

Le travail et l'emploi à l'épreuve de l'IA : Etat des lieux et analyse critique de la littérature

Flore Barcellini (CRTD – Le Cnam), Tamari Gamkrelidze (CRTD – Le Cnam), Nathalie Greenan (LIRSA – CEET – Le Cnam), Annie Jolivet (CRTD – CEET – Chercheure associée à l'IRES), Moustafa Zouinar (Le Cnam et Orange Labs)

Synthèse (Juin 2024)

IA : potentialités et limites

L'IA ne désigne pas une technologie ou un type particulier de système mais un ensemble hétérogène d'algorithmes, de systèmes informatiques, de machines (robots) et plus largement de technologies sous différentes formes (logiciels, robotiques etc.) qui visent à imiter des facultés cognitives humaines comme la perception visuelle ou auditive, la production et la compréhension du langage naturel, ou encore le raisonnement. Le terme « Intelligence » est utilisé de manière abusive alors même que les définitions sur ce qu'est l'intelligence humaine fait débats.

Il existe deux approches en IA : l'approche symbolique et l'approche par apprentissage comme le connexionnisme, qui est devenue dominante aujourd'hui. Dans cette approche, ce qui est appelé « apprentissage » est la production de modèles construits à partir d'un grand nombre de données structurées (i.e. annotées par un humain) et basées sur des optimisations mathématiques.

Les SIA actuels renvoient à 5 grandes fonctionnalités : recherches, reconnaissances et analyse d'informations ; prédiction et diagnostic, recommandation, exécution autonome d'action, génération de contenu nouveau. Ces systèmes présentent plusieurs limites : opacité et manque d'explicabilité de leur résultat, incertitude sur l'évolution dans le temps des systèmes, nature et biais des données utilisées pour l'apprentissage.

Au-delà des promesses d'augmentation de l'humain ou de l'automatisation de tâches routinières,

il existe un risque accru d'automatisation et de subordination du travail et la réelle collaboration entre humains et IA reste un mirage.

Transformations de l'emploi liées à l'IA

Les représentations des enjeux d'emploi dans le débat public se construisent au croisement des grandes narrations futurologiques et des concepts élaborés par la recherche académique qui ont été testés sur des jeux de données portant sur les périodes antérieures.

L'approche économique néoclassique traitant des effets sur l'emploi des technologies émergentes est techno-déterministe. Elle voit l'introduction des technologies comme « un choc exogène » qui apporte avec elle la direction de ses effets sur l'emploi et les revenus. L'introduction de technologies émergentes au travail – le progrès technologique – y est vue comme source de biais qui transforme le partage de la valeur ajoutée entre les facteurs de production. Selon les périodes et les technologies, ces biais favorisent le capital, le travail qualifié, la main d'œuvre moins âgée, les tâches non routinières ou les tâches non automatisables.

Pour une approche alternative, il s'agit de considérer les choix possibles en matière d'organisation du travail. Plutôt que d'adopter une approche de la production centrée sur les tâches, nous examinons une approche de la production centrée sur la production de connaissances nouvelles. Les IA, en association avec la R&D et la capacité d'apprentissage des organisations, contribuent à cette production des connaissances nouvelles sources d'innovations pour l'entreprise. L'investissement dans la capacité d'apprentissage est un choix d'organisation aux effets protecteurs sur l'emploi dans un contexte d'adoption de technologies nouvelles.

Pour comprendre les effets du progrès technique sur l'emploi, il faut également suivre ce qui se passe au niveau sectoriel où les entreprises sont en concurrence avec d'autres, au niveau des zones d'emploi où se manifestent les excès et les pénuries de main d'œuvre et s'interroger sur les enjeux de financement des investissements nouveaux.

Effets sur le travail et l'emploi : quelles articulations ?

Les travaux qui permettent d'articuler les effets sur l'emploi et les effets sur le travail sont relativement peu nombreux. Deux voies d'articulation ont été repérées. La première passe par

l'analyse d'éléments de contexte plus larges. Par exemple pour comprendre en quoi l'introduction d'un SIA peut modifier le travail des personnes qui l'utilisent, la répartition des tâches, réduire les emplois ou accompagner leur réduction, créer des emplois. La seconde passe par les modes de régulation, leurs champs d'application et leur prise en compte ou non des questions d'emploi et de travail.

Les approches monographiques ou par études de cas, au niveau d'entreprises ou de secteurs, sont les plus à même d'envisager les effets fins sur le travail, sur l'organisation du travail et de permettre de les relier aux effets sur les emplois. Une partie des effets sur l'emploi est peu visible (délocalisation, sous-traitance, départs naturels, réduction des effectifs déjà engagée).

Les travaux sur la régulation de l'IA soulignent la nécessité de penser une régulation d'ensemble pour pallier les incohérences et les lacunes des instruments juridiques actuels. La possibilité de réguler les conséquences des SIA sur le travail et l'emploi dépend beaucoup du modèle de relations professionnelles. De grandes différences existent entre les pays.

Usages de l'IA et transformations du travail : de la projection aux usages en situation réelle de travail

Dans l'ensemble, les études examinées montrent que les travailleurs ont globalement des points de vue plutôt « réalistes » et nuancés sur l'IA :

- Ils projettent des apports potentiels de l'IA, parfois en lien avec les discours de promotions de l'IA, par exemple l'automatisation de tâches répétitives, le gain de temps, une amélioration de la prise de décision. La valeur a priori des tâches automatisées par l'IA resterait cependant à questionner et à confronter à une valeur en action, basée sur des expériences concrètes.
- Ils identifient de nombreux risques et limites concernant la performance du travail (qualité, fiabilité) mais aussi la santé : perte d'emploi, opacité des systèmes, perte d'expertise et d'autonomie ; perte du sens du travail ; biais d'automatisation - sur confiance, pertinence, éthique des résultats de l'IA.
- Ils s'interrogent sur l'utilité « réelle » des systèmes qui sont proposés dans les secteurs professionnels qui les concernent. Sur ce dernier constat, il convient de noter que ce questionnement de l'utilité réelle n'est pas seulement une projection a priori mais repose sur

des expériences vécues de certains des acteurs interrogés comme dans l'étude de Gamkrelidze (2022).

Concernant la « prévention » des risques de dégradation de la performance du travail réalisé avec un SIA, plusieurs éléments sont soulignés :

- Le besoin de transparence sur les apports et les limites de l'IA (ce qu'elle permet de faire et de ne pas faire) ;
- Le besoin de compréhension des résultats fournis par l'IA (explicabilité) ;
- La prévention d'un risque de normalisation / standardisation des raisonnements mis en œuvre par les professionnels ;
- La nécessité d'un pouvoir de contrôle sur l'IA en lien avec des questions de fiabilité du travail, de responsabilité et d'indépendances des actions entreprises à l'aide des résultats produits par une IA ;
- La nécessité de réflexions sur les compétences à agir avec une IA qui concerne à la fois des démarches de développement de nouvelles compétences et de remobilisation d'expériences et de savoir-faire ;
- Des réflexions sur la confiance / la méfiance envers l'IA.

Ces résultats renforcent l'importance, pour aller plus loin, de la compréhension fine de l'activité dans ses dimensions physiques, cognitives et sociales – afin d'identifier et d'évaluer les effets transformatifs des SIA, et de soutenir le développement de nouvelles façons de faire et/ou la mobilisation d'expériences passées, ce qui est souvent un impensé de l'introduction des systèmes techniques dans les environnements de travail.

Plusieurs études expérimentales dans différents secteurs mettent en avant une augmentation de la performance (en termes de temps, d'efficacité, de précision, de productivité, etc.) dans différents types de tâches (diagnostic médical, résolution de problème, production de textes ou de programmes) réalisés à l'aide de l'IA.

Cependant, cette performance accrue n'est pas systématique et universelle :

- Elle bénéficie surtout aux travailleurs les moins experts ;
- Elle est meilleure lorsqu'elle est précédée de « tutoriels » ;

- Elle ne concerne que des tâches qui sont « à la portée de l'IA » ;
- Elle est dépendante du niveau de fiabilité du système et de ses possibilités de soutenir la construction d'une distance critique vis-à-vis de ce qu'il produit.

En conséquence, et dans certaines situations, il peut y avoir des effets négatifs voire pas d'effet positif sur la performance humaine ou des risques de diminution de la créativité / d'uniformisation de la pensée et des produits de l'action et donc une baisse de performance qualitative.

Les études expérimentales présentent des limites qui amènent à considérer avec prudence leurs résultats en situations de travail réelles : vision de la performance réduite à une performance technique pour certaines tâches du travail et excluant les effets sur la santé des professionnels, absence de prise en compte du rôle de la dimension collective et systémique du travail sur la performance.

Les études empiriques sur les transformations réelles du travail liées à l'introduction des SIA nuancent les résultats expérimentaux quant à la performance des IA. Le SIA introduit, comme second lecteur dans certaines situations, peut constituer un « filet de sécurité », mais surtout pour les professionnels novices et comme palliatif dans des situations de travail dégradées. Seule une analyse fine et systémique du travail réel ancrée dans l'expérience des professionnels permet d'appréhender la valeur réelle de l'activité – avec ou sans un SIA – au-delà des projections théoriques sur ce qui fait valeur.

Ces travaux identifient des effets sur la santé des professionnels (augmentation de la charge de travail, perte de sens au travail, entrave au travail collectif), une transformation potentielle de la construction de l'expérience professionnelle et des risques sur la fiabilité des actions menées avec une IA, notamment en lien avec l'opacité des systèmes. On observe un effet des contextes et des choix organisationnels sur les développements.

Les usages pertinents impliquent la mise en œuvre de « pratiques interrogatives » soutenant la construction d'une distance critique pour comprendre les résultats de l'IA, afin de donner du sens entre les interprétations des professionnels et les résultats de l'IA. Ces pratiques plus que nécessaires ne sont souvent pas soutenues dans des organisations du travail « sous-contraintes » et se révèlent donc coûteuses en temps pour des professionnels qui ont déjà une charge de travail importante.

Des propositions de démarches de conception d'IA existent, elles reprennent des principes éthiques à respecter, des recommandations pour assurer la compatibilité de l'IA avec l'infrastructure technologique existante, ou des principes de modèles de conduite du changement (communication, pédagogie dans les projets) ou d'une conception d'IA centrée sur l'humain.

Les intentions derrière les projets d'introduction d'IA au travail sont opaques dans les études et cas mobilisés. Lorsqu'elles sont mises à jour, on retrouve des intentions parfois marquées par des logiques de rationalisation financière, de course technologique, de solutionnisme technologique. Les conduites de projets sont le plus souvent impensées.

Le développement de démarches de conduite de projets d'IA centrées travail sont nécessaires pour enrichir les logiques prises en compte dans la définition des projets d'IA et permettre aux professionnels de pouvoir penser et pouvoir débattre des transformations de leur travail et de ce qui fait valeur pour eux.

Le développement de ces démarches de conduites de projet implique de repenser les rapports de force dans les prises de décisions et dans la conduite des projets et de porter une vision émancipatrice du travail dans les projets : un travail qui soit source de développement de l'expérience, de santé et porteur d'un sens.