AI regulation without theater

AI Regulation Without Theater

Une politique peut décrire une intelligence artificielle responsable avec des mots impeccables sans modifier une seule opération dangereuse. Elle exige parfois la transparence sans dire quels journaux conserver, promet une supervision humaine sans donner à quelqu’un le pouvoir d’arrêter un système, ou réclame des évaluations impossibles à reproduire. La régulation devient alors une représentation plutôt qu’un dispositif de contrôle.

La gouvernance effective commence là où une décision peut être observée et corrigée. Chaque usage important possède un responsable, une finalité, des entrées et sorties autorisées, un historique des versions du modèle et des instructions, des essais de défaillance et une voie d’escalade. On distingue explicitement les suggestions des actions qui nécessitent une approbation ou un blocage technique.

Les essais doivent correspondre au service réel. Un assistant documentaire, un outil de tri médical et un dispositif intervenant dans un procédé industriel n’exposent pas les mêmes risques. L’évaluation réunit donc scénarios de métier, tests adversariaux, examen des données personnelles, surveillance et réponse aux incidents.

Le marché public ou privé doit aussi conserver des droits utiles : accès aux traces, annonce des changements de modèle, export des données, limites de responsabilité, traitement des incidents et possibilité de sortir du contrat. Une promesse commerciale importe moins qu’un essai de réception que le client peut refaire.

Une réglementation se juge à la conduite effective des systèmes et à la capacité de les corriger. Si une exigence ne devient ni responsabilité vérifiable, ni essai, ni contrôle technique, ni recours applicable, la protection demeure incomplète.