

Hadopi

Consultation du Conseil d'Etat sur l'intelligence artificielle et l'action publique

1. Prénom NOM

Pauline BLASSEL

2. Entité

Hadopi

3. Fonction

Secrétaire générale

4. Adresse mail professionnelle

pauline.blassel@hadopi.fr

Question n°5 : Avez-vous développé ou financé - ou projetez-vous de le faire à brève échéance des projets recourant à l'intelligence artificielle - IA (ou dont vous estimez qu'elles utilisent l'intelligence artificielle) pour l'accomplissement des missions qui vous sont confiées (ou aux missions qui sont confiées aux opérateurs/établissements dont vous assurez la tutelle) ?

Oui

Question n°6 : Si vous avez répondu "oui" à la première question, veuillez préciser les objectifs du projet, la méthode employée (association des personnels...), les techniques et algorithmes utilisés (apprentissage machine – supervisé, non supervisé, par renforcement... - apprentissage profond...), les moyens associés, le coût de conception et de maintenance de l'outil, le ou les partenaires et prestataires publics ou privés sollicités (tâches internalisées/externalisées), l'état d'avancement (en projet, en cours de conception, déployé, abandonné...) et les résultats obtenus.

L'Hadopi, dont les missions s'exercent principalement dans l'univers numérique, s'est intéressée au recours à l'intelligence artificielle dans trois domaines :

- l'observation des usages culturels numériques, notamment pour qualifier et quantifier les contenus culturels hébergés par les services numériques ;

- le dialogue avec les internautes, notamment dans le cadre de la mise en œuvre de la procédure de réponse graduée ;
- la lutte contre les services illicites, pour faciliter leur identification et leur caractérisation.

L'observation des usages

Les débats qui ont présidé à la création de l'Hadopi avaient mis en évidence une carence importante de données fiables sur le fonctionnement des services numériques et leur utilisation. L'Hadopi a ainsi été chargée par le législateur d'une mission générale d'observation des pratiques culturelles en ligne dans le cadre de laquelle ont été conduits de nombreux travaux faisant appel à des technologies innovantes et, ponctuellement, à l'intelligence artificielle.

En 2012 et 2013, l'Hadopi a souhaité qualifier et quantifier les contenus culturels présents sur les grandes plateformes audiovisuelles (YouTube, Dailymotion) et sur certains services manifestement illicites (comme Megaupload ou Rapidgator par exemple).

Le volume considérable de contenus hébergés par ces différentes plateformes n'en permettait pas une évaluation manuelle. L'Hadopi a donc réalisé des échantillonnages de ces contenus, selon différentes méthodes (marche aléatoire ou *Random Walk*, génération d'URL aléatoires, par exemple), tendant à reproduire le comportement normal d'un utilisateur naviguant sur ces services.

La qualification et la quantification des contenus de l'échantillon ont été réalisées parfois de façon manuelle (visionnage) et parfois (dans le cas de Megaupload) par recours à l'IA : sur la base de quelques mots clés prédéfinis, le programme a été entraîné à classer les contenus à partir de leur titre (apprentissage machine supervisé).

Le recours à cette méthode a permis de réaliser un inventaire des contenus hébergés par Megaupload avec très peu de moyens : aucun budget identifiable n'a été nécessaire, les travaux ont été réalisés en interne par deux informaticiens sous contrôle des services juridiques.

Le site Megaupload ayant été rapidement fermé par la justice américaine, les résultats ont été peu exploités. En revanche, les résultats obtenus sur YouTube ont concouru à mettre en évidence dès 2013 le rôle déterminant de cette plateforme dans la diffusion des œuvres culturelles sur internet.

Le dialogue avec les internautes

Un projet de *chatbot* est en cours de programmation à l'Hadopi. Cet outil, se présentant sous la forme d'une fenêtre de chat (ou discussion) sera accessible sur le site de l'Hadopi, et permettra de répondre automatiquement aux questions d'internautes relatives à l'Hadopi, à ses missions, aux contenus et services proposés.

Cet outil devrait notamment permettre d'apporter de premières réponses aux nombreux abonnés rendus destinataires de recommandations de l'Hadopi.

Actuellement en phase de test, il repose sur le recours à l'IA (reconnaissance de textes lors de la formulation des questions).

Ce projet a associé les services en charge de la procédure de réponse graduée et ceux en charge de la communication. Le développement de l'outil a été assuré par un prestataire spécialisé.

La lutte contre les services illicites

Un algorithme visant à faciliter la caractérisation des sites susceptibles d'être illicites est actuellement en cours de développement. Cet algorithme, reposant sur une technique de *machine learning* supervisée (ou apprentissage automatique), notamment les *RandomForest*, est basé sur l'entraînement d'une machine en utilisant une base de données existantes de plusieurs milliers de sites illicites déjà caractérisés.

Ce projet, réalisé en interne, est actuellement dans sa première phase : programmation de l'algorithme avant test de celui-ci sur le corpus existant de sites illicites. Il associe les services juridiques et techniques. Le recours à un prestataire n'est pas envisagé à ce stade.

Question n°7 : Si vous avez répondu "oui" à la première question, quels sont les facteurs clés de succès et d'échec que vous avez identifiés ?

Facteurs de succès :

La diversité des compétences en interne est un premier facteur de succès. En effet, la réussite de la mise en place d'une solution d'apprentissage nécessite le recours à différentes compétences : expertise sectorielle, expertise technique en programmation informatique, expertise de traitement de données. Ces trois compétences sont disponibles au sein de l'institution.

Le dimensionnement des bases de données peut conditionner la bonne marche de tels projets. La taille de la base de données utilisée (environ plusieurs milliers de nom de domaine) est relativement modérée, son faible volume permet un traitement en interne, sans passer par des serveurs externes.

Facteurs d'échec :

Le potentiel d'automatisation des analyses est central. La réussite du projet étant conditionnée à la qualité des données récoltées et à leur adaptation à la solution développée (identification de la licéité du site), il n'est pas encore certain que les données automatiquement analysables sur un site illicite (titres de films, de catégories, etc.) soient suffisantes pour caractériser celui-ci.

Le cadre juridique doit être pris en compte.

La collecte automatisée de données destinées à alimenter des programmes recourant à l'intelligence artificielle peut se heurter à plusieurs obstacles juridiques.

La collaboration des services de l'institution avec ceux de la Commission nationale de l'informatique et des libertés (CNIL) permet de tenir compte des enjeux liés à la protection des données personnelles. Sous réserve de mettre en place des procédés d'anonymisation robustes et de respecter des règles de sécurité (machine dédiée, codes d'accès, etc.), la protection des données personnelles ne s'est pas avérée faire obstacle à la réalisation de travaux de collecte et d'analyse de données.

Mais la collecte de données ne doit ni enfreindre les conditions générales d'utilisation du service, ni représenter une atteinte à son système automatisé de traitement de données. Ainsi, hors autorisation expresse par la loi, en l'absence de sécurité juridique suffisante pour l'autorité, de nombreux travaux d'observation pertinents ne peuvent être envisagés, même lorsqu'il s'agit pourtant de données collectées et traitées en permanence à des fins d'indexation par les moteurs de recherche. Il n'est que rarement possible d'obtenir l'accord du service, notamment lorsqu'il s'agit d'acteurs internationaux et *a fortiori* lorsqu'il s'agit de services manifestement illicites.

La taille réduite de l'équipe (un seul agent spécialisé en *data science* et programmation informatique) peut rallonger la durée du projet.

Question n°8 : Si vous avez répondu "oui" à la première question, est-ce que ce ou ces projets s'inscrivent dans une stratégie numérique ou IA plus vaste ? Si oui, quels sont les piliers de cette stratégie ? Veuillez nous orienter vers la documentation présentant cette stratégie.

L'Hadopi n'a pas mis en place une stratégie numérique dédiée exclusivement au recours à l'intelligence artificielle, l'utilisation d'une technologie étant un moyen, et non une fin en soi.

Dans une volonté d'amélioration de ses capacités d'études et d'analyse, l'Hadopi multiplie cependant le recours à des technologies innovantes. Ainsi, d'autres projets d'étude de l'Hadopi, sans pour autant recourir explicitement à l'intelligence artificielle, reposent sur des technologies numériques récentes, telles que le *webscraping* – ou recueil automatique de données -, afin de comparer la disponibilité d'œuvres de cinéma sur des sites licites ou illicites (projet en cours de développement, réalisation interne).

Un projet réalisé en partenariat avec le Pôle d'expertise de la régulation numérique (PEReN), visant à quantifier et qualifier les vidéos hébergées sur YouTube, repose sur un recours soutenu à l'API de YouTube, afin de recueillir une grande quantité de vidéos (technique pouvant s'apparenter à du *webscraping* encadré par le recours à l'API).

Question n°9 : Si vous avez répondu "oui" à la première question, ce ou ces projets vous paraissent-ils reproductibles dans d'autres administrations ? Si oui, à quelles conditions ?

Ce type de projet (caractérisation de sites illicites) peut tout à fait être adopté par d'autres administrations, à des fins similaires : détection de faux comptes de réseaux sociaux ou de comptes litigieux (par exemple dans le cadre de la lutte contre les contenus haineux).

L'Hadopi et le Conseil supérieur de l'audiovisuel (CSA) devant fusionner le 1^{er} janvier 2022 pour créer l'Autorité de régulation de la communication audiovisuelle et numérique (Arcom), il pourra être envisagé de décliner ce projet de l'Hadopi dans le cadre de la lutte contre les contenus haineux, mission dont le CSA a la charge (sous réserve de son succès).

Dans ce cas, la duplication du projet paraît envisageable, compte tenu des similarités entre les deux besoins. Mais plus largement, chaque besoin et chaque projet étant uniques, il paraît difficile, en tout cas actuellement, de décliner simplement un projet mené par une administration pour une autre. Tout algorithme ou programme informatique nécessitera des aménagements plus ou moins importants et nombreux pour s'adapter aux besoins de chacun.

En revanche, des échanges poussés entre les administrations, directement entre elles ou via des intermédiaires – à l'exemple du PEReN – favorisent la circulation des bonnes pratiques.

Question n°10 : Quels obstacles avez-vous identifiés dans le déploiement de projets recourant à l'IA ?

	obstacle				
	Aucun obstacle	1/4 (très faible)	2/4 (faible à moyen)	3/4 (moyen à fort)	4/4 (fort)
Utilité non avérée	x				
Manque d'idées		x			
Manque de temps			x		
Manque de données ou données de qualité insuffisante			x		
Manque de moyens budgétaires		x			
Manque d'accompagnement / de formation		x			
Absence de prestataires compétents				x	
Résistance ou manque d'appropriation des personnels		x			
Résistance des partenaires nécessaires à la conduite du projet IA		x			
Risque juridique (RGPD...) ou incertitudes du cadre juridique					x
Règles de la commande publique					x
Problème de sécurisation de l'outil (risque cyber...)			x		

Question n°11 : Veuillez expliciter vos réponses à la question précédente

Les obstacles liés aux risques juridiques apparaissent comme les plus importants.

Les exigences liées au respect du RGPD, comme la réalisation d'une analyse d'impact relative à la protection des données (AIPD), qui doit obligatoirement être menée préalablement au traitement quand celui-ci est susceptible d'engendrer un risque élevé pour les droits et libertés des personnes concernées, nécessitent des équipes et des délais relativement importants à l'échelle de l'Hadopi mais peuvent être satisfaites.

Le respect de ces exigences peut néanmoins dissuader certains prestataires, potentiellement sous-traitant des administrations, de répondre aux appels d'offre.

Par ailleurs, comme indiqué supra (Q7), l'évaluation des risques juridiques pris par l'administration au regard des conditions générales d'utilisation du service et d'une éventuelle atteinte à son système automatisé de traitement de données demeure très incertaine.

Question n°12 : Si vous avez identifié d'autres risques que ceux mentionnés ci-dessus, veuillez les indiquer.

La difficulté à recruter des profils d'ingénieurs et de data analyst, encore rares, fortement courtisés par les entreprises privées – du numérique, mais pas seulement – est un problème majeur pour l'administration. Les règles de rémunération des agents publics ne permettent que rarement aux administrations d'être attractives.

Certains jeunes profils, qui restent susceptibles d'être intéressés, s'inquiètent souvent, en outre, des risques d'incompatibilité qui contraindraient la poursuite de leur carrière.

Plus largement, la conduite de projets technologiques innovants nécessite d'être rapide, flexible, de bénéficier de circuits de décision courts et d'une certaine émulation, ce qui n'est pas encore l'apanage de l'administration.

Un deuxième obstacle est technique : l'importance des capacités informatiques, en termes de serveurs et de puissance de calcul, doit être pris en compte dans la mise en place et la réussite des projets. Des équipements supplémentaires (ordinateurs dédiés aux calculs, serveurs spécifiques) peuvent être nécessaires dans le cas de projets de grande ampleur, ce qui présente un coût.

Question n°13 : Pensez-vous que l'IA puisse accroître sensiblement la performance de votre entité ?

Oui

Question n°14 : Si vous avez répondu "oui" à la question précédente, à quel(s) titre(s) ? (plusieurs réponses possibles)

Plusieurs réponses possibles.

X	Amélioration de la qualité / accessibilité / continuité du service public
	Elargissement du périmètre de l'action publique
	Economies budgétaires
X	Revalorisation des fonctions des agents publics
X	Image du service public
X	Fiabilité des données
	Autre

Question n°15 : Quels types d'outils recourant à l'IA vous paraît-il utile de concevoir pour les besoins de votre entité ?

	Pas utile	utilité faible	utilisé moyenne	utilité forte	Je ne sais pas
Gestion et exploitation des documents et informations disponibles (« data management ») ;				X	
Identification des priorités et des risques, ciblage des contrôles ou des interventions			X		
Prise de décision automatisée (sans intervention humaine)				X	
Aide à la décision (instruction de demandes, évaluation des chances de succès d'une action...)					X
Reconnaissance de textes, d'objets ou de personnes				X	
Automatisation de la relation avec les usagers (robot conversationnel...)			X		
Evaluation des politiques				X	

Question n°16 : Si vous avez identifié d'autres outils que ceux mentionnés ci-dessus, veuillez les indiquer.

Les outils recourant à l'IA peuvent aussi servir à mieux reconnaître des contenus, objets ou personnes, comme par exemple :

- des sites internet et leurs similarités afin de pouvoir les regrouper en galaxies (regroupement d'un site « originel » et des sites miroirs) ;
- des contenus de toutes natures, en particulier audio ou vidéo, des images fixes, des documents (afin d'analyser l'efficacité des mesures techniques d'identification mises en place par les services de partage de contenus en ligne).

Plus généralement, même si les outils recourant à l'IA à des fins d'analyse des contenus et des usages n'ont pas forcément d'utilité directe pour la Haute Autorité, leur adoption par les industries culturelles, et plus encore par les plateformes numériques, est susceptible d'entraîner des progrès et des évolutions majeures en matière de protection des contenus sur internet, ainsi qu'en matière de lutte contre les pratiques indésirables et illicites. Ces changements sont de nature à avoir un impact sur les activités de l'institution.

À titre d'exemple, l'Hadopi s'est vu confier la tâche d'évaluer l'efficacité des technologies mises en place par les plateformes de partage de contenus sur internet dans le but de lutter contre la diffusion interdite de ces contenus. L'adoption progressive par ces plateformes de technologies reposant en grande partie sur l'IA sera donc un élément clé que l'Hadopi devra prendre en compte dans sa mission d'évaluation.

Question n°17 : Quelles politiques et actions devraient selon vous être mises en place par l'Etat pour développer le recours à l'IA dans le service public ?

Une seule réponse possible par ligne.

	Pas une priorité	Priorité faible	Priorité moyenne	Priorité forte
Information sur les opportunités offertes par l'IA, soutien à l'émergence d'idées			X	
Accompagnement individualisé des projets		X		
Equipes dédiées / développement de pôles de compétences au sein des ministères			X	
Centre de ressources dédié pour les administrations		X		
Formation générale des agents publics à l'IA				X
Formation d'experts IA au sein de l'administration				X
Renforcement des liens avec la recherche			X	
Financements			X	
Faciliter le recours à des prestataires privés			X	

Question n°18 : Si vous avez identifié d'autres politiques et actions que l'Etat, ou une autre administration, pourrait mettre en œuvre pour promouvoir le déploiement de l'IA dans le secteur public, veuillez les indiquer

L'État au sens large (administrations, ministères, etc.) pourrait mettre en place différentes actions de moyen et long terme pour promouvoir le déploiement de l'IA dans le secteur public.

Dans le domaine de compétence de l'Hadopi, il convient en premier lieu de rappeler que la France a souvent fait figure de pionnière en matière de politiques publiques à l'égard de la culture, des arts, et plus largement du monde des médias et de la communication.

Les acteurs français de ces secteurs ont une longue tradition d'innovation – la France est par exemple un leader reconnu des effets spéciaux audiovisuels de l'animation, de la reconnaissance de contenus (avec l'INA). Elle dispose notamment d'entreprises de premier ordre dans l'univers des jeux vidéo. Ce sont des secteurs au sein desquels l'innovation concourt à notre rayonnement culturel, ce qui pose la question de l'intervention de l'Etat avec acuité.

Il paraît donc particulièrement important d'accompagner ces pratiques d'innovation, en renforçant ou en adaptant aux secteurs des industries culturelles les dispositifs actuels de soutien, tels que le crédit d'impôt pour la compétitivité et l'emploi (CICE), les aides de la BPI, la mise en place d'un fond dédié à des projets innovants dans le secteur de la culture, etc.

De manière concomitante, il serait pertinent de renforcer les liens entre le monde de la recherche publique et les initiatives privées dans le secteur de l'IA en particulier. La France dispose de chercheurs de notoriété mondiale en ce domaine (on peut citer par exemple Aurélie Jean, enseignante aux États-Unis et consultante ou Yann Le Cun, titulaire du prix Turing en 2017, actuellement scientifique en chef de l'IA au sein de Facebook), mais ceux-ci ne travaillent pas toujours pour l'université française ou des acteurs industriels français, leur préférant parfois des acteurs internationaux, ce qui est dommageable pour la recherche et l'industrie française.

Par ailleurs, dans une stratégie de long terme de renforcement des compétences technologiques de la France, il conviendrait de mieux valoriser les études scientifiques et techniques et les parcours d'ingénieurs (et ce en particulier auprès des lycéennes, alors que la féminisation de la filière scientifique ne progresse que faiblement et que celles-ci restent minoritaires dans l'enseignement supérieur scientifique). De nombreuses actions sont prises en ce sens et devraient être maintenues et renforcées, en mettant en avant le fort potentiel que présente les métiers en lien avec l'IA et le traitement des données.

Enfin, un renforcement de l'attractivité de l'administration en tant que « marque-employeur », auprès des futurs ingénieurs et spécialistes de l'IA, permettrait de compléter ce plan global de promotion du déploiement de l'IA dans le secteur public. Comme évoqué précédemment dans la question Q12, et pour pallier les difficultés de l'administration à recruter ces profils recherchés, la richesse des missions et emplois potentiels au sein de l'administration, sur des sujets en lien avec l'IA, pourrait être mise en avant, dans des campagnes de communication auprès des candidats potentiels – à l'image par exemple des campagnes de recrutement que fait le ministère de la Défense pour l'armée.

Question n°19 : Quels risques juridiques associez-vous à l'utilisation de l'IA pour l'exercice des missions qui vous sont confiées ?

La nécessité d'encadrer le développement de l'intelligence artificielle fait consensus.

Comme le souligne la Commission dans sa proposition de règlement ayant pour but d'établir un cadre réglementaire pour l'intelligence artificielle du 21 avril 2021, l'intelligence artificielle « *recouvre un large champ de technologies en évolution rapide et peut procurer de nombreux avantages économiques et sociétaux dans l'ensemble des secteurs économiques et des activités sociales* ». Néanmoins, « *les éléments et techniques qui rendent possibles les bénéfices socio-économiques de l'IA peuvent aussi être à l'origine de nouveaux risques ou de conséquences négatives pour les personnes ou la société.* ».

La nécessité de légiférer dans le domaine en introduisant une « approche équilibrée » se pose avec acuité.

Cette recherche d'équilibre et la volonté affichée des pouvoirs publics, tant à l'échelon national qu'europpéen, d'une plus grande régulation est aujourd'hui encore renforcée avec la présence dominante des GAFAM dans l'univers numérique, notamment au vu de la manière dont les systèmes algorithmiques façonnent les flux d'informations en ligne. L'intervention d'un régulateur afin de lutter contre les biais anticoncurrentiels, pro-piratage et pro-désinformation semble faire consensus tout comme la nécessité de favoriser une transparence accrue des algorithmes (Etude académique sur la transparence algorithmique, Alain Strowel).

En France, s'agissant plus précisément du rôle prescripteur des algorithmes destinés à la personnalisation, au classement et au référencement, l'article 49 de la loi n°2016-1321 du 7 octobre

2016 pour une République numérique implique que les plateformes sont tenues de « *délivrer au consommateur une information loyale, claire et transparente sur les conditions générales d'utilisation du service d'intermédiation qu'il propose et sur les modalités de référencement, de classement et de déréférencement des contenus, des biens ou des services auxquels ce service permet d'accéder* ».

L'Hadopi a également souligné dans le rapport « *Les outils de reconnaissance des contenus sur les plateformes numériques de partage. Propositions pour la mise en œuvre de l'article 17 de la directive européenne sur le droit d'auteur | Hadopi* » qu'elle a réalisé aux côtés du CSPLA et du CNC sur les outils de reconnaissance de contenus du 19 janvier 2021, l'importance d'améliorer la transparence sur les conditions d'application des algorithmes par les plateformes.

En effet, le paramétrage des algorithmes qui sont utilisés pour mettre en œuvre les règles conditionnent à la fois l'efficacité du système pour les ayants droit et son acceptabilité pour les utilisateurs. Une meilleure information sur les algorithmes mis en œuvre par les plateformes pour garantir la protection du droit d'auteur pourrait être encouragée. Cette transparence permettrait de vérifier que les outils et algorithmes utilisés ne comprennent pas de biais en la défaveur des ayants droit ou des utilisateurs. Le rapport présentait une proposition en ce sens visant à permettre au régulateur d'obtenir la communication d'informations de la part des plateformes.

Cet accès à des informations relatives aux algorithmes s'inscrit plus largement dans une régulation que l'on pourrait qualifier de « cyclique », et qui s'applique particulièrement bien aux contrôles des algorithmes, soit une régulation en trois temps : (1) le régulateur émet des recommandations relatives au fonctionnement des algorithmes, qui imposent notamment (2) des obligations de transparence aux acteurs régulés, ces derniers étant (3) audités et contrôlés par le régulateur en cas de difficulté, afin d'évaluer la bonne application des recommandations émises au préalable. À l'issue de ce processus, le régulateur peut émettre de nouvelles recommandations.

Des avancées législatives en ce sens peuvent déjà être constatées, comme l'a fait récemment l'ordonnance n°2021-580 du 12 mai 2021 transposant la directive droit d'auteur qui confie à l'Hadopi la mission de formuler des recommandations sur le niveau de transparence requis s'agissant des mesures de protection des œuvres mises en place par les fournisseurs de service de partage de contenus.

Le projet de loi relatif à la régulation et à la protection de l'accès aux œuvres culturelles à l'ère numérique prévoit à ce titre la possibilité pour l'Arcom de solliciter toutes informations utiles auprès des fournisseurs de service, des titulaires de droit et des concepteurs des mesures de protection.

En outre, le développement de l'intelligence artificielle invite à s'interroger sur la place et le rôle qu'elle occupe dans le processus créatif au regard des règles du droit d'auteur. Il importe notamment de savoir si une production générée par une IA relève de la protection du droit d'auteur ce qui implique de s'intéresser aux conditions d'éligibilité y afférentes (création, originalité, auteur). Si *a priori*, la conception française personnaliste pourrait exclure de la protection la production d'un contenu culturel généré par une IA, la doctrine semble conclure à l'absence d'obstacles dirimants tout en examinant les conditions d'une protection par un droit spécial ou encore un droit *sui generis* (« Intelligence artificielle et culture » A. Bensamoun et J. Farchy, 27 janvier 2020).

Enfin, il importe de rappeler que les enjeux de l'IA sont étroitement liés à ceux des données, puisque ces données nourrissent les systèmes d'IA, comme les données à caractère personnel ou les données issues d'un large corpus dans le cadre du *data mining*. En ce sens, la question de la protection des données à caractère personnel, en particulier dans les algorithmes de recommandation, est très

sensible notamment en ce qu'elle « *risque de produire des résultats discriminatoires fondés sur des données biaisées* », comme le relève la Résolution du Parlement européen du 19 mai 2021 précitée.

Question n°20 : Quels risques éthiques associez-vous à l'utilisation de l'IA pour l'exercice des missions qui vous sont confiées ?

Il ressort des différents travaux de l'Hadopi que le débat lié au développement de l'IA dans le secteur culturel est particulièrement clivant : entre les nouvelles possibilités de création et de diffusion, les évolutions sociétales induites par l'usage croissant des réseaux sociaux et les craintes d'enfermement des internautes, les questions éthiques sont nombreuses et complexes.

Les principales préoccupations semblent se cristalliser sur les algorithmes de recommandation à l'égard des internautes, qui influent sur la mise en valeur des contenus en fonction des individus et de leur volonté supposée de découverte.

Ces nouveaux outils sont présentés par les plateformes comme facilitant la circulation des œuvres et la découvrabilité des contenus, y compris les plus confidentiels. Ils pourraient ainsi répondre aux enjeux de diversité culturelle.

Néanmoins, comme pour tous les outils et usages innovants, les risques éthiques ne peuvent pas être pleinement écartés. Ainsi, parmi les critiques dirigés à l'encontre des outils de recommandations, il est notamment souligné le risque d'enfermement des consommateurs, de limitation de l'exposition effective des individus à la diversité culturelle et d'une offre culturelle neutralisée par la machine. En ce sens, la résolution du Parlement européen du 19 mai 2021 sur l'intelligence artificielle dans les domaines de l'éducation, de la culture et de l'audiovisuel « *invite la Commission à réaliser des études et à envisager des solutions stratégiques pour contrer l'effet néfaste du contrôle par l'IA des services de streaming en ligne conçus pour limiter la diversité et/ou maximiser les bénéfices tirés de l'inclusion de certains contenus, ou de la priorité donnée à ceux-ci, dans leurs offres aux consommateurs, et la manière dont cela nuit à la diversité culturelle et aux revenus des créateurs* ».

Ces évolutions technologiques répondent à une appétence des internautes tant elles leur permettent de se repérer dans l'immensité des offres de services sur internet et face à la multitude de contenus disponibles sur ces services. En effet, ces avancées technologiques permettent d'automatiser des traitements répétitifs à grande échelle, de trier des centaines de milliers de résultats ou bien de formuler des recommandations visant à mettre en valeur un contenu après avoir examiné des millions de décisions passées.

Par exemple, d'après l'observatoire de l'offre légale réalisée par l'Hadopi dont les résultats devraient être publiés prochainement, 84 % des internautes visionnent des séries recommandées en fonction de leur consommation et 57 % visionnent des films également en fonction de leur consommation. La recommandation personnalisée est ainsi devenue un argument concurrentiel et de différenciation pour les différents services.

Toutefois, il ressort un sentiment ambigu à l'égard de ces technologies, entre la recherche de simplicité et de sécurité d'usage et la méfiance, dès lors que ces technologies sont le plus souvent déployées de manière opaque par des acteurs en position dominante sur le marché. Ce sont également les biais de sélection et de classement des services et des œuvres qui – bien que correspondant aux souhaits des utilisateurs – en agrégeant les préférences émises par les utilisateurs lors de leur navigation sur ces plateformes – peuvent avoir une forte influence sur leurs usages et les enfermer dans une offre restreinte et un écosystème fermé.

Sur le plan de la sensibilisation aux usages numériques, il existe par ailleurs des risques d'accentuation des fractures numériques. L'utilisateur averti et aguerri peut jouer un rôle important dans son propre parcours de consommation, selon son degré de connaissance des services disponibles et leur possibilité de configuration. La question du paramétrage est ainsi devenue centrale, puisque pour une même commande, un service peut varier selon les réglages effectués. Or, nombre d'utilisateurs n'ont pas réellement conscience de l'existence des algorithmes de recommandations et se retrouvent « guidés » à leur insu, leur accordant même un haut niveau de confiance comme cela ressort de l'étude « assistants vocaux et enceintes connectées » précitée.

En outre, les dispositifs de recommandation, en participant à la mise en valeur des contenus sont susceptibles d'avoir une incidence sur la détermination de la rémunération de créateurs étant rappelé que les données produites par l'exploitation des œuvres représentent des actifs stratégiques pour les modèles économiques des acteurs de la filière.

Question n°21 : Si vous souhaitez compléter vos réponses et analyses ou porter d'autres informations à la connaissance de la mission :

Document en annexe (pièce jointe) : panorama du recours à l'intelligence artificielle pour la protection du droit d'auteur