Evaluer et homologuer des systèmes de mobilité automatisés et autonomes intégrant des briques à base d'intelligence artificielle

Guillaume Perrin, Emmanuel Arbaretier, Christophe Bohn, Rafaël de Sousa Fernandes, Cédric Gava, Dominique Gruyer, Abdelmename Hedhli, Sio-Song Ieng, Thibaut Jonville, Pierre Jouve, et al.

Evaluer et homologuer des systèmes de mobilité automatisés et autonomes intégrant des briques à base d'intelligence artificielle

Author : E. Arbaretier, C. Bohn, R. De Sousa Fernandes, C. Gava, D. Gruyer, A. Hedhli, S.S. Ieng, T. Jonville, P. Jouve, J-F. Marlière, L. Maisonobe, K. Quintero, G. Perrin, R. Regnier.

Date : Septembre 2024

PRISSMA

Table des matières

Chapitre 1 Introduction

L'idée de véhicules autonomes remonte à plusieurs décennies. Les premières recherches sérieuses sur le sujet ont débuté dans les années 1980, avec des projets principalement expérimentaux visant à développer des systèmes de guidage automatique. Mais ce n'est qu'au début des années 2000, avec l'avènement des technologies de capteurs avancés et des algorithmes d'apprentissage automatique, que le concept de véhicule automatisé et connecté (VA) a pris son envol. Depuis lors, les constructeurs automobiles et les centres de recherche du monde entier investissent massivement dans le développement de ces technologies. Quelques prototypes ont commencé à émerger, des tests en environnements contrôlés et réels ont été menés, et des progrès significatifs ont été réalisés pour rendre le VA plus sûr et plus fiable. Il devient alors primordial d'anticiper et d'accompagner leur arrivée progressive sur le marché, en travaillant sur leur évaluation et leur potentielle homologation.

Pour appréhender les progrès dans le domaine du VA, il peut être intéressant de comprendre les niveaux d'automatisation établis par la Society of Automotive Engineers (SAE). Ces niveaux vont de zéro (aucune automatisation) à cinq (automatisation complète) et sont souvent pris pour base de réflexion par les différentes parties prenantes de l'écosystème :

- **Niveau 0** : automatisation pour des actions ou des alertes ponctuelles (aide au maintien dans la voie par exemple).
- **Niveau 1** : Assistance à certaines fonctions, comme le régulateur de vitesse adaptatif. L'automatisation de la conduite devient continue pour le longitudinal OU le latéral sous supervision du conducteur.
- **Niveau 2** : Automatisation partielle. Le véhicule peut gérer simultanément la direction et l'accélération, mais le conducteur doit rester attentif. En particulier, l'automatisation de la conduite devient continue pour le longitudinal ET le latéral sous supervision du conducteur.
- **Niveau 3** : Automatisation conditionnelle. Le véhicule peut gérer toutes les fonctions sous certaines conditions, mais le conducteur doit être prêt à reprendre le contrôle en cas de demande de reprise en main par le système de conduite automatisée.
- **Niveau 4** : Automatisation élevée. Le véhicule peut fonctionner de manière autonome grâce à un système de conduite automatisée dans des environnements prédéfinis ou des situations spécifiques.
- **Niveau 5** : Automatisation complète. Le véhicule peut fonctionner de manière autonome dans n'importe quel environnement ou situation.

Les attentes sur les VA sont nombreuses : il est par exemple attendu que ces nouveaux véhicules permettent de réduire le nombre d'accidents de la route en limitant les erreurs humaines, d'améliorer l'efficacité énergétique en proposant des conduites optimisées, ou encore d'offrir une mobilité accrue aux personnes âgées ou handicapées.

Cependant, ces innovations ne sont pas sans défis, et la transition vers une délégation des tâches de conduite du conducteur vers le véhicule devra passer par une évaluation rigoureuse de ces technologies, en particulier celles en lien avec l'IA qui sont souvent utilisées par les VA. Et c'est par l'application d'un cadre réglementaire adapté et le respect de critères de performance définis, qu'il devrait être rendu possible pour les autorités de garantir la sécurité, la fiabilité et l'acceptation sociétale de ces VA. C'est pourquoi, l'Union Européenne a adopté le règlement UE ADS 2022/1426 établissant des procédures uniformes et des spécifications techniques pour la réception par type des systèmes de conduite automatisée (ADS) des véhicules entièrement automatisés (destinés

à être intégrés sur des véhicules du niveau 4 d'automatisation selon la SAE).

La sécurité reste une des principales préoccupations de l'évaluation d'un VA. Les autorités doivent en effet s'assurer que ces nouveaux véhicules puissent naviguer en toute sécurité, réagir de manière adéquate aux imprévus, et minimiser les risques d'accidents. Il peut alors être pertinent d'être en mesure d'évaluer la fiabilité des capteurs dont ils dépendent, la précision des algorithmes (à base d'IA ou non) de perception et de prise de décision sur lesquels ils reposent, ainsi que leur capacité à interagir de manière sécurisée avec les autres usagers plus ou moins vulnérables de la route.

Les problématiques sociétales restent également nombreuses, et l'introduction des VA soulève par exemple des questions complexes en matière d'assurance et de responsabilité en cas d'accident impliquant un VA. En France par exemple, dans le cas d'incident impliquant un véhicule de niveau 4, il a été décidé que ce serait le constructeur qui serait jugé responsable des délits d'atteinte involontaire à la vie ou à l'intégrité des personnes impliquées (ordonnance n° 2021-443 du 14 avril 2021 relative au régime de responsabilité pénale applicable en cas de circulation d'un véhicule à délégation de conduite et à ses conditions d'utilisation). La situation est différente dans le cas des véhicules de niveau 3 : en phase de délégation de conduite ou en phase de reprise en main, c'est à nouveau le constructeur qui est jugé responsable, mais en phase de conduite manuelle, ce sera cette fois le conducteur qui sera tenu responsable. Par ailleurs, il est important que les décisions prises par les VA soient conformes aux normes éthiques et aux attentes sociétales. Quelles exigences formuler sur la façon dont un VA doit réagir sur la route, suite à un arbitrage toujours contestable du point de vue de l'éthique et de la morale ? Enfin, il semble utile que les processus d'homologation du VA soient en mesure d'intégrer des études sur l'acceptabilité et la perception publique de ces technologies, en encourageant notamment une transparence accrue sur leur fonctionnement et leurs limitations. En conclusion, l'évaluation des VA va bien au-delà de la simple conformité technique ; elle est censée façonner l'avenir de la mobilité tout en garantissant des niveaux de sécurité élevés, en intégrant les problèmes d'assurance et de responsabilité, en favorisant l'acceptation sociétale et en intégrant les défis éthiques et sociaux.

Les premières expérimentations de VA lancées en France intègrent le plus souvent non seulement des VA mais aussi des installations techniques (centre de supervision, feux connectés, caméras de détection de piétons, unités de bord de route...) qui permettent au système de mobilité dans son ensemble de fonctionner de manière sûre. Par système de mobilité automatisé et/ou autonome, il est entendu dans le présent livre, un système de transport routier automatisé et/ou autonome (STRA). Le périmètre d'étude va donc au-delà du véhicule seul. La définition de STRA est donnée dans le code des transports, au titre V du livre Ier de la troisième partie de la partie réglementaire (voir Décret n° 2021-873 du 29 juin 2021).

Cette notion de STRA est née du cadre législatif français, et correspond à un ensemble de véhicules hautement ou totalement automatisés et d'installations techniques permettant une intervention à distance (par exemple, le poste de supervision) ou participant à la sécurité (par exemple, des feux connectés déployés sur le parcours ou des caméras connectées déployées sur des intersections interagissant avec le système de conduite automatisé), déployés sur un parcours ou des zones de circulation prédéfinis, et complété de règles d'exploitation, d'entretien et de maintenance, aux fins de fournir un service de transport routier public collectif ou particulier de personnes, ou de service privé de transport de personnes.

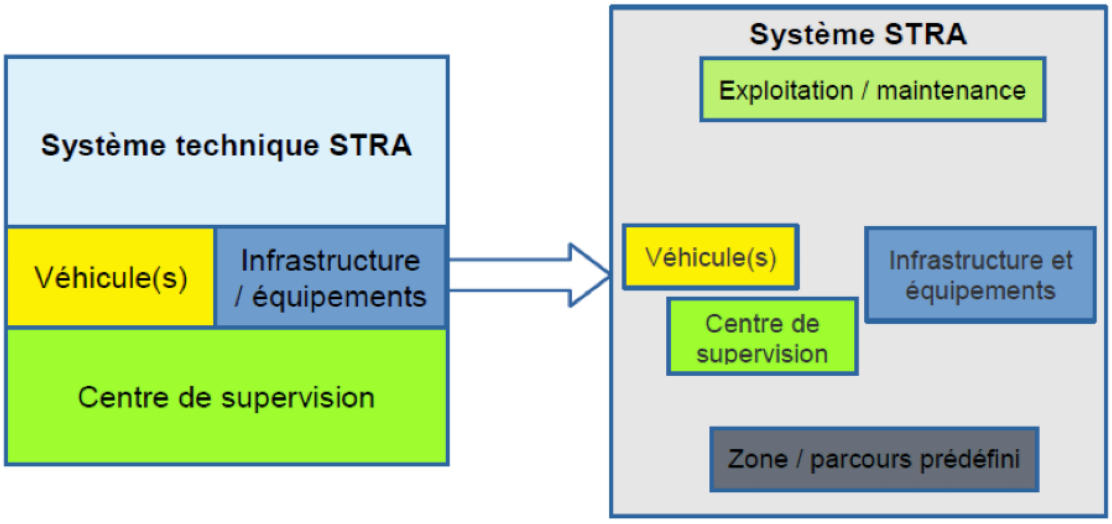

Figure 1.1 – Périmètre d'un STRA et de son système technique.

Préalablement à la mise en service d'un tel système, une démonstration de sécurité à l'échelle du système doit être menée par des entités responsables. Celle-ci est ensuite évaluée par un ou plusieurs organismes qualifiés agréés qui émettent un avis sur les dossiers techniques et de sécurité supportant la démonstration. En pré-requis de cette démonstration de sécurité de niveau système, les véhicules inclus dans le STRA doivent être homologués. L'homologation des systèmes de conduite automatisée équipant les véhicules entièrement automatisés (assimilables au niveau SAE4), est quant à elle cadrée par le règlement européen 2022/1426 du 5 août 2022. Le périmètre de ce livre blanc couvre l'ensemble des systèmes à base d'IA, qu'ils soient implantés directement dans les véhicules ou dans les installations techniques débarquées.

C'est dans ce contexte qu'a été lancé en 2021 le projet PRISSMA (Plateforme de Recherche et d'Investissement pour la Sûreté et la Sécurité de la Mobilité Autonome). Ce projet est la réponse proposée par la filière de la mobilité automatisée/autonome au Pilier 2 du Grand Défi "Sécuriser, certifier et fiabiliser les systèmes fondés sur l'intelligence artificielle" en partenariat avec le ministère de la Transition écologique (au travers notamment de la DGITM et de la DGEC), afin de travailler sur la sécurisation, la fiabilisation, et à terme, de délimiter les contours d'une homologation liée à l'utilisation de l'IA pour la mobilité automatisée/autonome. Les objectifs principaux de ce projet auront été d'identifier et de recenser les objectifs de sécurité et de sûreté pour les systèmes de mobilité automatisé intégrant des systèmes à base d'IA, puis de développer les processus complets de validation de la fiabilité en vue de leur potentielle mise en exploitation commerciale.

L'objectif de ce livre blanc est d'en recenser les principales conclusions et recommandations. Pour plus de détails, le lecteur intéressé pourra se reporter à l'ensemble des livrables produits par le projet, téléchargeables sur le site https ://prissma.univ-gustave-eiffel.fr et listées en Section 4.4, ainsi qu'aux références associées. Dans la suite de ce livre blanc, nous nous focaliserons ainsi sur les VA de niveau 3 ou 4 intégrant au moins une fonction à base d'IA, intégrés dans des systèmes de transport routier automatisés (STRA).

Définitions clés

Afin de faciliter la lecture de la suite de ce document, nous précisons ici plusieurs définitions importantes et tirées de la taxonomie commune construite dans le cadre des activités du projet PRISSMA :

- **Action** : Ensemble des opérations effectuées par le système (freinage, accélération, changement de voie, demande de reprise en main, activation, clignotant . . .)
- **Cas d'usage** : Spécification d'un domaine d'utilisation, comportant éventuellement les informations suivantes sur le système :
  - Un ou plusieurs scénarios,
  - Son domaine de fonctionnement,
  - Son comportement attendu,
  - Ses limites de fonctionnement.
- **Danger** : Circonstance pouvant mener à un accident
- **Domaine de conception opérationnelle (ODD)** : Conditions de fonctionnement dans lesquelles un système de conduite automatisé donné ou une de ses caractéristiques est spécifiquement conçu pour fonctionner, y compris :
  - L'infrastructure physique ;
  - Le scénario ;
  - Conditions environnementales ;
  - Conditions de circulation ;
  - L'infrastructure numérique ;
  - Capacités des véhicules.
- **Évènement** : Tout ce qui se produit ou apparaît dans un système (ou dans son environnement extérieur) à un temps-espace donné et qu'on peut détecter ou caractériser par un dispositif logiciel ou matériel du système. Généralement, un évènement doit être pris en compte par le système s'il influe sur la procédure de prise de décision.
- **Infrastructure** : L'infrastructure est l'ensemble des équipements constituant l'infrastructure routière et l'infrastructure numérique.
- **Manoeuvre à risque minimal** : Manoeuvre ayant pour finalité la mise à l'arrêt du véhicule en situation de risque minimal pour ses occupants et les autres usagers de la route, automatiquement effectuée par le STRA, suite à un aléa non prévu dans ses conditions d'utilisation, à une défaillance grave ou un défaut de reprise en main à expiration de la période de transition.
- **Parcours ou zone de circulation prédéfini** : Ensemble des sections routières ou espace dont les limites géographiques sont définies, sur lesquelles est prévue la circulation ou l'arrêt d'un ou plusieurs véhicules d'un STRA.
- **Précision** : capacité d'un système à produire des résultats exacts et fiables. Pour un algorithme d'IA, cela signifie que l'algorithme doit pouvoir détecter, classer et réagir correctement aux différents éléments de l'environnement avec un haut degré de fiabilité.
- **Résilience** : capacité d'un système à se remettre rapidement et efficacement de perturbations ou de défaillances. Pour un algorithme d'IA, il doit démontrer qu'il peut non seulement tolérer des erreurs et des incidents imprévus, mais aussi récupérer rapidement sans compromettre la sécurité ou les performances du véhicule. La résilience implique également que l'algorithme puisse continuer à fonctionner dans des conditions dégradées et prendre des mesures correctives pour revenir à un état normal.

- **Robustesse** : capacité d'un système à maintenir ses performances et à fonctionner correctement malgré des conditions variables et imprévues. Cela implique que le système doit être capable de gérer des perturbations, des bruits dans les données d'entrée et des scénarios exceptionnels sans défaillance significative.
- **"Dynamic Driving Task" ou Tâche de Conduite Dynamique (DDT)** : ensemble des activités liées à la conduite d'un véhicule.
- **Réponse à la détection d'objets et d'évènements (OEDR)** : Sous-tâches de la DDT qui inclut la surveillance de l'environnement de conduite (détection, reconnaissance et classification des objets et des événements, et préparation de la réponse adéquate), ainsi que l'exécution d'une réponse appropriée à de tels objets et événements (i.e. la réponse nécessaire pour poursuivre la DDT et/ou effectuer une manoeuvre à risque minimal).
- **Scénario** : Séquence temporelle de scènes entrecoupées incluant des actions et événements.
- **Scène** : État donné du système considéré et de son environnement (objets, acteurs, infrastructure routière, climat, . . .) à un instant "t" choisi ou observé. Elle définit les éléments de décors statiques (scénographie) et dynamiques, ainsi que les acteurs du scénario par leurs paramètres spécifiants et leurs valeurs.
- **Système** : Ensemble d'éléments (matériels, logiciels ou humains) reliés entre eux, considérés comme un tout dans un contexte défini et organisé de sorte à atteindre un objectif donné, dans certaines conditions.
- **Système de conduite automatisée (ADS)** : Système associant des éléments matériels et logiciels, permettant d'exercer le contrôle dynamique d'un véhicule de façon prolongée.

Chapitre 2 L'intelligence artificielle au service des STRA

Dans ce livre blanc, nous nous limitons aux systèmes de mobilité routière autonomes/automatisés (STRA) intégrant une ou plusieurs briques IA. Ces briques IA sont dans la suite considérées comme fonctionnelles et ayant fait l'objet d'une validation préalable rigoureuse. En particulier, le processus d'apprentissage est considéré comme terminé, et les algorithmes embarqués sont supposés déterministes, au sens où s'ils sont soumis plusieurs fois exactement aux mêmes sollicitations, ils répondent à chaque fois exactement de la même manière.

2.1 Composants clés des IAs dans les VAs

Les STRA intègrent le plus souvent un ou plusieurs composants clés à base d'IA pour assurer leur fonctionnement sûr. Ces composants (que l'on peut trouver au niveau du véhicule et/ou de l'infrastructure) peuvent être regroupés en quatre catégories principales : les capteurs et sources d'information, les systèmes de perception et de localisation, les processus de décision et de suivi de trajectoire, et les systèmes de contrôle/supervision.

2.1.1 Capteurs de perception et sources d'information

Les êtres humains utilisent leurs différents sens pour percevoir (sens visuel, auditif, tactile, gustatif, olfactif), se déplacer (organes proprioceptifs), et interagir avec l'environnement. Les capteurs extéroceptifs et proprioceptifs embarqués dans les véhicules autonomes/automatisés ont le même objectif. Les différentes technologies de capteurs utilisées permettent de collecter les données nécessaires à la perception et à l'interprétation de l'environnement qui entoure le véhicule. Les capteurs proprioceptifs permettent de pouvoir estimer à tout moment l'état interne et dynamique du véhicule. Les technologies de capteurs couramment utilisés pour le développement des STRA incluent les LiDAR, les RADARs, les caméras (visuelle, fish-eye, omnidirectionnelle, neuromorphique, infra-rouge, polarisée , ...) et les capteurs ultrasons. Les LiDARs émettent des faisceaux laser dans les longueurs d'ondes proche-infrarouge (850 nm à 1550 nm) pour créer des cartes d'impacts 3D précises de l'environnement immédiat, tandis que les RADARs utilisent des ondes radio (24 GHz à 79 GHz pour les applications automobiles) pour détecter la vitesse et la distance des objets. Les caméras fournissent des images dans différentes longueurs d'ondes en fonction de la technologie utilisée. Généralement les caméras fonctionnent dans le spectre visible. Cette technologie, très riche en information produite, permet de mettre en oeuvre des applications de détection et de reconnaissance des objets, des panneaux de signalisation, des marquages au sol et des conditions de visibilité. Les capteurs ultrasons utilisent des fréquences beaucoup plus basses et sont généralement utilisés pour des tâches de proximité, comme le stationnement. Pour tous ces capteurs et pour toutes les tâches et traitements présentés dans la figure 2.1, l'IA joue de plus en plus un rôle crucial. Le traitement et la transformation des informations reçues (potentiellement bruitées ou perturbées, voir figure 2.2) permettent de produire des données plus synthétiques et haut-niveau utilisables par les autres systèmes aval du véhicule (comme la prise de décision ou la supervision).

2.1.2 Systèmes de perception et de localisation

Les systèmes de perception constituent la partie amont du "cerveau" du véhicule autonome/automatisé. Comme présenté dans la figure 2.3, les traitements impliqués dans la perception vont permettre de filtrer, de

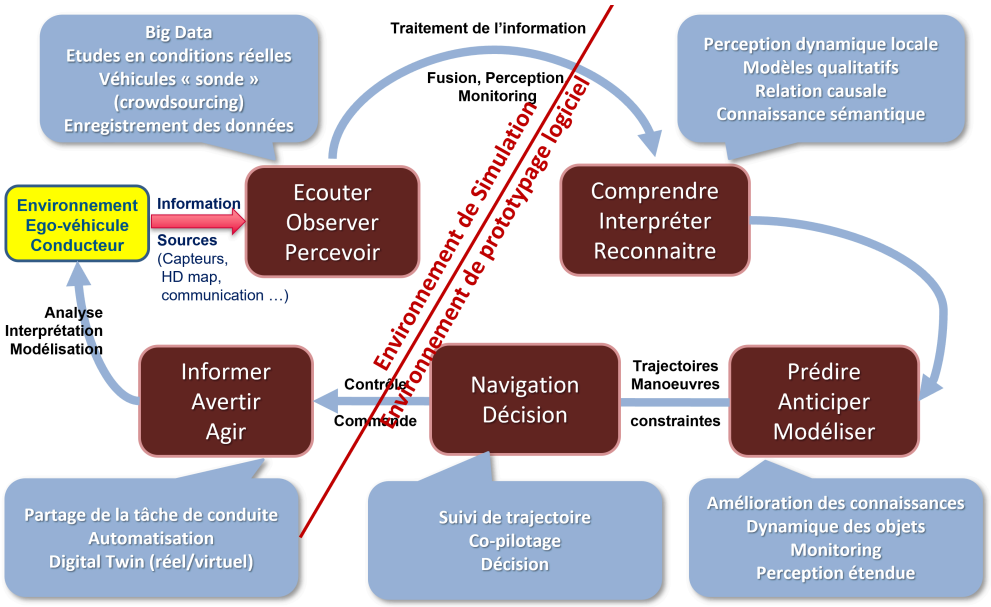

FIGURE 2.1 – Composants, fonctions, et propagation de l'information dans le développement d'un système d'automatisation de la conduite.

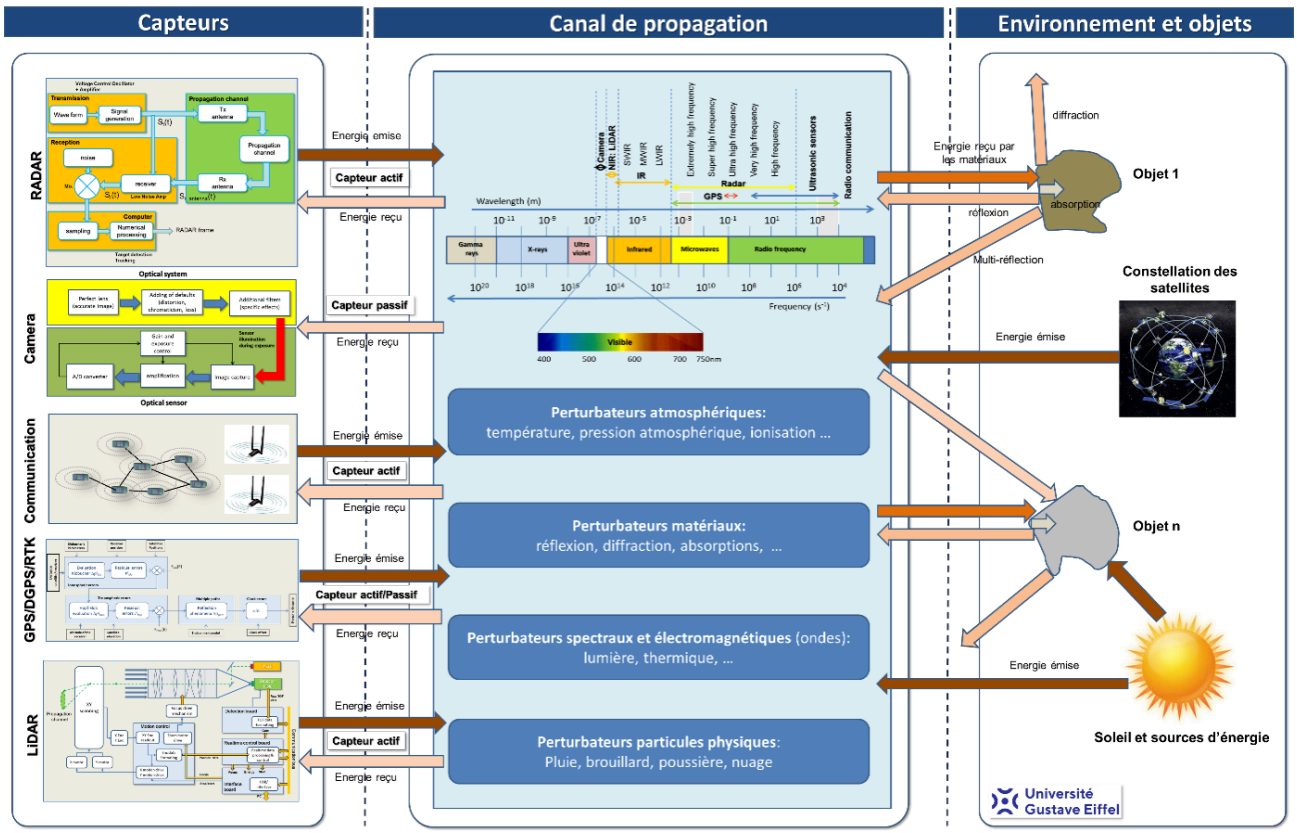

FIGURE 2.2 – Les capteurs, le canal de propagation, les matériaux, et les sources d'énergie extérieures : un écosystème compliqué impliquant quatre classes de perturbateurs impactant la qualité des données et du fonctionnement des capteurs.

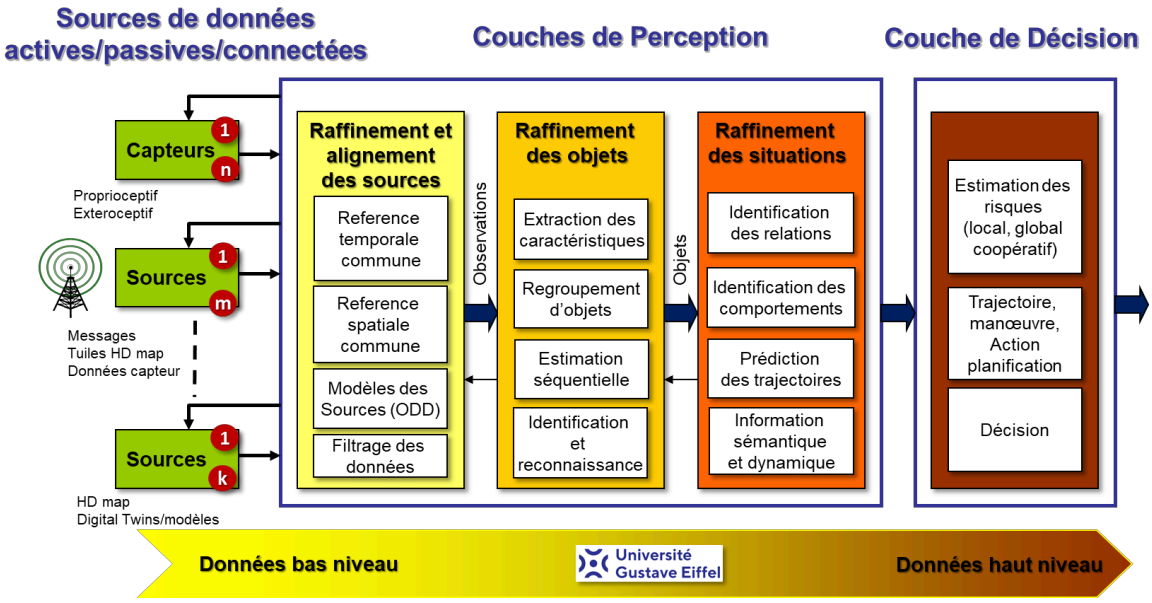

FIGURE 2.3 – Perception : des données brutes à la compréhension et à la modélisation de l'environnement, des comportements, et des situations.

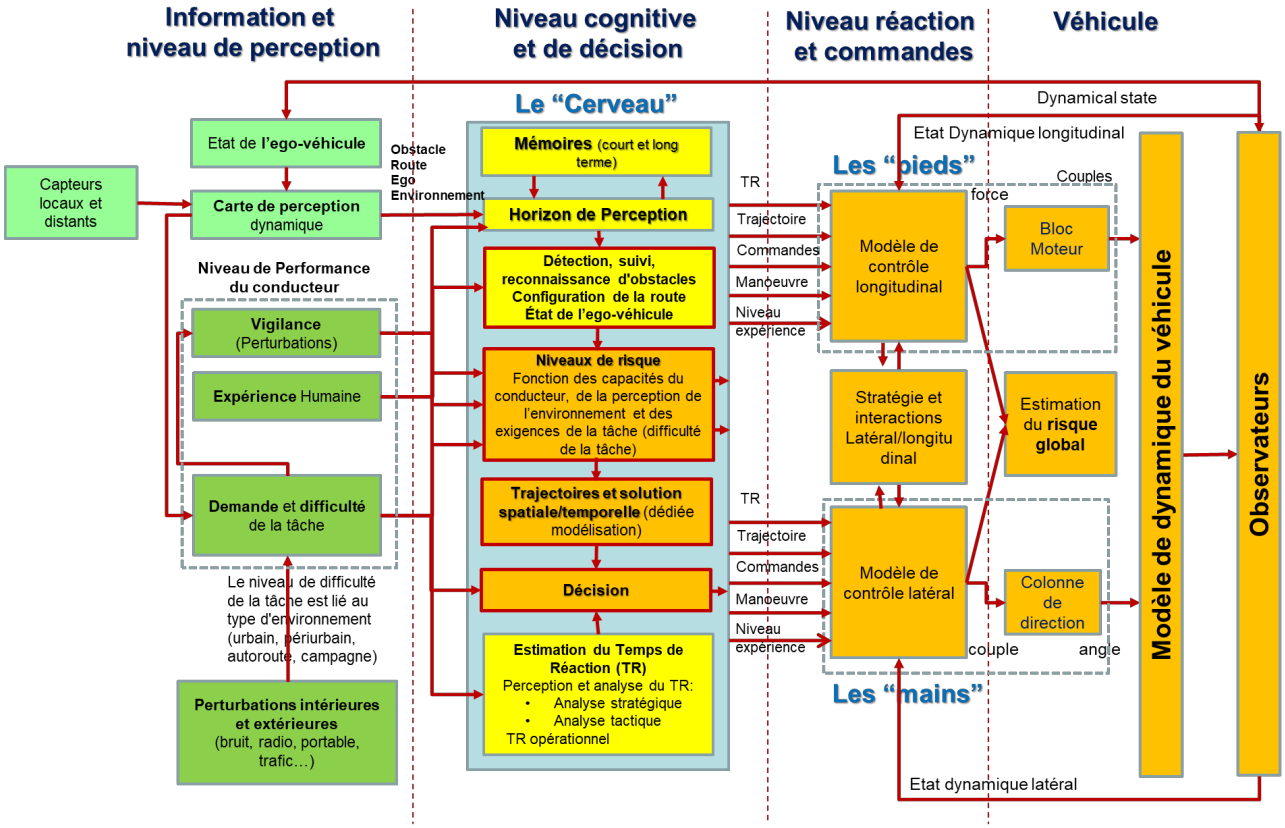

FIGURE 2.4 – Adéquation entre copilote virtuel et pilote humain, avec le "cerveau", les "pieds", et les "mains".

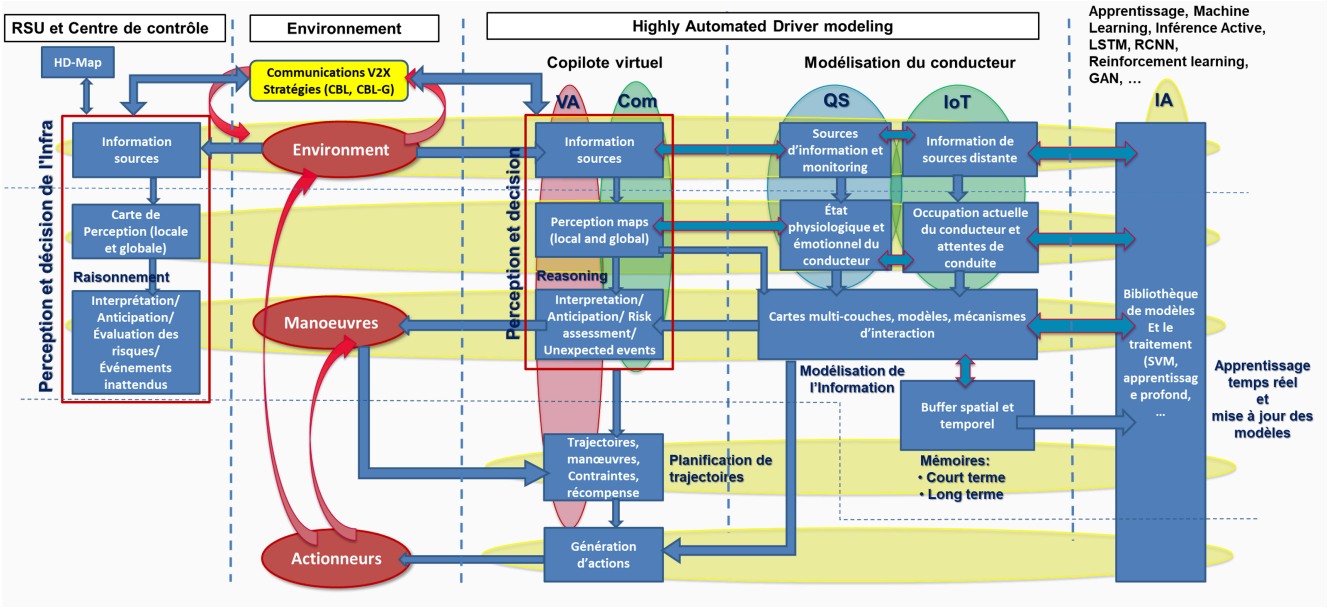

FIGURE 2.5 – Vision d'ensemble d'une architecture de copilote virtuel pour la conduite automatisée. Dans toutes les couches, les méthodes à base d'IA sont de plus en plus présentes.

mettre en forme, et de recaler les données brutes afin de pouvoir appliquer des opérateurs de fusion, de combinaison, d'association, de regroupement, de régression, de classification permettant de construire des cartes de perception dynamique locale contenant les informations nécessaires à la compréhension de l'environnement. Ces systèmes utilisent des algorithmes d'IA de traitement d'image, de reconnaissance d'objets et de fusion de données pour identifier les éléments en lien avec la trajectoire du VA tels que les autres véhicules, les piétons, les feux de signalisation et les obstacles. Les réseaux de neurones convolutifs (CNN) et les techniques de vision par ordinateur sont couramment employés pour ces tâches. La fusion de données combine les informations provenant de différents capteurs pour créer une représentation cohérente et précise de l'environnement, essentielle pour la navigation et la prise de décision.

2.1.3 Processus de Décision et de planification de trajectoire et de manoeuvres

Les processus de décision ont la responsabilité de déterminer les actions que le véhicule doit entreprendre en fonction de l'environnement perçu. Ces systèmes peuvent utiliser des algorithmes d'IA de planification de trajectoire et des modèles de prédiction pour anticiper les mouvements des autres usagers de la route et choisir la meilleure route à suivre. L'apprentissage par renforcement et les réseaux de neurones profonds sont souvent utilisés pour optimiser ces décisions en temps réel. Ces algorithmes doivent être capables de gérer des situations complexes et dynamiques, en prenant en compte les règles de la circulation, les conditions de circulation et les comportements des autres usagers. La figure 2.4 présente une vue d'ensemble des niveaux de traitement à aborder dans le développement des STRA. Les systèmes à base d'IA peuvent aussi bien intervenir dans le niveau "perception" que dans les niveaux "cognitif" et "décision", et dans la couche "réaction" et "commande". Néanmoins, mettre en oeuvre différentes méthodes à base d'IA dans les 3 couches de traitement pour aboutir à un contrôle des manoeuvres et du comportement d'un véhicule pose clairement un problème d'interprétabilité et d'explicabilité du comportement de ces algorithmes. La figure 2.5 montre le niveau de complexité en cas d'extension de l'ADS avec l'intégration de modèle de comportement cognitif humain et un accès à des informations distantes provenant de l'infrastructure et des autres véhicules.

2.1.4 Systèmes de contrôle et de supervision

Les systèmes de contrôle et de supervision exécutent les décisions prises par les processus de décision en contrôlant les divers composants du véhicule, tels que la direction, l'accélération et le freinage. Les algorithmes de contrôle utilisent des modèles de contrôle prédictifs et des systèmes de régulation pour garantir que le véhicule suive la trajectoire planifiée de manière sûre et efficace. Ces systèmes doivent également surveiller en continu les performances du véhicule et ajuster les commandes en réponse aux changements de l'environnement et/ou aux imprévus. Les systèmes de supervision incluent des mécanismes de redondance et de diagnostic pour détecter et gérer les pannes potentielles, assurant ainsi la sécurité et la fiabilité du véhicule automatisé/autonome.

En conclusion, les véhicules automatisés/autonomes reposent sur une intégration complexe et sophistiquée de diverses briques IA, chacune jouant un rôle crucial dans la perception, la décision et le contrôle. Une synergie entre ces composants est primordiale pour permettre aux véhicules automatisés/autonomes de naviguer de manière sûre et la plus autonome possible dans des environnements réels complexes, en interaction notamment avec une grande diversité d'acteurs.

2.2 Cadre Juridique et Normatif

2.2.1 Réglementations existantes

Pour mieux comprendre le cadre législatif dans lequel doit s'inscrire un STRA qui viserait une mise en circulation sur les routes françaises, nous prenons ici pour référence le cadre de réception pour les véhicules en Europe synthétisé en figure 2.6. Le règlement européen 2018/858 établit effectivement les règles administratives et techniques pour l'homologation et la mise sur le marché de nouveaux véhicules, systèmes, composants et équipements. Il inclut notamment des dispositions pour les pièces et équipements pouvant poser un risque pour le bon fonctionnement des véhicules en plus de définir les exigences de surveillance du marché pour ces éléments.

C'est également dans ce règlement cadre que sont détaillées les procédures à destination des constructeurs pour la demande d'homologation, y compris la soumission de la documentation technique, des échantillons et des rapports d'essai.

Les constructeurs de véhicules sont tenus de veiller à ce que les véhicules et les composants répondent aux exigences du règlement européen 2018/858 et des règlements/directives appelés par ce dernier. Ils doivent également tenir à jour la documentation et fournir les informations nécessaires aux autorités lorsque demandé (Homologation, Conformité de Production, Surveillance de marché...). Ils sont également responsables, envers l'autorité compétente en matière de réception, de tous les aspects de la procédure de réception, ainsi que de la conformité de la production tout au long de la durée de vie du véhicule.

En plus du cadre général pour la réception d'un véhicule en Europe, le règlement UE 2018/858 établit également la liste des règlements auxquels il est nécessaire de répondre pour la réception européenne par type des véhicules entièrement automatisés et produits en petites séries conformément à l'article 41 de l'UE2018/858.

Enfin, toujours dans ce cadre de réception des véhicules, la Commission Européenne a publié le règlement UE2022/1426, qui établit des règles, des procédures et des spécifications techniques pour la réception par type des systèmes de conduite automatisée (ADS) des véhicules entièrement automatisés. Il définit les cas d'utilisation des véhicules entièrement automatisés qui sont couverts par une homologation et les exigences pour les scénarios de trafic critiques, où le système ADS doit être capable d'éviter des collisions raisonnablement prévisibles et de minimiser les risques pour la sécurité pour la sécurité des occupants et des autres usagers de la route.

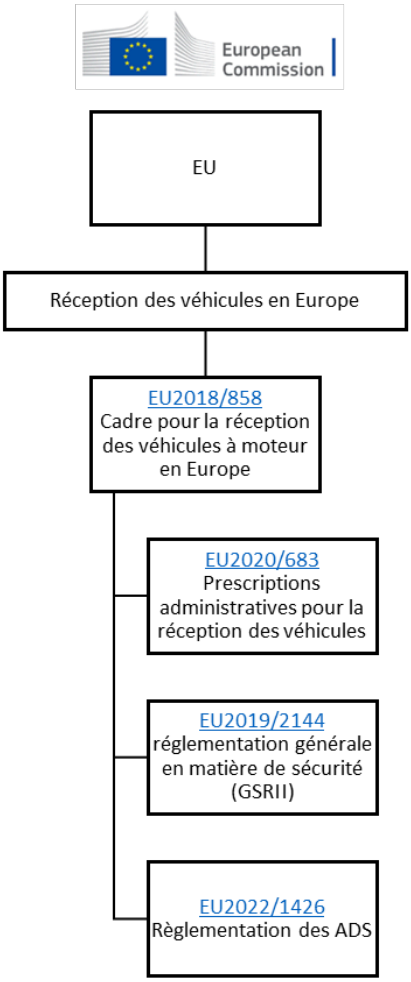

FIGURE 2.6 – Cadre réglementaire en vigueur actuellement pour l'homologation d'un véhicule équipé d'un ADS

Concernant l'homologation des composants à base d'IA dans les véhicules mis sur le marché européen, le règlement UE 2024/1689 DU PARLEMENT EUROPÉEN ET DU CONSEIL du 13 juin 2024 établissant des règles harmonisées concernant l'intelligence artificielle (appelé aussi "IA Act") a été publié le 12 juillet 2024 et entre en vigueur le 1er aout 2024, et sera pleinement applicable 24 mois après, à l'exception

- des interdictions de pratiques prohibées, qui s'appliqueront six mois après la date d'entrée en vigueur,
- des codes de pratique (neuf mois après l'entrée en vigueur),
- des règles générales en matière d'IA, y compris la gouvernance (12 mois après l'entrée en vigueur),
- des obligations pour les systèmes à haut risque (36 mois).

L'automobile entre dans la catégorie des systèmes à haut risque.

Enfin, il est prévu que l'IA Act soit applicable à l'industrie automobile par le biais des règlements UE, UE 2018/858 et UE 2019/2144 (Global Safety Regulation II - GSRII). Lorsque des systèmes d'IA doivent être déployés en tant que composants de sécurité d'un type de véhicule, les exigences énoncées au Titre III, Chapitre 2, de l'IA Act doivent être prises en compte. Les articles suivants doivent en particulier être considérés avec attention :

- Article 09 : Système de gestion des risques ;
- Article 10 : Données et Gouvernance des données ;
- Article 11 : Documentation technique ;
- Article 12 : Conservation des archives ;
- Article 13 : Transparence et fourniture d'informations aux utilisateurs ;
- Article 14 : Supervision humaine ;
- Article 15 : Précision, Robustesse et Cybersécurité.

C'est donc avec ce cadre réglementaire, en place et à venir, en tête, que le projet PRISSMA vise à proposer une plateforme de validation suffisamment robuste pour l'homologation et la surveillance continue des véhicules et de leurs composants IA, afin de s'assurer que les véhicules entièrement automatisés puissent être mis sur le marché européen en toute sécurité.

2.2.2 Défis et lacunes pour les nouveaux systèmes de mobilité à base d'IA

Le cadre de déploiement des STRA à base d'IA présente plusieurs manques, tant techniques et réglementaires que sociétaux et éthiques.

Les premiers sont d'ordre technique.

- **Perception et compréhension de l'environnement** : Les systèmes de capteurs, tels que les caméras, les LiDARs et les RADARs, doivent fournir des données précises en temps réel. Cependant, les conditions météorologiques défavorables, comme la neige ou la pluie, la variété des événements possibles ou des infrastructures et la multitude d'acteurs sur la route peuvent grandement affecter leur performance.
- **Navigation et localisation** : La précision de la navigation GPS peut être insuffisante dans des environnements urbains denses. Les véhicules doivent également être capables de se repérer dans des zones où les signaux GPS sont faibles ou inexistants. Cela nécessite notamment des cartes numérique de très haute précision (HD map) précises et mises à jour très régulièrement pour pallier les modifications du système routier (travaux notamment).
- **Planification et prise de décision** : Les algorithmes d'IA doivent prendre des décisions complexes en temps réel, ce qui nécessite une puissance de calcul élevée et une capacité à anticiper les comportements des autres usagers de la route, ce qui peut être complexe notamment pour les piétons ou animaux.

- **Cybersécurité des systèmes** : Assurer la cybersécurité des STRA pour prévenir les piratages et les interférences malveillantes est un défi majeur surtout si le véhicule est connecté.
- **Coût de développement et de déploiement** : Le développement de technologies de conduite automatisée/autonome est coûteux, et le cout de production des véhicules peut être élevé.
- **Infrastructures adaptées** : Les infrastructures routières doivent évoluer pour mieux s'adapter aux véhicules automatisés/autonomes, avec des routes intelligentes, des panneaux de signalisation numériques et des systèmes de communication véhicule-infrastructure (V2I).
- **Collecte et gestion des données** : Les STRA génèrent une quantité énorme de données. La collecte, le stockage notamment dans un but de traçabilité, et l'analyse de ces données soulèvent des questions sur la gestion des données et la protection de la vie privée.
- **Interopérabilité des systèmes** : Assurer que différents systèmes de véhicules automatisés/autonomes peuvent interagir de manière fluide et sécurisée est crucial pour un déploiement harmonieux.

Mais d'autres difficultés plus sociétales se posent.

- **Cadre réglementaire** : Les lois et réglementations varient considérablement d'un pays à l'autre et même entre les différentes juridictions d'un même pays. Il est ainsi nécessaire de développer un cadre réglementaire harmonisé pour faciliter le déploiement à grande échelle.
- **Normes de sécurité** : Établir des normes de sécurité universelles pour les véhicules automatisés/autonomes est crucial, mais il n'existe actuellement pas de consensus global sur ces normes.
- **Acceptation par le public** : Le degré de confiance du public dans les véhicules automatisés/autonomes est variable. Les utilisateurs doivent être convaincus de la sécurité et de la fiabilité de ces technologies.
- **Impacts sur l'emploi** : Le remplacement potentiel des chauffeurs professionnels par des véhicules automatisés/autonomes soulève des préoccupations concernant l'emploi et la reconversion professionnelle.
- **Décisions éthiques** : Les véhicules automatisés/autonomes doivent être programmés pour prendre des décisions éthiques complexes dans des situations d'urgence. Par exemple, comment un véhicule automatisé/autonome doit-il prioriser la sécurité des passagers par rapport aux piétons en cas de collision inévitable ?

2.2.3 Initiatives internationales

Si le cadre réglementaire existe, l'intégration de systèmes à base d'IA dans les véhicules automatisés/autonomes pose néanmoins une série de défis nouveaux et spécifiques. Dans ce contexte, plusieurs initiatives ont vu le jour avec une prise en compte plus ou moins poussée de la notion d'IA.

Par exemple, l'ISO (International Organization for Standardization) travaille, par l'intermédiaire du comité technique TC 22 sur les véhicules routiers, à la standardisation du processus de prise en compte de la sécurité, de la conception à la validation, par l'intermédiaire des documents ISO TR 4804 (Technical Report) et de son successeur ISO TS 5083 (Technical Specification) en cours d'élaboration (ISO, 2020). L'ISO travaille également sur la standardisation des scénarios de test pour les systèmes automatisés par l'intermédiaire des normes ISO 34501 à 34505. La norme ISO 34503 porte notamment sur une taxonomie des ODD (Operational Design Domain), initiative la plus récente d'un sujet ayant connu de multiples occurrences.

Si l'Europe est déjà dotée d'un cadre réglementaire pour les ADS (ADS Act - règlement UE2022/1426), la CEE-ONU (Commission Economique Européenne pour les Nations Unies), travaille également à la rédaction d'un règlement des Nations Unies (UNR) et d'un GTR (Global Technical Regulation), applicables respective-

ment par les parties contractantes des accords de 1958 et 1998. Ce travail de rédaction se base principalement sur les travaux déjà réalisés par les groupes de travail du FRAV (Functionnal Requirements for Automated Vehicles) et du VMAD (Validation Method for Automated Driving), qui avaient déjà alors servi de base à la rédaction du règlement européen 2022/1426. Ces travaux ont notamment donné lieu à la publication de la NATM (New Assessment Test Method), définissant les nouvelles pratiques à mettre en oeuvre pour la validation d'un système de conduite automatisé (ECE, 2022).

Le corps de la méthode NATM s'appuie sur cinq piliers qui ont été repris dans le projet PRISSMA (pour en développer les parties traitant spécifiquement les briques IA) :

- Trois piliers abordent les différents moyens d'essai que sont la simulation, les tests sur piste et les essais routiers, utilisés de manière combinée.
- Le quatrième pilier relève de l'audit et de la production d'un dossier de sécurité explicitant les éléments caractéristiques du système, les résultats d'analyse et les processus suivis pour son développement.
- Le cinquième définit les éléments à mettre en place pour le suivi en opération ainsi que la remontée d'informations vers les autorités.

A noter que la base de données de scénarios utilisée pour la validation est également considérée comme un atout fondamental qui peut être présenté comme un pilier supplémentaire dans certaines publications du VMAD. Néanmoins, ces documents comportent peu d'éléments facilitant la mise en place de moyens de démonstration de conformité aux exigences qu'ils décrivent, et ceci est encore plus vrai pour les fonctions reposant sur des briques IA.

2.3 Méthodes existantes de validation des performances de l'IA

2.3.1 Procédures classiques

Les méthodes classiques de validation des performances de l'IA reposent sur une série d'approches quantitatives et qualitatives visant à évaluer l'efficacité, la précision ou la robustesse des modèles. Parmi ces méthodes, les tests sur des ensembles de données de validation et de test sont couramment utilisés. Ces ensembles de données, séparés du jeu d'entraînement, permettent de mesurer la capacité du modèle à généraliser ses prédictions à des données inédites. Les métriques telles que la précision, le rappel, la F-mesure et l'aire sous la courbe ROC sont souvent utilisées pour évaluer les performances des modèles de classification. Pour les modèles de régression, des métriques comme l'erreur quadratique moyenne (aussi appelée MSE pour "mean-squared error") et le coefficient de détermination (R^2) sont couramment appliquées. Outre ces approches quantitatives, des méthodes qualitatives comme l'analyse des erreurs et les études de cas spécifiques aident à comprendre les limitations et les points forts des modèles d'IA dans des contextes particuliers. Des techniques de validation croisée, où le jeu de données est divisé en plusieurs sous-ensembles pour des validations multiples, et des techniques de régularisation pour éviter le sur-apprentissage, sont également couramment employées pour assurer une évaluation rigoureuse et robuste des modèles d'IA.

Au fil des années, la validation des IA s'est articulée autour de plusieurs écoles de pensée distinctes.

- L'**"évaluation du meilleur cas"** se concentre sur l'évaluation des performances des IA dans des conditions idéales, souvent avec des ensembles de données optimisés, pour démontrer le potentiel maximal des modèles. Cette approche est très utilisée lors des premiers essais en conception et est particulièrement appliquée pour l'évaluation de l'IA générative, notamment à travers le "Prompt engineering", mais ne permet pas d'assurer la sécurité pour un système critique.

- Au contraire, l'approche **IVVQ** (Intégration, Vérification, Validation et Qualification)) est orientée vers les tests, et s'appuie sur des processus rigoureux de certification pour garantir que les systèmes d'IA respectent des normes strictes de fiabilité et de sécurité. Cette école découle directement du référentiel classique de l'ingénierie système qui consiste à mettre en oeuvre des provessus de validation en cohérence avec les spécifications initiales. Définir de telles spécifications pratiques pour des algorithmes d'IA est néanmoins loin d'être trivial. C'est notamment pour cela que le projet PRISSMA a dédié une partie de ses travaux à raffiner les exigences et les spécifications autour de l'utilisation de l'IA pour les STRA.
- Une troisième école, l'école du **benchmark**, utilise des ensembles de données standardisés et des critères de performance pour comparer différents modèles d'IA. Cette approche facilite ainsi les évaluations comparatives, mais rien n'indique que les critères de comparaison ou le choix des bases de données ne soient pas biaisés pour favoriser un système par rapport à un autre.
- L'**"évaluation du pire cas"** met quant à elle l'accent sur l'analyse des performances des IA dans des conditions défavorables ou extrêmes, afin de comprendre les limites et les vulnérabilités des systèmes.
- Enfin, l'école dite **cognitive** vise à évaluer la capacité des IA à simuler des processus cognitifs humains, en se concentrant sur des aspects tels que la compréhension, la prise de décision et l'apprentissage adaptatif.

Chacune de ces approches apporte une perspective unique et complémentaire pour assurer une validation exhaustive et robuste des systèmes d'IA mais celle qui semble la plus proche d'être adoptée par les industriels de l'automobile reste l'approche IVVQ.

2.3.2 Défis et Lacunes du cadre existant pour la validation des systèmes critiques

Malgré les avancées significatives dans les méthodes de validation des performances de l'IA, le cadre existant présente plusieurs défis et lacunes, particulièrement pour les systèmes critiques. L'un des principaux défis réside dans la complexité et l'opacité des modèles d'IA, souvent manipulés en "boîtes noires" même pour le développeur, ce qui rend difficile l'interprétation et la compréhension des décisions qu'ils prennent. Cette opacité pose également des problèmes de traçabilité et de responsabilité, essentiels pour les systèmes critiques tels que ceux utilisés dans les transports. De plus, les ensembles de données de validation actuels peuvent ne pas représenter adéquatement toutes les situations et les scénarios extrêmes que ces systèmes peuvent rencontrer dans le monde réel, ce qui soulève des questions sur leur robustesse et leur fiabilité. Il existe également un manque de normes et de réglementations spécifiques pour encadrer l'utilisation de l'IA dans les systèmes critiques, ce qui complique la mise en place de protocoles de validation rigoureux. Le règlement européen UE 2024/1689 (IA Act) est sensé donner un cadre, mais le manque de normes harmonisées ou le manque de procédures pour permettre la mise en conformité à ce règlement se fait toujours sentir. Nous en avons déjà parlé, mais les défis liés à l'éthique, notamment en matière de biais et de discrimination, constituent toujours un autre obstacle majeur. Les modèles d'IA peuvent en effet reproduire et amplifier les biais présents dans les données d'entraînement, entraînant des décisions potentiellement injustes ou dangereuses dans des contextes critiques. Enfin, l'intégration de ces systèmes dans des environnements existants pose des problèmes de compatibilité et de stabilité, nécessitant des méthodes de validation capables d'assurer une cohabitation sans faille avec les systèmes traditionnels, ce qui est particulièrement important dans le domaine automobile afin de rendre possible la cohabitation avec les usagers de véhicules non automatisés. Pour surmonter ces défis, il est ainsi crucial de développer des approches de validation plus transparentes, plus robustes et plus normées, intégrant de manière systématique des considérations éthiques et de gestion des risques.

Chapitre 3 Méthodologie PRISSMA d'évaluation des briques IA dans les STRA

3.1 Stratégie générale

Pendant longtemps, l'évaluation et la validation des composants et des applications ont suivi le cycle de conception classique appelé "cycle en V". Dans ce cycle de conception, la phase descendante consiste à définir et à analyser les besoins, les exigences, et la faisabilité du système. Ensuite vient l'étape de spécification de la solution. Puis une architecture est proposée, et une conception des composants nécessaires à la mise en oeuvre du système est réalisée, conduisant à un premier prototype fonctionnel. La phase ascendante va alors consister à mettre sous test, à évaluer, et valider les composants et le système conçu. C'est au niveau de cette phase ascendante, et plus particulièrement au niveau de la définition d'une stratégie d'évaluation et de validation d'un STRA intégrant une ou plusieurs briques IA, que s'est positionné le projet PRISSMA.

En conditions réelles, les étapes d'évaluation et de validation ne sont pas triviales à mettre en oeuvre et impliquent de traiter et d'identifier un grand nombre de scénarios reproduisant deux types de configuration. La première concerne les scénarios nominaux, et la seconde s'intéresse aux situations critiques, comme les quasi-accidents et les situations potentiellement accidentogènes. Afin d'identifier ces configurations, il est nécessaire de pouvoir identifier, utiliser, faire évoluer et gérer les variables permettant d'évaluer le bon fonctionnement d'un STRA.

Pour identifier ces configurations, et ensuite, tester les systèmes de mobilité dans des conditions choisies, il semble indispensable de disposer d'environnements et d'outils de simulation (physique et/ou logiciel) efficaces et adaptés qui permettent de réaliser des tests en amont avant de passer à des tests dans des environnements réels plus ou moins contrôlés.

Dans ce contexte, les étapes d'évaluation et de validation doivent non seulement s'appliquer aux algorithmes et applications impliqués dans les briques IA du STRA, mais aussi à ces outils de test (garantie de la répétabilité des scénarios, gestion de variantes contrôlées pour gérer les scénarios critiques dans des situations dégradées, etc).

Les figures 3.1 et 3.2 présentent la méthodologie d'évaluation et de validation générale proposée dans le cadre PRISSMA. Dans cette méthodologie, quatre étapes clés sont identifiées en faisant le lien avec les types d'environnement de test utilisés et surtout avec les phases impliquées dans le processus d'évaluation et de validation.

La **première étape** (voir 3.2) concerne la définition et la mise en oeuvre du véritable composant, système, application dans un démonstrateur fonctionnel complet. Cette première étape est une étape amont du processus d'évaluation et de validation. Plus concrètement, cette étape consiste à :

- Définir le cadre générique du service en caractérisant l'OD, l'ODD (fourni par l'OEM ("Original Equipment Manufacturer"), le fabricant, le client), ainsi que l'OEDR (on pourra se référer à la sous sections Définitions clés de l'introduction pour un rappel de ces acronymes).
- Définir les scénarios (proposé par l'équipe d'évaluation et de validation).

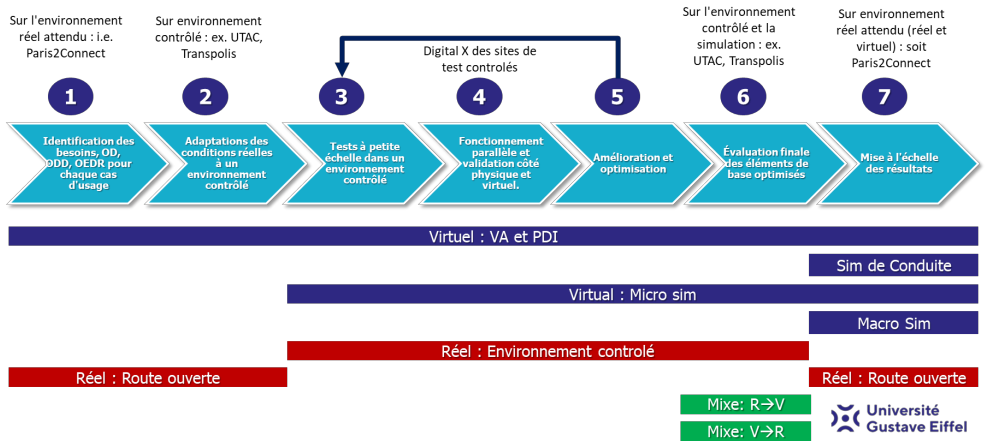

Figure 3.1 – Vue d'ensemble des différentes étapes entrant dans la méthodologie proposée d'évaluation et de validation des systèmes de mobilité automatisée. Ce schéma permet de voir les différents types de moyens d'essais applicables pour chaque étape.

- Définir le service, les systèmes et les composants pour le service de mobilité automatisée/autonome à déployer, avec ou sans systèmes à base d'IA. Cette partie doit aborder les étapes de pré-traitement des données brutes, les modules de perception, la localisation, la prise de décision, la planification de trajectoire, et la couche basse de contrôle/commande (fourni par l'OEM, le fabricant, le client).
- Définir l'architecture embarquée réelle du véhicule (matériel et logiciel) incluant l'accès au bus de communication (CAN) et l'accès aux actionneurs.
- Définir la topologie et la configuration des capteurs aussi bien intrinsèque qu'extrinsèque avec leurs types, leur nombre, leur couverture de perception, leur paramètres principaux, et la définition du processus de calibration (fournie par l'OEM, le fabricant, le client).
- Définir le processus d'enregistrement des données comprenant la définition des données et du format de la base de données. Cette étape implique également de traiter la génération des références et des vérités terrain. (proposé par l'équipe d'évaluation et de validation).
- Définir la procédure d'évaluation et de validation avec l'identification des indicateurs de risques, des KPI ("Key performance indicators"), et des métriques pour l'évaluation des composants. Il est également nécessaire de définir les critères et le seuil de validation (proposé par l'équipe d'évaluation et de validation).

La **seconde étape** va avoir pour principal objectif d'appliquer une instance du démonstrateur et du système dans un environnement réel contrôlé représentatif de l'environnement réel final. Un exemple d'architecture réelle et virtuelle pour l'évaluation d'un STRA est présenté en figure 3.3. Cette étape va permettre de contrôler l'exécution des tests et surtout des scénarios en limitant les situations inattendues. Les tests sur des sites de test vont garantir un plus haut niveau de répétabilité et de reproductibilité des essais. Cette étape permet également d'impliquer un humain dans la boucle d'évaluation en minimisant le risque. La répétabilité est généralement obtenue en utilisant un robot de conduite. Comme dans la première étape, il est nécessaire de définir le processus d'enregistrement des données avec la définition des formats de données et des références à générer. Initialement, cette définition des bases de données est réalisée dans la première étape. Il en va de même pour la définition des indicateurs de risques, des KPI, et des métriques pour l'évaluation des composants et des systèmes. Cependant, au contraire des environnements ouverts, il est possible d'étendre le nombre de KPI et de métriques utilisables en profitant du fait que l'environnement contrôlé dispose généralement de nombreux capteurs et équipements générant des observations et des références aussi bien en embarqué que provenant de l'infrastructure.

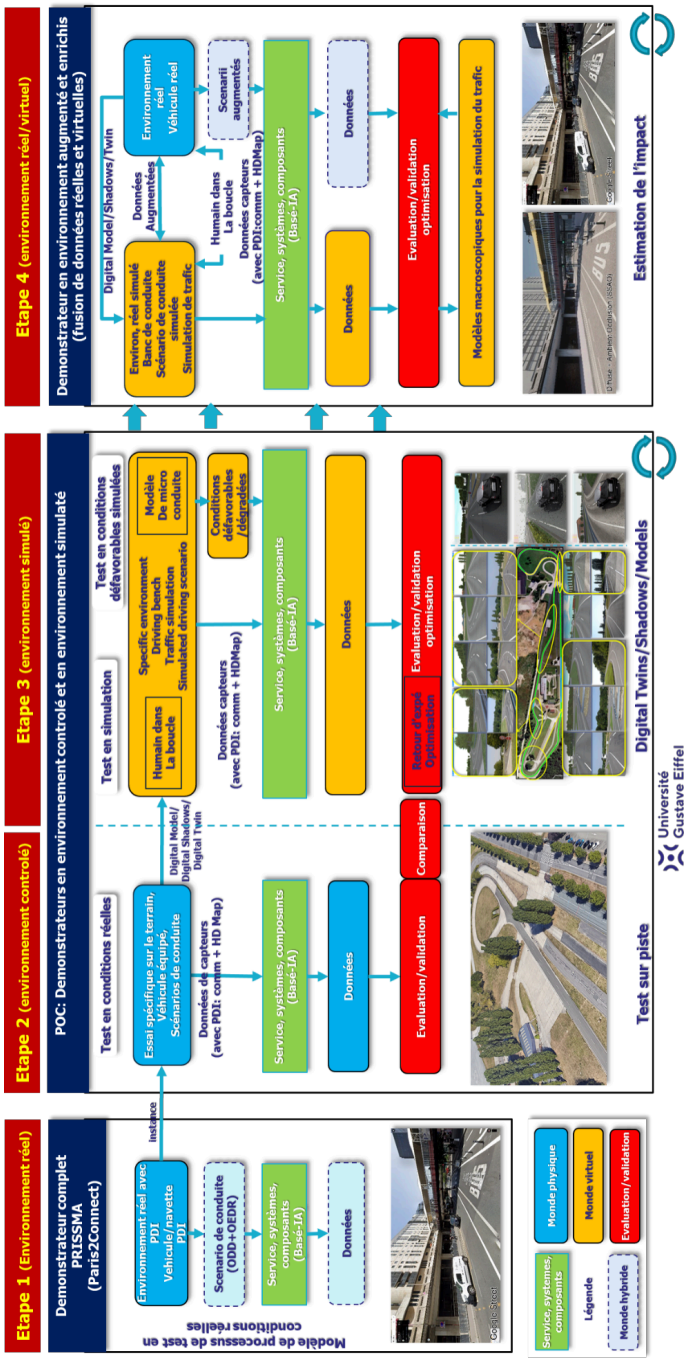

FIGURE 3.2 – Méthodologie complète d'évaluation et de validation proposée dans PRISSMA, en incluant les interactions et l'enchaînement des étapes en conditions réelles, en environnement réel contrôlé, et en simulation.

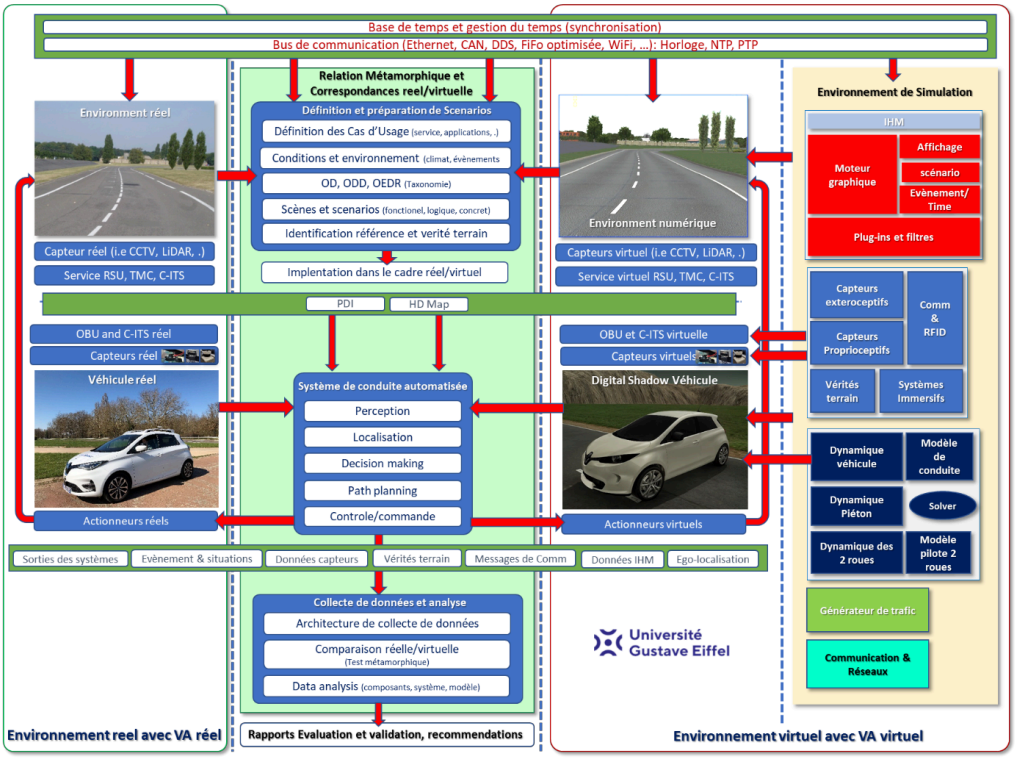

FIGURE 3.3 – Vue générique et simplifiée d'une architecture réelle et virtuelle pour l'évaluation et la validation des systèmes à base d'IA pour la mobilité automatisée.

Bien entendu, il est également nécessaire de choisir des environnements contrôlés permettant de respecter les exigences et les ODD définis dans les phases clés 1 et 2 (voir figure 3.1).

La **troisième étape** consiste à faire une instance du démonstrateur et du système dans un environnement de simulation. Pour réaliser cette opération, il est nécessaire d'utiliser un modèle numérique (ou Digital Model) de l'environnement contrôlé et des ombres numériques (ou Digital Shadows) des composants impliqués dans le fonctionnement du système (surtout la modélisation dynamique et physique des véhicules, des autres agents vulnérables, des capteurs, des moyens de communication, . . .). Dans cette troisième étape, deux types de scénarios seront mis en oeuvre : les scénarios nominaux et les scénarios "critiques". Pour cette seconde catégorie de scénario, il sera nécessaire d'être en capacité de générer de nombreux types de perturbateurs et des conditions défavorables et/ou dégradées comme des conditions météorologiques réduisant les performances des capteurs. Afin de pouvoir garantir la représentativité de ces tests en environnement simulé, il est impératif de proposer et d'utiliser des scores et des métriques quantifiant le niveau de représentativité, de fidélité, et de qualité des outils et modèles de simulation utilisés. Bien entendu, les environnements de simulation peuvent être hybrides et impliquer l'utilisation de bancs physiques et dynamiques (Bancs Human in the loop (HiTL) et Vehicle in the loop (ViL)). Comme pour l'étape 2, des opérateurs humains peuvent intervenir et être impliqués dans l'exécution des scénarios comme conducteurs ou acteurs de l'environnement. L'utilisation d'un robot de conduite est aussi envisageable. Il faut également définir le processus d'enregistrement des données comprenant les références et les vérités terrain. Il est important de souligner que les intérêts majeurs de l'utilisation de la simulation concernent sa capacité à :

- Contrôler l'exécution des scénarios,
- Générer des scénarios rares, critiques et à haut risque (couverture plus large de l'ODD),

- Garantir un haut niveau de reproductibilité et de répétabilité,
- Permettre la génération de références précises et relativement exhaustives,
- Pouvoir contrôler précisément les conditions de luminosité et météorologiques.

Dans un environnement simulé, on notera également que la quantité et la diversité des KPI et métriques utilisables est beaucoup plus grande.

La **quatrième étape** consiste à faire évoluer le système sous test dans les environnements réels de sa future mission, qu'ils soient physiques et virtuels (en utilisant notamment les Digital Model et Digital Shadow de l'environnement réel). Cette étape permet de mixer les mondes réel et virtuel à travers la génération de réalité enrichies et augmentées. Ces réalités augmentées peuvent consister à générer des données virtuelles de capteurs pour un enrichissement des données réelles des capteurs embarqués dans le vrai système de mobilité. Dans ce contexte nous avons à disposition des architectures et des plateformes d'évaluation et de validation ViL. L'interaction temps réel des composants des mondes réels et virtuels produit alors une architecture que nous pouvons apparenter à un vrai Digital Twin.

Pour finir, la méthodologie proposée recommande l'ajout à ces quatre étapes d'évaluation du STRA d'une étape d'évaluation des moyens mis en oeuvre pour assurer le maintien en conditions opérationnelles du STRA et des briques IA sur lequel il repose (identification des situations singulières, introduction de maintenances correctives, déploiement sécurisé de mises à jours,...).

3.2 Spécificités IA et adaptation des exigences

L'utilisation d'une fonctionnalité IA par un STRA implique plusieurs nouvelles exigences pour en garantir la sécurité ou la performance. A minima, l'homologation de cette fonctionnalité devra couvrir les prérogatives de la réglementation européenne UE 2024/1689 (IA Act) sur le système de gestion des risques, les données et gouvernance des données, la documentation technique, la conservation des archives, la transparence et la fourniture d'informations aux utilisateurs, la supervision humaine et l'analyse des performances de l'IA notamment sur la précision, la robustesse et la cybersécurité. Or il n'existe pas (encore) de consensus sur la façon d'appliquer ces exigences pour un STRA. C'est pourquoi le projet PRISSMA fournit une première formulation de telles exigences, afin de permettre l'application de méthodes d'évaluation en suivant une approche IVVQ. Une liste exhaustive de ces exigences est consultable sur https ://prissma.univ-gustave-eiffel.fr. Nous en résumons les composantes clé.

3.2.1 Exigences d'explicabilité et de traçabilité

Lors des tests, le système doit effectivement pouvoir fournir des éléments explicatifs lors d'une décision automatique. Ces éléments doivent être justifiés au regard de la réglementation, des spécifications contractuelles et de la criticité. De même, en phase de conception, le concepteur du système doit être en mesure de conserver une trace (journal ou autre) permettant de donner des explications sur les éléments fournis par la fonctionnalité d'IA à la suite d'une décision automatique. Toutes ces explications (en phase de conception et en phase de test) seront sauvegardées pendant une durée dépendant de la criticité de la fonctionnalité d'IA, notamment à des fins d'analyse a posteriori, en cas d'accident.

A cela s'ajoutent des exigences de confidentialité et d'éthique (respect du Règlement Général sur la Protection des Données (RGPD) par exemple).

3.2.2 Exigences de sécurité

Les spécificités de l'IA entraînent d'autres complications et d'autres exigences pour assurer un minimum de performance, de sécurité, de robustesse ou de résilience. Ces exigences sont très liées au type d'IA étudié, de sorte qu'un algorithme d'apprentissage ne sera pas évalué de la même manière qu'un système expert, par exemple, parce qu'il a ses propres particularités, telles que la base de données d'apprentissage, les notions d'indice de confiance, de testabilité ou des problèmes de convergence de l'algorithme.

Concernant les exigences liées à la sécurité, une stratégie possible de démonstration de sécurité est de vérifier que le STRA est au moins aussi sûr qu'un conducteur humain dans une situation équivalente (principe GAME - Globalement Au Moins Equivalent). Cette démonstration doit s'appuyer sur un critère objectif qui reste le même dans toutes les situations. Une situation dangereuse impliquant un STRA peut entraîner des blessés graves ou des morts. L'entité responsable doit donc prouver que le STRA atteint un niveau de risque résiduel acceptable pour l'ensemble des situations dangereuses rencontrées.

Devant la multiplicité des indicateurs susceptibles de caractériser la notion de "Safety", il paraît illusoire d'envisager la faisabilité technique ou scientifique de preuve ou de démonstration "exacte" de ces indicateurs, du moins en absolu et ce pour les deux raisons principales suivantes :

- l'univers des cas d'usage est **infini**, ce qui est lié entre autres à la capacité "inépuisable" de l'environnement à générer une diversité plus ou moins prévisible de scénarios opérationnels.
- même si certaines méthodes peuvent adresser cet "infini", les postulats auxquels il faut adosser leurs résultats (qu'il faut démontrer au cas par cas) deviennent extrêmement restrictifs par rapport à la réalité.

Face à ces deux difficultés majeures, il est intéressant d'avoir recours à la démonstration de mise en oeuvre d'activités "vertueuses" pour la "Safety", en conjuguant tous les points de vue possibles :

- Organisationnel,
- Qualité,
- Normatif,
- Méthodologique,
- Sûreté de Fonctionnement avec les outils associés.

C'est ce qu'on appelle un processus de justification par rapport à des exigences de moyen, et il est en général porté par un processus d'audit susceptible d'adresser des thématiques complémentaires :

- Comment le processus d'Ingénierie Système est-il organisé ?
- Comment l'IA est-elle spécifiquement gérée dans le référentiel d'Ingénierie Système ?
- Qui en sont les différents acteurs et quels sont leurs rôles ?
- Comment peut on jutifier la pertinence du processus d'ingénierie par rapport à l'intégration de briques logicielles à base d'IA ?
- Comment peut-on évaluer et mesurer spécifiquement l'apport de l'IA dans le système ?
- Comment le référentiel d'Ingénierie Système est-il structuré ?
- Quels sont les outils et ateliers d'Ingénierie Système ?
- Comment est effectuée la gestion de configuration du système ?

3.2.3 Exigences liées à la cybersécurité

La validation de la cybersécurité d'un STRA est à la fois critique et complexe. En France, les deux méthodologies de certification reconnues par l'ANSSI, que sont la CSPN (Certification de sécurité de premier niveau) et les Critères Communs, représentent des efforts très conséquents pour un résultat statique et limité en termes de taille de cible. Ces méthodologies, qui fournissent les plus hauts niveaux de confiance en cybersécurité, n'ont encore jamais été appliquées à des systèmes à base d'IA.

Pour pallier ce problème, la méthodologie PRISSMA propose d'étudier la cybersécurité en se plaçant au niveau du système évalué, en se concentrant sur les éléments les plus critiques après une analyse de risque initiale du système complet. Le méthodologie proposée permet notamment (i) de mener une analyse et une méthodologie d'assurance au niveau système et non composant, (ii) de fournir différents niveaux d'assurance (niveaux d'évaluation) selon les types de composants et leur sensibilité (audit de site, évaluation, auto-déclaration), (iii) d'adapter la méthodologie d'évaluation à une technologie précise, (iv) de relever le défi de l'évaluation de la sécurité des IA, en intégrant de nouvelles exigences telles que l'explicabilité et la démontrabilité, et en enrichissant les catalogues existants de recommandations de tests de sécurité et de vulnérabilité à partir de la définition d'une taxonomie des attaques existantes pour les IA.

3.2.4 Exigences sur les outils de simulation

Afin d'être utilisables pour les étapes d'évaluation et de validation, les moteurs de simulation et les moteurs graphiques doivent respecter un ensemble d'exigences et de contraintes (capacités, réalisme du niveau de rendu, réalisme du moteur physique, gestion du temps, partage des bandes de roulement, contrôle et accès à la mémoire, contrôle CPU et GPU, ...), que l'on peut regrouper en plusieurs familles. De même que pour la cybersécurité, ces exigences, que l'on résume ci-dessous, ne sont pas nécessairement guidées par des spécificités IA, et peuvent être modulées en fonction de l'état d'avancement du projet.

Exigences spécifiques à l'IA

Dans cette partie sont recensées sous la forme de recommandations, les caractéristiques des outils de simulation qu'il s'agit de contrôler en priorité car elles sont susceptibles d'impacter directement le bon fonctionnement de briques IA de perception et de prise de décision dans un STRA.

- **Rendu multispectral** : Utiliser des modèles physiques, des matériaux et des textures spécifiques pour reproduire l'interaction entre les capteurs et l'environnement dans plusieurs spectres (visible, infrarouge, etc.).
- **Fidélité de la modélisation visuelle** : Représenter avec précision les éléments de simulation avec une haute fidélité pour garantir que les capteurs d'IA perçoivent un environnement réaliste.
- **Anti-aliasing et filtrage** : Réduire les artefacts visuels pour améliorer la clarté visuelle, essentielle pour la précision des algorithmes de perception d'IA.
- **Éclairage et ombres dynamiques** : Créer des effets d'éclairage réalistes pour simuler avec précision les interactions lumineuses et les ombres, influençant les capteurs d'IA.
- **Effets de particules et effets spéciaux** : Simuler des phénomènes environnementaux réalistes comme la fumée ou la pluie pour tester la robustesse des capteurs d'IA.
- **Capacité de lancer de rayons** : Simuler des réflexions, des réfractions et des collisions pour améliorer la perception des capteurs d'IA et la simulation des émissions électromagnétiques des LiDAR, RADAR, et

GPS.

- **Niveaux de texture et réflexion** : Utiliser des textures haute résolution, des mappages de relief, et des techniques de réflexion pour un rendu précis des surfaces et des matériaux.
- **Intégration de l'interface utilisateur et multi-vues** : Offrir des interfaces et des vues multiples pour analyser les données des capteurs et configurer les simulations.

Exigences générales pour le moteur graphique

Les exigences ci-dessous sont également proposées pour s'assurer que le moteur graphique offre des performances de rendu en temps réel de haute qualité, avec une compatibilité multiplateforme, une personnalisation et une intégration avec d'autres composants de simulation, tout en répondant aux besoins spécifiques de l'IA.

- **Qualité du rendu** : Fournir un rendu de haute qualité pour garantir une visualisation réaliste.
- **Cohérence entre les plates-formes et l'API (Application Programming Interface)** : Assurer une qualité de rendu cohérente sur différentes plateformes et systèmes d'exploitation.
- **Compatibilité matérielle et prise en charge multiplateforme** : Supporter une large gamme de configurations matérielles et divers systèmes d'exploitation.
- **Rendu dynamique et optimisé** : Adapter la qualité du rendu en fonction des ressources disponibles pour maintenir des performances fluides.
- **Effets de post-traitement** : Intégrer des effets pour améliorer l'esthétique visuelle.
- **Niveau de détail** : Ajuster dynamiquement les détails des objets pour optimiser les performances de rendu.
- **Shaders et matériaux étendus** : Utiliser les shaders et les GPU pour optimiser le temps de rendu et prendre en charge des matériaux avancés pour le rendu multispectral.
- **Éclairage et ombres dynamiques** : Créer des effets d'éclairage réalistes.
- **Effets de particules et effets spéciaux** : Ajouter du dynamisme et du réalisme à l'environnement de simulation.
- **Génération de terrain** : Créer des paysages réalistes pour divers environnements de circulation.
- **Intégration physique et prise en charge de l'animation** : Simuler les interactions d'objets, les collisions et les animations dans l'environnement de simulation.
- **Architecture de plugin** : Fournir une architecture flexible pour étendre les fonctionnalités du moteur graphique.
- **Surveillance des performances** : Inclure des outils pour suivre les performances du rendu et optimiser les ressources.
- **Documentation et support** : Offrir une documentation complète et des ressources de support technique pour les utilisateurs de l'outil de simulation.

3.3 Analyse du monde réel et génération de scénarios

Comme expliqué en section 3.1, l'évaluation de la sécurité d'un STRA nécessite de spécifier l'environnement de circulation dans lequel ce système doit effectuer la tâche de conduite dynamique en sécurité, afin de le confronter aux scénarios représentatifs couvrant le domaine de manière exhaustive.

En amont, la description de l'environnement structure la conception du système au travers de la définition de l'espace des possibles qui devra être pris en compte dans les spécifications des différents éléments du système.

PRISSMA propose ainsi une méthodologie de description du domaine de conception opérationnel (ODD)

mais aussi du domaine opérationnel (OD) d'un système, comme support de la construction des scénarios, articulée sur une taxonomie de description adaptée à la présence de brique IA dans le système de mobilité considéré. Cette taxonomie propose un langage commun de description de l'environnement de circulation, que ce soit au travers des scénarios de circulation ou au travers des caractéristiques des parcours de circulation. Ce langage permet aux développeurs de délimiter les capacités opérationnelles des systèmes, de spécifier les composants et d'élaborer les scénarios utilisés dans le processus d'évaluation. Il permet ensuite aux opérateurs de décrire de manière homogène l'environnement opérationnel du système de conduite automatisé (et plus largement du STRA), rendant possibles les premières vérifications de l'adéquation de la conception du système technique sur lequel il s'appuie.

3.3.1 Taxonomie de description

La taxonomie proposée doit permettre de décrire le monde réel de manière non ambiguë et avec une granularité adaptée. Elle a été définie sur la base d'un état de l'art issu des documents académiques, normatifs, institutionnels et de groupes de travail.

La taxonomie de description est organisée selon une arborescence à trois niveaux, et structurée selon six familles de premier niveau :

- **Infrastructure physique** (configuration, état et équipement de l'infrastructure physique) : type de route, revêtement de la route, bord de la route, géométrie de la route, carrefours, structures temporaires, structures fixes environnantes, structures et caractéristiques spéciales, signalisation ;
- **Scène** (contexte de circulation au-delà de l'infrastructure physique) : zones géographiques, zones spécifiques pouvant amener à des scénarios de circulation particuliers, région/état, géo-clôture ;
- **Conditions environnementales** (conditions météorologiques, conditions d'éclairement, conditions routières induites par les conditions météorologiques, ...) ;
- **Conditions de circulation** : densité de circulation potentielle, caractéristiques des autres usagers de la route, conditions particulières ponctuelles quant à la sécurité routière ;
- **Infrastructure numérique** : infrastructure numérique et infrastructure de connectivité requises par le système pour exécuter en toute sécurité la tâche de conduite dynamique ;
- **Contraintes opérationnelles** : capacités de manoeuvre, géométrie, vitesse de circulation, ... ;

Les descripteurs de niveau 1 du parcours sont ensuite décomposés en des descripteurs de niveau 2 et 3. A chaque descripteur de niveau 3 est associée une métrique qui peut être numérique, logique ou descriptive.

3.3.2 Description de parcours

A partir de la taxonomie proposée, il est ensuite nécessaire de proposer une méthodologie de description du parcours visant à spécifier l'OD.

La description proposée pour le parcours s'appuie sur un découpage du linéaire du parcours en sections homogènes sur lesquelles les scénarios pourront se développer. Les sections sont définies comme des fractions du parcours, homogènes en termes de scénarios rencontrés.

Cette description s'organise sur 2 niveaux :

- **Au niveau du parcours** : description des caractéristiques globales à l'ensemble du parcours : zone géographique, type de conditions météorologiques, type de luminosité, infrastructure numérique présente, etc.

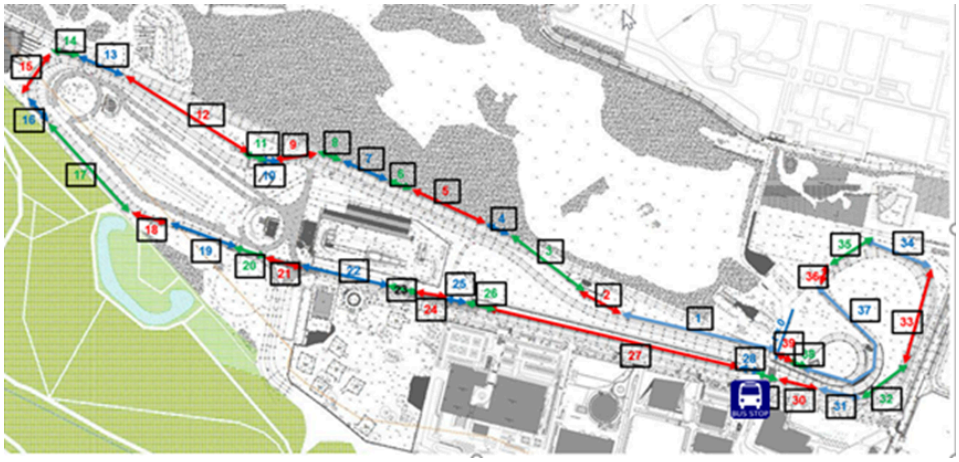

Figure 3.4 – Exemple de décomposition en sections homogènes du parcours ”Bus Station Automated Service”.

- **Au niveau des sections** : description des caractéristiques spécifiques à la section : géométrie, configuration des voies, environnement, etc.

Différents outils sont également proposés en support de la méthodologie de description :

- Support de description des sections ;
- Conventions de repérage des voies ;
- Conventions de mesures ;
- Exemples de description des intersections ;
- Exemple d'utilisation des métriques pour différents descripteurs.

La méthodologie a été éprouvée sur deux parcours réels : le parcours ”Bus Station Automated Service” de 2.4 km sur circuit (figure 3.4) et le parcours ”Paris2Connect” de 3.5 km en zone urbaine dense. Même si l'étendue des configurations rencontrées ne peut être exhaustive, ces cas d'application ont montré que la méthodologie proposée permet de structurer une analyse fine des configurations rencontrées, dans l'objectif de construire des scénarios raisonnablement prévisibles.

3.3.3 Génération des scénarios

Dans la méthodologie d'évaluation proposée, l'approche par scénarios est un moyen de répondre aux exigences d'évaluation de la solution de mobilité et de sa brique IA sur le parcours considéré. La modélisation par scénarios s'appuie sur plusieurs niveaux d'abstraction :

- Les scénarios fonctionnels qui explicitent des situations générales auxquelles le système est soumis,
- Les scénarios logiques qui spécifient les acteurs et les plages de paramètres dynamiques et environnementaux à prendre en compte,
- Les scénarios concrets qui instancient chaque combinaison de paramètre avec des valeurs précises.

La méthodologie PRISSMA permet de décrire les scénarios selon la même taxonomie que celle utilisée pour la description de l'ODD et pour la description du parcours. Cela permet d'assurer le lien entre ces différents outils d'évaluation du STRA. La génération de scénarios vise à couvrir l'ensemble du champ des possibles sur lequel va porter l'évaluation. La Figure 3.5 montre les éléments structurants pour la création des scénarios.

Un scénario logique est construit à partir :

- Des éléments statiques liés à l'infrastructure sur la section de parcours considérée,

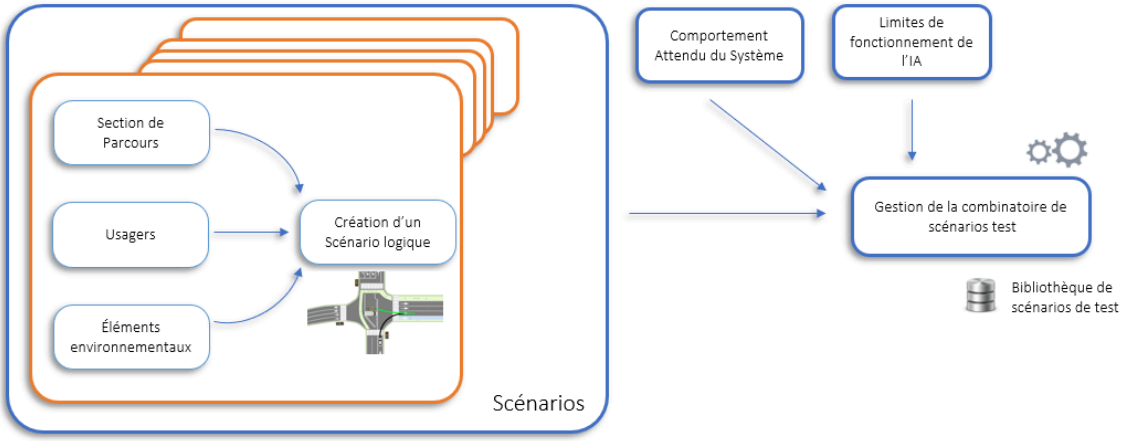

FIGURE 3.5 – Schématisation de la gestion de scénarios.

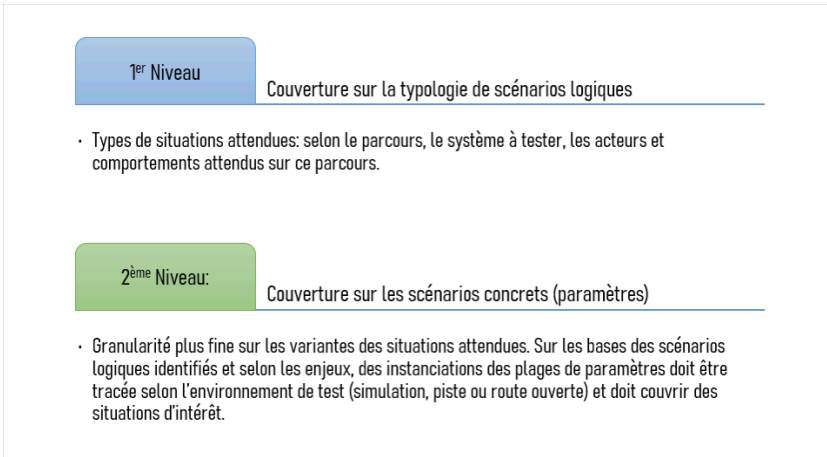

FIGURE 3.6 – Enjeux de couverture des scénarios gérés .

- Des éléments dynamiques liés aux usagers en interaction avec le STRA,
- Des conditions environnementales.

Pour l'infrastructure, les acteurs et les éléments environnementaux, des plages de paramètres sont définies pour décrire l'ensemble des situations couvertes. L'ensemble des scénarios logiques sont structurés par des scénarios fonctionnels de plus haut niveau et regroupés dans une bibliothèque afin d'en assurer la traçabilité. Cette méthode permet de gérer la combinatoire des situations rencontrées. Pour pouvoir passer à l'étape de test, il est nécessaire d'enrichir les scénarios par le comportement attendu du système dans chaque situation décrite ; on parle alors de "cas d'usage". Enfin, les contraintes liées aux limites de fonctionnement de l'IA sont prises en compte pour enrichir la base de scénarios et faire porter les tests sur les situations les plus critiques ; cette approche s'inspire de l'analyse des Triggering Conditions du SOTIF (ISO 21448).

La gestion de la combinatoire constitue un verrou vis-à-vis du besoin de preuves devant être apportées lors de l'évaluation. La bibliothèque de scénarios permet d'appréhender le sujet de la couverture de cette combinatoire pour les deux niveaux schématisés dans la Figure 3.6.

A titre illustratif, la Figure 3.7 montre une situation de conduite issue d'une des démonstrations de faisabilité du projet PRISSMA qui est modélisée sous forme de scénario logique sur la Figure 3.8. Il s'agit d'une

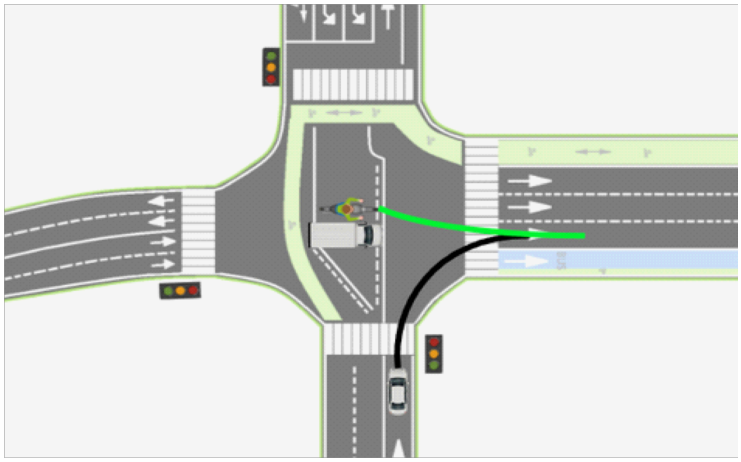

Figure 3.7 – Vue globale d'une situation de conduite pouvant être modélisée sous forme de scénario .

intersection sur laquelle le véhicule ego (VA du système) arrive par le segment sud et se dirige vers le segment est. Un camion est sur l'intersection et respecte le cédez-le passage. Un vélo est caché par le camion. Le vélo ne respectant ici pas le code de la route, ne s'arrête pas, continue sans céder le passage alors que l'ego véhicule a le feu vert. L'objet de ce scénario est d'évaluer le comportement de l'ego véhicule face à cette situation à risque. Le scénario logique dans la Figure 3.8 montre les paramètres dynamiques considérés dans le scénario pour un des acteurs, ici l'ego véhicule.

La méthodologie PRISSMA prévoit une utilisation de la modélisation par scénarios cohérente avec les référentiels établis sur le sujet :

- le code des transports (article R. 3152-2. – II. 1. et III. 1) au niveau national,
- le règlement UE 2022/1426 au niveau européen,
- les normes ISO (comme l'ISO 34503 :2023), le règlement UN R157 sur l'ALKS ou les documents des groupes de travail de l'ONUCE comme le GRVA au niveau international.

Cette modélisation permet la structuration des données clés pour la conception, le développement, l'évaluation et la validation des systèmes de mobilité incluant des briques IA. Dans le cadre de la méthodologie PRISSMA, qui s'intéresse plus particulièrement à l'évaluation et la validation, une gestion adéquate des scénarios permet de vérifier les données clés issues de la phase de conception du système et permet la prescription des tests à effectuer. La gestion des scénarios via un catalogue est cohérente avec le document ONU sur les NATM (New Assessment Test Methods) qui s'appuie sur un catalogue de scénarios pour recenser les situations de conduite pertinentes à la validation de STRA.

La méthodologie PRISSMA propose également d'utiliser une bibliothèque de scénarios pour réaliser la traçabilité entre les spécifications initiales et les résultats d'essai (notamment pour la documentation et d'audit) ; cette gestion permet le recueil et l'analyse statistique des résultats obtenus au niveau des scénarios logiques.

La méthodologie par scénarios est applicable quelle que soit l'allocation sur les modalités d'essai : en environnement contrôlé (tests sur pistes), en simulation et sur route ouverte. Cette phase d'allocation dépendra notamment du scénario, des potentielles conséquences sur la sécurité et de la facilité de mise en oeuvre de moyens de test. La bibliothèque de scénario permet de faire le lien entre les différents essais et de s'assurer

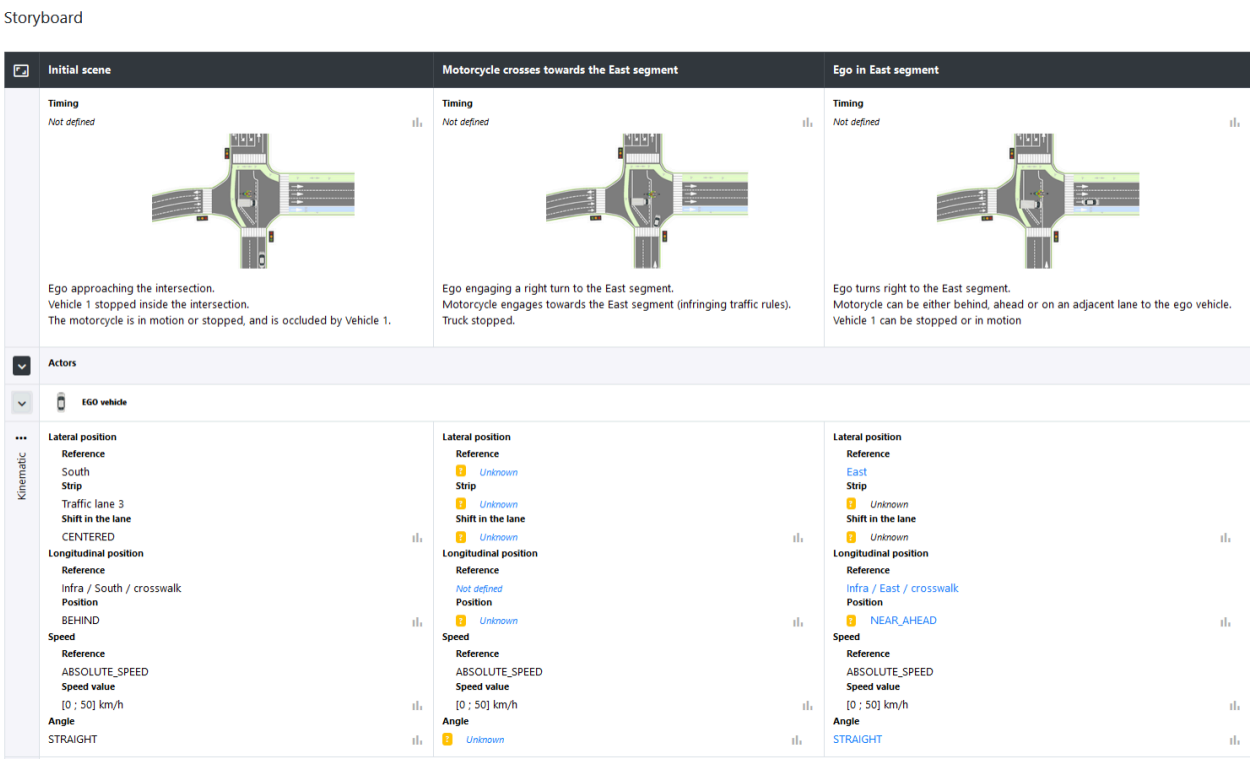

Figure 3.8 – Scénario logique : paramètres dynamiques pour un acteur dans une suite de trois scènes .

qu'au global, tous les scénarios sont couverts.

3.4 Tests en environnement contrôlé

Les essais sur piste et sur bancs d'essais sont cruciaux pour évaluer l'IA embarquée d'un véhicule autonome/automatisé, car ils permettent de tester les systèmes dans des environnements contrôlés avant leur déploiement, et de fournir des données essentielles pour les essais en simulation sous forme de vérités terrain. Les essais sur piste offrent un cadre de simulation réaliste pour les tests VIL (Vehicle In the Loop) notamment en termes de scénarios de conduite, permettant de vérifier les capacités de l'IA à réagir à diverses situations, comme les changements de trafic, les piétons imprévus, et les conditions météorologiques variées. Les travaux normatifs et réglementaires achevés ou en cours abordent la problématique des véhicules avec IA par des essais, des audits de conception et de la démonstration de sécurité. Mais ce n'est sans doute pas suffisant car les véhicules à base d'IA, notamment sur des fonctions de sécurité, pourraient avoir de nouvelles faiblesses et de nouveaux risques sécuritaires comparés à des véhicules n'embarquant pas d'IA, notamment par rapport aux questions liées à la répétabilité des tests, la robustesse aux limites, l'évaluation de l'anticipation de l'IA, et le sur-apprentissage des essais officiels d'homologation.

Afin de combler ce manque, trois approches ont été proposées au sein du projet PRISSMA.

- La première consiste à s'inspirer des essais d'homologation existants pour proposer de nouveaux tests dédiés à l'évaluation de l'IA. Par exemple, des essais de répétabilité et d'anticipation (sur des traversées de piéton par exemple, voir figure 3.9), de robustesse (par rapport aux salissures caméra ou aux aspects piétons par exemple, voir figure 3.10), ainsi que des essais pseudo-aléatoires (voir figure 3.11) ont ainsi

Figure 3.9 – Exemples de tests de répétabilité et d'anticipation pour la traversée de piétons.

Figure 3.10 – Exemples de tests de robustesse aux salissures caméra ou aux aspects piétons.

été proposés.

- La seconde est de proposer de nouvelles approches ou équipements de tests sur piste, notamment pour analyser l'interopérabilité des équipements (figure 3.12) ainsi que la résilience aux cyber-attaques (figure 3.13).
- La troisième est de renforcer le couplage entre essais sur banc et/ou piste et la simulation (cf figure 3.14), et le recours à la réalité augmentée (figure 3.15).

3.5 Tests en environnement simulé

3.5.1 Intérêts des tests en environnement simulé

Préalablement à la réalisation des tests en environnement simulé, il s'agit de s'assurer que les chaînes d'outils de simulation sont bien génériques, automatiques, interopérables et évolutives, et sont bien en mesure de mettre en oeuvre les procédures et protocoles recommandés. Cette étape est primordiale pour valider (de manière formelle ou semi-formelle) le fait que la simulation puisse être utilisée comme preuve acceptable dans le cadre d'un processus réglementaire de conception, d'évaluation et de pré-certification. Cette étape doit impérativement prendre en compte les retours d'expériences incluant le fonctionnement parallèle du système physique réel pour adapter les scénarios, les paramètres testés, et les types d'événements à gérer et à traiter. Afin de se rapprocher au plus près des conditions réelles, il est nécessaire d'utiliser des fonctions de simulation hybrides avec l'injection de données réelles.

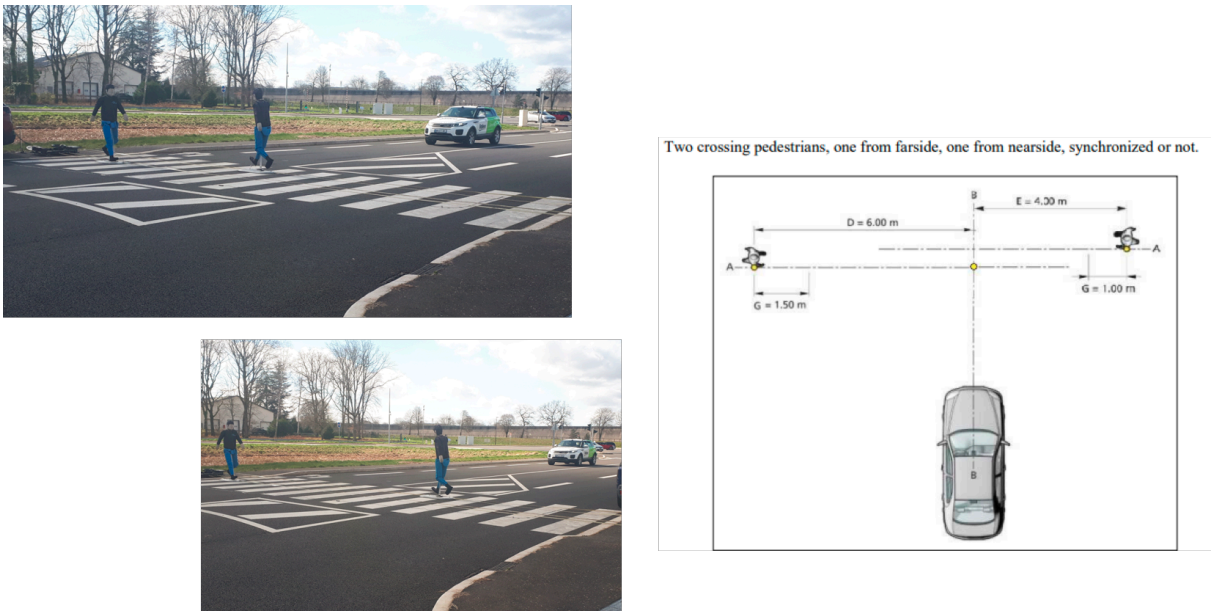

FIGURE 3.11 – Exemples de tests pseudo-aléatoires de traversée de piétons.

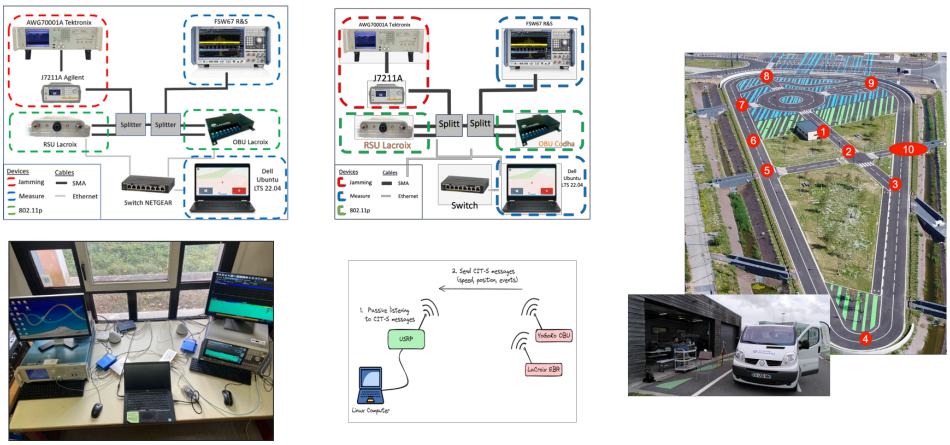

FIGURE 3.12 – Exemple de test d'interopérabilité de communication entre l'infrastructure et le véhicule.

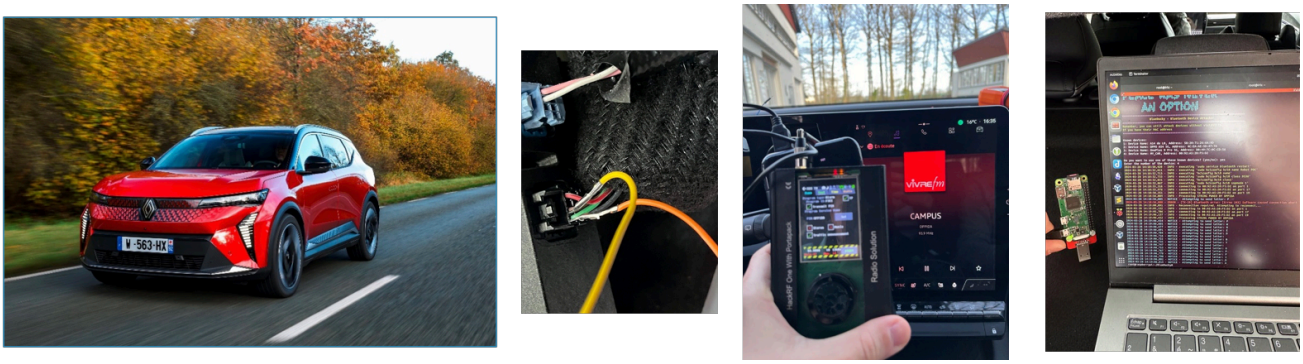

FIGURE 3.13 – Exemple de test sur piste d'essai de résilience du véhicule à une cyber-attaque.

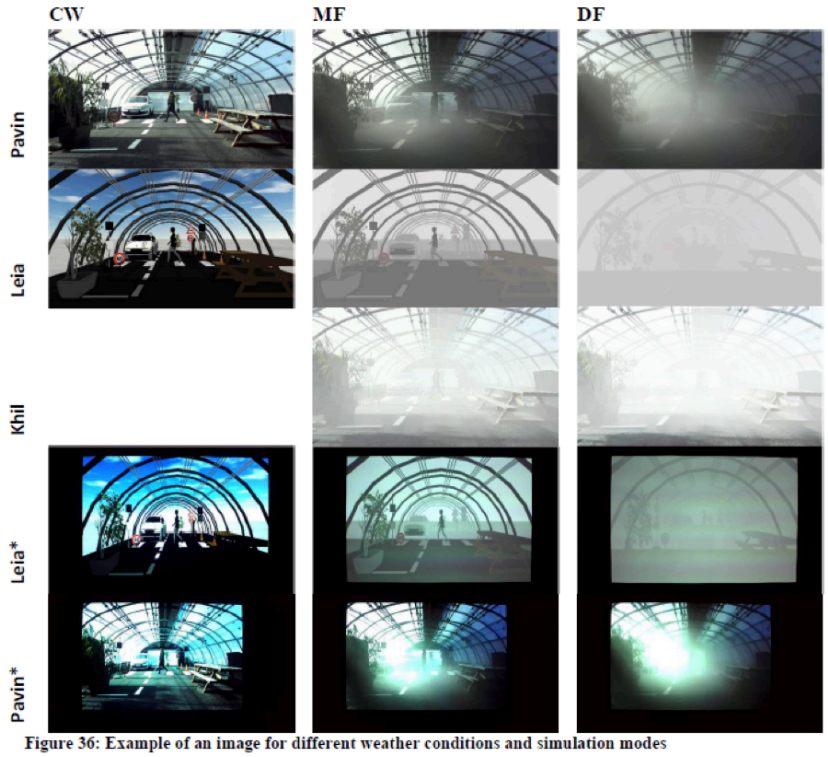

Figure 36: Example of an image for different weather conditions and simulation modes

FIGURE 3.14 – Exemple de test en banc d'essai pour la simulation de conditions météorologiques dégradées dans la plateforme PAVIN.

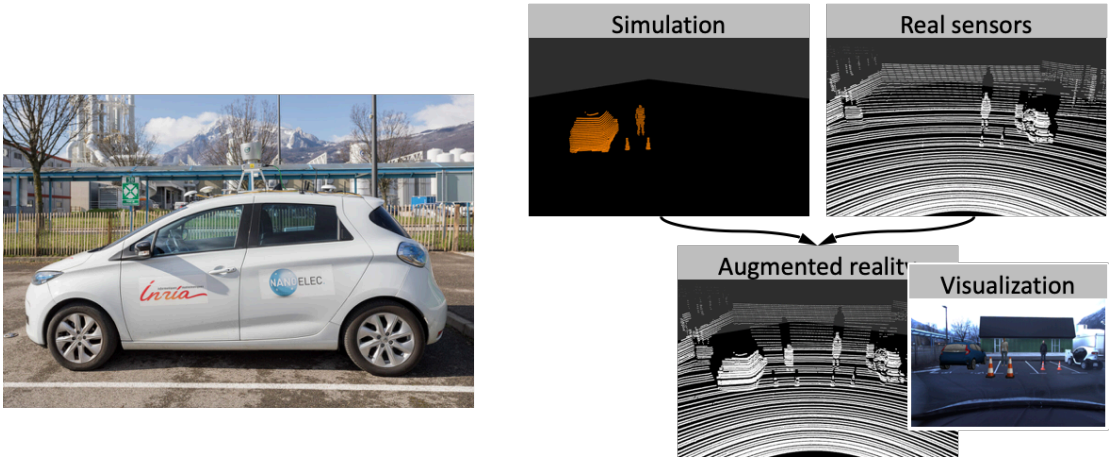

FIGURE 3.15 – Exemple de test réalisé en réalité augmentée, où des obstacles virtuels sont ajoutés au monde réel dans lequel évolue le véhicule.

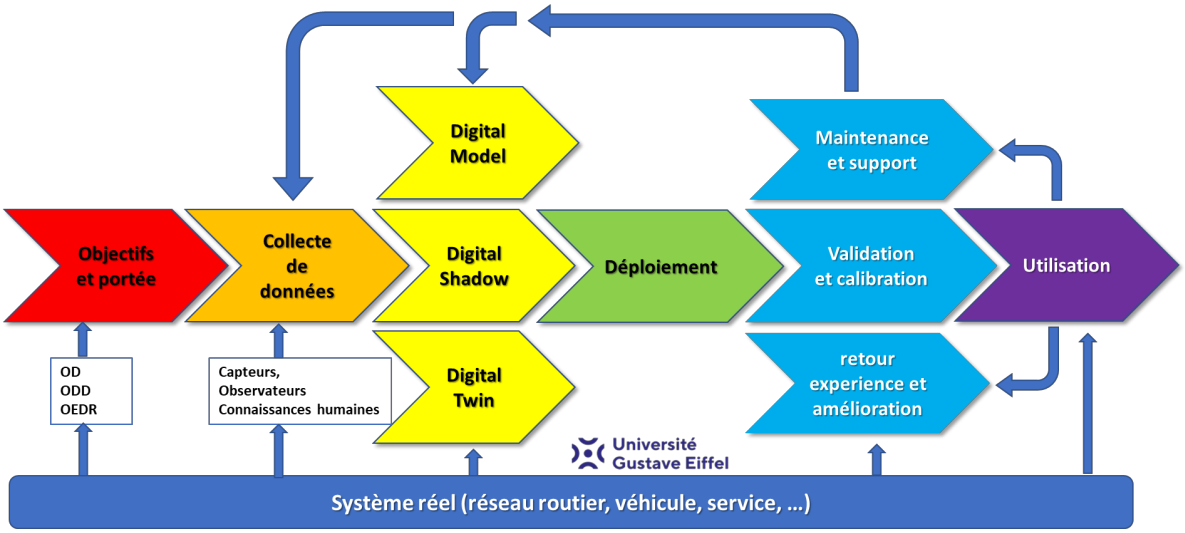

FIGURE 3.16 – Méthodologie générique de création et de développement des Digital-X .

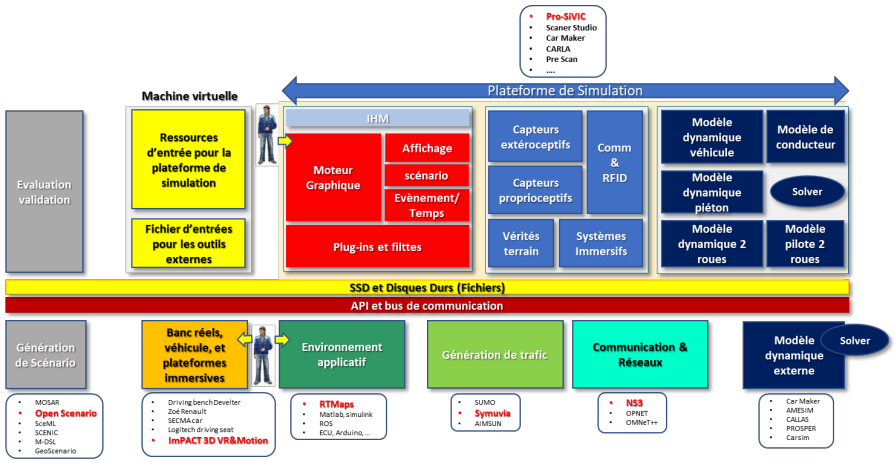

FIGURE 3.17 – Taxonomie et cadre générique pour le développement de plateforme et d'environnement de simulation.

Toutes les étapes et notions mentionnées ci-dessus se réfèrent, sans les avoir encore mentionnées, à l'utilisation des concepts de Jumeau Numérique (Digital Twin - DT), d'Ombres Numériques (Digital Shadows - DS) et de Modèles Numériques (Digital Model - DM). Ces différents concepts, avec leurs différences et leurs similitudes, ainsi que les possibilités et les limitations qu'ils apportent pour de l'évaluation de STRA ont été précisément étudiés dans le cadre du projet PRISSMA. La figure 3.16 synthétise les différentes phases de création et de développement de ces DT, DS et DM.

Concernant la méthodologie de test en simulation, un cadre générique a été proposé et une architecture fonctionnelle servant de taxonomie a été construite (cf figure 3.17). Cette taxonomie permet de générer un grand nombre de plateformes et d'environnements de simulation adaptés à divers types de besoin (Simulateur de conduite, simulation de mobilités automatisées/autonomes, simulation de mobilités connectées, architecture HiTL et ViL, ...).

Les tests d'évaluation en environnement simulé sont décomposés en deux processus.

- Le premier, présenté dans la figure 3.18, est dédié à la définition, la génération, et l'exécution des scénarios

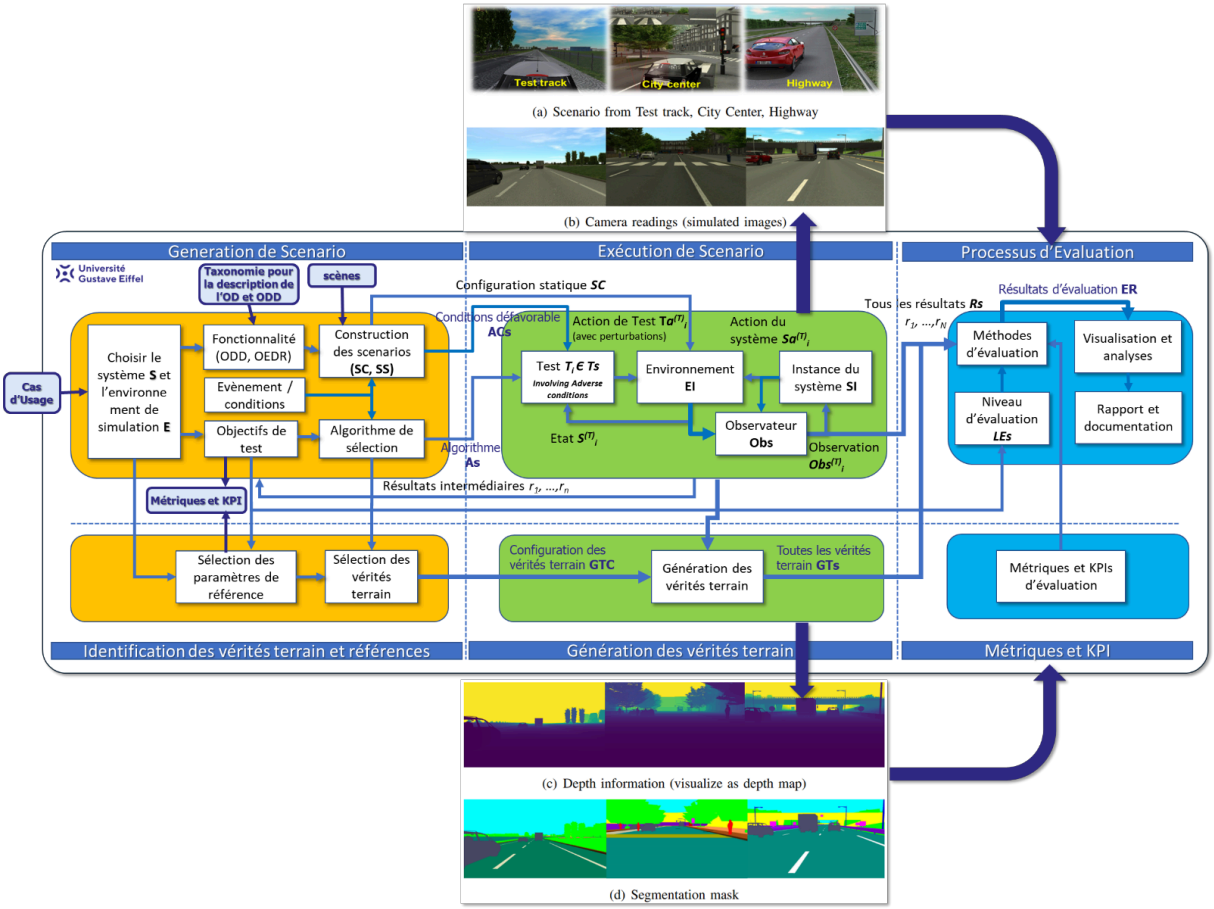

Figure 3.18 – Cadre générique synthétisant les étapes nécessaires dans la mise en oeuvre d'une procédure d'évaluation et de validation des systèmes à base d'IA pour la mobilité automatisée.

et à l'analyse des résultats pour estimer la performance du système à base d'IA mis sous test.

- Le second, présenté dans la figure 3.19, concerne les étapes à mettre en oeuvre pour la génération des Datasets et des bases de données améliorées.

Ces tests impliquent la génération de liens et d'interactions entre les mondes réel et simulé. Ces échanges sont utilisés pour vérifier la validité des méthodes/fonctions/algorithmes/applications en examinant si les relations attendues entre les entrées et les sorties sont maintenues après des transformations spécifiques nécessaires pour obtenir un lien entre les mondes réel et virtuel. Cela permet de tester des propriétés importantes des programmes sans avoir besoin de connaître les résultats exacts à l'avance.

Une fois validée, la simulation permet enfin l'exploration de l'espace des possibles de manière quasi-exhaustive, et l'identification d'un maximum de scénarios critiques représentatifs de situations accidentogènes ou de presqu'accident.

3.5.2 Premiers retours d'expériences

Cinq expérimentations ont été réalisées en milieu simulé dans le cadre du projet PRISSMA, faisant intervenir cinq sites d'essai (les sites d'UTAC, de Transpolis, de Paris2Connect, de Satory et la plateforme PAVIN) sous leur forme physique et numérique (voir figure 3.20), trois types de véhicules (Navette, Zoé robotisée, et véhicule VALEO Drive4U), et de nombreuses suites logicielles (SCaNER studio, Suite ANSYS, ProSIVIC (ver-

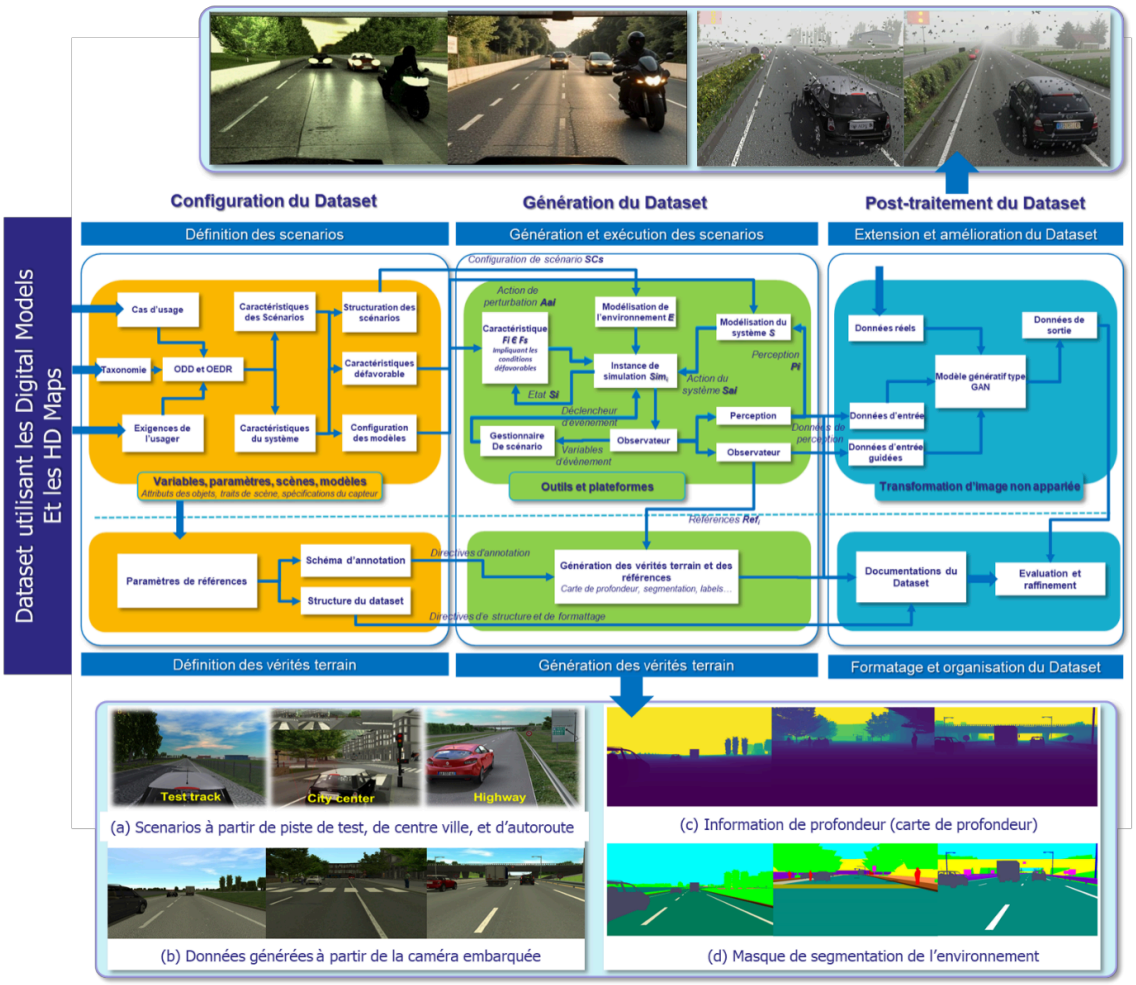

FIGURE 3.19 – Cadre générique synthétisant les étapes nécessaires à la production des Datasets et des bases de données intégrant de la réalité augmentée et de la réalité améliorée.

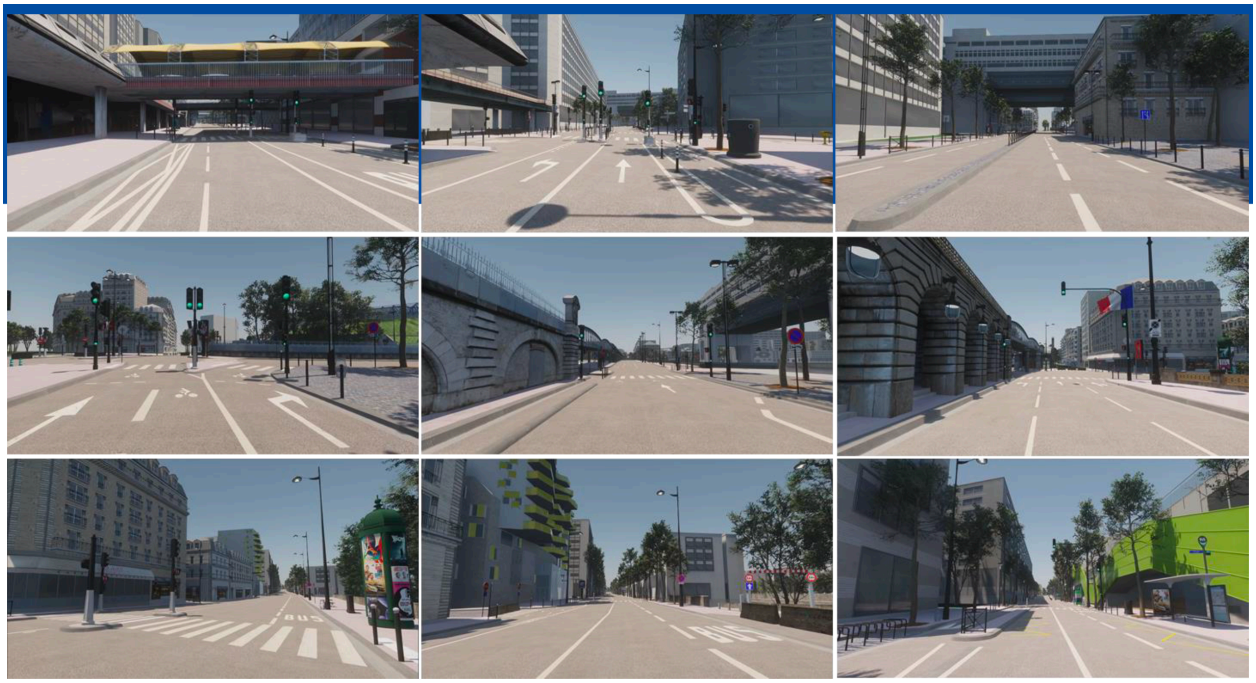

Figure 3.20 – Comparaison de vues physiques et numériques du site Paris2Connect.

sion UGE), CARLA, MOSAR, Symuvia, U-TEST, RTMaps, NS3, LEIA). Ces expérimentations ont permis de tester le caractère générique de la méthode de vérification et de validation proposée dans le cadre de PRISSMA, en testant le comportement du véhicule dans une grande variété de conditions, incluant des situations dégradées et perturbées.

Plusieurs enseignements peuvent être tirés de ces expérimentations. Pour commencer, l'investissement dans des environnements de simulation haute-fidélité en temps réel se révèle essentiel mais complexe. La reproduction fidèle des conditions du monde réel, incluant des situations routières variées et des comportements inattendus d'usagers de la route, pose un défi majeur pour assurer l'exactitude des évaluations et détecter toute défaillance potentielle avant le déploiement des véhicules.

La qualité des données simulées représente un autre défi crucial. Il est impératif de valider rigoureusement que les résultats obtenus en simulation correspondent fidèlement aux données réelles et aux scénarios du monde réel. Cette validation nécessite l'application de procédures formelles ou quasi-formelles afin d'assurer que les simulations répondent aux normes de qualité et de fidélité exigées, un processus souvent complexe nécessitant une collaboration étroite entre partenaires industriels, académiques et réglementaires.

La nécessité de développer des plateformes de simulation open-source et interopérables soulève également des défis. Cette initiative vise à stimuler l'innovation et la collaboration au sein d'une communauté élargie de développeurs et d'utilisateurs finaux. Cependant, garantir la cohérence et la compatibilité entre les différentes plateformes, tout en assurant la sécurité et la performance des STRA, représente une tâche complexe nécessitant des outils de validation et de vérification robustes.

En outre, la génération et l'analyse de grandes quantités de données restent une difficulté majeure. Les simulations doivent pouvoir traiter efficacement des données de capteurs complexes, des trajectoires de véhicules et des indicateurs de performance du système pour tester la robustesse des algorithmes impliqués et s'assurer des performances globales des systèmes sous test.

Enfin, pour répondre aux exigences de conformité réglementaire et d'assurance de sécurité, les platesformes

de simulation doivent fournir des outils adéquats pour documenter les résultats de simulation, mener des évaluations de sécurité et démontrer la conformité aux normes industrielles en constante évolution, ce qui n'est pas toujours facile à réaliser. L'adaptation continue des environnements de simulation pour intégrer de nouvelles technologies et méthodologies reste un défi constant pour maintenir la pertinence et l'efficacité des tests pour les STRA.

3.6 Tests en environnement réel

3.6.1 Intérêt des tests en environnement réel

En conditions réelles, lorsque l'on fait rouler un véhicule automatisé/autonome, la première des préoccupations et la première des priorités sont d'assurer la sécurité. C'est pourquoi, dans cette phase expérimentale, pour un niveau de sécurité maximum, une personne appelée "safety driver" doit être présente afin de reprendre en main le véhicule en cas de risque de collision ou de danger.

L'objectif des tests en environnement réel est de faire évoluer un véhicule automatisé/autonome à base d'IA dans l'environnement réel de sa mission, et de l'évaluer en le confrontant à des situations identifiées et standardisées. Ces situations sont notamment issues de l'exploration par la simulation du générateur de scénarios basé sur la description du parcours. Lorsqu'un scénario répertorié est identifié lors des tests en environnement réel, la situation peut être évaluée selon les critères de réussite définis par les indicateurs du règlement UE ADS 2022/1426, en comparant le comportement attendu et le comportement effectivement observé.

On remarquera que les tests sont réalisés en mode "boîte-noire", sans aucune connaissance préalable des capacités intrinsèques du véhicule sous test.

3.6.2 Premiers retours d'expérience

Dans le cadre du projet PRISSMA, une expérimentation (voir figure 3.21) a pu être menée avec un prototype de niveau 4 d'automatisation (un Range Rover Evoque modifié par Valeo) spécifiquement conçu pour opérer dans des conditions de conduite urbaine dense. Le parcours sélectionné était le parcours inter-gares de Paris2Connect, un itinéraire de 2.5 kilomètres reliant les gares d'Austerlitz, Lyon et Bercy dans les 12e et 13e arrondissements de Paris. Ce parcours urbain inclut divers défis tels que la gestion du trafic dense, les intersections complexes et les piétons. L'infrastructure du parcours comprend des équipements connectés qui permettent de suivre précisément le passage du véhicule automatisé/autonome sur certaines sections spécifiques du trajet.

Des technologies telles que le LiDAR et les caméras thermiques ont été utilisées pour analyser les situations et événements rencontrés par le véhicule pendant les tests. Des caméras installées à bord du véhicule ont également permis d'obtenir des vidéos internes et externes détaillées du comportement du véhicule, enrichissant ainsi l'analyse des données recueillies par l'infrastructure connectée.

Les tests ont été structurés autour de scénarios prédéfinis, visant à vérifier la capacité du véhicule à naviguer de manière autonome tout en respectant les règles de sécurité et les exigences réglementaires. Les principaux défis rencontrés lors de l'expérimentation comprenaient la gestion du temps et des ressources limitées, ainsi que la quantité importante de données générées. Le traitement (pour le moment manuel) des vidéos pour identifier et classer les scénarios pertinents s'est notamment avéré particulièrement chronophage, soulignant la nécessité de développer des méthodes adaptées pour accélérer ce processus à l'avenir.

Bien que l'expérimentation n'ait atteint que partiellement la phase de marche à blanc, elle a permis de tester progressivement la méthodologie d'évaluation utilisée, et permis de mettre en évidence la forte complexité de

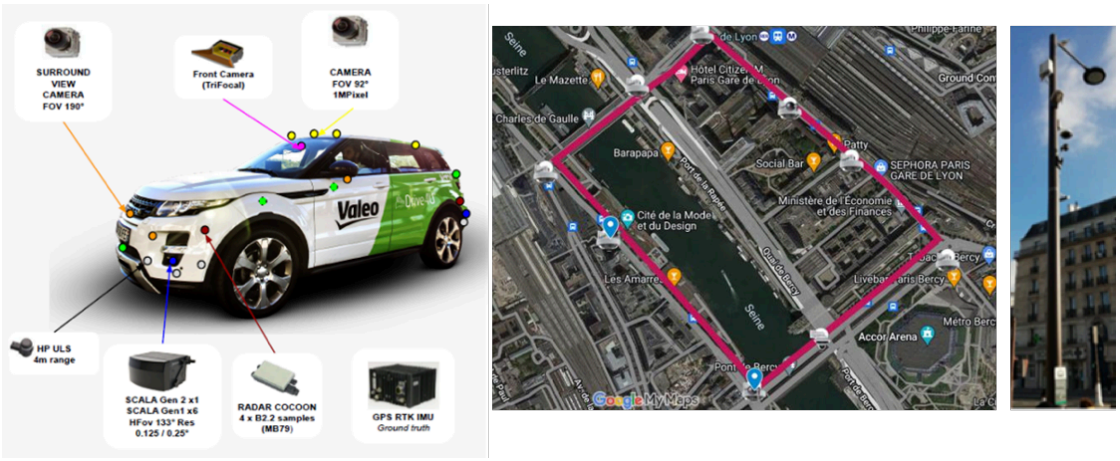

FIGURE 3.21 – Expériences de roulage dans l'environnement de Paris2Connect.

l'évaluation des STRA intégrant des technologies d'IA.

3.7 Non régression et amélioration continue

Gérer le cycle de vie d'un STRA à base d'IA représente un défi considérable, si l'on considère les propriétés spécifiques des technologies d'IA.

Pour les STRA, il s'agit ainsi de couvrir la phase d'exploitation du système, d'organiser les retours d'expérience pertinents ainsi que les actions correctives associées, notamment lorsqu'une situation opérationnelle inacceptable a été constatée pendant le cycle opérationnel du système. Et à chaque apparition de dysfonctionnement ou d'accident, il faut pouvoir s'assurer que le processus de correction suive un processus d'amélioration continue. On remarquera que pour les STRA, le décret STRA 2021-873 précise la nature des événements que l'exploitant doit faire remonter aux autorités.

Lorsque qu'une situation inacceptable est rencontrée, il est tout d'abord nécessaire d'identifier la cause simple ou les causes multiples du (ou des) comportement(s) inacceptable(s), et ensuite mettre en oeuvre les corrections pertinentes. Différents types de corrections peuvent être envisagées, en fonction de la nature des causes diagnostiquées :

- Si l'une des causes est un mode de défaillance d'un composant physique ou d'un module, une action de maintenance corrective pertinente peut être mise en oeuvre, en accord et conformité avec le plan de maintenance du système : ce mode de défaillance renverra à un élément échangeable sur site qui pourra être échangé sur site, ou à un autre niveau de maintenance, en fonction du concept de maintenance.
- Si l'une des causes possibles est une erreur logicielle n'impliquant pas un modèle d'IA, une analyse de cause logicielle doit être appliquée ; ce pourra être une erreur de spécification ou une erreur de codage, et dans les deux cas, une actualisation du logiciel doit être faite, et une recherche doit être menée dans le processus de développement pour savoir pourquoi cette erreur logicielle est passée inaperçue. . .
- Si l'une des causes possibles est une erreur de logiciel basée sur l'IA, une analyse de causes doit également être appliquée au logiciel ; à l'issue de cette analyse, une ou plusieurs correction(s) doi(ven)t être proposée(s), ainsi qu'une analyse d'impact de cette (ces) correction(s) ; en parallèle, il est souhaitable d'effectuer un diagnostic sur ce qui a provoqué ces erreurs du fait de lacunes possibles du référentiel de développement.

Cette tâche est une tâche particulièrement délicate, du fait des propriétés intrinsèques des technologies d'IA et du domaine scientifique associé. Ces corrections doivent être qualifiées dans le monde réel, car le niveau de fidélité des modèles supportant les simulations reste partiel pour l'instant : la réplicabilité et la répétabilité des situations inacceptables auxquelles les composants IA ont contribué seront décisives par rapport au fait de qualifier ces comportements inacceptables ainsi que le processus de diagnostic associé. Autrement dit, il importera de qualifier le (les) comportement (s) inacceptable(s) constaté(s) ainsi que les conditions et les facteurs susceptibles de favoriser leur(s) émergence(s).

Lorsque la contribution d'une ou plusieurs brique(s) d'IA à un dysfonctionnement ou un accident est avéré, et afin d'identifier les corrections pertinentes de ces logiciels à base d'IA susceptibles de rétablir un comportement acceptable et pertinent du système global par rapport au cas d'usage référencé à l'origine, il est nécessaire de conduire une analyse approfondie consistant à isoler les parties du logiciel à corriger, ainsi que les éventuels autres éléments du système à échanger, actualiser ou supprimer. Par exemple, si un module logiciel à base d'apprentissage profond est concerné, il faut déterminer quelle est la couche du réseau de neurones à corriger, et quelles sont les valeurs des poids à modifier et à réajuster, pour obtenir la correction du comportement global du système au sein du cas d'usage adressé.

Contrairement aux logiciels classiques, il n'y a pas beaucoup de méthodes et d'outils permettant de diagnostiquer les briques IA. Les modèles d'apprentissage de briques IA d'un STRA doivent être soumis à un processus de diagnostic lorsque des défauts sont observées au cours du cycle d'exploitation. Ces modèles d'apprentissage doivent suivre un processus de test et d'homologation, afin d'éviter des accidents potentiels. Ce processus est très exigeant et peut atteindre 6 mois à 1 an d'itérations pour chaque niveau d'actualisation.

Cependant, il reste possible que les systèmes d'IA soient sujets à défaut dans des situations sous-représentées au sein des bases de scénarios d'apprentissage, ou pas du tout prises en compte dans les jeux de test, ou encore liées à des aspects erronés inhérents aux modèles d'apprentissage. C'est pourquoi, un enjeu majeur des STRA consiste en la mise en oeuvre d'un processus de maintenance et de correction des briques d'IA pour passer d'une version à une autre, afin de corriger la logique de comportement du module d'IA dans la situiation rencontrée, ou encore de rajouter une fonction manquante sur ce module, sans avoir à re-valider le système global depuis le tout début. Le diagnostic et la maintenabilité des modèles d'apprentissage sont des défis importants pour le déploiement sécurisé des STRA. La maintenabilité d'un STRA doit par exemple permettre de corriger des défauts du modèle d'apprentissage, sans modifier en quoi que ce soit le comportement de conduite sur les milliers de kilomètres où ce même système donnait satisfaction.

Dans ce contexte de correction de briques d'IA, il est ainsi nécessaire de traiter la question de non régression. En effet, les actions correctives apportées à un logiciel défaillant doivent supprimer les erreurs constatées sur les situations problématiques concernées, sans produire de dysfonctionnements pour les situations qui ne posaient pas problème auparavant. C'est un sujet insidieux qui n'est pas encore couvert par l'état de l'art, mais à propos duquel des solutions partielles sont en cours de mise au point.

Par ailleurs, outre la(les) brique(s) IA responsable(s) d'un dysfonctionnement ou d'un accident observé, il faut éventuellement pouvoir remonter à l'étape du processus d'Ingénierie Système qui a rendu possible l'activation de ce défaut qui est en fait passé à travers les mailles du filet des nombreuses campagnes de tests virtuels, contrôlés ou réels.

Plusieurs composantes du référentiel d'Ingénierie Système peuvent être adressées pour expliquer pourquoi le cas d'usage qui a "dysfonctionné" n'a pas pu être testé ou anticipé pendant le processus d'ingénierie système, ce qui implique qu'un des aspects de ce processus n'a pas été correctement couvert ; ces questionnements peuvent être traduits par les interrogations suivantes :

- Les exigences ont-elles été bien formulées (peut-être les cas d'usage défaillants n'ont pas été suffisamment couverts par ces exigences) ?
- Les ODD ont-ils été correctement définis (peut-être les cas d'usages rencontrés allaient "au-delà" des ODD définis) ?
- Les OEDR ont-ils été correctement définis (peut-être les cas d'usages rencontrés allaient "au-delà" des OEDR définis) ?
- Quelles bibliothèques de cas d'usage ont été simulées dans les campagnes de test virtuel, au sein des campagnes de roulages contrôlés ou à travers les campagnes de roulages réels ? Les cas d'usages défaillants ont-ils manqué à ces bibliothèques ?
- Quelles métriques ont été utilisées dans ces campagnes virtuelles, contrôlées ou réelles ? Peut-être le cas d'usage défaillant n'a-t-il pas été évalué avec une métrique pertinente ?
- Quels critères ont-ils été utilisés pour caractériser les presqu'accidents ? Peut-être ces critères ont-ils été trop optimistes ?
- Quelle plateforme de simulation a été utilisée pour les campagnes virtuelles ? A-t-elle conduit à des traitements produisant des résultats trop optimistes ?

Chapitre 4 Conclusions et perspectives

4.1 Synthèse des principaux points

La validation des STRA intégrant des briques d'IA exige des compléments ou des adaptations des processus traditionnels de validation. La présence de ces briques IA introduit de nouveaux critères, notamment en matière de précision, de contrôle du sur-apprentissage, de répétabilité, de robustesse et de résilience, en particulier face aux cyber-attaques. Ces nouveaux critères exigent une évaluation rigoureuse pour garantir la fiabilité et la sécurité des systèmes automatisés.

En particulier, le besoin s'est fait jour dans le projet PRISSMA de définir de nouvelles exigences à différentes échelles : au niveau des capteurs, des sous-systèmes, des systèmes complets et des outils de simulation. Chaque niveau doit être scruté avec des exigences spécifiques pour assurer une performance optimale et sécurisée de l'ensemble.

Le projet a confirmé l'importance des exigences de complétude et de complémentarité entre les audits et les tests, menés dans des environnements simulés, contrôlés et réels. Cette approche globale garantit que toutes les facettes du système sont testées et vérifiées dans des conditions variées.

Un aspect crucial de cette validation est la gestion des données, incluant la certification des vérités terrain, qui assure l'authenticité et la fiabilité des données utilisées pour entraîner et tester les systèmes à base d'IA. L'enregistrement automatique des événements durant le fonctionnement du système est essentiel pour détecter et remonter les situations non prévues, facilitant ainsi une amélioration continue des performances du système.

Enfin, le projet a mis en exergue l'apport d'une approche probabiliste de la validation, intégrant des notions de moyenne et de forte probabilité ainsi que l'intégration des incertitudes. Cette formulation probabiliste permet de mieux appréhender et quantifier les risques et les performances du système dans des conditions réelles, en tenant compte des variabilités et des imprévus possibles. L'utilisation de cette approche ne doit néanmoins pas priver la conception et la validation de leur capacité à identifier et traiter les cas extrêmes par leur sévérité et leur rareté. Ces cas extrêmes devront continuer à être scrutés dans le cadre de l'homologation et de la validation de sécurité des véhicules.

4.2 Difficultés rencontrées

La construction de cette proposition de méthodologie PRISSMA a été confrontée à de nombreuses difficultés, reflétant la complexité et l'innovation inhérentes à l'intégration de l'IA dans les STRA.

Tout d'abord, les technologies censées permettre l'évolution autonome d'un véhicule ne sont pas toujours complètement matures.

La gestion de la combinatoire et la preuve de couverture de l'OD par l'approche par scénario reste également un défi majeur. Assurer que tous les scénarios pertinents sont bien couverts est une tâche complexe, surtout lorsque les scénarios critiques représentatifs doivent être générés. Un scénario accidentel, par exemple, peut ne pas être représentatif des conditions réelles, compliquant ainsi la validation.

Les travaux portant sur la définition de preuves de non-régression et/ou d'amélioration continue pour des système à base d'IA ne cessent de s'améliorer, sans pouvoir pour le moment être directement applicables en pratique. De même, la détection automatique d'une potentielle sortie de l'ODD nécessite des systèmes de surveillance sophistiqués et fiables, qui restent pour le moment à l'état de prototype.

La définition des seuils de validation a également soulevé des questions cruciales : qu'est-ce qui peut être considéré comme "suffisant" et quels critères d'arrêt doivent être appliqués pour déterminer la validation finale ?

L'un des défis les plus complexes a été l'évaluation en "boîte noire" des algorithmes d'IA et des failles de cybersécurité. La nature opaque de ces algorithmes rend leur validation difficile, notamment pour détecter les vulnérabilités potentielles et assurer la sécurité contre les cyber-attaques.

Certains systèmes ont par ailleurs démontré une performance exceptionnelle en environnement numérique mais ont échoué en environnement réel, soulignant des différences persistantes entre les simulations et les conditions du monde réel. Ce phénomène a révélé la nécessité de capteurs de référence pour valider les capteurs embarqués, garantissant ainsi la précision et la fiabilité des données collectées en conditions réelles.

Enfin, l'interprétation statistique des tests en environnements contrôlés, qui sont souvent en nombre limité, reste difficile. Les échantillons limités ne permettent en effet pas toujours d'obtenir des conclusions statistiquement significatives, pouvant ainsi compliquer la validation finale des systèmes sous test.

En résumé, la construction de la méthodologie proposée a dû surmonter une série de défis technologiques, méthodologiques et pratiques, et force est de constater qu'il en reste plusieurs pour envisager à court terme une validation complète d'un STRA.

4.3 Perspectives futures

Le projet PRISSMA ouvre un certain nombre de perspectives d'études pour l'évaluation et la validation des STRA intégrant des systèmes à base d'IA, que nous recensons ici.

L'une des perspectives futures de ce projet consiste à intégrer et valoriser les travaux réalisés dans les règlements et travaux de normalisation en cours de rédaction. Cette intégration permettra aux avancées réalisées dans ce projet d'être valorisées au sein des pratiques de l'industrie, garantissant ainsi que les meilleures pratiques et les innovations technologiques soient largement adoptées.

Une autre orientation clé est de décliner l'approche basée sur des scénarios aux cas d'utilisation (Use-Case) précis. Cet enrichissement de l'approche scénarios permettra de mieux représenter les situations réelles et d'adapter les tests et les validations aux contextes spécifiques dans lesquels les STRA seront déployés.

Travailler à l'augmentation progressive de l'ODD est également une priorité. Cela inclut la gestion améliorée des situations dégradées par le biais de simulations multi-conditions, ainsi que l'amélioration des interactions du VA avec les autres usagers de la route grâce à des simulations multi-modales. En élargissant l'ODD, les systèmes pourront fonctionner de manière plus fiable et sécurisée dans une gamme plus étendue de conditions et de scénarios.

L'amélioration des modèles de simulation pour les rendre toujours plus réalistes est une autre perspective cruciale. Cela passe par l'augmentation des ensembles de données (datasets) et l'amélioration de leur fidélité à l'aide de modèles génératifs, ainsi que par l'utilisation de preuves formelles pour valider les résultats des simulations. Ces avancées permettront de s'assurer que les simulations reflètent fidèlement la réalité et que les résultats obtenus sont robustes et fiables.

La prise en compte automatique des mises à jour de l'environnement est une autre voie prometteuse, avec l'objectif de passer d'un modèle numérique (digital model) à un jumeau numérique (digital twin). Cela permettra aux outils de simulation de s'adapter en temps réel aux changements de l'environnement, améliorant ainsi leur réactivité et leur performance.

L'exploitation des tests en réalité augmentée et réalité améliorée dans les processus d'évaluation constitue également une perspective importante. Les tests Vehicle-in-the-Loop (VIL) et Human-In-the-Loop (HITL)

offrent des opportunités uniques pour évaluer les systèmes dans des conditions proches de la réalité, tout en bénéficiant des contrôles offerts par les environnements simulés.

Appliquer la méthode de validation développée sur un système complet sera également une future étape clé. Cela permettra de tester et de valider l'efficacité et la robustesse des approches développées dans un contexte intégral et opérationnel, offrant des indications précieuses pour des déploiements à grande échelle.

Enfin, une question cruciale pour l'avenir est de comprendre comment utiliser l'IA pour valider les modèles d'IA. Cette réflexion portera sur l'élaboration de méthodes et d'outils permettant de garantir que les systèmes d'IA peuvent être validés de manière rigoureuse et fiable, assurant ainsi leur sécurité et leur efficacité dans des applications réelles.

4.4 Appels à l'action

Pour continuer à faire progresser les STRA intégrant des briques d'IA, il semble essentiel de maintenir et d'investir dans des environnements de simulation temps réel de haute fidélité. Cela inclut également la mise en place d'outils avancés de génération de données haute fidélité afin d'optimiser la couverture des situations potentielles rencontrées. En améliorant la précision et la représentativité des simulations, nous renforcerons la fiabilité et la sécurité des STRA.

Il est également crucial d'encourager la mutualisation des efforts entre partenaires académiques et industriels sur les problématiques de validation. Cela comprend la génération de données labellisées certifiées (vérité terrain), la mise à jour de jumeaux numériques et le maintien de sites d'expérimentation en milieu ouvert. Cette collaboration renforcée permettra de développer des standards communs et de partager les meilleures pratiques pour une validation efficace et robuste des STRA.

De manière connexe, il semble nécessaire de soutenir le développement d'outils permettant la construction rapide et de qualité de représentations numériques telles que les modèles numériques, les ombres numériques (digital shadows) et les jumeaux numériques (digital twin). Ces outils sont essentiels pour simuler, tester et valider les STRA dans des environnements virtuels avant leur déploiement dans le monde réel.

Il nous semble par ailleurs important de partager la méthodologie développée dans le projet PRISSMA auprès de l'International Working Group on Automated Driving Systems (IWG ADS) et du Working Party on Automated Driving Systems (WS on ADS) de la CEE-ONU. Cette action favorisera l'adoption généralisée de normes de validation et de pratiques recommandées au niveau international, assurant ainsi une approche harmonisée et efficace pour les STRA.

Pour promouvoir une adoption plus large et faciliter l'innovation, il est ensuite essentiel d'encourager l'interopérabilité et le caractère open-source des outils de simulation. Cela permettra une collaboration plus étroite et une adoption facilitée des normes communes et des meilleures pratiques, soutenant ainsi le développement harmonieux des technologies automatisées.

De même, le partage continu de bonnes pratiques et de cas tests "intéressants" est également recommandé pour assurer le maintien en condition opérationnelle et la non-régression des performances des systèmes automatisés. Cette initiative favorisera l'amélioration continue et l'adaptation aux évolutions technologiques et réglementaires.

Enfin, pour stimuler l'innovation et l'adoption à long terme des systèmes automatisés/autonomes, il semble crucial d'arriver à lancer des expérimentations sous surveillance à long terme (par exemple dans une ou plusieurs ville(s) européenne(s)). Ces expérimentations permettront de tester et de valider les technologies dans des environnements réels, tout en recueillant des données précieuses pour améliorer la conception et la sécurité des

STRA.

Références

Cette partie recense l'ensemble des livrables produits par le projet PRISSMA sur lesquels se base le présent document. Ces livrables sont consultables et téléchargeables sur le site https ://prissma.univ-gustave-eiffel.fr.

L1.3 WP1 STATE OF THE ART

L1.5 TESTS AND AUDIT REQUIREMENTS

L1.6 EVALUATION METHODOLOGY AND QUALIFICATION OF TEST EQUIPMENT

L2.3 WP2 STATE OF THE ART

L2.5 DEFINITION OF INTERFACES AND SIMULATION ENVIRONMENT : STRENGTH, WEAKNESS, ADVANTAGE, RISK, LIMIT, CONSTRAINTS

L2.6 DEFINITION OF PROCEDURES OF SCENARIO MANAGEMENT AND RESULTS ANALYSIS – INITIAL REPORT

L2.7 METHODOLOGY, PROCEDURES AND PROTOCOLS FOR EVALUATION OF APPLICATIONS AND VIRTUAL TEST FACILITIES

L2.9 PROOFS-OF-CONCEPT : DEVELOPMENT OF PLATFORMS MEETING THE DESIRED OBJECTIVES OF EVALUATING MEANS OF AUTOMATED MOBILITY

L3.1 WP3 STATE OF THE ART AND INVENTORY OF THE EXISTING SITUATION

L3.2 HOMOLOGATION TESTS PROTOCOLS & STRATEGY

L3.3 SECOND CONTROLLED ENVIRONMENT TEST CAMPAIGN AS PART OF THE GLOBAL HOMOLOGATION PROCESS FOR AUTONOMOUS MOBILITY

L3.4 APPROVAL PROCEDURE FOR CONTROLLED ENVIRONMENT TEST EQUIPMENT

L3.5 CARRYING OUT THE FIRST EVALUATION TEST CAMPAIGN IN A CONTROLLED ENVIRONMENT AND PRODUCING TEST REPORTS

L3.6 CARRYING OUT THE SECOND EVALUATION TEST CAMPAIGN IN A CONTROLLED ENVIRONMENT AND PRODUCING TEST REPORTS

L4.1 TESTS IN REAL CONDITIONS STATE OF THE ART

L4.2 REAL CONDITION TESTS METHODS : INFRASTRUCTURE ANALYSES, PATHWAY SELECTION CRITERIA AND DEFINITION OF RELEVANT SCENARIOS

L4.4 METHODOLOGY AND PROCEDURE FOR TESTING IN REAL CONDITION

L4.5 POC IMPLEMENTATION IN REAL CONDITIONS

L5.1 REPORT ON CONNECTIVITY PERFORMANCES CONS- TRAINTS IN AUTONOMOUS VEHICULES ENVIRONMENT

L5.2 REPPORT ON CYBER-THREAT ANALYSIS IN AUTONOMOUS VEHICULAR ECOSYSTEMS

L5.3 CYBERSECURITY OBJECTIVES FOR AUTONOMOUS VEHI- CLE SYSTEMS

L5.4 INTEROPERABILITY AND PERFORMANCE OBJECTIVES FOR AUTONOMOUS VEHICLE SYSTEMS

L5.5 RELIABILITY OF AUTONOMOUS VEHICLE SYSTEMS IMPLEMENTATIONS

L5.6 SECURITY TESTS REPORT OF AI BASED AUTONOMOUS SYSTEMES

L5.7 TESTS REPORT ON THE SECURITY AND EFFICIENCY OF THE COMMUNICATIONS

L6.3 WP6 STATE OF THE ART RISK ASSESSMENT AND CERTIFICATION FOR AI

L6.5 WP6 INTEGRATION OF AI SPECIFICATIONS INTO DESIGN STANDARDS

L6.6 EVIDENCE TO BE PROVIDED ON THE DEVELOPMENT PROCESS, WITH IDENTIFICATION OF

POSSIBLE COVERAGE OF APPROVAL REQUIREMENTS

L7.1 SAFE AND SECURE PROCESS FOR DATA COLLECTION

L7.2 DATA ANALYSIS PROCESS AND IDENTIFICATION OF SINGLE SITUATIONS

L7.3 MISE A JOUR CONTROLEE DES ELEMENTS DU DOSSIER (DONT LES SCENARIOS ET LES OUTILS)

L7.4 SECURE DEPLOYMENT PROCESS OF UPDATES

L8.1 COMMON TERMINOLOGY FOR THE PRISSMA PROJECT

L8.2 HIGH-LEVEL SAFETY RULES AND IDENTIFICATION OF THEIR DOMAINS

L8.3 SUPPORT TO VALIDATION AND APPROVAL

L8.4 REFERENCE REPORT ON VALIDATION PRINCIPLES AND PROCESSES AS WELL AS THE ACTORS' REPARTITION

L8.8 PROGRESS OF THE POSITIONS EXPRESSED BY THE IDENTIFIED GROUPS AND PROVISION OF ESSENTIAL ELEMENTS TO THE ACTORS CONCERNED IN CHARGE OF FORMULATING THE FRENCH POSITIONS

L8.9 OPERATIONAL DESIGN DOMAIN

L8.10 OPERATIONAL DOMAIN METHODOLOGY FOR SPLITTING A PREDETERMINED ROUTES INTO SECTIONS BASED ON PRISSMA'S TAXONOMY

L8.11 OPERATIONAL DESIGN DOMAIN

L8.12 USING PRISSMA'S TAXONOMY FOR ODD DEFINITION AND SCENARIO IMPLEMENTATION

L8.13 REFERENCE REPORT ON SYSTEM ENGINEERING

L8.15 FINAL REPORT ON THE IMPACT OF AI IN SYSTEM ENGINEERING CHOICES