

## GENERATIVE AI AND COPYRIGHT

### Part 2: Computational Data Analysis Exception and Fair Use

#### [2023] SAL Prac 25

Generative artificial intelligence (“AI”) applications require large quantities of authorial works for machine learning purposes in order to produce the appropriate outputs in response to human text prompts. Some of the outputs may substitute for the market demand for the original works or reduce their licensing potential. Part 2 of this two-part article on copyright and generative AI discusses how the computational data analysis exception and fair use provision in the Singapore Copyright Act 2021 may be relevant to understanding liability for copyright infringement in input and output scenarios.

David **TAN**

*Professor, Faculty of Law, National University of Singapore;  
Co-Director, Centre for Technology, Robotics, Artificial Intelligence & the Law;  
Head (Intellectual Property), EW Barker Centre for Law & Business.*

#### **I. Introduction**

1 Part 1 of this two-part article discusses how Singapore copyright law is poised to tackle the issue of whether the use of copyright-protected works for machine learning (“input”) and the works created from natural language commands (“output”) are infringing copyright. It concludes that it is generally difficult to prove wholesale copying of millions of works as the various generative artificial intelligence (“AI”) applications (“GAIAs”) do not disclose the training datasets, and one would have to proceed on a classic substantial similarity analysis in respect of *each* output text/image, *vis-à-vis*, the original work. This Part evaluates how the computational data analysis exception

and fair use provision in the Singapore Copyright Act 2021<sup>1</sup> (the “Copyright Act 2021”) may be applicable to the determination of liability in input and output scenarios.

## II. Does the computational data analysis exception apply?

2 In Singapore, copyright law can provide a defence for infringing uses if these uses fall under either the computational data analysis exception<sup>2</sup> or the fair use provision.<sup>3</sup> Under the computational data analysis exception which is most relevant for use of copyright-protected works for machine learning, five stringent conditions must be fulfilled, and they include “lawful access” to the copyrighted content and that the making of a copy cannot be used for any other purpose except for identifying, extracting or analysing information/data *and* using that to improve the functioning of a program in relation to that type of information/data.<sup>4</sup> In implementing Proposal 8 of the Copyright Review Report,<sup>5</sup> s 243 of the Copyright Act 2021 introduces a specific exception for reproduction of works made for the purpose of *computational data analysis*, provided that a number of conditions in s 244 are met. Computational data analysis is defined non-exhaustively as “using a computer program to identify, extract and analyse information or data from the work”<sup>6</sup> – which is synonymous with text and data mining (“TDM”). The UK has already in place a TDM exception – albeit narrower than the Singapore version – that “a person who has lawful access to the work may carry out a computational analysis of anything recorded in the work for the sole purpose of research for a non-commercial purpose”.<sup>7</sup> However, TDM exceptions tend to be narrowly circumscribed and are unlikely to encompass the

---

1 2020 Rev Ed.

2 Copyright Act 2021 (2020 Rev Ed) ss 243 and 244.

3 Copyright Act 2021 (2020 Rev Ed) ss 190 and 191.

4 For an analysis of the computational data analysis exception, see David Tan & Thomas Lee Chee Seng, “Copying Right in Copyright Law: Fair Use, Computational Data Analysis and the Personal Data Protection Act” (2021) 33 SAcLJ 1032.

5 Ministry of Law & Intellectual Property Office of Singapore, *Singapore Copyright Review Report* (17 January 2019).

6 Copyright Act 2021 (2020 Rev Ed) s 243.

7 Copyright, Designs and Patents Act 1988 (c 48) (UK) s 29A.

way that generative AI makes use of the copyright-protected works in order to generate new works.

3 Under the Copyright Act 2021, the five conditions that must all be satisfied are:

- (a) the user proving that the copy is made for the purpose of computational data analysis;
- (b) the user proving that the copy is not used for any other purpose;
- (c) the user not supplying the copy to any person other than for the purpose of verifying the results of the computational data analysis carried out by the user;
- (d) the user has lawful access to the material (“the first copy”) from which the copy is made; and
- (e) that the first copy is not an infringing copy; and if it is an infringing copy, it has to satisfy other requirements.

The Singapore legislation gives an example that “X does not have lawful access to the first copy if X accessed the first copy by circumventing paywalls”<sup>8</sup> and that the use of images to train a computer program to recognise images, such as facial recognition software, as a permissible purpose. The latter scenario is likely deemed to be permissible since the software does not use the copyright-protected expression that inheres in a photograph, but rather utilises the “facts” (which are not protected) of a face to map the facial points typically using 68 co-ordinates. Furthermore, it is stated in the Copyright Act 2021 that “X does not have lawful access to the first copy if X accessed the first copy in breach of the terms of use of a database”.<sup>9</sup>

4 For machine learning purposes, the scraping of the Internet for text and images will often circumvent paywalls or violate the terms of use, hence failing the “lawful access” requirement under s 244(2)(d). The way that ChatGPT, as well as its precursor models, is allegedly trained using datasets from

---

8 Copyright Act 2021 (2020 Rev Ed) s 244(2)(d), illus (a).

9 Copyright Act 2021 (2020 Rev Ed) s 244(2)(d), illus (b).

Common Crawl, BookCorpus, Books1 and Books2 from shadow libraries, suggests that the access is “unlawful”. Furthermore, the making of a copy, which will involve the conversion of authorial works into a machine-readable format or, in some GAIAs, data storage, will not be for the sole purpose of analysing the data to improve the functioning of the AI in relation to that data;<sup>10</sup> it will be for the purpose of generating new expressive works based on that data, which is an impermissible purpose. It is also worth noting that the UK government’s consultation on a proposed TDM exception which allows TDM for any purpose received 88 written submissions, with only 13 in favour.<sup>11</sup>

### III. Is it fair use?

#### A. Relevant principles

5 But is it nonetheless fair use? In Singapore, s 191 of the Copyright Act 2021 enumerates a non-exclusive list of four factors to be weighed to determine whether an unauthorised use is fair, and hence a permitted use, much akin to the US fair use provision.<sup>12</sup> In the US, fair use has allowed Google Books, acting without permission of rights holders, to make digital copies of tens of millions of books to establish a publicly available Internet search function.<sup>13</sup> An important feature is that an Internet user can use this function to search without charge to determine

---

10 Copyright Act 2021 (2020 Rev Ed) s 244(2)(b).

11 UK Intellectual Property Office, “Consultation Outcome – Artificial Intelligence and IP: Copyright and Patents” (28 June 2022) <<https://www.gov.uk/government/consultations/artificial-intelligence-and-ip-copyright-and-patents>> (accessed 6 November 2023). The Cyberspace Administration of China released a draft version of its Administrative Measures for Generative Artificial Intelligence Services for public consultation; the deadline for submitting comments was 10 May 2023. Article 7 of the draft Measures stipulates that providers of generative AI must ensure that data used for training and optimization is obtained through legal means, and must, *inter alia*, comply with requirements stipulated by the Cybersecurity Law and not contain content that infringes intellectual property. See, *eg*, Jeremy Daum, “Overview of Draft Measures on Generative AI”, *China Law Translate* (14 April 2023) <<https://www.chinalawtranslate.com/en/overview-of-draft-measures-on-generative-ai/>> (accessed 6 November 2023).

12 17 USC (US) § 107.

13 *Authors Guild v Google Inc* 804 F 3d 202 (2nd Cir, 2015).

whether the book contains a specified word or term and also see snippets of text containing the searched-for terms.

6 This decision on “transformative use” by the US Second Circuit Court of Appeals was also cited by the Singapore Court of Appeal with approval in *Global Yellow Pages v Promedia Directories Pte Ltd*.<sup>14</sup> In evaluating the extent to which a work is transformative, the court will typically consider the *purpose* of the original and infringing secondary works; the secondary use should be plainly different from the original purpose for which the works were created. In a recent decision, the US Supreme Court commented: “the first factor relates to the problem of substitution—copyright’s *bête noire*. The use of an original work to achieve a purpose that is the same as, or highly similar to, that of the original work is more likely to substitute for, or ‘supplan[t]’ the work”.<sup>15</sup>

7 It was important that Google Books augmented public knowledge by making available information about the books without providing the public with a substantial substitute for matter protected by the copyright interests in the original works.<sup>16</sup> In May 2023, the majority opinion of the US Supreme Court’s decision of *Andy Warhol Foundation for the Visual Arts Inc v Goldsmith*<sup>17</sup> (“*Andy Warhol Foundation for the Visual Arts*”) found that Lynn Goldsmith’s original photograph of the pop singer Prince, and the Andy Warhol Foundation (“AWF”)’s copying use of that photograph in an image licensed to a special edition magazine devoted to Prince, shared substantially the same purpose, and the use was of a commercial nature; as a result, the first factor weighed against fair use.<sup>18</sup> The majority commented that a use

---

14 *Global Yellow Pages Ltd v Promedia Directories Pte Ltd* [2017] 2 SLR 185 at [81].

15 *Andy Warhol Foundation for the Visual Arts Inc v Goldsmith* 598 US \_\_\_ (2023) at 15.

16 *Andy Warhol Foundation for the Visual Arts Inc v Goldsmith* 598 US \_\_\_ (2023) at 15, citing *Authors Guild v Google Inc* 804 F 3d 202 at 207 (2nd Cir, 2015).

17 598 US \_\_\_ (2023).

18 *Andy Warhol Foundation for the Visual Arts Inc v Goldsmith* 598 US \_\_\_ (2023) at 4. The appeal to the Supreme Court was an unusual one as it was only on the finding of the first factor against AWF by the Second Circuit; the other three factors also weighed in favour of the plaintiff Lynn Goldsmith. With the majority of the Supreme Court deciding that the first factor weighed against AWF, the Second Circuit’s judgment was affirmed. It is pertinent to  
(cont’d on the next page)

may be justified because copying is reasonably necessary to achieve the user's *new* purpose:<sup>19</sup>

An independent justification like this is particularly relevant to assessing fair use where an original work and copying use share the same or highly similar purposes, or where wide dissemination of a secondary work would otherwise run the risk of substitution for the original or licensed derivatives of it.

8 In respect of the first factor, the Supreme Court majority opinion was of the view that one should be careful not to adopt an overly broad notion of transformativeness, and that “the meaning of a secondary work, as reasonably can be perceived, should be considered to the extent necessary to determine *whether the purpose of the use is distinct from the original*, for instance, because the use comments on, criticizes, or provides otherwise unavailable information about the original” [emphasis added].<sup>20</sup>

9 Regarding GAIAs, two fair use factors that are likely to carry the greatest weight in the analysis are: (a) the purpose/character of the use, namely whether the use by generative AI is “transformative”, *ie*, whether it changes the purpose or the nature of the original work in some way; and (b) the impact of the generative AI’s use on the market, *ie*, whether it threatens the livelihood of the original creator by competing with their works or the licensing market for their works.

10 In respect of the first factor, *Authors Guild Inc v HathiTrust*<sup>21</sup> (“*HathiTrust*”) is also instructive – the issue was whether the digitisation of copyrighted works by 13 universities and other organisations in creating the HathiTrust Digital Library (“HDL”) without authorisation may constitute fair use.<sup>22</sup> The US Second Circuit Court of Appeals found that the first factor weighed in

---

note that “the Court expresses no opinion as to the creation, display, or sale of any of the original Prince Series works [by Andy Warhol]”: *Andy Warhol Foundation for the Visual Arts Inc v Goldsmith* 598 US \_\_\_ (2023) at 21.

19 *Andy Warhol Foundation for the Visual Arts, Inc v Goldsmith* 598 US \_\_\_ (2023) at 19.

20 *Andy Warhol Foundation for the Visual Arts, Inc v Goldsmith* 598 US \_\_\_ (2023) at 32.

21 755 F 3d 87.

22 *Authors Guild Inc v HathiTrust* 755 F 3d 87 at 92 (2nd Cir, 2014).

favour of fair use as HDL’s enabling of full-text search “serves a new and different function from the original” and is socially beneficial.<sup>23</sup> Additionally, the dealing was found to carry a “non-profit educational” purpose as the HDL was a project started by educational and non-profit institutions targeted at providing greater access to works without any “purely commercial” motive.<sup>24</sup> Even if there is a commercial motivation, the Second Circuit Court of Appeals in the later decision of *Authors Guild Inc v Google Inc* (“*Authors Guild v Google*”) saw “no reason ... why Google’s overall profit motivation should prevail as a reason for denying fair use over its highly convincing transformative purpose, together with the absence of significant substitutive competition, as reasons for granting fair use”.<sup>25</sup> The court held that similar to *HathiTrust*, the purpose of Google’s copying of the original copyrighted books is “to make available significant information *about those books*, permitting a searcher to identify those that contain a word or term of interest, as well as those that do not include reference to it” [emphasis in original]<sup>26</sup> which is significantly different from the purposes of the original books.

11 Regarding the output, the raw material must have been transformed in “the creation of new information, new aesthetics, new insights and understanding”<sup>27</sup> but it does not follow that “any secondary work that adds a new aesthetic or new expression to its source material is necessarily transformative” especially where the new elements did not obscure the original image.<sup>28</sup> The secondary work itself must reasonably be perceived as embodying an entirely distinct artistic purpose, one that conveys a “new meaning or message” entirely separate from its source material.<sup>29</sup> In a recent decision handed down in 2021,

---

23 *Authors Guild Inc v HathiTrust* 755 F 3d 87 at 96 (2nd Cir, 2014). See also William F Patry, *Patry on Copyright* vol 4 (West, Online, 2015) at section 10:21 (observing that the use in *HathiTrust* is “socially beneficial, serves a different purpose than the original, and is in no way substitutional”).

24 *Authors Guild Inc v HathiTrust* 755 F 3d 87 at 90–91 (2nd Cir, 2014).

25 *Authors Guild v Google Inc* 804 F 3d 202 (2nd Cir, 2015).

26 *Authors Guild v Google Inc* 804 F 3d 202 at 217 (2nd Cir, 2015).

27 *Cariou v Prince* 714 F 3d 694 at 706 (2nd Cir, 2013).

28 *Andy Warhol Foundation for the Visual Arts Inc v Goldsmith* 992 F 3d 99 at 111 and 113 (2nd Cir, 2021).

29 *Andy Warhol Foundation for the Visual Arts Inc v Goldsmith* 992 F 3d 99 at 113 (2nd Cir, 2021).



the Second Circuit emphasised that “the judge must examine whether the secondary work’s use of its source material is in service of a ‘fundamentally different and new’ artistic purpose and character, such that the secondary work stands apart from the ‘raw material’ used to create it”.<sup>30</sup>

12 Parodies fall squarely within this definition of transformative use,<sup>31</sup> and parody is also an autonomous concept in European Union law.<sup>32</sup> According to the European Court of Justice in *Johan Deckmyn and Vrijheidsfonds VZW v Helena Vandersteen*,<sup>33</sup> the “essential characteristics of parody are, first, to evoke an existing work while being noticeably different from it, and, secondly, to constitute an expression of humour or mockery”.<sup>34</sup> The US courts adopt a similar definition, allowing the parodist to quote from existing material in creating a new one that, at least in part, comments on that author’s work for comic effect or ridicule.<sup>35</sup> These approaches would be relevant when evaluating whether the outputs by GAIAs may be considered parodies, thus qualifying more easily for fair use protection, compared to non-parodies.

13 The Ninth Circuit’s decision in *Kelly v Arriba Soft Corp*<sup>36</sup> is also useful in understanding how the evaluation of the third factor (ie, the amount and substantiality of the portion used in relation to the copyrighted work as a whole)<sup>37</sup> could be applied

---

30 *Andy Warhol Foundation for the Visual Arts Inc v Goldsmith* 992 F 3d 99 at 114 (2nd Cir, 2021). The decision was affirmed by the Supreme Court on appeal.

31 See, eg, *Campbell v Acuff-Rose Music Inc* 510 US 569 (1994); *Leibovitz v Paramount Pictures Corp* 137 F 3d 109 (2nd Cir, 1998); *Suntrust Bank v Houghton Mifflin Co* 268 F 3d 1257 (11th Cir, 2001); *Dr Seuss Enterprises LP v ComicMix LLC* 983 F 3d 443 (9th Cir, 2020).

32 *Johan Deckmyn and Vrijheidsfonds VZW v Helena Vandersteen*, Case C-201/13, EU:C:2014:2132. According to the Court of Justice of the European Union, “the concept of ‘parody’, which appears in a provision of a directive that does not contain any reference to national laws, must be regarded as an autonomous concept of EU law and interpreted uniformly throughout the European Union”: *Johan Deckmyn and Vrijheidsfonds VZW v Helena Vandersteen*, Case C-201/13, EU:C:2014:2132 at [15].

33 Case C-201/13, EU:C:2014:2132.

34 *Johan Deckmyn and Vrijheidsfonds VZW v Helena Vandersteen*, Case C-201/13, EU:C:2014:2132 at [20].

35 *Campbell v Acuff-Rose Music Inc* 510 US 569 at 580 (6th Cir, 1994).

36 336 F 3d 811 (9th Cir, 2002).

37 Copyrights 17 USC (US) § 107(3).



to generative AI uses. There, it was held that the use of entire copyrighted works was *necessary* in situations involving search engines since copying only a part of the copyrighted work would create practical difficulties for users, thereby reducing the usefulness of the search engine. In the same vein, even if entire works were copied by web robots in the TDM context, it could be reasoned that such a taking is reasonable, considering the *different* purpose of the dealing (*ie*, to identify patterns in vast amounts of raw data); thus, the third factor might not necessarily weigh against fair use. But if the purpose of generative AI is to analyse the specific expression of particular artists, and then replicate portions of that expression in response to a text prompt, then it does not appear to be a different purpose. The Second Circuit has observed that the courts have rejected any categorical rule that a copying of the entirety cannot be a fair use, especially when the copying was reasonably appropriate to achieve the copier's transformative purpose and was conducted in a manner that did not offer a competing substitute for the original.<sup>38</sup>

14 Again, the Supreme Court's decision in *Andy Warhol Foundation for the Visual Arts* is instructive. The majority opinion observed that "[w]hether the purpose and character of a use weighs in favor of fair use is, instead, an objective inquiry into what use was made, *i.e.*, what the user does with the original work".<sup>39</sup> In that case, the use was AWF's commercial licensing of Warhol's Orange Prince (which was based on Lynn Goldsmith's original photograph) to appear on the cover of Condé Nast's special commemorative edition. The purpose of that use was to illustrate a magazine about Prince with a portrait of Prince, and an infringing work that portrays Prince somewhat differently from Goldsmith's photograph (yet has no critical bearing on her photograph) was insufficient for the first factor to favour AWF, given the specific context of the use. The majority emphasised:<sup>40</sup>

---

38 *Authors Guild v Google Inc* 804 F 3d 202 at 220 (2nd Cir, 2015).

39 *Andy Warhol Foundation for the Visual Arts Inc v Goldsmith* 598 US \_\_\_ (2023) at 33.

40 *Andy Warhol Foundation for the Visual Arts Inc v Goldsmith* 598 US \_\_\_ (2023) at 33.

To hold otherwise would potentially authorize a range of commercial copying of photographs, to be used for purposes that are substantially the same as those of the originals. As long as the user somehow portrays the subject of the photograph differently, he could make modest alterations to the original, sell it to an outlet to accompany a story about the subject, and claim transformative use.

These observations are especially pertinent for images produced by GAIAs such as DALL·E, Stable Diffusion and Midjourney. If a user is looking for an image for illustrative purposes for a magazine, book, annual report or marketing brochure, and provides specific text prompts to a generative AI system to produce such an image – as opposed to licensing one directly from the original author – then the first factor is unlikely to weigh in favour of fair use.

15 The application of the fourth factor (*ie*, the effect of the use upon the potential market for or value of the copyrighted work)<sup>41</sup> is also highly dependent on the finding of the first factor. The US Supreme Court in *Campbell v Acuff-Rose Music Inc* (“*Campbell*”) had emphasised the close linkage between the first and fourth factors, in that the more the copying is done to achieve a purpose that differs from the purpose of the original, the less likely it is that the copy will serve as a satisfactory substitute for the original.<sup>42</sup> The Second Circuit noted that even if the *purpose* of the copying was for a valuably transformative purpose, such copying might nonetheless harm the value of the copyrighted original if done in a manner that resulted in widespread revelation of sufficiently significant portions of the original as to make available a significantly competing substitute.<sup>43</sup> Generally, copyright-protected works copied for data mining purposes will require extensive processing and analysis before knowledge is derived and shared. Miners must ensure that they do not reveal significant portions of the original copyrighted works to the public. Although one could argue that data mining could limit the rights owners’ expansion into a

---

41 Copyrights 17 USC (US) § 107(4).

42 *Campbell v Acuff-Rose Music Inc* 510 US 569 at 591 (6th Cir, 1994).

43 *Authors Guild v Google Inc* 804 F 3d 202 at 223 (2nd Cir, 2015).

potential market (eg, a lost opportunity to license the works)<sup>44</sup> since markets are dynamic and change over time to meet new demands, the US Circuit Courts have universally dismissed this argument where only a small portion of the original works was revealed to the public. In the *Authors Guild v Google* litigation, the Second Circuit held that a mere revelation of “16% of the text of Plaintiffs’ books overstates the degree to which snippet view can provide a meaningful substitute”.<sup>45</sup> In generative AI scenarios where a significant portion of an original work is reproduced in an output in response to a user’s text prompt, then one may more confidently discern a substitutive impact.

16 In Barton Beebe’s recent empirical analysis of US fair use judicial opinions, he commented that: “In instructing courts to assess ‘the effect of the use upon the potential market for or value of the copyrighted work,’ factor four requires courts to define the limits of the ‘potential market’ and ‘value’ that the copyright owner should have the exclusive right to exploit.”<sup>46</sup> While the judicial attention to transformativeness under the first factor seemed to be waning by the end of 2005, it has now returned to “very high levels”.<sup>47</sup> Beebe cautioned that although a finding of transformativeness is nearly sufficient to result in an overall finding of fair use, “it is far from necessary to trigger such a finding” and that the data “indicate that the transformativeness inquiry has not in fact replaced the overall fair use analysis”.<sup>48</sup> The first and the fourth factors continue to be the most influential in driving the fair use test, and the recent data indicate that the fourth factor has remained the single dominant factor in the US courts’ adjudication of fair use. Thus it would appear that

---

44 This was an argument the plaintiffs raised. *Authors Guild Inc v HathiTrust* 755 F 3d 87 at 99 (2nd Cir, 2014).

45 *Authors Guild v Google Inc* 804 F 3d 202 at 223 (2nd Cir, 2015).

46 Barton Beebe, “An Empirical Study of U.S. Copyright Fair Use Opinions Updated, 1978–2019” (2020) 10 *New York University Journal of Intellectual Property and Entertainment Law* 1 at 34.

47 Barton Beebe, “An Empirical Study of U.S. Copyright Fair Use Opinions Updated, 1978–2019” (2020) 10 *New York University Journal of Intellectual Property and Entertainment Law* 1 at 25.

48 Barton Beebe, “An Empirical Study of U.S. Copyright Fair Use Opinions Updated, 1978–2019” (2020) 10 *New York University Journal of Intellectual Property and Entertainment Law* 1 at 27.

“economic analysis continues ultimately to define fair use in the American copyright system”.<sup>49</sup>

## **B. Fair use in respect of input and output**

17 ChatGPT, Stable Diffusion, DreamStudio and many other comparable GAIAs are not search engines. A number of them are highly successful commercial enterprises, with Stability AI valued at US\$1bn, and some charging a user fee for their services. There is also little transformative purpose to be found as the AI would be accessing and reproducing the creative expression in these works in the outputs, *ie*, the works would have been appropriated for their creative elements rather than their underlying facts.

18 It is not easy to apply the fair use analysis separately to the issue of input independently of the output, as the *purpose* of the defendant’s use of the training input is often discernible only when considered in light of the output. In *Authors Guild v Google*, the Second Circuit observed that “copying from an original for the purpose of criticism or commentary on the original or provision of information about it, tends most clearly to satisfy *Campbell*’s notion of the ‘transformative’ purpose involved in the analysis of Factor One”.<sup>50</sup> In the claim, *Authors Guild v OpenAI Inc*,<sup>51</sup> filed in September 2023, it was alleged that when prompted, ChatGPT can, *inter alia*, generate summaries of books and infringing unauthorised detailed outlines for the next purported instalment of certain books.<sup>52</sup> In the *Chabon v OpenAI Inc*<sup>53</sup> claim, when ChatGPT was prompted to produce a screenplay in the style of David Henry Hwang’s *The Dance and the Railroad*, it generated a script written in Hwang’s style.<sup>54</sup> Ideas, facts, style and genre generally do not attract copyright protection. As fair use is a fact-specific inquiry, courts will have to assess each

---

49 Barton Beebe, “An Empirical Study of U.S. Copyright Fair Use Opinions Updated, 1978–2019” (2020) 10 *New York University Journal of Intellectual Property and Entertainment Law* 1 at 4.

50 *Authors Guild v Google Inc* 804 F 3d 202 at 216 (2nd Cir, 2015).

51 Case 1:23-cv-08292 (SDNY, 2023).

52 *Authors Guild v OpenAI Inc* Case 1:23-cv-08292 (SDNY, 2023).

53 Case 3:23-cv-04625 (ND Cal, 2023).

54 *Chabon v OpenAI Inc* Case 3:23-cv-04625 (ND Cal, 2023) at [50].

particular infringing use to determine whether it is a fair use. In response to a user’s prompt, ChatGPT’s generation of a summary of a book that it had ingested for training purposes may not be fair use (as both training input and output). However, its critical essay of the book would qualify as the equivalent of a book review and may be fair use, especially in light of the public benefit and absence of market substitutability. Moreover, the output from GAIAs that generate text commands a very different analysis from those that generate images. A summary or critique of a book is less likely to substitute for the original book compared to an AI-generated image that would substitute for one that could have been purchased from Getty Images or Shutterstock.

19 In *Google LLC v Oracle America Inc*<sup>55</sup> (“*Google v Oracle*”), the US Supreme Court emphasised that the copier’s use must add “something new, with a further purpose or different character, altering the copyrighted work with new expression, meaning or message”.<sup>56</sup> GAIAs are trained essentially with existing creative works, and in a significant number of outputs, it appears the works have been simply remixed or collaged to derive more works of the same kind based on our text prompts. Depending on the text prompts, it may be reproducing certain copyright-protected works without the requisite degree of transformativeness that will tilt the first factor in favour of fair use. Furthermore, the majority’s decision in *Andy Warhol Foundation for the Visual Arts*, as well as *Google v Oracle*, suggests that the first factor analysis may depend upon whether there is a market for licensing content for training data (fourth factor); commercial copying to generate training data input for GAIAs would likely fail the first factor analysis absent a “particularly compelling justification” for the copying.<sup>57</sup>

---

55 141 S Ct 1183.

56 *Google LLC v Oracle America Inc* 141 S Ct 1183 at 1202 (2021).

57 *Andy Warhol Foundation for the Visual Arts Inc v Goldsmith* 598 US \_\_\_ (2023) at 35.



**Figure 1: Mickey Mouse in the style of Picasso in battle with Captain America (generated by DALL-E 2)**

20 Consider, for instance, DALL-E 2’s response to the author’s text prompt “Mickey Mouse in the style of Picasso in battle with Captain America” (Figure 1).<sup>58</sup> Although it does not share a similar style as Picasso’s paintings of *Guernica* and *La Femme qui pleure* (*Weeping Woman*), there is substantial similarity with two copyright-protected works, *ie*, Captain America and Mickey Mouse. There does not seem to be sufficient visual transformation (in the manner required in the US Second and Ninth Circuit decisions like *Blanch v Koons*<sup>59</sup> and *Dr Seuss Enterprises LP v ComicMix LLC*<sup>60</sup>),<sup>61</sup> and it does not appear to be a parody that is “a spoof, send-up, caricature, or comment on” the original

---

58 DALL-E 2 at <<https://openai.com/dall-e-2>>.

59 467 F 3d 244 (2nd Cir, 2006).

60 983 F 3d 443 (9th Cir, 2020).

61 The collage *Niagara* by Jeff Koons included new visual elements and used Andrea Blanch’s image as fodder for his commentary on the social and aesthetic consequences of mass media; his purposes in using Blanch’s image were also reasonably perceived to be sharply different from Blanch’s goals in creating it: *Blanch v Koons* 467 F 3d 244 at 252–253 (2nd Cir, 2006). But some collages which are mere juxtapositions or mash-ups may not be sufficiently transformative and will not be fair use. *Eg*, *Dr Seuss Enterprises LP v ComicMix LLC* 983 F 3d 443 (9th Cir, 2020) (a cartoon mash-up of *Star Trek* and Dr Seuss’ *Oh, the Places You’ll Go!*).

works or holds the style of the original works to ridicule.<sup>62</sup> In interpreting the Singapore Copyright Act, the courts in Singapore are likely to find the US decisions persuasive, especially when the Court of Appeal had previously cited the US Supreme Court’s decision of *Campbell* with approval, and its comment that “parody had an obvious claim to transformative value”.<sup>63</sup> Although the issue of who is the author(s) of the work *Mickey Mouse in the style of Picasso in battle with Captain America* remains unsettled,<sup>64</sup> the text prompt neither requested for the creation of this work as a “parody” nor as “an expression of humour or mockery”.<sup>65</sup> It would be tenuous to argue *ex post facto* that this was a parody.

21 In another scenario, let us assume an author is publishing a book titled “Andy Warhol and the Pop Art Movement”; one that contains commentaries on Warhol’s illustrious iconography and his influence on other artists. The author could approach the AWF to pay a licence fee for the use of one of Warhol’s *Shot Marilyns* works for the cover of the book and its related marketing materials. Instead, using Stable Diffusion Online,<sup>66</sup> the author entered the text prompt “Marilyn Monroe in the style of Andy Warhol”, and the image generated resembles the original “Marilyn” silkscreens by Andy Warhol (Figure 2). The output image with the pink facial hue and golden hair – albeit without the blue background –

---

62 *Dr Seuss Enterprises LP v ComicMix LLC* 983 F 3d 443 at 452 (9th Cir, 2020). The Supreme Court found that 2 Live Crew’s rap song *Pretty Woman* was a parody of Roy Orbison’s *Oh, Pretty Woman* as it “juxtaposes the romantic musings of a man whose fantasy comes true, with degrading taunts, a bawdy demand for sex, and a sigh of relief from paternal responsibility. The later words can be taken as a comment on the naiveté of the original of an earlier day, as a rejection of its sentiment that ignores the ugliness of street life and the debasement that it signifies”. It is this “joinder of reference and ridicule” that defines a parody: *Campbell v Acuff-Rose Music Inc* 510 US 569 at 583 (6th Cir, 1994).

63 *Global Yellow Pages Ltd v Promedia Directories Pte Ltd* [2017] 2 SLR 185 at [81].

64 In this regard, see Wee Liang Tan & David Tan, “AI, Author, Amanuensis” (2022) 5(2) *Journal of Intellectual Property Studies* 1 at 16–30. Some pertinent issues include whether ChatGPT and the person providing the text prompt may be joint authors, thus resulting in a situation of joint liability, and vicarious copyright infringement (eg, *Metro-Goldwyn-Mayer Studios Inc v Grokster Ltd* 545 US 913 at 930 (2005); *Perfect 10 Inc v Visa International Service Association* 494 F 3d 788 at 802 (9th Cir, 2007)).

65 *Johan Deckmyn and Vrijheidsfonds VZW v Helena Vandersteen*, Case C-201/13, EU:C:2014:2132 at [33].

66 Stable Diffusion Online at <<https://stablediffusionweb.com/>>.



appears to have been reproduced from the original *Shot Sage Blue Marilyn* by Andy Warhol.<sup>67</sup> Since one can freely use the images from Stable Diffusion Online in a variety of commercial settings under the CreativeML OpenRAIL-M licence, one can assume that the generated image may be used in advertisements, creation and sale of non-fungible tokens, commercial merchandising – and book covers. Applying the reasoning of the majority opinion of the US Supreme Court, as well as the Second Circuit, in *Andy Warhol Foundation for the Visual Arts*, the purpose of the use is neither likely to be sufficiently distinct from the original nor offering a significantly different meaning or message. As the Stable Diffusion output reproduces a significant proportion of the original Warhol work, it can be distinguished from the snippets produced by Google Books in *Authors Guild v Google* which were adjudged to be fair use. The image generated by Stable Diffusion Online will also compete with the licensing market for Andy Warhol’s *Shot Marilyns*, even though it would not be a substitute for the original artwork by Warhol. In summary, both the input (assuming Warhol’s “Marilyn” was used for machine learning) and output are unlikely to be fair use.



**Figure 2: Marilyn Monroe in the style of Andy Warhol (generated by Stable Diffusion Online)**

---

<sup>67</sup> “Warhol’s Marilyn Monroe Painting Sold for Record-Breaking \$195m”, *BBC News* (10 May 2022) <<https://www.bbc.com/news/world-us-canada-61339179>> (accessed 6 November 2023).

22 Last but not least, ChatGPT's replies to our text prompts are not based on a process of reasoning or akin to human comprehension; it is based on the probabilities of certain words occurring together, and may generate paragraphs of text from copyrighted literary works in its response. To be clear, some of the output generated by AI may be highly transformative (eg, they may resemble Jeff Koons' collage *Niagara* in the *Blanch v Koons*<sup>68</sup> case which was reasonably perceived to be a new work of art with a distinct and new meaning, message and character), but it is the use of the creative works in the machine learning process that is arguably not transformative. Such unrestricted and widespread use would have a substantially adverse impact on the licensing markets of these copyright-protected works.<sup>69</sup>

#### IV. Conclusions

23 The different creative, cultural and economic imperatives that underlie the copyright system in Singapore are not significantly different from the US approach.<sup>70</sup> Generative AI can equally inspire awe and concern. The arrival of ChatGPT in November 2022 and its ensuing global success with users precipitated a chaotic hustle amongst developers and corporations to refine their existing generative AI systems and launch new ones. Our lives may have been enriched by generative AI, and our secret dreams of being a writer or artist can now be realised with the aid of GPT-4 or Stable Diffusion, but it does not mean that these benefits should not come at a price. Program developers, and the organisations that invest significant resources in generative AI, will have the benefit of access to two generous and distinct provisions in the Singapore Copyright Act 2021 as defences to their unauthorised uses of copyright-protected works. However, they must be mindful that the judicial interpretation of the fair use provision in Singapore likely remains guided by paradigmatic cases where an open-universe search engine makes a digital copy for the purpose of enabling searches for books that

---

68 467 F 3d 244 (2nd Cir, 2006).

69 *Global Yellow Pages Ltd v Promedia Directories Pte Ltd* [2017] 2 SLR 185 at [84].

70 See, eg, *Andy Warhol Foundation for the Visual Arts Inc v Goldsmith* 598 US \_\_\_ (2023) at 13.

contain a term of interest and involves a highly transformative purpose, and where copying from an original for the purpose of criticism, commentary, caricature, parody, or pastiche would be transformative.

24 Copyright is, at its core, driven by an economic ethos. Copyright exists to reward original creative efforts, and any subsequent uses that draw on and reproduce creative content ought to pay an appropriate fee. Perhaps the best things in life are not for free.