Reform of the Computational Data Analysis Exception

Disclaimer

Whilst every effort has been made to ensure that the information contained in this report is correct, all authors, their organisations, and the Singapore Academy of Law disclaim all liability and responsibility for any error or omission in this report, and in respect of anything, or the consequences of anything, done or omitted to be done by any person in reliance, whether wholly or partially, upon the whole or any part of the contents of this report.

Acknowledgments

For insightful discussions relating to this report and its topics, the Singapore Academy of Law would like to record its gratitude to:[1]

- Diyanah Baharudin
- Cheah Yew Kuin
- Chng Bee Peng, VP, Legal Regulatory & Secretariat, Mediacorp Pte. Ltd.
- Byron Karuppiah, General Counsel, Ahrefs
- Shameek Kundu
- Dr Stanley Lai, SC, Head of Intellectual Property Practice, Co-Head of Cybersecurity & Data Protection Practice and Partner of Litigation & Dispute Resolution Department, Allen & Gledhill LLP
- Lam Chung Nian, Partner & Head, IP, Technology & Data Group, WongPartnership LLP, Senior Accredited Specialist (Data & Digital Economy), SAL
- Professor David Llewelyn, David Llewelyn & Co LLC
- Jeth Lee, Regional Legal Director, APAC, Microsoft
- Jeffrey Lim, Co-Managing Director, Joyce A. Tan & Partners LLC, Senior Accredited Specialist (Data & Digital Economy), SAL
- Sujoy Majumdar, Counsel
- Professor Ng-Loy Wee Loon, SC (hon), Faculty of Law, National University of Singapore
- Oh Pin-Ping, Partner, Bird & Bird ATMD LLP
- Derek Ow, Corporate Counsel, SPH Media Limited
- Jared Ragland, PhD, Senior Director, Policy – APAC, Business Software Alliance
- Roe Yun Song, VP, Legal Regulatory & Secretariat, Mediacorp Pte. Ltd.
- Peter Schoppert, Vice President, Singapore Book Publishers Association
- Cheryl Seah, Director, Drew & Napier LLC
- Simon Seow, Consultant
- Dr. Kelvin Sum Ying Yew
- Dr. Eric Tan, Assistant Director, BizTech Group, Infocomm Media Development Authority
- James Tan, President, Visual, Audio, Creative Content Professionals Association (Singapore)
- Tham Shen Hong, Senior Policy Manager (APAC), Business Software Alliance
- Dr. William Tjhi
- Alvaro Vigil, Senior AI Solutions Architect, LawNet Technology Services

[1] In alphabetical order of surname. The Singapore Academy of Law similarly extends its gratitude to many others who have participated in discussions but prefer to remain anonymous. Participation in discussions by these individuals does not indicate their endorsement of, or agreement with, the views or recommendations expressed in this report, nor their organisations' endorsement of, or agreement with, the views or recommendations expressed in this report.

Reform of the Computational Data Analysis Exception

First published January 2026; last updated February 2026

Table of content

Background

1. On 21 November 2021, the Copyright Act 2021 (**Act**) came into force in Singapore. The Act repealed and re-enacted the previous Copyright Act that had been in effect since 1987.

2. Incorporated into Part 5 ("Permitted Uses of Copyright Works and Protected Performances") of the Act is Division 8 titled "Computational data analysis". Division 8 comprises two sections: sections 243 ("Interpretation: what is computational data analysis") and 244 ("Copying or communicating for computational data analysis").

3. As indicated by its title, section 244 establishes two permitted uses of material that is subject to copyright: copying and communication to the public. As a result, Division 8 of Part 5 is commonly referred to as the "computational data analysis *exception*".

4. During the Second Reading of the Bill comprising Division 8 of Part 5, the Second Minister for Law explained section 244 as follows:[2]

> Another change, Sir, is found in clause 244 of the Bill, which allows the use of works and recordings of protected performances for computational data analysis.
>
> To give Members an example, this permitted use can be relied on when training an Artificial Intelligence programme or using computers to analyse large databases of materials, without needing to seek permission from each rights owner.
>
> It will apply in both commercial and non-commercial contexts.
>
> However, the user must have lawfully accessed the relevant materials, for example, he or she should not have accessed the materials by circumventing a paywall.
>
> This supports our Smart Nation initiatives, our push towards data-driven innovation and Singapore's efforts to grow our Artificial Intelligence and technology sectors.

5. While the Government had artificial intelligence (**AI**) development in mind when introducing Division 8 of Part 5, much of what occurred in the years that followed could not have been predicted.

6. On 30 November 2022, approximately one year after the Act came into force, OpenAI launched ChatGPT. Following the launch of ChatGPT, various large language models (**LLMs**) have been made available to the public in rapid succession.

7. Given the chronology of events, one then queries whether Division 8 of Part 5, created before the generative AI boom, remains adequate in this new environment.

8. In August 2025, the Singapore Academy of Law held a roundtable under the Chatham House Rule to discuss this question and others. The roundtable was attended by a range of stakeholders including rightsholders and developers from both local and international organisations, as well as individuals involved in the creative industries.[3]

9. In the report that follows, assisted by comments made during the roundtable, we assess Division 8 of Part 5, particularly regarding the development of generative AI models. In doing so, we have not sought to reopen the policy choice behind the introduction of this Division.

---

2 "Second Reading of the Copyright Bill", *Parliament of Singapore*, 13 September 2021 (Edwin Tong Chun Fai, Second Minister for Law).

3 Refer to the Acknowledgement Page for a fuller list of roundtable participants. Note that some participants prefer to remain anonymous.

Rather, our assessment is limited to whether the Division achieves the policy aims stated above.

10. To that end, we have made a small number of recommendations for the Government's consideration.

Summary of recommendations

11. Our recommendations are as follows:

    a. that consideration be given to amending paragraphs (2)(c) and 4(b) of section 244 of the Act to include an additional exception such that *X* may supply a copy of a work or a recording of a protected performance to another person for the purpose of that other person performing computational data analysis on *X*'s behalf (see "Outsourcing computational data analysis" section below); and

    b. that consideration be given to amending section 244 of the Act to include a provision with words to the effect that "X *is deemed not to have lawful access to the first copy if access to the first copy is obtained in breach of a contract (ignoring any terms that are void by virtue of section 187)*" (see "Access in breach of contract" section below).

Definition of computational data analysis

> 243. In this Division, "computational data analysis", in relation to a work or a recording of a protected performance, includes —
>
> (a) using a computer program to identify, extract and analyse information or data from the work or recording; and
>
> (b) using the work or recording as an example of a type of information or data to improve the functioning of a computer program in relation to that type of information or data.
>
> *Illustrations*
>
> An example of computational data analysis under paragraph (*b*) is the use of images to train a computer program to recognise images.

12. The phrase "computational data analysis" is defined by reference to two uses, specified in paragraphs (a) and (b) of section 243. Those two uses do not appear to be exhaustive. This is evidenced by the use of the word "includes" at the end of the chapeau to section 243. So, the two types of computational data analysis in paragraphs (a) and (b) of section 243 do not appear to be the sole forms of such analysis that a person is permitted to undertake.

13. Both uses involve a "computer program". A "computer program" is defined in section 13(3) of the Act as follows:

> A "computer program" is an expression (in any language, code or notation) of a set of instructions (whether with or without related information) intended to —
>
> (a) directly cause a device with information processing capabilities to perform a particular function; or
>
> (b) cause a device with information processing capabilities to perform a particular function after —
>
> (i) converting the instructions into another language, code or notation;
>
> (ii) copying the instructions in a different material form; or
>
> (iii) both of those acts.

Using a computer program to identify, extract and analyse information or data

14. The first type of computational data analysis is expressed as follows:

> (a) using a computer program to identify, extract and analyse information or data from the work or recording;

15. Information and data are not things that can be the subject of copyright.[4] So, copying information and data is not an act which can infringe copyright.

---

4 *Asia Pacific Publishing Pte Ltd v Pioneers & Leaders (Publishers) Pte Ltd* [2011] 4 SLR 381; [2011] SGCA 37 at [37].

16. One may then ask why paragraph (a) was included in section 243. Materials may contain both protected elements (the expressive elements) and unprotected elements (like information and data). To identify, extract and analyse the unprotected elements of the material, a person, ***X***, may first need to make a copy of the material as a whole. Absent the permitted use, *X* would, in so doing, infringe the exclusive right of reproduction in respect of the expressive elements of the material.

Improving the functioning of a computer program

17. The second type of computational data analysis is expressed as follows:

> (b) using the work or recording as an example of a type of information or data to improve the functioning of a computer program in relation to that type of information or data.

18. Appended to paragraph (b) is an illustration of this form of computational data analysis. That illustration is as follows:

> An example of computational data analysis under paragraph (b) is the use of images to train a computer program to recognise images.

19. This illustration is not to be taken as exhaustive.[5]

Use of the expressive elements of a work or recording

20. During the second reading of the Copyright Bill, the Second Minister for Law stated the following in answer to a question about computational data analysis:[6]

> Computational data analysis treats the works as data points and does not make use of the expressive nature of the works, which is what copyright fundamentally seeks to protect. In short, because computational data analysis does not itself make use of the expressive nature of the works, the rights owners' reputational and commercial interests are not adversely affected.

21. We observe that Division 8 of Part 5 does not expressly restrict a person from making use of the expressive elements of a work or recording that it has copied.

22. Using a computer program to identify, extract and analyse *information or data* from the work or recording is one type of computational data analysis. But it is not the sole form of such analysis.

23. The other form of computational data analysis provided is as follows:[7]

> using the work or recording as an example of a type of information or data to improve the functioning of a computer program in relation to that type of information or data

---

5 Interpretation Act 1965, section 7A(a).

6 "Second Reading of the Copyright Bill", *Parliament of Singapore*, 13 September 2021 (Edwin Tong Chun Fai, Second Minister for Law).

7 Copyright Act 2021, section 243(b).

24. Here the person does not identify, extract or analyse unprotected elements of a work or recording. Rather, the person treats the work "as an example of a type of information or data" to "improve the functioning of a computer program". The illustration of this second type of computational data analysis provides further context:

> An example of computational data analysis under paragraph (b) is the use of images to train a computer program to recognise images.

25. Depending on the nature of an image, the expressive elements of it may need to be used in order for the computer program to effectively recognise that type of image. The same, we imagine, could be said of other types of works (see paragraph 29 below).

26. We also repeat our observation that these two types of computational data analysis are not the sole types of such analysis as section 243 is expressed inclusively ("'computational data analysis' … *includes* —"), not exhaustively.

27. If computational data analysis were restricted to using the unprotected elements of a work or recording, then the effect of Division 8 of Part 5 on rightsholders would be limited. *X* would be permitted to engage in an otherwise infringing act (copying a work or a recording of a protected performance, including the supply of that copy for the limited purposes of section 244(4)(b)(i) and (ii)) but he or she would not be permitted to use the protected expressive elements of that work or recording.

28. However, section 243 is not expressly restricted in this way. Nor does it appear reasonable to read such a restriction into the provision in light of paragraph (b) of section 243 (including its illustration).

    (We observe that the equivalent provision of the United Kingdom's Copyright, Designs and Patents Act 1988 appears clearer in this regard, stating that a person "may carry out a computational analysis *of anything recorded in the work* …" (emphasis added).[8])

29. Indeed, as a practical matter, it may be difficult for a person to: (a) discern the protected expressive elements of a work or recording from the work or recording's unprotected elements, particularly as this involves questions of law; and (b) limit its computational data analysis to those unprotected elements, thereby undermining the purpose of the Division. Take, for example, a literary work. Computational data analysis of a literary work may involve identifying and recording word frequency, word order, word choice, grammar and other syntactical elements from the work. But those things are the methods by which an author of a literary work uses to express his or her ideas.[9] So, in this regard, it may be impossible for a person to limit his or her computational data analysis to the work's unprotected elements.

    (To the extent one were to argue that the label "computational data analysis" necessarily limits Division 8 of Part 5 to *analysis* of *data* by a *computer*, the foregoing illustrates that such an argument is unfounded. One must interpret the *meaning* of the label as expressed in the text of section 243 consistent with the purposive approach required by section 9A of the Interpretation Act 1965.)

---

8 Copyright, Designs and Patents Act 1988 (1988 c. 48) (United Kingdom), section 29A(1)(a).

9 See, for example, the comments in *Kadrey v Meta Platforms, Inc.*, 3:23-cv-03417, (N.D. Cal. Jun 25, 2025) ECF No. 598 (pdf) at pp. 23–24.

30. In light of the above, there appears to be some inconsistency between the answer of the Second Minister for Law and the words of section 243, if not also the purpose of Division 8 of Part 5 (as expressed by the Second Minister for Law in paragraph 4 above). However, regard is only to be had to such extraneous material when interpreting a statute in limited circumstances. Those limited circumstances are set out in section 9A(2) of the Interpretation Act 1965 and were conveniently summarised by the Court of Appeal of Singapore as follows:[10]

> (a) under s 9A(2)(a), to *confirm* that the ordinary meaning deduced as aforesaid is, after all the correct and intended meaning having regard to any extraneous material that further elucidates the purpose or object of the written law;
>
> (b) under s 9A(2)(b)(i), to *ascertain* the meaning of the text in question when the provision on its face is *ambiguous or obscure*; and
>
> (c) under s 9A(2)(b)(ii), to *ascertain* the meaning of the text in question where having deduced the ordinary meaning of the text as aforesaid, and considering the underlying object and purpose of the written law, such ordinary meaning is *absurd or unreasonable*.

31. In respect of (c), there does not appear to be anything absurd or unreasonable in a person being permitted to use the expressive elements of a work when performing computational data analysis. As explained above, such use may be necessary to enable computational data analysis at all.

32. As to (b), the text of section 243, on its face, is not ambiguous (that is, open to one or more interpretations) or obscure. Even if computational data analysis were (contrary to the statutory language) limited to the two uses in paragraphs (a) and (b) of that section, the second use in paragraph (b) recognises that the work or recording may be used. There is no indication in paragraph (b) that only the unprotected elements of the work or recording can be used. To the extent any ambiguity is created, it is created by the extraneous material rather than the text of the provision itself.

33. With respect to (a), the Court of Appeal of Singapore has queried whether there is a "real point" in considering extraneous material in such circumstances. As the court explained:[11]

> If the extraneous material does not confirm the ordinary meaning—or even calls that ordinary meaning into question—the court is not permitted to use that extraneous material as a basis for departing from the ordinary meaning, as that is only permissible when reference is made under s 9A(2)(*b*). If instead the extraneous material does confirm the ordinary meaning, that too would not alter the result: the court would have had to apply the ordinary meaning in any event since s 9A(2)(*b*) was not invoked. It may seem from this that there is no point in referring to the extraneous material either way.

---

10 *Attorney-General v Ting Choon Meng and another appeal* [2017] 1 SLR 373; [2017] SGCA 6 at [65].

11 *Tan Cheng Bock v Attorney-General* [2017] 2 SLR 850; [2017] SGCA 50 at [48].

34. The court's answer to this issue was as follows:[12]

> In our judgment, the explanation for this is a practical one: even though extraneous material referred to under s 9A(2)(a) alone cannot alter the outcome of a decision, it is useful for demonstrating the soundness—as a matter of policy—of that outcome.

35. The court also added:[13]

> It also bears mentioning that extraneous material cannot be used "to give the statute a sense which is contrary to its express text" (*Seow Wei Sin v PP* [2011] 1 SLR 1199 at [21]) save perhaps in the very limited circumstances identified in s 9A(2)(b)(ii) of the [Interpretation Act 1965] …). This echoes the broader principle that the proper function of the judge when applying s 9A of the [Interpretation Act 1965] is to interpret a given statutory provision. Although purposive interpretation is an important and powerful tool, it is not an excuse for rewriting a statute … . The authority to alter the text of a statute lies with Parliament, and judicial interpretation is generally confined to giving the text a meaning that its language can bear. Hence, purposive interpretation must be done with a view toward determining a provision's or statute's purpose and object "as reflected by and in harmony with the express wording of the legislation": *PP v Low Kok Heng* [2007] 4 SLR(R) 183 at [50].

36. In respect of Parliamentary debates on Bills containing the legislative provision in question, including the speech made by the Second Minister for Law when the Bill was moved, the court said:[14]

> The court should guard against the danger of finding itself construing and interpreting the statements made in Parliament rather than the legislative provision that Parliament has enacted.

37. In light of the above, it does not appear reasonable to adopt an interpretation of section 243 that restricts computational data analysis to the unprotected elements of a work or a recording of a protected performance.

---

12 *Tan Cheng Bock v Attorney-General* [2017] 2 SLR 850; [2017] SGCA 50 at [49].

13 *Tan Cheng Bock v Attorney-General* [2017] 2 SLR 850; [2017] SGCA 50 at [50].

14 *Tan Cheng Bock v Attorney-General* [2017] 2 SLR 850; [2017] SGCA 50 at [52(b)], citing *Attorney-General v Ting Choon Meng and another appeal* [2017] 1 SLR 373; [2017] SGCA 6 at [70]).

Subject matter of computational data analysis

38. There are two types of material that may be the subject of computational data analysis: "works" and "recordings of protected performances".

Works

39. A "work" is:[15]

    a. an "authorial work";

    b. a published edition of an "authorial work";

    c. a sound recording;

    d. a film;

    e. a broadcast;

    f. a cable programme.

40. An "authorial work" is a literary, dramatic, musical or an artistic work.[16]

Recordings of protected performances

41. A "protected performance" means a performance protected under Part 4 of the Act.[17]

42. A performance is protected under Part 4 of the Act if it is:[18]

    a. a "qualifying performance"; and

    b. given live (i) in Singapore; or (ii) by a "qualified individual".

43. A "qualifying performance" is:[19]

    a. a performance (including an improvisation and a performance that uses puppets) of a dramatic work or part of the dramatic work;

    b. a performance (including an improvisation) of a musical work or part of the musical work;

    c. the reading, recitation or delivery of a literary work or part of the literary work, or the recitation or delivery of an improvised literary work;

    d. a performance of a dance;

    e. a performance of a circus act or a variety act or any similar presentation or show.

44. The Act also expressly stipulates certain performances that are not "qualifying performances".[20]

---

15 Copyright Act 2021, section 8.

16 Copyright Act 2021, section 9.

17 Copyright Act 2021, section 7(1), definition of "protected performance".

18 Copyright Act 2021, section 173.

19 Copyright Act 2021, section 37(1)(a).

20 Copyright Act 2021, section 37(1)(b).

45. A “qualified individual” is a Singapore citizen or a “Singapore resident”.[21] A “Singapore resident” is an individual who is (a) resident in Singapore; or (b) residing in Singapore under a valid pass lawfully issued to him or her under the Immigration Act 1959 to enter and remain in Singapore for any purpose other than a temporary purpose.[22]

46. A “recording”, in relation to a protected performance, means a sound recording of the performance or a substantial part of the performance, and includes a copy of such a recording.[23]

[21] Copyright Act 2021, section 77.

[22] Copyright Act 2021, section 79.

[23] Copyright Act 2021, section 38(1).

Permitted uses

47. Where an act in relation to a work is a permitted use, the act is not an infringement of any copyright in the work.[24] Similarly, where an act in relation to a protected performance is a permitted use, the act is not an infringing use of the performance.[25]

48. Division 8 of Part 5 of the Act permits *X* to perform two acts in respect of works and recordings of protected performances: (a) to make a copy of the work or performance; and (b) to communicate the work or performance to the public. Each permitted use is subject to conditions (see the "Conditions to the permitted uses" section below).

Making more than one copy

49. Section 244(1) states that it is a permitted use for *X* to "make *a copy*" of a work or a recording of a protected performance. On its face, this suggests that *X* is only permitted to make a single copy. However, in the Act, words in the singular include the plural.[26] So, "make a copy" includes "make *copies*".

Communication of a prepared version of a copy

50. We observe that section 244(4) states that it is a permitted use if the communication "is made *using a copy* made in circumstances to which subsection (1) [of section 244] applies" (emphasis added).

51. Does this mean that *X* cannot communicate a version of the copy that was prepared for computational data analysis under section 244(2)(a)(ii) for the purposes of section 244(4)?

52. The text of section 244(4) states that the communicated copy must be made in circumstances "to which subsection (1) [of section 244] applies". Subsection (1) of section 244 comprises the permitted use of copying. But that permitted copying is expressly subject to the conditions of subsection (2) of section 244 being met. Making a copy for the purpose of preparing a work or recording for computational data analysis is found in subsection (2). As a result, it does not appear that *X* is prevented from communicating a prepared copy of work for the purposes of section 244(4).

53. Putting the purely textual analysis aside, such a restriction may undermine the purpose of section 244(4). To verify the results of *X*'s computational data analysis, it may be necessary for the verifier to review how the copy of the work or recording was prepared for such analysis. That is, the performance of *X*'s computational data analysis may be tied to how *X* prepared the copy of the work or recording for such analysis, with poor preparation leading to poor results.

54. Moreover, the prepared version of the copy of the work or recording itself may not itself be subject to copyright and thus communication of it to a third party, whether in compliance with section 244(4) or not, is not an act of infringement (see the "Data cleaning and pre-processing" section below).

---

24 Copyright Act 2021, section 183(1).

25 Copyright Act 2021, section 183(2).

26 Interpretation Act 1965, section 2(1).

Conditions to the permitted uses

Copying

55. There are five conditions to the permitted copying of a work or a recording of a protected performance under section 244(1) of the Act.

56. *First*, *X* must make the copy for the purpose of computational data analysis or preparing the work for such analysis.[27]

57. *Second*, *X* must not use the copy for any other purpose.[28]

58. *Third*, *X* cannot supply the copies to any other person, save for two exceptions.[29] Those two exceptions are:

    a. verifying the results of *X*'s computational data analysis;[30] or

    b. collaborative research or study relating to the purpose of *X*'s computational data analysis.[31]

    (These two exceptions mirror the two requirements to permitted communication of a work or recording under section 244(4), discussed below.)

59. *Fourth*, *X* must have had lawful access to the relevant material when making the copy.[32]

60. *Fifth*, the relevant material must not be an "infringing copy".[33] *X* may not know that the material is an infringing copy. So, there is an exception to deal with that circumstance.[34]

Communication to the public

61. There are two conditions to the permitted use of communicating a work or recording to the public under section 244(4) of the Act.

62. *First*, the communication must be made using the copy of the work or recording made under section 244(1).[35] That is, the copy was made for the purpose of computational data analysis or preparing the work or recording for such analysis.[36]

63. *Second*, *X* must not supply (whether by communication or otherwise) the copy to any person other than for the purpose of:

---

27 Copyright Act, section 244(2)(a).

28 Copyright Act, section 244(2)(b).

29 Copyright Act 2021, section 244(2)(c) chapeau.

30 Copyright Act 2021, section 244(2)(c)(i).

31 Copyright Act 2021, section 244(2)(c)(ii).

32 Copyright Act 2021, section 244(2)(d).

33 Copyright Act 2021, section 244(2)(e)(i). An "infringing copy" is defined in section 98 of the Copyright Act 2021.

34 Copyright Act 2021, section 244(2)(e)(ii).

35 Copyright Act 2021, section 244(4)(a).

36 Copyright Act 2021, section 244(1) read with section 244(2)(a).

a. verifying the results of *X*'s computational data analysis;[37] or

b. collaborative research or study relating to the purpose of *X*'s computational data analysis.[38]

X's purposes

64. The phrase "for the purpose of" is used several times in section 244 of the Act. There is the condition that *X*'s copying of the material is "for the purpose of" computational data analysis (or preparing the material for such analysis),[39] the prohibition on *X* using the copies made "for any other purpose"[40] and the exception to the prohibition on supplying the copies other than "for the purpose" of verifying the results of *X*'s computational data analysis[41] or collaborative research or study relating "to the purpose" of *X*'s computational data analysis.[42]

65. There is no express limitation on the purpose for which *X* undertakes computational data analysis. For example, *X* is not limited to performing computational data analysis for a non-commercial purpose.[43] Where *X's* purpose for undertaking computational data analysis may be relevant, it is in respect of the "collaborative research or study" exception to the prohibition on supplying copies of material to any other person. That is, *X* is permitted to supply a copy of a work or recording (made under section 244(1)) to another person for the purpose of collaborative research or study "relating to the purpose of the computational data analysis carried out by *X*.[44]

66. In respect of *X*'s copying of material, we observe that the first condition expresses alternative individual purposes: *X* may make a copy of a work or a recording of a protected performance for the purpose of computational data analysis *or* for the purpose of *preparing the work or recording* for such analysis. If *X* is permitted to make a copy of a work or recording for the purpose of preparing the work or recording for the subsequent computational data analysis, then it appears implied that *X* is also permitted to use that prepared version of the work or recording for the purpose of computational data analysis. It would not make sense if *X* could make a copy of a work or recording for the purpose of preparing the work or recording for computational data analysis but could not use that prepared version for computational data analysis because it is not expressly stated in section 244(2)(a).

67. Save for section 244(2)(e)(ii), there does not appear to be any reason to suggest that the use of "purpose" in section 244 of the Act is to be understood other than objectively.

---

37 Copyright Act 2021, section 244(4)(b)(i).

38 Copyright Act 2021, section 244(4)(b)(ii).

39 Copyright Act 2021, section 244(2)(a).

40 Copyright Act 2021, section 244(2)(b).

41 Copyright Act 2021, section 244(2)(c)(i) and (4)(b)(i).

42 Copyright Act 2021, section 244(2)(c)(ii) and (4)(b)(ii).

43 *Cf.,* Copyright, Designs and Patents Act 1988 (United Kingdom) section 29A(1)(a).

44 Copyright Act 2021, section 244(2)(c)(ii) and (4)(b)(ii).

Supply

68. The word "communication" has a defined meaning under section 61 of the Act: see **Annexure 1**: Permitted uses. The word "supply" is not, however, defined by the Act. It appears that "supply" may have a broader meaning than "communicate". *First*, in the respective chapeaus to sections 244(2)(c) and 244(4)(b), "communication" is expressed, parenthetically, as an instance of "supply". *Second*, some permitted uses go beyond the acts comprised in copyright and what would otherwise be an infringing use of a protected performance.[45] As "supply" is not one of the expressed exclusive rights in a work or recording of a protected performance, the permitted use of supplying a work or recording appears to be an instance of a permitted use going beyond the acts comprised in copyright.

69. Section 244(5)(a) states that the supply of copies of any material in circumstances to which section 244 applies "is not to be treated as publishing the material (or any work or recording included in the material". This provision appears to have been included because the *supply of copies* of material to the public may constitute *publication* of that material.[46] Paragraph (b) of section 244(5) appears to have been included for a similar reason: duration of copyright in the material may be affected by the date of its publication.

Collaborative research or study

70. The phrase "research or study" (with or without the prepended adjective "collaborative") is not defined in the Act.

71. Other than the inclusion of the word "collaborative", the phrase is expressed without qualification (for example, the noun "study" is not modified by the adjective "private"[47]).

72. In ordinary usage, "study" refers to the systematic examination, learning, or analysis of existing material, while "research" refers to the structured investigation or experimentation directed toward the discovery, validation, or refinement of new knowledge. These concepts are not mutually exclusive and frequently overlap, particularly in technical and scientific domains.

73. In respect of the qualifier "collaborative", collaboration commonly occurs where multiple persons contribute different expertise, resources, or functions toward a shared research or study objective. In the context of generative AI development, collaboration may include joint work between subject-matter experts, data specialists, and technical developers, whether within a single organisation or across organisational boundaries.

Consequence of non-compliance with a condition to a permitted use

74. Non-compliance with one or more of the above conditions of a permitted use, of itself, is not actionable. That is, such non-compliance *per se* does not establish in the rightsholder a

---

45 Copyright Act 2021, section 185.

46 See Subdivision (3) ("Publishing") of Division 3 ("Acts relating to works and performance") of Part 2 ("Interpretation") of the Copyright Act 2021.

47 *Cf.,* for example, Copyright Act 1987 (Singapore), section 35 ("Fair dealing for purpose of research or study") where the phrase "research or *private* study" was used in subsection (1) of the section and "research" was defined, in subsection (5) of the section, as *excluding* "industrial research, research carried out by bodies corporate (not being bodies corporate owned or controlled by the Government), companies, associations or bodies of persons carrying on any business". The legislature subsequently repealed subsection (5): Copyright (Amendment) Act 1998, section 5. This was done by the legislature "to make the defence [of fair dealing for the purpose of research or study] available for commercial research provided that the dealing was fair": *Global Yellow Pages v Promedia Directories Pte Ltd* [2016] 2 SLR 165; [2016] SGHC 9 at [393].

statutory cause of action against *X* for that act of non-compliance. Nor does it result in an offense under the Act.

75. Nor does such non-compliance necessarily mean that the (non-compliant) use is an act that infringes copyright. The test for copyright infringement is not whether a person has failed to comply with the conditions of a permitted use. Rather, the test is whether (a) a person does, or authorises the doing of, in Singapore an act comprised in copyright; and (b) the person neither owns the copyright nor has the license of the copyright owner.[48]

Territorial scope of computational data analysis

76. It is a "well-established" principle that "a statute generally operates within the territorial limits of the Parliament that enacted it".[49] This follows the "near universal" rule of international law that legislative sovereignty is territorial, meaning that legislative sovereignty "may be exercised only in relation to persons and things within the territory of the state".[50]

77. Consistent with these principles, Division 8 of Part 5 of the Act is limited to things and the activities of persons that have a connection to Singapore.

Territorial scope of the materials to which the permitted uses apply

78. In respect of each type of material to which computational data analysis may be applied, there is a connection (aka territorial nexus) with Singapore. There are two main types of connecting factors.

79. The first type of connecting factor is where the material is made by a "qualified person". A person is a qualified person if the person is (a) a "qualified individual"; or (b) a body corporate incorporated *in Singapore* under any written law.[51] An individual is a "qualified individual" if he or she is (a) a *Singapore citizen*; or (b) a "Singapore resident".[52] An individual is a Singapore resident if he or she is (a) resident *in Singapore*; or (b) residing *in Singapore* under a valid pass lawfully issued to him or her under the Immigration Act 1959 to enter and remain in Singapore for any purpose other than a temporary purpose.[53]

---

48 Copyright Act 2021, section 146(1).

49 *Ng Kok Wai v Public Prosecutor* [2024] 3 SLR 1516; [2023] SGHC 306 at [25], quoting *Public Prosecutor v Taw Cheng Kong* [1998] 2 SLR(R) 489; [1998] SGCA 37 at [66].

50 *The Republic of the Philippines v Maler Foundation and others and other appeals* [2014] 1 SLR 1389; [2013] SGCA 66 at [60], quoting *Eram Shipping Company Ltd & Ors v Hong Kong and Shanghai Banking Corporation Ltd* [2003] UKHL 30 at [80].

51 Copyright Act 2021, section 78.

52 Copyright Act 2021, section 77.

53 Copyright Act 2021, section 79.

80. The second type of connecting factor is a group of alternative sub-factors. This is where the material is made or first published (a) *in Singapore*, (b) by or under the direction or control of the *Government of Singapore*; or (c) on or after 10 April 1987 by or under the direction or control of a prescribed international organisation. (A list of "prescribed international organisations" is provided in regulation 28 of the Copyright Regulations 2021, which includes the Asian Development Bank, Southeast Asian Ministers of Education Organisation and the United Nations—being organisations of which Singapore is a member.)

81. In **Annexure 2**: Connecting factors between Singapore and a work or a recording , we explain the connecting factors between Singapore and each of the types of material to which computational data analysis can apply.

Territorial scope of the permitted uses

82. As noted above, Division 8 of Part 5 permits *X* to perform two acts (subject to certain conditions) that would otherwise infringe copyright:

    a. making a copy of a work or a recording of a protected performance; and

    b. communicating a work or a recording of a protected performance to the public.

83. The question is then *where* these activities can occur. That is, must the copying or communication occur in Singapore, or can these activities occur outside Singapore? It appears the answer is that these activities, to be permitted uses, must occur in Singapore.

84. Under the Act, copyright is infringed if (a) a person does *in Singapore*, or authorises the doing *in Singapore* of, any act comprised in the copyright; and (b) the person neither owns the copyright nor has the licence of the copyright owner.[54]

85. In relation to "protected performances", the "infringing uses" of making a copy of a recording of the performance[55] and making a recording of the performance available to the public[56] are only infringing uses if they are done *in Singapore*.[57]

86. The permitted uses permit that which would otherwise be an infringement. So, by parity of reasoning, the permitted uses similarly are limited to uses that occur *in Singapore*.

Extension of Division 8 of Part 5 to reciprocating countries

87. Subsistence of copyright under the Act is extended to works (excluding cable programmes) and qualifying performances of a "reciprocating country" which is a party to the Berne Convention or a member of the World Trade Organisation.[58] We observe that over 180 nations are party to the Berne Convention, and the World Trade Organisation has over 160 members.

88. This extension is achieved in two ways.[59] The first way is by extending the definition of "qualified individual" or "qualified person" to include the (a) citizens or nationals; (b) residents; or (c) body corporates incorporated under the law, of a reciprocating country. The second way is by reading any reference to "Singapore", in respect of where a work or qualifying performance is made, as including a reference to a reciprocating country.

---

54 Copyright Act 2021, section 146(1).

55 Copyright Act 2021, section 175(a)(ii).

56 Copyright Act 2021, section 175(a)(iv).

57 Copyright Act 2021, section 175(b)(ii).

58 Copyright Regulations 2021, regulation 6(1).

59 See Copyright Regulations 2021, regulations 9–16.

89. The question then is whether Division 8 of Part 5 is also extended to these foreign materials. That is, can *X* copy and communicate these foreign materials under section 244, the copyright in which is recognised under the Act, without engaging in copyright infringement *under Singapore law*? The answer appears to be "yes". Regulation 8(2) of the Copyright Regulations 2021 states that "the application of the other provisions of the Act is extended accordingly." So, while subsistence of copyright in foreign materials is recognised under Singapore law, such materials are also subject to the permitted uses in Division 8 of Part 5.

(The outcome is that a rightsholder in a reciprocating country gets the benefit of protection under the Act but is equally subject to the permitted uses under the Act.

If the permitted uses under the Act did not apply to reciprocally-recognised copyright, it would mean that a rightsholder in a reciprocating country would have greater protection under Singapore law than a local Singapore rightsholder.

As discussed in "Singapore as an AI development hub" below, if a copyright infringement claim is brought against *X* in a reciprocating country, it is unlikely that *X* could raise Division 8 of Part 5 of the Act as a defence to that claim. The Division is a foreign law from the perspective of that other country and to enforce it in that other country may be against the public policy of that other country.)

Outsourcing computational data analysis

Recommendation. We recommend that consideration be given to amending paragraphs (2)(c) and 4(b) of section 244 of the Act to include an additional exception such that *X* may supply a copy of a work or a recording of a protected performance to another person for the purpose of that other person performing computational data analysis on *X*'s behalf.

90. Many individuals and businesses may desire to provide products and services that leverage AI, but not all will have the ability (or it may simply be inefficient for them) to develop the requisite knowledge and expertise in-house. As a result, computational data analysis is a task that may need (or is simply more efficient) to be procured from an expert third party.

91. We observe that there is no express provision in section 244 of the Act to permit *X* to supply copies of a work or a recording of a protected performance to another person for the purpose of that other person performing computational data analysis on *X*'s behalf (i.e., *X* outsourcing computational data analysis to *Y*).

> (1) If the conditions in subsection (2) are met, it is a permitted use for a person (*X*) to make a copy of any of the following material:
>
> (a) a work; and
>
> (b) a recording of a protected performance.
>
> . . .
>
> (2) The conditions are —
>
> . . .
>
> (c) *X* does not supply (whether by communication or otherwise) the copy to any person other than for the purpose of —
>
> (i) verifying the results of the computational data analysis carried out by *X*; or
>
> (ii) collaborative research or study relating to the purpose of the computational data analysis carried out by *X*;
>
> . . .
>
> (4) It is a permitted use for *X* to communicate a work or a recording of a protected performance to the public if —
>
> (a) the communication is made using a copy made in circumstances to which subsection (1) applies; and
>
> (b) *X* does not supply (whether by communication or otherwise) the copy to any person other than for the purpose of —
>
> (i) verifying the results of the computational data analysis carried out by *X*; or

(ii) collaborative research or study relating to the purpose of the computational data analysis carried out by *X*.

92. Consider the following scenario. *X* wants to perform computational data analysis on copies of the relevant material to which it has lawful access. But *X* lacks the technical ability to perform such analysis itself. So, *X* engages *Y*, an expert in machine learning, to perform computational data analysis for a fee. We assume that for *Y* to perform such analysis on *X*'s behalf, *X* would need to perform two acts: (a) make copies of the works or recordings of the protected performances that will be the subject of the computational data analysis; and (b) supply those copies, by whatever means, to *Y* so that *Y* may perform such analysis on *X*'s behalf.

(There are two common commercial arrangements for this to happen. In the first arrangement, *Y* is a consultancy firm. *X* engages *Y* to provide computational data analysis services. *X* makes copies of the relevant material and supplies them to *Y* so that *Y* may perform such analysis. *Y* provides the results of the analysis for *X* to use or deploy. In the second arrangement, *Y* is a systems integrator. *X* engages *Y* to develop an IT system. *X* makes copies of the relevant material and hosts those copies on the computing infrastructure it owns or controls, such as servers in its data centre or cloud computing infrastructure that it has subscribed to, and provides *Y* with access to such computing infrastructure (i.e., the development environment) so that *Y* can access the copies.[60] *Y* performs computational data analysis as part of the system development services it provides for *X*. Unlike the first arrangement, *X* here does not supply *Y* with the copies of the relevant material. Instead, it provides *Y* with access to those copies.)

93. The first act that *X* needs to perform, i.e., to copy the relevant material, is an act of copyright infringement absent ownership, a license or a statutory permitted use because making a copy is one of the exclusive rights comprised in the copyright of a work.[60] Section 244, however, permits *X* to make copies of the relevant material for the purpose of computational data analysis or preparing those copies for such analysis.

94. The second act that *X* needs to perform, i.e., to supply, by whatever means, the copies made to *Y*, may not infringe copyright as such supply may come within the definition of "communicate" in the Act. (In the second arrangement, *X* only provides *Y* with access to the copies made.) The Court of Appeal of Singapore has held that communication made to a recipient with an existing contractual relationship with the communicator is not communication *to the public*[61] and that communication made to third parties "privately and individually" is also not communication *to the public*.[62]

95. However, one of the conditions to the permitted copying for the purpose of computational data analysis or preparing the copies made for such analysis (i.e., the first act that *X* needs to perform) is that *X* cannot supply those copies "to any person" save for that other person: (a) to verify the results of the computational data analysis *carried out by X*; or (b) for collaborative research or study relating to the purpose of the computational data analysis *carried out by X* (emphasis added).[63]

---

60 Copyright Act 2021, section 112(1)(a) (literary, dramatic and musical works); section 113(a) (artistic works); section 118 (published editions of an authorial work); section 121(a)(ii) (sound recordings); section 124(a) (films); section 127(a) (broadcasts) and section 131(1)(a) (cable programmes) read with Copyright Act 2021, section 146(1).

61 *RecordTV Pte Ltd v MediaCorp TV Singapore Pte Ltd and others* [2011] 1 SLR 830; [2010] SGCA 43 at [25].

62 *RecordTV Pte Ltd v MediaCorp TV Singapore Pte Ltd and others* [2011] 1 SLR 830; [2010] SGCA 43 at [26].

63 Copyright Act 2021, section 244(2)(c).

96. Even where computational data analysis is performed by *Y*, we assume *X* will, in most cases, nevertheless be involved in such analysis. For example, in the development of a generative AI model, *X* will invariably be involved in selecting the dataset for model training or analysis (see the "Data collection" section below), reviewing the results of data cleansing and pre-processing (see the "Data cleaning and pre-processing" section below), preparing samples for "fine-tuning" (see the "Fine-tuning" section below)and ultimately evaluating and benchmarking the model's outputs (see the "Evaluation and benchmarking" section below). Such involvement is consistent with the definition of computational data analysis under section 243. In such a scenario, the computational data analysis is not *exclusively* performed by *Y*, as *X* would, for much of the process (if not all of it) be the ultimate decision-maker regarding the analysis.

97. Separately, the development of a generative AI model necessarily involves both "study", such as the examination and analysis of the material to identify linguistic, structural, or sematic patterns, and "research", such as experimentation with model architecture, training methods, parameters and evaluation techniques to generate new functional capabilities or insights. During commercial R&D activities, while computational data analysis is performed by *Y*, *X* will still ordinarily be involved in research and study because it needs to do the activities described above (i.e., select dataset for training, review results of data cleansing, rate outputs, etc). The involvement by *X* may amount to "collaboration" between *X* and *Y* in research or study relating to the purpose of computational data analysis.

98. The activities by *Y* are all intended to support *X*'s computational data analysis. That said, difficulty arises in identifying clear textual support for this in section 244.[64] The provision says computational data analysis *carried out by X*. It is unclear what level of involvement of *X* in the computational data analysis is required in order to qualify such analysis as being carried out by *X*.

99. As the Second Minister for Law explained above, Division 8 of Part 5 was enacted to support Singapore's "Smart Nation initiatives", the nation's "push towards data-driven innovation" as well as Singapore's efforts to grow its AI and technology sectors. Restrictions on the outsourcing of computational data analysis may undermine those efforts.

100. As such, we recommend that consideration be given to amending paragraphs (2)(c) and 4(b) of section 244 of the Act to include an additional exception such that *X* may supply a copy of a work or a recording of a protected performance to another person for the purpose of that other person performing computational data analysis on *X*'s behalf. This will provide clarity for situations where computational data analysis needs to be *fully* outsourced to a third party.

---

64 It has been argued that outsourcing to a service provider is permissible where the service provider carries out the computational data analysis "on behalf and to the benefit of" the principal. See Yeong Zee Kin, *Technology Regulation in the Digital Economy* (Singapore Academy of Law, 2023) at [8.50].

Lawful access

Recommendation. We recommend that consideration be given to amending section 244 of the Act to include a provision with words to the effect that "X *is deemed not to have lawful access to the first copy if access to the first copy is obtained in breach of a contract (ignoring any terms that are void by virtue of section 187)*".

101. *X* cannot make a copy of a work or a recording of a protected performance for computational data analysis unless it has "lawful access" to that work or recording.[65]

102. The phrase "lawful access" is not defined, although two illustrations are provided of circumstances where *X* does not have lawful access (discussed below). "Lawful" is generally contrasted with "unlawful". An act is "unlawful" if it is not permitted by law. Thus, an act is "lawful" if it is permitted by law.

103. The phrase appears directed at the way by which *X* obtains the material (rather than *X*'s rights in respect of the material) that is subsequently copied. This is supported by the illustrations. In the first illustration, *X* obtains the material by circumventing paywalls (in contravention of the law) and, in the other, *X* obtains the material in breach of the terms of use of a database (in breach of contract).

104. It appears that *X* must have lawful access to the material *at the time of copying* that material. Thus, if *X*'s access to the relevant material was lawful at the time of copying but subsequently became unlawful before the computational data analysis was performed, there does not appear to be any prohibition on *X* nevertheless performing such analysis using that copy. This situation may arise, for example, where the material was copied from an open source, but was subsequently moved behind a paywall before computational data analysis was performed.

Access in contravention of law

105. The first illustration of the absence of lawful access is as follows:

> (a) *X* does not have lawful access to the first copy if *X* accessed the first copy by circumventing paywalls.

106. A paywall restricts access to online content to users who have paid for such access. Paywalls are enforced using a combination of *client-side* and *server-side* technologies. In web architecture, *client-side* refers to operations that occur in the user's web browser, while *server-side* refers to processes performed on the website's server before any content is sent to the browser.

107. The question is then: *what makes that act of circumventing a paywall unlawful?*

108. In **Annexure 3**: Statutory provisions prohibiting unlawful access, we attempt to provide the answer to this question by reference to the statutory prohibitions against: (a) circumvention of an access control measure;[66] and (b) unauthorised access to computer material.[67]

---

65 Copyright Act 2021 section 244(2)(d).

66 Copyright Act 2021, section 439.

67 Computer Misuse Act 1993, section 3(1).

Access in breach of contract

109. The second illustration of the absence of lawful access is as follows:

> (b) *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 (ignoring any terms that are void by virtue of section 187).

110. Terms of use (also known as "terms of service" or "terms and conditions") are contracts between service providers and their customers (the service being provided in the illustration is "use of a database").

111. To use the service, the customer must agree to abide by the terms of service. Where the service is an online service, those terms of service are often expressed in a "click wrap" agreement—an online contract where a user signifies their agreement to the terms of the service by clicking a button or checking a box, often labelled "I agree" or "I accept".

112. If, after acceptance of the terms of service, the customer does not abide by those terms, he or she will be in breach of contract. That breach may have consequences at law (e.g., liability to pay damages) or under the contract (e.g., termination of use of the service).

Policy rationale

113. One may query why the legislature is concerned with whether *X* complies with its private law obligations. If *X* engages in a breach of contract, then the obligee will have a *prima facie* right of action for breach of contract against *X*.

(As discussed in the "Contractual exclusion or restriction of computational data analysis" section below, the legislature's primary concern is that *rightsholders* use contract terms in a way that would exclude or restrict the two permitted uses stipulated in Division 8 of Part 5. In anticipation of this potential problem, the legislature introduced section 187 of the Act to render such terms void.)

114. We observe that the legislature does not appear concerned with whether *X* complies with its contractual obligations *in general*, including *X's* compliance with its contractual obligations to the obligee regarding subsequent *use* of the material copied pursuant to section 244(1) of the Act.

(We observe that one of the conditions to the permitted copying under section 244(1) is that "*X* does not use the copy for any other purpose".[68] We do not take this requirement as overriding any agreement between the rightsholder and *X* which permits *X* to use a copy for a non-computational data analysis purpose.)

115. Rather, the legislature appears concerned with circumstances where in *accessing* the material (which is subsequently copied), *X* breaches the terms of a contract.

(For example, assume a term of a contract contained words to the effect that "*you may not distribute copies of our content, save for legally compliant computational data analysis*". Such a term would not appear to fall foul of section 187 as it does not exclude or restrict *X* from making a copy, or supplying that copy, for the purpose of computational data analysis. If *X* distributed the works to a third party in breach of this term (i.e., distribution other than per the permitted use in section 244(4)), such a breach would not appear to fall within section 244(2)(d) because, while there is a breach by *X* of the contract, the act that constituted the breach was not regarding *X*'s *access* to the material, but *X*'s subsequent *use* of that material.)

---

68 Copyright Act 2021, section 244(2)(b).

116. The question is then why the legislature is concerned with this particular form of contractual non-compliance?

117. By way of illustration, consider the following examples:

    a. a term of the contract prohibits access to non-public areas of the online service and *X* accesses those areas and copies relevant materials;

    b. a term of the contract prohibits access to an online service by a means other than that provided by the service provider and *X* accesses the services to copy relevant materials via an alternative means; and

    c. a term of the contract prohibits use of an online service in any manner that could disable, overburden, damage, or impair the service and *X* uses an automated means to copy relevant materials from the service that overburdens the service, resulting in an outage.

118. Assume in each example that *X* copies the content for the purpose of computational data analysis (or preparing the copy for such analysis) in accordance with section 244(2)(a) and *X* is otherwise compliant with the conditions in paragraphs (b) and (c) of section 244(2).

119. In each case, *X* has *accessed* the content in breach of a terms of the contract. The service provider may be able to take legal action against *X* for the relevant breach. However, absent the condition in section 244(2)(d), the service provider could not take legal action against *X* for copyright infringement. *X* would be free to use the copies it obtained, in breach of the contract, for computational data analysis.

120. A rightsholder may feel aggrieved by this outcome, particularly as the legislature has nullified its ability to exclude or restrict computational data analysis by virtue of section 187.

121. So, it appears that the legislature included the condition expressed in section 244(2)(d) in response to such situations.

    (A rightsholder or its exclusive licensee may make *access* to material conditional on compliance with terms that govern subsequent *use* of the material. For example, the contract may include words to the effect that "*Your access to our service is conditional on compliance with our terms regarding use of our material*". In other words, *X*'s *access* is subject to one or more *conditions subsequent* regarding *use*.

    In this example, breach by *X* of a condition subsequent regarding *use* may not necessarily establish a breach of contract falling within section 244(2)(d). That is, there is an argument that the example term here simply specifies an *outcome*—termination of the service—triggered on breach of a term regarding *use*; rather than imposing an obligation on *X* regarding *access* itself. The answer will depend on the construction of the term read in light of the contract as a whole.)

The unlawfulness of a breach of contract

122. Breach of contract is a civil (or private) wrong, not a criminal (or public) wrong. If an obligor breaches a contract, the obligee may have a (private) cause of action for the breach; but the obligor is not exposed to a potential action by the state for that breach. As such, it does not appear to be correct to describe breach of contract as an unlawful activity.

123. As Phang JCA (as he then was) explained in *Tan Ng Kuang and another v Jai Swarup Pathak* (emphasis added):[69]

> We recognise that, *absent any criminal behaviour, it is a person's prerogative to decide whether he wishes to breach a private obligation*, and there are even legal scholars who have propounded the theory of "efficient breach" of contract (which, in a nutshell, argues that damages are preferable to contractual performance when the latter provides less utility than the former (see, for example, Gregory Klass, "Efficient Breach" in *The Philosophical Foundations Of Contract Law* (Gregory Klass, George Letsas and Prince Saprai eds) (Oxford University Press, 2014) at pp 362–387)).

124. His Honour observes that an obligor enjoys a "prerogative"—that is, a right or privilege—to elect non-performance of a contractual obligation. In support of the observation, His Honour refers to the theory of efficient breach. The theory posits that a breach may be economically rational where the obligor's cost of performance exceeds the utility derived from it, such that paying damages is preferable to performance. Some criticise the theory because it appears to legitimise deliberate non-performance. Yet, as His Honour implies, the law does not prohibit such conduct *per se*.

125. However, where the act or omission constituting breach also amounts to a crime, the obligor's "prerogative" ceases to exist. The obligor cannot claim a right to engage in criminal conduct.

126. So, it is difficult to characterise breach of contract, in and of itself, as unlawful. To do so would appear conceptually inconsistent: in circumstances where the obligor's conduct is not criminal, it cannot simultaneously be said that the breach is both: (a) the exercise of a right or privilege of the obligor; and (b) an act prohibited by law.

    (For completeness, we observe that breach of contract is also not a conduct that is *per se* unlawful (aka a *malum in se*), such as murder, rape or theft; nor a conduct that is unlawful only because it is prohibited by statute (aka a *malum prohibitum*), as it appears no statute prohibits an obligor from breaching a contract.[70])

Illustration (b) to section 244(2)(d)

Editorial note

The policy intention – as discussed in paragraphs 113 to 121 above – appears to be that while excluding the right of computational data analysis is prohibited, rights holders have the liberty to constrain the manner of accessing copyrighted materials. However, there is case law to the effect that a breach of contract is not *per se* unlawful. That case law may defeat the policy intention that breaches of contractual terms constraining the manner of access should be unlawful for the purpose of section 244(2)(d) of the Act. That much is clear from the illustration.

127. The question is then: what is the effect of the illustration? Does it have the effect of deeming that a breach of contract is unlawful for the purpose of section 244(2)(d)? The answer appears to be "no".

---

69 [2022] 3 SLR 788; [2021] SGHC 232 at [87].

70 For a discussion on these concepts, see *Tan Chor Jin v Public Prosecutor* [2008] 4 SLR(R) 306; [2008] SGCA 32 at [31].

128. As the High Court of Singapore explained in *Shaikh Farid v Public Prosecutor and other appeals*:[71]

> Illustrations are only "examples of how it was anticipated that the law would apply to a given factual situation". They "do not, therefore, have the effect of altering the scope of the law as defined in the substantive provision and are not 'binding'". They also "do not curtail or expand the ambit of the provision itself". This means that if any inconsistency emerges between the substantive provision and the illustrations, the substantive provision "will prevail" … . These principles of statutory construction are well-settled.

Recommendation

129. In light of the above, we recommend that consideration be given to amending section 244 of the Act to include a provision with words to the effect that "X *is deemed not to have lawful access to the first copy if access to the first copy is obtained in breach of a contract (ignoring any terms that are void by virtue of section 187)*".

(Our recommendation should be read in light of our observations regarding the difficulty in distinguishing between the permitted contractual regulation of *access* to works and recordings of protected performances and the impermissible contractual exclusions or restrictions of the two permitted uses in Division 8 of Part 5: see "indirect exclusions and restrictions" below.)

Machine-readable directives regarding use of website contents

130. Some website operators use a robots.txt file to deny access to web crawlers collecting training data for generative AI, such as OpenAI's GPTBot.[72]

131. A web crawler is an Internet bot that accesses, downloads and indexes Internet content. Web crawlers have traditionally been used by search engines like Google to find web pages to be returned in search results.

132. A robots.txt is a text file that implements the Robots Exclusion Protocol. The protocol is a set of rules regarding the use of robot that web crawlers are requested to honour when crawling a website. The text file is placed in the root of a website hierarchy and specifies whether crawling (by all crawlers or specific ones) is `allow`ed or `disallow`ed in respect of the entire website, specific directories or specific files.

133. OpenAI, Anthropic and others use web crawlers for the purpose of crawling content for use in training their respective foundational models. By indicating that crawling by OpenAI's GPTBot or Anthropic's ClaudeBot is disallowed, a website operator seeks to prevent those companies from using the contents of its website for AI training.

134. The limitation with this approach is that one must know the name of the relevant bot. As a result, organisations have sought to extend the protocol by including specific declarations regarding AI training. For example, Clouldfare introduced a non-standard extension to the protocol called Content Signals, which allows a website operator to declare whether the contents of the website can be crawled for AI training (by use of the syntax `ai-train=yes` or `ai-train=no`). A similar approach is taken by the Really Simple Licensing open content licensing standard.

135. But the protocol (in original or extended form) relies on voluntary compliance. A robots.txt file does not prevent web crawling. In the language of section 423 of the Act, a robots.txt file does

---

71 [2017] 5 SLR 1081; [2017] SGHC 239 at [25], quoting Stanley Yeo, Neil Morgan and Chan Wing Cheong, *Criminal Law in Malaysia and Singapore* (LexisNexis, 2nd Ed, 2012) at [1.39]–[1.40].

72 David Pierce, "The text file that runs the internet", *The Verge*, 14 February 2024.

not "effectively control access to" website contents and thus a robots.txt is not an "access control measure".

136. The question is then whether legal effect should be given to robots.txt files or other machine-readable directives so that non-compliance with such files or directives is unlawful.[73]

137. There are several arguments in favour of such an approach.

138. *First,* a rightsholder may wish their material to be exposed to the public at large but simultaneously wish to control which items of content can be scraped (i.e., copied). Access control measures (e.g., a paywall) generally hide the material from public view. The content can only be seen if access is provided. With advisory measures, the rightsholder can (if such measures are respected or given legal effect) achieve both the public exposure of its material as well as control over whether the material can be copied.

139. *Second*, managing large repositories of materials may be easier using advisory measures. A rightsholder can indicate whether crawling for AI training is `allow`ed or `disallow`ed for individual items, or collections, of materials.

140. *Third,* regulating use via such methods may be more efficient than regulating access through contract terms. Automated Internet bots cannot interpret contract terms but can read and respond to instructions regulating whether copying is `allow`ed or `disallow`ed.

141. However, there are several countervailing arguments.

142. *First*, the alleged efficiency gains that such measures have over contracts in respect of automated web scraping may be overstated. There are different levels of control that rightsholders may wish to exercise regarding use of their materials over and above a simple binary choice of `allow`ing or `disallow`ing copying (including for specific purposes like AI training). This is evident by public copyright licences like the Creative Commons Licence.

143. This reality is recognised by those who have created such advisory measures, which incorporate URLs that link to *human-readable* copyright licenses: see, for example, §5.2 of the TDM Reservation Protocol (which seeks to give effect to Article 4(3) of EU Directive 2019/790 that allows a rightsholder to opt out of the European Union's text and data mining regime using "machine-readable means") and the Really Simple Licensing open content licensing standard.

144. *Second*, there may be a risk that operators of websites simply `disallow` crawling for AI training by default. One then queries whether such an outcome may undermine the permitted use of copying for the purpose of computational data analysis.

145. By way of explanation, contract terms that exclude or restrict the permitted uses under Division 8 of Part 5 are void. The Second Minister for Law explained that this was because the benefit of the permitted use could be "drastically reduced or even [rendered] non-existent if we allow contracts to modify its application".[74]

146. A robots.txt file is unlikely to satisfy the requirements of a legally enforceable contract (as noted above, compliance is voluntary, not contractual). Thus, extensions which `disallow` crawling for AI training would be beyond the scope of section 187.

---

73 See, for example, EU Directive 2019/790, Article 4(3).

74 "Second Reading of the Copyright Bill", *Parliament of Singapore*, 13 September 2021 (Edwin Tong Chun Fai, Second Minister for Law).

147. However, if one were to incorporate such extensions as terms of a contract, then those terms would probably fall foul of the prohibition on contract terms excluding or restricting computational data analysis in section 187.

148. So, it would seem unusual for the legislature to decide to give effect to machine-readable directives regarding the copying of works and recordings of protected performances for AI training when if those directives were expressed as contract terms, those terms would offend section 187.

Additional technical safeguards

149. Generative AI models may facilitate copyright infringement by reproducing copyrighted materials, particularly where a model has been trained on copyrighted materials. This is evidenced by numerous examples of ongoing litigation around the world.[75]

150. As computational data analysis may be used to develop a generative AI model, the question arises as to whether additional safeguards of a technical nature should be implemented to ameliorate the risk that a model reproduces the material that was subject to the initial computational data analysis (namely during pre-training and fine-tuning).

151. Our observations on this question are as follows.

152. While beyond the scope of this report, one might reasonably argue that the appropriate way to deal with infringing outputs generated by generative AI models is by way of an infringement action. That is, there is no need to impose technical safeguards during the computational data analysis stage when a legal response is available, in the form of an infringement action, to deal with infringing outputs.

153. In attempting to ameliorate the "downstream" problem of infringing outputs by implementing "upstream" technical safeguards on computational data analysis, there is a risk that the legislature unintentionally curtails the permitted uses (and thus the policy aims of Division 8 of Part 5) or introduces unintended compliance-like costs on Singapore model developers (particularly smaller ones) that put those developers at a comparative disadvantage to their (bigger) international peers.

154. Further, there is evidence that suggests the industry is aware of concerns regarding the outputting of infringing outputs (and the associated threat of litigation) and is responding in various ways.[76]

155. Moreover, changes to technology vastly outpace the legislative drafter's pen. By the time safeguards are imposed by the legislature, particularly safeguards of a technical nature, both the technology and the industry are very likely to have already moved on.

156. For these reasons, we do not think the imposition of additional technical safeguards on computational data analysis is necessary or desirable.

    (We query whether Division 8 of Part 5 of the Act may, in fact, be used to negate the risk that a generative AI model produces outputs that reproduce existing copyrighted works.

---

[75] See, for example, *Andrea Bartz et al v Anthropic PBC.,* 3:24-cv-05417-WHA, (N.D. Cal. Jun 23, 2025) (pdf), *Kadrey v Meta Platforms, Inc.*, 3:23-cv-03417, (N.D. Cal. Jun 25, 2025) ECF No. 598 (pdf) and *GEMA vs OpenAI* (ref. 42 O 14139/24, 42nd Civil Chamber of the Munich I Regional Court, 11 November 2025).

[76] See, for example, the "mitigations" deployed by Meta in respect of its open-source large language model, LLaMA (*Kadrey v Meta Platforms, Inc.*, 3:23-cv-03417, (N.D. Cal. Jun 25, 2025) ECF No. 598 (pdf) at pp.12–13) and Adobe's indemnification of enterprise users of its Firefly software (Stephen Nellis, "Adobe pushes Firefly AI into big business, with financial cover", *Reuters*, 9 June 2024; and "Firefly Legal FAQs – Enterprise Customers", *Adobe*, 10 May 2024 (pdf)).

Techniques are available to prevent, or at least reduce the possibility that, models memorise and regurgitate copyrighted material upon which the models have been trained. Some techniques involve comparing a model's outputs against the material that was used in the model's training.

We observe that one of the two types of computational data analysis provided in section 243 of the Act is "using the work or recording as an example of a type of information or data to improve the functioning of a computer program in relation to that type of information or data". The illustration provided of this type of computational data analysis is "the use of images to train a computer program to recognise images."

In light of the above, it appears that it may be possible for the developer of a generative AI model, who employs computational data analysis in developing the model, to also employ such analysis to test and develop measures to prevent the model from memorising and regurgitating the material that was subject to the initial computational data analysis.)

Contractual exclusion or restriction of computational data analysis

157. Section 187 governs contract terms that purport to exclude or restrict computational data analysis.

> (1) Any contract term is void to the extent that it purports, directly or indirectly, to exclude or restrict any permitted use under any provision in —
>
> . . .
>
> (c) Division 8 (computational data analysis); …
>
> . . .
>
> (2) Without limiting subsection (1), a contract term is void to the extent that it purports, directly or indirectly, to prevent or restrict the doing of any of the following acts in circumstances that constitute a permitted use under the provisions mentioned in subsection (1):
>
> (a) making a copy of a work or a recording of a protected performance;
>
> (b) supplying (whether by communication or otherwise) a copy of a work or a recording of a performance;
>
> . . .

158. During the second reading of the Copyright Bill, the Second Minister for Law stated the following in answer to a question about section 187:[77]

[77] "Second Reading of the Copyright Bill", *Parliament of Singapore*, 13 September 2021 (Edwin Tong Chun Fai, Second Minister for Law).

> … many permitted uses can be modified or excluded via reasonable contract terms. However, some permitted uses are mandatory for public policy reasons because the benefit of that particular permitted use could be drastically reduced or even non-existent if we allow contracts to modify its application. So, we protect certain categories which are necessary for public policy reasons, and computational data analysis is one such category.

Policy rationale

159. The drafting of Division 8 of Part 5 of the Act suggests that the legislature recognised that there are two primary ways that rightsholders might seek to manage the acts, comprised in the copyright of their works or recordings of protected performances, that they authorise others to perform: by way of a contractual provision or by way of an access control measure.

160. An access control measure is (as the name suggests) a method of controlling *access* to such materials. As explained above (see "Access in contravention of law"), circumvention of an access control measure will mean that *X*'s access to materials is not lawful and thus *X's* copying is not a permitted use, the result of which is that *X* is exposed to liability for copyright infringement. (*X*'s circumvention of an access control measure is, of itself, also actionable.[78])

161. Contractual provisions are a method which can be used to regulate (a) *access* to material (particularly access to a service by which the material is made available for potential copying); and (b) the exercise of the exclusive rights in that material by others (that is, regulate *use* of the material). This distinction is recognised in the legislation.

162. In respect of terms governing *access* to material, the legislature has sought to make *access* to material in breach of contract unlawful (see "Access in breach of contract" above).

    (A rightsholder could make circumvention of an access control measure a breach of contract. This way, the rightsholder may have two rights of action against *X* arising out of the same conduct: (a) a right of action under section 436 of the Act for circumvention of an access control measure under section 425 of the Act; and (b) a right of action under the common law for breach of contract.

    Alternatively, the rightsholder could make circumvention of an access control measure an event which gives it a right to terminate the contract. This is an example of a contractual "self-help" remedy: the rightsholder can remedy *X's* breach of contract without needing to commence legal proceedings to obtain a court order against *X*.)

163. As to terms that govern *use* (including copying and communication to the public) of material, the legislature is concerned that rightsholders may oust or restrict the *use* of their material for computational data analysis and thereby render Division 8 of Part 5 nugatory. In anticipation of this potential outcome, the legislature introduced section 187 to render void contract terms that exclude or restrict the two permitted uses in Division 8 of Part 5.

    (Not all contractual exclusions or restrictions on *use* of works and recordings of protected performances are prohibited. Only those that have the effect of excluding or restricting the two permitted uses.)

164. As this is a matter of government policy, which has been passed into law by the Singapore legislature, this report does not reconsider the balance struck.

---

78 Copyright Act 2021, section 436 read with section 425.

Direct and indirect exclusions and restrictions

165. Section 187 renders a contract term void to the extent that it "directly or indirectly" excludes or restricts the uses permitted under Division 8 of Part 5.

    (Section 187 states that a contract term is void to the extent that it "*purports … to*" exclude or restrict a permitted use. We do not think that these words should be interpreted as being concerned with anything other than determining the substantive effect of a term—determined through a process of contractual construction—such that if a term does not, on its proper construction, exclude or restrict a permitted use, then that term is not rendered void by section 187.)

166. The dividing line between what is "direct" and "indirect" may not always be easy to draw.

Direct exclusions and restrictions

167. A direct *exclusion* would appear to be one that expressly excludes one or more of the permitted uses. For example, a contract term with words to the effect that "*Making copies of our content for the purpose of computational data analysis is not permitted under any circumstances*" amounts to a direct exclusion.

168. A direct *restriction* would appear to be one that expressly limits or conditions one or more of the permitted uses. For example, Division 8 of Part 5 of the Act places no express restriction on the purpose for which *X* undertakes computational data analysis. For example, there is no restriction on *X* performing computational data analysis for commercial purposes.[79]

169. So, a contract term with words to the effect that "*You may make copies of our content for the purpose of non-commercial computational data analysis only*" would appear to be a direct restriction of computational data analysis because, while it permits computational data analysis for a non-commercial purpose, it necessarily precludes such analysis for a commercial purpose.

Indirect exclusions and restrictions

170. An indirect exclusion or restriction would appear to be one that effectively excludes or restricts one or more of the permitted uses but without expressly saying so. For example, a contract term with words to the effect that "*You must obtain prior written permission before making any copies of our content*" may be an indirect restriction of a permitted use as it makes the permitted copying conditional on consent being provided. Division 8 of Part 5 does not require consent to be obtained from the rightsholder before computational data analysis can be performed.

171. The dividing line between a (permissible) contract term that regulates *access* to a work or recording of a protected performance and a (prohibited) contract term that *indirectly* excludes or restricts the permitted use of *copying* a work or a recording of a protected performance for computational data analysis may be difficult to draw, particularly with respect to works and recordings made available on the Internet. *Access* is, in such cases, upstream from downstream *use.* So, exclusions or restrictions on *access* may in many cases, if not necessarily, *indirectly* exclude or restrict the permitted *use* of copying for the purpose of computational data analysis.

    (This does not of course mean that one cannot regulate access to works and recordings of protected performances: one can, as noted above, employ an access control measure.)

---

79 *Cf*, for example, Copyright, Designs and Patents Act 1988 (United Kingdom), section 29A(1)(a).

Doctrine of severance

172. Division 1 of Part 5 of the Act does not address whether, in circumstances where *part* of a contract term contravenes section 187(1)(c) but the remainder of the term does not, *the part* of a contract term that contravenes section 187(1)(c) can be severed from the remainder.

173. While the provision uses the modifying phrase "to the extent", this may not necessarily mean that severance is permitted.[80]

Anti-avoidance

174. A rightsholder may attempt to avoid the operation of section 187 by including in its contract with *X* a governing law clause that selects the laws of a place other than Singapore to govern the contract. Section 188 responds to that scenario by rendering such clauses void. Our observations of section 188 are set out in **Annexure 4**: Anti-avoidance and consumer protection.

Division 8 of Part 5 and AI model development

175. There are several steps to developing an AI model, some of which are listed below:

    a. defining the problem or use case;

    b. data collection;

    c. data processing;

    d. pre-training;

    e. fine-tuning;

    f. alignment and safety training;

    g. evaluation and benchmarking; and

    h. model deployment.

176. Below we address the stages where Division 8 of Part 5 may be used to assist in the development of a generative AI model.

Data collection

177. Data collection involves gathering the raw materials that will "teach" the model the relevant domain. This could include web pages, books, scientific papers, computer codes and so on.

178. Division 8 of Part 5 helps with the data collection step as it permits that which would otherwise be an act of infringement: copying of works and recordings of protected performances.

179. In addition, Division 1 of Part 5 renders void any contract terms that purport to exclude or restrict copying for the purpose of computational data analysis.

---

80 See, for example, *Koh Lin Yee v Terrestrial Pte Ltd and another appeal* [2015] 2 SLR 497; [2015] SGCA 6 at [65] in relation to section 3(2) of the Unfair Contract Terms Act 1977 which uses an analogous modifier "except in so far as".

Data cleaning and pre-processing

180. Once collected, the raw data need to be cleaned and prepared for machine learning. This involves identifying and correcting or removing inaccurate, incomplete, improperly formatted, or duplicate data from a dataset.

181. After cleaning, the data will be generally pre-processed. In respect of images, this may involve a decoded image array (pixels as numbers) or a compressed or resized version of the original. In respect of text, the data may be "tokenised". Tokenisation converts text into words, sub-words and characters, collectively called "tokens". Tokens are the smallest units of meaning that can be processed by an LLM. This conversion allows LLMs to work with more manageable and understandable chunks of data. Tokenisation is generally performed using an algorithm.

182. Following tokenisation, a process called "vectorisation" may occur. Vectorisation is the process of creating numerical representations (called "embeddings") of the tokens. Vectorisation is also performed using an algorithm, called an "embedding model".

183. Illustrations of tokenisation and vectorisation are provided in **Annexure 5**: Practical demonstration of tokenisation and vectorisation**.**

184. Division 8 of Part 5 assists with the data cleaning and pre-processing stage as *X* is permitted to make copies of works and recordings of protected performances for the purpose of preparing those works and recordings for computational data analysis.[81] The processes of cleaning and pre-processing are all performed to prepare the works or recordings for subsequent "pre-training".

    (Some have queried whether the processes of tokenisation and vectorisation, which result in the creation of new data structures, result in the making of "adaptions" which are not expressly permitted by Division 8 of Part 5.[82] Making an adaptation is one of the exclusive rights of literary, dramatic and musical works.[83] We make the following observations regarding this issue.

    First, it appears this issue would only arise in relation to literary works, where the question is whether such data structures amount to a "translation" of a literary work.[84] In that regard, it is not clear that a collection of tokens or a vector thereof is a *language* into which a literary work has been *translated*.

    Second, as noted above, section 244(2)(a)(ii) permits *X* to make a copy of a work for the purpose of *preparing* the work for subsequent pre-training. If pre-training satisfies the definition of "computational data analysis", then tokenisation and vectorisation are steps that prepare the work for that analysis.

    Third, one might reasonably argue that an adaptation must itself satisfy the requirements of a work.[85] In that regard, for copyright to subsist in a work, the work must bear a sufficient causal

---

81 Copyright Act 2021, section 244(2)(a)(ii).

82 See, for example, Jeffrey Lim, "Did Singapore's Copyright Act 2021 Solve Copyright Problems in the Training of AI", *Joyce A Tan & Partners LLC,* 25 August 2023; and Jeffrey Lim, "Copyright and AI Model Training: The Balancing Act Seen From Singapore", *Singapore Business Review*, June 2025.

83 Copyright Act 2021, section 112(1)(e).

84 Copyright Act 2021, sections 17 and 18.

85 See, for example, *Computer Edge Pty Ltd v Apple Computer Inc* (1986) 161 CLR 171; [1986] HCA 19 in relation to Australia's Copyright Act 1968.

connection with the engagement of the human intellect.[86] The products of tokenisation and vectorisation are created by automated processes (see **Annexure 5**: Practical demonstration of tokenisation and vectorisation) rather than human ones. So, the products of tokenisation and vectorisation cannot be works.)

Pre-training

185. During the pre-training phase, a model's parameters are created and iteratively adjusted through large-scale exposure to data.

186. At the outset, the model's weights—numerical values defining the strength of connections in the network—are typically initialised randomly.[87]

187. The developer then copies the training dataset to high-performance storage and "shows" the data to the model in batches.[88]

188. As the model processes each batch, an optimisation algorithm measures its performance against a defined objective and updates the weights accordingly.[89]

189. Through repeated iterations, the model forms high-dimensional latent representations that encode patterns, similarities, and structural or stylistic features not visible in the raw form.[90]

190. Contemporary pre-training commonly uses *self-supervised learning*. Here, the model learns from unlabelled data by predicting masked or missing portions of the input—such as the next token or a concealed region of an image—using surrounding context. By repeatedly solving these tasks, the model adjusts its parameters to capture statistical relationships, syntactic and semantic regularities, and broader stylistic or compositional structures embedded in the corpus.

191. These processes appear to fall within the definition of "computational data analysis".

192. Pre-training generally uses versions of the original materials prepared under section 244(2)(a)(ii). As such, copyright infringement is unlikely to be engaged (see the parenthetical comments in the "Data cleaning and pre-processing" section above).

193. We observe that there are arguments that pre-training goes beyond mere "automated analytical techniques" involved in traditional text and data mining.[91] Consequentially, one may argue that the process of pre-training goes beyond the scope of section 243(a) of the Act.

194. Nevertheless, when a developer exposes model weights to training data to improve predictive performance—including the loading of the training examples or their processed representations into memory during batch processing, and the iterative updating of weights—the developer uses the training data as examples to improve the functioning of the model (a

---

86 *Global Yellow Pages Ltd v Promedia Directories Pte Ltd and another matter* [2017] 2 SLR 185; [2017] SGCA 28 at [24].

87 *Getty Images (US) Inc & Ors v Stability AI Ltd (Rev1)* [2025] EWHC 2863 (Ch) at [5].

88 *Copyright and Artificial Intelligence Part 3: Generative AI Training* (Register of Copyrights, US Copyright Office; Pre-Publication Version, May 2025) (pdf) pp. 27–30.

89 *Copyright and Artificial Intelligence Part 3: Generative AI Training* (Register of Copyrights, US Copyright Office; Pre-Publication Version, May 2025) (pdf) pp. 17–19.

90 Nicola Lucchi, *Generative AI and Copyright: Training, Creation and Regulation* (European Parliament, Policy Department for Justice, Civil Liberties and Institutional Affairs, Study PE 774.095, July 2025) (pdf) p.31 and fn. 69–71.

91 Nicola Lucchi, *Generative AI and Copyright: Training, Creation and Regulation* (European Parliament, Policy Department for Justice, Civil Liberties and Institutional Affairs, Study PE 774.095, July 2025) (pdf) pp.37–42.

computer program), thus falling within paragraph (b) of section 243 of the Act.

195. Even if the original material were used or, if one were to take the view (contrary to the parenthetical comments in the "Data cleaning and pre-processing" section above) that the pre-processed versions "encode" protected expressive elements of the original work (so that those elements may be subsequently recombined),[92] this is not an infringing use. As explained above ("Use of the expressive elements of a work or recording"), use of the expressive elements of a work or recording is permitted, and indeed required, to fulfil the expressed policy objective of "training an Artificial Intelligence programme".

    (To the extent pre-training involves the storage of original materials, or pre-processed versions thereof, such storage is permitted.[93])

Fine-tuning

196. To optimise the pre-trained model for specific tasks or domains, developers employ a technique called "fine-tuning". This generally involves the use of labelled examples to "teach" the model what outputs are expected for particular inputs.

197. See, for example, the following fine-tuning methods supported by the OpenAI platform:[94]

METHOD	HOW IT WORKS	BEST FOR	USE WITH

Supervised fine-tuning (SFT)	Provide examples of correct responses to prompts to guide the model's behavior.
	
	Often uses human-generated "ground truth" responses to show the model how it should respond.
		• Classification
		• Nuanced translation
		• Generating content in a specific format
		• Correcting instruction-following failures
			`gpt-4.1-2025-04-14`
			`gpt-4.1-mini-2025-04-14`
			`gpt-4.1-nano-2025-04-14`

Vision fine-tuning	Provide image inputs for supervised fine-tuning to improve the model's understanding of image inputs.	• Image classification • Correcting failures in instruction following for complex prompts	`gpt-4o-2024-08-06`

Direct preference optimization (DPO)	Provide both a correct and incorrect example response for a prompt. Indicate the correct response to help the model perform better.	• Summarizing text, focusing on the right things
		• Generating chat messages with the right tone and style
			`gpt-4.1-2025-04-14`
			`gpt-4.1-mini-2025-04-14`
			`gpt-4.1-nano-2025-04-14`

92 See, for example, Nicola Lucchi, *Generative AI and Copyright: Training, Creation and Regulation* (European Parliament, Policy Department for Justice, Civil Liberties and Institutional Affairs, Study PE 774.095, July 2025) (pdf) pp. 31, 42-45 and 52).

93 Copyright Act 2021, section 244(3).

94 OpenAI, "Model Optimization", *OpenAI Platform*, undated.

Reinforcement fine-tuning (RFT)	Generate a response for a prompt, provide an expert grade for the result, and reinforce the model's chain-of-thought for higher-scored responses. Requires expert graders to agree on the ideal output from the model. **Reasoning models only.**	• Complex domain-specific tasks that require advanced reasoning • Medical diagnoses based on history and diagnostic guidelines • Determining relevant passages from legal case law	`o4-mini-2025-04-16`

198. We observe that several of these methods involve providing the model with examples of content to improve its functioning. As such, these methods appear to fit within paragraph (b) of section 243. Indeed, we observe that "vision fine-tuning" appears to fit squarely within the illustration to paragraph (b) of section 243.

Alignment and safety training

199. Alignment and safety training involves "teaching" the model to follow instructions appropriately, refuse harmful requests, acknowledge limitations, and behave according to intended values and constraints.

200. Part of safety training may also involve ensuring that the fine-tuned model does not regurgitate its training data. As noted in the parenthetical comments to the "Additional technical safeguards" section above, we query whether Division 8 of Part 5 of the Act may be used in this regard.

Evaluation and benchmarking

201. Evaluation involves testing the fine-tuned model's performance against various metrics. Benchmarking is the process of measuring the model's performance on certain tasks against other existing models to provide a comparative assessment.

202. Division 8 of Part 5 assists in evaluation and benchmarking as *X* is permitted to supply copies of works and recordings of protected performances used in computational data analysis to another person for the purpose of "verifying" the results of that computational data analysis.[95]

Singapore as an AI development hub

203. Some have argued that Division 8 of Part 5 of the Act establishes Singapore as "a safe harbour for AI development, ensuring that companies can innovate without the fear of infringement claims if the developer complies with the conditions of Division 8 of Part 5".[96]

204. Given that it appears that Division 8 of Part 5 can extend to materials created by persons in reciprocating countries (as discussed above), there is evidence in support of that claim.

205. There are, however, two limitations.

---

95 Copyright Act 2021, section 244(4).

96 Adil Al-Busaidi *et al*, "Redefining boundaries in innovation and knowledge domains: Investigating the impact of generative artificial intelligence on copyright and intellectual property rights" (2024) 9(4) *Journal of Innovation & Knowledge* 100630.

206. *First*, it is unlikely that if a copyright infringement claim is brought against *X* in a foreign country that *X* could raise Division 8 of Part 5 of the Act as a defence to that claim. The Division is a foreign law from the perspective of that other country and to enforce it in that other country may be against the public policy of that other country.

207. Whether such a claim could be brought against *X* would, no doubt, turn on whether there is a sufficient territorial nexus between the dispute and the other country. For example, if no copying of the material occurs in the territory of the other country, there may not be a sufficient territorial nexus.

(For an illustration of this issue, see *Getty Images (US) Inc & Ors v Stability AI Ltd*[97] regarding "the Location Issue". We observe that during closing arguments of the subsequent hearing of the claim, the rightsholder, Getty Images, dropped its copyright infringement claim regarding the training of Stable Diffusion, acknowledging the challenges in proving that the alleged copyright infringement occurred within the United Kingdom.[98])

208. Given that this issue turns on an assessment of foreign copyright and private international law, it is beyond the scope of this report.

(For completeness, we observe that even where there is a nexus, the claim may be dismissed on *forum non-conveniens* grounds.[99])

209. *Second*, while Division 8 of Part 5 may establish Singapore as an AI *development* safe harbour, it does not establish Singapore as an AI *deployment* safe harbour. That is, if *X*'s generative AI model produces outputs that infringe copyright, whether that copyright is established under Singapore law or foreign law, *X* may be exposed to potential litigation.

---

97 [2023] EWHC 3090 at [43]–[62].

98 See "Getty Images v Stability AI: main copyright claims dropped", *Pinsent Masons*, 25 June 2025; Antony Craggs, "Getty drops primary copyright claim against stability AI", *Shoosmiths*, 3 July 2025 and, the substantive hearing of the proceeding: *Getty Images (US) Inc & Ors v Stability AI Ltd (Rev1)* [2025] EWHC 2863 (Ch) at [9].

99 See, for example, the majority decision in *Creative Tech, Ltd v Aztech Sys. Pte Ltd* 61 F.3d 696 (9th Cir. 1995), but noting the criticism of this approach provided in Jane Ginsburg, "Extraterritoriality and Multiterritoriality in Copyright Infringement" (1997) 37 *Virginia Journal of International Law* 587 at pp. 592–595 (pdf).

References

Statutes

- Computer Misuse Act 1993 (Singapore).
- Copyright Act 1987 (Singapore).
- Copyright Act 2021 (Singapore).
- Copyright (Amendment) Act 1998 (Singapore).
- Copyright, Designs and Patents Act 1988 (United Kingdom).
- Copyright Regulations 2021 (Singapore).
- EU Directive 2019/790 (European Union).
- Interpretation Act 1965 (Singapore).
- Unfair Contract Terms Act 1977 (Singapore).

Judgments

- *Amin Rasheed Shipping Corp v Kuwait Insurance Co* [1983] 2 All ER 884; [1984] 1 AC 50; [1983] 3 WLR 241.
- *Andrea Bartz et al v Anthropic PBC.,* 3:24-cv-05417-WHA, (N.D. Cal. Jun 23, 2025) (pdf).
- *Asia Pacific Publishing Pte Ltd v Pioneers & Leaders (Publishers) Pte Ltd* [2011] 4 SLR 381; [2011] SGCA 37.
- *Attorney-General v Ting Choon Meng and another appeal* [2017] 1 SLR 373; [2017] SGCA 6.
- *Computer Edge Pty Ltd v Apple Computer Inc* (1986) 161 CLR 171; [1986] HCA 19.
- *Creative Tech, Ltd v Aztech Sys. Pte Ltd* 61 F.3d 696 (9th Cir. 1995).
- *Eram Shipping Company Ltd & Ors v Hong Kong and Shanghai Banking Corporation Ltd* [2003] UKHL 30.
- *GEMA vs OpenAI* (ref. 42 O 14139/24).
- *Getty Images (US) Inc & Ors v Stability AI Ltd (Rev1)* [2025] EWHC 2863 (Ch).
- *Global Yellow Pages v Promedia Directories Pte Ltd* [2016] 2 SLR 165; [2016] SGHC 9.
- *Global Yellow Pages Ltd v Promedia Directories Pte Ltd and another matter* [2017] 2 SLR 185; [2017] SGCA 28.
- *Kadrey v Meta Platforms, Inc.*, 3:23-cv-03417, (N.D. Cal. Jun 25, 2025) ECF No. 598 (pdf).
- *Koh Lin Yee v Terrestrial Pte Ltd and another appeal* [2015] 2 SLR 497; [2015] SGCA 6.
- *Ng Kok Wai v Public Prosecutor* [2024] 3 SLR 1516; [2023] SGHC 306.
- *Public Prosecutor v Taw Cheng Kong* [1998] 2 SLR(R) 489; [1998] SGCA 37.
- *RecordTV Pte Ltd v MediaCorp TV Singapore Pte Ltd and others* [2011] 1 SLR 830; [2010] SGCA 43.
- *Shaikh Farid v Public Prosecutor and other appeals* [2017] 5 SLR 1081; [2017] SGHC 239.
- *Tan Cheng Bock v Attorney-General* [2017] 2 SLR 850; [2017] SGCA 50.

- *Tan Chor Jin v Public Prosecutor* [2008] 4 SLR(R) 306; [2008] SGCA 32.
- *Tan Ng Kuang and another v Jai Swarup Pathak* [2022] 3 SLR 788; [2021] SGHC 232.
- *The Republic of the Philippines v Maler Foundation and others and other appeals* [2014] 1 SLR 1389; [2013] SGCA 66.

Secondary sources

- "About CC Licenses", *Creative Commons*, undated.
- *Copyright and Artificial Intelligence Part 3: Generative AI Training* (Register of Copyrights, US Copyright Office; Pre-Publication Version, May 2025) (pdf).
- "Firefly Legal FAQs – Enterprise Customers", *Adobe*, 10 May 2024 (pdf).
- "Getty Images v Stability AI: main copyright claims dropped", *Pinsent Masons*, 25 June 2025.
- "Model Optimization", *OpenAI Platform*, undated.
- "Second Reading of the Copyright Bill", *Parliament of Singapore*, 13 September 2021.
- Adil Al-Busaidi *et al*, "Redefining boundaries in innovation and knowledge domains: Investigating the impact of generative artificial intelligence on copyright and intellectual property rights" (2024) 9(4) *Journal of Innovation & Knowledge* 100630.
- Antony Craggs, "Getty drops primary copyright claim against stability AI", *Shoosmiths*, 3 July 2025.
- Jane Ginsburg, "Extraterritoriality and Multiterritoriality in Copyright Infringement" (1997) 37 *Virginia Journal of International Law* 587 at pp. 592–595 (pdf).
- Jeffrey Lim, "Copyright and AI Model Training: The Balancing Act Seen From Singapore", *Singapore Business Review*, June 2025.
- Jeffrey Lim, "Did Singapore's Copyright Act 2021 Solve Copyright Problems in the Training of AI", *Joyce A Tan & Partners LLC,* 25 August 2023.
- Nicola Lucchi, *Generative AI and Copyright: Training, Creation and Regulation* (European Parliament, Policy Department for Justice, Civil Liberties and Institutional Affairs, Study PE 774.095, July 2025) (pdf).
- Stephen Nellis, "Adobe pushes Firefly AI into big business, with financial cover", *Reuters*, 9 June 2024.
- David Pierce, "The text file that runs the internet", *The Verge*, 14 February 2024.
- Yeong Zee Kin, *Technology Regulation in the Digital Economy* (Singapore Academy of Law, 2023).

Models

- all-MiniLM-L6-v2 (available from HuggingFace).

Annexure 1: Permitted uses under Division 8 of Part 5

Copying

Works

1. Making a copy is one of the exclusive rights comprised in the copyright of a work.[100]

2. If a person makes (or authorises the making of) a copy of a work in Singapore and that person neither owns the copyright nor has the licence of the copyright owner, that person infringes the copyright in the work.[101]

3. However, making a copy of a work for the purpose of computational data analysis (or preparing a work for such analysis) is a permitted use.[102]

   (Explanations of what comprises a "copy" in respect of each type of work is set out in Subdivision (2) ("Copying") of Division 3 ("Acts relating to works and performances") of Part 2 ("Interpretation") of the Act.)

Recordings of protected performances

4. A performer of a "protected performance" may bring an action in the General Division of the High Court of Singapore against any person who makes an "infringing use" of the performance.[103]

5. It is an "infringing use" if a person makes a "copy" of a "recording" of a protected performance.[104]

6. A "recording", in relation to a protected performance, is a sound recording of the performance or a substantial part of the performance and includes a copy of such a recording.[105]

7. A "copy" of a recording of a protected performance is a record (a) embodying a recording of the performance or a substantial part of the performance; and (b) derived, directly or indirectly, from a record produced upon the making of a recording of the performance or a substantial part of the performance.[106] Making a copy of a recording of a protected performance that is temporary or is incidental to some other use of the recording is to be treated as making a copy of the recording.[107]

8. However, making a copy of a recording of a protected performance for the purpose of computational data analysis (or preparing a work for such analysis) is a permitted use.[108]

---

100 Copyright Act 2021, section 112(1)(a) (literary, dramatic and musical works); section 113(a) (artistic works); section 118 (published editions of an authorial work); section 121(a)(ii) (sound recordings); section 124(a) (films); section 127(a) (broadcasts) and section 131(1)(a) (cable programmes).

101 Copyright Act 2021, section 146(1).

102 Copyright Act 2021, section 244(1)(a), read with section 244(2)(a).

103 Copyright Act 2021, section 177 read with section 7(1), definition of "Court".

104 Copyright Act 2021, section 175(a)(ii) and (b).

105 Copyright Act 2021, section 38(1).

106 Copyright Act 2021, section 51.

107 Copyright Act 2021, section 52.

108 Copyright Act 2021, section 244(1)(b), read with section 244(2)(a).

Communication to the public

9. "Communicate" in relation to a work or a protected performance means to transmit the work or performance by electronic means.[109] "Communicate" expressly includes making the work or performance available (on a network or otherwise) in a way that it may be accessed by any person on demand.[110] "Communication" has a corresponding meaning.[111]

Works

10. Communication to the public is one of the exclusive rights comprised in the copyright of a work.[112]

11. If a person communicates (or authorises the communication of) a work to the public in Singapore and that person neither owns the copyright nor has the licence of the copyright owner, that person infringes the copyright in the work.[113]

12. However, it is a permitted use for *X* to communicate a work to the public if:

    a. the communication is made using a copy of the work that complies with the conditions of section 244(2) of the Act;[114] and

    b. the communication is made for the purpose of (i) verifying the results of computational data analysis performed on the work carried out by *X*; or (ii) collaborative research or study relating to the purpose of the computational data analysis performed on the work carried out by *X*.[115]

Recordings of protected performances

13. It is an "infringing use" if a person makes a recording of a protected performance available to the public (on a network or otherwise) in a way that the recording may be accessed by any person on demand.[116]

14. However, it is a permitted use for *X* to communicate a recording of a protected performance to the public if:

    a. the communication is made using a copy of the performance that complies with the conditions of section 244(2) of the Act;[117] and

    b. the communication is made for the purpose of (i) verifying the results of computational data analysis performed on the performance carried out by *X*; or (ii) collaborative research or study relating to the purpose of the computational data analysis performed on the performance carried out by *X*.[118]

---

109 Copyright Act 2021, section 61(1).

110 Copyright Act 2021, section 61(1)(c).

111 Copyright Act 2021, section 61(3).

112 Copyright Act 2021, section 112(1)(d) (literary, dramatic and musical works); section 113(c) (artistic works); section 121(a)(iv) (sound recordings); section 124(d) (films); section 127(b) (broadcasts) and section 131(1)(b) (cable programmes).

113 Copyright Act 2021, section 146(1).

114 Copyright Act 2021, section 244(4)(a), read with section 244(1).

115 Copyright Act 2021, section 244(4)(b).

116 Copyright Act 2021, section 175(a)(iv) and (b).

117 Copyright Act 2021, section 244(4)(a), read with section 244(1)).

118 Copyright Act 2021, section 244(4)(b).

Annexure 2: Connecting factors between Singapore and a work or a recording of a protected performance

Works

Authorial works

1. In addition to being "original", copyright subsists in an *unpublished* authorial work if the author is a "qualified individual" when the work is made; if the work is made over a period, the author is a "qualified individual" for a substantial part of that period; or the work is made by or under the direction or control of the *Government of Singapore* or on or after 10 April 1987 by or under the direction or control of a "prescribed international organisation".[119]

2. Where an authorial work is published and copyright subsists in the work immediately before its first publication, copyright continues to subsist in the work if the author of the work is a "qualified individual" when the work is first published or dies before the work is first published but is a qualified individual immediately before his or her death.[120] Alternatively, copyright continues to subsist in the work if the work is first published in Singapore, by or under the direction or control of the *Government of Singapore*, or on or after 10 April 1987 by or under the direction or control of a "prescribed international organisation".[121]

3. Where an authorial work is published and there is no copyright in the work immediately before its first publication, copyright subsists in the work if the work is "original" and either of the following criteria is satisfied:[122]

    a. the author of the work is a "qualified individual" when the work is first published or dies before the work is first published but is a qualified individual immediately before his or her death; or

    b. the work is first published in Singapore, by or under the direction or control of the *Government of Singapore*, or on or after 10 April 1987 by or under the direction or control of a "prescribed international organisation".

Published editions of authorial works

4. Copyright subsists in a published edition of an authorial work if the person who first published the edition is a "qualified person" at the date when the edition is first published.[123] Alternatively, copyright subsists in a published edition of an authorial work if the edition is first published *in Singapore*, by or under the direction or control of the *Government of Singapore*, or by or under the direction or control of a "prescribed international organisation".[124]

119 Copyright Act 2021, section 109(1).

120 Copyright Act 2021, section 110(1)(a)(i).

121 Copyright Act 2021, section 110(1)(a)(ii).

122 Copyright Act 2021, section 110(2).

123 Copyright Act 2021, section 117(1)(a).

124 Copyright Act 2021, section 117(1)(b).

Sound recordings

5. Copyright subsists in a sound recording if the maker is a "qualified person" when the recording is made.[125] Alternatively, copyright subsists if the recording is made or first published *in Singapore*, by or under the direction or control of the *Government of Singapore*, or on or after 10 April 1987 by or under the direction or control of a "prescribed international organisation".[126]

Films

6. Copyright subsists in a film if the maker of the film is a "qualified person" for the whole or a substantial part of the period during which the film is made.[127]

7. Alternatively, copyright subsists if the film is made or first published *in Singapore*, by or under the direction or control of the *Government of Singapore*, or on or after 10 April 1987 by or under the direction or control of a "prescribed international organisation".[128]

Broadcasts

8. Copyright subsists in a broadcast if the broadcast is made from a place *in Singapore* by the holder of a broadcasting licence.[129]

Cable programmes

9. Copyright subsists in a cable programme if the programme is included in a cable programme service that is provided by a "qualified person" *in Singapore*.[130]

Recordings of protected performances

10. As noted above, a "recording of a protected performance" is a sound recording of a "protected performance" (or a substantial part thereof) and includes a copy of such a recording.[131] In turn, a "protected performance" is a "qualifying performance" (as defined by section 37(1) of the Act) that was either given live *in Singapore* or given live by a *Singapore citizen* or an individual resident or residing *in Singapore*.[132]

---

125 Copyright Act 2021, section 120(a).

126 Copyright Act 2021, section 120(b).

127 Copyright Act 2021, section 123(1)(a).

128 Copyright Act 2021, section 123(1)(b).

129 Copyright Act 2021, section 126.

130 Copyright Act 2021, section 130(1).

131 Copyright Act 2021, section 38(1).

132 Copyright Act 2021, section 173 read with sections 77 ("Who is a qualified individual") and 79 ("Who is a Singapore resident").

Annexure 3: Statutory provisions prohibiting unlawful access

Circumvention of an access control measure

1. It is an offence if a person circumvents an "access control measure".[133]

2. An "access control measure" is "any technology, device or component that, in the normal course of its operation, effectively controls access to" a work (or a copy thereof) or a recording of a protected performance.[134]

3. The elements of the offence are as follows:

    a. the "rights owner" (or someone with the rights owner's authority) applied the access control measure to the work (or a copy thereof) or a recording of a protected performance in connection with the exercise of the copyright in the work (or a copy thereof) or any right in the performance;[135]

    b. the person "wilfully" circumvents the access control measure;[136]

    c. the person does so to obtain a commercial advantage or private financial gain;[137] and

    d. the circumvention is done without the rights owner's authority.[138]

    ("Rights owner" is defined in section 96 of the Act.)

4. There are several exceptions to the offence.[139] But Division 8 of Part 5 of the Act is not one of them.

5. As such, circumvention of an access control measure in the circumstances set out above is an offence and thus will be unlawful for the purpose of section 244(2)(d) of the Act.

    (We observe that a rights owner of a work (or a copy thereof) or a recording of a protected performance may bring an action in the General Division of the High Court of Singapore against any person who infringes the prohibition on circumventing access control measures.[140] The elements of that action are different from the criminal offence set out above.[141] Namely, the rights owner does not need to show that the circumventor "willfully" circumvented the access control measure. Rather, the circumventor must have performed an act that the circumventor knew or ought reasonably to have known circumvents the access control measure. In addition, there is no requirement that the circumventor performed the act so as to obtain a commercial advantage or private financial gain.

---

133 Copyright Act 2021, section 439.

134 Copyright Act 2021, section 423, definition of "access control measure" read with section 422, definition of "protected copy".

135 Copyright Act 2021, section 439(1)(c) read with section 425(2)(a).

136 Copyright Act 2021, section 439(1)(a)(i).

137 Copyright Act 2021, section 439(1)(b).

138 Copyright Act 2021, section 439(1)(c) read with 425(2)(c).

139 see Copyright Act 2021, sections 428–435.

140 Copyright Act 2021, section 436 read with section 425.

141 Copyright Act 2021, section 425.

It is not clear, however, whether such infringing activity is unlawful for the purpose of section 244(2)(d) of the Act. That is, there is an argument that all section 436 does is to establish a private cause of action (and associated remedies in section 438) which a rights owner can choose to avail itself of, but the infringing act is not unlawful unless the elements of the criminal offence are established.)

Unauthorised access to computer material

6. It is an offence if a person knowingly causes a computer to perform any function for the purpose of securing access, without authority, to any program or data held in any computer.[142]

7. In this context, the following definitions apply:

    a. "data" means "representations of information or of concepts that are being prepared or have been prepared in a form suitable for use in a computer";[143]

    b. "function" includes "logic, control, arithmetic, deletion, storage and retrieval and communication or telecommunication to, from or within a computer";[144]

    c. a person "secures access" to any program or data held in a computer if by causing a computer to perform any function the person "copies or moves it to any storage medium other than that in which it is held …";[145]

    d. access of any kind by any person to any program or data held in a computer is unauthorised or done "without authority" if the person:

        i. is not himself or herself entitled to control access of the kind in question to the program or data;[146] and

        ii. does not have consent to access by him or her of the kind in question to the program or data from any person who is so entitled.[147]

8. If a work or recording of a protected performance satisfies the definition of "data", then circumvention of a paywall may be an offence under section 3(1) of the Computer Misuse Act 1993 and thus would appear to be unlawful for the purpose of section 244(2)(d) of the Act.

142 Computer Misuse Act 1993, section 3(1).

143 Computer Misuse Act 1993, section 2(1), definition of "data".

144 Computer Misuse Act 1993, section 2(1), definition of "function".

145 Computer Misuse Act 1993, section 2(2)(b).

146 Computer Misuse Act 1993, section 2(5)(a).

147 Computer Misuse Act 1993, section 2(5)(b).

Annexure 4: Anti-avoidance and consumer protection

1. Section 188 of the Act regulates a contract term that purports to apply the law of a country other than Singapore. Section 188 is in the following terms:

> (1) A contract term that purports to apply the law of a country other than Singapore is void if —
>
> (a) the application of that law has the effect of excluding or restricting the operation of any permitted use; and
>
> (b) either —
>
> (i) the term is imposed wholly or mainly for the purpose of evading the operation of any permitted use; or
>
> (ii) in the making of the contract one of the parties dealt as consumer, and he or she was then a Singapore resident, and the essential steps for the making of the contract were taken in Singapore (whether by him or her or by others on his or her behalf).
>
> (2) For the purposes of subsection (1)(b) —
>
> (a) the interpretation of section 27(2)(b) of the Unfair Contract Terms Act 1977 must be considered; and
>
> (b) if a person claims that a party does not deal as a consumer, the burden is on the person to prove this.
>
> (3) This section applies to any contract made before, on or after 21 November 2021.

2. An express term of a written contract that specifies that the laws of a particular legal system apply to any disagreement between the parties about their legal relationship is called a "governing" or "choice of" law clause.

3. Section 188(1) applies to a governing law clause that selects the law of a country other than Singapore. There are two requirements that must be satisfied for the section to apply.

4. *First*, the application of the selected governing law must have the effect of excluding or restricting the operation of any permitted use.

5. *Second,* one of two alternative conditions is satisfied:

   a. the clause is imposed wholly or mainly for the purpose of evading or restricting the operation of any permitted use; or

   b. in the making of the contract one of the parties dealt as consumer, and he or she was then a Singapore resident, and the essential steps for the making of the contract were taken in Singapore (whether by him or her or by others on his or her behalf).

6. Our observations on section 188 are as follows.

7. *First*, we query whether the requirement that "the application of [the foreign] law has the effect of excluding or restricting the operation of any permitted use" may mean that evidence must be led about the foreign law selected in the governing law clause. We observe that the approach taken by the drafters of section 27(2) of the Unfair Contract Terms Act 1977 (quoted below), upon which section 188 appears to be modelled, does not have a similar requirement and thus would appear to avoid this potential outcome.

8. *Second*, the first alternative condition of section 188(1)(b) appears to be directed to anti-avoidance of any permitted use under the Act, while the second alternative condition appears to be directed to consumer protection—to ensure that Singapore resident consumers can engage in any permitted uses under the Act.

9. *Third*, subsection 2(a) of section 188 states that, for the purpose of subsection 1(b) of section 188, "the interpretation of section 27(2)(b) of the Unfair Contract Terms Act 1977 must be considered". This appears to be because section 27(2)(b) of the Unfair Contract Terms Act 1977 (quoted below) is, for all intents and purpose, identical to section 188(1)(b)(ii).

   (Although we observe that section 27(2)(b) of the Unfair Contract Terms Act 1977 uses the words "habitually resident in Singapore", whereas section 188(1)(b)(ii) of the Act uses the phrase "Singapore resident" which, in turn, is defined in section 79 of the Act).

10. While not expressed directly, the direction in section 188(2)(a) of the Act would appear to mean that the phrase "dealt as consumer" used in section 188(1)(b)(ii) of that Act adopts the definition of that term used in section 12 of the Unfair Contract Terms Act 1977.

11. *Fourth*, section 188(1)(b)(i) of the Act states that the governing law clause "*is* imposed … for the purpose of evading the operation of any permitted use" (emphases added) while the equivalent requirement in section 27(2)(a) of the Unfair Contract Terms Act 1977 states that the governing law clause "*appears* to … have been imposed … for the purpose of enabling the party imposing it to evade the operation of this Act" (emphasis added). We query whether the difference in phraseology here indicates that the Act adopts a higher evidentiary threshold than the Unfair Contract Terms Act 1977.

12. *Fifth*, if a governing law clause falls within section 188(1) of the Act, then the governing law clause "is void".[148] We observe that this effect is different from the effect of section 27(2) of the Unfair Contract Terms Act 1977. That is, while both sections appear to be directed to similar concerns —the use of governing law clauses to oust the application of mandatory public law—the approach taken by the drafters of each section is different.

13. Section 27(2) of the Unfair Contract Terms Act 1977 is in the following terms (emphasis added):

148 Copyright Act 2021, section 188(1), chapeau.

> (2) *This Act has effect notwithstanding* any contract term which applies or purports to apply the law of some country outside Singapore, where (either or both) —
>
> (a) the term appears to the court, or arbitrator or arbiter to have been imposed wholly or mainly for the purpose of enabling the party imposing it to evade the operation of this Act; or
>
> (b) in the making of the contract one of the parties dealt as consumer, and he was then habitually resident in Singapore, and the essential steps necessary for the making of the contract were taken there, whether by him or by others on his behalf.

14. Unlike section 188(1) of the Act, section 27(2) of the Unfair Contract Terms Act 1977 does not render an offending governing law clause void. Section 27(2) simply states that the Unfair Contract Terms Act 1977, with its regulation of unfair contract terms, continues to have effect in the face of a governing law clause that selects the law of another place provided that one of the conditions in paragraph (a) or (b) of section 27(2) of the Unfair Contract Terms Act 1977 is satisfied: "The Act has effect notwithstanding …".

15. If a governing law clause is void, then the court will have to apply the principles of private international law to determine what the proper law of the contract is. This is because a contract cannot exist in a legal vacuum. As Lord Diplock of the House of Lords of the United Kingdom explained in *Amin Rasheed Shipping Corp v Kuwait Insurance Co*:[149]

> My Lords, contracts are incapable of existing in a legal vacuum. They are mere pieces of paper devoid of all legal effect unless they were made by reference to some system of private law which defines the obligations assumed by the parties to the contract by their use of particular forms of words and prescribes the remedies enforceable in a court of justice for failure to perform any of those obligations … .

16. A dispute about the proper law of a contract is obviously undesirable as it results in the expenditure of time and cost by the parties to the contract, as well as the court in resolving that dispute. The approach taken by the drafters of section 27(2) of the Unfair Contract Terms Act 1977 avoids this outcome.

17. The application of the principles of private international law in determining the proper law of the contract may not necessarily lead to the conclusion that Singapore law (including its mandatory laws) is the proper law of the contract.

18. However, the conditions which give rise to a governing law clause being rendered void under section 188(1) of the Act may inevitably lead to a conclusion that Singapore law has the closest and most real connection with the contract and the parties thereto and thus the proper law of the contract is Singapore law.

19. To repeat, a governing law clause only falls within section 188(1) of the Act if:

    a. the clause is imposed wholly or mainly for the purpose of evading the operation of any permitted use; or

---

149 [1983] 2 All ER 884; [1984] 1 AC 50; [1983] 3 WLR 241 at 249-250.

b. in the making of the contract one of the parties dealt as consumer, and he or she was then a Singapore resident, and the essential steps for the making of the contract were taken in Singapore (whether by him or her or by others on his or her behalf).

20. In respect of the first condition, a contract party presumably only imposes a governing law clause selecting a foreign legal system for the purpose of evading the operation of one or more permitted uses under the Act because, absent the clause, Singapore law would apply.

21. As to the second condition, there are various connections to Singapore. One of the parties must have been a *Singapore* resident when making the contract and the essential steps for the making of the contract were taken in *Singapore* by that person (or others on his or her behalf).

Annexure 5: Practical demonstration of tokenisation and vectorisation

Tokenisation

1. Tokenisation converts text into words, sub-words and characters, collectively called “tokens”. Tokens are the smallest units of meaning that can be processed by an LLM. This conversion allows LLMs to work with more manageable and understandable chunks of data.
2. By way of illustration, let’s take the following quote from American jurist Learned Hand:

> Justice, I think, is the tolerable accommodation of the conflicting interests of society,
> and I don't believe there is any royal road to attain such accommodations concretely.

3. We can tokenise this quote using WordPiece tokenisation (an algorithm developed by Google), which:
   a. splits text into sub-word units;
   b. handles out-of-vocabulary words by breaking them into smaller pieces;
   c. adds special tokens like `[CLS]` (classification) at the start and `[SEP]` (separator) at the end; and
   d. identifies sub-words by adding a `##` prefix.
4. Each token gets converted to a numerical ID from the model’s vocabulary. The result, comprising 38 tokens, looks as follows (represented in JSON for clarity):

```
{
   “[CLS]”: 101,
   “justice”: 3425,
   “,”:1010,
   "i": 1045,
   "think": 2228,
   ",": 1010,
   "is": 2003,
   "the": 1996,
   "to": 2000,
   "##ler": 3917,
   "##able": 3085,
   "accommodation": 11366,
   "of": 1997,
   "the": 1996,
   "conflicting": 19326,
   "interests": 5426,
   "of": 1997,
```

```
    "society": 2554,
    ",": 1010,
    "and": 1998,
    "i": 1045,
    "don": 2123,
    "’": 1521,
    "t": 1056,
    "believe": 2903,
    "there": 2045,
    "is": 2003,
    "any": 2151,
    "royal": 2548,
    "road": 2346,
    "to": 2000,
    "attain": 18759,
    "such": 2107,
    "accommodations": 26167,
    "concrete": 5509
    "##ly": 2135,
    ".": 1012,
    "[SEP]": 102
}
```

5. We can see how the words “concretely” and “tolerable” are broken into smaller sub-words. This shows how WordPiece tokenisation breaks out-of-vocabulary words into familiar pieces that capture meaning.
6. There are different tokenisation algorithms, and each will lead to different results.

Vectorisation

7. Vectorisation is the process of creating numerical representations of the tokens that capture semantic meaning. That is, similar meanings result in similar numerical representations.
8. To illustrate the process of vectorisation, we can use the all-MiniLM-L6-v2 sentence transformers model.
9. In this example, each of the 38 tokens above is represented in 384-dimensional vector space. The model processes all 38 token vectors simultaneously through its transformer layers. This means the final vector for “justice” gets influenced by seeing “society”, “conflicting interests”, “accommodation”, etc. in the same sentence. This is what helps capture the semantic meaning of the sentence.
10. Then the token vectors are averaged into a single 384-dimensional sentence representation (through a process called mean pooling).

11. The final output (represented as a collection of floating-point numbers) is as follows:

```
[0.07458861917257309, 0.0006146560190245509, -0.0623638853430748, -0.06244099140167236,
-0.004005978349596262, 0.051988475024700165, -0.004588609095662832, 0.0042947023175656796,
-0.05273613706231117, 0.05855022743344307, -0.01967121660709381, 0.011409603990614414,
0.01211056299507618, -0.0022557047195732594, 0.005550748202949762, -0.024863259866833687,
0.02932574786245823, 0.02042216993868351, -0.017853228375315666, 0.11067699640989304,
0.013865277171134949, 0.0032497684005647898, 0.027873987331986427, -0.01908906176686287,
-0.10732760280370712, -0.024462811648845673, 0.027442820370197296, -0.030126696452498436,
-0.019660454243421555, -0.02618454582989216, 0.02790532261133194, -0.018403606489300728,
0.013834626413881779, 0.07506126165390015, 0.048941340297460556, 0.017126448452472687,
0.04622885212302208, -0.024395188316702843, 0.02614935114979744, -0.09184816479682922,
0.04603464901447296, 0.04002352058887482, 0.010621483437716961, -0.06980514526367188,
0.07177559286355972, -0.019948715344071388, 0.0013339087599888444, -0.07745711505413055,
-0.05582940950989723, -0.11384270340204239, 0.026502283290028572, 0.0512222945690155,
-0.0019516414031386375, 0.0051658437587320805, -0.07797197252511978, 0.025840122252702713,
-0.050092101097106934, -0.014738618396222591, 0.04084969684481621, -0.049629345536231995,
0.04713389649987221, 0.05306895449757576, 0.046664897352457047, 0.0369376540184021,
-0.004832553677260876, -0.06483535468578339, 0.025464775040745735, 0.036428168416023254,
0.013131577521562576, -0.011332261376082897, -0.0365191213786602, 0.06379121541976929,
0.059588972479104996, -0.014071146957576275, -0.10133934020996094, 0.0006736447685398161,
0.03490763157606125, -0.05642223358154297, 0.04293235391378403, 0.037702448666095734,
0.020057322457432747, 0.024795732707452774, -0.0429159551858902, 0.02266268990933895,
-0.007341562770307064, -0.11654847115278244, -0.042339954525232315, -0.07171238213777542,
0.020531564950942993, -0.0008371563162654638, 0.03482414782047272, -0.006039132364094257,
0.09137049317359924, -0.0026237349957227707, -0.014351235702633858, 0.0159728042781353,
0.015543978661298752, 0.013685005716979504, -0.0005773277371190488, 0.04155626520514488,
-0.05582648143172264, 0.0287795327603817, 0.007339983247220516, 0.03322676569223404,
-0.009720380418002605, -0.022812407463788986, -0.05575656518340111, 0.004470643121749163,
-0.04194048419594765, -0.10623884201049805, -0.03845750913023949, -0.05990926921367645,
0.043270744383335114, -0.02943069115281105, 0.058609381318092346, 0.08658333867788315,
0.024556871503591537, 0.017050432041287422, 0.06823268532752991, -0.030839838087558746,
0.049751266837120056, 0.04854470491409302, -0.058896176517009735, 0.0353364460170269,
0.012661945074796677, -0.02375902235507965, -0.019438527524471283, -3.0312849350188576e-33,
-0.07705353200435638, -0.013456915505230427, -0.03544251248240471, -0.06593599915504456,
0.004024474881589413, -0.0215990599244833, -0.08042627573013306, 0.018202099949121475,
0.042572274804115295, 0.03790391981601715, 0.09714262187480927, -0.07736253734840332,
0.031834136694669724, -0.10878819227218628, -0.045284915715456 01, 0.07234477251768112,
-0.07044664025306702, 0.0411069318652153, -0.02124498412013054, 0.03560535982251167,
-0.00310097960755229, 0.024753190577030182, 0.04104592278599739, 0.08659981191158295,
-0.09509127587080002, -0.05259421840310097, 0.08184072375297546, -0.051204781979322433,
0.07454293221235275, 0.019924286752939224, -0.0002902782871387899, 0.07328936457633972,
0.012244547717273235, 0.025570789352059364, 0.052642855793237686, 0.06886183470487595,
-0.006775807589292526, -0.009569620713591576, 0.033908262848854065, -0.05491733178496361,
-0.07811368256807327, -0.0247524194419384, 0.020443173125386238, 0.01654699817299843,
0.029280735179781914, 0.023420874029397964, -0.02904752641916275, -0.05114677920937538,
-0.08699558675289154, 0.08593441545963287, 0.01600094884634018, -0.03908746689558029,
-0.018089503049850464, -0.0721554234623909, -0.057781800627708435, -0.0444064624607563,
-0.03921045362949371, 0.13811585307121277, 0.0006387022440321743, 0.008476483635604382,
-0.0445462167263031, -0.08224783092737198, -0.029969405382871628, -0.008437123149633408,
0.0264445748180151, -0.022518720477819443, -0.03837176039814949, -0.03622477874159813,
-0.007167581468820572, -0.02749662473797798, -0.0332656055688858, 0.12032051384449005,
0.034578509628772736, 0.05442064255475998, -0.060157012194395065, -0.046806819736957 55,
-0.021419420838356018, 0.0025741076096892357, 0.006759037263691425, -0.02807648666203022,
-0.04417530819773674, 0.04646331071853638, -0.006580034736543894, 0.05701450631022453,
0.09510452300310135, -0.12876009941101074, -0.008530572056770325, 0.011226705275475979,
0.09581608325242996, 0.020157204940915108, 0.014205063693225384, -0.08372935652732849,
0.06615494936704636, 0.018033921718597412, -0.031435467302799225, 6.759345064842588e-34,
0.02447868324816227, -0.09170076996088028, -0.05788874253630638, 0.05970682203769684,
0.02238883264362812, 0.07861942797899246, -0.03657889738678932, -0.08320563286542892,
-0.039876360446214676, -0.007982741110026836, -0.1201513260602951, 0.021276108920574188,
0.08724131435155869, 0.04245664179325104, 0.018419252708554268, -0.028671713545918465,
0.08568086475133896, -0.015816504135727882, -0.036640703678131104, 0.06657736003398895,
0.08822507411241531, 0.0053008077666163445, 0.00985805131494999, -0.008028536103665829,
-0.005221199244260788, 0.10164976865053177, -0.0046463217586278915, -0.07752513140439987,
-0.059497907757759094, 0.022404292598366737, -0.018198223784565926, 0.0810912474989891,
-0.09413513541221619, -0.1250005066394806, -0.004304336383938789, 0.03008984588086605,
0.04963436722755432, 0.003272248199209571, -0.02009822428226471, 0.040185362100601196,
-0.12060292065143585, -0.025783967226743698, -0.0556632652878761 3, -0.01073357742279768,
0.056052371859550476, 0.007773055229336023, 0.019149016588926315, -0.004585872404277325,
0.06790769845247269, -0.04694104194641113, -0.0421314463019371, 0.018116947263479233,
-0.017447074875235558, 0.02599960006773472, 0.05007162317633629, -0.010132252238690853,
-0.04034622758626938, -0.02950265258550644, -0.0005605426849797368, 0.02268131449818611,
```

```
0.04679446294903755, -0.007136932574212551, -0.036618128418922424, 0.008421838283538818,
-0.02669495902955532, 0.1233430728316307, -0.0846262276172638, -0.012008724734187126,
-0.05043768137693405, 0.006873901002109051, 0.01444748044013977, -0.11077801138162613,
-0.019904153421521187, 0.0092436783015728, 0.025796694681048393, 0.052212394773960114,
0.10609571635723114, 0.03501278534531593, -0.07007404416799545, -0.0245696809142828,
0.06635146588087082, -0.099256128072738 65, 0.06207851320505142, -0.0648115947842598,
-0.005298446398228407, -0.01027386263012886, -0.0038324708584696054, -0.017424114048480988,
-0.021247928962111473, -0.024782607331871986, -0.05610724538564682, -0.08085664361715317,
0.016458451747894287, -0.08900026232004166, 0.06285323947668076, -3.334240616936768e-8,
-0.03653063252568245, -0.05633283033967018, -0.010023062117397785, 0.0664278194308281,
-0.07236675918102264, -0.007924930192530155, 0.015706369653344154, -0.01408134214580059,
-0.06825344264507294, 0.0281261894851923, 0.08093864470720291, -0.0015228250995278358,
0.039140958338975906, 0.028151294216513634, 0.030856117606163025, 0.007061198819428682,
0.01379757933318615, -0.060205549001693726, -0.07392124086618423, 0.10334544628858566,
-0.033088598400354385, -0.029764683917164803, -0.02184351161122322, -0.09793960303068161,
-0.01702742837369442, -0.01859866827726364, 0.002813448430970311, -0.06446780264377594,
0.004948142915964127, 0.029209036380052567, 0.021648138761520386, 0.0580034106969834,
-0.05144011601805687, -0.004751311149448156, 0.011855087243020535, -0.023282591253519058,
0.0037422312889248133, 0.03716428950428963, 0.0005374533357098699, -0.020539091899991035,
0.00624116463586688, -0.007780138403177261, 0.014025136828422546, 0.08107354491949081,
0.10113450139760971, -0.03616445139050484, 0.006556411739438772, 0.13924580812454224,
-0.00765759963542223, 0.01580081135034561, 0.019034819677472115, -0.04163332283496857,
0.009859669953584671, -0.05536218360066414, 0.050885360687971115, 0.026110874488949776,
0.12086042761802673, 0.048098307102918625, -0.1548193097114563, 0.017976470291614532,
0.07562381774187088, -0.04460503160953522, -0.0014124884037300944, -0.018680419772863388]
```

12. These 384 numbers are not random, nor interpretable by humans. Each dimension captures different aspects of the meaning of the sentence learned from billions of text examples. Together, the 384 numbers create a unique “fingerprint” that captures the sentence’s semantic meaning about justice and society that allows for similarity comparisons.

13. Here, we have used the all-MiniLM-L6-v2 sentence transformer model. There are other embedding models trained on different datasets that use a greater number of dimensions. For example, OpenAI’s text-embedding-3-large model generates embeddings with up to 3,072 dimensions—eight times more than the model used in this demonstration. Using a different embedding model will lead to different results.