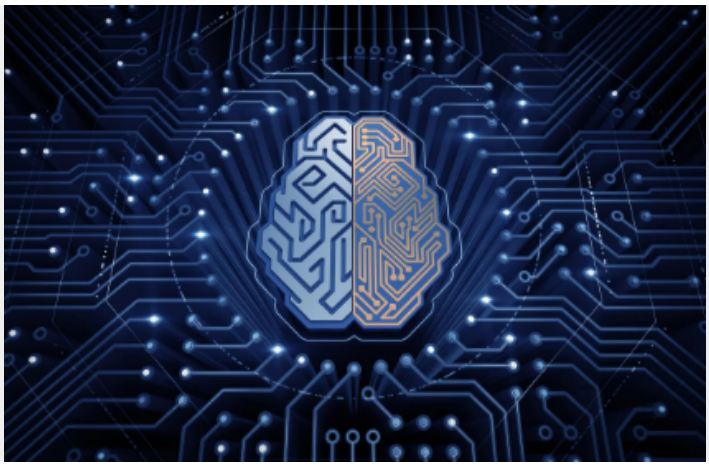
THOMSON REUTERS C. ROSS INTELLIGENCE : UNE AFFAIRE CLÉ EN MATIÈRE DE DROIT D'AUTEUR ET D'IA

Posted on mars 18, 2025 by Philippe Gilliéron



Le 11 février 2025, le Tribunal de district des États-Unis pour le district du Delaware a statué dans l'affaire *Thomson Reuters et al. c. Ross Intelligence, Inc.* (No. 1:20-cv-613-SB) que la reproduction d'œuvres protégées par le droit d'auteur pour l'entraînement d'une IA ne relevait pas de l'exception de fair use. Cette décision a des implications majeures pour la recherche juridique assistée par IA et le cadre plus large de la propriété intellectuelle.

Contexte de l'affaire : Un Conflit de Copyright dans la LegalTech

Ross Intelligence, un concurrent de Westlaw, a développé un moteur de recherche juridique basé sur l'intelligence artificielle. Pour entraîner son système, Ross avait besoin d'une base de données de questions-réponses juridiques. Après que Thomson Reuters eut refusé de lui accorder une licence, Ross s'est tourné vers LegalEase, qui a fourni des documents d'entraînement (*Bulk Memos*). Cependant, ces mémos étaient indirectement issus des résumés de jurisprudence de Westlaw, ce qui a conduit Thomson Reuters à engager des poursuites pour violation du droit d'auteur.

Problématiques juridiques clés sur le droit d'auteur et l'IA

Le Tribunal a examiné deux questions fondamentales :

1. Ross a-t-il directement enfreint le droit d'auteur ?
2. Si oui, pouvait-il invoquer la défense du fair use conformément à l'article 17 USC § 107 ?

Violation du droit d'Auteur : L'examen des données d'entraînement de l'IA

Le Tribunal a conclu que l'entraînement du modèle d'IA de Ross remplissait les trois conditions

retenues en droit américain pour admettre une violation directe du droit d'auteur :

- **Titularité** : Thomson Reuters détenait des enregistrements valides de droits d'auteur.
- **Originalité** : Les résumés de jurisprudence répondaient au seuil d'originalité (*Feist*, 499 US 345).
- **Reproduction** : les *Bulk Memos* (dont le 100% était utilisé que ce soit à des fins d'entraînement ou de validation) ressemblaient davantage aux résumés de Westlaw qu'à des décisions judiciaires, soit des oeuvres protégées par le droit d'auteur dont la reproduction en violait les droits.

L'exception de "fair use" : Un argument rejeté pour l'entraînement de l'IA

Le Tribunal a examiné les quatre facteurs pour apprécier de l'existence d'un cas de "fair use" :

1. **Finalité et nature de l'utilisation** : L'outil d'IA de Ross était conçu à des fins commerciales et son usage des contenus protégés n'était pas transformateur (*Warhol*, 598 US 508).
2. **Nature de l'œuvre protégée** : Les résumés n'étaient pas particulièrement créatifs, ce qui favorisait Ross.
3. **Quantité et importance de l'extrait utilisé** : Ross n'avait pas reproduit les résumés en accès public (soit comme résultats) mais uniquement pour entraîner le système, un point en sa faveur.
4. **Effet sur le marché** : L'outil développé par Ross entrait en concurrence avec Westlaw, ce qui excluait le "fair use" dès lors que le lancement d'un produit concurrent était à l'évidence susceptible d'avoir un effet sur le marché.

Le Tribunal a également rejeté l'argument de Ross selon lequel la copie intermédiaire (comme dans *Google c. Oracle*, 593 US 1) justifierait le "fair use", puisque cette exception n'avait été admise qu'en raison du fait que la reproduction du code l'avait été pour permettre l'interopérabilité entre les systèmes. Or, ce souci d'assurer l'interopérabilité était inexistant dans le cas d'espèce, excluant une

utilisation à des fins différentes (qui aurait permis de reconnaître le caractère transformateur).

Conséquences pour l'IA, le "fair use" et le droit d'auteur

Cette décision peut être mise en parallèle avec celle rendue dans l'affaire LAION en Allemagne. En septembre 2024, la Cour de Hambourg avait jugé que la constitution d'un ensemble de données IA basé sur des œuvres protégées était un usage scientifique permettant de se prévaloir de l'exception de texte et fouille de données (TDM). Le Tribunal du district du Delaware en a décidé différemment dans l'affaire *Thomson Reuters c. Ross*, en mettant en évidence une distinction clé :

- Si un outil d'IA entre en concurrence avec l'œuvre protégée d'origine, le "fair use" ne s'applique pas.
- Si un modèle d'IA a des applications plus larges et non concurrentielles, l'argument du "fair use" est plus recevable.

Vers de nouvelles batailles juridiques sur l'IA et le droit d'auteur

Le Tribunal a précisé que cette décision ne concernait pas directement les IA génératives (GenAI), soulevant ainsi la question de savoir si l'entraînement des modèles d'IA sur des données protégées pourrait constituer un cas de "fair use". A mon sens, de deux choses l'une: soit l'outil de GenAI est suffisamment spécifique pour constituer une alternative aux oeuvres concernées, soit il est général et ne peut dès lors être considéré comme un produit concurrent. Dans le premier cas, l'exception de "fair use" ne serait pas applicable alors qu'elle pourrait l'être dans le second.

Points clés pour l'IA et le droit d'auteur

1. L'entraînement des IA avec des contenus protégés ne relève pas automatiquement du "fair use".
2. Les modèles d'IA juridique entrant en concurrence avec des bases de données propriétaires (et entraînées sur de telles bases) s'exposent à des risques de violation du droit d'auteur.
3. Si le produit final une fois entraîné entre en concurrence avec les oeuvres utilisées pour cet entraînement, le "fair use" est peu probable.
4. L'avenir juridique des IA génératives reste une question ouverte.