Principes directeurs du CRIC sur l'utilisation de l'IA dans les études de marché

Introduction

Ces principes directeurs ont pour objectif de fournir un cadre régissant l'utilisation de l'intelligence artificielle[1] (IA) dans le domaine des études de marché, compte tenu de la généralisation croissante de sa mise en œuvre et de son utilisation. Ils visent à assurer une utilisation stratégique, éthique et responsable des applications d'IA.

Depuis sa publication en 2024, ce cadre a rempli sa mission, qui consistait à fournir des orientations pratiques conciliant innovation et responsabilité. Le sous-comité de l'éthique de l'IA du CRIC a ajouté des addendas à plusieurs sections dans le but de renforcer les principes plutôt que de les modifier en profondeur.

Principes directeurs

1. **Transparence :** assurer la transparence et la responsabilité dans l'utilisation de l'IA. Toute organisation qui utilise l'IA doit communiquer ouvertement et clairement avec les clients, les répondants et le grand public au sujet de l'intégration de l'IA dans les processus, et expliquer comment, pourquoi et quand l'IA est utilisée.

Addenda sur la transparence

Les organisations doivent faire la distinction entre les informations issues exclusivement de l'IA et celles générées par une analyse humaine (qu'elle soit assistée par l'IA ou non). Les détails comprennent :

- Expliquer le rôle de l'IA dans le travail de terrain (notamment l'utilisation de « personnes synthétiques[2] » et le rôle de l'IA dans l'épuration des données), ainsi que dans les processus d'analyse et de production de rapports (par exemple, tri préliminaire des données, détermination des tendances, extraction des thèmes, génération de rapports).
- Préciser clairement si des données synthétiques[3] ont été utilisées et dans quelle mesure.
- Fournir des informations contextuelles sur la manière dont un contrôle humain a été mis en œuvre pour vérifier ou affiner le contenu et les conclusions générés par l'IA.
- Déterminer les limites du contenu et des conclusions issus de l'IA.

2. **Sécurité des données :** le recours à des applications d'IA peut soulever des questions relatives à la sécurité des données. Il faut assurer le respect des normes pertinentes du CRIC, des pratiques essentielles en matière de sécurité indiquées dans la Trousse d'outils du CRIC sur la sécurité de l'information, ainsi que des contrats des clients en ce qui concerne la transmission, la conservation et la sécurité des données. Les chercheurs doivent accorder une attention particulière à la manière dont les informations saisies dans les applications d'IA seront utilisées par les applications pour l'apprentissage à partir de données ou la production d'autres résultats. Dans les situations où plusieurs normes ou contrats peuvent s'appliquer, le document de référence doit être celui qui fixe la barre la plus haute en matière de sécurité des données.

Annex relative á la sécurité des données

Surveillance de la sécurité et gestion des risques

- Les organisations doivent mettre en place des systèmes de surveillance continue afin de garantir la résilience quant à la sécurité de l'IA, notamment :
    - Effectuer régulièrement des vérifications de sécurité visant notamment à cerner les risques liés à l'utilisation de l'IA.
    - S'assurer que les procédures d'intervention en cas d'incident prennent en compte les failles de sécurité liées à l'IA.

Intégration de l'IA provenant de tiers fournisseurs

Lorsqu'elles intègrent des applications d'IA externes, les organisations doivent respecter des normes de sécurité rigoureuses :

- Effectuer des vérifications préalables et examiner régulièrement les politiques de sécurité et le niveau de conformité des fournisseurs tiers d'IA.
- Établir et consigner des protocoles de transmission sécurisée des données entre les systèmes de recherche et les applications d'IA externes.

3. **Protection des participants (sujets des données) contre tout préjudice :** comme pour tous les outils de recherche, il convient d'adhérer aux principes d'un programme rigoureux de gestion de la vie privée, comme l'exigent les normes du CRIC et comme l'indique la Trousse d'outils du CRIC sur la protection de la vie privée, lorsque vous avez recours à l'IA. Il faut veiller à ce que les données et les intrants que les applications d'IA utilisent soient obtenus, utilisés et divulgués en toute légalité. Il faut également tenir compte de la protection de la vie privée des répondants ainsi que de leur droit à comprendre comment leurs données sont recueillies et conservées. Plus précisément, les sujets des données doivent être clairement informées de l'utilisation d'une « personne synthétique » aux fins de collecte de données dès le début de la recherche. Il est essentiel d'indiquer clairement aux participants lorsqu'ils interagissent avec une application d'IA et non avec un être humain.

Addenda sur la protection des participants contre tout préjudice : les interactions facilitées par l'IA ainsi que l'autonomie et le bien-être des participants

Lorsque des applications d'IA sont utilisées dans le cadre d'une interaction directe ou indirecte avec les participants à la recherche (comme des « personnes synthétiques », notamment les modérateurs IA, les robots conversationnels, les systèmes de questionnement adaptatifs ou les invites automatisées), les organisations doivent prendre des mesures supplémentaires pour préserver l'autonomie, le bien-être et la dignité des participants.

Plus précisément :

- **Sensibilisation des participants et attentes**
  Les participants doivent disposer d'informations claires et accessibles sur la nature des interactions facilitées par l'IA, notamment sur les capacités et les limites du système d'IA. Les applications d'IA ne doivent pas être présentées d'une manière susceptible d'induire raisonnablement les participants en erreur en leur faisant croire qu'ils interagissent avec un être humain ou que le système possède une compréhension, un jugement ou une empathie humains.

- **Limites de l'autonomie de l'IA dans des contextes délicats**
  Le recours à une interaction facilitée par l'IA doit être soigneusement évalué dans le cadre de recherches portant sur des sujets délicats, personnels ou chargés d'émotion. Lorsqu'il existe un risque accru de détresse ou de malentendu, les organisations doivent mettre en place des mesures de protection telles qu'une supervision humaine, des protocoles de recours hiérarchique ou des restrictions concernant les interactions entièrement autonomes avec l'IA.

- **Contrôle et choix des participants**
  Dans la mesure du possible, les participants doivent se voir proposer des options claires concernant leur participation à des recherches impliquant l'IA, notamment la possibilité de refuser toute interaction avec l'IA ou de demander l'intervention d'un être humain sans encourir de sanction ni subir de désavantage.

- **Prévention des préjudices liés à l'expérience**
  Les systèmes d'IA doivent faire l'objet d'une surveillance afin de réduire le risque de préjudice pour les participants résultant d'un ton inapproprié, d'une mauvaise interprétation des réponses, de questions répétitives ou intrusives, ou d'un renforcement involontaire de contenus sensibles. Les résultats générés par l'IA qui ont une incidence directe sur l'expérience des participants doivent faire l'objet d'un examen et d'une amélioration.

- **Responsabilité et recours**
  Les organisations restent entièrement responsables de l'expérience et des résultats des participants lorsque l'IA est utilisée. Des procédures claires doivent être mises en place

pour traiter les préoccupations, les réclamations ou les événements indésirables des participants découlant d'interactions impliquant l'IA lors de la recherche.

4. **Efforts pour minimiser les préjugés :** il faut comprendre et surveiller les préjugés potentiels de l'IA et accorder la priorité aux besoins des personnes et des communautés, notamment des groupes défavorisés sur le plan de l'équité. Les chercheurs doivent rester vigilants quant aux préjugés et aux limites naturelles de l'IA et prendre des mesures pour minimiser leurs conséquences tout en favorisant le principe de responsabilité. Il faut évaluer les résultats des applications d'IA, notamment les outils génératifs, afin de réduire les préjugés et les inexactitudes. De plus, lorsque le contenu est principalement ou entièrement généré par l'IA, une mention claire de ce fait doit être ajoutée afin de maintenir la transparence. Ainsi, les clients sont entièrement au courant de la source et de la méthode de création du contenu.

5. **Assurer la surveillance :** les membres du CRIC doivent s'assurer que des mesures de contrôle efficaces sont en place. Ils sont encouragés à procéder régulièrement à des vérifications de la partialité de leurs applications d'IA, à créer des environnements de test et à mettre en applicables des mécanismes de surveillance humaine des applications d'IA afin d'assurer la responsabilité. Les équipes doivent être multidisciplinaires et comprendre des spécialistes des données, de l'éthique et du droit afin d'assurer une compréhension globale.

Addenda sur la supervision

Les organisations doivent étendre leurs processus existants en matière d'éthique de la recherche afin d'y intégrer explicitement les considérations liées à l'IA, en veillant à ce que la gouvernance de l'IA soit soumise au même niveau de rigueur que celui appliqué à l'éthique de la recherche classique. Les détails comprennent :

- S'appuyer sur les cadres éthiques de recherche existants, notamment en matière de détermination des risques, d'évaluation des effets sur les participants et le public, de clarification des rôles et des responsabilités, ainsi que d'examen par les comités d'éthique ou de gouvernance compétents. Intégrer dans ces cadres les considérations spécifiques à l'IA, comme les biais, les exigences de transparence et les obligations en matière de sécurité des données.
- Veiller à ce que les projets faisant appel à l'IA fassent l'objet d'un examen éthique proportionné à leur portée, à leurs risques et à leur incidence potentielle, en veillant au respect des principes canadiens d'éthique de la recherche et des normes du CRIC.

Définitions

1. **Intelligence artificielle :** ensemble de technologies conçues pour simuler l'intelligence humaine et offrir des capacités de résolution de problèmes (ce qui comprend les grands modèles de langage/GML).
2. **Données synthétiques :** informations générées dans le but de reproduire les caractéristiques des données du monde réel.
3. **Personnes synthétiques :** représentation numérique d'une personne générée pour imiter les comportements, les préférences et les caractéristiques de vraies personnes ou de groupes réels.