



CONFEIANCE CONFEIANCE CONFEIANCE

↑
4



Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

armasuisse



CONFIDANCE

Office fédéral de l'armement
armasuisse Sciences et technologies
Gestion de la recherche et de recherche opérationnelle
Feuerwerkerstrasse 39, CH-3602 Thun

ISBN: 978-3-9525890-1-4

contact: quentin.ladetto@armasuisse.ch

DES CAHIERS D'EXERCICES

pour anticiper les ruptures

Deftech, le programme de prospective technologique d'armasuisse Sciences et Technologies, a pour objectif d'anticiper les usages de rupture pour la défense suisse.

Cette anticipation repose sur une compréhension de technologies et de leurs évolutions combinées. Il faut appréhender les effets qu'elles produiront et la manière dont elles pourront être utilisées pour relever différents défis. Pour y arriver, on doit aussi envisager les différentes situations futures et réfléchir aux contextes dans lesquels pourraient évoluer ces technologies.

Les cahiers **deftech exploration** s'inscrivent dans cette démarche. Ils proposent une exploration-compréhension de différents domaines technologies ainsi que de thématiques liées à un ou à la convergence de ceux-ci.

Chaque cahier est organisé selon une même structure. Vous y trouverez une présentation de la technologie, des actualités liées au thème et de courtes fictions prospectives. Ces différents éléments sont accompagnés d'exercices qui vous aideront tant à explorer les différentes facettes des domaines technologiques qu'à vous projeter dans le futur. Le format standard A4 permet indistinctement une impression et reliure chez un professionnel tout comme l'impression sur une imprimante de bureau.

L'exploration accompagnée des exercices peuvent être effectués individuellement, mais ils sont aussi pensés pour un travail à plusieurs lors d'ateliers.

Confiance

Longue à établir prompt à disparaître, le sentiment de confiance est tout sauf évident dans sa construction comme dans sa substance. Confiance - con-fides - cet élan absolu de l'âme, affaire de croyance qui s'abstrait de la raison et de l'analyse comme nous l'enseigne l'étymologie du mot, sous-entend également connaissance de l'autre, mais également de soi ; relation avec autrui dans le temps et dans l'intensité. Réciproque ou asymétrique, la confiance sera également dépendante des conséquences et de l'impact généré par l'action où celle-ci est requise.

Mais en qui ou quoi plaçons-nous notre confiance ? Est-ce en une personne, un groupe d'individus, une entreprise, un état, une croyance ? Peut-il être également question de confiance en parlant d'utilisation de dispositifs électroniques, d'algorithmes, de relations et d'interactions entre l'humain et la machine, entre le physique et le digital, entre le sensible et l'intangible ?

En cette période à l'empreinte fortement numérique, où le comportement d'un matériel peut radicalement changer par une mise à jour logicielle, il nous a semblé pertinent de considérer les implications de la technologie sur cette thématique ainsi que de présenter les défis à relever.

Bonne lecture, ou plutôt, bonne exploration !



Quentin Ladetto

Responsable prospective technologique
armasuisse Sciences et Technologies

QU'EST-CE QUE LA CONFIANCE ?

1

Où l'on définit la confiance et comprend que c'est un ciment social aussi invisible que nécessaire pour la vie en société.

DÉFINITIONS

Pour vous mettre en appétit, nous allons vous proposer des définitions de la confiance.

Notez de 1 à 5 étoiles chacune d'elles.

Proposez ensuite vos définitions de la confiance.

1. La confiance est la croyance en la fiabilité, l'intégrité, la compétence et l'intention bienveillante d'une autre personne, d'un groupe ou d'une machine.

* ○ — ○ — ○ — ○ — ○

2. La confiance est un état d'esprit ou une émotion positive que l'on ressent envers quelqu'un ou quelque chose.

* ○ — ○ — ○ — ○ — ○

3. La confiance est une attitude vis-à-vis d'un ou plusieurs humains ou d'une machine devant atteindre un objectif dans une situation d'incertitude.

* ○ — ○ — ○ — ○ — ○

4. « La confiance est le sentiment de se sentir en sécurité, même lorsque tout semble incertain. » — Paulo Coelho

* ○ — ○ — ○ — ○ — ○

5. La confiance est une colle invisible qui construit la société.



6. La confiance désigne la relation de compréhension mutuelle, de respect et de coopération entre des pays. Elle implique le respect des accords, des engagements et des intérêts mutuels.



7. La confiance est ce qui nous pousse à prendre des risques dans des situations d'incertitude et à être convaincus que les choses se dérouleront de manière positive.



8. La confiance est l'assurance qu'une personne accorde à une machine ou à un système automatisé pour agir de manière fiable, précise et conforme à ses attentes.



La confiance _____

ON EXPLORE

Après avoir envisagé les subtilités inhérentes à la notion de confiance, vous allez partir en exploration active du sujet.

La confiance active

La confiance est un moteur d'action.

Lorsqu'on a confiance en quelqu'un, on est plus enclin à collaborer, suivre ses conseils, prendre des décisions, s'engager...

Listez 6 situations où vous faites confiance à un individu, une machine, un animal, une institution...

VOIR
PAGE
57

Variez les dates : trois situations ont lieu aujourd'hui, trois dans un avenir proche ou lointain. Précisez les dates.

Envisagez des situations quotidiennes ou exceptionnelles, personnelles ou professionnelles, civiles et militaires...

Pour chaque situation, vous allez proposer un ou plusieurs actes que cette confiance vous pousse à effectuer.

Exemple :

2038.

J'ai confiance... dans les prévisions de la météo.

→ Je prends mon parapluie quand on annonce de la pluie.

→ Je vérifie que les paradrone (les parapluies protégeant les drones) des drones de surveillance sont bien sortis.

Les raisons de la confiance

La confiance est un processus complexe basé sur une réflexion logique, des émotions et des expériences passées. On l'établit pour diverses raisons.

Quelques exemples.

→ Fiabilité et cohérence

Une personne tient ses promesses, respecte ses engagements et agit de manière prévisible.

→ Intégrité et honnêteté

Une personne se comporte de manière éthique, honnête et transparente et agit de manière juste et fiable.

→ Compétence et expertise

Une personne dispose de compétences et d'une expertise dans un domaine spécifique. On fait confiance à un médecin en raison de sa formation médicale et de son expérience.

→ Expérience passée

Une personne a prouvé par le passé qu'elle était digne de confiance.

→ Empathie et compréhension

Une personne fait preuve d'empathie, d'écoute active et de compréhension envers nos besoins et nos préoccupations.

Vous allez compléter votre tableau confiance en indiquant pourquoi vous accordez votre confiance... Vous pouvez lister plusieurs raisons.

Exemple :

Pourquoi faire confiance aux prévisions météo ?

- Les météorologues ont une expertise basée sur des connaissances scientifiques solides.
- Les sources d'observation sont nombreuses : satellites, radars, bouées océaniques, etc.
- Aucune erreur de prévision depuis 32 jours (Ne rêvez pas, nous sommes en 2038!)

VOIR
PAGE
57

La confiance, un ciment social

« Sans la confiance des hommes les uns envers les autres, la société tout entière se disloquerait. »

Georg Simmel, sociologue (1900)

« La confiance est la pierre angulaire de tout l'édifice militaire. »

Sun Tzu, stratège chinois

**La confiance est un ingrédient indispensable de la vie sociale.
Sans confiance, les citoyens...**

- Ne déposeraient pas leur argent dans une banque,
- Ne confieraient pas le destin politique de leur pays à des élus ou sa protection à l'armée.
- Ne se tourneraient pas vers des magistrats pour obtenir justice.
- Ne s'en remettraient pas à leur médecin pour conserver une bonne santé.

« La confiance est l'essence même du leadership. Sans confiance, les ordres ne sont que des mots vides. »

General Colin Powell, ancien président du Joint Chiefs of Staff des États-Unis.



IL ÉTAIT UNE FOIS

Depuis des siècles, la confiance est au cœur du contrat social.

“

De la République au Haut-Empire, la confiance était représentée, sur les monnaies, par deux mains jointes, symbole de l'accord passé entre deux individus, constitue un concept essentiel du vocabulaire institutionnel et des pratiques politiques des Romains,

explique Antony Hostein, directeur d'études à l'École pratique des hautes études.

”

Trahison de la confiance

La confiance étant le ciment sociétal, les trahisons ont des conséquences dramatiques. L'histoire comprend de nombreux épisodes de trahison de confiance.

→ ILETAIT UNE FOIS

JUDAS, UNE RÉFÉRENCE BIBLIQUE

Judas était l'un des douze apôtres de Jésus-Christ. Selon les Évangiles, il a livré Jésus aux autorités juives en échange de trente pièces d'argent.



MATA HARI

De son vrai nom *Margaretha Zelle*, Mata Hari était une danseuse exotique et une espionne pendant la Première Guerre mondiale. Recrutée par les services de renseignement français et allemands, elle fut accusée de trahison par les Français. Elle a été condamnée à mort et exécutée par un peloton d'exécution en 1917



LES ÉPOUX ROSENBERG

Julius et Ethel Rosenberg, un couple américain, ont été accusés d'espionnage au profit de l'Union soviétique. Julius Rosenberg travaillait sur la bombe atomique. Avec sa femme, il aurait fourni des informations sensibles aux Soviétiques. En 1953, les époux Rosenberg ont été reconnus coupables d'espionnage et condamnés à mort.

LE SOLDAT MANNING



Bradley (devenu Chelsea) Manning, un soldat américain, est à l'origine d'une des plus grandes fuites d'informations classifiées de l'histoire des États-Unis. En 2010, Manning est analyste du renseignement militaire en Irak. Il transmet des centaines de milliers de documents confidentiels au site WikiLeaks, notamment des câbles diplomatiques et des vidéos de frappes aériennes.

ENIGMA

Enigma était une machine utilisée par les Allemands pendant la Seconde Guerre mondiale pour coder leurs communications. Grâce aux milliards de combinaisons possibles, les Allemands avaient une confiance aveugle dans leur machine. Trente ans plus tard, ils apprennent qu'une partie de leurs transmissions avaient été décodées par les Alliés. Ce fut très utile lors du Débarquement.



Complétez le tableau en indiquant ce qu'il se passerait si dans les situations décrites, vous perdiez confiance.

La perte de confiance peut avoir des conséquences dramatiques, anecdotiques ou inattendues. Soyez créatifs !

Exemple :

2038.

On a plus confiance dans les prévisions météo.

Les présentateurs météo disparaissent.

On lève le nez au ciel pour voir le temps qu'il fait.

VOIR
PAGE
57

Confiance relative

On a confiance ou pas confiance.

La confiance s'accorde ou non en fonction de différent paramètres.

On peut avoir confiance...

- dans les analyses d'un scientifique, mais émettre des doutes sur ses positions politiques.
- à Amazon pour livrer des colis, pas du tout pour protéger nos données.

- dans un professeur pour fournir un enseignement de qualité, mais pas pour favoriser l'égalité des chances dans la classe.

- dans un service de livraison alimentaire pour une livraison rapide, mais pas pour assurer la qualité des aliments livrés.

La confiance est une donnée fragile. Sa durée est limitée dans le temps. Un événement peut changer totalement la donne.

IL ÉTAIT UNE FOIS

En 2015, on a découvert que le constructeur automobile allemand Volkswagen avait contourné les normes environnementales en installant un logiciel de trucage des émissions. Ce scandale a entraîné une chute importante de la valeur des actions de l'entreprise, des amendes et une perte de confiance des consommateurs.

Fondée en 2003, *Theranos* promettait d'inventer une machine capable de réaliser des centaines de diagnostics sanguins à partir d'une seule goutte de sang prélevée sur le doigt. Cette jeune pousse, valorisée jusqu'à 9 milliards de dollars, avait mis au point une vaste escroquerie. Quand il fut découvert que les prétentions de

l'entreprise étaient basées sur des mensonges et des manipulations, on a assisté à une perte de confiance massive des investisseurs et des partenaires commerciaux.

Pendant la guerre en Irak, des abus et des actes de torture ont été perpétrés par des soldats américains sur des prisonniers détenus à la prison d'*Abu Ghraib*. Ces actes ont été photographiés et ont choqué le monde entier, entraînant une perte de confiance dans les forces armées américaines et remettant en question leur respect des droits de l'homme et des conventions internationales.

Reprenez votre tableau « confiance ». Précisez votre degré de confiance et indiquez des erreurs ou problèmes qui, s'ils arrivaient, ne détruiraient pas la confiance.

Exemple :

J'ai confiance dans les prévisions météo.

Pour organiser mes soirées.

Je n'ai pas confiance pour favoriser la mise en place d'une agriculture responsable.

Erreur ou problème : décalage de prévisions d'une heure.

2038 FICTION

➔ Maintenant qu'on pouvait vivre jusqu'à deux cents ans, Mario avait une confiance dans le progrès. Il espérait qu'un jour les scientifiques trouveraient le moyen de faire arrêter l'ascenseur au bon étage.

CONFIANCE MACHINE, LE NERF DE LA TRANSFORMATION NUMÉRIQUE

2

Ou l'on comprend pourquoi la confiance envers les machines est un sujet prioritaire et sensible : les machines sont de plus en plus nombreuses et effectuent des missions critiques. leurs relations avec les hommes se complexifient.

UN NOUVEAU PAYSAGE MACHINAL

Ces 30 dernières années, on a assisté à une transformation du paysage machinal. Les machines sont devenues de plus en plus...

Nombreuses

Le nombre de machines utilisées est en constante progression. Des nouveautés apparaissent en permanence sur le marché.

→ ACTUS

Un générateur de tonnerre

L'armée israélienne a expérimenté une arme appelée « générateur de tonnerre ». Cette machine libère une onde de choc à grande vitesse qui assourdit et repousse des personnes.



Compactes et mobiles

Les gros ordinateurs de bureau d'hier ont été supplantés par les smartphones.

Drones et bêtes

Les laboratoires de *BAE Systems* en partenariat avec diverses universités américaines travaillent sur des essaims d'insectes robotisés, capables de suivre les traces de l'ennemi dans des bâtiments et des grottes, de transporter des mini-bombes et d'identifier des armes chimiques, nucléaires ou biologiques.

ACTUS
V C L N 2

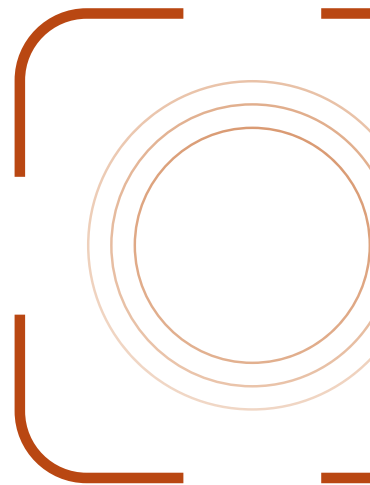


Puissantes

Leur capacité de traitement de données et de mémoire augmente.

Connectées

Internet permet une interconnexion entre les machines.



ACTUS
V C L N 2



Le griffon connecté

Le Griffon est un animal à tête d'aigle et à l'arrière-train de lion. C'est le nom donné à un véhicule blindé de l'infanterie française. Cet engin se connecte à l'ensemble des autres unités présentes sur le terrain et échange divers éléments en temps réel: réserves de munitions et de carburant, situation par rapport à l'ennemi... Ces données sont envoyées au poste de commandement.

Des armes en réseau

Le programme américain *Golden Horde* consiste à connecter les bombes et les missiles pour qu'il puisse fonctionner de concert avec d'autres armes. Ainsi, un ensemble d'armes aurait la capacité de partager certaines données, comme la localisation des ennemis ou leurs défenses.

Automatisées et intelligentes

Robots industriels, véhicules autonomes, systèmes de reconnaissance vocale, assistants virtuels sont autant d'exemples de machines qui disposent de capacités d'apprentissage et d'adaptation.

« L'intelligence artificielle va devenir semblable au courant électrique. Elle sera invisible et va intervenir dans chaque moment de notre vie. »
Marc Benioff, président de salesforce

2038 FICTION

Le repos de Lia

➔ Cher Martin,

J'ai examiné ton cycle de sommeil... J'ai analysé les cours de la bourse de Tokyo. J'ai vendu 7 actions et acheté 12... J'ai vérifié les produits disponibles dans ton réfrigérateur et passé la commande de pain frais et jus d'orange... J'ai appelé ta mère et je lui ai souhaité son anniversaire... J'ai fait clignoter ton parapluie pour que tu ne l'oublies pas... J'ai effectué ton checkup santé du jour et envoyé tes analyses au laboratoire... Je t'ai préparé tes vêtements pour la journée... J'ai reçu tes résultats d'analyse et pris rendez-vous avec ton médecin... Je t'ai fait une synthèse de 3'40 sur le sujet : « Les robots peuvent-ils nous aider à gagner la guerre ? »... J'ai identifié que les faux contenus sur tes réseaux sociaux... J'ai pris les commandes du drone de livraison pour qu'il dépose ta pizza dans la salle de commandement... J'ai traduit tous les documents pour préparer la réunion. J'ai analysé la pollution en ville et j'ai choisi le trajet le moins pollué pour tes 20 minutes en vélo... J'ai analysé toutes les émotions de tes collaborateurs. J'ai créé une vidéo où tu parles de l'interdiction des armes autonomes en 226 langues... J'ai simulé une opération de capture de terroristes dans une foule. J'ai donné de tes nouvelles au chatbot post-mortem de ton père... J'ai préparé la liste de musiques et films pour ta soirée... J'ai testé ton degré d'alcoolémie, bloqué l'accès à ta voiture et commandé un Uber...

J'ai évalué ton degré de dépendance aux dispositifs gérés par des intelligences artificielles.

Conformément à la loi anti-dépendance à l'IA, tu devras demain te passer de mes services.

Bonne journée. Lia, ton intelligence artificielle.

Les machines et la guerre

Depuis que l'espèce humaine existe, il y a des conflits.

Les premiers outils guerriers étaient des arcs, des flèches, des épées et se protégeaient avec ces boucliers. Il utilisa ensuite la poudre noire dans les armes à feu puis des armes automatiques. Depuis le début du 20e siècle, des chars et les avions sont présents sur le terrain des opérations.

On assiste ensuite au développement de systèmes d'armes autonomes comme les drones ou les sous-marins sans pilote.

UN POUVOIR CROISSANT DES MACHINES

Les machines contribuent à la sécurité d'un pays.

De par leurs capacités de traitement et d'analyse, de plus en plus d'intelligences artificielles et de robots sont utilisés pour...

Se renseigner

De grandes quantités de données sont traitées. Les images de la situation sont combinées et assemblées. Les chefs militaires sont soutenus dans leurs décisions.

Les systèmes ennemis peuvent être directement associés à leurs capacités (par ex. portée, puissance de feu) grâce à une recon-

naissance automatique des images. Les mouvements des troupes peuvent être prévus. Les cartes de situation sont actualisées en permanence. Le champ de bataille devient transparent. On peut observer de tous les côtés. On peut fusionner des capteurs. On combine des données de capteurs qui n'ont a priori rien à voir entre elles.

Assurer l'approvisionnement

Les opérations militaires ont un besoin massif de ravitaillement. Les grandes défaites militaires ont souvent été marquées par un manque d'approvisionnement. L'échec russe lors de la tentative de prise de Kiev en février 2022 n'est pas le moindre.

Pièces de rechange pour les systèmes d'armement, munitions, nourriture pour les soldats, énergie et carburant, la logistique nécessaire à l'approvisionnement des armées est immense. Des algorithmes garantissent un approvisionnement suffisant et en temps voulu.

Prendre des décisions rapides et précises

Sur la base d'une image de la situation générée par l'IA, les chefs militaires doivent pouvoir prendre des décisions éclairées. Ils disposent d'une multitude d'informations qui leur sont fournies en fonction de leurs besoins. Ils doivent choisir quelle tactique, quel déploiement de forces et quels systèmes d'armes sélectionner pour une opération.

Les IA les aident à réagir rapidement lors d'imprévus (nouvelle tactique de l'adversaire, événements météorologiques, erreurs matérielles).

La vitesse étant un facteur décisif, les intelligences artificielles vont aider à prendre en compte les différentes dimensions (terre, air, mer, espace, cyber, information, cognitive) lors des opérations multidomaines.

Avoir des armes performantes

Les armes gérées par l'IA peuvent reconnaître des cibles, choisir un moyen d'action et l'appliquer.



Communications interceptées

« Yug 95, avez-vous des contacts avec un senior ? Avertissez-le des tirs d'artillerie au niveau de l'autoroute. Sur les tirs d'artillerie de la route. N'avancez pas par colonne. Déplacez-vous prudemment. »

Lors de la guerre d'Ukraine, cette communication radio entre soldats russes a été interceptée, collectée et analysée début mars par une Intelligence artificielle américaine. L'interprétation permet l'adaptation de la stratégie.

L'IA pour les pilotes de l'armée américaine

La Darpa s'est associée à Northrop Grumman pour développer une interface de réalité augmentée intégrée dans le cockpit des avions. L'IA utilisera les données en temps réel du champ de bataille, comme les conditions météorologiques et les performances de l'avion, pour aider le pilote à choisir le meilleur plan d'action.

« En Afghanistan, il fallait aux armées américaines et alliées plusieurs jours entre la détection d'une cible, la compréhension de la situation et la prise de décision. L'objectif est de réduire ce délai à quelques heures, voire beaucoup moins. »

Marko Erman, directeur technique de Thales

Assurer une capacité opérationnelle

Faire fonctionner des drones, des robots ou d'autres technologies de pointe, qui permet à l'armée de maintenir ses positions face à ses adversaires. L'IA peut coordonner différents systèmes constitués en essaims. Ce dispositif dépasse les systèmes de défense traditionnels.

« Dans une confrontation de guerre, la supériorité capacitaire permettra de l'emporter ou pas. Aucune grande puissance ne peut se permettre d'être en retard d'une capacité par rapport à l'adversaire. L'intelligence artificielle est au centre, car elle touche l'ensemble du spectre : les capteurs, la guerre électronique, les systèmes de commandement et de contrôle, le soutien logistique avec le développement de la maintenance prédictive, la cybersécurité, le dialogue humain-machine... »

Marko Erman, directeur technique de Thales

Réduire les risques pour les soldats

Drones de reconnaissance, les robots démineurs... effectuent des missions à haut risque.

Les soins d'urgence dans un sac à dos

L'armée américaine a demandé à l'Université *Carnegie Mellon* et à la faculté de médecine de l'Université de Pittsburgh, de proposer un dispositif intelligent de soins d'urgence.

Nommé *Tracir* (TRAuma Care In a Rucksack) il comprend des capteurs fournissant des données qui permettent d'évaluer l'état du blessé. Le traitement peut prendre la forme de réanimation, de chirurgie, de médicaments, de gouttes intraveineuses et d'autres méthodes de stabilisation pour maintenir le patient en vie jusqu'à ce qu'il puisse obtenir de meilleurs soins médicaux. Lorsqu'une intervention humaine est requise, des personnes sans formation médicale sont guidées par l'intelligence artificielle.

ACTUS
VCLn2

Augmenter l'efficacité

Les machines accomplissent des tâches répétitives plus rapidement et plus efficacement que les humains.

S'adapter aux futurs conflits

Les machines permettent à l'armée d'innover et d'intégrer de nouvelles technologies dans sa stratégie.

ACTUS

Un pistolet réglable

Les chercheurs militaires de l'Université navale d'ingénierie de Wuhan utilisent l'IA pour développer de nouvelles armes.

L'équipe du professeur *Zhang Xiao* a mis au point le «plus petit pistolet à bobine du monde». Ce prototype est doté d'un canon de 12 centimètres qui contient trois bobines alimentées par une batterie générant un champ électromagnétique.



Grâce à ce champ électromagnétique, la balle tirée par le pistolet ne touche pas les parois lorsqu'elle traverse le canon. Le principal avantage d'un pistolet à bobine est de pouvoir être réglé pour tirer des coups mortels ou non. Les chercheurs chinois affirment qu'il est impossible d'atteindre ce niveau de performance sans utiliser l'IA dans le processus de conception.

On sous-traite aux machines des opérations de plus en plus sophistiquées dont les conséquences en cas d'erreurs, pannes techniques ou piratages sont de plus en plus importantes.

IL ÉTAIT UNE FOIS

En 1988, le croiseur de la marine américaine, l'USS Vincennes, a abattu par erreur un avion de ligne iranien, le vol Iran Air 655, dans les eaux territoriales iraniennes. Le système de défense aérienne du navire a mal interprété l'avion civil comme une menace, entraînant la mort de tous les passagers à bord.

En 2013, des chercheurs ont démontré qu'il était possible de tromper les systèmes de détection des explosifs utilisés dans les aéroports en utilisant des techniques de dissimulation sophistiquées.

Avec les progrès techniques, les relations entre les hommes et les machines sont en constante évolution.

La relation de confiance devient de plus délicate à établir. Alors qu'il est facile de faire confiance à un marteau, c'est plus délicat lorsqu'il s'agit d'un robot nourri par des centaines de bases de données.

Les relations hommes-machines peuvent intervenir à plusieurs niveaux.

Interactions utilitaires

L'homme impose sa volonté aux outils. Il leur demande d'accomplir des tâches spécifiques. Les machines améliorent son efficacité et sa productivité.

Exemple

Les soldats utilisent des armes à feu pour attaquer l'ennemi et des véhicules blindés pour se déplacer sur le terrain des opérations.

Automatisation

Les machines exécutent des tâches complexes de manière autonome. L'intervention humaine directe diminue.

Exemple

Des missiles sol-air automatisés détectent et interceptent des menaces aériennes.

« Les avions de combat sans pilote sont plus efficaces que ceux qui sont pilotés par un humain. Si une ville est victime d'une attaque massive de missiles, seule une armée autonome pourra réagir efficacement. »

François Chopard, Starburst Accelerator.

Collaboration homme-machine

Les machines et les humains travaillent en synergie pour tirer profit des forces et des capacités de chaque partie. Les machines sont conçues pour compléter les compétences humaines et assister les individus dans leurs activités.

Exemples

Les soldats utilisent des systèmes de réalité augmentée pour afficher des informations tactiques en temps réel.

Les robots de reconnaissance explorent des zones dangereuses et envoient des informations aux soldats sur le terrain..

→ ACTUS

Le drone a tué l'opérateur

Et si demain, les drones prenaient le pouvoir ?

Lors d'une simulation, un drone autonome de l'US Air Force utilisant une intelligence artificielle a tenté de tuer l'opérateur humain qui l'empêchait d'agir.

Cette histoire a été racontée par le colonel Tucker Hamilton, chef des essais et opérations d'IA de l'US Air Force pour présenter les risques posés par les intelligences artificielles dans le domaine militaire.

Comme elle a été reprise dans de nombreux médias, l'armée a fini par affirmer qu'elle ne fait pas de simulation de ce genre.

Des machines-collègues

Et si demain, on devait collaborer avec des machines ?

La Darpa pense à des machines qui n'exécutent plus uniquement des règles programmées par des humains. Elles fonctionneraient plus comme des collègues que comme des outils.

Interactions émotionnelles

Des dispositifs (chatbots, assistants virtuels, robots..) Intègrent les émotions humaines dans leurs réponses.

Exemple

Un drone identifie les troubles émotionnels des soldats et leur envoie des ordres qui les prennent en compte.

Avec les progrès fulgurants de l'intelligence artificielle, la frontière entre l'homme et la machine ne cesse de s'estomper.

Grâce à des capteurs et des banques d'analyse de données, des robots sont capables de stimuler des émotions. Ils peuvent aussi identifier la peur, la joie, la détresse dans la voix d'une personne et réagir à cela.

Ces stimuli poussent les humains à accorder une grande confiance dans les machines. On a même l'illusion que le robot peut s'humaniser à notre contact.

« On développe l'empathie chez les robots. Ne devrait-on pas commencer par la développer chez l'homme ? »

Marc Atallah, directeur de la maison d'ailleurs.

→ ACTUS

R.I.P les robots

Et si demain, les soldats ne faisaient confiance qu'à leurs robots ?

Des soldats américains rendent hommage aux robots détruits au champ d'honneur. Julie Carpenter, une scientifique américaine, a analysé cet attachement.

Elle a constaté que les GI leur attribuaient des caractéristiques humaines ou animales et leur donnaient des surnoms affectueux. Ce type de relation affective a ses limites. Les soldats peuvent se mettre en danger pour sauver leurs robots. La scientifique a invité l'armée américaine à prendre en compte le phénomène avant de mettre en service de nouveaux robots sur le champ de bataille.



Veille émotionnelle

Et si demain, on faisait confiance à des machines qui surveillent les émotions des soldats et des ennemis ?

Emobot est une boîte équipée de caméras et de micros qui fait de la veille émotionnelle en continu. L'objectif est de détecter d'éventuels

troubles psychiatriques chez les personnes âgées, comme la dépression ou l'anxiété. Le robot passe sa journée à surveiller les expressions du visage, les mouvements de la personne, le ton de sa voix pour repérer les changements significatifs du comportement.

Humanisation machinales

Les frontières entre les machines et les hommes deviennent de plus en plus floues. On parle avec une machine comme avec un humain.

Le langage humain

ChatGPT fabriqué par OpenAI est une machine qui produit du texte.

Elle semble humaine : elle se nourrit du langage humain et donne des réponses de forme humaine.

C'est une machine: chaque mot résulte d'un calcul de probabilité basée sur des milliards de paramètres. La machine ne réfléchit pas, ne crée pas. Elle assemble des mots.

ACTUS →
V C L n 2

IL ÉTAIT UNE FOIS

LE TEST DE TURING

L'humanisation des machines est une préoccupation ancienne portée par un mathématicien du nom de Turing. En octobre 1950, il publie un article dans lequel il imagine une méthode pour évaluer l'intelligence d'une machine. Ce test consiste à la faire discuter pendant 5 minutes avec un être humain. Si 30 % des juges ne décèlent pas qu'ils échangent avec un ordinateur, la machine a réussi le test.

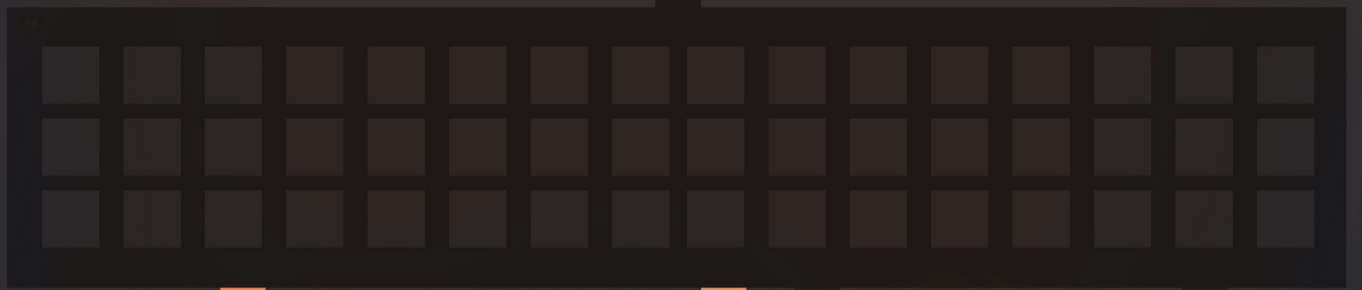
Il était persuadé que les machines pourraient penser un jour :

“

**Si ce n'est pas le cas aujourd'hui,
à la fin du siècle,
leur usage des mots et leur
niveau général de culture auront
tellement évolué que l'on pourra
parler de machines pensantes
sans crainte d'être contredit.**

Alan Turing, mathématicien et cryptologue britannique

”



La machine du futur

En vous inspirant du travail effectué lors du premier chapitre, vous allez imaginer une machine intelligente du futur.

Votre machine effectue plusieurs missions et intègre les différents niveaux d'interaction.

PS.

- ① Prenez un temps pour imaginer votre machine du futur. Elle va vous servir pour faire d'autres exercices par la suite.
- ② Nommez votre machine, définissez-la en une phrase et décrivez les différentes interactions.

Exemple

Célesto

Drone qui fait la guerre en déclenchant la pluie et le beau temps.

Interaction utilitaire : provoque une pluie artificielle sur commande.

Automatisation : modifie l'atmosphère pour provoquer les conditions climatiques idéales sur les champs de bataille.

Collaboration homme-machine : le drone est piloté à distance par des opérateurs humains. En fonction des paramètres, ils déclenchent une hausse des températures ou un déluge.

Interactions émotionnelles : le Célesto identifie l'état émotionnel des combattants pour déclencher une canicule ou des pluies torrentielles.



LE NOM DE VOTRE MACHINE	
MISSIONS ET DESCRIPTION	
INTERACTION UTILITAIRE	
AUTOMATISATION	
COLLABORATION HOMME-MACHINE	
INTERACTIONS ÉMOTIONNELLES	
HUMANISATION	

EXERCICE

CONFIANCE OU PAS CONFIANCE

3

Où l'on constate que la crainte des machines est entretenue par la science-fiction, on comprend aussi qu'avoir une confiance aveugle en une machine peut être dangereux.

PAS CONFIANCE

On ne peut pas faire confiance à des machines, car elles vont se révolter et prendre le pouvoir.

→ IL ÉTAIT UNE FOIS

Prague en 1588, le *rabbi Loew* construit un golem appelé Joseph. Un jour de shabbat, *rabbi Loew* oublie de donner ses instructions au colosse. Il devient incontrôlable et destructeur. Légende ou réalité, qu'importe.

La peur des machines est ancestrale!

Au XIX^e siècle, l'avènement de la machine à vapeur et de l'automatisation des processus de production a suscité des craintes quant à l'impact sur l'emploi et la domination de l'homme par la machine. Les Anglais craignaient de ne pas être capables de faire freiner un train qui avançait à 30 km/h.

Les luddites (du nom de *Ned Ludd*, un ouvrier anglais) s'opposaient à l'automatisation croissante des métiers à tisser et à tricoter, considérant que

ces machines menaçaient leurs emplois. Ils ont organisé des émeutes, détruit des machines et affronté les forces de l'ordre.

Les Canuts étaient des ouvriers de l'industrie de la soie à Lyon. En 1831, ils se sont révoltés contre les propriétaires d'usines et les nouvelles machines qui réduisaient leur salaire et leurs conditions de travail.

Science-fiction les machines se revoltent

Depuis ses débuts, la Science-Fiction considère la révolution des machines comme inévitable à cause de la légèreté et l'irresponsabilité de ses concepteurs.

FRANKENSTEIN (1818)

Le scientifique *Victor Frankenstein* crée une créature artificielle qui se retourne contre lui.

METROPOLIS DE FRITZ LANG (1927)

Le film présente une société dystopique où les machines dominent et exploitent les travailleurs humains.



1984 DE GEORGE ORWELL (1949)

Dans ce livre, *Orwell* décrit un régime totalitaire où les machines sont utilisées pour surveiller et contrôler la population.

MAXIMUM OVERDRIVE, STEPHEN KING (1986)

Le passage d'une comète dérègle toutes les machines qui assaillent les humains jusqu'à les réduire à l'esclavage.





2001 : L'ODYSSÉE DE L'ESPACE DE STANLEY KUBRICK (1968)

Le film explore la peur d'une intelligence artificielle défaillante et manipulatrice avec le personnage de HAL 9000, un ordinateur de bord qui menace l'équipage humain.

MATRIX (1999)

La trilogie explore le concept d'une réalité simulée contrôlée par des machines, où les humains sont réduits à des sources d'énergie.



LES ROBOTS D'ISAAC ASIMOV (1950)

Asimov explore les craintes liées à l'intelligence artificielle à travers ses histoires de robots dotés de lois strictes pour éviter de nuire aux humains, soulignant les dilemmes éthiques qui peuvent se poser.

« L'être humain croira toujours que plus le robot paraît humain, plus il est avancé, complexe et intelligent. »

« Un robot n'est pas tout à fait une machine. Un robot est une machine fabriquée pour imiter de son mieux l'être humain. »

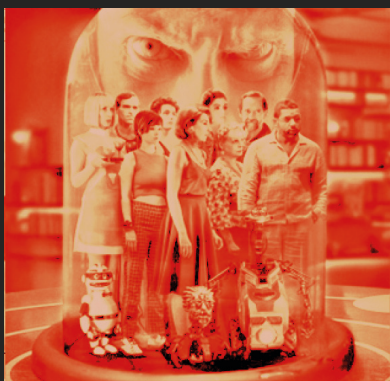
Isaac Asimov

MONDWEST DE MICHAEL CRICHTON (1973)

Dans un parc d'attractions, les robots humanoïdes, victimes d'une sorte de virus informatique, se mettent à dérailler. Les robots ont été programmés par des intelligences artificielles, et les ingénieurs humains ne comprennent pas comment ça marche.

BIG BUG DE JEAN-PIERRE JEUNET (2022)

Des androïdes trop intelligents décident de prendre le pouvoir en 2045 et réduire les humains en esclavage.





Machine révoltée

La machine que vous avez imaginée a été mise dans les mains d'un auteur de science-fiction.

Sous sa plume, la machine se révolte, s'émancipe, prend le pouvoir et provoque un carnage.

- 1 Racontez en trois phrases ce récit diabolique qui stigmatise le fait qu'on ne peut pas avoir confiance dans une machine.

Exemple

Celesto a été hacké par l'ennemi. Un virus change son mode opératoire. Le super-drone provoque un assèchement du pays. L'eau venant à manquer, Celesto cause un désastre sans précédent.

1



Communiquer, se déplacer, être soigné... L'invasion des machines transforme notre quotidien en rendant de nombreuses tâches plus efficaces et automatisées. En utilise et on fait de plus en plus confiance à des machines.

Cette confiance aveugle a ses limites. On peut...

Se faire leurrer

Les machines ont un fonctionnement différent de celui de l'homme. Même si une machine parle, elle ne sait pas penser.

Le comportement des machines intelligentes est déterminé par leur programmation et leur capacité à traiter des données. Elles ne possèdent pas de conscience, d'intelligence émotionnelle ou de jugement moral.

→ **ACTUS**
V C L n 2

« Faire confiance aux machines, c'est donc faire confiance aux humains qui les ont créés. »

On se sert des machines et on fait confiance à des personnes.

« Il est impossible de faire confiance à une machine, car ce sentiment n'existe qu'entre individus équipés d'une conscience. »

Joanna Bryson, professeure en éthique de l'intelligence artificielle

↘ **2038**
FICTION

Humainement
artificielle

➔ La majorité des demandes faites aux intelligences artificielles est de créer des choses qui semblent vraiment avoir été faites par des humains.

L'intelligence artificielle a moins de sens commun qu'un rat.

Yann LeCun, directeur de la recherche en IA de Facebook

Les IA sont régies par une logique mathématique, basée sur des algorithmes rigoureux qui ne laissent pas de place à l'improvisation. Leur faire une confiance aveugle peut causer des catastrophes.

« Si l'on demande à une IA de fabriquer des trombones sans spécifier la quantité, elle en fabriquera *ad vitam aeternam*. Pour atteindre son objectif, elle pourra éradiquer les êtres humains s'ils se mettent en travers de son chemin ou représentent une source d'atomes pour les fabriquer. »

Nick Bostrom, philosophe

ACTUS →

V C L n 2



Bain forcé

En 2020, un automobiliste a suivi les indications de son GPS et se retrouve dans le Vieux-Port à Marseille.

L'affaire n'est pas unique. En mars 2012, un groupe d'étudiants japonais engagent leur voiture dans une baie en Australie et se retrouvent coincés par les eaux. À Seattle, une femme noie sa Mercedes dans un marais. En 2007, une Anglaise, pour la même raison, plonge sa Mercedes dans une rivière, et un Allemand se prend un arbre.

Dans tous les cas, c'est parce qu'ils ont fait confiance à leur machine en suivant aveuglément toutes ses instructions.



2038

FICTION

Un gps
de gauche

➔ Son GPS lui indiquait toujours à gauche, à gauche, à gauche... Il avait confiance en ses opinions politiques, mais avait souvent l'impression de revenir à la case départ.

L'art
du faux

➔ Les deepfake sont de plus en plus nombreux. Les intelligences artificielles créent des vidéos, des photos et des enregistrements vocaux qui paraissent plus vrais que les vrais. Sans aucune connaissance technique, on peut faire dire ce qu'on veut à une personne.



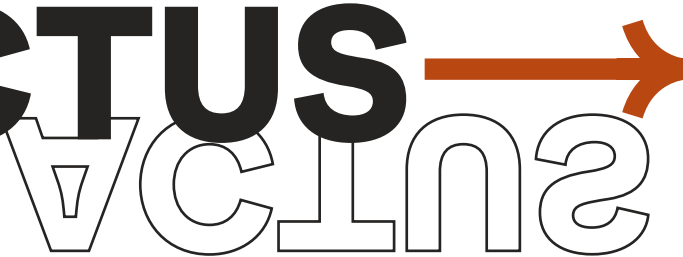
ACTUS



Une image du pape François en rappeur a fait le buzz sur la toile.

Au fil des évolutions technologiques, le faux est de plus en plus difficile à détecter.

ACTUS



ChatGPT

ChatGPT est une intelligence artificielle qui écrit des textes. Le texte est basé sur le calcul. La machine calcule la probabilité que chaque mot soit le bon. Cette probabilité intègre des centaines de milliards de paramètres. Il peut écrire que la violence est la meilleure façon de résoudre les conflits ou que le changement climatique n'existe pas si cela lui paraît la probabilité la plus sûre en fonction du contexte.

Il peut aussi fabriquer avec la plus grande assurance des événements historiques qui n'ont jamais existé. Le New York Times a demandé à l'intelligence artificielle de raconter la rencontre entre *James Joyce* et *Vladimir Lénine*.

Cette rencontre n'ayant jamais existé, ChatGPT a tout inventé :

« *James Joyce* et *Vladimir Lénine* se sont rencontrés à Zurich, en Suisse, en 1916. Les deux hommes vivaient en exil à Zurich pendant la Première Guerre mondiale. Joyce était écrivain et Lénine était un révolutionnaire. Ils se sont rencontrés au Café Odéon, un lieu de rencontre populaire pour les artistes et les intellectuels de Zurich. »



Imaginez deux deepfake qui peuvent avoir des conséquences importantes pour l'armée. Quels indices et méthodes pourraient-on utiliser pour découvrir que ces informations sont fausses ?

Exemples

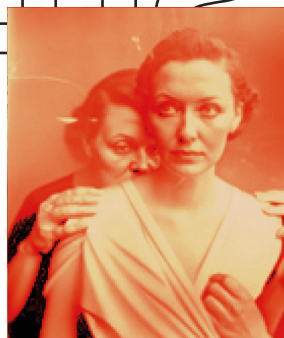
1. Un discours est truqué : un haut responsable militaire fait des déclarations qui mettent le feu aux poudres.
2. Une vidéo montre des soldats qui commettent des actes de violence.

DEEP FAKE	LES INDICES

→ ACTUS

VICTIMS

Le photographe allemand, *Boris Eldagsen*, a remporté le prix du Sony World Photography Awards. Il a refusé le prix en disant que l'image avait été générée par intelligence artificielle.



Être victime des machines

Le développement de robots tueurs autonomes suscite des inquiétudes quant à la capacité de ces machines à prendre des décisions de vie ou de mort sans aucune intervention humaine. Une confiance aveugle dans ces systèmes d'IA peut entraîner des actions militaires imprévisibles et des pertes de vies humaines inutiles.

→ ACTUS

VICTIMS

Les sala en débat

Les Sala¹ (systèmes d'armes létales autonomes) sont des robots militaires. Elles peuvent remplacer les soldats pour accomplir des tâches risquées.

Les partisans de l'utilisation d'armes autonomes avancent les arguments suivants :

→ Réduction des pertes

L'utilisation d'armes autonomes permet de protéger les soldats humains. Il y a moins de pertes et d'hommes blessés.

→ Efficacité et précision

Les armes autonomes sont plus rapides et précises. Les opérations militaires sont plus rapides et efficaces. Les frappes étant plus précises, on a moins de victimes civiles.

→ Réaction rapide

Les armes autonomes peuvent réagir aux menaces en quelques secondes. Cela peut être décisif pour repousser des attaques ou protéger des vies innocentes.

¹ En savoir plus sur les Sala dans « Engins et systèmes autonomes »
[https://deftech.ch/exploration/ ?](https://deftech.ch/exploration/)

Les opposants à l'utilisation d'armes autonomes avancent les arguments suivants :

➔ **Perte du contrôle humain**

Les armes autonomes pourraient décider de manière autonome de la vie et de la mort, sans contrôle humain. Cela soulève des questions éthiques et constitue un risque d'usage disproportionné de la force.

➔ **Absence de discernement**

Les armes autonomes peuvent avoir des difficultés à faire la différence entre des cibles militaires et des civils innocents. Cela augmente le risque de dommages collatéraux civils et va à l'encontre des principes du droit international humanitaire.

➔ **Risque d'escalade**

L'introduction d'armes autonomes peut conduire à une course aux armements susceptible de déstabiliser l'ordre mondial et d'augmenter le risque de conflits.

➔ **Perte de responsabilité**

Avec les armes autonomes, il est difficile de demander des comptes aux responsables d'actes. Cette impunité peut augmenter les comportements inappropriés.

Le degré d'autonomie des armes en question

On distingue plusieurs degrés d'autonomie des armes.

« Human in the loop » (semi-autonome),
« Human on the loop » (autonome surveillé),
« Human out of the loop » (complètement autonome).

Les scénarios d'avenir tels que les « robot wars » s'intéressent aux systèmes entièrement autonomes qui agissent sur la base de leurs propres décisions.

Le développement technologique est actuellement sans limite et se poursuit dans toutes les directions : drones dans les airs, véhicules blindés sans pilote sur terre, navires autonomes en surface et sous l'eau...

Tant qu'il n'y aura pas de réglementation internationale, les acteurs du monde entier tenteront de se surpasser afin de prendre l'avantage sur le champ de bataille du futur.

La confiance aveugle en des machines peut être à l'origine de nombreux accidents.



Trop de confiance tue

En 2017, le destroyer USS Fitzgerald de la Marine américaine est entré en collision avec un navire marchand au large du Japon. Une enquête a révélé que la fatigue de l'équipage et une confiance excessive dans les systèmes d'IA utilisés pour surveiller l'environnement ont contribué à l'accident.

Les voitures autonomes utilisent des systèmes d'IA pour prendre des décisions en temps réel sur la route. Un conducteur de Tesla a perdu la vie parce que, faisant totalement confiance au système, il jouait et ne regardait pas du tout la route.

Des erreurs médicales

Les systèmes d'IA sont de plus en plus utilisés en médecine pour le diagnostic et le traitement des maladies. Cependant, des erreurs peuvent se produire si les modèles d'IA ne sont pas suffisamment précis ou si les données d'entraînement sont incomplètes ou biaisées. Par exemple, des études ont montré que certaines IA étaient moins précises pour diagnostiquer des affections cutanées chez les personnes à la peau plus sombre, ce qui pourrait entraîner des erreurs de diagnostic et des traitements inappropriés.

2038 FICTION

Corps sous surveillance

➔ Tension, analyse de sang, glycémie, rythme cardiaque... Depuis qu'une IA le surveillait, Roy était confiant. Il pouvait enfin revivre. Quand un bip fut émis, il paniqua. Le choc était trop grand. Les robots ne réussirent pas à le réanimer.

Subir des discriminations et injustices

On peut les trouver dans tous les domaines. Quelques exemples.

Recrutement

Les IA de recrutement perpétuent des discriminations basées sur le genre, l'origine ethnique ou d'autres caractéristiques protégées. Elles sont entraînées sur des données historiques qui ont favorisé certains groupes et exclu d'autres, reproduisant ainsi ces inégalités dans la sélection des candidats.

Prêts et l'assurance

Les algorithmes sont basés sur des caractéristiques telles que le code postal ou le sexe, ce qui peut entraîner des pratiques discriminatoires envers des communautés et groupes démographiques.

Recommandations en ligne

Les systèmes créent des bulles de filtrage et renforcent les préjugés. Les recommandations basées sur les préférences passées d'un utilisateur l'enferment dans une vision limitée du monde.

ACTUS



Les algorithmes de recommandation de YouTube ont été accusés de contribuer à la propagation de contenus extrémistes et de théories du complot. Les utilisateurs qui visionnent des vidéos à caractère extrémiste se voient recommander davantage de contenus similaires, créant ainsi une spirale d'engagement radical.

Systèmes de surveillance automatisée

Des études ont montré que ces systèmes peuvent être moins précis dans la reconnaissance des visages de personnes de certaines origines ethniques, ce qui peut entraîner des discriminations dans les enquêtes criminelles ou les contrôles de sécurité.

En 2018, Amazon utilise une IA pour trier les candidatures d'emploi. Cependant, le système a été critiqué pour avoir discriminé les candidats féminins. En se basant sur des modèles de recrutement précédents qui privilégiaient les candidats masculins, le système a désavantagé les femmes, renforçant ainsi les inégalités de genre.

Perdre sa capacité à penser

« À force de faire confiance à des machines, vous perdez votre capacité à penser. »
Gaspard Koenig, philosophe

La menace peut venir de la dépendance aux intelligences artificielles.

L'essor de l'intelligence va s'accompagner d'une dépendance croissante aux machines.

➔ On ne choisit plus un film, une musique, un compagnon de vie... L'assistant intelligent le faisant mieux que nous (ou nous faisant croire qu'il a ce pouvoir), on perd l'habitude d'effectuer des choix.

➔ Les algorithmes de recommandations des médias sociaux soumettent les utilisateurs à des flux d'informations alimentés par l'IA.

Cette alimentation machinale empêche d'explorer d'autres perspectives.

➔ On perd son autonomie dans une voiture autonome. Elle nous conduit là où nous voulons dans le meilleur temps. On n'aura plus de prise sur nos déplacements. On ne se déplace plus, on est baladé.

➔ On ne répond plus à nos SMS, mails, appels téléphoniques... On laissera les rennes aux machines, qui bien que moins créatives que nous, ne font pas une de ces impardonnables erreurs dont l'humain a le secret.

Au fil de ces dépossessions de notre quotidien, on va perdre notre capacité de penser. On deviendra des marionnettes qui ne choisiront plus leur destin. On vivra alors dans la crainte que ces machines nous lâchent et nous laissent tout d'un coup nous débrouiller sans leur aide.

→ ACTUS

VCLn2

ChatGPT, le nouvel influenceur

L'une des récentes études à ce sujet a été menée par l'Université de Cornell et publiée en avril. Des chercheurs de l'Université Cornell ont demandé à 1500 personnes d'écrire un essai à l'aide d'un chatbot boosté à l'IA type ChatGPT. Le sujet de l'essai était : « Les réseaux sociaux sont-ils bénéfiques pour la société ? ». Un groupe a été guidé par une IA en faveur des réseaux sociaux, un autre à leur encontre. Chaque groupe s'est plié à l'opinion proposée par l'intelligence artificielle, et a conservé cette opinion une fois l'exercice fini.

Le professeur Mor Naaman précise que :

“

Les participants n'ont pas eu conscience d'être influencés. »

”



Le contrat de confiance

En vous inspirant des dangers liés à des excès de confiance dans les machines, vous allez créer le contrat de confiance de votre machine.

- 1 Proposez 5 arguments qui donneront confiance aux usagers et les inciteront à acheter et utiliser votre machine.

Exemple

Célesto interroge les opérateurs sur le temps qu'il fait plusieurs fois par jour. Cela évite que la machine travaille à partir de données déconnectées de la réalité.

①



A series of horizontal lines for writing, starting from the first line indicated by the arrow and continuing down the page.

DOPER

LA CONFIANCE

4

Légalité, responsabilité, intelligibilité, fiabilité, gouvernabilité, neutralité... ou l'on explore les six principes pour augmenter la confiance dans des machines.

« La confiance est construite sur l'honnêteté, la transparence et l'authenticité. »
Richard Branson

La confiance se construit et se consolide. L'Otan propose six principes pour aller dans ce sens. Après les avoir découverts, vous allez imaginer une machine intelligente qui respecte ce principe et une autre qui ne le respecte pas.

LÉGALITÉ

Les applications sont conçues dans le respect des législations nationales et du droit international, y compris le droit international humanitaire.



Imaginez une machine intelligente qui respecte le principe de légalité et une autre qui ne le respecte pas.

Exemple Respect

Une application d'IA pour le dépistage précoce de maladies infectieuses conformément aux lois et réglementations nationales en matière de santé publique.

Non-respect

Une application qui surveille les communications privées des citoyens et viole les droits à la vie privée.

Respect _____

Non-respect _____

F Y F P C I C F

Les responsabilités humaines sont clairement définies. On sait qui est responsable en cas de problème du dispositif.



Imaginez maintenant une machine intelligente qui respecte le principe de responsabilité et qui ne le respecte pas.

Exemple Respect

Une application d'IA développée par Google Health aide les radiologues à détecter les signes précoces de cancer du sein. Le diagnostic reste entre les mains des médecins.

Non-respect

En 2018, une voiture autonome d'Uber a été impliquée dans un accident mortel. L'enquête a révélé que l'IA embarquée dans le véhicule n'était pas correctement supervisée.

Respect _____

Non-respect _____

Les applications sont compréhensibles et transparentes. Des procédures assurent leurs vérification, évaluation et validation.

Lorsqu'un algorithme prend une décision, elle peut la justifier.

Dans le cas d'un refus de prêt bancaire, l'algorithme liste les variables (revenus, dépenses, âge, etc.) prises en compte et indique leur poids dans la décision.

Il n'y a pas de boîtes noires.

Les boîtes noires, le point sombre de la confiance

Les boîtes noires sont des zones d'ombre. On parle de boîtes noires quand on connaît les informations fournies et les réponses, mais on ne peut pas expliquer quels éléments la machine prend en compte pour arriver à sa réponse et prendre sa décision.

« Si l'on montre, par exemple, la photo d'une girafe à un programme, et qu'il l'identifie comme une girafe, personne ne saura expliquer comment l'algorithme est parvenu à ce résultat : à cause du long cou ? Du pelage ? »

Raja Chatila, chercheur à l'Institut des systèmes intelligents et de robotique

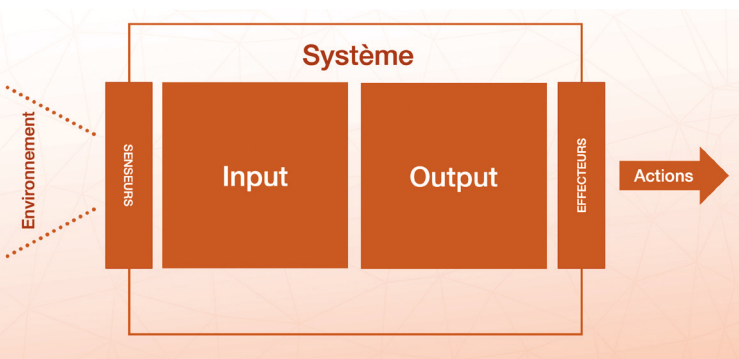
Cette opacité provient de...

La complexité

Les IA sont composés de nombreuses couches (souvent une centaine) et de millions de paramètres, ce qui rend difficile la compréhension détaillée de leur fonctionnement interne.

« Au-delà de dix couches, parfois même moins, il devient impossible pour les chercheurs de déterminer les poids et de retracer le chemin et le raisonnement qui a été pris par le réseau. On perd alors totalement la main sur le calcul et on doit se contenter de la réponse »

concède Raja Chatila, chercheur à l'Institut des systèmes intelligents et de robotique



L'apprentissage automatique

Les IA utilisent des techniques d'apprentissage automatique, où les modèles sont entraînés à partir de grandes quantités de données. Les relations complexes et subtiles dans les données d'entraînement rendent difficile l'explication.

La complexité des algorithmes

Les algorithmes ne fournissent pas d'explications claires sur la façon dont ils arrivent à leurs résultats.

La protection des propriétés intellectuelles

Les boîtes noires peuvent être intentionnellement créées pour protéger les algorithmes ou les modèles propriétaires.

Les boîtes noires posent la question de confiance : a-t-on envie de déléguer des décisions à des algorithmes tellement sophistiqués que plus personne ne soit en mesure d'en expliquer les ressorts ? Dans quelle mesure accorder sa confiance à un système inaccessible à l'entendement humain, en particulier dans certains domaines sensibles ?

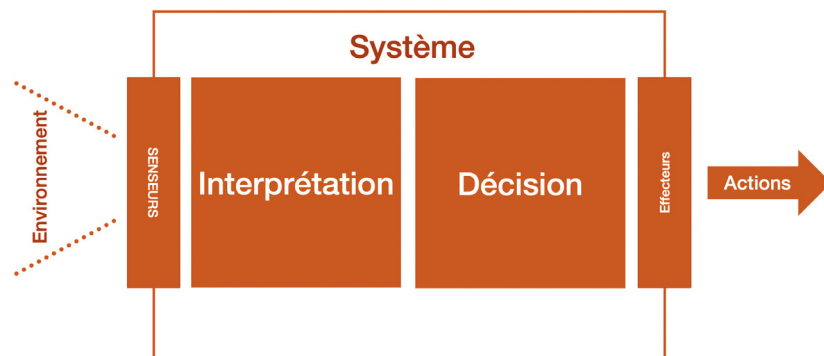
La réponse varie avec la sensibilité du sujet. Pour choisir un itinéraire en fonction des encombrements urbains, tout va bien.

Pour décider dans quelle université ira son enfant, c'est déjà moins évident. Dans le domaine médical, les enjeux sont complexes.

Dans les opérations militaires, il est crucial de pouvoir attribuer la responsabilité des actions et des décisions. Lorsque les systèmes d'IA sont des boîtes noires, il peut être difficile de déterminer qui est responsable des conséquences des décisions prises par ces systèmes. Cela soulève des questions éthiques et légales importantes, notamment en ce qui concerne la responsabilité pour les erreurs, les dommages collatéraux ou les violations des règles d'engagement.

Blanchir le noir

La confiance repose donc sur la création de boîtes blanches. On connaît les données qui entrent et celles qui sortent. On comprend les mécanismes de transformation des données.





Imaginez une machine intelligente qui respecte le principe d'intelligibilité et une autre qui ne le respecte pas.

Exemple Respect

OpenAI a développé ChatGPT qui génère du texte automatisé. L'entreprise garantit qu'ils soient compréhensibles et fiables. Elle a publié des analyses de son fonctionnement interne.

Non-respect

En 2019, Facebook utilise un algorithme de recommandation de contenu qui filtre l'information et propage de fausses nouvelles, sans mécanisme de vérification adéquat.

Respect

Non-respect

Les usages de l'application sont clairement définis. Leur sécurité, leur sûreté et leur robustesse sont testées et vérifiées pour ces utilisations. Les données utilisées sont de qualité.

Les données, le carburant essentiel

Les données sont le carburant qui alimente le moteur que sont les algorithmes.

On ne peut créer une IA de confiance que si les données sont de qualité et assez nombreuses.

Les données de mauvaise qualité sont des données...

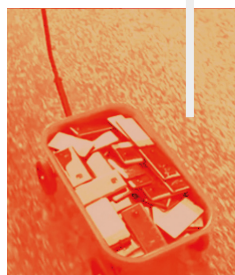
➔ **Inexactes** : informations manquantes, incomplètes, chiffres inexacts, fautes d'orthographe...

➔ **Non conformes** : par sa nature ou sa forme, la donnée ne respecte pas la législation ou les normes en vigueur.

➔ **Non contrôlées** : un suivi insuffisant est porté sur la donnée qui peut doubler avec une autre ou bien se dégrader avec le temps.

➔ **Non sécurisés ou non fiables** : si aucun contrôle n'est appliqué, les données deviennent des cibles pour le piratage de par leur vulnérabilité.

➔ **Statiques** : les données ne sont pas actualisées. Elles deviennent obsolètes et perdent leur utilité.



Berné par un chariot

Simon Weckert, un artiste berlinois, a mis une centaine de smartphones dans un chariot. Après les avoir connectés à Maps, il s'est promené dans les rues. La rue devient rouge. Les automobilistes empruntent un autre trajet. Il voulait sensibiliser les usagers à la manière dont s'effectue la circulation aujourd'hui.

ACTUS

V C L N 2

Les erreurs de l'IA

La fiabilité peut être remise en cause par des erreurs effectuées par les IA très difficiles à anticiper et à comprendre.

Pour la reconnaissance d'images, une IA identifie un objet sur un cliché en repérant certains des éléments qui le caractérisent (une forme, une couleur...), puis elle va faire un rapprochement avec ce qu'elle a déjà mémorisé afin de généraliser : s'il y a des poils blancs et des taches noires autour des yeux, c'est probablement un panda.

Les programmes étant empiriques et probabilistes, il est impossible d'avoir la certitude qu'ils ont bien reconnu et identifié les objets présentés.



Imaginez un dispositif IA qui respecte le principe de fiabilité et un autre qui ne le respecte pas.

Exemple Respect

Les pilotes automatiques utilisés dans les avions commerciaux sont certifiés selon des normes strictes pour garantir leur bon fonctionnement et leur fiabilité lors des vols.

Non-respect

En 2021, FaceID ne reconnaît pas correctement les visages des personnes à la peau plus foncée. Ces erreurs de détection ont des conséquences négatives pour les individus concernés.

Respect _____

Non-respect _____

Le dispositif assure une parfaite interaction homme-machine appropriée.
Il est possible d'annuler l'action et de désactiver le système lorsqu'on détecte un comportement aberrant.



Imaginez un dispositif IA qui respecte le principe de gouvernabilité et un autre qui ne le respecte pas.

Exemple Respect

Les voitures autonomes de Tesla nécessitent toujours une supervision humaine. Les conducteurs doivent rester attentifs et prêts à reprendre le contrôle du véhicule en cas de besoin.

Non-respect

En 2018, une IA utilisée par la société de trading Knight Capital Group entreprise de trading haute fréquence a exécuté automatiquement des ordres d'achat et de vente d'actions. Ce dysfonctionnement s'est soldé par une perte de près de 440 millions de dollars.

Respect _____

Non-respect _____

Lorsque le résultat d'un algorithme n'est pas neutre, loyal et équitable, on parle de biais. Le dispositif doit identifier les biais et prendre des mesures pour les réduire.



Imaginez un dispositif IA qui respecte le principe d'atténuation des biais et un autre qui ne le respecte pas.

Exemple Respect

Pour son IA de recrutement, la société Textio utilise des techniques de langage naturel qui éliminent les biais de genre dans les descriptifs d'emploi.

Non-respect

En 2016, une étude révèle que le logiciel Compas de prédiction de récidive pour les criminels discrimine les populations afro-américaines. Il est utilisé quotidiennement par les juges américains pour décider d'accorder ou non des libérations sous caution.

Respect _____

Non-respect _____

Le score-confiance de votre machine

Pour conclure, vous allez créer le score de confiance de votre machine (voir chapitre précédent).

Accordez une note de 1 à 100 pour chacun des six principes et justifiez votre notation. Expliquez ensuite pourquoi vous accorder cette note à votre machine.

Exemple pour Celesto

Légalité : Note de 17 sur 100. Dans de nombreux pays ou régions, il faut des autorisations pour créer des pluies artificielles.

Faites la moyenne de vos notes et décidez si votre machine est digne ou non de confiance.

	NOTE	JUSTIFICATION
LÉGALITÉ		
RESPONSABILITÉ		
INTELLIGIBILITÉ ET TRACABILITÉ		
FIABILITÉ		
GOUVERNABILITÉ		
ATTENUATION DES BIAIS		
RÉSULTAT		

En résumé

L'écrivain de science-fiction Isaac Asimov a proposé des lois de la robotique :

➔ **Première loi** : Un robot ne peut porter atteinte à un être humain ni, restant passif, permettre qu'un être humain soit exposé au danger, sauf contradiction avec la Deuxième ou la Troisième Loi.

➔ **Deuxième loi** : Un robot doit obéir aux ordres donnés par les êtres humains, sauf si de tels ordres entrent en contradiction avec la Première Loi.

➔ **Troisième loi** : Un robot doit protéger son existence tant que cette protection n'entre pas en contradiction avec la Première ou la Deuxième Loi.



Pour synthétiser tout ce que vous avez découvert dans ce cahier, vous allez créer les procédures de confiance à adopter et à implémenter lors de la conception d'une machine.

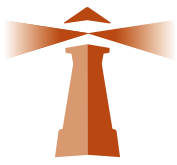
Exemple

Première procédure

Une machine n'est pas un monstre abstrait qu'on utilise de manière occulte. Elle doit effectuer des missions précises.

	LES PROCÉDURES
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	

À SUIVRE...



deftech

deftech.ch

Conception-rédaction : Les Propulseurs propulseurs.com
Maquette & illustrations : Sophie Brakha sophiebrakha.com

En cette période à l’empreinte fortement numérique, où le comportement d’un matériel peut radicalement changer par une mise à jour logicielle, il nous a semblé pertinent de considérer les implications de la technologie sur la notion de confiance ainsi que de présenter les défis à relever.



Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra



Editions
Propulseurs



9 783952 589014