

Il n'y a pas d'éthique de l'intelligence artificielle sans éthique du travail réel

À mesure que l'intelligence artificielle s'installe dans les organisations, les chartes éthiques se multiplient. Transparence, explicabilité, robustesse, non-discrimination : jamais les entreprises n'ont autant réfléchi aux conditions d'un usage responsable des technologies.

Cette mobilisation est indispensable mais elle laisse dans l'ombre une question plus fondamentale.

Une organisation peut-elle prétendre développer une intelligence artificielle éthique lorsqu'elle méconnaît le travail réel de celles et ceux qui devront vivre avec elle ?

Notre réponse est non.

Nous faisons même une hypothèse plus radicale : le premier problème éthique de l'intelligence artificielle n'est peut-être pas l'intelligence artificielle elle-même. C'est notre ignorance persistante du travail réel.

Pourquoi ? Parce qu'une intelligence artificielle n'agit jamais directement sur le réel. Elle agit toujours à travers le travail humain. Elle modifie les prescriptions, les procédures, les indicateurs, les critères d'évaluation et les modes de décision. Avant de transformer les organisations, elle transforme le travail. Or le travail ne se réduit jamais aux modèles qui le décrivent.

Depuis plusieurs décennies, les sciences du travail convergent sur ce point. Entre le travail prescrit et le travail réellement accompli subsiste toujours un écart irréductible. Les procédures définissent ce qu'il conviendrait de faire, les professionnels réalisent ce qu'il faut effectivement faire pour que le travail soit efficace malgré les contingences du réel. Cet écart n'est pas un dysfonctionnement. Il est la condition même du fonctionnement des organisations, du travail bien fait et producteur de santé.

Aucune procédure ne s'applique d'elle-même. Aucun référentiel ne décide à la place des professionnels. Aucun algorithme ne peut anticiper la totalité des situations rencontrées. Dans le travail humain, entre la règle et son application intervient toujours le jugement.

Interpréter une consigne, arbitrer entre plusieurs objectifs contradictoires, composer avec l'imprévu, coopérer avec d'autres acteurs, assumer une responsabilité : voilà ce qui constitue le travail réel. C'est précisément parce que cette activité demeure largement invisible qu'elle est si souvent oubliée et donc niée.

L'intelligence artificielle accroît considérablement notre capacité à produire des représentations du travail. Les données deviennent plus abondantes, les modèles plus précis, les prédictions plus performantes. Cette évolution est une avancée considérable mais une représentation, aussi sophistiquée soit-elle, ne devient jamais le réel.

Le danger n'est pas que les modèles soient imparfaits. Aucun modèle ne pourra jamais épuiser la richesse des situations concrètes. Le véritable risque apparaît lorsque nous oublions cette limite et finissons par gouverner les organisations à partir de leurs représentations plutôt qu'à partir du travail effectivement réalisé.

C'est là que surgit la question éthique.

Une organisation devient injuste lorsqu'elle demande aux salariés de compenser silencieusement les insuffisances de ses procédures. Elle devient injuste lorsqu'elle ne reconnaît que ce qui est mesurable en oubliant ce qui rend effectivement le travail possible. Elle devient injuste lorsqu'elle attribue aux individus les défaillances de modèles qui ne correspondent plus aux réalités du terrain.

À l'inverse, une organisation véritablement éthique accepte que ses modèles soient continuellement confrontés au réel. Elle considère que les procédures existent pour servir le travail et non que le travail existe pour confirmer les procédures. Cette distinction change profondément notre manière d'aborder l'éthique de l'intelligence artificielle.

Aujourd'hui, les débats portent essentiellement sur les qualités intrinsèques des systèmes : leurs biais, leur transparence, leur sécurité ou leur explicabilité. Ces exigences sont évidemment nécessaires mais elles demeurent insuffisantes.

Une intelligence artificielle peut satisfaire à toutes ces exigences tout en dégradant profondément le travail réel : en rigidifiant les décisions, en invisibilisant les arbitrages des professionnels, en réduisant leur autonomie ou en affaiblissant leur capacité de jugement. Autrement dit, il est impossible d'évaluer l'éthique d'une intelligence artificielle indépendamment de ses effets sur l'activité humaine.

C'est pourquoi il devient nécessaire de reconnaître l'existence d'une éthique du travail réel.

Cette éthique ne vient pas concurrencer l'éthique de l'intelligence artificielle. Elle lui est antérieure. Elle en constitue le fondement car avant de s'interroger sur les qualités morales d'un algorithme, encore faut-il comprendre le travail dans lequel cet algorithme prendra place.

À mesure que les modèles deviennent plus puissants, une responsabilité nouvelle apparaît : préserver ce qu'aucun modèle ne pourra jamais entièrement contenir. Le jugement, l'expérience, la coopération, la responsabilité et l'intelligence des situations.

L'intelligence artificielle continuera de produire des modèles toujours plus performants. C'est une excellente nouvelle mais aucune organisation ne sera véritablement éthique si elle oublie que les modèles n'agissent jamais seuls. Entre les modèles et l'action demeurera toujours le jugement des femmes et des hommes.

C'est pourquoi il n'y a pas d'éthique de l'intelligence artificielle sans éthique du travail réel.

Ibrahima Fall est docteur en sciences de gestion, président de l'Institut du Travail Réel et dirigeant du cabinet Hommes & Décisions. Il est l'auteur de *L'Entreprise contre la connaissance du travail réel ?* (L'Harmattan).