

Titre de la thèse (sur plusieurs lignes si nécessaire)

*Traduction du titre de la thèse (sur plusieurs lignes si
nécessaire)*

**Thèse de doctorat de l'université Paris-Saclay et de l'université XXX (si
cotutelle - sinon enlever cette seconde partie)**

École doctorale n° d'accréditation, dénomination et sigle

Spécialité de doctorat : voir annexe

Graduate School : voir annexe. Référent : voir annexe

Thèse préparée dans la (ou les) unité(s) de recherche **Nom(s)** (voir annexe), sous la
direction de **Prénom NOM**, titre du directeur ou de la directrice de thèse, la co-direction
de **Prénom NOM**, titre du co-directeur ou de la co-directrice de thèse, le co-encadrement
de **Prénom NOM**, titre, du co-encadrant ou de la co-encadrante ou la co-supervision de
Prénom NOM, titre, du tuteur ou de la tutrice (en cas de partenariat industriel)

Thèse soutenue à Paris-Saclay, le JJ mois AAAA, par

Prénom NOM

Composition du jury

Membres du jury avec voix délibérative

Prénom NOM
Titre, Affiliation
Prénom NOM
Titre, Affiliation
Prénom NOM
Titre, Affiliation
Prénom NOM
Titre, Affiliation
Prénom NOM
Titre, Affiliation

Président ou Présidente
Rapporteur & Examineur / trice
Rapporteur & Examineur / trice
Examineur ou Examinatrice
Examineur ou Examinatrice

Titre : titre (en français).....

Mots clés : 3 à 6 mots clefs (version en français)

Résumé : Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci

eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Title : titre (en anglais).....

Keywords : 3 à 6 mots clefs (version en anglais)

Abstract : Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci

eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Table des matières

1	Introduction	5
2	Problématique	7
3	État de l'art : PLAN (à supprimer après rédaction)	9
3.1	Introduction	9
3.2	IA générative	9
3.3	Méthodes pour le traitement de séquence	9
3.4	Les améliorations	10
4	État de l'art	11
4.1	Introduction	11
4.1.1	Cadre Conceptuel : Environnement Virtuel et Jumeau Numérique	11
4.1.2	Injection 1 : Génération et modélisation de l'environnement	13
4.1.3	Injection 2 : Simulation de phénomènes physiques	14
4.1.4	Injection 3 : Adaptation et interaction	15
4.2	IA générative	16
4.2.1	L'approche probabiliste explicite et les VAE (2013)	17
4.2.2	La révolution antagoniste : Les GAN (2014)	17
4.2.3	Le paradigme séquentiel et l'autorégression (2016-2018)	18
4.2.4	La génération par raffinement itératif : Les Modèles de Diffusion (2020-Présent)	18
4.3	IA générative	20
4.3.1	L'approche probabiliste explicite : Les VAE (2013)	20
4.3.2	La révolution antagoniste : Les GAN (2014)	21
4.3.3	Le paradigme séquentiel et l'autorégression (2016-2018)	21
4.3.4	La génération par raffinement itératif : Les Modèles de Diffusion (2020-Présent)	22
4.4	IA Générative	23
4.4.1	L'approche probabiliste explicite et les VAE (2013)	23
4.4.2	La révolution antagoniste : Les GAN (2014)	24
4.4.3	Le paradigme séquentiel et l'autorégression (2016-2018)	24
4.4.4	La génération par raffinement itératif : Les Modèles de Diffusion (2020-Présent)	25
5	AVERTISSEMENT	27
6	COMPOSITION GÉNÉRALE, CHARTE GRAPHIQUE	29
6.1	COMPOSITION DU DOCUMENT	29
6.2	QUELS LOGOS FAIRE FIGURER?	29
6.3	POLICES DE CARACTÈRES ET COULEURS	29

7	INFORMATIONS GÉNÉRALES SUR LA PAGE DE COUVERTURE	31
7.1	TITRE DE LA THÈSE ET LANGUE(S)	31
7.2	SPÉCIALITÉ DE DOCTORAT	31
7.3	UNITÉ DE RECHERCHE	32
7.4	LE RÉFÉRENT	32
7.5	GRADUATE SCHOOL	33
7.6	ÉCOLE DOCTORALE	33
7.7	LIEU ET DATE DE SOUTENANCE	34
8	CIVILITÉ, FÉMINISATION DES TITRES ET FONCTIONS	37
9	PRÉSENTATION DE LA DIRECTION DE LA THÈSE OU DE L'ÉQUIPE D'ENCADREMENT	39
10	COMPOSITION DU JURY	41
10.1	A QUOI SERVENT CES INFORMATIONS?	41
10.2	LÉGITIMITÉ ACADÉMIQUE	41
10.3	INDÉPENDANCE	42
10.4	FONCTION DANS LE JURY ET ORDRE DE CITATION	43
10.4.1	Ordre de citation	43
10.4.2	Les rapporteurs	43
11	BIEN CITER SES SOURCES	45
11.1	S'INFORMER SUR LE PLAGIAT	45
11.2	LES IMAGES	45
11.3	ARTICLES JOINTS A LA THÈSE	45
12	DÉPOSER ET DIFFUSER SA THÈSE	47
12.1	LES RESSOURCES A CONSULTER	47
12.2	LES DÉMARCHES	47
13	ANNEXE : LES LOGOS INSTITUTIONNELS	49
13.1	LE LOGO DE L'UNIVERSITÉ PARIS-SACLAY	49
13.2	LOGOS, NUMÉROS D'ACCRÉDITATION ET DÉNOMINATIONS DES ÉCOLES DOCTORALES	49

1 - Introduction

Le chapitre introduction comprendra les éléments suivants :

- Introduction du domaine de la guerre électronique
- Les grandes lignes du rôle du capteur ESM
- Explications de l'utilité de l'environnement numérique
- La problématique d'accélération
- Introduction succincte sur le domaine de l'IA
- Le plan de notre approche

2 - Problématique

Le chapitre sur la problématique contiendra les éléments suivants :

- Description du fonctionnement du capteur ESM dans l'environnement (dans la limite de ce qu'on peut dire), intégré dans la chaîne algorithmique de traitement de l'information.
- Description du fonctionnement de l'environnement numérique, avec l'explication des modélisations de chaque traitement.
- Spécification du goulot d'étranglement et commentaire sur les données I/O.
- Commentaire sur la complexité du problème pour l'apprentissage automatique : double problématique génération et traitement de séquence.

3 - État de l'art : PLAN (à supprimer après rédaction)

Le chapitre sur l'état de l'art se découpe en 4 parties.

3.1 . Introduction

Cette section aborde les aspects suivants :

- Environnements numériques
- Injection 1 : génération et modélisation de l'environnement
- Injection 2 : simulation de phénomènes physiques
- Injection 3 : adaptation et interaction

3.2 . IA générative

Cette section présente le domaine de l'IA générative. Notre problème peut y être naïvement associé mais en réalité quasiment aucune des méthodes ne sera applicable. Les aspects présentés sont :

- Le concept d'IA générative. En notant que n'importe quelle fonction génère une sortie à partir d'une entrée et que la dérive de tout appeler IA générative est tentante.
- Les VAE et spécialement VAE conditionnels
- Les GAN et spécialement GAN conditionnels
- Les modèles de diffusion et spécialement ceux conditionnels
- Les modèles de langage et GPT

3.3 . Méthodes pour le traitement de séquence

Cette section présente les architectures connues pour leurs capacités à traiter des séquences, de leurs formes les plus simples aux formes les plus complexes. Par ordre d'apparition :

- Le concept de séquence : notion de proximité dans un ensemble. Série temporelle, image, texte.
- Réseau de convolution :
 - Histoire de son apparition : dans l'image
 - Comment la convolution interagit avec la séquence
 - La convolution dans l'image (vue comme une séquence)
 - La convolution dans le texte
 - La convolution ailleurs

- Réseau de neurones récurrents :
 - Histoire de son apparition
 - Comment un RNN interagit avec la séquence
 - Variante SSM
 - RNN dans le texte
 - RNN dans les systèmes temporels (chaîne de Markov)
- Transformer :
 - Histoire de son apparition : dans le langage
 - Comment le Transformer interagit avec la séquence
 - Transformer dans le texte (traduction, GPT, ...)
 - Transformer dans les systèmes temporels (chaîne de Markov)
 - Transformer dans l'image
 - Transformer ailleurs (généralisation)

3.4 . Les améliorations

Cette section met en avant les difficultés liées à l'apprentissage automatique, entre complexité calculatoire, mémorielle et instabilité en entraînement. À cette occasion, nous montrons les propositions existantes visant à résoudre ces problèmes. Par ordre d'apparition :

- Compréhension des architectures : Mechanistic Interpretability
- Présentation des soucis de performances
- Présentation des solutions aux soucis de performances
 - Positional Encoding
 - Certains mécanismes d'attention
 - Pre-Training
 - Embedding et tokenization
- Présentation des soucis de stabilité
- Présentation des solutions aux soucis de stabilité :
 - Layer-norm
 - Initialisation
 - Structure (hyper-paramètre de manière générale)
- Présentation des soucis d'efficacité et leurs solutions
 - Complexité mémoire et calcul : mécanisme d'attention
 - Vitesse d'entraînement : MAMBA

4 - État de l'art

4.1 . Introduction

L'émergence de l'industrie 4.0 et le développement rapide de l'intelligence artificielle ont propulsé l'utilisation de représentations virtuelles pour simuler, analyser et optimiser des systèmes physiques. Dans ce paysage technologique en pleine effervescence, les termes « jumeau numérique » et « environnement numérique » sont souvent employés de manière interchangeable, engendrant une confusion sémantique préjudiciable à la précision scientifique. Cette confusion tend à oblitérer les distinctions fondamentales qui existent entre un modèle générique et un avatar individualisé d'un actif physique. Afin de poser les bases conceptuelles solides nécessaires à ce travail, cette section a pour objectif de démêler ces notions. Nous retracerons dans un premier temps l'origine et les définitions, tant idéales que pragmatiques, du jumeau numérique. Dans un second temps, nous présenterons une définition unificatrice et fonctionnelle de l'environnement numérique. Enfin, une synthèse comparative nous permettra d'établir une distinction claire basée sur les flux de données et le critère d'individualisation, et de justifier le positionnement terminologique adopté dans le cadre de cette étude.

4.1.1 . Cadre Conceptuel : Environnement Virtuel et Jumeau Numérique

Le concept de jumeau numérique, popularisé et formalisé dès le début des années 2000 par les travaux de Michael Grieves dans le domaine de la manufacturing [5], puis théorisé comme un pilier des systèmes cyber-physiques (CPS) par des auteurs comme Negri et al. [11], a connu une adoption rapide et variée à travers l'industrie.

Si le terme de "jumeau numérique" s'est imposé dans le paysage technologique, sa définition précise fait l'objet d'un débat animé entre une vision idéale et une approche pragmatique. D'un côté, les puristes, s'appuyant sur les travaux fondateurs de la NASA et de Grieves [5], défendent l'idée qu'un véritable jumeau numérique se caractérise par un couplage bidirectionnel et dynamique avec son homologue physique. Dans cette perspective exigeante, le jumeau n'est pas une simple représentation ; il est un système cyber-physique qui s'enrichit continuellement des données du physique et, en retour, pilote, optimise et prédit son comportement [11]. Cette boucle fermée est considérée comme la condition sine qua non pour distinguer le jumeau numérique d'un simple modèle ou d'une simulation. De l'autre, une approche plus pragmatique, largement répandue dans l'industrie, adopte une définition évolutive et par niveaux de maturité. Dans cette vision, une maquette 3D enrichie de

données, parfois qualifiée de "digital shadow", peut déjà être labellisée "jumeau numérique". Cette flexibilité sémantique, bien que source de confusion, reflète la réalité des projets industriels où la complexité et le coût d'une intégration parfaite imposent une progression par étapes. Malgré tout, une ligne de démarcation essentielle fait consensus : l'existence d'un transfert de données automatique du système physique vers son représentant virtuel. Sans ce flux, la représentation demeure une simulation ou un modèle générique, que nous qualifierons ici d'« environnement numérique ». Par exemple les simulateurs de conduite autonome comme CARLA [3] sont des environnements numériques essentiels pour l'entraînement des algorithmes d'IA, mais ils simulent un monde routier générique non couplé à un véhicule physique unique, et ne sont en ce sens pas des jumeaux numériques. En revanche, certains simulateurs de moteur d'avion, comme ceux déployés par General Electric [17], qui est alimenté en temps réel par les données de vol de l'équipement spécifique, incarnent la définition minimale du jumeau numérique, souvent appelée « Digital Shadow ». Ils permettent un suivi individualisé de l'état de santé et de l'usure de chaque moteur de la flotte.

Pour désigner les représentations numériques qui ne sont pas couplées à une instance physique unique, nous proposons de recourir au terme plus large et unificateur d'Environnement Numérique (Virtual Environment - VE).

La notion d'Environnement Virtuel est interdisciplinaire, et sa définition varie selon que l'on se place dans la communauté de la Réalité Virtuelle, de l'Ingénierie Système ou de l'Intelligence Artificielle. La recherche en Réalité Virtuelle, historiquement focalisée sur l'immersion sensorielle et l'interaction humain-machine, définit souvent les VE comme des « mondes synthétiques générés par ordinateur dans lesquels l'utilisateur a un sentiment d'être présent et d'y interagir » [15]. Cette perspective met l'accent sur les aspects perceptuels et cognitifs. En revanche, dans les domaines de l'ingénierie et de l'IA, l'accent est davantage porté sur la fonction de simulation et de cadre d'expérimentation. Ici, un VE est vu comme un « modèle informatique exécutable d'un système » [4] ou un « cadre de simulation qui permet le test et la validation d'algorithmes dans des conditions contrôlées et reproductibles » [1]. Cette vision est moins concernée par l'immersion de l'utilisateur que par la fidélité de la modélisation des processus et des interactions.

Pour englober ces différentes finalités – de la formation immersive au banc d'essai algorithmique – nous proposons la définition unificatrice suivante : Un Environnement Virtuel (VE) désigne une simulation numérique interactive modélisant un ensemble d'entités et de phénomènes, dans le but d'observer, d'analyser ou d'expérimenter des comportements au sein d'un cadre contrôlé.

Ainsi, la notion de VE s'étend du monde immersif interactif au simulateur de système, selon l'objectif visé. Dans le contexte spécifique du développe-

ment algorithmique, qui est le nôtre, un VE est principalement un outil de prototypage et de validation : il permet de reproduire des situations expérimentales, de générer des données synthétiques et de tester des modèles ou des algorithmes de manière intensive, sûre et économique, sans recourir initialement à des dispositifs physiques.

La distinction fondamentale entre le jumeau numérique et l'environnement numérique réside donc dans le principe d'individualisation par les données et la nature du couplage à un actif physique. Le jumeau numérique, qu'il soit envisagé sous sa forme idéale de couplage bidirectionnel ou sous sa forme minimale de « Digital Shadow », se définit intrinsèquement comme l'avatar numérique d'une instance physique unique, tel le moteur d'avion portant un numéro de série spécifique ou une ligne de production particulière. Son essence et sa valeur opérationnelle sont indissociables du lien data continu avec son jumeau physique. En revanche, l'environnement numérique se conçoit comme une représentation générique d'une classe de systèmes, un modèle de cœur humain standard ou un simulateur de réseau routier type. Son essence réside dans la modélisation fidèle de comportements et de lois physiques au sein d'un cadre contrôlé et reproductible. Par conséquent, et afin d'éviter toute ambiguïté terminologique, ce mémoire utilise de manière exclusive et justifiée le terme d'environnement numérique pour désigner le cadre de simulation générique qui constitue son objet d'étude central. Le système modélisé que nous analysons, à l'instar du simulateur de conduite CARLA [3], est un banc d'essai virtuel destiné au développement et à la validation algorithmique.

L'intégration de l'Intelligence Artificielle au cœur des VE ne constitue pas une approche monolithique, mais se déploie selon trois grands axes d'intervention complémentaires. Ils adressent respectivement le défi de la création des VE, l'amélioration des performances de la simulation elle-même et les capacités d'interaction et d'adaptation du VE.

4.1.2 . Injection 1 : Génération et modélisation de l'environnement

Cette première injection d'IA répond à une problématique économique et logistique majeure : le coût marginal de création des mondes virtuels. Historiquement, la construction d'un Environnement Numérique (VE) de haute fidélité est un processus manuel, artisanal et peu scalable, nécessitant une expertise pointue en modélisation 3D, en texturage et en éclairage. L'IA intervient ici pour basculer d'une logique de construction explicite à une logique de génération implicite ou procédurale, permettant de peupler des environnements vastes et sémantiquement cohérents avec une intervention humaine minimale.

L'avancée la plus spectaculaire concerne la représentation neurale des scènes. Le paradigme classique des maillages polygonaux est remis en cause

par les représentations implicites. Les travaux fondateurs sur les Neural Radiance Fields (NeRF) par Mildenhall et al. (2020) [9] ont démontré qu'un simple perceptron multicouche (MLP) pouvait encoder la géométrie et l'apparence d'une scène complexe en apprenant une fonction continue mappant une coordonnée spatiale et une direction de vue vers une densité et une couleur. Si le NeRF original souffrait de temps d'entraînement et de rendu prohibitifs, les développements récents ont levé ces verrous. L'introduction du Hash Encoding multi-résolution (Instant-NGP) par Müller et al. (2022) [10] a permis d'accélérer l'apprentissage de plusieurs ordres de grandeur. Plus récemment, le 3D Gaussian Splatting de Kerbl et al. (2023) [6] a marqué une rupture en abandonnant le ray-marching coûteux des NeRF au profit d'une méthode de rasterisation de gaussiennes 3D, permettant pour la première fois un rendu photoréaliste en temps réel (> 100 FPS), compatible avec les exigences d'un simulateur interactif.

Parallèlement à la reconstruction du réel, l'IA générative permet la création ex nihilo d'actifs 3D. En s'appuyant sur les puissants modèles de diffusion texte-image pré-entraînés (comme Stable Diffusion), des méthodes comme DreamFusion (Poole et al., 2022) [12] ont introduit le concept de Score Distillation Sampling (SDS). Cette technique permet d'optimiser une représentation 3D (NeRF ou maillage) pour qu'elle corresponde à une description textuelle, sans avoir besoin d'un jeu de données de modèles 3D massif, qui reste rare. Les travaux les plus récents, tels que ProlificDreamer (Wang et al., 2023) [20], améliorent encore cette approche via le Variational Score Distillation (VSD), générant des scènes d'une fidélité et d'une complexité géométrique inédites. Ces outils permettent désormais d'envisager des pipelines où la description sémantique d'un environnement ("une rue urbaine pluvieuse avec des obstacles imprévus") suffit à instancier un VE complet pour le test.

4.1.3 . Injection 2 : Simulation de phénomènes physiques

La seconde injection d'IA vise à lever le verrou du coût computationnel inhérent aux solveurs numériques classiques (Éléments Finis, Volumes Finis). Pour accélérer la simulation des phénomènes physiques au sein des VE, l'état de l'art se divise actuellement en deux paradigmes majeurs : les approches guidées par les données (Data-Driven) et les approches guidées par la physique (Physics-Informed).

Le premier paradigme, guidé par les données, considère le simulateur comme une "boîte noire" dont il faut apprendre la dynamique à partir d'observations. Dans ce cadre, les Graph Neural Networks (GNN) se sont imposés comme l'architecture de référence pour les systèmes particuliers ou maillés, grâce à leur capacité à modéliser les interactions locales non-linéaires. Les travaux de Sanchez-Gonzalez et al. (2020), Learning to Simulate Complex Physics with Graph Networks [14], illustrent la puissance de cette approche : en ap-

prenant sur des trajectoires pré-calculées, le modèle (GNS) parvient à prédire l'évolution de fluides ou de tissus déformables avec une accélération significative à l'inférence et une bonne généralisation à des géométries inédites. Cependant, cette méthode reste tributaire de la quantité et de la qualité des données d'entraînement et ne garantit pas, par construction, le respect strict des lois de conservation (masse, énergie), pouvant mener à des dérives physiques sur le long terme.

Le second paradigme, guidé par la physique, est dominé par les Physics-Informed Neural Networks (PINNs), introduits par Raissi et al. (2019) [13]. Contrairement à l'approche précédente, les PINNs ne cherchent pas seulement à imiter des données, mais à résoudre les Équations aux Dérivées Partielles (EDP) régissant le système. En intégrant les résidus des équations (comme Navier-Stokes) directement dans la fonction de coût du réseau, cette méthode permet de s'affranchir de données d'étiquetage massives et agit comme un solveur sans maillage (mesh-free). Si les PINNs offrent une garantie théorique plus forte et permettent de résoudre des problèmes inverses (retrouver des paramètres physiques à partir d'observations), leur entraînement reste complexe et instable face à des dynamiques chaotiques ou multi-échelles.

Enfin, une troisième voie émerge pour tenter de réconcilier rapidité et précision : l'apprentissage d'opérateurs, notamment via les Fourier Neural Operators (FNO) [8]. Ces modèles apprennent l'opérateur de résolution dans un espace fonctionnel, permettant une invariance à la résolution (resolution independence), ce qui constitue un avantage décisif par rapport aux méthodes classiques dépendantes du maillage.

4.1.4 . Injection 3 : Adaptation et interaction

La troisième injection d'IA redéfinit la dynamique interactionnelle au sein du VE. L'environnement ne se limite plus à une physique passive ; il devient un écosystème peuplé d'agents intelligents et un cadre d'adaptation pour le système sous test. Cette injection vise à enrichir le réalisme comportemental et à fiabiliser le transfert vers le réel.

Premièrement, l'IA est utilisée pour peupler l'environnement d'entités autonomes (personnage non-joueur ou PNJ, adversaires, trafic) afin de créer des scénarios d'interaction complexes et réalistes. Au lieu de reposer sur des scripts comportementaux figés (machines à états finis), ces entités sont pilotées par des politiques apprises via l'Apprentissage par Renforcement Multi-Agents (MARL). L'exemple emblématique est le défi DARPA AlphaDogfight (2020), où des agents IA ont surclassé des pilotes humains experts en combat aérien [2]. Dans un contexte de simulation, intégrer de tels agents "adversariaux" permet de soumettre le système testé à des stratégies optimales ou imprévues, accélérant considérablement la validation de la robustesse par rapport à des scénarios statiques. Cette approche s'appuie sur les mécanismes de

Self-Play, théorisés par Silver et al. (2017) [16], où l'environnement se complexifie automatiquement à mesure que les agents adverses s'améliorent.

Deuxièmement, l'environnement numérique devient le lieu critique du transfert d'apprentissage (Sim-to-Real). Pour qu'une politique apprise en simulation soit applicable dans le monde physique, le VE doit surmonter le Reality Gap. L'approche classique de Domain Randomization (Tobin et al., 2017) [18] utilise l'IA pour perturber aléatoirement les paramètres physiques (frottements, luminosité), forçant l'agent à apprendre une politique robuste invariante aux erreurs de modélisation.

Enfin, les avancées récentes poussent cette logique vers une adaptation active. Plutôt qu'une randomisation aveugle, des méthodes comme l'Automatic Curriculum Learning (ex : POET, Wang et al., 2019 [19]) utilisent des algorithmes évolutionnaires pour générer des configurations d'environnement qui ciblent spécifiquement les faiblesses actuelles de l'agent. De plus, des techniques comme RMA (Kumar et al., 2021) [7] permettent à l'agent d'apprendre, au sein du VE, à estimer implicitement les paramètres physiques invisibles, lui conférant une capacité d'adaptation en temps réel lors du déploiement.

Ainsi, l'IA transforme le simulateur : d'un simple banc d'essai physique, il devient un partenaire d'entraînement actif, capable de générer des opposants redoutables et d'adapter sa propre complexité pour guider l'apprentissage.

4.2 . IA générative

Le concept d'Intelligence Artificielle générative, bien qu'omniprésent dans la littérature récente, nécessite une définition formelle pour être distingué des approches discriminatives classiques. Fondamentalement, là où un modèle discriminatif cherche à modéliser la probabilité conditionnelle d'une étiquette y sachant une entrée x (dans un but de classification ou de régression), les modèles génératifs visent à capturer la distribution conjointe $P(x, y)$ ou la distribution marginale $P(x)$ des données elles-mêmes. L'objectif est d'apprendre la topologie de la variété des données afin de pouvoir échantillonner de nouvelles instances plausibles. Dans le contexte spécifique de l'accélération de simulation, nous nous intéressons particulièrement aux modèles génératifs conditionnels, capables de produire une sortie structurée complexe y (tel un champ physique ou un état futur) correspondant à une condition initiale x . Cette section explore l'évolution chronologique de ces architectures, depuis les approches probabilistes explicites jusqu'aux modèles de diffusion actuels.

4.2.1 . L'approche probabiliste explicite et les VAE (2013)

La première avancée significative dans l'apprentissage profond de distributions complexes fut l'introduction des Auto-encodeurs Variationnels (VAE) par Kingma et Welling en 2013. Contrairement aux auto-encodeurs classiques qui compressent l'information en un point déterministe de l'espace latent, les VAE imposent une structure probabiliste à cet espace, généralement sous la forme d'une distribution gaussienne multivariée. L'innovation majeure réside dans l'introduction de l'astuce de reparamétrisation (reparameterization trick), qui rend le processus d'échantillonnage différentiable et permet l'optimisation du modèle par descente de gradient en maximisant la borne inférieure de la vraisemblance (ELBO). Cette capacité à structurer l'espace latent est particulièrement pertinente pour les problèmes de simulation où une même condition initiale peut mener à plusieurs résultats possibles (stochasticité). C'est ce qu'ont démontré Sohn, Lee et Yan (2015) dans leur article Learning Structured Output Representation using Deep Conditional Generative Models. En introduisant les VAE Conditionnels (C-VAE), ils ont prouvé qu'il était possible de modéliser des sorties structurées multimodales en conditionnant la génération à la fois par une variable latente aléatoire et par une observation d'entrée. Bien que théoriquement élégants, les VAE souffrent historiquement d'une limitation qualitative : l'utilisation d'une fonction de perte de reconstruction type L2 tend à produire des résultats moyennés et flous. L'état de l'art a depuis évolué pour pallier ce défaut, notamment avec les VQ-VAE-2 (Razavi et al., 2019) qui utilisent un espace latent discret quantifié vectoriellement pour générer des données d'une fidélité nettement supérieure.

4.2.2 . La révolution antagoniste : Les GAN (2014)

Pour répondre au manque de piqué et de réalisme des méthodes précédentes, Goodfellow et al. ont introduit en 2014 une rupture paradigmatique avec les Réseaux Antagonistes Génératifs (GAN). Cette approche délaisse l'estimation explicite de la densité de probabilité au profit d'une méthode implicite fondée sur la théorie des jeux. Le processus d'apprentissage est modélisé comme un jeu minimax à somme nulle entre deux réseaux : un générateur qui tente de créer des données indiscernables du réel, et un discriminateur qui tente de distinguer les échantillons générés des données d'entraînement. Comme le soulignent Mohamed et Lakshminarayanan dans Learning in Implicit Generative Models (2016), cette formulation permet de s'affranchir des contraintes liées à la définition d'une fonction de vraisemblance traitable, autorisant le générateur à apprendre des distributions de données extrêmement complexes et détaillées. Dans le cadre de la "traduction" d'environnement, les variantes conditionnelles telles que les cGANs et les architectures type Pix2Pix se sont imposées pour transformer une représentation en une autre, produisant des structures fines souvent inaccessibles aux méthodes basées sur la minimisation de l'erreur moyenne. Néanmoins, les GAN sont

notoirement difficiles à entraîner, souffrant d'instabilités et du phénomène d'effondrement de mode (mode collapse). Malgré ces défis, des architectures abouties comme StyleGAN3 (2021) représentent aujourd'hui l'apogée de cette famille, capables de synthétiser des images haute résolution avec une cohérence géométrique presque parfaite.

4.2.3 . Le paradigme séquentiel et l'autorégression (2016-2018)

Parallèlement aux avancées en vision par ordinateur, le domaine du Traitement du Langage Naturel (NLP) a formalisé une vision de la génération fondée sur la séquence, particulièrement pertinente pour la simulation temporelle. Dans cette optique, la génération d'une donnée complexe Y est décomposée en une série de décisions discrètes, où la probabilité de la séquence complète est le produit des probabilités conditionnelles de chaque élément (ou token) sachant l'historique précédent. C'est le principe de l'autorégression. L'article Google's Neural Machine Translation System : Bridging the Gap between Human and Machine Translation de Wu et al. (2016) a marqué un tournant industriel en validant l'architecture Encodeur-Décodeur pour la traduction de séquences de longueur variable. Ce paradigme "Sequence-to-Sequence" (Seq2Seq) a introduit l'idée que la génération est conditionnée par un vecteur de contexte compressé représentant l'entrée. Cependant, la limitation de la mémoire des réseaux récurrents (RNN) a conduit à l'adoption des mécanismes d'attention, puis des architectures Transformers, permettant de gérer des dépendances à très long terme. L'évolution majeure est survenue avec le concept de pré-entraînement génératif, théorisé par Radford et al. dans Improving Language Understanding by Generative Pre-Training (2018), donnant naissance à la lignée des modèles GPT. Ce travail a démontré qu'un modèle entraîné massivement sur l'objectif simple de prédire le prochain élément d'une séquence (Next Token Prediction) acquiert une capacité de généralisation et de compréhension structurelle émergente. Aujourd'hui, les Grands Modèles de Langage (LLM) comme GPT-4 (Achiam et al., 2023) illustrent la puissance de ce paradigme : si une simulation physique peut être discrétisée en une séquence d'états ou de tokens, l'approche autorégressive permet de "traduire" une condition initiale en une trajectoire future cohérente, transformant le problème de la simulation en un problème de génération de langage.

4.2.4 . La génération par raffinement itératif : Les Modèles de Diffusion (2020-Présent)

La dernière vague d'innovation, qui définit l'état de l'art actuel, puise son inspiration dans la physique statistique. Les modèles de diffusion probabilistes proposent de construire la génération comme l'inversion d'un processus de destruction d'information. L'idée, initialement proposée par Sohl-Dickstein et al. dans Deep Unsupervised Learning using Nonequilibrium Thermodynamics (2015), consiste à détruire progressivement la structure des données par

l'ajout successif de bruit gaussien, puis d'entraîner un réseau de neurones à inverser ce processus temporel pour reconstruire la donnée originale étape par étape. Ce concept a été porté à maturité par Ho et al. avec les Denoising Diffusion Probabilistic Models (DDPM) en 2020. Ils ont démontré que cette approche permettait d'atteindre une qualité d'échantillonnage supérieure à celle des GANs, tout en offrant une couverture de la distribution des données bien plus large (diversité) et une stabilité d'entraînement comparable à celle des méthodes supervisées. Pour répondre aux contraintes de temps de calcul inhérentes à ce processus itératif (nécessitant parfois des milliers de passes), Song et al. ont proposé dans Denoising Diffusion Implicit Models (2020) des méthodes d'échantillonnage non-markoviennes accélérées (DDIM), rendant ces modèles exploitables en production. L'état de l'art contemporain, incarné par les Latent Diffusion Models (Rombach et al., 2022), combine ces mécanismes de diffusion avec la compression des VAE pour opérer dans un espace latent réduit. Cette synergie est au cœur des systèmes génératifs modernes comme Stable Diffusion ou Midjourney, prouvant que les modèles de diffusion sont aujourd'hui les candidats les plus robustes pour la génération d'environnements dynamiques complexes.

Bibliographie de la section

- salut
- **Kingma, D. P., and Welling, M. (2013).** Auto-Encoding Variational Bayes. *International Conference on Learning Representations (ICLR)*.
- **Sohn, K., Lee, H., and Yan, X. (2015).** Learning Structured Output Representation using Deep Conditional Generative Models. *Advances in Neural Information Processing Systems (NeurIPS)*.
- **Razavi, A., van den Oord, A., and Vinyals, O. (2019).** Generating Diverse High-Fidelity Images with VQ-VAE-2. *Advances in Neural Information Processing Systems (NeurIPS)*.
- **Goodfellow, I., et al. (2014).** Generative Adversarial Networks. *Advances in Neural Information Processing Systems (NeurIPS)*.
- **Mohamed, S., and Lakshminarayanan, B. (2016).** Learning in Implicit Generative Models. *arXiv preprint arXiv :1610.03483*.
- **Karras, T., et al. (2021).** Alias-Free Generative Adversarial Networks (StyleGAN3). *Advances in Neural Information Processing Systems (NeurIPS)*.
- **Wu, Y., et al. (2016).** Google's Neural Machine Translation System : Bridging the Gap between Human and Machine Translation. *arXiv preprint arXiv :1609.08144*.
- **Radford, A., et al. (2018).** Improving Language Understanding by Generative Pre-Training. *OpenAI Technical Report*.
- **Achiam, J., et al. (2023).** GPT-4 Technical Report. *arXiv preprint arXiv :2303.08774*.
- **Sohl-Dickstein, J., et al. (2015).** Deep Unsupervised Learning using No-

nequilibrium Thermodynamics. *International Conference on Machine Learning (ICML)*.

- **Ho, J., Jain, A., and Abbeel, P. (2020).** Denoising Diffusion Probabilistic Models. *Advances in Neural Information Processing Systems (NeurIPS)*.
- **Song, J., Meng, C., and Ermon, S. (2020).** Denoising Diffusion Implicit Models. *International Conference on Learning Representations (ICLR)*.
- **Rombach, R., et al. (2022).** High-Resolution Image Synthesis with Latent Diffusion Models. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*.

4.3 . IA générative

Cette section dresse un panorama des architectures génératives, structuré selon leur apparition historique et les paradigmes mathématiques qu'elles introduisent. Nous nous concentrons ici sur les modèles capables d'apprendre et d'échantillonner des distributions de données complexes, en distinguant les approches explicites, implicites, autorégressives et par diffusion.

4.3.1 . L'approche probabiliste explicite : Les VAE (2013)

La première avancée significative dans l'apprentissage profond de distributions complexes fut l'introduction des Auto-encodeurs Variationnels (VAE). Contrairement aux auto-encodeurs classiques qui compressent l'information en un point déterministe, les VAE imposent une structure probabiliste à l'espace latent, généralement une gaussienne multivariée. L'innovation majeure réside dans l'astuce de reparamétrisation (reparameterization trick), qui rend le processus d'échantillonnage différentiable et permet l'optimisation par descente de gradient en maximisant la borne inférieure de la vraisemblance (ELBO). [1] Kingma, D. P., and Welling, M. (2013). Auto-Encoding Variational Bayes. ICLR.

Cette capacité à structurer l'espace latent est cruciale pour les problèmes de simulation où une même condition initiale peut mener à plusieurs résultats possibles (stochasticité). C'est précisément ce point qu'adressent les VAE Conditionnels (C-VAE), qui permettent de modéliser des sorties structurées multimodales en conditionnant la génération à la fois par une variable latente et par une observation d'entrée. Cette architecture est fondamentale pour les problèmes de prédiction structurée où l'incertitude doit être quantifiée. [2] Sohn, K., Lee, H., and Yan, X. (2015). Learning Structured Output Representation using Deep Conditional Generative Models. NeurIPS.

Bien que théoriquement élégants, les VAE souffrent historiquement d'une limitation qualitative : l'utilisation d'une fonction de perte de reconstruction type L2 tend à produire des résultats moyennés et flous. L'état de l'art a depuis évolué pour pallier ce défaut via la quantification vectorielle de l'espace latent (VQ-VAE), permettant de générer des données d'une fidélité nettement supé-

rière tout en conservant les propriétés probabilistes du modèle. [3] Razavi, A., van den Oord, A., and Vinyals, O. (2019). Generating Diverse High-Fidelity Images with VQ-VAE-2. NeurIPS.

4.3.2 . La révolution antagoniste : Les GAN (2014)

Pour répondre au manque de piqué et de réalisme des méthodes variationnelles, une rupture paradigmatique a été introduite avec les Réseaux Antagonistes Génératifs (GAN). Cette approche délaisse l'estimation explicite de la densité de probabilité au profit d'une méthode implicite. Le processus d'apprentissage est modélisé comme un jeu minimax à somme nulle entre deux réseaux : un générateur qui tente de créer des données indiscernables du réel, et un discriminateur qui tente de distinguer les échantillons générés des données d'entraînement. [4] Goodfellow, I., et al. (2014). Generative Adversarial Networks. NeurIPS.

Cette formulation permet de s'affranchir des contraintes liées à la définition d'une fonction de vraisemblance traitable. Comme le soulignent les travaux théoriques sur les modèles implicites, cela autorise le générateur à apprendre des distributions de données extrêmement complexes sans avoir à spécifier une forme fonctionnelle de la densité, capturant ainsi des statistiques d'ordre supérieur souvent ignorées par les méthodes classiques. [5] Mohamed, S., and Lakshminarayanan, B. (2016). Learning in Implicit Generative Models. arXiv.

Dans le cadre de la "traduction" d'environnement ou de simulation, les variantes conditionnelles (cGAN) sont essentielles. Elles permettent de diriger la génération selon une contrainte d'entrée, transformant le bruit aléatoire en une structure cohérente avec le contexte. L'architecture Pix2Pix a notamment démontré que ces modèles restaient une référence absolue pour générer des structures à haute fréquence spatiale (détails fins et nets), là où les méthodes L2 échouent. [6] Mirza, M., and Osindero, S. (2014). Conditional Generative Adversarial Nets. arXiv. [7] Isola, P., et al. (2017). Image-to-Image Translation with Conditional Adversarial Networks. CVPR.

Néanmoins, les GAN sont notoirement difficiles à entraîner, souffrant d'instabilités et du phénomène d'effondrement de mode (mode collapse). Malgré ces défis, des architectures abouties comme StyleGAN3 représentent aujourd'hui l'apogée de cette famille, capables de synthétiser des images haute résolution avec une cohérence géométrique quasi parfaite. [8] Karras, T., et al. (2021). Alias-Free Generative Adversarial Networks (StyleGAN3). NeurIPS.

4.3.3 . Le paradigme séquentiel et l'autorégression (2016-2018)

Parallèlement à la vision par ordinateur, le domaine du Traitement du Langage Naturel a imposé une vision de la génération fondée sur la séquence, particulièrement pertinente si l'on considère la simulation comme une trajectoire d'états. Ici, la génération est modélisée comme une prédiction condi-

tionnelle du prochain élément discret (token), où la probabilité de la séquence complète est factorisée comme le produit des probabilités de chaque étape. [9] Wu, Y., et al. (2016). Google's Neural Machine Translation System : Bridging the Gap between Human and Machine Translation. arXiv.

L'évolution majeure de ce paradigme réside dans le concept de pré-entraînement génératif (GPT). Il a été démontré qu'un modèle entraîné massivement sur l'objectif simple de prédire le prochain élément d'une séquence acquiert une capacité de généralisation et de compréhension structurelle émergente. Ce principe, initialement appliqué au texte, est agnostique à la nature des données, pourvu qu'elles puissent être sérialisées. [10] Radford, A., et al. (2018). Improving Language Understanding by Generative Pre-Training. OpenAI.

Aujourd'hui, les Grands Modèles de Langage (LLM) comme GPT-4 illustrent la puissance de ce paradigme. Ils démontrent que l'approche autorégressive permet de "traduire" une condition initiale en une trajectoire future cohérente en respectant des dépendances à très long terme, ce qui en fait une méthode de choix pour les problèmes de simulation temporelle discrétisée. [11] Achiam, J., et al. (2023). GPT-4 Technical Report. arXiv.

4.3.4 . La génération par raffinement itératif : Les Modèles de Diffusion (2020-Présent)

La dernière vague d'innovation, qui définit l'état de l'art actuel, puise son inspiration dans la physique statistique hors équilibre. Les modèles de diffusion proposent de construire la génération comme l'inversion d'un processus de destruction d'information. L'idée consiste à détruire progressivement la structure des données par l'ajout successif de bruit gaussien, puis d'entraîner un réseau de neurones à inverser ce processus temporel pour reconstruire la donnée originale. [12] Sohl-Dickstein, J., et al. (2015). Deep Unsupervised Learning using Nonequilibrium Thermodynamics. ICML.

Ce concept a été porté à maturité avec les Denoising Diffusion Probabilistic Models (DDPM). Cette approche permet d'atteindre une qualité d'échantillonnage supérieure à celle des GANs, tout en offrant une couverture de la distribution des données bien plus large (diversité) et une stabilité d'entraînement comparable à celle des méthodes supervisées, résolvant ainsi le dilemme historique entre fidélité et diversité. [13] Ho, J., et al. (2020). Denoising Diffusion Probabilistic Models. NeurIPS.

Pour répondre aux contraintes de temps de calcul inhérentes à ce processus itératif, des méthodes d'échantillonnage non-markoviennes (DDIM) ont été proposées, rendant ces modèles exploitables en production grâce à un processus de génération déterministe et accéléré. [14] Song, J., et al. (2020). Denoising Diffusion Implicit Models. ICLR.

L'état de l'art contemporain combine ces mécanismes de diffusion avec des espaces latents compressés (Latent Diffusion Models) pour réduire la dimensionnalité du problème. C'est cette architecture qui propulse les sys-

tèmes génératifs modernes, prouvant que les modèles de diffusion sont aujourd'hui les candidats les plus robustes pour la génération d'environnements dynamiques complexes. [15] Rombach, R., et al. (2022). High-Resolution Image Synthesis with Latent Diffusion Models. CVPR.

4.4 . IA Générative

Le concept d'Intelligence Artificielle générative, bien qu'omniprésent dans la littérature récente, nécessite une définition formelle pour être distingué des approches discriminatives classiques. Fondamentalement, là où un modèle discriminatif cherche à modéliser la probabilité conditionnelle d'une étiquette y sachant une entrée x (dans un but de classification ou de régression), les modèles génératifs visent à capturer la distribution conjointe $P(x, y)$ ou la distribution marginale $P(x)$ des données elles-mêmes. L'objectif est d'apprendre la topologie de la variété des données afin de pouvoir échantillonner de nouvelles instances plausibles. Dans le contexte spécifique de l'accélération de simulation, nous nous intéressons particulièrement aux modèles génératifs conditionnels, capables de produire une sortie structurée complexe y (tel un champ physique ou un état futur) correspondant à une condition initiale x . Cette section explore l'évolution chronologique de ces architectures, depuis les approches probabilistes explicites jusqu'aux modèles de diffusion actuels.

4.4.1 . L'approche probabiliste explicite et les VAE (2013)

La première avancée significative dans l'apprentissage profond de distributions complexes fut l'introduction des Auto-encodeurs Variationnels (VAE) par Kingma et Welling en 2013. Contrairement aux auto-encodeurs classiques qui compresse l'information en un point déterministe de l'espace latent, les VAE imposent une structure probabiliste à cet espace, généralement sous la forme d'une distribution gaussienne multivariée. L'innovation majeure réside dans l'introduction de l'astuce de reparamétrisation (reparameterization trick), qui rend le processus d'échantillonnage différentiable et permet l'optimisation du modèle par descente de gradient en maximisant la borne inférieure de la vraisemblance (ELBO). Cette capacité à structurer l'espace latent est particulièrement pertinente pour les problèmes de simulation où une même condition initiale peut mener à plusieurs résultats possibles (stochasticité). C'est ce qu'ont démontré Sohn, Lee et Yan (2015) dans leur article Learning Structured Output Representation using Deep Conditional Generative Models. En introduisant les VAE Conditionnels (C-VAE), ils ont prouvé qu'il était possible de modéliser des sorties structurées multimodales en conditionnant la génération à la fois par une variable latente aléatoire et par une observation d'entrée. Bien que théoriquement élégants, les VAE souffrent historiquement

d'une limitation qualitative : l'utilisation d'une fonction de perte de reconstruction type L2 tend à produire des résultats moyennés et flous, manquant de détails haute fréquence. L'état de l'art a depuis évolué pour pallier ce défaut, notamment avec les VQ-VAE-2 (Razavi et al., 2019) qui utilisent un espace latent discret quantifié vectoriellement, permettant de générer des données d'une fidélité nettement supérieure tout en conservant les propriétés probabilistes du modèle.

4.4.2 . La révolution antagoniste : Les GAN (2014)

Pour répondre au manque de piqué et de réalisme des méthodes précédentes, Goodfellow et al. ont introduit en 2014 une rupture paradigmatique avec les Réseaux Antagonistes Génératifs (GAN). Cette approche délaisse l'estimation explicite de la densité de probabilité au profit d'une méthode implicite fondée sur la théorie des jeux. Le processus d'apprentissage est modélisé comme un jeu minimax à somme nulle entre deux réseaux : un générateur qui tente de créer des données indiscernables du réel, et un discriminateur qui tente de distinguer les échantillons générés des données d'entraînement. Comme le soulignent Mohamed et Lakshminarayanan dans *Learning in Implicit Generative Models* (2016), cette formulation permet de s'affranchir des contraintes liées à la définition d'une fonction de vraisemblance traitable, autorisant le générateur à apprendre des distributions de données extrêmement complexes et détaillées. Dans le cadre de la "traduction" d'environnement ou de simulation, les variantes conditionnelles telles que les cGANs et les architectures type Pix2Pix se sont imposées comme des standards pour transformer une représentation (par exemple une carte sémantique ou une condition initiale) en une autre (une image photoréaliste ou un état physique), produisant des structures fines souvent inaccessibles aux méthodes basées sur la minimisation de l'erreur quadratique moyenne. Néanmoins, les GAN sont notoirement difficiles à entraîner, souffrant d'instabilités et du phénomène d'effondrement de mode (mode collapse), où le générateur se contente de produire une variété limitée d'échantillons. Malgré ces défis, des architectures abouties comme StyleGAN3 de NVIDIA (2021) représentent aujourd'hui l'apogée de cette famille, capables de synthétiser des images haute résolution avec une cohérence géométrique et texturale quasi parfaite, bien que leur application stricte à la physique nécessite une vigilance quant aux hallucinations visuelles.

4.4.3 . Le paradigme séquentiel et l'autorégression (2016-2018)

Parallèlement aux avancées en vision par ordinateur, le domaine du Traitement du Langage Naturel (NLP) a développé une vision de la génération fondée sur la séquence. Dans cette optique, générer une donnée équivaut à prédire séquentiellement le prochain élément discret (token) conditionnellement à l'historique précédent. C'est le principe de l'autorégression. L'article

Google's Neural Machine Translation System : Bridging the Gap between Human and Machine Translation de Wu et al. (2016) a marqué un tournant industriel en démontrant l'efficacité des architectures Encodeur-Décodeur pour traduire une séquence source en une séquence cible de longueur variable. Bien que ce système reposât initialement sur des réseaux récurrents (RNN), il a posé les bases de la modélisation de problèmes complexes sous forme de traduction. Cette approche a été transcendée par l'apparition des Transformers (détaillés dans la section suivante) et le concept de pré-entraînement génératif introduit par Radford et al. dans Improving Language Understanding by Generative Pre-Training (2018). Ce travail fondateur sur les modèles GPT a prouvé qu'un modèle entraîné massivement à prédire le prochain élément d'une séquence acquiert une capacité de généralisation et de compréhension de la structure sous-jacente des données sans précédent. Aujourd'hui, ce paradigme domine l'IA via les Grands Modèles de Langage (LLM) comme GPT-4 ou Gemini. Ces modèles démontrent que si un problème (y compris une simulation physique) peut être discrétisé sous forme de séquence, l'approche autorégressive permet de générer des solutions complexes respectant des dépendances à long terme, justifiant l'intérêt de traiter la simulation comme un problème de traduction d'état.

4.4.4 . La génération par raffinement itératif : Les Modèles de Diffusion (2020-Présent)

La dernière vague d'innovation, qui définit l'état de l'art actuel, puise son inspiration dans la physique statistique. Les modèles de diffusion probabilistes proposent de construire la génération comme l'inversion d'un processus de destruction d'information. L'idée, initialement proposée par Sohl-Dickstein et al. dans Deep Unsupervised Learning using Nonequilibrium Thermodynamics (2015), consiste à détruire progressivement la structure des données par l'ajout successif de bruit gaussien jusqu'à obtenir un bruit pur, puis d'entraîner un réseau de neurones à inverser ce processus temporel pour reconstruire la donnée originale. Ce concept a été porté à maturité par Ho et al. avec les Denoising Diffusion Probabilistic Models (DDPM) en 2020. Ils ont démontré que cette approche permettait d'atteindre une qualité d'échantillonnage supérieure à celle des GANs, tout en offrant une couverture de la distribution des données bien plus large (diversité) et une stabilité d'entraînement comparable à celle des méthodes supervisées. Pour répondre aux contraintes de temps de calcul inhérentes à ce processus itératif, Song et al. ont proposé dans Denoising Diffusion Implicit Models (2020) des méthodes d'échantillonnage accélérées et déterministes (DDIM), rendant ces modèles exploitables en production.

L'état de l'art contemporain, incarné par des systèmes comme Stable Diffusion ou Midjourney, combine souvent ces mécanismes de diffusion avec des espaces latents (Latent Diffusion) pour réduire la dimensionnalité. Ré-

cemment, des modèles comme Sora (OpenAI) ont étendu ce principe à la génération vidéo cohérente temporellement, prouvant que les modèles de diffusion sont aujourd'hui les candidats les plus sérieux pour la génération d'environnements dynamiques complexes et de phénomènes physiques haute fidélité.

5 - AVERTISSEMENT

La composition de la page de couverture doit être respectée pour la diffusion de la thèse sur www.theses.fr et pour le dépôt légal de la thèse, qui est obligatoire pour l'obtention du diplôme (cf. articles 24 et 25 de l'arrêté du 25 mai 2016 fixant le cadre national de la formation et les modalités conduisant à la délivrance du diplôme national de doctorat).

Les consignes et les recommandations ci-après ont pour objet d'assurer une **homogénéité graphique** pour toutes les thèses soutenues à l'université Paris-Saclay et de les rendre **immédiatement reconnaissables**.

Elles ont également pour objet de donner un cadre de référence permettant d'éviter qu'un lecteur futur puisse avoir des **doutes sur la conformité de la thèse ou du jury**. L'université reçoit régulièrement des demandes d'informations, au sujet de thèses, pour lesquelles il y a des questionnements sur la conformité du jury ou bien des incohérences entre les informations qui figurent sur la couverture de la thèse, d'une part, et les méta-données de la thèse visibles sur www.theses.fr, d'autre part.

Il est rappelé que ces consignes et recommandations ne s'appliquent que pour le dépôt légal de la thèse et sa diffusion via le portail www.theses.fr. **Ce canal de diffusion n'est pas exclusif**. D'autres formats de page de couverture peuvent être librement utilisés par les auteurs sur d'autres canaux de diffusion (par exemple : pour afficher le nom et le logo d'une organisation qui aurait co-financé la thèse et pour la diffusion au sein de cette organisation), à condition que les informations requises pour la citation complète de la thèse de doctorat figurent. C'est-à-dire : au minimum : nom et prénom de l'auteur, titre de la thèse, date, lieu et établissement de soutenance (université Paris-Saclay et le cas échéant un établissement partenaire en cas de cotutelle internationale de thèse), ainsi que le logo de l'université Paris-Saclay et le cas échéant d'une université étrangère partenaire en cas de cotutelle internationale de thèse.

6 - COMPOSITION GÉNÉRALE, CHARTE GRAPHIQUE

6.1 . COMPOSITION DU DOCUMENT

Les deux premières pages sont consacrées aux informations institutionnelles.

Une troisième page peut être ajoutée pour compléter les informations institutionnelles réglementaires des deux premières pages. Par exemple, pour donner des informations sur l'organisme d'accueil ou financeur et afficher leurs logos, pour décrire brièvement un cadre partenarial, pour fournir les noms de personnalités invitées à siéger aux côtés du Jury pour la soutenance, pour afficher le logo du laboratoire etc.

La page des remerciements est alors placée en 3^e ou 4^e page, selon qu'une 3^e page a été ajoutée ou non pour apporter ces compléments d'informations.

6.2 . QUELS LOGOS FAIRE FIGURER ?

Il ne doit figurer sur la **page de couverture de thèse**, aucun autre logo que le **logo de l'université Paris-Saclay** et, en cas de cotutelle internationale de thèse, le logo de l'université partenaire étrangère qui délivre également le diplôme de doctorat pour cette thèse.

Il ne doit figurer sur la **seconde page**, aucun autre logo que le **logo de l'école doctorale**. Les logos institutionnels en vigueur de l'université Paris-Saclay et des écoles doctorales sont fournis au paragraphe 7.2.

Les autres logos, comme celui du laboratoire, d'une entreprise, d'une composante, d'un établissement-composante, d'une université membre associée, d'un organisme de recherche ou de toute autre organisation partenaire de la thèse, peuvent être regroupés dans une troisième page intérieure, avant la page des remerciements, mais ne doivent pas figurer pas sur les deux premières pages.

6.3 . POLICES DE CARACTÈRES ET COULEURS

Les polices de caractère à utiliser sont : Open Sans ou Segoe UI ou Tahoma ou Ebrima. Il ne faut utiliser qu'**une seule police de caractère**.

RVB 99 0 60	RVB 49 62 72	RVB 124 135 143	RVB 213 218 223
RVB 198 11 70	RVB 237 20 91	RVB 238 52 35	RVB 243 115 32
RVB 124 42 144	RVB 125 106 175	RVB 198 103 29	RVB 254 188 24
RVB 0 78 125	RVB 14 135 201	RVB 0 148 181	RVB 70 195 210
RVB 0 128 122	RVB 64 183 105	RVB 140 198 62	RVB 213 223 61

Figure 6.1 – Palette de couleurs de la charte graphique

Sur les 3 premières pages, seules deux couleurs de police sont utilisées, noir et prune (R : 99 V : 0 B : 60). Dans le reste du document, vous pouvez utiliser d'autres couleurs de police, si nécessaire, en veillant à ce qu'elles appartiennent à la palette de couleurs de la charte graphique de l'université Paris-Saclay. D'autres nuances de couleurs peuvent être utilisées parmi les nuances de la palette de l'UPSaclay.

La [charte graphique de l'Université](#) peut être téléchargée sur l'intranet pour plus d'information.

Sur la couverture de la thèse, le **titre** est en police normale de taille 20, de couleur prune et la **traduction du titre** est en police normale de taille 12, de couleur noire et en italique. Si le titre et sa traduction sont très longs, la police peut éventuellement être réduite, mais sans descendre en dessous d'une police 14 pour le titre et d'une police 10 pour la traduction du titre.

7 - INFORMATIONS GÉNÉRALES SUR LA PAGE DE COUVERTURE

Les informations figurant sur la page de couverture de la thèse doivent être cohérentes avec le diplôme et avec les métadonnées de la thèse sur le portail nationale des thèses www.theses.fr.

7.1 . TITRE DE LA THÈSE ET LANGUE(S)

Le **titre de la thèse** doit être fourni en **français** et en **anglais**. Par défaut, le titre est en français et la traduction du titre est en anglais. Cependant, lorsque la thèse est rédigée en anglais, le titre peut être fourni en anglais et la traduction en français.

Les affiliations (université de rattachement...) peuvent, le cas échéant, être fournies en anglais pour des membres étrangers du jury. La langue par défaut restant le français.

Tous les autres éléments de la couverture de la thèse sont en français, les noms des entités (école doctorale, unité de recherche, référent etc.) ainsi que les titres des membres du jury (Professeur, Maître de Conférences etc.). Les correspondances entre titres étrangers et français peuvent être trouvées sur le site du ministère ([GALAXIE](#))¹.

7.2 . SPÉCIALITÉ DE DOCTORAT

La spécialité de doctorat doit faire partie des spécialités pour lesquelles l'école doctorale est accréditée (en pratique : cela implique que vous devez pouvoir la sélectionner dans le menu déroulant des spécialités dans Adum).

La spécialité de doctorat retenue, via le menu déroulant dans Adum, sera celle qui figurera sur le diplôme.

Si votre spécialité n'apparaît pas, il faut contacter le directeur de votre école doctorale.

1. https://www.galaxie.enseignementsup-recherche.gouv.fr/ensup/pdf/EC_pays_etrangers/Tableau_comparaison_au_26_septembre_2012.pdf

7.3 . UNITÉ DE RECHERCHE

L'unité de recherche dans laquelle la thèse a été préparée est précisée sur la couverture de la thèse. Le nom de l'unité est cité en respectant les règles de signature officielles, telles qu'elles ont été convenues entre les tutelles des unités de recherche liées à l'université Paris-Saclay.

Pour les trouver : il faut sélectionner votre unité de recherche via la barre de sélection depuis cette page web : <https://www.universite-paris-saclay.fr/fr/signature> et copier-coller l'adresse de l'unité de recherche sur la couverture de thèse. Puis mettre l'acronyme officiel en premier et le nom des tutelles ensuite, entre parenthèses, dans l'ordre où elles sont indiquées sur <https://www.universite-paris-saclay.fr/fr/signature>. Par exemple, pour IJCLab :

- Voici ce qu'on récupère par un copié-collé depuis l'adresse ci-dessus : « *Université Paris-Saclay, CNRS, IJCLab, 91405, Orsay, France* ».
- Voici comment faire la citation sur la couverture de thèse : « *IJCLab (Université Paris-Saclay, CNRS)* ».

Si la thèse a été préparée dans deux unités de recherche (travaux interdisciplinaires, cotutelle internationale, mobilité...) merci de citer les deux unités de recherche.

Si vous êtes doctorant de l'université Paris-Saclay mais ne trouvez pas votre unité dans la liste, votre unité ne fait probablement pas partie de l'université. Dans ce cas, et à défaut d'une recommandation commune entre l'université et votre unité, complétez la ligne "unité de recherche" de votre page de titre en suivant les recommandations de votre unité de recherche.

La mention de l'université Paris-Saclay comme établissement de soutenance de votre thèse sera automatique en utilisant le modèle de page de couverture de l'université Paris-Saclay.

7.4 . LE RÉFÉRENT

Les référents sont à choisir, en cohérence avec ce qui figure dans votre dossier d'inscription, parmi les composantes, établissements-composantes et universités membres associés de l'Université Paris-Saclay :

- Faculté de droit, économie et gestion,
- Faculté de médecine
- Faculté de pharmacie
- Faculté des sciences d'Orsay
- Faculté des sciences du sport

- AgroParisTech
- Institut d'Optique
- ENS Paris-Saclay
- CentraleSupélec
- Université de Versailles-Saint-Quentin-en-Yvelines
- Université d'Évry Val d'Essonne
- École Nationale d'Architecture de Versailles

7.5 . GRADUATE SCHOOL

La Graduate School est à choisir en cohérence avec votre sujet de thèse et ce qui figure dans votre dossier d'inscription, parmi la ou les Graduate Schools de l'Université Paris-Saclay de rattachement de votre école doctorale ou de votre pôle d'école doctorale :

- Biosphera
- Chimie
- Informatique et sciences du numérique
- Droit
- Économie - Management
- Géosciences, climat, environnement et planètes
- Humanités et Sciences du Patrimoine
- Life Sciences and Health
- Mathématiques
- Physique
- Santé et médicaments
- Santé publique
- Sciences de l'ingénierie et des systèmes
- Sociologie et Science Politique
- Sport, mouvement et facteurs humains

7.6 . ÉCOLE DOCTORALE

- n°127 : astronomie et astrophysique d'Île-de-France (AAIF)
- n°129 : sciences de l'environnement d'Île-de-France (SEIF)
- n°564 : physique en Île-de-France (PIF)
- n°566 : sciences du sport, de la motricité et du mouvement humain (SSMMH)
- n°567 : sciences du végétal : du gène à l'écosystème (SEVE)
- n°568 : signalisations et réseaux intégratifs en biologie (Biosigne)
- n°569 : innovation thérapeutique : du fondamental à l'appliqué (ITFA)
- n°570 : santé publique (EDSP)
- n°571 : sciences chimiques : molécules, matériaux, instrumentation et

biosystèmes (2MIB)

- n°572 : ondes et matière (EDOM)
- n°573 : interfaces : matériaux, systèmes, usages (INTERFACES)
- n°574 : mathématiques Hadamard (EDMH)
- n°575 : electrical, optical, bio : physics and engineering (EOBE)
- n°576 : particules hadrons énergie et noyau : instrumentation, imagerie, cosmos et simulation (PHENIICS)
- n°577 : structure et dynamique des systèmes vivants (SDSV)
- n°579 : sciences mécaniques et énergétiques, matériaux et géosciences (SMEMaG)
- n°580 : sciences et technologies de l'information et de la communication (STIC)
- n°581 : agriculture, alimentation, biologie, environnement, santé (ABIES)
- n°582 : cancérologie : biologie - médecine - santé (CBMS)
- n°629 : Sciences sociales et humanités (SSH)
- n°630 : Droit, Économie, Management (DEM)

7.7 . LIEU ET DATE DE SOUTENANCE

Au moment de l'annonce de soutenance, le lieu et la date de soutenance, servent à donner au public toutes les informations nécessaires pour assister à la soutenance. Étant donné que les soutenances de doctorat doivent être publiques. Il faut donc une information détaillée permettant au public d'y accéder, précisant ainsi l'horaire de début de la soutenance, la salle, l'adresse physique en présentiel ou le lien d'accès à la salle virtuelle lorsque la soutenance se tient en visioconférence ou les deux.

En revanche, **pour la couverture de la thèse** et le dépôt légal de la thèse, le lieu et la date de soutenance ont une fonction « légale » : le lieu définit de quelle juridiction relève le dépôt légal de la thèse et la date est utile, par exemple pour définir l'antériorité ou la fin d'une période de confidentialité. L'information doit donc être donnée sous une forme beaucoup plus synthétique que dans l'annonce de soutenance.

Sur la couverture de la thèse, la date doit être fournie au format « **JJ Mois AAA** » et le lieu de soutenance est simplement la ville, la commune ou la communauté d'agglomérations où s'est tenue la soutenance. Lorsque la soutenance a eu lieu dans les locaux de l'université Paris-Saclay, le lieu à indiquer est celui de la communauté d'agglomérations où se trouve le siège de l'université Paris-Saclay, à savoir « **Paris-Saclay** », que la thèse ait eu lieu en présentiel ou en visioconférence.

Exemple : Thèse soutenue à Paris-Saclay, le 10 Mars 2021.

8 - CIVILITÉ, FÉMINISATION DES TITRES ET FONCTIONS

Il est recommandé de ne pas indiquer les civilités (Madame / Monsieur) ni pour le docteur ou la docteure, ni pour les membres du Jury ou de l'équipe d'encadrement.

Toutefois, si cette recommandation n'était pas suivie, il faudrait alors assurer l'homogénéité. La civilité devrait alors être précisée pour **toutes les personnes** qui figurent sur la couverture (docteur.e, membres du Jury ou de l'équipe d'encadrement) en utilisant les **mêmes conventions** pour tous (Madame / Monsieur ou Mme / M.)

Il est recommandé de féminiser les titres des membres du Jury ou de l'équipe d'encadrement (Professeur / Professeure, Maître ou Maîtresse de conférences etc.) ainsi que les fonctions tenues dans le Jury (examineur / examinatrice ou Présidente / Présidente).

Pour « rapporteur », la forme féminine n'est pas recommandée faute de stabilisation. Si la personne concernée souhaitait la forme féminine, il faudrait alors lui demander de préciser la forme (« rapporteure » ou « rapporteuse »?) qu'elle préfère voir figurer sur la couverture.

9 - PRÉSENTATION DE LA DIRECTION DE LA THÈSE OU DE L'ÉQUIPE D'ENCADREMENT

Toutes les informations sur la direction de la thèse et l'équipe d'encadrement sont précisées sur la couverture de thèse (par exemple : directeur ou directrice de thèse, co-directeur ou codirectrice de thèse, co-encadrant ou co-encadrante, le cas échéant tuteur ou tutrice ou superviseur en entreprise).

Il faut également faire figurer lisiblement, à ce niveau, leur rôle vis-à-vis du doctorant ou de la doctorante et dans la préparation de la thèse.

Le directeur de thèse ou la directrice de thèse est cité en premier et est suivi, le cas échéant, des autres membres de l'équipe d'encadrement de la thèse, co-directeur ou co-directrice par ordre alphabétique puis des co-encadrantes et co-encadrantes par ordre alphabétique, puis tuteur ou tutrices par ordre alphabétique.

Le directeur ou la directrice de thèse et toute autre personne ayant participé à la direction scientifique des travaux et à l'encadrement du doctorant ou de la doctorante ne prend pas part à la décision de Jury de soutenance de doctorat. Ils et elles ne sont ni président, ni rapporteurs, ni examinateurs dans le Jury. Aucun membre de l'équipe d'encadrement de fait partie des membres du Jury avec voix délibérative.

Puisqu'ils et elles n'apparaissent pas via le Jury, il est essentiel que leur rôle soit précisé clairement et lisiblement, sur la couverture de la thèse et sur www.theses.fr pour leur rôle dans l'équipe de direction de la thèse. Les autres personnes qui ont pu contribuer significativement à la thèse sans faire partie de l'équipe d'encadrement, ni du Jury, sont signalées dans la page des remerciements.

10 - COMPOSITION DU JURY

La soutenance de la thèse est une évaluation. Les travaux de recherche de doctorat devant être originaux à l'échelle internationale, le Jury est composé sur mesure pour chaque doctorant.e et chaque thèse de doctorat. La composition du Jury est essentielle, le doctorat est délivré par l'université, sous condition du dépôt légal de la thèse, sur avis conforme du Jury. **Le Jury est garant de la qualité de la thèse.**

Pour chacun des membres du Jury, il faut préciser le **titre**, l'**affiliation** et la **fonction dans le jury** sur la page de couverture.

10.1 . A QUOI SERVENT CES INFORMATIONS?

Ces informations doivent permettre, au premier regard, de vérifier la **conformité de la composition** du Jury :

- sa **légitimité** académique pour se prononcer sur l'obtention du plus haut diplôme universitaire, le doctorat (le jury comprend au moins la moitié de professeurs et assimilés et, sauf dérogation, les membres du Jury sont tous eux-mêmes docteurs).
- sa capacité à se prononcer en toute **indépendance** (au moins la moitié d'externes, à l'établissement de soutenance, à l'école doctorale, à l'équipe d'encadrement, au projet doctoral).

10.2 . LÉGITIMITÉ ACADÉMIQUE

Les **titres** des membres du Jury permettent de vérifier **qu'au moins la moitié des membres du Jury est professeur** des universités ou assimilé.

Les libellés exacts des titres français assimilés aux professeurs des universités (au moins la moitié du Jury) sont disponibles sur [legifrance](https://www.legifrance.gouv.fr/).

Le **président du Jury** est obligatoirement professeur des universités ou assimilé¹.

1. **Arrêté du 15 juin 1992** fixant la liste des corps de fonctionnaires assimilés aux professeurs des universités et aux maîtres de conférences pour la désignation des membres du Conseil national des universités : <https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=LEGITEXT000019860291>

Arrêté du 10 février 2011 relatif à la grille d'équivalence des titres, travaux et fonctions des enseignants-chercheurs mentionnée aux articles 22 et

Si l'un des rapporteurs n'était pas professeur des universités ou assimilé, il faudrait alors préciser, en plus, qu'il dispose bien de l'HDR (par exemple : Maître de conférences, HDR).

10.3 . INDÉPENDANCE

Les affiliations permettent de vérifier que le Jury est bien en **majorité externe** à l'établissement de soutenance, à l'école doctorale et à l'équipe d'encadrement. Pour cela, le nom ou l'acronyme du laboratoire ne suffit pas, en revanche, le nom de l'université ou de l'établissement délivrant le doctorat de rattachement du membre du jury suffit. Il n'est pas utile de préciser certains détails comme l'adresse postale complète ou le pays.

Exemple d'affiliation inadaptée car ambiguë : « IJCLab »

Lorsque le membre du jury est un chercheur d'un organisme national, fournir le nom de son organisme de rattachement ne suffit pas pour juger de son extériorité (CNRS par exemple). Dans ce cas-là ou dans d'autres cas où il y aurait une incertitude de cette nature, susceptible de susciter des interrogations sur le fait qu'au moins la moitié des membres du Jury est externe, il est alors demandé de préciser, en plus, l'université ou l'établissement où ce chercheur inscrit habituellement ses propres doctorants.

Exemple d'affiliation inadaptée car ambiguë : « CNRS »

Exemple d'affiliation adaptée : « CNRS, Université de Toulouse »

Lorsqu'il s'agit d'une entreprise ou d'une fondation ou d'une organisation qui n'est pas en lien direct avec un établissement d'enseignement supérieur pour l'inscription de doctorants, il faut alors le préciser.

Par exemple : « Saint Gobain recherche, entreprise »

Par exemple : « Moveo, Pôle de compétitivité »

43 du décret n° 84-431 du 6 juin 1984 fixant les dispositions statutaires communes applicables aux enseignants-chercheurs et portant statut particulier du corps des professeurs des universités et du corps des maîtres de conférences : https://www.galaxie.enseignementsup-recherche.gouv.fr/ensup/pdf/EC_pays_etrangers/Tableau_comparaison_au_26_septembre_2012.pdf

10.4 . FONCTION DANS LE JURY ET ORDRE DE CITATION

La fonction dans le Jury de chaque membre du Jury doit également être précisée sur la page de couverture.

Les fonctions possibles dans un Jury sont : président(e), examinateur ou examinatrice, rapporteur et directeur ou directrice de thèse.

10.4.1 . Ordre de citation

Le président du Jury est le premier de la liste. Il est immédiatement suivi des deux rapporteurs dans l'ordre alphabétique, puis des autres examinateurs dans l'ordre alphabétique.

Un membre du Jury peut avoir deux fonctions dans le Jury (par ex. rapporteur & examinateur).

10.4.2 . Les rapporteurs

Les rapporteurs participent à l'évaluation de la thèse et figurent donc sur la couverture de thèse, qu'ils soient présents ou non le jour de la soutenance.

Si un rapporteur était absent le jour de la soutenance, il figurerait alors en tant que rapporteur seulement, sinon il figure à la fois en tant que rapporteur & examinateur. Lorsqu'un membre du Jury autre qu'un rapporteur, n'a pas pu participer au Jury de soutenance, physiquement ou bien en visioconférence, son nom ne figure pas sur la page de couverture de la thèse. Dans ce cas, il faut veiller à ce que la composition du Jury reste conforme, malgré l'absence du membre du Jury désigné. Cela peut demander de faire passer un membre interne en invité, par exemple.

11 - BIEN CITER SES SOURCES

La citation des sources fait partie intégrante du travail scientifique et participe de sa qualité et de son intégrité.

11.1 . S'INFORMER SUR LE PLAGIAT

Des mauvaises pratiques de citation peuvent conduire, même sans le vouloir, au plagiat. Le copier/coller, la paraphrase, la réutilisation d'images ou d'idées sans citer la source sont des situations de plagiat (n'hésitez pas à regarder cette courte vidéo sur les différentes formes de plagiat, volontaires ou non : <https://infotrack.unige.ch/comment-reconnaitre-les-cas-de-plagiat>)

Chaque discipline possède ses propres normes en termes de citation des sources. Renseignez-vous auprès de vos pairs pour connaître le style bibliographique et le style de citation à privilégier. Nous vous encourageons vivement à utiliser un logiciel de gestion bibliographique tel que Zotero. Vous pouvez retrouver des supports de formation à ce logiciel dans l'espace eCampus Doctorat, ouvert à tou·te·s sur auto-inscription : <https://ecampus.paris-saclay.fr/course/view.php?id=36678>

11.2 . LES IMAGES

Vous pouvez réutiliser dans votre thèse des images provenant d'articles ou de livres protégés par un copyright. Cela relève en effet de l'exception pédagogique, une des exceptions au droit d'auteur. Attention cependant, vous ne pouvez pas faire ce que vous voulez de cette image ! La loi vous autorise à inclure jusqu'à 20 images (en 720 dpi) sans demander d'autorisation à l'auteur. En revanche, une autorisation est nécessaire à partir de la 21^{ème} image. Les sources des images doivent être mentionnées et aucune modification n'est autorisée.

Pour une présentation détaillée des différents cas d'utilisation d'images dans les thèses et les travaux universitaires, voir : <https://ethiquedroit.hypotheses.org/2947>

11.3 . ARTICLES JOINTS A LA THÈSE

Vous pouvez joindre vos articles à votre thèse. Toutefois, si votre thèse est diffusée en ligne (tout de suite après la soutenance, après un embargo ou

après une période de confidentialité), il convient de s'assurer que vous respectez bien les politiques des éditeurs. En effet, tous n'autorisent pas la diffusion en accès libre de la version éditeur des articles. Utilisez [Sherpa Romeo](#) pour connaître la politique des éditeurs.

Selon la [Loi pour une république numérique](#), si votre recherche est financée à au moins 50% par des fonds publics français, vous avez le droit, en tant qu'auteur, de diffuser la version acceptée de l'article (mais sans la mise en pages de l'éditeur) au bout de 6 mois après publication pour les articles en sciences, techniques et médecine et 12 mois pour les articles en sciences humaines et sociales. Et ce quelque soit la politique éditoriale de l'éditeur.

Pour plus d'informations sur cette question, consultez la section « Déposer dans une archive ouverte » du [Passeport pour la science ouverte](#).

12 - DÉPOSER ET DIFFUSER SA THÈSE

La thèse fait l'objet d'un dépôt légal, en deux étapes, avant la remise du manuscrit aux rapporteurs et après la soutenance, qui protège le droit d'auteur du docteur.

Elle fait ensuite l'objet d'une diffusion sur le portail national des thèses www.theses.fr et le portail européen des thèses [DART-Europe](#), sauf si la thèse présente un caractère confidentiel avéré.

12.1 . LES RESSOURCES A CONSULTER

- [Je publie, Quels sont mes droits ?](#)
- Le cadre réglementaire du [dépôt légal](#) sur Légifrance

12.2 . LES DÉMARCHES

Retrouvez le détail des démarches du dépôt de thèse dans cette [fiche explicative](#)

Si la thèse présente un caractère confidentiel avéré, le classement confidentiel de la thèse et, si nécessaire, une dérogation au caractère public de la soutenance (huis-clos) peuvent être [demandés au chef d'établissement](#).

13 - ANNEXE : LES LOGOS INSTITUTIONNELS

13.1 . LE LOGO DE L'UNIVERSITÉ PARIS-SACLAY



13.2 . LOGOS, NUMÉROS D'ACCREDITATION ET DÉNOMINATIONS DES ÉCOLES DOCTORALES

◆ n°127 : astronomie et astrophysique d'Île-de-France (AAIF)



◆ n°129 : sciences de l'environnement d'Île-de-France (SEIF)



◆ n°564 : physique en Île-de-France (PIF)



◆ n°566 : sciences du sport, de la motricité et du mouvement humain (SSMMH)



◆ n°567 : sciences du végétal : du gène à l'écosystème (SEVE)



ÉCOLE DOCTORALE

Sciences du végétal :
du gène à l'écosystème
(SEVE)

◆ n°568 : signalisations et réseaux intégratifs en biologie (Biosigne)



ÉCOLE DOCTORALE

Signalisations et réseaux
intégratifs en biologie
(BIOSIGNE)

◆ n°569 : innovation thérapeutique : du fondamental à l'appliqué (ITFA)



ÉCOLE DOCTORALE

Innovation thérapeutique
du fondamental à l'appliqué
(ITFA)

◆ n°570 : santé publique (EDSP)



ÉCOLE DOCTORALE

Santé Publique
(EDSP)

◆ n°571 : sciences chimiques : molécules, matériaux, instrumentation et bio-systèmes (2MIB)



ÉCOLE DOCTORALE

Sciences Chimiques: Molécules,
Matériaux, Instrumentation
et Biosystèmes (2MIB)

◆ n°572 : ondes et matière (EDOM)



ÉCOLE DOCTORALE

Ondes et matière
(EDOM)

◆ n°573 : interfaces : matériaux, systèmes, usages (INTERFACES)



ÉCOLE DOCTORALE

Interfaces:
matériaux, systèmes, usages

◆ n°574 : mathématiques Hadamard (EDMH)



ÉCOLE DOCTORALE

de mathématiques
Hadamard (EDMH)

◆ n°575 : electrical, optical, bio : physics and engineering (EOBE)



ÉCOLE DOCTORALE

Physique et ingénierie:
Electrons, Photons,
Sciences du vivant (EOBE)

◆ n°576 : particules hadrons énergie et noyau : instrumentation, imagerie, cosmos et simulation (PHENIICS)



ÉCOLE DOCTORALE

Particules, hadrons, énergie et noyau:
instrumentation, imagerie, cosmos
et simulation (PHENIICS)

◆ n°577 : structure et dynamique des systèmes vivants (SDSV)



ÉCOLE DOCTORALE

Structure et dynamique
des systèmes vivants
(SDSV)

◆ n°579 : sciences mécaniques et énergétiques, matériaux et géosciences (SMEMaG)



ÉCOLE DOCTORALE

Sciences mécaniques et
énergétiques, matériaux
et géosciences (SMEMAG)

◆ n°580 : sciences et technologies de l'information et de la communication (STIC)



ÉCOLE DOCTORALE

Sciences et technologies
de l'information et de
la communication (STIC)

◆ n°581 : agriculture, alimentation, biologie, environnement, santé (ABIES)



ÉCOLE DOCTORALE

Agriculture, alimentation,
biologie, environnement,
santé (ABIES)

◆ n°582 : cancérologie : biologie - médecine - santé (CBMS)



ÉCOLE DOCTORALE

Cancérologie : biologie -
médecine - santé (CBMS)

◆ n°629 : Sciences sociales et humanités (SSH)



ÉCOLE DOCTORALE

Sciences Sociales
et Humanités (SSH)

◆ n°630 : Droit, Économie, Management (DEM)



ÉCOLE DOCTORALE

Droit, Économie,
Management (DEM)

Bibliographie

- [1] Greg Brockman et al. *OpenAI Gym*. 5 juin 2016. doi : [10.48550/arXiv.1606.01540](https://doi.org/10.48550/arXiv.1606.01540). arXiv : [1606.01540\[cs\]](https://arxiv.org/abs/1606.01540). url : <http://arxiv.org/abs/1606.01540> (visité le 21/11/2025).
- [2] Christopher R DeMay et al. "AlphaDogfight Trials : Bringing Autonomy to Air Combat". In : *Johns Hopkins APL Technical Digest* 36.2 (2022).
- [3] Alexey Dosovitskiy et al. *CARLA : An Open Urban Driving Simulator*. 10 nov. 2017. doi : [10.48550/arXiv.1711.03938](https://doi.org/10.48550/arXiv.1711.03938). arXiv : [1711.03938\[cs\]](https://arxiv.org/abs/1711.03938). url : <http://arxiv.org/abs/1711.03938> (visité le 20/11/2025).
- [4] Peter Fritzon. "Principles of Object-Oriented Modeling and Simulation with Modelica". In : ().
- [5] Michael Grieves. "Digital Twin : Manufacturing Excellence through Virtual Factory Replication". In : (1^{er} mars 2015).
- [6] Bernhard Kerbl et al. "3D Gaussian Splatting for Real-Time Radiance Field Rendering". In : *ACM Trans. Graph.* 42.4 (26 juill. 2023), 139 :1-139 :14. issn : 0730-0301. doi : [10.1145/3592433](https://doi.org/10.1145/3592433). url : <https://dl.acm.org/doi/10.1145/3592433> (visité le 21/11/2025).
- [7] Ashish Kumar et al. *RMA : Rapid Motor Adaptation for Legged Robots*. 8 juill. 2021. doi : [10.48550/arXiv.2107.04034](https://doi.org/10.48550/arXiv.2107.04034). arXiv : [2107.04034\[cs\]](https://arxiv.org/abs/2107.04034). url : <http://arxiv.org/abs/2107.04034> (visité le 21/11/2025).
- [8] Zongyi Li et al. *Fourier Neural Operator for Parametric Partial Differential Equations*. 17 mai 2021. doi : [10.48550/arXiv.2010.08895](https://doi.org/10.48550/arXiv.2010.08895). arXiv : [2010.08895\[cs\]](https://arxiv.org/abs/2010.08895). url : <http://arxiv.org/abs/2010.08895> (visité le 21/11/2025).
- [9] Ben Mildenhall et al. "NeRF : representing scenes as neural radiance fields for view synthesis". In : *Commun. ACM* 65.1 (17 déc. 2021), p. 99-106. issn : 0001-0782. doi : [10.1145/3503250](https://doi.org/10.1145/3503250). url : <https://dl.acm.org/doi/10.1145/3503250> (visité le 21/11/2025).
- [10] Thomas Müller et al. "Instant neural graphics primitives with a multiresolution hash encoding". In : *ACM Transactions on Graphics* 41.4 (juill. 2022), p. 1-15. issn : 0730-0301, 1557-7368. doi : [10.1145/3528223.3530127](https://doi.org/10.1145/3528223.3530127). url : <https://dl.acm.org/doi/10.1145/3528223.3530127> (visité le 21/11/2025).

- [11] Elisa Negri, Luca Fumagalli et Marco Macchi. "A Review of the Roles of Digital Twin in CPS-based Production Systems". In : *Procedia Manufacturing*. 27th International Conference on Flexible Automation and Intelligent Manufacturing, FAIM2017, 27-30 June 2017, Modena, Italy 11 (1^{er} jan. 2017), p. 939-948. issn : 2351-9789. doi : [10.1016/j.promfg.2017.07.198](https://doi.org/10.1016/j.promfg.2017.07.198). url : <https://www.sciencedirect.com/science/article/pii/S2351978917304067> (visité le 20/11/2025).
- [12] Ben Poole et al. *DreamFusion : Text-to-3D using 2D Diffusion*. 29 sept. 2022. doi : [10.48550/arXiv.2209.14988](https://doi.org/10.48550/arXiv.2209.14988). arXiv : 2209.14988[cs]. url : <http://arxiv.org/abs/2209.14988> (visité le 21/11/2025).
- [13] M. Raissi, P. Perdikaris et G.E. Karniadakis. "Physics-informed neural networks : A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations". In : *Journal of Computational Physics* 378 (fév. 2019), p. 686-707. issn : 00219991. doi : [10.1016/j.jcp.2018.10.045](https://doi.org/10.1016/j.jcp.2018.10.045). url : <https://linkinghub.elsevier.com/retrieve/pii/S0021999118307125> (visité le 13/11/2025).
- [14] Alvaro Sanchez-Gonzalez et al. *Learning to Simulate Complex Physics with Graph Networks*. 15 sept. 2020. doi : [10.48550/arXiv.2002.09405](https://doi.org/10.48550/arXiv.2002.09405). arXiv : 2002.09405[cs]. url : <http://arxiv.org/abs/2002.09405> (visité le 13/11/2025).
- [15] William R. Sherman et Alan B. Craig. *Understanding Virtual Reality : Interface, Application, and Design*. San Francisco, CA, USA : Morgan Kaufmann Publishers Inc., août 2002. 608 p. isbn : 978-0-08-052009-4.
- [16] David Silver et al. "Mastering the game of Go without human knowledge". In : *Nature* 550.7676 (oct. 2017). Publisher : Nature Publishing Group, p. 354-359. issn : 1476-4687. doi : [10.1038/nature24270](https://doi.org/10.1038/nature24270). url : <https://www.nature.com/articles/nature24270> (visité le 21/11/2025).
- [17] Fei Tao et al. "Digital Twin in Industry : State-of-the-Art". In : *IEEE Transactions on Industrial Informatics* 15.4 (avr. 2019), p. 2405-2415. issn : 1941-0050. doi : [10.1109/TII.2018.2873186](https://doi.org/10.1109/TII.2018.2873186). url : <https://ieeexplore.ieee.org/document/8477101> (visité le 20/11/2025).
- [18] Josh Tobin et al. "Domain randomization for transferring deep neural networks from simulation to the real world". In : *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). ISSN : 2153-0866. Sept. 2017, p. 23-30. doi : [10.1109/IROS.2017.8206648](https://doi.org/10.1109/IROS.2017.8206648) (visité le 20/11/2025).

1109/IR0S.2017.8202133. url : <https://ieeexplore.ieee.org/abstract/document/8202133> (visité le 21/11/2025).

- [19] Rui Wang et al. *Paired Open-Ended Trailblazer (POET) : Endlessly Generating Increasingly Complex and Diverse Learning Environments and Their Solutions*. 21 fév. 2019. doi : [10.48550/arXiv.1901.01753](https://doi.org/10.48550/arXiv.1901.01753). arXiv : [1901.01753\[cs\]](https://arxiv.org/abs/1901.01753). url : <http://arxiv.org/abs/1901.01753> (visité le 21/11/2025).
- [20] Zhengyi Wang et al. "ProlificDreamer : High-Fidelity and Diverse Text-to-3D Generation with Variational Score Distillation". In : ().