THALES
Building a future we can all trust

Quels enjeux technologiques pour une IA de défense ?

Juliette Mattioli
Senior Experte IA

www.thalesgroup.com

Pour mémoire…

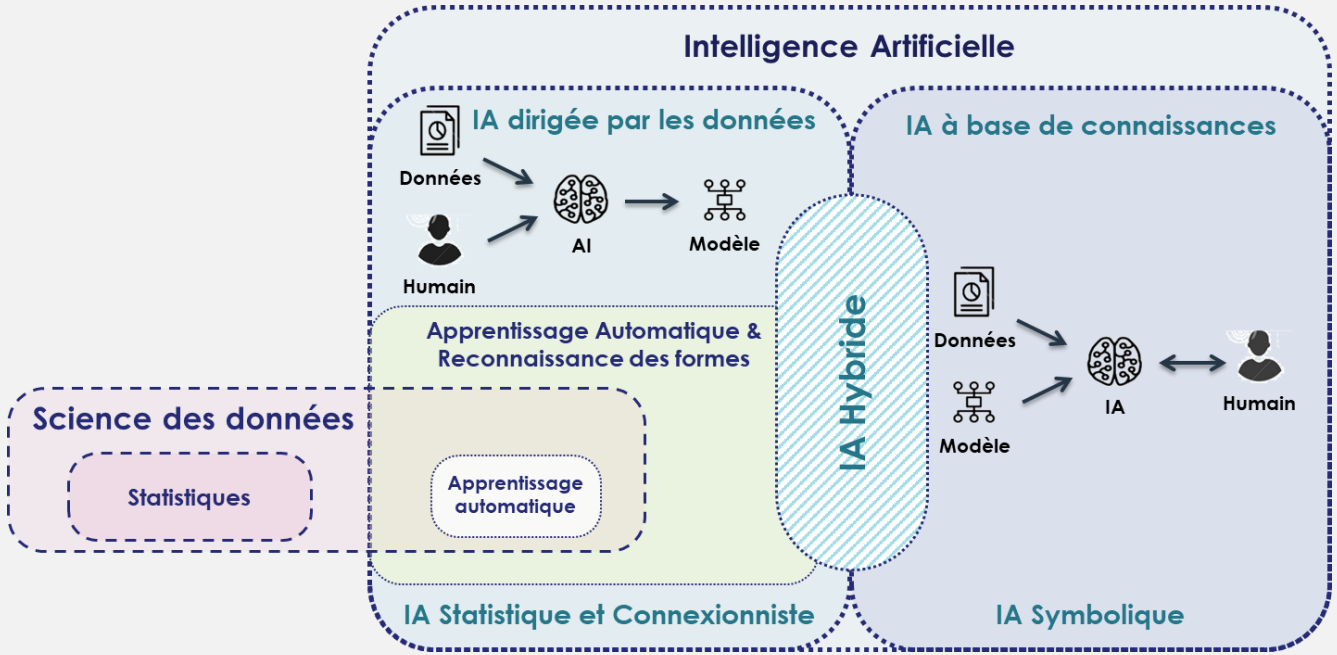
Inputs

Données

Informations

Connaissances

Les différents paradigmes



Outputs

> Perception
informations riches, complexes et imparfaites

> Apprentissage
à partir d'exemples

> Abstraction
création de sens

> Raisonnement
découverte de connaissances, planification et décision

> Communication
Dialogue naturel

> Action
pour atteindre un objectif rationnel

Première illustration de la zoologie des technologies d'IA

Communication : Capacité à dialoguer et à comprendre l'intention du dialogue

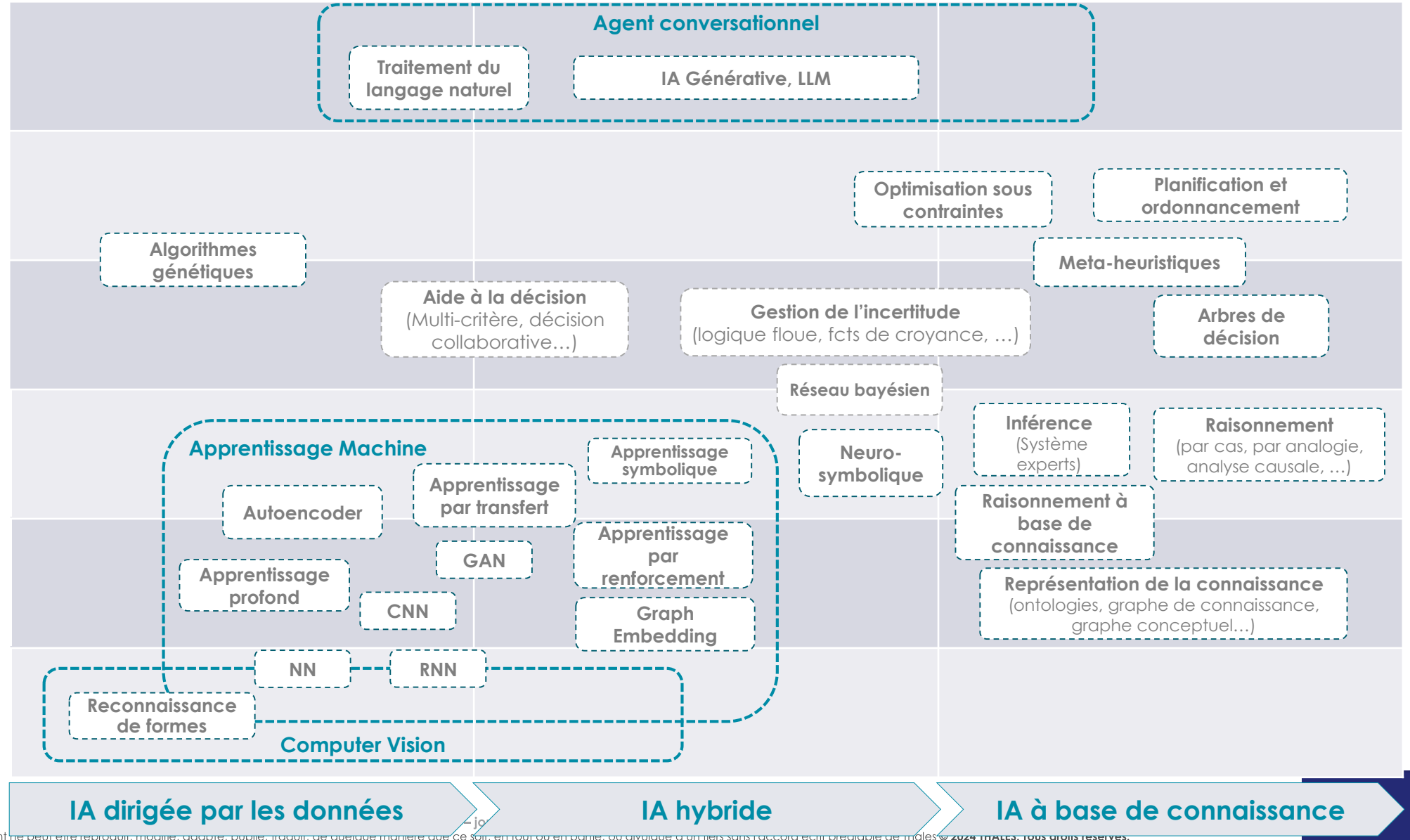
Planification : Capacité à fixer et à atteindre des objectifs

Décision : Processus consistant à faire des choix parmi des alternatives possibles

Raisonnement : Capacité à résoudre des problèmes

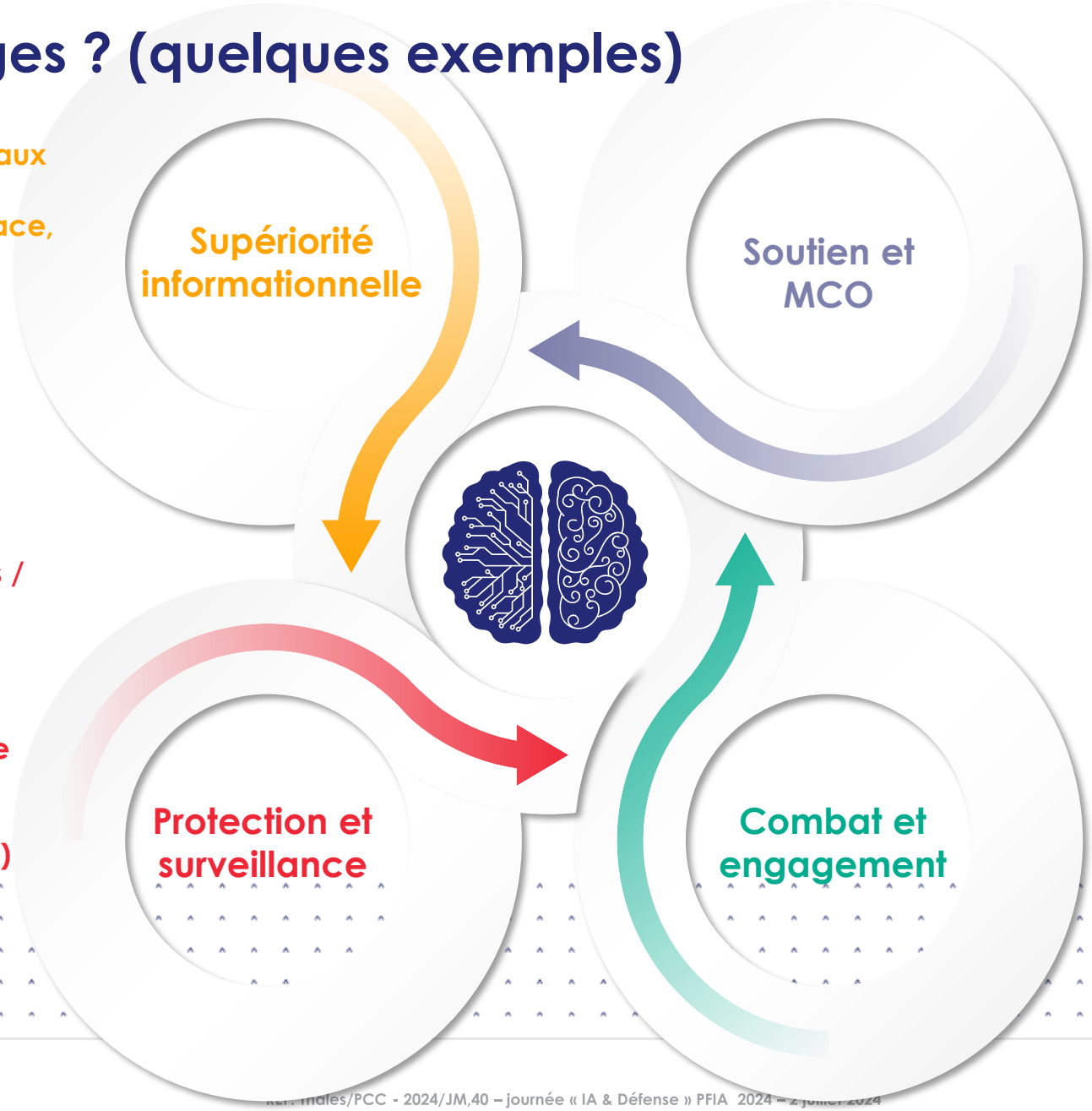
Connaissances : Capacité à représenter et à comprendre le monde

Perception : Capacité à transformer des données brutes (images, sons…) en informations utilisables



Pour quels usages ? (quelques exemples)

- Traitement des données liées aux capteurs spécifiques défense.
- Analyse du niveau de la menace, compréhension de l'intention
- Renseignement (incl. OSINT)
- PsyOps
- Cybersécurité (incl. cyber-sécurité de l'IA) et influence numérique
- …



- Gestion de flotte de véhicules (incl. Cannibalisme)
- Maintenance (prescriptive, prédictive)
- Débriefing de mission
- Logistique
- Soutien en santé
- …

- Aide au déploiement capteurs / effecteurs
- Allocation des ressources de veille, de reconnaissance et de poursuite active
- Détection et caractérisation de menace
- Fusion d'informations pour la tenue de situation (IMINT, ELINT…)
- Protection de convois
- Alerte avancée
- …

- Détection et engagement de cibles de faible SER
- Aide à la décision en planification et en conduite
- Allocation des moyens en fonction du contexte et de la doctrine
- Combat mono-milieu
- Combat multi-milieu
- Combat collaboratif
- Lutte anti-drones, essaims de drones
- …

{OPEN}

Facteurs d'optimisation et d'acceptation des systèmes à base d'IA

> Données vs. connaissances

- Données : condition nécessaire aux techniques d'IA connexionniste et statistiques
- Connaissances métier pour une meilleure prise en compte des règles d'emploi et d'engagement

> Compromis entre performances vs. tempo des actions/opérations

- Décision dans un environnement dynamique et incertain (frugalité en données, informations imparfaites, temps réel)
- Prise en compte de contraintes d'embarquabilité (SWaP, Edge AI…)

> Collaboration humain – système efficiente

- Explicabilité adaptée à l'usage et au contexte mais aussi aux règles d'emploi et d'engagement
- Approche UX pour optimiser la place de l'humain dans la boucle

> Validation et qualification du système

- Méthodologie de bout en bout pour valider, qualifier et maintenir un système d'IA dans son domaine de conception opérationnelle (ODD)
- Formation des ingénieurs d'essai (ex. EPNER) pour comprendre et minimiser les incidents

Quelques contraintes de l'IA à base de données

> Les données

- Qualité de la donnée (label, précision, biais, fraicheur, complétude)
- Frugalité des jeux de données (transfer learning, few shot learning, active learning…)
- Préservation de la confidentialité et de la propriété des données dans une coopération internationale (Federated learning, multi-party computation, homomorphic encryption…)

> Apprentissage continu en fonction de l'évolution du contexte (menace, doctrine, …)

- Vers des systèmes plus ouverts
- Méthodologie de validation/vérification incrémentale
- Mise en place de méthodologie de déploiement des mises à jour.

> Au-delà de l'IA connexionniste et statistique

{OPEN}

L'IA de confiance : condition nécessaire au déploiement de l'IA dans les systèmes critiques

Validité

Garantir qu'un système à base d'IA fait ce qu'il doit faire, tout ce qu'il doit faire et seulement ce qu'il doit faire

Transparence et explicabilité

Fournir des justifications et des explications compréhensibles et adaptées au contexte.

Responsabilité

Respecter les cadres éthiques, juridiques et réglementaires et être conformes aux standards

Sécurité et robustesse

Garantir la robustesse et la résilience aux conditions adverses, telles que le leurre et les cyber-attaques, mais aussi au mauvais usage

Gouvernance de la donnée

(RGPD et qualité de la donnée)

Fiabilité et sûreté

Contrôler le risque de défaillances inacceptables et d'insuffisances fonctionnelles

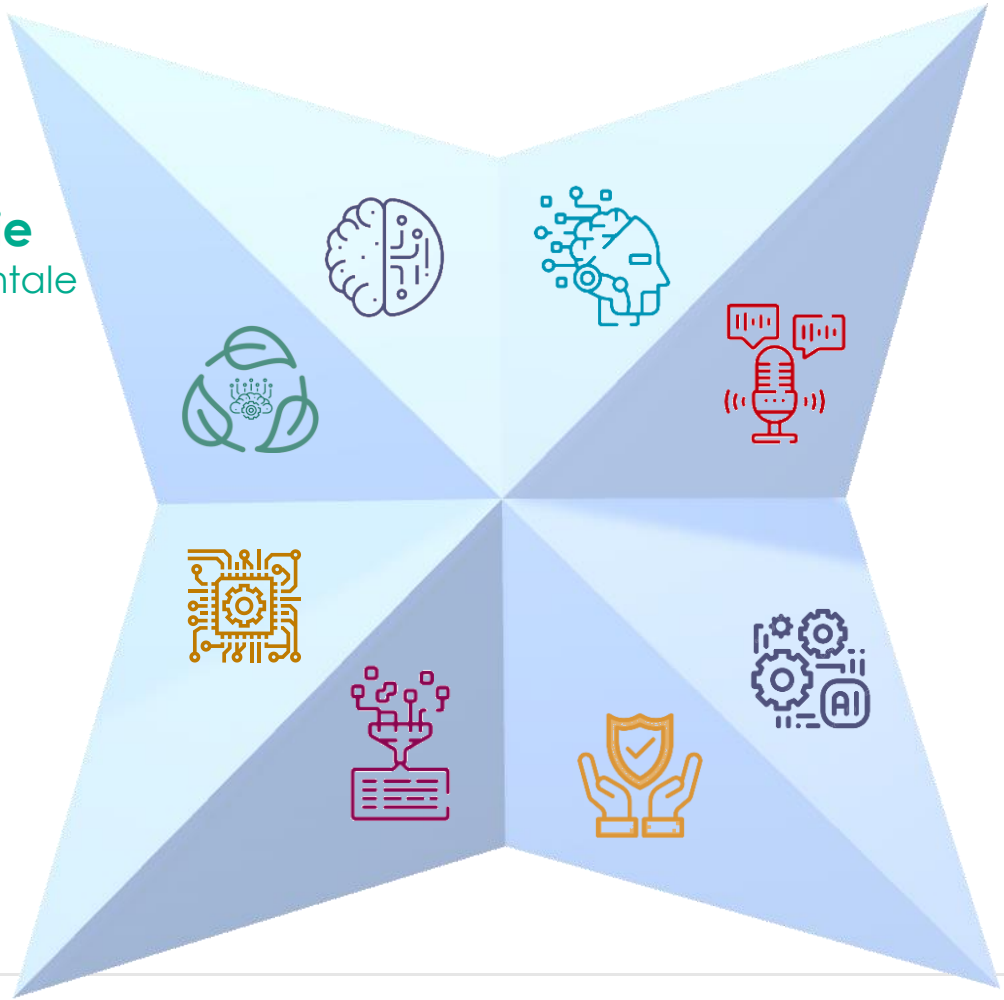
Les techonologies d'IA différentiantes (@ Thales)

IA hybride
IA connexionniste et statistique
⊕ IA symbolique

IA frugale en donnée et énergie
Vers une IA à faible impact environnementale
Données simulées et synthétiques
"Smart data" versus "Big data"

IA embarquée
Prise en compte des contraintes
SWaP (Taille, poids et puissance)

IA Générative
LLM pour les systèmes critiques
IA générative de confiance,
fiable et responsable



Apprentissage par renforcement et automomie
Environnement de simulation
Digital Twin

Dialogue Humain-Machine
Interaction contextuelle et intuitive
IA auto-explicable

Intelligence Collaborative
Systèmes multi-agents
IA distribuée

Ingénierie de l'IA de confiance
Conception, Développent,
Qualification, Certification

REF: Thales/PCC - 2024/JM,40 – journée « IA & Défense » PFIA 2024 – 2 juillet 2024

{OPEN}

Développer et déployer un système critique à base d'IA nécessite

des méthodes et outils d'ingénierie sur le cycle de vie de bout-en-bout

pour assurer la qualification et la conformité aux réglementations, standardisations et normes

{OPEN}

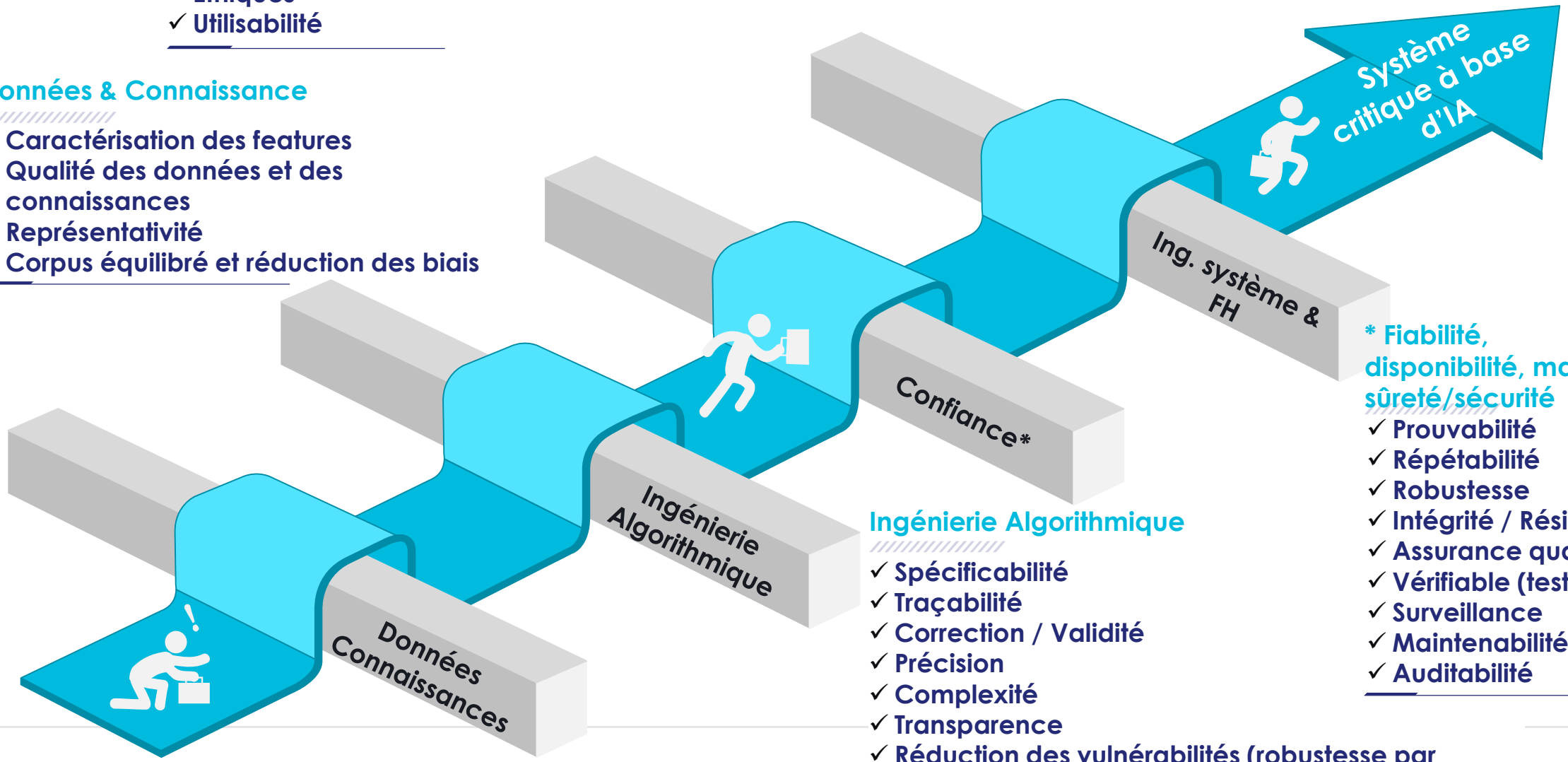
Ce qui induit de nouveaux enjeux d'ingénierie

Ing. système et facteurs humains

- ✓ Performance
- ✓ Interprétabilité/ Explicabilité/ Transparence
- ✓ Dialogue Humain-IA
- ✓ Ethiques
- ✓ Utilisabilité

Données & Connaissance

- ✓ Caractérisation des features
- ✓ Qualité des données et des connaissances
- ✓ Représentativité
- ✓ Corpus équilibré et réduction des biais



Ingénierie Algorithmique

- ✓ Spécificabilité
- ✓ Traçabilité
- ✓ Correction / Validité
- ✓ Précision
- ✓ Complexité
- ✓ Transparence
- ✓ Réduction des vulnérabilités (robustesse par conception)

* Fiabilité, disponibilité, maintenabilité, sûreté/sécurité

- ✓ Prouvabilité
- ✓ Répétabilité
- ✓ Robustesse
- ✓ Intégrité / Résilience
- ✓ Assurance qualité
- ✓ Vérifiable (test)
- ✓ Surveillance
- ✓ Maintenabilité
- ✓ Auditabilité

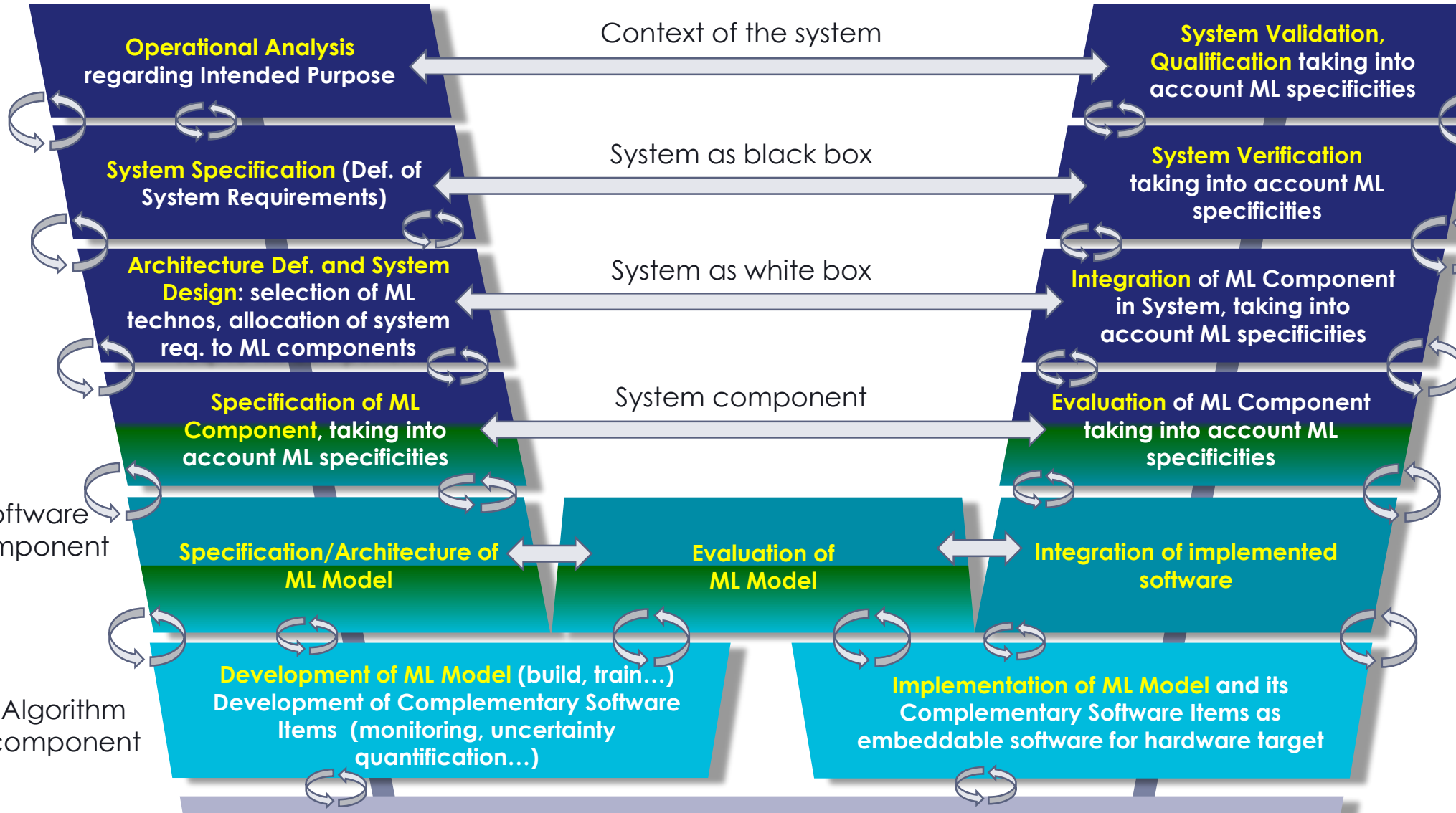
Sans oublier les nouvelles réglementations

RGPD *Depuis mai 2018*	Gouvernance des données *Depuis septembre 2023*	DATA ACT *Septembre 2025*	AI ACT *Mai 2025 pour les systèmes à haut risque*	Responsabilité civile de l'IA *En discussion*
Responsabilité **Droit d'accès** **Notification des violations** **Amendes harmonisées**	**Intermédiaires de données** **L'altruisme des données** **Gouvernance de l'UE en matière de données** **Souveraineté des données**	**Accès aux données générées par l'IoT** **Partage des données avec le secteur public** **Changement plus facile de fournisseur de services cloud** **Souveraineté des données**	**Approche basée sur le risque** **Robustesse, précision et cyber-sécurité** **Transparence et explicabilité** **Gouvernance et qualité des données** **Surveillance humaine** **Évaluation de la conformité**	**Adapter la responsabilité civile à l'IA** **Révision des règles relatives à la responsabilité en fonction des produits**
Traitement des données à caractère personnel de manière responsable pour un meilleur contrôle	**Règlement visant à faciliter le partage des données au sein de l'UE**	**Règlement visant à faciliter le partage des données au sein de l'UE**	**Des exigences et des obligations adaptées à une "approche fondée sur les risques".**	**Garantie pour les personnes lésées par l'IA bénéficient du même niveau de protection que celles lésées par d'autres technologies**

REF: Thales/PCC - 2024/JM,40 – journée « IA & Défense » PFIA 2024 – 2 juillet 2024

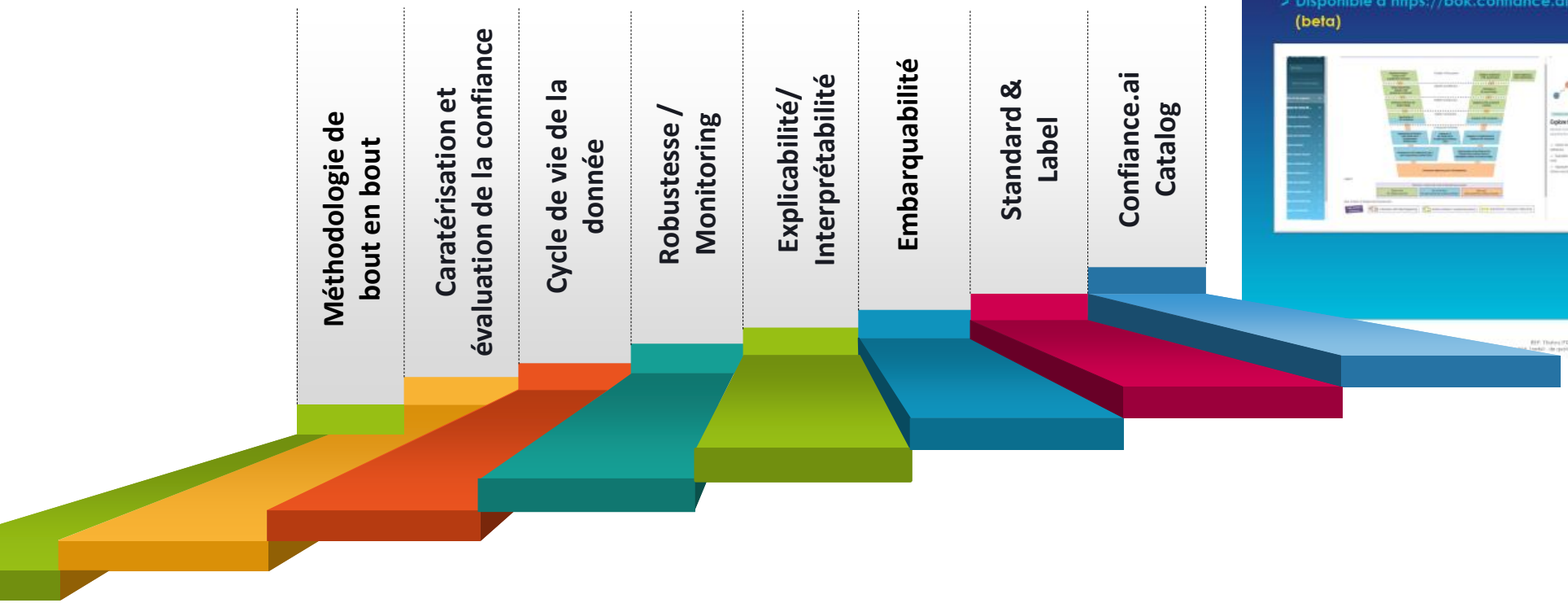
{OPEN}

Cycle de vie des systèmes/logiciels/algos/données pour concevoir un système basé sur l'apprentissage



Leveraging the Concepts of Design Assurance for Neural Networks – CoDANN [EASA]

Le programme confiance.ai (système critique à base d'apprentissage)



Open-source building blocks

32

Full Confiance.ai IP building blocks

26

Méthodologies & Guides

34

Etat de l'art

28

Benchmark et cas d'usage

44

Publications

45

Les enjeux de souveraineté

> Quelle stratégie ?

- La stratégie nationale de l'IA (comité d'évaluation ANR IA Cluster, comité l'IA générative)
- Les infrastructures vs. Cloud souverain, supercalculateur à la « Jean Zay », AMIAD
- Les composants (NVIDIA) vs. l'Open Source HW (Greenwaves)
- Les logiciels et l'Open Source SW (PyTorch, Yolo, OpenAI…) vs. Scikit Learnt (Probabl:) / AMIAD Recherche

> Expertise technique de l'IA

- Une maitrise des algorithmes de modélisation (au-delà des approches à base d'IA connexionniste)
- Vers une IA de confiance et responsable outillée (incl. Friendly Hacking de l'IA)
- Formation à la gestion de programme complexe à composante d'IA (ingénierie de l'IA, campagne expérimentale de tests, évaluation de l'IA – LNE/INRIA…)

> Expertise métier

- Définition de nouveaux concepts et usages opérationnels
- Gestion des données de défense (collecte, stockage, qualité des données)
- Campagne d'essais

Conclusion

Animation de la communauté

Acculturer et accompagner l'ensemble des acteurs de défense pour leur permettre d'identifier de nouvelles capacités, d'induire de nouveaux concepts opérationnels…

Excellence

Garantir l'excellence et l'efficacité des solutions à base d'IA, étayée sur une connaissance opérationnelle.

Méthodologie

Se doter d'une approche méthodologique outillée de bout en bout pour la conception, le déploiement et le maintien en condition opérationnelle et de sécurité d'une IA de confiance et responsable

Ecosystème

Fédérer un écosystème de partenaires (civil, académique, industriel) français élargis avec des partenaires européens ou d'accords binationaux.

Souveraineté

Des infrastructures, de composants et/ou open-source HW, des logiciels et/ou open-source mais aussi expertises technologique fortement couplée au métier

{OPEN}

THALES
Building a future we can all trust

Thank you

www.thalesgroup.com