

Communiqué de presse

Paris, 7 juillet 2026

IA : le GESTE lance un Observatoire des bots et crawlers pour mesurer l'exploitation des contenus des éditeurs et reprendre le contrôle sur la chaîne de valeur

Alors que les systèmes d'intelligence artificielle générative reposent de plus en plus sur l'exploitation massive de contenus éditoriaux, le GESTE lance, avec ses partenaires Botscorner et [255hex.ai](#), un **Observatoire de l'IA et des crawlers** destiné à objectiver les flux de collecte, identifier les acteurs impliqués et mieux comprendre les mécanismes de captation de valeur à l'œuvre.

L'Observatoire repose sur l'analyse des logs de connexion fournis par les éditeurs participants, croisés avec la base de connaissance de Botscorner permettant d'identifier les différents acteurs du trafic B2B : IA génératives, agents IA, fournisseurs de datasets, acteurs du SEO, de l'ad-tech, de la veille média ou encore du scraping industriel.

L'expertise d'Olivier Martinez, fondateur de [255hex.ai](#), vient compléter cette approche technique par une analyse comportementale et économique des usages observés. Au-delà de l'identification des acteurs et des flux, les travaux de l'Observatoire visent à mieux comprendre les comportements des bots, agents et crawlers, leurs interactions avec les audiences humaines, ainsi que leurs impacts sur les modèles d'audience, la monétisation publicitaire, les infrastructures techniques et, plus largement, sur l'équilibre économique de l'écosystème numérique.

L'ambition du GESTE est ainsi de dépasser la seule mesure du trafic automatisé pour documenter les transformations profondes induites par l'essor des usages de l'IA sur les modèles économiques des éditeurs et sur le partage de la valeur entre producteurs de contenus, intermédiaires techniques et acteurs de l'intelligence artificielle.

Les premiers résultats sont sans appel : **près de la moitié du trafic observé sur les sites de presse est aujourd'hui généré par des acteurs automatisés**. Derrière les moteurs de recherche traditionnels et quelques grands modèles d'IA identifiés se déploie un écosystème beaucoup plus vaste et beaucoup plus opaque : scrapers anonymes, brokers d'adresses IP, plateformes de datasets, outils de veille, agents IA, services de monitoring, intermédiaires techniques et places de marché de données.

Ces acteurs collectent, structurent, enrichissent, revendent ou réutilisent des contenus éditoriaux qui constituent pourtant le cœur de la valeur créée par les éditeurs. Une partie croissante de cette valeur alimente désormais des chaînes d'approvisionnement complexes destinées à entraîner des modèles, alimenter des services d'intelligence artificielle ou développer de nouveaux produits commerciaux, le plus souvent sans visibilité, sans contractualisation et sans rémunération pour les producteurs de contenus.

L'Observatoire met également en évidence les limites des mécanismes historiques de gouvernance du web. Seule une faible proportion des bots consultent les fichiers robots.txt,

et leur consultation ne préjuge en rien du respect des règles exprimées par les éditeurs. Dans le même temps, le recours croissant à des infrastructures d'anonymisation, à des proxys rotatifs ou à des intermédiaires spécialisés rend l'identification des utilisateurs finaux toujours plus difficile.

Selon Yan Gilbert fondateur de Botscorner : *« Depuis plus de vingt ans, des marchés organisés de la donnée et des datasets se sont développés autour des contenus produits par les éditeurs. L'IA générative accompagne ce phénomène de la collecte à l'utilisation, en augmentant les usages commerciaux associés. Avant de pouvoir penser à des tarifs, négocier des licences, mettre en place des mécanismes de rémunération ou définir des stratégies de protection, les éditeurs doivent d'abord savoir qui collecte leurs contenus, par quels procédés, dans quelles quantités et pour quels usages. »*

L'objectif de l'Observatoire est double :

- fournir aux éditeurs une vision précise des usages de leurs contenus ;
- constituer une base factuelle commune permettant d'alimenter les discussions sur les modèles de licence, les mécanismes de rémunération et les évolutions réglementaires nécessaires à un partage plus équilibré de la valeur créée par l'IA.

Pour le GESTE, la question n'est désormais plus de savoir si les contenus des éditeurs sont utilisés par les systèmes d'IA, mais de déterminer dans quelles conditions économiques, juridiques et techniques cette utilisation peut être encadrée.

À propos du GESTE

Le GESTE rassemble les principaux éditeurs de contenus et services numériques français et défend les conditions d'un écosystème numérique ouvert, innovant et équilibré.

A propos de Botscorner

Botscorner analyse le trafic B2B depuis 2017. Nos clients se trouvent en France, au Canada, en Allemagne.

Chaque jour, des milliers de robots extraient les données sur les sites médias : podcasts, photos, vidéos, textes, publicités.

Ils structurent le big data et proposent les bases de données nécessaires à leurs activités (datasets, IA, veille, analyses, etc.)

Botscorner est destiné aux services commerciaux, juridiques et techniques des médias et des bases de données.

A propos de 255hex.ai

255hex.ai accompagne les entreprises et leurs dirigeants, notamment dans les secteurs des médias et de la communication, dans la compréhension, l'adoption et l'intégration concrète des outils d'intelligence artificielle, ainsi que dans la définition de stratégies digitales adaptées à leurs enjeux, leurs métiers et leur organisation.