Response to the Public Consultation Paper:
Copyright and Artificial Intelligence

Our Perspectives on Copyright Challenges in AI-Generated Works

Prepared on September 8, 2024

Table of Contents

1. About Votee Limited and Our Commitment to Copyright

Votee Limited is a Hong Kong-based startup established in 2013, specializing in AI-driven consumer intelligence. Over the years, we have gathered primary data from consumers across Southeast Asia, leveraging this information to develop our proprietary AI algorithms. Our expertise includes automated speech recognition, optical character recognition, and sentiment analysis, with a strong focus on Cantonese language processing. This has positioned us as a leader in our field.

In 2018, we expanded our offerings to include enterprise-level generative AI solutions, providing customized AI solutions and consulting services to a diverse range of organizations, currently serving over 200 clients, including the Hong Kong government, financial institutions, management consulting firms, and FMCG brands.

What distinguishes us is our comprehensive approach to AI projects. We deliver full-stack AI solutions that encompass application layers, platform layers, and infrastructure layers. After more than a year of collaborative research and development with field experts and academia, we are excited to announce the upcoming launch of our curated community version of a Cantonese Large Language Model (LLM) in June 2024. We are committed to further developing our Cantonese models and exploring low-resource language models, enhancing our capabilities to provide advanced and tailored solutions.

Additionally, our expertise in LLM security sets us apart. We have successfully served clients in Hong Kong's government and financial services sectors, which require stringent data security measures. Our in-house LLM security experts address these needs, offering a level of service that many competitors do not provide.

At Votee, we recognize the importance of copyright issues in the ever-evolving landscape of AI-generated works. As a leader in AI innovation, we care deeply about ensuring that both creators and users are protected and respected. We believe that establishing clear guidelines and frameworks surrounding copyright issues is essential for fostering creativity and innovation while safeguarding intellectual property rights. By contributing to the discourse on copyright in AI, we aim to support a balanced approach that encourages growth, collaboration, and ethical use of AI technologies.

This consultation response has been crafted by a team of seasoned LLM experts and AI engineers, all of whom are integral members of our AI/LLM research team. In addition, we are proud to have team members with backgrounds in business law, possessing formal education in copyright and intellectual property. This diverse expertise allows us to provide a comprehensive and nuanced perspective on these vital issues.

2. Copyright protection of works generated by generative AI ("AI-generated works")

2.1 Evaluating Copyright Protection for AI-Generated Works under the CO in Hong Kong

The overarching objective of the CO is to foster creativity and investment in artistic endeavors while balancing the rights of copyright holders with public interest. In the context of AI-generated works, it is equally important to promote the advancement of AI technology alongside creativity as it contributes to the public interest.

Whether the current provisions are deemed sufficient will hinge entirely on the objectives of this review. The following content assesses whether our existing copyright framework can effectively navigate the evolving landscape of generative AI and address the emerging challenges.

2.1.1 Assessment of Current Protections and Their Impact

AI-generated works are protected under the current CO, covering both literary, dramatic, musical, and artistic (LDMA) works and non-LDMA works. While it is clear that the current provisions recognize AI-generated works as eligible for copyright protection, several issues within the existing provisions fail to meet the overarching objectives of the CO:

(A) Issues with Originality Requirements

(i) Ambiguity of the Originality Requirement in Protecting CG LDMA Works

We believe that current provisions are insufficient to CG LDMA works, potentially discouraging creators from producing new content with Generative AI models.

The cornerstone of the CO requiring originality in LDMA works is to ensure that copyright protection is granted only to creations that exhibit a sufficient level of creativity and independent effort. However, Generative AI models are capable of producing deterministic outputs through specific configurations, such as setting a fixed random seed, adjusting the temperature parameter to zero, and employing greedy decoding. These controls diminish the inherent randomness normally associated with Generative AI models, resulting in predictable and reproducible outcomes:

- **Creativity Concerns:** When outputs are deterministic, the creative contribution from the user may be minimal.
- **Similarity Among Works:** Multiple authors can generate similar works using the same configurations, challenging the originality threshold. The law may struggle to protect fundamentally similar outputs.
- **Increased Legal Disputes:** Deterministic outputs may result in more copyright disputes, requiring legal clarification on protectable works.

(ii) Inclusion of Original Works in Generative AI Models

Copyright exists in non-LDMA works as long as they are not copies of previous works, with no originality requirement for protection. However, the inclusion of original works in the training dataset raises significant concerns, applicable to CG LDMA works as well.

This situation prompts questions about whether AI-generated outputs may inadvertently replicate elements of the original training data, even if they do not qualify as direct copies. The proximity of these original works within the training dataset could affect the originality of the resultant outputs.

If AI-generated outputs can be shown to be transformative—meaning they introduce new expression or meaning—this bolsters the argument for fair use. However, demonstrating this transformation poses challenges for creators. The underlying mechanics of generative AI models are often opaque, making it difficult to ascertain how specific outputs are produced.

The impact of copyright concerns on CG LDMA and non-LDMA works lies in the potential legal risks associated with inadvertently replicating elements of original training data, complicating the demonstration of transformative use and hindering innovation and creativity.

(iii) Assessing the Originality Requirement for AI-Generated Works

Some copyright jurists suggest that the skill, labor, and judgment of the person arranging the creation of the work should be considered, or they propose evaluating whether the AI-generated work involves sufficient skill, labor, and judgment to meet the originality threshold as if it had been created by a human. However, this perspective raises concerns.

Requiring that copyright protection depends on the arranger's level of skill, labor, and judgment risks sidelining the contributions of users who leverage AI tools. This approach can discourage exploration and limit the potential of generative AI as a medium for creative innovation.

(iv) Discrepancies in Protection of CG LDMA and CG non-LDMA Works

The CO recognizes a fundamental distinction between CG LDMA works and CG non-LDMA works. While CG LDMA works are subjected to shorter copyright protection durations and more restrictive moral rights than ordinary LDMA works due to the perceived lower effort involved in their creation. This rationale purportedly aligns with the CO's overarching goal: to reward creators in proportion to their contributions.

However, the same protective framework does not apply to CG non-LDMA works, which are accorded copyright protection equal to their traditional counterparts, regardless of their creation process. This discrepancy raises important questions regarding equity and justification.

While some may perceive CGWs (whether LDMA or non-LDMA) as requiring less effort, this viewpoint may disregard the complexity and creativity involved in digital creation. Awarding lesser protection could discourage innovative approaches by digital creators who already face skepticism regarding the value of their work.

The CGWs provisions state that the author of a CG LDMA work is characterized as the "necessary arranger." In the context of AI-generated works, ambiguity arises regarding who qualifies as the necessary arranger—whether it be the developer, operator, or user of the AI system. This uncertainty complicates the determination of authorship and copyright ownership.

To shed light on who qualifies as the "necessary arranger" for AI-generated works, the producer analogy proves to be beneficial:

- **Determining Authorship:** The analogy highlights that individuals who facilitate the creation of a work may retain rights, even if their contributions are not traditionally creative in nature.
- **Dispute Resolutions:** Courts may draw upon the producer analogy during disputes over ownership, considering the various roles played by the parties involved in the creation of CGWs.

The recognition of these roles will be crucial in resolving disputes and clarifying ownership:

- **Developer/Programmer:** Similar to a producer, the developer/programmer of the AI may make necessary arrangements for the ability of the AI to generate creative outputs. Their contributions could mirror the organizational aspects of a producer.
- **Operator of AI System:** The operator, like a producer managing a team, might control the system under which creative generation occurs without directly creating the output.
- **User Input:** When a user prompts the AI model, their input may parallel the role of a creative contributor (e.g., writer or director), providing the essential direction intended for the output.

2.1.2 Fact-Specific and Case-Sensitive Evaluations

The interpretation of the originality requirement for CGWs is poised to evolve through the development of case law, mirroring the historical shaping of copyright principles for traditional LDMA works. This evolution permits a flexible adaptation to rapid advancements in technology and creative methodologies.

While we acknowledge that copyright issues concerning CG works are inherently fact-specific and case-sensitive, the absence of a coherent framework may jeopardize Hong Kong's ambitions in advancing AI and nurturing creativity. Several key points underscore the need for clarity:

- **Navigating Uncertainties:** The lack of a defined framework creates uncertainties for creators, investors, and technologists, deterring investment and innovation.
- **Fostering Innovation:** A clear copyright framework improves understanding of rights and obligations, encouraging collaboration between technologists and traditional creators. This synergy can lead to groundbreaking projects that integrate human creativity with AI.
- **Enhancing Hong Kong's Competitive Edge:** Clarifying copyright protections for CG works can enhance Hong Kong's competitiveness, attracting talent and investment as a hub for technological and artistic innovation.
- **Encouraging Responsible Use of AI:** A defined legal structure helps creators navigate ethical considerations, promoting responsible innovation by clarifying fair use and engagement with existing works.

2.1.3 Exploring Copyright Protection for User Prompts in Generative AI

As the Copyright Office (CO) considers granting copyright protection for outputs generated by Generative AI models, it's essential to evaluate the role of user inputs—specifically, user prompts—in shaping these results. User prompts provide critical direction and creativity, significantly influencing the nature of the final AI-generated work. Therefore, we must assess whether these prompts should be eligible for copyright protection under the same criteria applied to traditional literary and dramatic works (LDMA). This discussion not only addresses the intricate dynamics of creative input and output but also raises important questions about authorship and ownership in the rapidly evolving landscape of AI technologies.

In a conversational interface, user prompts can vary widely—ranging from unique, highly creative phrases to generic, commonly-used expressions. This variability complicates the assessment of which prompts significantly contribute to the AI-generated output. Determining the specific impact of each prompt within a multi-turn conversation presents a challenge. It is often difficult to isolate which user input was instrumental in shaping a particular output, especially when the interaction is iterative and context-dependent.

For copyright eligibility, prompts must meet a standard of originality. However, assessing originality can be challenging when prompts are commonplace or lack a distinctive character. Establishing a clear threshold for what constitutes an original prompt within the Generative AI context will be essential.

Under the current CO, if user prompts are deemed protected, there exists an inherent risk of potential infringement for AI systems. The ambiguity surrounding the originality requirement complicates the assessment of AI-generated outputs, leaving developers and system operators vulnerable to legal repercussions. Without clear guidelines on what constitutes sufficient originality in the generated content, AI creators may inadvertently infringe upon copyright, exposing themselves to serious legal consequences. This uncertainty not only stifles innovation in AI development but also raises significant concerns about liability and the future of creative technologies in Hong Kong.

Developers are particularly vulnerable under the current copyright framework because generating outputs inherently relies on user prompts, which may be protected by copyright. In scenarios where both the user prompts and the AI-generated outputs are utilized in subsequent training of the AI model to enhance its performance, the risk of infringement escalates. This reliance creates a cycle where copyrighted user prompts could be inadvertently incorporated into the training data, leading to further outputs that potentially replicate or derive from the original prompts. Consequently, developers may find themselves facing serious legal liabilities for copyright infringement, as the lack of clarity regarding originality in the generated content leaves them exposed to claims from copyright holders, thereby hindering innovation and responsible AI development.

While implementing terms that require users to grant necessary licenses for the AI system to generate responses or improve its functionality may address potential copyright infringement issues, it simultaneously poses significant challenges for creators. Such provisions may deter individuals from utilizing AI tools to enhance their works, as they may be reluctant to relinquish rights or feel uncertain about the implications of granting licenses. This hesitance can stifle creativity and innovation within the creative industry, ultimately hindering the development and evolution of artistic expression. A balance must be struck between protecting copyright interests and fostering an environment where creators feel empowered to leverage AI as a collaborative tool in their creative processes.

As an AI company based in Hong Kong, we oppose copyright protection for user prompts in Generative AI for several reasons:

- **Community Knowledge Exchange:** In the tech community, prompts are often exchanged to promote learning and skill advancement among prompt engineers and AI developers. This collaborative spirit is vital for fostering innovation and improving the tools available to a larger audience.
- **Plain Language:** Unlike programming code, which is a specific language governed by structured syntax, prompts are typically expressed in plain English. They serve as intuitive requests or queries to the AI rather than distinct creative works. Imposing copyright protection on such expressions would impede the fluidity of knowledge sharing.
- **Hindrance to Innovation:** Allowing copyright protection on user prompts could severely slow down the development of AI technologies. Developers and engineers would face hurdles in accessing and utilizing existing prompts for research and development purposes. This stifling of creativity could lead to a stagnation in the evolution of AI capabilities.
- **Advocating for Open Access:** An open environment where prompts can be freely shared fosters a culture of experimentation and advancement. It encourages diverse contributions that can lead to groundbreaking developments in AI and related fields.

We recognize that artists and creators may seek to copyright user prompts as part of their products. This desire stems from the importance of preserving the unique expressions and creativity contributing to their work. Balancing the rights of creators with the need for openness in prompt sharing is a complex challenge that requires careful consideration.

2.2 Consideration of the Need for Statutory Enhancements

The overarching objective of copyright law is to foster creativity and investment in artistic endeavors while balancing the rights of copyright holders with the public interest. However, this objective appears to be incompatible with the rapid development of AI, which serves to benefit the public. As we grant increasing copyright protections to the creative industry, we inadvertently hinder the growth of AI.

Under the current framework, if we do not implement statutory enhancements to clarify the protection criteria for CGWs, it seems that there is a prioritization of AI development over the creative industries. Copyright law is fundamentally designed to reward human creativity; however, as AI takes on a larger role in content creation, the rationale for rewarding human creators diminishes.

Given this premise, we agree that the existing provisions sufficiently cover the protection of AI-generated works. The terms are broad enough to allow for interpretation, and case-by-case analysis appears to be adequate and fair.

However, this lack of clarity poses challenges not only for AI-generated works but also for AI-assisted artists, as there are no clear guidelines on qualifying for copyright protection. While market contracts provide some relief, public education is crucial to safeguarding the creative industry. Artists contribute significantly to AI development through their domain expertise, and it is essential that they are equipped to protect their work.

We propose that we establish guidelines instead of statutory enhancements to educate individuals who wish to have their CGWs protected by copyright. This should include instruction on how to document their creative processes when using AI, thus demonstrating originality. Additionally, it is imperative that they understand copyright terms and how to navigate the platforms they use as tools, enabling them to better protect themselves and assert their rights.

2.3 Our Experience in Copyright Claims for AI-Generated Works

We have not needed to rely on the CGWs provisions of the Copyright Ordinance for AI-generated works, so we have no relevant experiences to share.

2.4 Evaluating Contractual Arrangements as Practical Solutions

We advocate for market self-regulation in the rapidly evolving landscape of AI development. However, we recognize the critical need for public education to ensure this approach functions effectively:

- **Informing Creators of Rights:** Many creators lack a comprehensive understanding of their copyright rights. By providing educational resources, we empower them to protect their intellectual property when entering into contracts.
- **Understanding Contracts:** Contracts for AI-generated content can be complex. Educational programs can help creators comprehend terms and identify potential issues.
- **Best Practices for Negotiation:** Training on negotiation can lead to fairer terms. Understanding what rights to retain and how to negotiate royalties is crucial.
- **Fostering Ethical Standards:** Education can promote responsible AI use, ensuring creators understand the implications of AI technologies on their rights.

3. Copyright infringement liability for AI-generated works

3.1 Evaluation on Existing Legal Position and Potential Copyright Infringement

As the Copyright Ordinance recognizes computer-generated works (CGWs) as eligible for protection under existing provisions, the burden of infringement increasingly falls on AI developers and operators. This raises significant questions regarding accountability for the outputs of AI models, regardless of whether such outputs are copyright protected. Determining who should bear responsibility for AI-generated content remains a complex and unresolved issue.

If an AI model generates content that inadvertently replicates or derives from existing works, responsibility is likely to fall on the AI developer or operator. This discussion can be divided into two key aspects:

- **Use of Copyrighted Material in AI Model Training:** If the developer has incorporated copyrighted material into the model's training data, they should bear some responsibility for any resulting copyright infringement. This accountability hinges on the data collection and cleaning processes implemented by the developer. If these processes failed to adequately filter out copyrighted materials, it raises questions about the developer's diligence in ensuring compliance with copyright laws.
- **Absence of Copyrighted Material in AI Model Training:** If the developer has not used copyrighted material in the training process, the outputs of the AI can still be deterministic based on how the model is configured. In such cases, it becomes challenging to assign responsibility to the developer for the generated content. Given that the model operates based on algorithms and user prompts, holding the developer accountable for every output may not be reasonable, particularly when the model is functioning as intended.

The uncertainty surrounding liability for AI-generated content hinders AI development in several ways:

- **Deterrence of Innovation:** Fear of legal repercussions may prevent developers from exploring new algorithms and training methods.
- **Increased Compliance Costs:** Developers face higher expenses related to legal consultations, stringent data processes, and acquiring licenses.
- **Limited Data Access:** Caution in sourcing training data can reduce the diversity and quality of datasets, impacting output effectiveness.
- **Chilling Effect on Collaboration:** Uncertainty may inhibit partnerships between AI developers and content creators, limiting innovation opportunities.
- **Impediments to Commercialization:** Startups may struggle to enter the market due to perceived copyright risks, reducing competition.
- **Erosion of Public Trust:** Concerns about AI generating infringing content can diminish public confidence and adoption of AI technologies.

As mentioned in Section 2.1.2, the protection of user prompts adds another layer of complexity to the discussion on liability for AI-generated content. User prompts significantly influence the outputs of AI models, raising questions about authorship and ownership. If user prompts are deemed copyrightable, developers may face additional challenges related to potential infringement claims, especially if they incorporate such prompts into their training data.

3.2 Evaluation on Existing Legal Position and Potential Moral Rights Infringement

Under the current provisions of the CO, several concerns arise regarding the existing legal position and potential moral rights infringement:

- **Ambiguity in Moral Rights:** The Copyright Ordinance recognizes moral rights, including attribution and integrity. AI-generated content that resembles existing works raises concerns about infringing original creators' moral rights, particularly if alterations harm their reputation. The deterministic nature of AI outputs complicates this, leaving developers, operators, and users uncertain about how to protect themselves. Without clear guidelines on moral rights in the context of AI, all parties face legal uncertainties regarding attribution and integrity claims.
- **Attribution Challenges:** The lack of clear authorship for AI-generated content complicates the attribution of moral rights. If an AI model generates a work that is derivative of multiple sources, identifying the original creators and properly attributing their contributions becomes problematic.
- **Integrity Rights Concerns:** The potential for AI-generated outputs to distort or misrepresent the original works can lead to violations of integrity rights. If an AI-generated piece is perceived as derogatory or damaging to the original creator's reputation, it could result in moral rights infringement claims.
- **User Prompts and Moral Rights:** As discussed in previous sections, if user prompts are considered protected, there is a risk that user moral rights may also come into play. Users may seek recognition for their contributions, complicating the relationship between AI developers and users.

The existing legal framework does not adequately address the implications of AI on moral rights, leading to uncertainty for both creators and developers. This ambiguity can result in disputes and potential legal liabilities, further hindering AI development.

3.3 Consideration of the Need for Statutory Enhancements

As outlined in Section 2.2, we acknowledge that the current provisions, while broad and flexible, fail to provide clear guidance for concerned parties on how to protect their rights in the context of AI-generated works. Instead, we rely on a fact-specific, case-by-case analysis framework, allowing case law to evolve over time.

We recognize that amending the law could introduce chaos, particularly given the rapid pace of technological advancement and the uncertainty it brings. However, we maintain our position that issuing official guidelines is essential. These guidelines should convey best practices, instilling confidence in both the AI and creative industries to effectively navigate the use of AI technologies.

3.4 Our Experience in Legal Claims for Copyright Infringement

We have not needed to rely on the CGWs provisions of the Copyright Ordinance for AI-generated works, so we have no relevant experiences to share.

3.5 Evaluating Contractual Arrangements as Practical Solutions

The ambiguity of current provisions may lead businesses and AI platforms to implement stricter terms to protect themselves against potential claims. This trend could discourage users from producing and commercializing AI-generated works, as they may feel inadequately rewarded for their contributions.

When platforms impose harsher terms, users might perceive the risks as outweighing the benefits, ultimately stifling creativity and innovation. If users are uncertain about their rights or the potential financial returns from their work, they may be reluctant to engage fully with AI technologies. This unintended consequence not only hampers individual creativity but also limits the overall growth of the AI-driven creative economy.

Clearer guidelines and more equitable terms are essential to foster an environment where users feel empowered to collaborate with AI and reap the rewards of their efforts.

4. Possible introduction of specific copyright exception

4.1 Whether to support the Proposed TDM Exception in the CO

As an AI solution company in Hong Kong, we fully support the introduction of the Proposed TDM Exception into the CO. As discussed in Section 2.2, the challenge of balancing the interests of creators with the needs of AI development is inherently contradictory; the advancement of one often comes at the expense of the other. While we deeply value and respect the contributions of creators, we advocate for a balanced solution that protects rights on both sides, which is why we endorse the TDM exception.

The justifications outlined in the consultation document are well-articulated. Additionally, we believe that the TDM exception would directly address concerns raised in Section 2.1.3. Under the current CO, AI developers face potential copyright infringement liability even when utilizing open and publicly available information for training purposes. Implementing the TDM exception would provide much-needed clarity and legal protection, fostering an environment conducive to innovation and the responsible development of AI technologies. By alleviating these legal uncertainties, we can encourage greater investment in AI research and development, ultimately benefiting both the creative and technology sectors.

4.2 Impact of The Proposed TDM Exception

In our experience, due to copyright restrictions, many materials cannot be included in our model training, limiting the effectiveness and scope of our AI applications. This situation not only hinders our ability to create robust models but also restricts innovation and competitiveness in the industry. The Proposed TDM Exception could help overcome these challenges by facilitating business and industry development in several ways:

- **Reduction in Licensing Complexity and Costs:** Obtaining licenses from multiple copyright owners for large datasets is often complex and costly. The Proposed TDM Exception would eliminate the need for extensive rights clearance, reducing transaction costs and administrative burdens, allowing businesses to focus resources on innovation.
- **Legal Certainty and Risk Mitigation:** The ambiguity surrounding the use of copyrighted materials for TDM creates risks of inadvertent infringement, deterring engagement in TDM activities. A clear legal framework from the Proposed TDM Exception would provide legal certainty and reduce litigation risks, encouraging investment in TDM projects.
- **Enhanced Access to Diverse Data:** Access to a wider range of copyrighted works, including text, images, and multimedia, is essential for training effective AI models. The Proposed TDM Exception would facilitate this access, benefiting industries that rely on large datasets, such as healthcare, finance, and marketing.
- **Support for Research and Development:** The exception would enable academic and industrial researchers to conduct TDM activities without fear of infringement, leading to comprehensive studies and technological advancements, such as identifying new drug candidates in pharmaceuticals.
- **Promotion of Collaborative Projects:** The Proposed TDM Exception would streamline access to copyrighted materials for collaborative projects involving various stakeholders, fostering open innovation and cross-disciplinary research.

4.3 The availability and effectiveness of copyright licensing for TDM activities

Is copyright licensing commonly available for TDM activities? If so, in respect of which fields/industries do these licensing schemes accommodate? Do you find the licensing solution effective?

Copyright licensing for TDM activities is less prevalent than expected. For example, when we sought licenses from major newspaper providers to use their published articles as training material, we faced considerable variability in costs and restrictive terms that hinder effective AI model training. These licenses are primarily designed for those looking to republish articles, often imposing limitations on the number of requests made to the organization's server for data retrieval over a specified timeframe. However, efficient model training necessitates access to data in bulk rather than in fragmented increments.

Additionally, establishing contact with the appropriate personnel within organizations to discuss copyright matters can be challenging. We frequently find ourselves communicating with PR or media teams, who often lack the necessary legal or technical expertise to engage in meaningful discussions. This disconnect can hinder negotiations for exceptions, as they may not fully understand our requests or their implications.

In other industries, particularly those involving domain-specific data, obtaining licenses can be challenging. Beyond concerns related to data security and proprietary information, which may include personal identifiable information, there is often significant scrutiny regarding the intended use of the data. Even when organizations are open to granting licenses, we frequently encounter numerous questions and hesitations about how the data will be utilized. This necessitates considerable effort in educating and persuading stakeholders about the legitimacy and value of our intended applications, making the licensing process more complex and time-consuming.

Consequently, the process of obtaining relevant licenses is often frustrating and inefficient.

4.4 Conditions for the Proposed TDM Exception

The Proposed TDM Exception should incorporate several key conditions to effectively balance the interests of copyright owners and users while serving the best interests of Hong Kong:

- **Lawful Access:** Users must obtain lawful access to the copyrighted materials being mined, ensuring that these materials are acquired through legitimate means and respecting copyright owners' rights.
- **Opt-Out Mechanism:** Copyright owners should have the option to opt-out of the exception, possibly by marking their works with a machine-readable code to indicate their preference. This empowers copyright owners to control whether their works can be utilized in TDM activities.
- **Purpose Limitation:** The exception should explicitly state that copied materials are used solely for TDM activities and not for redistribution or commercial exploitation beyond the scope of analysis.
- **Security and Confidentiality:** Data extracted and processed should be stored securely and used confidentially to prevent misuse and unauthorized access.
- **Retention Period:** The duration for which extracted data can be retained should be limited, with requirements for its destruction after TDM activities are completed or after a specified timeframe.

- **Acknowledgment and Attribution:** When feasible, users should be required to acknowledge the source of the copyrighted materials, maintaining a connection between the work and its creator.
- **Commercial and Non-Commercial Distinction:** The exception should differentiate between non-commercial and commercial TDM activities, potentially imposing stricter conditions on commercial uses to ensure fair compensation for copyright owners.

Practical Difficulties:

- **Complexity and Compliance:** Ensuring compliance with lawful access and opt-out mechanisms can be complex and burdensome, particularly for smaller entities or individual researchers.
- **Security Measures:** Implementing and maintaining robust security and confidentiality protocols may require significant resources.
- **Attribution Challenges:** In large-scale TDM activities, providing detailed acknowledgments for every piece of data used may be impractical.
- **Retention and Destruction:** Managing data retention and timely destruction can introduce additional administrative overhead.

Balancing these conditions necessitates careful consideration of the practical implications for both users and copyright owners, fostering an environment that supports innovation while respecting intellectual property rights.