

Les membres de la Cité de l'IA se sont réunis pour définir ensemble les grands enjeux éthiques liés à la conception ou à l'utilisation de l'intelligence artificielle, en explorant les différentes dimensions de cette technologie tout en tenant compte des préoccupations actuelles. Les réflexions ont permis de mettre en évidence plusieurs points clés qui illustrent la position de la Cité de l'IA sur cette question.

Premièrement, le principe de **non-discrimination** a été identifié comme un enjeu majeur. La Cité de l'IA souligne l'importance de garantir la diversité, l'équité, l'égalité et l'accessibilité dans le développement et l'utilisation de l'intelligence artificielle. Cela implique de veiller à ce que les systèmes d'IA ne reproduisent pas ou n'amplifient pas les biais ou préjugés existant dans la société.

La question de la **confiance** est également primordiale. Pour établir la confiance dans l'IA, il est nécessaire de mettre en place des normes et des **réglementations** claires. La Cité de l'IA souligne l'importance de la normalisation et de l'applicabilité des systèmes d'IA afin de fournir des preuves tangibles de leur fiabilité et de leur conformité aux principes éthiques. Cette notion de confiance s'applique pour tous les sujets, y compris ceux internes à l'entreprise.

La **transparence** est également considérée comme un pilier fondamental de l'éthique de l'IA. La Cité de l'IA encourage la mise en place de mécanismes permettant l'**explicabilité** des décisions prises par les systèmes d'IA. Il est essentiel que les individus puissent comprendre comment les décisions sont prises par les modèles, en particulier quand elles les concernent ou impactent leur travail.

Dans le cadre de l'éthique de l'IA, les parties prenantes (concepteur de l'IA ou utilisateur) ont des rôles complémentaires. Les attentes de l'utilisateur concernent notamment sa sûreté, la confidentialité de ses données, l'exposition de ses actions, sa responsabilité dans les prises de décision et sa valeur ajoutée en tant qu'être humain. Le point de vue du créateur comprend des préoccupations juridiques, la capacité à faire adopter la solution, la définition des limites éthiques "acceptables" compatibles avec les objectifs commerciaux et la volonté d'innover, ainsi que la maîtrise des risques liés à l'IA. **L'intention** – aussi bien des créateurs que des utilisateurs de l'IA – est un point clé : pour simplifier une IA créée et utilisée avec une intention éthiquement acceptable sera éthique.

En ce qui concerne la place de l'homme, la Cité de l'IA met en avant le **rôle central de l'humain** dans la prise de décision. L'IA doit être ici considérée comme un outil au service de l'homme. Les décisions finales doivent être prises par des individus en prenant en compte les valeurs morales et éthiques globales et locales. Il est fondamental de veiller à éviter le traitement déshumanisé de situations par l'IA et de bien distinguer les interactions avec un humain et avec une machine.

La dimension **environnementale** est aussi primordiale, compte tenu des crises climatique, démographique, et énergétique actuelles. Il est indispensable de définir la finalité d'utilisation de l'IA, la manière de l'utiliser (sobriété, Green IT, etc.) et de mesurer l'impact carbone d'un projet d'IA.

Enfin, la protection des données et la propriété intellectuelle sont également des préoccupations majeures pour la Cité de l'IA. Il est nécessaire de garantir la propriété des données des utilisateurs et de tracer la propriété intellectuelle des algorithmes développés. De plus, la certification par des organismes externes et les méthodologies pluri-opinion pour tester la non-discrimination des systèmes d'IA sont des mesures recommandées par la Cité de l'IA.

Ces principes éthiques sont essentiels pour promouvoir un développement et une utilisation responsables de l'IA, tout en préservant les valeurs et les intérêts de la société dans son ensemble. La Cité de l'IA encourage donc une réflexion continue et une gouvernance adaptée pour relever les défis éthiques liés à l'intelligence artificielle.