



## ToxTwin : Industrialisation d'un pipeline de prédiction toxicologique pré-Phase 1

### Du data brut au score toxicologique — architecture, pipeline et considérations réglementaires

Jérôme Vétillard, VP R&D Engineering / CPO, Qualees | Fondateur, Twingital Institute  
 (<https://twingital-ventures.com>)

#### Résumé

La prédiction toxicologique constitue l'un des goulots d'étranglement les plus coûteux du développement pharmaceutique. Les approches traditionnelles (essais in vitro, modèles animaux, études de toxicité aiguë...) interviennent tardivement dans le cycle de R&D et génèrent des délais et des coûts disproportionnés par rapport à leur valeur prédictive réelle.

Cet article présente ToxTwin, un pipeline end-to-end de prédiction toxicologique conçu pour scorer des candidats médicamenteux avant leur entrée en essais de Phase 1.

Le modèle central, ToxGNN-V1, est un Graph Neural Network (GIN 5 couches) pré-entraîné par AttrMasking sur 144 877 composés puis fine-tuné sur les 12 essais Tox21.

Il atteint un Mean AUC de 0.857 sur scaffold split Bemis-Murcko, le positionnant entre DeepTox (0.846) et AttentiveFP (0.863) dans le benchmark MoleculeNet.

Nous détaillons l'architecture Medallion de données, les sources publiques exploitées (PubChem, ChEMBL, Tox21), l'extraction NER par LLM local (Phi-4 14B), le framework de tracking expérimental, et les considérations réglementaires qui ont guidé chaque choix technique.

**Mots-clés** : toxicologie prédictive, jumeau numérique moléculaire, Graph Neural Network, Medallion architecture, MLflow, LLM local, MDR, EU AI Act

## **1. Introduction : le coût de la toxicité tardive**

### **1.1. Le problème économique**

Environ 90 % des candidats médicamenteux échouent en essais cliniques. Parmi les causes d'attrition, la toxicité représente approximativement 30 % des échecs (une proportion qui n'a pas significativement diminué depuis les estimations de DiMasi et al. (Tufts CSDD, 2016)).

Le coût moyen de développement d'un médicament, estimé entre 1,3 et 2,6 milliards de dollars selon les méthodologies, est en grande partie imputable à ces échecs tardifs : un signal de toxicité détecté en Phase 2 ou 3 représente un coût irrécupérable de plusieurs centaines de millions, sans compter le coût d'opportunité des années de développement perdues.

### **1.2. Les limites des approches conventionnelles**

Les essais in vitro offrent une couverture partielle des mécanismes de toxicité et ne capturent pas les effets systémiques. Les modèles animaux souffrent d'une transposabilité inter-espèces limitée (concordance humain-animal d'environ 70 % pour la toxicité hépatique, significativement inférieure pour d'autres organes).

Les approches QSAR (Quantitative Structure-Activity Relationship) classiques dépendent de descripteurs moléculaires prédéfinis et peinent à généraliser hors de leur domaine d'entraînement.

Les Graph Neural Networks (GNN) offrent une alternative en apprenant directement à partir de la structure moléculaire, mais les benchmarks publiés (DeepTox, AttentiveFP, GROVER) se concentrent sur la performance prédictive sans documenter le pipeline industriel nécessaire à leur déploiement dans un contexte réglementé.

### **1.3 L'hypothèse ToxTwin**

Construire un "jumeau numérique toxicologique" qui intègre des données multi-sources (bases structurées, littérature non structurée, essais biologiques) dans un pipeline industriel traçable, reproductible, et aligné avec les exigences réglementaires européennes, pour produire un score de risque toxicologique exploitable en amont de la Phase 1, accompagné d'une quantification de l'incertitude et d'un indicateur de domaine applicatif.

## **2. Architecture de données : le modèle Medallion appliqué à la toxicologie**

### **2.1 Logique architecturale**

L'architecture Medallion (Bronze / Silver / Gold), initialement popularisée par Databricks pour les data lakehouses, s'applique naturellement à la problématique toxicologique : les données brutes sont hétérogènes, multi-format, et nécessitent des transformations successives dont la traçabilité est critique dans un contexte réglementé.

L'ensemble du pipeline est stocké en Delta Lake 3.2 sur PySpark 3.5.3 (mode local), garantissant les transactions ACID et le time-travel pour l'audit trail complet (exigence alignée avec le MDR (Classe IIa) et l'EU AI Act).

## 2.2 Couche Bronze : ingestion brute

La couche Bronze constitue l'archive immuable des données sources. Aucune transformation chimique n'est appliquée à ce stade.

**PubChem (NIH/NCBI)** : 100 000 composés (CID 1 à 100 000) ingérés via l'API PUG-REST. Champs collectés : SMILES isomériques, InChI, InChIKey, masse moléculaire, XLogP, TPSA, donneurs/accepteurs de liaison hydrogène, liaisons rotatives, charge formelle. Stockage : table Delta bronze/pubchem\_compounds.

**ChEMBL (EMBL-EBI, v34)** : 100 000 composés (filtre MW ≤ 900 Da) et 500 000 activités biologiques (IC50, Ki, EC50 en nM). Les données hERG (canal potassique cardiaque) sont incluses pour la couverture cardiotoxicité. Tables Delta : bronze/chembl\_compounds + bronze/chembl\_activities.

**Tox21 (EPA/NIH/FDA, via MoleculeNet)** : 7 831 composés avec 78 000+ labels binaires (actif/inactif) couvrant 12 essais biologiques : SR-ARE, SR-ATAD5, SR-HSE, SR-MMP, SR-p53 (stress response) et NR-AhR, NR-AR, NR-AR-LBD, NR-Aromatase, NR-ER, NR-ER-LBD, NR-PPAR-gamma (nuclear receptors). Le ratio actifs/inactifs varie de 3 % à 16 % selon l'essai — déséquilibre géré au niveau Gold. Table Delta : bronze/tox21\_assays.

## 2.3. Enrichissement toxicologique : extraction NER par LLM local

*Ce pipeline constitue la différenciation principale de ToxTwin vis-à-vis des solutions concurrentes.*

Les monographies toxicologiques PubChem (format JSON PUG-View) sont soumises à Phi-4 14B (déployé localement via Ollama) pour extraction d'entités nommées.

Sur 7 345 composés traités, le pipeline a extrait : 2 453 LD50 orales rat (mg/kg normalisé), 189 LDLo (dose létale basse), 151 TDLo (dose toxique basse), 112 NOAEL, 3 138 annotations d'organes cibles et 876 données de toxicité chronique (avec un taux de succès NER de 99,99 % (5 400 profils complets)).

Un score composite tox\_score (0-5) est calculé selon une grille de sévérité croissante : 0 = absence de données, 1 = LD50 > 2000 mg/kg, 2 = LD50 500-2000, 3 = LD50 50-500 ou ≥ 2 organes cibles, 4 = LD50 < 50 ou mutagène IARC, 5 = LD50 < 10 + organes critiques (SNC, cœur, foie).

Le choix d'un LLM local plutôt qu'une API cloud repose sur trois contraintes convergentes :

- souveraineté des données dans un pipeline à vocation réglementaire,
- coût marginal nul à l'échelle (dizaines de milliers de résumés),
- et reproductibilité des extractions via un modèle versionné avec paramètres d'inférence figés.

## 2.4 Couche Silver : curation et normalisation

La couche Silver applique un pipeline de curation en 5 phases, produisant 144 879 composés dédupliqués.

**Phase 1 : Sanitization chimique (RDKit)** : chaque SMILES est soumis au pipeline RDKit (sanitize, neutralisation des formes ioniques, canonicalisation, calcul InChI-Key). Les SMILES non sanitisables (erreurs de kékularisation aromatique) sont rejetés (~2-3 % du Bronze).

**Phase 2 : Déduplication cross-source** : déduplication sur la clé InChI-Key (couche standard, 14 premiers caractères). Un composé présent dans PubChem, ChEMBL et Tox21 avec des SMILES légèrement différents est consolidé en un seul enregistrement Silver. Priorité de source : ChEMBL > PubChem. Résultat : 144 879 composés uniques à partir de ~200 000 entrées brutes.

**Phase 3: Filtrage drug-like** : filtre de Lipinski avec tolérance +20 % ( $MW \leq 500$  Da,  $\text{LogP} \leq 5$ ,  $\text{HBD} \leq 5$ ,  $\text{HBA} \leq 10$ ). Exclusion des sels, polymères et inorganiques purs. Quality scoring composite 0-1 basé sur la complétude et la validité SMILES.

**Phase 4 : Jointure tox\_ annotations Tox21** : les 77 733 labels Tox21 sont pivotés (format long → wide) et joints via InChI-Key. La table silver/tox\_ annotations contient un vecteur binaire de 12 dimensions par composé.

**Phase 5 : Jointure tox\_profiles enrichis** : les 5 400 profils toxicologiques enrichis par Phi-4 sont intégrés via le mapping CID → InChI-Key, produisant la table silver/tox\_profiles avec LD50, LDLo, TDLo, NOAEL, organes cibles (JSON), tox\_score 0-5 et données de toxicité chronique.

**Tables Silver produites** : silver/compounds (144 879 composés), silver/bioactivity (87 696 activités), silver/tox\_ annotations (77 733 labels), silver/tox\_profiles (5 400 profils enrichis), silver/quality\_log (rapport de run).

## 2.5 Couche Gold : dataset d'entraînement

**Featurization moléculaire** : Morgan Fingerprints ECFP4 (rayon 2, 2 048 bits, useChirality=True) et 14 descripteurs RDKit physicochimiques (MW, MolLogP, TPSA, NumHDonors, NumHAcceptors, NumRotatableBonds, NumAromaticRings,

FractionCSP3, NumHeteroatoms, NumStereocenters, BalabanJ, BertzCT, Chi0v, Kappa1). Vecteur total : 2 062 dimensions par composé.

**Scaffold split Bemis-Murcko** : séparation train/val/test par scaffold moléculaire garantissant l'absence de fuite structurelle. Les composés partageant un même scaffold sont systématiquement affectés à la même partition. Résultat : 4 074 composés (train, 80 %), 452 composés (validation), 502 composés (test holdout).

### **3. Modélisation : ToxGNN-V1**

#### **3.1 Baseline Random Forest**

Un algorithme de type Random Forest par endpoint toxicologique est entraîné sur les features Gold (Morgan FP 2 048 bits + 14 descripteurs, class\_weight=balanced) pour établir une référence de performance.

Mean ROC-AUC = 0.836, dans la fourchette haute de DeepTox (Mayr et al., 2016 : 0.846), validant la qualité des features. Les performances varient de 0.750 (NR-AR) à 0.944 (NR-AhR) pour certains labels de sortie.

#### **3.2 Architecture de l'encodeur GNN**

L'encodeur ToxGNNEncoder est un Graph Isomorphism Network (GIN) à 5 couches avec jump connections.

Chaque atome constitue un noeud du graphe (128 dimensions : numéro atomique, degré de valence, charge formelle, hybridisation, aromaticité, H implicites), chaque liaison une arête bidirectionnelle (11 dimensions : type de liaison, conjugaison, cycle).

Le modèle compte 2 132 038 paramètres entraînaibles, avec une dimension cachée de 512 et un dropout de 0.3.

#### **3.3 Pré-entraînement par AttrMasking**

Stratégie inspirée de GROVER (Rong et al., 2020) : 25 % des atomes sont masqués aléatoirement et le GNN doit reconstruire leurs attributs one-hot.

Cette stratégie est l'équivalent moléculaire du BERT masqué, le modèle apprend des représentations générales sans aucun label toxicologique.

Pré-entraînement sur 144 877 composés pendant 100 epochs (loss : 0.084 → 0.0045).

Les hyperparamètres ont été optimisés par recherche bayésienne Optuna (TPE Sampler, MedianPruner, 30 trials) : hidden\_dim=512, n\_layers=5, dropout=0.3, batch\_size=128, mask\_ratio=0.25, lr=4.0e-4.

#### **3.4 Fine-tuning supervisé multi-tâches**

L'encodeur pré-entraîné est augmenté d'une couche partagée (Linear → BatchNorm → ReLU → Dropout) et de 12 têtes de classification indépendantes (une par assay Tox21).

Stratégie freeze/unfreeze : les 10 premières epochs entraînent uniquement les têtes (encodeur gelé) pour préserver les représentations pré-entraînées ; à partir de l'epoch 11, l'encodeur est dégelé avec un lr réduit ( $\times 0.1$ ). La loss est une BCE pondérée ( $\text{pos\_weight} = \text{n\_neg}/\text{n\_pos}$ , cap à 20) avec masquage des labels manquants (-1). Trois versions ont été évaluées : v1 (dropout 0.3, AUC 0.846, overfitting), v2 (dropout 0.5, AUC 0.857, retenue), v3 (oversampling 30 %, AUC 0.629, rejetée car l'oversampling agressif sur Tox21 pousse au sur-apprentissage des actifs répétés).

### 3.5 Résultats : ToxGNN-V1 v2 vs baseline RF

Le GNN dépasse la baseline RF sur 7 endpoints sur 12, avec un gain moyen de +0.021 AUC. Les gains les plus significatifs concernent les endpoints complexes : SR-p53 (+0.077, de 0.834 à 0.911), NR-ER-LBD (+0.078, de 0.770 à 0.848), NR-AR-LBD (+0.044, de 0.808 à 0.852). Le RF conserve un léger avantage sur les endpoints à forte séparation linéaire (NR-AhR : 0.944 vs 0.931, SR-MMP : 0.927 vs 0.912).

Un point de vigilance : NR-AR (0.659 vs RF 0.750) ; c'est un endpoint à investiguer, potentiellement lié à l'insuffisance de données ChEMBL sur le récepteur androgène.

### 3.6 Positionnement benchmark MoleculeNet Tox21

ToxGNN-V1 (Mean AUC 0.857) se positionne entre DeepTox (0.846, DNN classique, Mayr et al. 2016) et AttentiveFP (0.863, GNN attention, Xiong et al. 2020). L'écart avec GROVER (0.876, GNN self-supervised, Rong et al. 2020) s'explique par la taille du corpus de pré-entraînement (144K composés vs 10M+ pour GROVER) et l'absence de multi-head attention dans l'architecture GIN. Ces deux axes constituent la roadmap v2.0.

## 4. Évaluation : KPIs post-entraînement

### 4.1 ROC-AUC et Average Precision

Mean ROC-AUC de 0.857 sur le validation set (452 molécules), dépassant la cible de 0.836 (baseline RF). Le meilleur endpoint est NR-ER-LBD (0.848), le plus faible NR-AR (0.659).

### 4.2 Calibration (ECE)

L'Expected Calibration Error (ECE) mesure la concordance entre probabilités prédites et fréquences observées. ECE mesuré : 0.149 (seuil cible :  $< 0.10$ ) FAIL. Le Temperature Scaling n'améliore pas le résultat (0.151 post-calibration). Cause racine identifiée : la BCE pondérée avec  $\text{pos\_weight}$  élevé décale les logits vers les positifs. Action prévue pour v1.1 : Isotonic Regression par endpoint.

### **4.3 Incertitude épistémique (Monte Carlo Dropout)**

30 passes stochastiques avec dropout activé sur le test holdout. Incertitude moyenne : 0.0459 (seuil < 0.05) PASS. 34/502 molécules (6.8 %) présentent une haute incertitude (> 0.15), ce qui est acceptable et attendu pour un scaffold split. En production, les endpoints à forte incertitude sont signalés à l'utilisateur comme requérant une interprétation prudente.

### **4.4 Domaine applicatif (Tanimoto)**

Similarité Tanimoto maximale de chaque composé test vis-à-vis de 2 000 molécules du training set. Un composé est déclaré hors domaine si `max_similarity < 0.3`. Résultat : 192/502 composés (38.2 %) hors domaine qui est la conséquence directe et intentionnelle du scaffold split. En production, ce flag constitue un mécanisme de protection critique : un utilisateur soumettant un composé structurellement très différent du training set en est informé explicitement.

## **5. Infrastructure souveraine**

Oui c'est synonyme d'une machine sous le bureau...

### **5.1 Stack technique**

Windows Server 2025 + WSL2 Ubuntu 24.04, SSD Bitlocké, NVIDIA RTX 5080 16 Go VRAM CUDA 12.8), conda env ml (Python 3.11, PyTorch 2.12 cu128, PyTorch Geometric 2.7.0), PySpark 3.5.3 + Delta Lake 3.2, MLflow (backend SQLite), Ollama (Phi-4 14B), Optuna 4.8.0, RDKit 2024.x, FastAPI + Uvicorn, JupyterLab.

### **5.2 La logique de choix d'une infrastructure locale**

Le choix d'une infrastructure souveraine, hors cloud public, repose sur l'alignement avec les exigences MDR (Classe IIa) et EU AI Act sur la maîtrise du lieu de traitement. L'ensemble du pipeline (de l'ingestion des données publiques à l'inférence) s'exécute sur un serveur physique dont la gouvernance est entièrement contrôlée.

L'absence de dépendance à un cloud provider élimine les risques de vendor lock-in, de transfert de données hors juridiction, et de rupture de service affectant la traçabilité. Le coût total de possession est favorable pour un workload intermittent mais intensif en compute.

Le "setup" de la machine a permis de démontrer que nous pouvions "off loader" de nombreuses tâches exécutées dans le cloud, sur une machine locale. Il a permis aussi de définir un set-up multi-utilisateurs, même si le pool "GPU + NVRAM" est partagé par tous les utilisateurs, ce qui peut poser des soucis... et amène à la recommandation d'avoir au plus, 2 data Scientists / machine. Enfin le set-up définitif sera réalisé directement en Ubuntu sans passer par WSL2 pour simplifier l'architecture logique de la machine.

## **6. Considérations réglementaires**

### **6.1 Positionnement MDR et EU AI Act**

Le score ToxTwin n'est pas un dispositif médical au sens du MDR (pas d'usage diagnostique ou thérapeutique direct sur un patient). Il constitue un outil d'aide à la décision en R&D pharmaceutique.

Sous l'EU AI Act, un système de scoring toxicologique pré-clinique ne relève pas a priori de la catégorie « haut risque » (Annexe III), sauf s'il est intégré dans un processus décisionnel ayant un impact direct sur la santé humaine.

Position adoptée : concevoir le pipeline comme si il relevait du haut risque pour anticiper une éventuelle reclassification et démontrer la maturité méthodologique.

### **6.2 Traçabilité et auditabilité**

Le Delta Lake time-travel permet de reconstituer l'état exact de chaque couche (Bronze, Silver, Gold) à n'importe quel point dans le temps. Les modèles sont versionnés via MLflow avec enregistrement systématique des hyperparamètres, métriques et artefacts.

Les prompts d'extraction NLP Phi-4 sont documentés et versionnés. L'incertitude épistémique (MC Dropout) et le domaine applicatif (Tanimoto) sont exposés dans chaque réponse API conformément à l'exigence de transparence de l'EU AI Act pour les systèmes à risque.

### **6.3 Différenciation réglementaire**

Quatre mécanismes distinguent ToxTwin des benchmarks publiés sur le plan réglementaire :

1. L'extraction NER Phi-4 sur monographies PubChem (absent des benchmarks publiés),
2. L'incertitude épistémique intégrée dans l'API (exigence MDR Classe IIa),
3. Le domaine applicatif avec flag automatique (traçabilité réglementaire),
4. Le pipeline Medallion auditable (conformité EU AI Act).

## **7. Validation fonctionnelle : résultats observés**

### **7.1 Aspirin (acide acétylsalicylique)**

Soumission par nom DCI, résolution PubChem PUG-REST → CID 2244, SMILES canonique CC(=O)Oc1ccccc1C(=O)O, MW 180.16 g/mol.

- Temps d'inférence : 2 017 ms.
- Signal le plus élevé : NR-ER 0.350 (récepteur œstrogène, attendu pour un AINS),

- Suivi de NR-AhR 0.191 et SR-ARE 0.168.
- Données expérimentales Silver : LD50 orale rat 200 mg/kg,
- tox\_score 4/5, organe cible foie.
- Domaine applicatif : Tanimoto 0.594 (dans le domaine).
- Le profil prédictif est cohérent avec la pharmacologie connue.

## 7.2 Ibuprofène (SMILES direct)

- Soumission par SMILES direct (CC(C)Cc1ccc(cc1)C(C)C(=O)O),
- résolution RDKit sans appel réseau, gain de ~430 ms.
- Signal le plus élevé : NR-PPAR-gamma 0.383 , activation PPAR-γ documentée dans la littérature (Lehmann et al., 1997, Nature Medicine).
- Le modèle ToxGNN-V1 a capturé cette relation structure-activité sans supervision explicite.
- NR-ER 0.318 et NR-ER-LBD 0.264 indiquent une perturbation endocrinienne modérée, cohérente avec la classe AINS.
- LD50 orale rat : 636 mg/kg (moins toxique que l'aspirine en aigu),
- tox\_score 2/5.
- Tanimoto : 0.600 (dans le domaine).

## 7.3 Enseignement principal

Le différentiel NR-PPAR-gamma entre aspirine (0.038) et ibuprofène (0.383) constitue un résultat non trivial : il correspond à une différence pharmacologique documentée qui émerge des représentations apprises par le GNN pré-entraîné, sans avoir été explicitement supervisée.

Ce type de signal valide la capacité du modèle à capturer des relations structure-activité au-delà de la simple corrélation avec les labels d'entraînement.

## 8. Accès à l'API ToxTwin

Pour permettre à la communauté scientifique et industrielle d'évaluer le potentiel de ToxTwin, **nous mettrons prochainement à disposition une API REST** (FastAPI) offrant un accès limité au scoring toxicologique.

- Modalités d'accès : - Authentification via compte Twingital (Supabase Auth) - 10 requêtes par semaine (quota gratuit, suffisant pour une évaluation)
- Input : nom DCI (résolution PubChem automatique) ou SMILES canonique

- Output : ScoreResponse complet, résolution moléculaire, 12 scores Tox21 avec incertitudes MC Dropout, données expérimentales (LD50, organes, mutagénicité), domaine applicatif (Tanimoto), avertissement de calibration

### Endpoints :

- POST /v1/score-toxicity pour le scoring d'un composé unique
- POST /v1/score-toxicity/batch pour le scoring batch jusqu'à 10 molécules
- GET /health : état de l'API et du modèle

### 9. Dette technique et roadmap

- v1.1 (court terme) : correction de la calibration ECE (Isotonic Regression, objectif < 0.05), passage inférence sur GPU RTX 5080 (objectif < 200 ms vs 1 500-2 000 ms CPU actuel), enrichissement données NR-AR via ChEMBL.
- Fork V1.2 (court terme) : Intégration dans le pipeline TweenMe pour bénéficier de l'Ontology Modeler Genius, du Smart Data Fertilizer et développement d'une version CINN (Clinically Informed Neural Network). Mais ToxTwin ne serait plus alors développé de façon 100% "souveraine".
- v2.0 (moyen terme) : élargissement du corpus de pré-entraînement (objectif 1M+ composés pour se rapprocher de GROVER), ajout de multi-head attention dans l'architecture, extension au profil ADMET complet, k-NN mean similarity pour affiner le domaine applicatif.

### 10. Conclusion

ToxTwin démontre qu'un pipeline de prédiction toxicologique de niveau benchmark (Mean AUC 0.857 sur Tox21) peut être construit et déployé sur infrastructure souveraine avec une traçabilité de bout en bout compatible avec les exigences réglementaires européennes. Les trois contributions principales (extraction NER par LLM local sur monographies toxicologiques, quantification systématique de l'incertitude et du domaine applicatif, architecture Medallion auditable) répondent à des lacunes identifiées dans les benchmarks publiés, qui se concentrent sur la performance prédictive sans adresser les conditions nécessaires au déploiement en contexte réglementé.

La validation fonctionnelle sur des molécules de référence (aspirine, ibuprofène) confirme la cohérence pharmacologique des prédictions, y compris la capacité du modèle à capturer des relations structure-activité non explicitement supervisées. La mise à disposition d'une API publique à accès limité permettra une évaluation indépendante par la communauté.

### Références

- DiMasi JA, Grabowski HG, Hansen RW. Innovation in the pharmaceutical industry: New estimates of R&D costs. J Health Econ. 2016;47:20-33.
- Mayr A, Klambauer G, Unterthiner T, Hochreiter S. DeepTox: Toxicity prediction using deep learning. Front Environ Sci. 2016;3:80.
- Xiong Z, Wang D, Liu X, et al. Pushing the Boundaries of Molecular Representation for Drug Discovery with the Graph Attention Mechanism. J Med Chem. 2020;63(16):8749-8760.
- Rong Y, Bian Y, Xu T, et al. Self-Supervised Graph Transformer on Large-Scale Molecular Data. NeurIPS. 2020.
- Lehmann JM, Lenhard JM, Oliver BB, et al. Peroxisome Proliferator-Activated Receptors  $\alpha$  and  $\gamma$  Are Activated by Indomethacin and Other Non-steroidal Anti-inflammatory Drugs. J Biol Chem. 1997;272(6):3406-3410.
- Règlement (UE) 2017/745 relatif aux dispositifs médicaux (MDR). - Règlement (UE) 2024/1689 établissant des règles harmonisées concernant l'intelligence artificielle (EU AI Act).
- MoleculeNet benchmark: Wu Z, et al. MoleculeNet: a benchmark for molecular machine learning. Chem Sci. 2018;9(2):513-530.
- Delta Lake documentation. <https://docs.delta.io/> - MLflow documentation.
- <https://mlflow.org/docs/latest/index.html> - Tox21 Data Challenge.
- <https://tripod.nih.gov/tox21/challenge/>

Cet article est publié par le Twingital Institute. Pour toute question ou collaboration : [contact@twingital-ventures.com](mailto:contact@twingital-ventures.com) ou *Via* mon profil linkedin.