (et parfois leur savoir-faire pratique) à optimiser les modèles. Ils ont surtout comparé des métriques telles que le score F1 entre SDDM et PDDM, ce qui présentait plusieurs limites (choix de deux classes de modèles différentes, variable cible sur données solo sans association évidente à la variable cible sur données mutualisées, etc.) mais a fait ressortir dans la majorité des cas un net gain de performance prédictive associé à la mutualisation de données.

Un travail particulièrement intéressant a été réalisé par deux participants utilisant des analyses non supervisées à base de graphes, conduisant à des observations instructives en termes de faux positifs et faux négatifs issues de la comparaison de leurs résultats avec les variables cibles fournies dans les données du PoC. Quoique sortant du périmètre du PoC, ces deux participants ont fourni une solution « tout en un » couvrant mutualisation de données, Machine Learning supervisé et analyses non supervisées, le tout appliqué aux données confidentielles.

Tâches complémentaires : échantillonnage en sortie et revue manuelle

Bien que la revue manuelle ne fasse pas à proprement parler partie des tâches exigées pour le PoC[12], l'expérimentation sur données réelles suivant le Tech Sprint MCD nécessitera vraisemblablement une telle étape d'annotation par opérateur humain des sorties des deux modèles de détection. Certaines variables d'entrée seront visibles en revue manuelle, tandis que d'autres (y compris des variables prédictives employées par les modèles) ne seront pas accessibles aux annotateurs car elles contiennent une information appartenant à un établissement pair et dont la confidentialité est à préserver.

La plupart des solutions du PoC ne proposaient aucune fonction d'exportation des sorties. Certaines permettaient une exportation très basique dans laquelle les attributs confidentiels étaient simplement supprimés (ce qui ne permet qu'une revue manuelle minimaliste, basée sur les données solo et non sur les données mutualisées). Quelques solutions enfin ont fourni une fonctionnalité d'exportation intéressante ouvrant la voie à une phase de revue manuelle, soit en réduisant le risque de ré-identification par une simple pseudonymisation, soit en quantifiant le risque résiduel associé au moyen de critères de « k-anonymat » ou par la confidentialité différentielle.

[12] Pour les besoins du PoC, un mécanisme rudimentaire d'évaluation automatique de la performance pouvait être utilisé, accompagné d'une simple fonctionnalité d'extraction d'un échantillon en sortie.

5. Les solutions du Tech Sprint en détail

Cette section présente chacune des 12 solutions de façon plus détaillée, en résumant leurs caractéristiques selon 4 dimensions : l'architecture, le modèle de sécurité et sa gouvernance, la capacité de mutualisation des données, et enfin les fonctionnalités LCB-FT à savoir ici de création et optimisation de modèles de détection.

Pour éviter les erreurs factuelles, chaque résumé a été soumis à la relecture de l'équipe concernée.

Un lien est fourni vers une brève présentation vidéo de chaque société (ou équipe de sociétés) et de la solution proposée. L'ensemble de ces présentations vidéo, ainsi qu'un reportage de 5 minutes résumant l'ensemble du Tech Sprint, sont regroupés au sein d'un playlist YouTube :

➔ **Lien vers le reportage du Tech Sprint sur YouTube**

5.1 Atos - Privitar - TigerGraph – IBM

➔Lien vers la présentation vidéo

Architecture de la solution	
Périmètre de Confidentialité	**Hébergement**
Le Périmètre de Confidentialité comprend un environnement *Virtual Private Cloud* (VPC) isolé, le *cloud* sécurisé IBM FS hébergeant d'autres nœuds (y compris la plateforme Privitar et le composant d'analyse TigerGraph).	*Le Cloud* sécurisé IBM FS (Financial Services) est le plus sécurisé mais la solution peut fonctionner sur tout autre Cloud
Participants et leurs rôles	
- Un intermédiaire : tiers de confiance agissant comme point d'entrée pour le traitement des données, mais sans accès aux données brutes - Un fournisseur de services pour la plateforme MCD, Atos, sans accès non plus aux données en clair - Fournisseurs de données : chaque établissement avec son propre jeu de données	

Modèle de sécurité et gouvernance
Modèle de menace
La solution inclut la coordination des contributeurs avec un intermédiaire qui joue le rôle d'un tiers de confiance. L'ensemble des contributeurs doit donc s'accorder sur la désignation de l'entité légale jouant le rôle de l'intermédiaire.
Gestion des clefs et processus de chiffrement
Lorsque les données sont envoyées dans la zone protégée, elles subissent un double chiffrement : le premier avec une clef du contributeur, le second avec celle de l'intermédiaire. Par la suite à réception : - Les données sont déchiffrées au moyen de la seule clef privée de l'intermédiaire. À ce stade, les données sont toujours protégées par la clef de chiffrement du contributeur. - Les données sont re-chiffrées avec une clef dite « aveuglante » utilisée pour le chiffrement homomorphe qui permet de reconstruire l'intégrité des données sur l'ensemble des transmissions des contributeurs. - Les données sont alors traitées variables par variable selon les exigences des Établissements Participants et les techniques de protection proposées par Privitar (Tokenisation RegEx, Perturbation, Généralisation, …).

Fonctionnalité de mutualisation de données	
Chargement des données	**Mutualisation des données**
Les données de chaque Établissement Participant (contributeurs) sont doublement chiffrées par le logiciel Privitar et chargées dans un VPC sur le *cloud* sécurisé IBM FS.	Croisement de données tabulaires (basé sur du PSI)
Matching flou (Fuzzy matching)	
Résolution et reconnaissance d'entités (*Entity Resolution*) basée sur l'analyse de graphes	
Génération de données synthétiques	
À la fois standard et *ad hoc* (c'est-à-dire spécialisée pour un schéma de données)	

Fonctionnalités de modélisation (et requête) sur données mutualisées	
Variables d'entrée possibles	**Modélisation avec garanties de confidentialité**
Soit la table des transactions, soit le graphe de transactions enrichi de données issues de l'analyses de graphes.	Modèles conventionnels ou à base de graphes (PyTigerGraph, PyTorch Geometric, Tensorflow, DGL, etc.). Les modèles peuvent ainsi être définis par des règles ou via du Machine Learning, sur la base de données tabulaires ou graphe, ou par une combinaison de ces options.
API	
Un environnement Python de data science est fourni avec accès à scikit-learn, PyTorch et PyTorch Geometric (pour les modèles opérant sur graphes), Tensorflow et la bibliothèque DeepGraph.	

5.2 Cleyrop - Cosmian – Ekimetrics

➔Lien vers la présentation vidéo

Architecture de la solution	
Périmètre de Confidentialité	**Hébergement**
Le Périmètre de Confidentialité comprend plusieurs *Data Hubs* (déployés dans chaque Établissement Participant), un *Pooling Hub* incluant un TEE dans lequel les jeux de données sont centralisés, et un *Compute Hub* sur lequel sont exécutés l'entraînement et l'inférence des modèles.	OVH
Participants et leurs rôles	
La solution Cleyrop / Cosmian basée sur un TEE définit les rôles suivants : - Environnement de déploiement du TEE : un tiers de confiance pour l'expérimentation - Propriétaire des calculs : rôle d'administrateur dans l'expérimentation - Fournisseurs de données : chaque établissement avec son propre jeu de données - Fournisseur de code: l'établissement exécutant son protocole avec son propre modèle de détection - Consommateur des résultats : le même établissement, seul à accéder aux résultats.	

Modèle de sécurité et gouvernance
Gestion des clefs et processus de chiffrement
La plupart des primitives cryptographiques utilisées sont celles de la famille NaCl : - X25519 (RFC 7748) pour "Elliptic Curve Diffie-Hellman" (ECDH) Key Exchange - Ed25519 (RFC 8032) pour la signature digitale - XSalsa20-Poly1305 (RFC 7539) pour la méthode "Authenticated Encryption with Associated Data" (AEAD)
Modèle de menace
La solution Secure Computation fournit à chaque participant les garanties suivantes: - ils exécutent le calcul sur une enclave matérielle sécurisée (avec une attestation générée par l'enclave garantissant la présence d'une enclave SGX valide et non d'un simulateur) ; - l'opérateur de l'enclave (Cosmian sur sa plateforme publique) n'a trafiqué aucun paramètre du calcul ; - leur code/données ne peut être lu que par l'enclave sécurisée, ni son opérateur, ni son hébergeur, ni aucun autre participant ; - la liste des participants est correcte (nombre et identités) ; - le code fourni par le fournisseur de code porte la signature attendue ; - le reste de la pile logicielle est correct et correctement identifié (donc vérifiable).
Procédure d'autorisation des requêtes
Le fournisseur de code doit approuver les calculs. En complément, l'identifiant de l'enclave sécurisée permet à chacun des participants de vérifier que le calcul est exécuté dans une enclave sécurisée.

Fonctionnalité de mutualisation de données	
Chargement des données	**Mutualisation des données**
Tous les participants commencent par partager un secret. Une fois la clef publique de chaque participant envoyée à l'enclave sécurisée, celle-ci génère son identifiant, puis chaque participant envoie ses données chiffrées ainsi que sa clef symétrique scellée avec la clef publique de l'enclave et signée avec sa propre clef privée.	Par une méthode d'unification de n-uplets
Matching flou	
Une méthode basée sur une distance de Levenshtein est exécutée au sein de l'enclave sécurisée.	

Fonctionnalités de modélisation (et requête) sur données mutualisées	
Variables d'entrée possibles	**Modélisation avec garanties de confidentialité**
Les mêmes champs que les jeux de données d'entrée	Aucune fonctionnalité de Machine Learning confidentiel (car les attributs sensibles sont simplement supprimés)
API	
Du code Python arbitraire peut être exécuté au sein de l'enclave Intel SGX. Le *Data Hub* de Cleyrop permet l'écriture de notebooks Python, la récupération des résultats, l'affichage des alertes dans un outil de dataviz et tous types de traitement à destination de contrôles post traitement.	

5.3 Decentriq

➔ **Lien vers la présentation vidéo**

Architecture de la solution	
Périmètre de Confidentialité	**Hébergement**
Le Périmètre de Confidentialité est matérialisé sous la forme d'une *Data Clean Room* créée par la solution Decentriq.	Par défaut sur des instances européennes d'Azure Confidential Compute. Également possible sur un *cloud* privé ou *on premise* si un TEE est disponible.
Participants et leurs rôles	
- Plateforme de mise à disposition des *Data Clean Rooms*, incluant des TEEs et gérée par Decentriq - Fournisseurs de données : chaque établissement avec ses propres jeux de données - Analystes de données: utilisateur de chaque établissement qui analyse les données combinées et extrait les résultats	

Modèle de sécurité et gouvernance	
Détails techniques (algorithme ou mise en œuvre)	
La technologie de calcul confidentiel est basée sur une architecture matérielle ne nécessitant pas la mise en œuvre d'algorithmes *ad hoc* de traitements de données dans une couche logicielle. Les ressources de calcul confidentiel disponibles actuellement sont Intel SGX et AMD SEV-SNP.	
Modèle de menace	
Il suppose qu'un attaquant distant a obtenu un accès complet à la plateforme Decentriq, y compris à l'infrastructure et aux colocataires du Périmètre de Confidentialité et de l'enclave. Ce modèle de menace découle des propriétés de sécurité de la technologie TEE et de sa capacité à fournir une attestation distante du code exécuté, afin de ramener la source de confiance au niveau matériel. Il convient néanmoins de noter que si un attaquant obtient un accès à la couche matérielle, un risque de manipulation physique du CPU existe qui peut conduire à une fuite de clefs de chiffrement. C'est pourquoi des fournisseurs de *cloud* bien établis et certifiés doivent être utilisés pour protéger les serveurs utilisés par la solution. Ainsi la source de confiance réside *in fine* chez les fabricants de matériel (Intel or AMD). De plus, la logique exécutée au sein du Périmètre de Confidentialité est définie par les participants, et toutes les données d'entrée sont fournies par ces mêmes participants. Ainsi une logique défectueuse ou une manipulation des données par un participant pourrait conduire à une potentielle fuite de données confidentielles, aussi convient-il que tous les participants valident la définition de la *Data Clean Room* avant d'y charger leurs données et de participer aux calculs, voire de mettre en œuvre des contrôles de vérification du chargement effectué.	
Procédure d'autorisation des requêtes	**Gestion des clefs et processus de chiffrement**
Tout traitement réalisé doit avoir été défini dans la *Data Clean Room* et approuvé par les fournisseurs de données avant le chargement de ces données.	Le protocole d'échange de clefs repose sur des clefs éphémères. La clef de chiffrement est générée lors du chargement des données, peut être unique à chaque jeu de données provisionné, et devient obsolète une fois les données chiffrées et provisionnées dans la *Data Clean Room*. Le processus de stockage des clefs est le suivant : - l'utilisateur chiffre la clef de chiffrement par une clef provisionnée par l'enclave via un canal sécurisé ; - la clef chiffrée est ensuite envoyée à l'enclave ; - l'enclave reçoit la clef, la déchiffre et la re-chiffre avec une clef uniquement connue de cette enclave particulière ; - enfin, l'enclave stocke la clef chiffrée dans un stockage permanent déployé dans le même *data center* que l'enclave (par exemple Azure CH ou Green.ch).

Auditabilité et traçabilité	Avis juridique ou d'expert, certification
Une trace d'audit log est générée au sein de l'environnement de calcul sécurisé, permettant une auditabilité et une transparence complètes vis-à-vis des actions réalisées dans chaque Data Clean Room. Les erreurs survenant lors de l'exécution des calculs sont remontées directement à l'utilisateur, tout en préservant la confidentialité des données.	- collaborations sur données client entre établissements financiers et éditeurs de presses pour permettre des campagnes marketing sans cookie tiers - conformité au RGPD validé par le cabinet MLL - collaborations sur données de patients entre hôpitaux et labo pharmas - conformité au RGPD validée par avocat expert - collaborations avec garanties de confidentialité en cybersécurité entre institutions financières orchestrée par ArmaSuisse - collaboration pour le financement de transport entre banques d'entreprises et transporteurs maritimes.

Fonctionnalité de mutualisation de données	
Chargement des données	**Mutualisation des données**
Chaque jeu de données d'entrée peut être chargé par un utilisateur accrédité via une interface web ou un SDK Python. Le chiffrement des données est réalisé localement et la clef de chiffrement n'est partagée avec aucune partie, pas même Decentriq. La transmission de la copie chiffrée des données est sécurisée par TLS et réalisée uniquement après identification du serveur comme sûr et fiable au moyen du protocole d'attestation distante du TEE.	Toute requête SQL réalisable via un moteur externe est possible, bien que seule la jointure exacte ait été réalisée sur le PoC. La solution comprend un moteur SQL propriétaire, conforme ANSI et basé sur la syntaxe SQLite, qui couvre un sous-ensemble limité de fonctions afin de fournir des garanties fortes sur les calculs et sur la confidentialité en sortie.
Matching flou	
Les VM Python confidentielles permettent aux utilisateurs de concevoir et mettre en œuvre la logique de matching flou de leur choix. Dans le moteur SQL propriétaire de Decentriq, une fonction permet aussi de réaliser un matching flou.	
Génération de données synthétiques	
Oui, en utilisant l'algorithme PATE-GAN	
Fonction d'exportation	
Oui, en utilisant des filtres de k-anonymat et de la DP	

Fonctionnalités de modélisation (et requête) sur données mutualisées	
Variables d'entrée possibles	**Modélisation avec garanties de confidentialité**
Table des transactions, enrichie de statistiques au niveau du compte client	Modèles à base de règles ou de Machine Learning, utilisant du code Python quelconque.
API	
Les utilisateurs peuvent interagir avec la plateforme Decentriq via soit une interface web soit un SDK Python ou JavaScript.	

5.4 Inpher

➔ **Lien vers la présentation vidéo**

Architecture de la solution	
Périmètre de Confidentialité	**Hébergement**
Le Périmètre de Confidentialité virtuel comprend une machine Inpher XOR installée sur chaque nœud participant, ainsi qu'un service d'orchestration XOR.	*Cloud* ou sur site, tant pour l'orchestrateur que pour les nœuds participants.
Participants et leurs rôles	
- Service d'orchestration XOR : routeur autogéré et sans accès aux données brutes, rendant inutile la présence d'un tiers de confiance - Fournisseurs de données : chaque établissement fournissant son propre jeu de données.	

Modèle de sécurité et gouvernance
Détails techniques (algorithme ou mise en œuvre)
La solution socle Inpher est basé sur des méthodes SMPC mais la plateforme permet plus généralement la mise en œuvre d'autres technologies telles que le FL, la DP ou le (F)HE. Son architecture flexible et fédérée permet aux machines XOR (qui sont des machines virtuelles ou *AMI*) d'être déployées virtuellement sur un environnement quelconque sans nécessiter de matériel spécialisé. Le protocole SMPC utilisé dans la plateforme XOR Secret Computing est basé sur le partage de secret et sur l'approximation de Fourier de fonctions réelles. Pour plus de détails, voir l'article dans Financial Cryptography de 2018 (soumis à *peer review*) ainsi que le papier sur Manticore.
Modèle de menace
Adversaire supposé honnête mais curieux, avec *dealer* (nœud réalisant le partage de secret) de confiance et doté d'un mécanisme de seuil (nombre de participants nécessaire au déchiffrement) complet.
Avis juridique ou d'expert, certification
Recommandations et analyse juridique indépendantes par le CEPD soutenant l'utilisation du SMPC pour la conformité des transferts de données et les collaborations soumises au RGPD.

Fonctionnalité de mutualisation de données	
Chargement des données	**Mutualisation des données**
Les données peuvent demeurer à leur emplacement initial, typiquement dans une zone de confidentialité. Il est toutefois possible (selon les exigences de sécurité, par exemple si un participant ne fait pas confiance au fournisseur de *cloud*) de chiffrer les données et de les charger dans une enclave sécurisée co-localisée avec la machine XOR.	Croisement de données tabulaires, basé sur un traitement SMPC distribué (PSI utilisant l'algorithme OPPRF et matching flou utilisant *MinHash*)
Matching flou	
Oui, basé sur un algorithme MinHash	
Génération de données synthétiques	
Oui, en utilisant la DP	

Fonctionnalités de modélisation (et requête) sur données mutualisées
Modélisation avec garanties de confidentialité
Régressions linéaires et logistiques, XGBoost et k-Means
API
- XOR-py: une bibliothèque Python client permettant d'interagir avec le *backend* XOR. Cette bibliothèque encapsule l'API REST de la plateforme XOR, qui exploite des abstractions de haut niveau pour faciliter son utilisation. Alternativement, les primitives de bas niveau de la machine XOR permettent d'interagir avec l'ensemble des fonctionnalités de la plateforme. - API REST XOR : facilite l'intégration et l'interaction avec d'autres applicatifs via des services web RESTful. - SDL XOR : une bibliothèque de composition de calculs avec garanties de confidentialité, permettant des calculs locaux sur données confidentielles, du SMPC ainsi que du FL avec agrégation sécurisée. - UI XOR : une interface graphique fournissant des fonctionnalités similaires aux API.

5.5 Roseman Labs

➔ **Lien vers la présentation vidéo**

Architecture de la solution	
Périmètre de Confidentialité	**Hébergement**
Le Périmètre de Confidentialité est virtuel : il comprend les *Virtual Data Lakes* (VDL) localisés à chaque nœud participant aux calculs et qui opèrent par partage de secret. La technologie sous-jacente repose sur du SMPC (calculs multipartites sécurisés).	Tout *Cloud*
Participants et leurs rôles	
- Au moins 3 VDL : les nœuds mettant en œuvre le protocole SMPC et participant aux calculs. Plusieurs clients (propriétaires de données ou analystes) peuvent se connecter à un tel nœud. - Fournisseurs de données : chaque établissement avec son propre jeu de données - Un représentant des participants : une personne physique habilitée à autoriser les requêtes	

Modèle de sécurité et gouvernance
Modèle de menace
La confidentialité des données est assurée par un algorithme de partage de secret durant l'intégralité des calculs. Le VDL fournit cette garantie même en présence de clients malicieux (appelés aussi « activement corrompus ») et d'un nœud de SMPC « honnête mais curieux » (on parle alors de SMPC à sécurité passive).
Procédure d'autorisation des requêtes
La couche d'autorisation garantit qu'un client ne peut exécuter que les requêtes explicitement approuvées par un ou plusieurs représentants de participants. Le mécanisme d'autorisation est flexible en ce qu'il permet d'autoriser des requêtes spécifiques ou des classes entières de requêtes via un système de *templates*. La couche d'autorisation fournit un contrôle de sécurité, qui empêche par exemple la fuite de données incontrôlée d'origine interne, où un employé malveillant tente de dérober ou consulter tout un jeu de données d'une autre société.

Fonctionnalité de mutualisation de données	
Chargement des données	**Mutualisation des données**
Réalisé par une procédure cryptographique, de sorte que les données parviennent aux nœuds SMPC exclusivement sous forme de « partage de secret ».	Les requêtes de jointure externe (« *outer join* » en SQL) ont été démontrées au PoC ; d'autres types de jointure exacte ou floue sont disponibles.
Fonction d'exportation	
Sur approbation des participants qu'un certain type de calcul peut être réalisé (et ses résultats révélés)	

Fonctionnalités de modélisation (et requête) sur données mutualisées	
Variables d'entrée possibles	**Modélisation avec garanties de confidentialité**
Table des transactions, ou toute autre donnée tabulaire	Des opérations arithmétiques et SQL classiques, ainsi que des opérations statistiques plus avancées telles que la régression logistique binaire, sont actuellement disponibles. Dans un futur proche, les régressions logistiques multinomiales et ordinales, la régression « *secure ridge* », les réseaux de neurones et les arbres de décision seront aussi disponibles.
API	
Une interface Python permet aux analystes d'interagir avec le VDL comme s'il s'agissait d'une base de données standard.	

5.6 Sarus - Microsoft – EY

➔ **Lien vers la présentation vidéo**

Architecture de la solution	
Périmètre de Confidentialité	**Hébergement**
Le Périmètre de Confidentialité est matérialisé et comprend un *tenant* Microsoft Azure tenant, qui héberge un service *Azure Key Vault Managed HSM* pour la gestion des clefs de chiffrement.	Microsoft Azure
Participants et leurs rôles	
- Un intermédiaire : la solution (incluant le composant de gestion de clefs et le composant Sarus) déployé sur Azure mais sans accès aux données brutes - Fournisseurs de données : chaque établissement avec son propre jeu de données - Les analystes ou data scientists pouvant accéder aux données mutualisées ou aux alertes, le risque de ré-identification ayant été réduit par DP	

Modèle de sécurité et gouvernance
Détails techniques (algorithme ou mise en œuvre)
La confidentialité des données entre l'entité et la base de consolidation repose : - sur la confidentialité et les restrictions d'accès aux clés de chiffrement - sur la confidentialité d'exécution des processus par l'utilisation de l'exécution des processus dans des environnements confidentiels chiffrés et dont l'exécution en clair se fait dans les enclaves processeurs - sur la définition des contrats attestés d'autorisation d'accès aux clés de chiffrement par les processus au sein d'une enclave.
Modèle de menace
- La confidentialité en entrée est couverte par l'utilisation d'une enclave sécurisée, accompagnée de la gestion par le partenaire bancaire de sa propre clef de chiffrement. - L'approche Sarus se concentre sur la sécurité des sorties de tous les calculs réalisés sur les données ainsi protégées. Le modèle de menace associé porte sur les personnes requêtant le service via l'API Sarus. Ces requêtes peuvent venir notamment des utilisateurs (humains ou informatiques) d'un établissement exécutant des requêtes susceptibles d'extraire l'information relative à un client d'autres établissements. Le composant de DP garantit que ce risque est minimal. - Les observables considérées chez l'attaquant sont les données légitimement accessibles par l'API. On ne considère pas les *timing attacks* exploitant une corrélation entre les données et le temps d'exécution des requêtes.

Procédure d'autorisation des requêtes	Gestion des clefs et processus de chiffrement
Toute l'information sortante est conforme à la politique de confidentialité définie par le propriétaire de données. Si la politique la plus stricte est choisie (DP, budget fixe pour un groupe d'utilisateurs), les garanties sont les plus fortes. Ce principe s'applique aussi bien aux échantillons de données synthétiques, aux agrégats et aux poids des modèles, mais une politique trop stricte peut rendre certaines tâches trop bruitées. L'utilisateur peut délibérément autoriser certaines requêtes dérogatoires (accès à la performance des modèles, aux poids des modèles, inférence sur un individu) pour certains utilisateurs.	Les données au repos seront chiffrées afin d'assurer leur protection en fonction de la solution de stockage utilisée : - Les données fournies par les sources restent chiffrées (clef de chiffrement du fournisseur) et stockées dans le boîtier HSM sous contrainte de la politique du consortium pour son utilisation dans la zone de stockage. - Un compte de stockage Azure avec le chiffrement de service de stockage (*Storage Service Encryption* ou SSE) activé par défaut (et ne pouvant être désactivé) où les données sont systématiquement chiffrées en AES 256 offrant une conformité avec le standard FIPS 140-2. Ces clés AES sont scellées en utilisant une clé asymétrique stockée dans le boîtier HSM géré par le consortium d'administration de la plateforme, qui fournira sa propre clef au lieu d'une clef par défaut générée par le fournisseur *cloud*. Cette configuration est appelée BYOK (*Bring Your Own Key*).

Fonctionnalité de mutualisation de données	
Chargement des données	**Mutualisation des données**
Le jeu de données de chaque Établissement Participant est chiffré localement puis chargé sur une instance confidentielle du *cloud* Azure.	Croisement de données tabulaires, réalisé en clair
Matching flou	
Oui, sur toutes données derrière l'API. La garantie de confidentialité est préservée sur tous les résultats provenant de l'API, y compris après du matching.	
Fonction d'exportation	
Oui, pour tous les résultats conformes à la politique de confidentialité (DP par défaut), y compris les données synthétiques.	

Fonctionnalités de modélisation (et requête) sur données mutualisées	
Variables d'entrée possibles	**Modélisation avec garanties de confidentialité**
La table des transactions, incluant les deux pattes (émetteur et bénéficiaire)	Les bibliothèques de Machine Learning standard sont gérées : - numpy and pandas quasiment dans leur intégralité - scikit-learn, TensorFlow et XGBoost partiellement
API	
L'API Sarus permet d'interagir avec les données de deux façons : - en soumettant des requêtes SQL dans le même dialecte que la technologie SQL retenue (en l'occurrence T-SQL d'Azure SQL dans cette démonstration) - en soumettant du code python standard via le SDK Sarus.	

5.7 Scalnyx

➔ **Lien vers la présentation vidéo**

Architecture de la solution	
Périmètre de Confidentialité	**Hébergement**
Le Périmètre de Confidentialité est la plateforme SCALTRUST sur laquelle les jeux de données d'entrée sont chargés après chiffrement.	*Cloud* public (OVH, bientôt EXAION) ou privé
Participants et leurs rôles	
- Fournisseur de service centralisé : le serveur central, déployé et maintenu par un tiers de confiance pour l'expérimentation. Le serveur central est autonome et fournit les services exécutés au sein du Périmètre de Confidentialité (c'est-à-dire dans le domaine chiffré dans lequel toutes les données restant cryptées et où aucune clef de déchiffrement homomorphe n'est accessible). Les deux services proposés sont d'une part les requêtes FHE (conformes au standard SQL), d'autre part l'utilisation conjointe de FHE et FL (voire DP). - Fournisseurs de données : chaque établissement provisionne son jeu de données dans SCALTRUST, utilise l'un des deux services pour réaliser les calculs et peut mettre à jour ses propres données périodiquement.	

Modèle de sécurité et gouvernance
Détails techniques (algorithme ou mise en œuvre)
L'agrégateur FL utilisé dans SCALTRUST réalise uniquement des calculs sur les paramètres de modèles. Ces paramètres sont les moyennes et déviations standard de variables pour les modèles causaux de type bayésien naïf, ou alternativement les poids et gradients pour les modèles de type réseau de neurones. (À l'inverse, une approche FHE standard au Machine Learning nécessite de calculer directement sur les données chiffrées l'ensemble des paramètres requis par le modèle, ce qui est exorbitant en temps de calcul dans le domaine FHE). Les bibliothèques cryptographiques utilisées par SCALTRUST sont BFV, SEAL et Additive.
Modèle de menace
- L'agrégation FHE est robuste au cas d'un adversaire « honnête mais curieux » mais des enjeux de confidentialité des modèles demeurent en sortie (un problème annexe pour l'expérimentation LCB-FT car le modèle n'est pas censé être partagé avec les pairs). - Dans le cas d'un adversaire malveillant, aux questions de confidentialité des sorties s'ajoutent des questions d'intégrité du FL. - Le FHE présuppose l'absence de collusion entre nœuds participant aux calculs. - Chaque participant étant en possession de la clef de déchiffrement FHE, ils ne doivent en aucun cas avoir accès au serveur central utilisé pour traiter les données dans le domaine FHE, au risque de compromettre leur confidentialité. En d'autres termes, le fournisseur de service central ne doit pas être un participant aux calculs.
Gestion des clefs et processus de chiffrement
La solution proposée ne traite pas le commissionnement des clefs de chiffrement : tous les Établissements Participants sont supposés partager les mêmes clefs via un protocole sécurisé de distribution de clefs, et les clefs supposées être générées localement par un participant élu par ses pairs.

Fonctionnalité de mutualisation de données	
Chargement des données	**Mutualisation des données**
Les données confidentielles de chaque Établissement Participant sont envoyées directement au serveur central, puis traitées par ce même serveur sans autre interaction avec les différents nœuds.	Mutualisation de modèle – et non de données – au moyen du FL (le FHE étant utilisé pour l'agrégation sécurisée des modèles)
Génération de données synthétiques	
Oui, via le modèle d'IA causale	
Fonction d'exportation	
Également possible par modèle d'IA causale	

Fonctionnalités de modélisation (et requête) sur données mutualisées	
Variables d'entrée possibles	**Modélisation avec garanties de confidentialité**
Soit les variables communes à tous les participants (FL horizontal) soit l'union des ensembles de variables de chaque participant (FL vertical)	Modèles causaux de type bayésien. Plusieurs publications scientifiques ont montré l'applicabilité de l'IA causale aux scénarios de détection de fraude, néanmoins à notre connaissance le seul produit commercial ayant adopté ce type d'approche (Hugin AML) ne remplit pas les critères de confidentialité du Tech Sprint MCD.
API	
La plateforme fournit 3 types d'environnement : - Un SDK facilitant l'intégration de la plateforme SCALTRUST dans un environnement de développement standard (par exemple en Python) - Un module de développement « no code » permettant de concevoir des applications préservant la confidentialité en assemblant les briques de calcul disponibles, sans nécessiter d'expertise de programmation ni de cryptographie - Des outils de développement FHE pour assister les programmeurs à mettre en œuvre de nouvelles briques de calcul dans le domaine FHE. Elle inclut aussi un outil de construction de requêtes FHE conforme au standard SQL.	

5.8 Secretarium – FutureFlow

➔Lien vers la présentation vidéo

Architecture de la solution	
Périmètre de Confidentialité	**Hébergement**
La couche de confidentialité ou *Secretarium Confidentiality Layer* (SCL) est un composant assurant la sécurité des données de bout en bout (c'est-à-dire en transit, au repos et en cours de traitement). La SCL stocke les données dans des registres inviolables et permet des analyses par des modèles à base de règles ou de Machine Learning.	*Cloud* public (OVH, SwissCom) ou sur site
Participants et leurs rôles	
- La SCL incluant un TEE - Fournisseurs de données : chaque établissement fournissant son propre jeu de données chiffré à la SCL.	

Modèle de sécurité et gouvernance	
Détails techniques (algorithme ou mise en œuvre)	
Secretarium utilise une enclave sécurisée offrant une fonction d'attestation distante. Le protocole conçu pour la solution (*Secretarium Connection Protocol*) requiert ainsi une signature d'enclave comprenant à la fois l'empreinte de l'enclave et différents éléments garantissant que du matériel sécurisé et revu par Secretarium est employé. Les Établissements Participants ont alors l'assurance qu'ils communiquent avec un service distant approuvé et fiable.	
Modèle de menace	
La solution réalise une dé-identification au sein de la SCL, ce qui réduit le risque de ré-identification accidentelle ou délibérée des données partagées par un établissement avec ses pairs via la plateforme. La SCL fournit des attestations de code indiquant aux établissements que seul du code pré-approuvé sera exécuté sur les données qu'ils ont fournies.	
Procédure d'autorisation des requêtes	**Gestion des clefs et processus de chiffrement**
Non, sauf dans la revue manuelle post-suspicion	Une clef secrète est générée et scellée dans l'enclave sécurisée, invisible à quiconque y compris aux Établissements Participants.
Avis juridique ou d'expert, certification, déploiements existants	
- Solution éprouvée depuis 2018 dans l'initiative DANIE - Participation au bac à sable réglementaire de l'ICO (autorité de protection des données britannique) - La méthode d'analyse a été appliquée lors du FCA AML TechSprint en 2019 - La méthode d'analyse a été appliquée sur des données réelles lors du pilote Deloitte TriBank en 2019 - Déjà audité de grandes institutions financières - Audité aussi par une société indépendante de sécurité informatique - Cas d'usage de réconciliation de données déployé en production	

Fonctionnalité de mutualisation de données	
Chargement des données	**Mutualisation des données**
Les données sont soumises avant chargement à de multiples validations, vérifications d'intégrité et transformations afin d'assurer la qualité des réconciliations.	Réconciliation sur la base de structures en réseau
Matching flou	
La solution est capable de détecter à la volée les erreurs dans les données et de réaliser un matching flou, ce qui contourne un obstacle majeur des méthodes de type hachage unidirectionnel aveugle coordonné, où la contrepartie centrale ne peut identifier les erreurs d'intégrité des données. Afin d'éviter l'automatisation de matching flou incorrects, les comptes imparfaitement réconciliés sont proposés aux institutions financières dans des enceintes de collaborations anonymes et sécurisées pour des confirmations manuelles.	

Fonction d'exportation
Exportation pseudonymisée, dont toute information identifiante est absente mais comporte néanmoins en théorie un risque résiduel de ré-identification

Fonctionnalités de modélisation (et requête) sur données mutualisées	
Variables d'entrée possibles	**Modélisation avec garanties de confidentialité**
Les données d'entrée sont les transactions pseudonymisées par le SCL. Ces données sont transformées en un ensemble de graphes des comptes clients dé-identifiés (enrichis de valeurs agrégées sur l'ensemble des transactions). Les variables d'entrée des modèles de Machine Learning sont les propriétés formelles de réseau de chaque compte.	La technologie FutureFlow est basée sur des méthodes d'analyse non supervisée. Les relations entre comptes bancaires sont analysés comme graphes, sans toutefois utiliser de base de données graphe : les données sont stockées dans une base de données non relationnelle et les graphes sont générés à la volée en fonction des analyses nécessaires. Les modèles d'apprentissage supervisé de type LightGBM sont aussi supportés en prenant en entrée les statistiques agrégées d'alertes, fournissant ainsi une possibilité d'évaluation *a posteriori* des résultats de l'analyse non supervisée.
API	
Aucune API n'est fournie, toutefois la solution permet de visualiser et de parcourir le graphe des comptes clients, en mettant en exergue les comptes de chaque établissement et de ses pairs, les alertes et indicateurs de suspicion, et avec la possibilité d'ouvrir une investigation à tout moment.	
Revue manuelle	
Cette solution est la seule parmi les solutions du Tech Sprint fournissant un processus robuste de revue manuelle, qui peut être utilisé pré-analyse lors de l'étape de matching flou, ainsi que post-analyse en vue d'une investigation conjointe sur un ensemble de comptes clients suspects, au besoin en partageant des données sensibles dans une phase post-suspicion. Les investigations sont menées au sein d'une enceinte sécurisée fournie par la SCL, initiée anonymement par un établissement participant. - Lors de la réconciliation, et une fois que tous les participants invités ont accepté de rejoindre l'enceinte sécurisée, la donnée imparfaitement réconciliée (matching flou) leur est révélée et chacun peut voter sur le résultat de cette réconciliation. En cas de consensus positif, la SCL retraite l'ensemble des transactions mutualisées en question et recalcule les métriques utilisées dans l'analyse. - De la même manière, une investigation post-analyse peut être ouverte dans une autre enceinte sécurisée sur un ensemble de comptes suspicieux, dans laquelle les institutions financières, si nécessaire et si elles le souhaitent, peuvent partager des informations sensibles.	

5.9 SnowPack – Sphinx

➔Lien vers la présentation vidéo

Architecture de la solution	
Périmètre de Confidentialité	**Hébergement**
Le Périmètre de Confidentialité comprend : - le client Sphinx installé dans chaque Établissement Participant (qui stocke les transactions pseudonymisées) ; - plusieurs (>=2) serveurs « aveugles » Sphinx sur une infrastructure couvrant plusieurs *data centers* ; - la « surcouche réseau d'invisibilité » de Snowpack (SNO pour *Secure Overlay Network*), utilisée pour fragmenter les paquets HTTPS et faire transiter les fragments par des routes réseau différentes.	De tout type : Linux ou Windows, conteneur LXC ou Docker, ou une VM quelconque.
Participants et leurs rôles	
- Plusieurs (>=2) intermédiaires, chacun étant un serveur aveugle Sphinx qui relaie au SNO des requêtes et réponses (et réalise aussi la surveillance des performances, le *reporting* sur l'utilisation ainsi que d'autres tâches opérationnelles) - Fournisseurs de données : chaque établissement avec son propre jeu de données pseudonymisé par un client Sphinx	

Modèle de sécurité et gouvernance
Détails techniques (algorithme ou mise en œuvre)
La technologie Snowpack s'appuie sur 5 années de R&D. Elle est protégée par 3 brevets pour lesquels elle détient un accord de licence exclusive quasiment mondial. SNO est opérationnel et est déployé dans 4 pays européens, sur 30 serveurs de 4 opérateurs clouds différents. La technologie Sphinx s'appuie sur 3 ans de travaux d'innovation au sein du projet Cybersec4Europe et exploite plusieurs technologies *PET* (SMPC, chiffrement homomorphe, confidentialité différentielle combinée à l'utilisation de données synthétiques).
Modèle de menace
Snowpack gère le modèle de menace le plus fort de l'ensemble des solutions du Tech Sprint : - secret parfait au niveau de chaque nœud du réseau (l'adresse IP et le chemin utilisé sont masqués par secret-partagé des nœuds participants et de chaque serveur) ; - architecture réseau « en anneau » fournissant des garanties de sécurité et d'intégrité ainsi qu'une preuve d'origine (c'est-à-dire que la donnée provient d'un membre du réseau) ; - résistance aux attaques « man in the middle » grâce à la distribution de ses nœuds de sortie, ne requérant ainsi aucun tiers de confiance au niveau du réseau. Sphinx protège la propriété des données au travers d'une architecture décentralisée en terme de stockage et de traitement, et apporte des mesures techniques (chiffrement des données à l'usage au niveau de la couche applicative) et organisationnelles (protocole de partage des données, processus de gestion des clés de chiffrement) pour se conformer au RGPD et au secret bancaire.
Gestion des clefs et processus de chiffrement
Un client Sphinx pseudonymise les données chargées en utilisant 3 secrets : - Sel de hachage : utilisé pour le hachage des variables de jointure avant chiffrement à des fins de mutualisation (partagé uniquement avec les autres participants et non avec le serveur Sphinx). - Clef de chiffrement privée : utilisée pour le chiffrement des variables de jointure après hachage et avant chiffrement (spécifique à chaque participant, stockée localement, et qui peut être confiée à un gestionnaire de clefs de chiffrement). - Clef de chiffrement partagée : utilisée pour chiffrer les résultats d'analyse confidentiels avant transit ou transfert sur le réseau (partagée avec tous les participants. Les secrets partagés (sel pour le hachage, clé de chiffrement partagée) sont créés, partagés la première fois, puis renouvelés au travers d'un protocole exploitant le protocole d'échange de clés de Diffie-Hellman ainsi que le schéma de partage de secret de Shamir au travers de Snowpack pour prévenir toute interception.

Fonctionnalité de mutualisation de données	
Chargement des données	**Mutualisation des données**
Les transactions de chaque Établissement Participant sont pseudonymisées et chargées sur le client Sphinx local. Une fois chargées, les données sont automatiquement disponibles pour réaliser des calculs confidentiels avec les données des autres participants	- Croisement de données tabulaires (basé sur du PSI pour le rapprochement de données et du chiffrement homomorphe pour des calculs sur les données rapprochées). - Croisement de données en masse exploitant des structures de données compressées et probabilistes

Fonctionnalités de modélisation (et requête) sur données mutualisées
Variables d'entrée possibles
Durant le chargement des données, chaque établissement choisit les données limitées aux opérations de jointure et celle qui peuvent être exploitées à des fins de prédiction. En fonction de ce choix, chaque information subit un traitement (chiffrement, inclusion dans une structure de données anonymisée, …) qui permet soit de l'impliquer uniquement sur des opérations simples telles que des jointures, soit sur des opérations plus complexes consistant à l'exploiter comme donnée d'entrée pour l'entraînement d'un modèle ou comme donnée permettant de confirmer/infirmer les prédictions dudit modèle.

5.10 ThetaRay – Duality

➔ **Lien vers la présentation vidéo**

Architecture de la solution	
Périmètre de Confidentialité	**Hébergement**
Duality propose deux architectures différentes : un déploiement en mode *hub* et un déploiement pair-à-pair. Le mode *hub* offre plusieurs avantages (élasticité vis-à-vis du nombre de participants, agrégation anonymisée des résultats, calculs à valeur ajoutée et analyses au niveau du réseau). Aussi un déploiement en *hub* est-il recommandé dans les cas où plus de deux participants collaborent. Le Périmètre de Confidentialité comprend un nœud Duality déployé au sein de chaque Établissement Participant, ainsi que le *hub* central éventuel.	Linux OS (soit Ubuntu 20.04 soit RHEL 7.8) avec machine Docker
Participants et leurs rôles	
- Chaque Établissement Participant exécute un SecureNode, contenant tous les composants nécessaires à l'exécution des tâches collaboratives. Chaque participant a donc un rôle potentiel de propriétaire de données (visualisation de ses propres données, définition du projet d'analyse collaborative et des validations/permissions nécessaires) ; d'analyste de données (optimisation, déploiement et évaluation des modèles de Machine Learning ou des requêtes SQL) ; voire de propriétaire de modèles (intégration dans la plateforme d'un modèle pré-entraîné). Chaque participant peut exercer un nombre quelconque de ces rôles. - Le *hub* central représente une tierce partie agissant comme orchestrateur (configuration du projet collaboratif, définition des rôles de chacun, distribution des calculs chiffrés aux propriétaires de données respectifs, agrégation des résultats et envoi aux analystes) et optionnellement aussi comme nœud de calcul sur les données mises en œuvre. Le *hub* n'est jamais exposé à des données confidentielles, si bien que toute entité peut jouer ce rôle (Établissement Participant, fournisseur de *cloud*, Duality, etc.)	

Modèle de sécurité et gouvernance	
Détails techniques (algorithme ou mise en œuvre)	
La plateforme exploite la bibliothèque de chiffrement totalement homomorphe OpenFHE, open-source et conforme aux standards, et la combine à des fonctionnalités de data science pouvant être invoquées par une interface graphique ou par API. Les requêtes préservant la confidentialité renvoient leurs résultats en un temps de l'ordre de quelques secondes. Duality participe aussi à HEBench.org, initiative de *benchmarking* permettant de comparer les performances de différentes mises en œuvre du chiffrement homomorphe.	
Modèle de menace	
Le type de chiffrement employé – de nature probabiliste et robuste post-quantique – garantit que chaque opération de chiffrement produit une représentation apparaissant à un adversaire potentiel comme différente des opérations précédentes, ce qui prémunit des attaques « man in the middle » (dans le cas de l'expérimentation ACPR, cette garantie signifie qu'un acteur malveillant ne peut corrompre le résultat qui sera présenté à un analyste LCB-FT). Cette approche est aussi résistante aux attaques menées à partir de machines quantiques.	
Procédure d'autorisation des requêtes	**Gestion des clefs et processus de chiffrement**
Les propriétaires de données sont en mesure de définir et gérer les permissions spécifiques de chaque participant vis-à-vis des requêtes exécutées. Par exemple, le type de requêtes, le type de réponses (numériques, binaires, etc.), ou encore le nombre de requêtes possibles pour un participant ou un rôle donné (analyste ou manager). En particulier, la définition appropriée de la structure des requêtes et des réponses permet de préserver la confidentialité des données et de prévenir toute fuite d'informations involontaire.	La plateforme expose une API *Data and Key Management*, qui fournit des méthodes de stockage de données et de clefs, de recherche de données, ainsi que de gestion des clefs de chiffrement.

Auditabilité et traçabilité
Un nœud Duality permet les calculs chiffrés avec une tenue en charge importante, le stockage et le transfert des données, la mise en œuvre de restrictions sur l'accès aux données, et l'audit de l'ensemble des opérations.

Fonctionnalité de mutualisation de données
Mutualisation des données
Les jeux de données peuvent être chiffrés chez le propriétaire de données, puis provisionnés via par jointure (c'est-à-dire des croisements) ou empilement (c'est-à-dire des unions). Ce mode de collaboration nécessitait, à date de la réalisation du Tech Sprint, un schéma de données commun approuvé par les participants.
Matching flou
La plateforme Duality offre différents types de prétraitement, incluant maching flou et normalisation : - Prétraitement ouvert : les utilisateurs peuvent charger tout script Python, les propriétaires de données conservant tout contrôle sur l'exécution du script, avec une fonctionnalité d'approbation ou de rejet de tout processus de prétraitement de leurs données - Normalisation de texte, afin d'améliorer le taux de croisement via différents algorithmes – standard ou propriétaires – de normalisation et de résolution d'entités - Chacune de ces fonctionnalités peut être combinée avec des approches heuristiques.
Génération de données synthétiques
Possible via l'utilisation de DP

Fonctionnalités de modélisation (et requête) sur données mutualisées
Modélisation avec garanties de confidentialité
Duality gère un certain nombre de modèles de Machine Learning, incluant les GLM (*Generalized Linear Models*) et GBT (*Gradient Boosted Trees*). Les type de requêtes à la demande gérés par la solution incluent des requêtes pandas arbitraires, des requêtes SQL pré-approuvées, ou de l'analyse (statistique ou à base de Machine Learning) chiffrée.
API
- Chaque nœud Duality est basé sur les capacités de calculs homomorphes de la bibliothèque OpenFHE, étendues par l'API *Computation*. Cette API facilite l'accès aux fonctions homomorphes, elle est utilisable par des programmeurs expérimentés pour le développement de nouvelles méthodes ou modules de data science. - Une autre API exposée est l'API *Data and Key Management*, qui fournit des méthodes de stockage et extraction de données et de clefs, de recherche de données, et de gestion des clefs de chiffrement. - La couche data science expose des fonctionnalités de data science incluant de l'inférence statistique, les modèles GBT et GLM, etc.

5.11 TripleBlind - Accenture

➔[Lien vers la présentation vidéo]

Architecture de la solution	
Périmètre de Confidentialité	**Hébergement**
Le Périmètre de Confidentialité est nommé *Virtual Data Pool* et comprend le point d'accès TripleBlind installé sur chaque nœud participant, ainsi qu'un nœud de routage TripleBlind.	*Cloud* public (GCP, AWS) ou sur site
Participants et leurs rôles	
- Routeur TripleBlind : n'a aucune façon d'accéder aux données, ce qui dispense de faire appel à un tiers de confiance - Fournisseurs de données : chaque établissement fournissant son propre jeu de données.	

Modèle de sécurité et gouvernance
Détails techniques (algorithme ou mise en œuvre)
La solution de calculs distribués TripleBlind ne requiert pas de transférer les données hors de leur emplacement initial. Lorsqu'une analyse est approuvée, les étapes suivantes sont effectuées : - Les données sont chiffrées au niveau de l'octet - L'algorithme d'analyse ou le modèle à entraîner est également chiffré - Ces éléments sont alors fragmentés, chaque fragment étant distribué à l'ensemble des participants aux calculs, qui sont réalisés sur les données toujours chiffrées - Les résultats sont combinés entre eux et seule la sortie agrégée est fournie par la solution.
Modèle de menace
La solution TripleBlind réside dans le data center de chaque Établissement Participant et peut donc se conformer à toute politique de cybersécurité ou de conformité. Les fonctionnalités de sécurité incluent les suivantes : - Une fois les analyses approuvées par les propriétaires de données, TripleBlind produit des rapports d'analyse exploratoire des jeux de données mutualisés par chaque participant afin d'en faciliter la compréhension par les parties prenantes. Ces rapports ne montrent aucunement les données brutes, mais fournissent des métadonnées associées à un jeu de données. - Des connexions sont établies entre participants à l'analyse collaborative sur des ports sécurisés et via des protocoles de transmission eux-mêmes chiffrés et validés. - Le point d'accès TripleBlind est durci pour résister aux attaques tierces, de plus il est déployé dans un périmètre lui-même généralement sécurisé (dans le cas de l'expérimentation ACPR, derrière le pare-feu d'une institution financière).
Gestion des clefs et processus de chiffrement
La solution est basée sur des méthodes de partage de secret, rendant superflu tout système de gestion de clefs de chiffrement.
Avis juridique ou d'expert, certification
- Le Comité européen de la protection des données (CEPD) considère que les calculs multipartites sécurisés avec partage de secret (*split processing*) fournissent « une mesure supplémentaire effective » pour sécuriser les données (Recommandations, 28 juin 2021, cas d'usage 5). - Les solutions de SMPC ont été soumises à une évaluation technique indépendante par le MITRE (*Massachusetts Institute of Technology Research and Engineering*).

Fonctionnalité de mutualisation de données	
Chargement des données	**Mutualisation des données**
Les Établissements Participants spécifient l'emplacement de leurs jeux de données : ceux-ci sont alors provisionnés sur un point d'accès donné, mais les données brutes ne quittent jamais leur emplacement initial.	La mutualisation est uniquement effectuée sur demande. Aucun schéma de données commun aux participants n'est requis.
Génération de données synthétiques	
Possible en utilisant la DP ou des GAN.	

Fonctionnalités de modélisation (et requête) sur données mutualisées	
Variables d'entrée possibles	**Modélisation avec garanties de confidentialité**
L'apprentissage horizontal ou vertical est possible. Pour le PoC, un modèle non supervisé (de type *embedding* de graphe de connaissances) a été entraîné sur les jeux de données partitionnés horizontalement et filtrés par requêtes SQL.	Forêts aléatoires, Ampligraph (*embedding* de graphe de connaissances) et l'intégralité de la bibliothèque scikit-learn
API	
TripleBlind est programmable via de multiples bibliothèques Python de Machine Learning (scikit-learn, XGBoost, PyTorch, Keras, Tensorflow, Ampligraph, etc.)	

5.12 Tune Insight

➔ **Lien vers la présentation vidéo**

Architecture de la solution	
Périmètre de Confidentialité	**Hébergement**
Le Périmètre de Confidentialité comprend l'agent Tune Insight, installé sur chaque nœud participant et communiquant avec les autres participants (en utilisant la méthode MHE c'est-à-dire le « chiffrement multipartite homomorphe ») et avec le service d'authentification.	Local ou sur le Cloud, dans toute VM avec machine Docker
Participants et leurs rôles	
- Fournisseurs de données : chaque établissement avec son propre jeu de données. Un agent Tune Insight est installé chez chaque fournisseur et connecté à leur base de données. - Un service central d'authentification et d'autorisation, connecté au service d'identité local ou fédéré pour les calculs collaboratifs et garantissant donc que les acteurs participant au réseau peuvent définir leurs propres règles d'autorisation en fonction de leurs besoins spécifiques, et gérer leurs propres utilisateurs et les rôles associés.	

Modèle de sécurité et gouvernance	
Détails techniques (algorithme ou mise en œuvre)	
Les calculs multipartites sont mis en œuvre sur la base de la bibliothèque FHE *open source* Lattigo. Lorsqu'un calcul est lancé, les étapes suivantes ont lieu : 1. La requête est transmise à chaque agent participant et validée par ces derniers. Une trace d'audit est alors produite. 2. Chaque agent charge ses données d'entrée et enregistre la requête ainsi que des métadonnées associées à ces données. 3. Les agents exécutent un calcul MHE : a. Les clés cryptographiques nécessaires sont générées collectivement. Typiquement, un agent crée une clé publique collective et chaque agent détient une partie de la clé privée associée. Ces clés peuvent être éphémères ou statiques (soumises à renouvellement périodique selon les politiques de sécurité des établissements participants). b. Les jeux de données locaux sont chiffrés avec la clé publique collective et le résultat est calculé dans le domaine homomorphe. c. Le résultat chiffré est re-chiffré collectivement avec la clé publique de l'agent propriétaire des données ou de l'utilisateur autorisé (en fonction du cas d'usage). d. Le résultat peut être déchiffré et éventuellement utilisé par l'utilisateur comme données d'entrée pour de nouveaux traitements.	
Procédure d'autorisation des requêtes	**Gestion des clés et processus de chiffrement**
Oui, en utilisant une procédure de contrôle d'accès par attributs et des protocoles d'accès automatisés ou semi-automatisés, intégrés dans le calcul MHE.	L'agent Tune Insight peut être configuré pour se connecter à des services de gestion de clés (KMS) tels que Microsoft Azure Key Vault afin de récupérer les identifiants nécessaires à l'accès local et au stockage en base de données.
Auditabilité et traçabilité	
Les agents Tune Insight emploient une base de données immuable capable de validations cryptographiques (utilisant des arbres de Merkel), ce qui permet de stocker les traces d'audit localement avec des garanties fortes d'intégrité tout en évitant la complexité des solutions Blockchain. Entre autres fonctionnalités, toute création ou modification de *logs* applicatifs est systématiquement enregistrée en base de données, fournissant une piste d'audit détaillée capturant l'ensemble des actions réalisées au cours du processus (accès aux données mutualisées, entraînement de modèles, protocole d'agrégation, etc.)	

Fonctionnalité de mutualisation de données	
Chargement des données	**Mutualisation des données**
L'étape de chargement des données consiste en une simple connexion de l'agent Tune Insight déployé localement à la source de données. Pour éviter tout stockage d'identifiants chez l'agent, la solution permet aussi une connexion à un KMS.	Mutualisation distribuée multipartite (basée sur PSI – *Private Set Intersection*). Les données sont mutualisées virtuellement, car les données brutes ne sont pas transférées. Une autre fonctionnalité est le calcul de statistiques partitionnées verticalement sur les données (virtuellement) mutualisées.

Fonctionnalités de modélisation (et requête) sur données mutualisées	
Variables d'entrée possibles	**Modélisation avec garanties de confidentialité**
Variables communes (dans le cas de jeux de données empilés) ou union des variables d'entrée de chaque participant (dans le cas de jointures entre jeux de données)	La fonctionnalité MHE de chaque agent Tune Insight permet les analyses suivantes : - Machine Learning chiffré collectivement pour entraîner des régressions linéaires et des perceptrons multicouches (MLP), ainsi que des solutions hybrides chiffrées et avec confidentialité différentielle pour l'entraînement de modèles complexes ; - protocoles de croisement tels que PSI et statistiques sur des données partitionnées verticalement ; - statistiques agrégées chiffrées ; - inférence pour de nombreux types de modèles de Machine Learning (GLM, SVM, CNN), en protégeant à la fois les données et le modèle ; - récupération d'information confidentielle (PIR ou *Private Information Retrieval*), en protégeant la requête.
API	
Trois modules permettent d'interagir avec la solution Tune Insight : - chaque agent Tune Insight agent expose aux clients locaux une API RESTful HTTPS et peut communiquer avec d'autres agents participants - l'application web Tune Insight fournit une interface graphique pour initier aisément des calculs collaboratifs et des analyses sécurisées - le SDK Python Tune Insight facilite l'intégration des fonctionnalités Tune Insight dans les processus habituels des data scientists.	

6. Glossaire

Confidentialité en entrée (input privacy): ensemble de garanties concernant les données en entrée d'un flux d'information ou d'un processus. Voir confidentialité en sortie pour un concept lié et complémentaire.

Confidentialité en sortie (output privacy) : ensemble de garanties concernant les données en sortie d'un flux d'information ou d'un processus. Voir confidentialité en entrée pour un concept lié et complémentaire.

DP (Differential Privacy, Confidentialité différentielle): l'une des familles de technologies possibles pour réaliser la MCD. La confidentialité différentielle, en bases de données et parfois associé à la cryptographie, est une propriété d'anonymisation pouvant être atteinte via différents mécanismes. Celle-ci vise à définir une forme de protection de résultats de requêtes faites à une de bases de données en minimisant les risques d'identification des entités qu'elle contient, si possible en maximisant la pertinence des résultats de la requête. Elle permet notamment l'exploitation statistique de données individuelles agrégées, sans compromettre la vie privée des individus. (Source: Wikipedia.)

Établissement (participant) : ce terme désigne l'ensemble des entités ayant des obligations en matière LCB-FT et ayant choisi de participer à l'expérimentation ACPR et donc à l'étape clef du Tech Sprint MCD. Ces entités sont principalement des établissements de crédit, avec aussi quelques organismes d'assurance. Ces établissements sont regroupés en équipes travaillant chacune sur un cas d'usage. Au sein d'une équipe donnée chaque établissement reste libre de définir son protocole expérimental et le fournisseur de technologie le plus approprié, ce dernier point faisant l'objet du Tech Sprint MCD.

FL (Federated Learning, apprentissage fédéré): l'une des familles de technologies possibles pour réaliser la MCD. En intelligence artificielle et en apprentissage machine, l'apprentissage fédéré est une méthode ou un paradigme qui consiste à entraîner un algorithme sur la machine des utilisateurs d'une application et à partager les apprentissages réalisés sur la machine de chaque utilisateur. Cette méthode s'oppose à l'apprentissage centralisé où l'apprentissage se fait sur les serveurs du fournisseur de service. (Source: Wikipedia.)

GAN (Generative Adversarial Network, réseau antagoniste génératif) : En intelligence artificielle, les réseaux antagonistes génératifs, parfois aussi appelés réseaux adverses génératifs, sont une classe d'algorithmes d'apprentissage non supervisé. Un GAN est un modèle génératif où deux réseaux sont placés en compétition dans un scénario de théorie des jeux. Le premier réseau est le générateur, il génère un échantillon, tandis que son adversaire, le discriminateur, essaie de détecter si un échantillon est réel ou bien s'il est le résultat du générateur. Ainsi, le générateur est entrainé avec comme but de tromper le discriminateur. L'apprentissage peut être modélisé comme un jeu à somme nulle. (Source: Wikipedia.)

GBT (Gradient-Boosted Tree): Le *gradient boosting* est une technique de Machine Learning technique utilisée entre autres pour les tâches de régression et de classification. Elle produit un modèle prédictif composé d'un ensemble de modèles de prédiction dit « faibles », typiquement des arbres de décision auquel cas l'algorithme résultant est appelé *gradient-boosted trees*.

GLM (Generalized Linear Model, modèle linéaire généralisé) : en statistiques, le modèle linéaire généralisé est une généralisation souple de la régression linéaire. Le GLM généralise la régression linéaire en permettant au modèle linéaire d'être relié à la variable réponse via une fonction lien et en autorisant l'amplitude de la variance de chaque mesure d'être une fonction de sa valeur prévue, en fonction de la loi choisie. (Source: Wikipedia.)

HE (Homomorphic Encryption, Chiffrement homomorphe): l'une des familles de technologies possibles pour réaliser la MCD. En cryptographie, un algorithme de chiffrement homomorphe est un système possédant des caractéristiques algébriques qui lui permettent de commuter avec certaines opérations mathématiques, c'est-à-dire qu'il permet d'opérer un obscurcissement sur des valeurs numériques tout en conservant les propriétés permettant lesdites opérations. Le déchiffrement du résultat desdites opérations sur des données chiffrées donnera ainsi le même résultat qu'en ayant effectué ces opérations sur les données non chiffrées ; cette

propriété permet de confier des calculs à un agent externe, sans que les données ni les résultats soient accessibles à cet agent. (Source: Wikipedia.)

Honnête mais curieux (adversaire): également appelé semi-honnête, l'adversaire honnête mais curieux est un participant autorisé à un protocole de communication distribué, qui ne va pas déroger au protocole prédéfini à ceci près qu'il tentera d'extraire autant d'information que possible des messages légitimement reçus. (Source: *Modelling and Automatically Analysing Privacy Properties for Honest-but-Curious Adversaries*, Andrew Paverd et al.)

KMS (Key Management System, système de gestion de clefs): désigne un système de sécurité, distribué ou centralisé, permettant de gérer la création, la distribution et la maintenance des clefs de chiffrement utilisées dans un système cryptographique.

LCB-FT (Lutte Contre le Blanchiment et le Financement du Terrorisme): le secteur financier est exposé au risque de blanchiment des capitaux et de financement du terrorisme. À ce titre, il est assujetti à des dispositions en matière de lutte contre le blanchiment des capitaux et le financement du terrorisme (LCB-FT). (Source: ACPR.)

MCD (Mutualisation Confidentielle de Données) : ce terme (invité par l'ACPR dans le cadre de l'expérimentation) vise à décrire toute technologie permettant le stockage, le croisement, les requêtes, et l'alimentation de modèles LCB-FT à partir de plusieurs jeux de données dont la confidentialité doit être préservée, c'est-à-dire tout type de données – à commencer par les transactions financières – auxquelles certaines parties prenantes à l'expérimentation associent des exigences de confidentialité et éventuellement d'intégrité. L'ACPR avait anticipé un large recouvrement du champ de la catégorie MCD avec le terme *PET* (*Privacy Enhancing Technologies*, voir ci-après), toutefois un terme dédié fut adopté afin de ne pas limiter le périmètre des techniques pouvant être proposées par les fournisseurs de technologie participants.

Mécanisme de mutualisation (pooling mechanism) : la méthode (et sa mise en œuvre) utilisée dans l'expérimentation ACPR par une équipe d'établissements participants pour mutualiser leurs jeux de données respectifs. Afin d'évaluer le gain maximal associé à l'analyse collaborative, l'ACPR a encouragé les participants à explorer les méthodes de mutualisation les plus sophistiquées. Celles-ci peuvent par exemple être fondées sur un nombre arbitraire de requêtes SQL incluant jointures exactes ou approchées, *a contrario* des techniques de mutualisation plus traditionnelles (par exemple des jeux de données aux schémas identiques simplement empilés, ou encore le partage d'information minimale telle qu'un simple indicateur binaire de suspicion).

ML (Machine Learning, apprentissage automatique) : L'apprentissage automatique est un champ d'étude de l'intelligence artificielle qui se fonde sur des approches mathématiques et statistiques pour donner aux ordinateurs la capacité d'apprendre à partir de données, c'est-à-dire d'améliorer leurs performances à résoudre des tâches sans être explicitement programmés pour chacune. Plus largement, il concerne la conception, l'analyse, l'optimisation, le développement et la mise en œuvre de telles méthodes. On parle d'apprentissage statistique car l'apprentissage consiste à créer un modèle dont l'erreur statistique moyenne est la plus faible possible. (Source: Wikipedia.)

Participant (au Tech Sprint) : tout fournisseur de technologie ayant répondu à l'appel à candidatures publié le 16 mai 2022 et ayant été ensuite sélectionné par l'ACPR pour concourir au Tech Sprint.

Périmètre de Confidentialité : dans le cadre d'une solution MCD, le périmètre de confidentialité désigne une zone de stockage et d'analyse, physique ou virtuelle, fournissant des garanties de confidentialité en son sein. Par exemple dans le cas de chiffrement de bout en bout, ce périmètre comprend l'ensemble du protocole expérimental (base de données et flux de données inclus) en aval du chiffrement initial des données confidentielles.

PDDM (Pooled Data Detection Model) : pour tout établissement participant à l'expérimentation ACPR (ou pour l'établissement fictif A du PoC du Tech Sprint), désigne un modèle de détection LCB-FT opérant sur ses propres données enrichies de données d'un ou plusieurs établissements pairs en collaboration.

PET (Privacy-Enhancing Technologies, technologies de maintien de la confidentialité) : aussi appelées technologies d'amélioration de la confidentialité, un ensemble de méthodes de protection des données permettant à des utilisateurs de protéger la confidentialité de leurs informations confidentielles, fournies et traitées par des services ou applicatifs. Les *PET* utilisent des algorithmes de minimisation ou chiffrement des données confidentielles ou en évitant autant que possible de diminuer leur utilité.

PSI (Private Set Intersection, intersection confidentielle): une méthode cryptographique de type SMPC permettant à deux acteurs possédant des ensembles d'éléments de comparer une version chiffrée de chacun de ces ensembles afin d'en calculer l'intersection. Dans ce scénario, aucune des deux parties ne révèle la moindre information à l'autre en-dehors des éléments figurant dans l'intersection.

SDDM (Solo Data Detection Model): pour tout établissement participant à l'expérimentation ACPR (ou pour l'établissement fictif A du PoC du Tech Sprint), désigne un modèle de détection LCB-FT opérant sur ses propres données.

SMPC (Secure Multi-Party Computation, calcul multipartite sécurisé): l'une des familles de technologies possibles pour réaliser la MCD. Le calcul multipartite sécurisé est une branche de la cryptographie dont l'objectif est de permettre aux agents d'un réseau de communication de calculer conjointement une fonction sur leurs entrées, afin que les entrées restent privées et que le résultat soit exact. Cela peut être réalisé, par exemple, par transfert inconscient ou par chiffrement homomorphe. (Source: Wikipedia.)

TEE (Trusted Execution Environment, Enclave sécurisée): l'une des familles de technologies possibles pour réaliser la MCD. Le TEE, ou enclave protégée d'exécution du code, garantit la confidentialité des données (les entités non-autorisées n'ont pas accès aux data pendant leur utilisation), l'intégrité des données (les entités non-autorisées ne peuvent pas supprimer ou modifier des données en utilisation) et l'intégrité du code (les entités non-autorisées ne peuvent pas supprimer ou modifier du code en utilisation). L'informatique confidentielle prend également d'autres processus en charge, comme le mécanisme d'attestation du TEE ou la migration de charges de calcul entre plusieurs TEE.

7. Annexe : modèle de fiche d'évaluation des solutions

Les critères d'appréciation étaient groupés en trois sections en fonction de la phase correspondante au sein de l'expérimentation LCB-FT, chaque section contenant plusieurs questions, pour un total de 10 questions (elles-mêmes composées de 3 à 5 items).

Première section

Cette section concerne uniquement les travaux pour le PoC sur données fictives, réalisés en 3 mois et présentés le 13 septembre.

Question 1 – garanties de sécurité

- modèle de menace
- garanties de confidentialité des données d'entrée selon leur niveau de sensibilité
- garanties d'intégrité du code et des données
- estimation du risque résiduel, certifications de sécurité telles que normes, standards, audits externes ou analyse juridique.

Question 2 – technologie

- type de technologie (PET ou autre, logiciel/matériel/hybride)
- tenue en charge en phase A (mutualisation de données)
- tenue en charge en phase B (modèles de détection).

Question 3 – couverture fonctionnelle de la solution

- fonctionnalités de mutualisation (jointures exactes ou floues, etc.)
- fonctionnalités analytiques (modèles à base de ML / règles écrites en SQL / autres)
- évaluation de la performance de détection des modèles, traçabilité et auditabilité.

Deuxième section

Cette section concerne la solution, construite bien sûr sur la base des travaux du Tech Sprint, envisagée pour réaliser l'expérimentation « one-shot » sur données réelles prévue pour (environ) 9 mois suivant le Tech Sprint, donc entre octobre 2022 et mi-2023.

Question 4 – modélisation LCB-FT

- expertise LCB-FT (pour compléter l'expertise maison / pour l'appuyer / pour tester de nouvelles approches)
- tenue en charge en phase A (mutualisation de données)
- tenue en charge en phase B (modèles de détection)

Question 5 – adéquation fonctionnelle aux protocoles expérimentaux envisagés

- données d'entrée (type de variables, schéma, volumétrie)
- mécanisme de mutualisation (jointure sur état civil ou numéro de compte, sur personnes physiques et morales)
- classe de modèles de détection (supervisé ou non, analyse relationnelle ou graphes, …)
- sorties de modèle (priorisation/routage de transactions, score de risque client, signaux faibles)

Question 6 – adéquation technologique

- appétence de votre établissement pour l'innovation (technologies encore peu adoptées par le secteur)
- contraintes d'intégration au SI (solutions applicatives internes ou externes) de votre établissement

- contraintes de déploiement imposées par votre établissement (compatibilité avec les infrastructures maison ou externes, ou au moins avec le standard du marché)

Question 7 – adéquation en matière de traitement des données

- règles de gouvernance des données de votre établissement
- politique de votre établissement en matière d'hébergement
- interprétation du cadre juridique dans votre établissement (par le DPD, vis-à-vis du secret bancaire, …)

Troisième section

Cette section propose une analyse des bénéfices mais aussi des coûts et obstacles associés à l'adoption éventuelle de la solution considérée, il s'agit donc de se projeter dans l'hypothèse d'une utilisation au-delà de l'expérimentation « one-shot » (cette adoption de plus long terme étant elle-même envisageable à titre expérimental ou en production)

Question 8 - valeur ajoutée pour le métier

- gains d'efficacité (réduction du taux de faux positifs, détection de nouveaux signaux ou schémas, ...)
- gains d'efficience (accélération de la construction de profils à risque, automatisation de la procédure de requête d'informations complémentaires...)
- évaluation de nouveaux modèles de détection

Question 9 – impact opérationnel

- impact sur les équipes métier (maintenance fonctionnelle de la solution - par exemple paramétrage -, ...)
- impact sur les équipes informatiques (changement ou évolutions de l'applicatif, ...)
- impact sur les équipes data science / innovation (workflow modifié, solution intrusive ou pas, limitation des modèles d'IA utilisables, capacité à faire du réentraînement, ...)
- impact sur les équipes sécurité et infrastructure (autonomie dans la maintenance technique de la solution, nouveaux processus tels que gestion du cycle de vie des clés de chiffrement, ...)

Question 10 – coûts et risques associés.

- coûts d'équipement matériel et logiciel
- ressources humaines (IT, SGBD, Data Science…)
- maturité de la solution (présence sur le marché, périmètre d'intervention en cas de partenariat au Tech Sprint, ...)
- risques associés à l'adoption (échec technique, blocage juridique, …)