

SUSTAINING INNOVATION AND AI IN A DATA-DRIVEN CLIMATE

Industry and Critical Reception to Singapore's Computational Data Analysis Exception

[2023] SAL Prac 28

Gavin **FOO**¹

*LLB (National University of Singapore);
Head, Copyright Unit and Principal Legal Counsel, Legal Department,
Intellectual Property Office of Singapore.*

Trina **HA**¹

*LLB (National University of Singapore);
Chief Legal Counsel and Director, Legal Department, Intellectual Property
Office of Singapore.*

I. Introduction

1 The deployment of data mining and data analytics tools has become increasingly commonplace in both the private and public sectors. The explosion of generative artificial intelligence (“AI”) onto the scene earlier this year has led to businesses and even government bodies worldwide developing such tools to improve efficiency, services and customer experience. For the fintech sector, areas in which such technology solutions are deployed include assisted know-your-client checks, fraudulent transaction detection, customer profiling, investment research and polling of investor recommendations. These tools necessarily involve using and copying materials such as articles, commentaries, blog posts, reports and graphs – all of which may be protected by copyright. Developers and companies, like others

1 Any views expressed in this article are the authors' personal views only and should not be taken to represent the views of their employer. All errors remain the authors' own.

who engage in text and data mining (“TDM”) activities, thus need to assess whether such uses would infringe copyright.

2 Singapore’s copyright exception² for computational data analysis (“CDA”) – often referred to as TDM – has been in force for two years since November 2021. This exception was introduced as part of a larger suite of amendments in the Copyright Act 2021³ (“the Copyright Act 2021”), which aimed to update Singapore’s laws to respond to the technological and market developments of the digital age and achieve a balance between public interest access and the protection of rights.⁴

3 Lawmakers had bold ambitions for the CDA exception to support a wider growth in AI and data-driven innovation in Singapore.⁵ In advocating for the exception during the Parliamentary debates on the Copyright Act 2021, Second Minister for Law Mr Edwin Tong explained that:⁶

... this permitted use can be relied on when training an Artificial Intelligence programme or using computers to analyse large databases of materials ... This supports [Singapore’s] Smart Nation initiatives, ... [Government’s] push towards data-driven innovation and ... efforts to grow [its] Artificial Intelligence and technology sectors.

2 Exceptions to copyright infringement are referred to as “permitted uses” under Pt 5 of the Copyright Act 2021 (2020 Rev Ed).

3 2020 Rev Ed.

4 Singapore Parl Debates; Vol 95, Sitting No 37 [13 September 2021] (Edwin Tong Chun Fai, Second Minister for Law).

5 See, eg, Singapore Parl Debates; Vol 95, Sitting No 37 [13 September 2021] (He Ting Ru, MP for Sengkang): “this is a helpful and necessary exception to have, particularly in the advent of more and more powerful AI capabilities and Singapore’s drive to be a Smart Nation”, and Singapore Parl Debates; Vol 95, Sitting No 37 [13 September 2021] (Raj Joshua Thomas (Nominated Member)): “This will allow for aggregation and analysis of data for such purposes as sentiment analysis or machine learning. These, ultimately, have a public interest aspect to synthesise data for better understanding of, for example, social or economic trends. Importantly, this will support the development of technologies in machine learning and deep learning, which, by their very nature, require the input of a voracious volume of data. The permitted use will allow access to this data without having to seek approval or licensing from each copyright owner, which may be a significant obstacle.”

6 Singapore Parl Debates; Vol 95, Sitting No 37 [13 September 2021] (Edwin Tong Chun Fai, Second Minister for Law).

4 There was strong support from certain quarters of industry as well; that the exception would “propel AI revolution” by assisting to foster increased innovation in Singapore and allowing developers to access far greater quantities of data for training AI algorithms and amassing big data without risking copyright infringement.⁷ Was this optimism well-founded? How have the ambitions of Singapore’s lawmakers panned out?

5 With the exception marking its second year anniversary in Singapore, we examined the industry and critical reception to the exception to understand whether it has indeed been useful for stakeholders and supporting innovation. We sought to understand the issue from two key perspectives:

- (a) scope of the exception – Is the exception broad enough to facilitate TDM uses while balancing rights owners’ interests?
- (b) clarity and certainty – Is the exception clear enough and does it provide adequate certainty to stakeholders?

This article concludes with some insights into what the future might hold for AI, machine learning and other possible applications of the CDA exception in Singapore.

II. Overview of the CDA exception

6 Prior to the introduction of the CDA exception, the main exception that would have been invoked to support TDM activities was Singapore’s general fair dealing exception (now called fair use under s 190 of the Copyright Act 2021). With the

7 “New Singapore Copyright Exception Will Propel AI Revolution”, *Inside Tech Law* (15 November 2021) <<https://www.insidetechlaw.com/blog/new-singapore-copyright-exception-will-propel-ai-revolution>> (accessed 30 August 2023). See also, Alban Kang & Pin-Ping Oh, “Coming Up in Singapore: New Copyright Exception for Text and Data Mining”, *Bird & Bird* (19 September 2021) <<https://www.twobirds.com/en/insights/2021/singapore/coming-up-in-singapore-new-copyright-exception-for-text-and-data-mining>> (accessed 21 August 2023), which highlights that the exception would be of greater utility to businesses due to its mandatory nature, significantly increase the availability of data for TDM and training AI, and significantly lower the cost of TDM research and AI development in Singapore.

CDA exception, TDM activities may potentially be exempted under fair use,⁸ the CDA exception, or *both* exceptions. This is due to the operation of s 184 of the Copyright Act 2021, which provides that a permitted use is generally independent of, and does not affect the application of, any other permitted use, even on the same facts.⁹

7 Part 5, Div 8 of the Copyright Act 2021 provides for the CDA exception. Section 243 defines “computational data analysis” non-exhaustively to include:

- (a) using a computer program to identify, extract and analyse information or data from [a] work or recording [of a protected performance]; 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.

Illustration

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

8 CDA can be carried out on all subject matter that is protected under the Copyright Act 2021, *ie*, works¹⁰ and recordings of protected performances. The exception does not discriminate between whether the CDA is commercial or non-commercial in nature. The exception is a mandatory one in that it cannot be excluded or restricted by contract; any contractual term is void to the extent that it purports, directly or indirectly, to exclude or restrict the exception.¹¹

8 For an analysis of whether TDM would constitute fair use in Singapore, see, *eg*, David Tan & Thomas Lee Chee Seng, “Copying Right in Copyright Law: Fair Use, Computational Data Analysis and the Personal Data Protection Act” (2021) 33 SAclJ 1032 at p 1063, paras 32–38.

9 Explanatory Statement to cl 184 of the Copyright Bill (Bill No 1/2021).

10 Section 8 of the Copyright Act 2021 (2020 Rev Ed) defines a work as an authorial work (*ie* literary, dramatic, musical or artistic work); a published edition of an authorial work; a sound recording; a film; a broadcast or a cable programme.

11 Copyright Act 2021 (2020 Rev Ed) s 187.

9 At the same time, there are important parameters and conditions that outline the contours of what is permissible and prevent the exception from being abused. If any of the conditions in s 244(2) is not satisfied, the exception will not apply and the use of the materials for CDA will be infringing.

10 First, the exception does not apply to all exclusive rights under the Copyright Act 2021; it applies only to acts of copying, and in very narrow circumstances, communication to the public.¹²

11 Second, a purpose-based limitation is built into the exception, in that a copy made with the benefit of the exception must be for the purpose of CDA (which could include preparing a work or recording for CDA). A user cannot use the copies for any other purpose.¹³

12 Third and crucially, a user must have “lawful access” to the material from which a copy is made. Access would not be lawful if the user accesses the material by circumventing paywalls or is in breach of the terms of use of a database (ignoring any terms that are void by virtue of s 187 of the Copyright Act 2021).¹⁴

13 There are further limitations if infringing copies are used for CDA.¹⁵ If a user does so knowingly, the exception will cease to apply.¹⁶ In addition, even if the user does not know that a particular copy is infringing, the exception will also not apply if the user knows or could have reasonably known that the copy was obtained from a flagrantly infringing online location.¹⁷

12 The only circumstances in which a work or recording of a protected performance may be so communicated is if the copy that is communicated was made in circumstances in which the CDA exception applies and the user either communicates the copy for the purpose of verifying the results of the user’s CDA, or for collaborative research or study relating to the user’s CDA: see s 244(4) of the Copyright Act 2021 (2020 Rev Ed).

13 Copyright Act 2021 (2020 Rev Ed) s 244(2)(a) and 244(2)(b).

14 Copyright Act 2021 (2020 Rev Ed) s 244(2)(d).

15 Copyright Act 2021 (2020 Rev Ed) s 244(2)(e).

16 The only instance where a user may use infringing copies knowingly is to conduct research or study relating to rights infringement: see s 244(2)(e)(iii) of the Copyright Act 2021(2020 Rev Ed) read with reg 57 of the Copyright Regulations 2021.

17 Copyright Act 2021 (2020 Rev Ed) s 244(2)(e)(ii).

14 Finally, a user cannot supply the copies made under the exception to any person except for the purpose of verifying the results of the user’s CDA, or for collaborative research or study relating to the user’s CDA.¹⁸

III. Reception to the CDA exception

15 As mentioned above, we examined views on the usefulness of the exception from two key perspectives:

- (a) scope of the exception; and
- (b) clarity of the exception and whether it provides certainty.

16 In examining these perspectives, we reviewed close to 30 commentaries published from 2021 to 2023 across a diverse range of sources, including articles in law and scientific journals, professional and trade publications, press releases, legal updates, blog posts, submissions to foreign governments and legal opinions.¹⁹ Commentators spanned the gamut – including legal practitioners, academics, researchers and research organisations, industry stakeholders and observers, and journalists – who offered both local and international perspectives.

A. Scope of exception

17 In 2022, a comparative study of research and TDM exceptions in more than 190 countries found Singapore, with its CDA exception, to be the world’s second-most open country (behind only Japan).²⁰ The exception was also characterised as “robust” by the Israeli Ministry of Justice in a comparative

18 Copyright Act 2021 (2020 Rev Ed) s 244(2)(c). These are aligned with the same circumstances in which communicating a work or recording may be permitted under the exception.

19 This article does not purport to be a comprehensive review of all commentaries on the exception and, in the interest of brevity, only highlights select commentaries that are emblematic of the issues raised.

20 Sean Flynn *et al*, “Research Exceptions in Comparative Copyright” (2022) PIJIP/TLS Research Paper Series no. 75 at p 36 <<https://digitalcommons.wcl.american.edu/research/75>> (accessed 14 August 2023).

analysis of jurisdictions that allow the use of copyright works for machine learning.²¹

18 It would have been easy for Singapore to have resorted to a broad and unqualified exception in the name of encouraging and supporting TDM uses. However, aside from the three-step test considerations,²² an exception that tilted too far in favour of granting public access would have risked disregarding rights owners' interests and in fact, harming innovation. The challenge was creating an exception that would fall within the proverbial *Goldilocks zone*: broad enough to facilitate TDM uses while balancing rights owners' legitimate interests and concerns.

19 From the commentaries examined for this article, the middle ground was that the exception is indeed broad enough to support TDM uses. Research communities validated the importance of an exception that applies to uses of all kinds of works and enables sharing of materials for the purpose of collaboration and validation.²³ The exception was also said to reflect the reality of today's research ecosystem by not artificially distinguishing between commercial and non-commercial research.²⁴ For the European Alliance for Research Excellence, the exception is "balanced" and serves as "a blueprint for other jurisdictions to create an enabling legal environment for their research and innovation ecosystems" – even as it permits TDM uses, it has safeguards such as lawful access, which takes into account rights owners' interests and "ensures that the process

21 State of Israel, Ministry of Justice, "Opinion: Uses of Copyrighted Materials for Machine Learning" (18 December 2022) at p 31.

22 Under international copyright law, the three-step test requires limitations and exceptions to exclusive rights to be confined to certain special cases which do not conflict with a normal exploitation of a work and do not unreasonably prejudice the legitimate interests of the right holder. See, eg, Berne Convention for the Protection of Literary and Artistic Works Art 9(2) and TRIPS Agreement Art 13.

23 Sean M Fill-Flynn *et al*, "Legal Reform to Enhance Global Text and Data Mining Research" 378(6623) *Science* 951.

24 European Alliance for Research Excellence, "Singapore's New Text and Data Mining Exception Will Support Innovation in the Digital Economy" (20 July 2021) <<https://eare.eu/singapores-new-text-and-data-mining-exception-will-support-innovation-in-the-digital-economy/>> (accessed 16 August 2023).

and results of TDM do not interfere with the economic value or business models associated with publications”.²⁵

20 A few legal commentators and stakeholders opined that the exception could be even more permissive. In particular, Prof David Tan observed that the exception is safeguarded by “five stringent conditions”²⁶ that some users may in fact, still regard as “onerous”.²⁷ In an empirical study reported in an article co-authored by Prof Tan in 2021, one data miner:²⁸

... analogised the strict requirements of the [CDA exception] with an example in patent law – that if contact tracing in a COVID-19 pandemic was patentable by a private company in Singapore, this will result in public harm on a global scale if the only applicable defence against patent infringement requires other countries to get ‘lawful access’ to the patent.

21 By and large however, where commentators disagreed, it was generally on the grounds that the exception is too broad (rather than too restrictive). These views came from certain foreign associations representing rights owners, who generally preferred having no exception at all and instead having recourse to licensing solutions to facilitate TDM.

22 The aspect of the exception that appeared to be their greatest concern was its mandatory nature. As part of a briefing to the UK Intellectual Property Office in response to the UK’s proposed text and data mining exception, the Publishers Association referenced Singapore’s exception and opined that the “extremely broad exception ... without rightsholder opt-out ... created a major business risk in the eyes of our members and

25 European Alliance for Research Excellence, “Singapore’s New Text and Data Mining Exception Will Support Innovation in the Digital Economy” (20 July 2021) <<https://eare.eu/singapores-new-text-and-data-mining-exception-will-support-innovation-in-the-digital-economy/>> (accessed 16 August 2023).

26 David Tan, “The Best Things in Life are Not for Free: Copyright and Generative AI Learning”, *The Singapore Law Gazette* (April 2023).

27 David Tan, “Data Analytics in Fintech and the New Copyright Act” [2022] SAL Prac 4 at para 20.

28 David Tan & Thomas Lee Chee Seng, “Copying Right in Copyright Law: Fair Use, Computational Data Analysis and the Personal Data Protection Act” (2021) 33 SAJLJ 1032 at 1063, para 59.

other creative industry rightsholders”.²⁹ The Copyright Alliance expressed similar sentiments, commenting that the exception “overbroadly permits unauthorised TDM ... for any purpose with no ability for rightsholders to opt out or contract around the exception”.³⁰ This concern has been further amplified by the recent lawsuits filed by various artists in the US against companies that create image generators, alleging copyright infringement and claiming compensation for data they scraped for training.³¹

23 Numerous international legal commentators have, however, suggested that there are limitations to an opt-out approach:³²

TDM research often requires use of massive datasets with works from many publishers, including copyright owners that cannot be identified or are unwilling to grant licences. Forcing researchers to use only licensed or public domain content ... can restrict topics of study, hamper reproducibility and validation, bias results, and dissuade researchers from undertaking projects. [references omitted]

There are also views from industry that requiring permission to train an AI model on copyright works may actually “have the perverse effect of limiting this technology development to the very largest companies, who can assemble a data set by compelling their workers to assign the ‘training right’ as a condition of employment or content creation”.³³

29 “Publishers Association Briefing on Text and Data Mining (TDM)” (26 August 2022) <<https://www.publishers.org.uk/wp-content/uploads/2022/08/22-8-Briefing-note-for-IPO-on-TDM.pdf>> (accessed 16 August 2023) at para 4.3.9.

30 Rachel Kim, “AI and Copyright: AI Policies Must Respect Creators and Their Creatives” (8 December 2022) <<https://copyrightalliance.org/ai-copyright-policies-must-respect-creators>> (accessed 16 August 2023).

31 There is also an open letter by the Authors Guild of America, signed by more than 15,000 writers including well-known authors, urging tech companies responsible for generative AI applications to cease using their works without authorisation or compensation: “Open Letter to Generative AI Leaders” <<https://actionnetwork.org/petitions/authors-guild-open-letter-to-generative-ai-leaders>> (accessed 20 November 2023).

32 Sean M Fill-Flynn *et al*, “Legal Reform to Enhance Global Text and Data Mining Research” 378(6623) *Science* 951 at 951.

33 Katherine Trendacosta & Cory Doctrow, “AI Art Generators and the Online Image Market”, *EFF* (3 April 2023) <<https://www.eff.org/>
(cont’d on the next page)

24 While these comments were not made specifically with reference to Singapore, such international experience appears to resonate with Singapore’s justification for making the CDA exception mandatory. In the Parliamentary debates, it was explained that:³⁴

The benefit of data analysis improves ... [when] the set of underlying data used is as complete as possible. Having specific databases locked out of the analysis because of contractual restrictions would make the results less useful, or worse, create bias or inaccuracies.

Therefore, it is important not to allow private contractual provisions to override this permitted use. It supports the development of useful applications of data analysis, which is indispensable to the current digital economy.

25 Overall, the views of stakeholders and commentators exemplify the delicacy in creating a TDM exception that can accommodate a plurality of views while achieving its goals. The diverse voices that constantly seek to renegotiate the appropriate “balance” in the exception reflect a deeper divide in the interests and affiliations of rights owners versus users, as well as the growing impact of AI technology on larger issues surrounding jobs, business, and economic development – typical of any large-scale disruption.

B. Clarity and certainty

26 In the period preceding the coming into force of the CDA exception, there was lingering scepticism from local legal commentators about the utility of the exception and whether it would be superfluous in light of fair use. There were doubts if a hybrid legal defence strategy (involving both fair use and the exception) would be employed in practice, even if it were theoretically possible. One view was that such a strategy would be time-consuming, expensive and may not reflect favourably on a defendant, and that a defendant would more likely plead fair

[deeplinks/2023/04/ai-art-generators-and-online-image-market>](#) (accessed 20 November 2023).

34 Singapore Parl Debates; Vol 95, Sitting No 37 [13 September 2021] (Edwin Tong Chun Fai, Second Minister for Law).

use given the preponderance of US fair use cases that appear to favour data miners.³⁵

27 However, even with the body of apparently favourable fair use cases in the US, fair use alone might not give Singapore’s stakeholders sufficient certainty. It has been said to be “an unreliable basis to build any broad AI exceptions on or to make general claims that ... AI practices are excused from ... copyright infringement”.³⁶ Indeed, on 30 August 2023, the US Copyright Office published a notice of inquiry in response to “mount[ing]” concerns and uncertainties from stakeholders with diverse views on the use of copyright works to train AI models and other issues. Comments were sought on, among others, the circumstances in which such a use might constitute fair use, and how fair use for the purpose of training AI models should be evaluated in light of the US Supreme Court’s recent decisions.³⁷ As American technologist Andy Baio remarked: “I see people on both sides ... extremely confident in their positions, but the reality is nobody knows ... And anyone who says they know confidently how this will play out in court is wrong.”³⁸

28 Consistent with this, in many commentaries, there was a strong sense that the CDA exception, being a specific exception, is necessary and helpful because it provides much-needed legal certainty in terms of whether TDM activities would infringe copyright.

35 David Tan & Thomas Lee Chee Seng, “Copying Right in Copyright Law: Fair Use, Computational Data Analysis and the Personal Data Protection Act” (2021) 33 SAclJ 1032 at 1063, para 18.

36 Rachel Kim, “AI and Copyright: AI Policies Must Respect Creators and their Creatives”, *Copyright Alliance* (8 December 2022) <<https://copyrightalliance.org/ai-copyright-policies-must-respect-creators>> (accessed 16 August 2023).

37 US Copyright Office, “Artificial Intelligence and Copyright” (2023) 66(167) Federal Register 59942 at 59943 and 59946 (30 August 2023) <<https://www.govinfo.gov/content/pkg/FR-2023-08-30/pdf/2023-18624.pdf>> (accessed 2 September 2023).

38 James Vincent, “The Scary Truth About AI Copyright is Nobody Knows What Will Happen Next”, *The Verge* (15 November 2022) <<https://www.theverge.com/23444685/generative-ai-copyright-infringement-legal-fair-use-training-data>> (accessed 2 September 2023).

29 Singapore’s hybrid TDM framework was cited as a legislative model linking generality and specificity,³⁹ given that it accommodates both fair use and a “highly specific”, “very useful” exception for CDA.⁴⁰ Commenting on the framework, Prof Daryl Lim at Penn State Dickinson Law opined that the problem with fair use as an affirmative defence is that parties cannot know in advance if the use of works for TDM is permissible. This “forces parties to make difficult calculations and deter[s] socially productive conduct”.⁴¹ His view was emphatically encapsulated in the following words:⁴²

When the fuel for the race to be technologically competitive is data, idiosyncratic outcomes are risky and unwise. A statutory exemption like Singapore’s brings the balance and clarity ... [that] AI stakeholders need.

30 Key industry stakeholders and international commentators endorsed this sentiment in welcoming the exception. The Software Alliance (BSA) – which counts global technology companies like Microsoft and IBM among its members – hailed Singapore’s exception as a “game-changer” that “provides much-needed legal clarity for organisations that are developing AI”.⁴³

39 Sean M Fill-Flynn *et al*, “Legal Reform to Enhance Global Text and Data Mining Research” 378(6623) *Science* 951 at 953.

40 Sean Flynn *et al*, “Submission to Canadian Government Consultation on a Modern Copyright Framework for AI and the Internet of Things” (17 September 2021) <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3952238> (accessed 24 August 2023).

41 Daryl Lim, “Life After Google v. Oracle: Three Reflections on a Theme”, *IP Theory 2022* (2 November 2022) at p 47.

42 Daryl Lim, “Life After Google v. Oracle: Three Reflections on a Theme”, *IP Theory 2022* (2 November 2022) at p 56.

43 “The incredible advances in AI capabilities ... in recent years have been enabled by machine learning. At its core, machine learning involves the computational analysis of large data sets to identify correlations and patterns that can be used to make predictions about future data inputs. Many of the most promising applications of machine learning – including computer vision and natural language processing – rely on training data that may be subject to copyright protection, so the exception helps resolve a source of uncertainty for the Singaporean AI research and development community.”: Espie Angelica A de Leon, “Computational Data Analysis Exception in Singapore’s Copyright Act 2021 a Game Changer”, *AsiaIP* (7 December 2021) <<https://asiaiplaw.com/section/news-analysis/computational-data-analysis-exception-in-singapores-copyright-act-2021-a-game-changer>> (accessed 21 August 2023).

31 The exception was also said to provide certainty by “explicitly protect[ing] computational uses of lawfully acquired, non-infringing works”.⁴⁴ It was suggested that such an exception is the “minimum requirement for any copyright law that hopes to keep pace with developments in the digital world”, and that by giving researchers clarity, they would be less hesitant in carrying out TDM and machine learning and hence would achieve greater progress in those fields.⁴⁵

32 Nonetheless, certain local legal commentators opined that specific elements of the exception could benefit from greater clarity and/or prescriptiveness. In Prof David Tan’s view, the exception’s illustrations are of limited usefulness⁴⁶ and there is ambiguity in what lawful access means in a number of situations, such as whether bypassing Robots Exclusion Protocols (“REPs”) would be unlawful, especially since REPs are frequently used by owners to protect their works and are easily adoptable from online repositories.⁴⁷ A client update from Bird & Bird’s Singapore office similarly noted that there is uncertainty in the ambit of the terms “lawful access” and “computational data analysis”.⁴⁸

33 Perhaps – as indeed was acknowledged by the same commentators – some clarity may be found by looking beyond the strict confines of the legislative text, such as through a

44 Mike Palmedo “Singapore’s Copyright Act 2021: New Exception for Computational Uses and Updates to Fair Use and Educational Exceptions” (27 November 2021) <<https://infojustice.org/archives/43799>> (accessed 21 August 2023).

45 Glyn Moody “Singapore Starts Making its Copyright Law Fit for the Digital World; Others Need to Follow its Example” (2 December 2021) <<https://walledculture.org/singapore-starts-making-its-copyright-law-fit-for-the-digital-world-others-need-to-follow-its-example/>> (accessed 21 August 2023).

46 David Tan & Thomas Lee Chee Seng, “Copying Right in Copyright Law: Fair Use, Computational Data Analysis and the Personal Data Protection Act” (2021) 33 SAclJ 1032 at 1063, para 76.

47 David Tan & Thomas Lee Chee Seng, “Copying Right in Copyright Law: Fair Use, Computational Data Analysis and the Personal Data Protection Act” (2021) 33 SAclJ 1032 at 1063, para 40.

48 Alban Kang & Pin-Ping Oh, “Coming Up in Singapore: New Copyright Exception for Text and Data Mining”, *Bird & Bird* (19 September 2021) <<https://www.twobirds.com/en/insights/2021/singapore/coming-up-in-singapore-new-copyright-exception-for-text-and-data-mining>> (accessed 21 August 2023).

comparative approach to interpretation (given that the term “lawful access” is also found in other TDM exceptions like the UK’s and the EU’s) and with the passage of time (as Singapore’s own jurisprudence develops).⁴⁹

IV. Looking beyond ChatGPT

“AI could contribute up to \$15.7 trillion to the global economy in 2030, more than the current output of China and India combined.”⁵⁰

34 With the intense media attention and extensive public discourse around the world on the implications of large language models (LLMs) on the future of the creative industries, it would be easy to lose sight of the other purposes of TDM and AI applications (including generative AI). While generative AI has raised socio-economic implications (that are beyond the remit of this article), the CDA exception (and the like) does have a broader utility in supporting TDM and machine learning in ways that can benefit business and society – from identifying new insights through analysing large volumes of text for medical research, to gathering customer insights and sentiments through perusing social media content for business improvement purposes.

35 In Singapore, the future for machine learning and other forms of data-driven innovation does seem promising. While a longer-term review will be required to make more definitive conclusions on the trajectory of legislative change, we already see AI tools involving TDM being developed and used locally by both the private and public sectors. These include SG Translate Together, a machine translation engine developed by A*STAR’s Institute for InfoComm Research in partnership with the Ministry

49 Alban Kang & Pin-Ping Oh, “Coming Up in Singapore: New Copyright Exception for Text and Data Mining”, *Bird & Bird* (19 September 2021) <<https://www.twobirds.com/en/insights/2021/singapore/coming-up-in-singapore-new-copyright-exception-for-text-and-data-mining>> (accessed 21 August 2023).

50 PwC, “Sizing the Prize –What’s the Real Value of AI For Your Business and How Can You Capitalise?” at p 3 <<https://www.pwc.com/gx/en/issues/analytics/assets/pwc-ai-analysis-sizing-the-prize-report.pdf>> (accessed 20 November 2023).

of Communications and Information. SG Translate Together is powered by AI algorithms and neural networks and has been used by the public since 2022 to generate quality localised translations in Singapore's official languages.⁵¹

36 In the fintech sector, OCBC Securities recently launched AI bots to provide customers with stock recommendations. These recommendations are generated by the AI bots from daily news articles, market movements and relevant news feeds.⁵² OCBC has also rolled out its own generative AI chatbot, OCBC GPT, to assist their employees around the world with writing, research and ideation.⁵³ In an engagement with the Intellectual Property Office of Singapore, Mr Donald MacDonald, Head of OCBC's Group Data Office, affirmed the relevance of the CDA exception:

We anticipate that the computational data analysis exception will be very useful for OCBC's future developments in the AI space. The exception gives us more flexibility and confidence in our use of third-party data as we enhance OCBC's in-house AI capabilities, as part of OCBC's broader aim to increase the adoption of AI across banking services and deliver greater customer value through AI-driven solutions.

37 Generative AI may be the latest TDM application to have emerged from a rapidly evolving technological landscape, but it certainly will not be the last. At the same time, businesses that depend on customer trust and reputation are embarking on technological and business solutions ahead of legislative reform, to seek the balance between innovation and protection of rights owners' legitimate interests.⁵⁴ OpenAI, who sparked the

51 SG Translate Together at <<https://www.sgtranslatetogether.gov.sg/>>.

52 Chew Sue-Ann, "AI Powers Investing Tips, Raises Productivity at OCBC", *The Straits Times* (30 July 2023) <<https://www.straitstimes.com/business/ai-powers-investing-tips-raises-productivity-at-ocbc>> (accessed 30 September 2023).

53 "OCBC is First Singapore Bank to Roll Out Generative AI Chatbot to All Employees Globally" (24 October 2023) <<https://www.ocbc.com/group/media/release/2023/ocbc-is-first-singapore-bank-to-roll-out-generative-ai-chatbot-to-all-employees-globally>> (accessed 20 November 2023).

54 Adobe's generative AI-powered image maker tool, Firefly, only uses Adobe Stock images, openly licensed content and out of copyright images to avoid potential copyright infringement, and offers an intellectual property indemnity in its user agreement: Jon Gold, "Adobe Offers Copyright (cont'd on the next page)

generative AI boom with ChatGPT, announced a deal with The Associated Press (“AP”) to license AP’s archive of news stories.⁵⁵ Is there a future for businesses to adopt a more human/creator-centric view of AI technology as part of their ethical obligations to society, irrespective of the positions in law? These emerging models, together with the larger policy issue of balancing technological innovation against the potential risks to certain skillsets and sectors of the economy, will continue to impact the evolution of the IP regime.

“AI is here, and we will embrace it responsibly together with our music partners.”⁵⁶

Neal Mohan, CEO of YouTube

Indemnification For Firefly AI-based Image App Users”, *Computerworld* (8 June 2023) <<https://www.computerworld.com/article/3699053/adobe-offers-copyright-indemnification-for-firefly-ai-based-image-app-users.html>> (accessed 6 October 2023). Getty Images recently announced that it is partnering with Nvidia to launch a new generative AI tool that lets people create images using its library of licensed photos: Emilia David, “Getty Made an AI Generator That Only Trained on its Licensed Images”, *The Verge* (25 September 2023) <<https://www.theverge.com/2023/9/25/23884679/getty-ai-generative-image-platform-launch>> (accessed 4 October 2023). The latest entrant to this model is Microsoft, who announced a new Copilot Copyright Commitment whereby it will indemnify customers against copyright infringement claims for use of its Copilot service and the output generated: Brad Smith, “Microsoft Announces New Copilot Copyright Commitment for Customers” (7 September 2023) <<https://blogs.microsoft.com/on-the-issues/2023/09/07/copilot-copyright-commitment-ai-legal-concerns/>> (accessed 4 October 2023).

55 Matt O’Brien, “ChatGPT-maker OpenAI Signs Deal With AP to License News Stories”, *The Associated Press* (13 July 2023) <<https://apnews.com/article/openai-chatgpt-associated-press-ap-f86f84c5bcc2f3b98074b38521f5f75a>> (accessed 4 October 2023).

56 Neal Mohan, “Our Principles For Partnering With the Music Industry on AI Technology”, (21 August 2023) <<https://blog.youtube/inside-youtube/partnering-with-the-music-industry-on-ai>> (accessed 6 October 2023), announcing that Universal Music and YouTube have formed a partnership to jointly develop AI tools that offer “safe, responsible and profitable” opportunities to music rightsholders, the first tool being a “Music AI Incubator” at YouTube.