

MONOPOLY RIGHTS VS FREEDOM OF ACCESS: THE COPYRIGHT BALANCE IN A DATA-DRIVEN ECONOMY*

Trina HA[†]

LLB (Hons) (National University of Singapore)

Gavin FOO[‡]

LLB (Hons) (National University of Singapore)

I. Introduction

1 In February 2017, the Committee on the Future Economy (“CFE”) published its recommendations to guide Singapore’s economic strategies for the next decade and achieve an average of 2–3% per year growth for Singapore.¹ The harnessing of data as an asset to generate value was highlighted as an increasingly important source of comparative advantage,² with the CFE urging data-driven innovation.³ This is in line with the increasing worldwide recognition of the potential of data to generate significant financial value and become a key basis of competition, underpinning new waves of productivity growth and innovation.⁴

* 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.

† Director, Legal Department and Head, IP Management (Gov), Intellectual Property Office of Singapore.

‡ Legal Counsel, Legal Department, Intellectual Property Office of Singapore.

1 Report of the Committee on the Future Economy (February 2017) at p 1.

2 Report of the Committee on the Future Economy (February 2017) at paras 37 and 112.

3 Report of the Committee on the Future Economy (February 2017) at paras 111–112.

4 McKinsey Global Institute, *Big Data: The Next Frontier for Innovation, Competition and Productivity* (2011) Executive Summary <<https://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/big-data-the-next-frontier-for-innovation>> (accessed 25 February 2018).

2 In the same year, the Court of Appeal dismissed all but the entirety of an appeal filed by telephone directory publisher Global Yellow Pages Ltd (“GYP”) in respect of its action against its competitor, Promedia Directories Pte Ltd (“Promedia”) for infringing the alleged copyright in GYP’s telephone directories. The High Court found no infringement and this was upheld by the Court of Appeal on the fundamental basis that data are not copyrightable material, even if they are a “valuable commodity” that represent the fruit of an investment.⁵

3 *Global Yellow Pages Ltd v Promedia Directories Pte Ltd*⁶ (“GYP Case”) exemplifies copyright law’s fundamental premises and principles, which have resulted in limited (if any) protection over data and data products, such as data that have been further enhanced, derived or analysed,⁷ compilations of data and databases. Such material often lie “at the very edge of copyright protection”.⁸ As a result, certain jurisdictions have taken the view that a *sui generis* database right is required to safeguard the position of database makers against misappropriation of the fruits of their investment, and in turn, incentivise investments in and exploitation of such compilations.⁹ In the absence of such a right in Singapore, do the limits

5 *Global Yellow Pages Ltd v Promedia Directories Pte Ltd* [2017] 2 SLR 185 at [22] and [34].

6 [2017] 2 SLR 185.

7 For example, Global Yellow Pages Ltd (hereinafter “GYP”) claimed that copyright subsisted in the “enhanced data” in its directories, *ie*, individual business listings after they had been verified, embellished, arranged and classified, see *Global Yellow Pages Ltd v Promedia Directories Pte Ltd* [2017] 2 SLR 185 at [11].

8 *Global Yellow Pages Ltd v Promedia Directories Pte Ltd* [2017] 2 SLR 185 at [89].

9 Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases, see, for example, Recital 39: “Whereas ... this Directive seeks to safeguard the position of makers of databases against misappropriation of the results of the financial and professional investment made in obtaining and collection the contents by protecting the whole or substantial parts of a database against certain acts by a user or competitor”, and Recital 40: “Whereas the object of this *sui generis* right is to ensure protection of any investment in obtaining, verifying or presenting the contents of a database for the limited duration of the right;

(continued on next page)

imposed by copyright law therefore impede the achievement of national aspirations to harness data as an asset and catalyse data-driven innovation?

4 This article posits that conferring stronger protection of data and data products through granting monopolies (whether under copyright law or a *sui generis* regime) may be a blunt tool for achieving data-driven innovation and growth – particularly given the realities and complexities of the present data landscape. Far from being an impediment, the limits set by copyright law may in fact spur innovation and growth in the present landscape. The way forward is not to “stretch”¹⁰ copyright’s fundamental premises to accommodate protection of data and data products, but to make principled extensions within the boundaries of copyright orthodoxy – one current example is the copyright exception for text and data mining, which is discussed in the last part of this article. These considerations are symptomatic of a broader, perennial tension that both afflicts and informs much of intellectual property law in general – balancing the goal of providing adequate incentives for continued creation (by conferring monopoly rights on database makers, producers and the like) against the competing goal of ensuring adequate access (to data and data products).¹¹

II. *Global Yellow Pages Ltd v Promedia Directories Pte Ltd*: Elucidating principles of copyright protection

A. *Brief facts and summary of decision*

5 In the *GYP Case*, GYP (the plaintiff) alleged that its competitor Promedia (the defendant) infringed GYP’s copyright in GYP’s print telephone directories (the *Business Listings*, the *Yellow Pages Business* and the *Yellow Pages Consumer*) and online telephone directory (the *Internet Yellow Pages*). GYP’s copyright had been allegedly infringed by Promedia in the

whereas such investment may consist in the deployment of financial resources and/or the expending of time, effort and energy”.

10 *Global Yellow Pages Ltd v Promedia Directories Pte Ltd* [2017] 2 SLR 185 at [26]. At [35], the Court of Appeal rejected GYP’s argument that a lower standard of creativity should be adopted in the case of factual compilations, in comparison to other authors’ works.

11 See, eg, Tan Tee Jim SC, “New Law for Compilations and Databases in Singapore?” (2012) 24 SAclJ 745 at para 5.

print, digital (CD-ROM) and/or online telephone directories produced and/or maintained by Promedia. GYP also alleged that its copyright was infringed by Promedia's use and deployment of material from GYP's directories in Promedia's temporary database, including by photocopying or scanning the *Business Listings* onto the database.¹² The High Court found no infringement and GYP appealed.

6 The Court of Appeal essentially found GYP's appeal to be unmeritorious because fundamentally, what Promedia took from GYP were data, which are not copyrightable material.¹³

7 The Court of Appeal did not disturb the High Court's finding that copyright did not subsist in the individual business listings in GYP's directories. The High Court found that copyright did not subsist because the form of expression of the individual listings was unoriginal. The selection or arrangement of information within each listing lacked the minimum level of creativity or intellectual effort necessary to make it an original work. Effort, skill and judgment in collecting, verifying, enhancing and arranging the data were not directed at a particular form of expression. Instead, these were directed at ensuring the accuracy of the underlying facts. In the circumstances, granting copyright would amount to granting an impermissible monopoly over the use of bare facts.¹⁴

8 The Court of Appeal considered whether copyright subsisted in the following three classes of GYP's works:

- (a) The listings of businesses arranged within each business classification in GYP's directories.
- (b) The listings in the *Business Listings*.
- (c) The "seeds" in GYP's directories, which were fictitious listings designed principally to detect and prove copying.¹⁵ GYP proved that

12 *Global Yellow Pages Ltd v Promedia Directories Pte Ltd* [2017] 2 SLR 185 at [2], [3] and [8].

13 *Global Yellow Pages Ltd v Promedia Directories Pte Ltd* [2017] 2 SLR 185 at [22].

14 *Global Yellow Pages Ltd v Promedia Directories Pte Ltd* [2017] 2 SLR 185 at [18(a)].

15 *Global Yellow Pages Ltd v Promedia Directories Pte Ltd* [2017] 2 SLR 185 at [50].

Promedia had copied the data in GYP's directories because these seeds were found in Promedia's digital and online directories.¹⁶

9 The Court of Appeal found that copyright did not subsist in any of these, save for a very narrow instance where the *Business Listings* in (b) was concerned:¹⁷

(i) As regards (a), the Court of Appeal agreed with the High Court that copyright cannot subsist in the listings arranged within each classification, because the selection of the contents lacked creativity or was in reality a fact-discovery exercise.¹⁸

(ii) As regards (b), the Court of Appeal held that the *selection* of the listings of businesses in the *Business Listings* lacked creativity, but the *arrangement* exhibited sufficient creativity, albeit only barely so.¹⁹ Even though the listings were presented in alphabetical order,²⁰ GYP applied specific sorting rules to the listings that made minor tweaks to the arrangement of the listings within the narrow confines of an alphabetical arrangement.²¹ The Court of Appeal took the view that the sorting rules taken together, were neither entirely obvious nor inevitable, and this crossed the creativity threshold but just barely.²²

(iii) As regards (c), the court observed that the very purpose of the "seeds" was to be copied. The court held that copyright cannot subsist in a seed as it is not a literary work to begin with, and in any event,

16 *Global Yellow Pages Ltd v Promedia Directories Pte Ltd* [2017] 2 SLR 185 at [10].

17 *Global Yellow Pages Ltd v Promedia Directories Pte Ltd* [2017] 2 SLR 185 at [37].

18 *Global Yellow Pages Ltd v Promedia Directories Pte Ltd* [2017] 2 SLR 185 at [38]–[42].

19 *Global Yellow Pages Ltd v Promedia Directories Pte Ltd* [2017] 2 SLR 185 at [43].

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

21 GYP's sorting rules were expressed in five propositions, which were set out in *Global Yellow Pages Ltd v Promedia Directories Pte Ltd* [2017] 2 SLR 185 at [48].

22 *Global Yellow Pages Ltd v Promedia Directories Pte Ltd* [2017] 2 SLR 185 at [43]–[49].

the copying of the seeds did not cause GYP any loss or Promedia any gain.²³

10 Given that the arrangement of the listings in the *Business Listings* exhibited sufficient creativity but only barely so, the *Business Listings* was found to be “cloaked” with copyright, although the resulting protection was “extremely thin”.²⁴ This meant that infringement would only be established by nothing less than near-wholesale taking of the listings in the *Business Listings*, arranged exactly as they were in the *Business Listings*.²⁵ This was found to be the case as Promedia had photocopied or scanned the entire *Business Listings* onto its temporary database.²⁶ Its *modus operandi* was to systematically copy the *Business Listings*, which was a key source of information for its master database.²⁷

11 Notwithstanding this, Promedia was not liable for infringement because the Court of Appeal held that its photocopying or scanning of the *Business Listings* into the temporary database amounted to fair dealing within s 35 of the Copyright Act²⁸ (“Copyright Act”).²⁹ The *GYP Case* is significant for the Court of Appeal’s elucidation of the approach to assessing fair dealing under s 35 and application to the specific facts in that case, but this aspect of the decision will not be discussed in this article.

12 The *GYP Case* elucidates the following fundamental copyright principles and their implications for protection of data and data products, which are examined in greater detail below:

23 *Global Yellow Pages Ltd v Promedia Directories Pte Ltd* [2017] 2 SLR 185 at [52].

24 *Global Yellow Pages Ltd v Promedia Directories Pte Ltd* [2017] 2 SLR 185 at [49].

25 *Global Yellow Pages Ltd v Promedia Directories Pte Ltd* [2017] 2 SLR 185 at [49].

26 *Global Yellow Pages Ltd v Promedia Directories Pte Ltd* [2017] 2 SLR 185 at [71].

27 *Global Yellow Pages Ltd v Promedia Directories Pte Ltd* [2017] 2 SLR 185 at [65].

28 Cap 63, 2006 Rev Ed.

29 *Global Yellow Pages Ltd v Promedia Directories Pte Ltd* [2017] 2 SLR 185 at [71].

- (a) Data *per se* are not copyrightable but a compilation of data is.
- (b) Data compilations must be “original”, *ie*, data must be selected and/or arranged with sufficient creativity.
- (c) Data compilations must have human authors.
- (d) Copyright in data compilations is “thin”.

B. Data per se are not copyrightable but a compilation of data is

13 It is axiomatic that copyright does not protect ideas and facts but the form of expression of such. Correspondingly, data *per se* are not protected; a compilation of data however, may attract copyright. Article 10(2) of the Agreement on Trade-Related Aspects of Intellectual Property Rights (“TRIPS Agreement”) demonstrates the international consensus on this distinction and provides:

Compilations of data or other material, whether in machine readable or other form, which by reason of the selection or arrangement of their contents constitute intellectual creations shall be protected as such. Such protection, which shall not extend to the data or material itself, shall be without prejudice to any copyright subsisting in the data or material itself.

14 The position under the Copyright Act complies with Art 10(2) of the TRIPS Agreement. The Copyright Act protects a compilation of data as an original literary work, but any copyright subsisting in such a compilation is “limited to the selection or arrangement of its contents which constitutes an intellectual creation”; copyright does not subsist in the underlying data or facts.³⁰

15 Examples of compilations of facts and/or data that have been protected by copyright law include tables, databases, telephone directories, street directories, railway tables, examination papers, trade catalogues, a racing information service, football fixtures, betting lists and listings of television broadcast programmes.³¹

30 Copyright Act (Cap 63, 2006 Rev Ed) ss 4, 7A and 27. As the Court of Appeal stressed in *Global Yellow Pages Ltd v Promedia Directories Pte Ltd* [2017] 2 SLR 185 at [15]: “Copyright protects not ideas, facts or data, but the expression thereof.”

31 Tan Tee Jim SC, “New Law for Compilations and Databases in Singapore?” (2012) 24 SAclJ 745 at para 12.

16 By protecting only expressions of data and other facts, copyright law ensures that the underlying data and facts always remain in the public domain and free to use.

C. *Data compilations must be “original”, ie, data must be selected and/or arranged with sufficient creativity*

17 For copyright to subsist in a compilation of data, the compilation must be “original”.³² There are two main approaches to determine whether a compilation is original. These approaches have differing views on the quality and the object of the effort in producing a work, and were summarised by the Court of Appeal in the *GYP Case* as follows:³³

(a) The “sweat of the brow” approach. This considers the author’s labour and industry to be relevant, and admits consideration of the effort taken in the preparatory steps leading to the reduction of the work to its final form.

(b) The “creativity” approach. This places emphasis on the end product of the work and considers efforts applied towards the formulation of how that end product will be expressed. This approach does not protect preparatory efforts or the process of gathering facts, and an expression of data that involves little ingenuity or skill beyond mechanical labour or routine programming may well not attract copyright protection.³⁴

18 The Court of Appeal observed that there has been a “noticeable retreat from the ‘sweat of the brow’ approach that once featured in the decisions of the Australian and English courts”, and the creativity approach applies in Singapore.³⁵ For copyright to subsist in any literary work, there must be an

32 Copyright Act (Cap 63, 2006 Rev Ed) s 7A read with s 27.

33 *Global Yellow Pages Ltd v Promedia Directories Pte Ltd* [2017] 2 SLR 185 at [23].

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

35 *Global Yellow Pages Ltd v Promedia Directories Pte Ltd* [2017] 2 SLR 185 at [26].

“authorial creation that is causally connected with the engagement of the human intellect”.³⁶ This means:³⁷

... the application of intellectual effort, creativity, or the exercise of mental labour, skill or judgment. Effort (even intellectual) that is applied not towards the authorial creation but towards other ends such as the verification of facts will not be relevant ... even if such verified facts might be the eventual subject of the authorial creation ... It follows from this that the compiler must exercise sufficient creativity in selecting or arranging the material within the compilation; and if the compiler does so, the resulting copyright will only protect the original expression in the form of the selection or arrangement of the material, as the case may be ...

19 There is no prescribed minimum level of creativity required for copyright to subsist in a compilation.³⁸ Instead, the sufficiency of creativity is a question of fact and degree, *ie*, whether in each case the human acts done and choices made in reducing the universe of raw information and data into a compilation demonstrate sufficient creativity and bear sufficient causal nexus with the final work (*ie*, the compilation) such that the compilation is found to be “original” for the purpose of copyright subsistence.³⁹

20 This exercise is heavily fact-centric and can be a “nuanced” one, as the Court of Appeal acknowledged in the *GYP Case*.⁴⁰ At one end of the spectrum, there are managerial decisions which clearly do not merit copyright protection – such as whether to publish or update a phone directory. Then there are purely mechanical tasks that, taken alone, would not cross the creativity threshold – such as the algorithmic collection or arrangement of data. Then there are choices in between that require human judgment and do impact the selection or arrangement of the material, and

36 *Global Yellow Pages Ltd v Promedia Directories Pte Ltd* [2017] 2 SLR 185 at [24].

37 *Global Yellow Pages Ltd v Promedia Directories Pte Ltd* [2017] 2 SLR 185 at [24].

38 *Global Yellow Pages Ltd v Promedia Directories Pte Ltd* [2017] 2 SLR 185 at [28].

39 *Global Yellow Pages Ltd v Promedia Directories Pte Ltd* [2017] 2 SLR 185 at [28]–[30].

40 *Global Yellow Pages Ltd v Promedia Directories Pte Ltd* [2017] 2 SLR 185 at [30]–[31].

which might more properly be regarded as authorial in nature.⁴¹ In essence, the spark of creativity must extend beyond mere data aggregation and input.

D. Data compilations must have human authors

21 For copyright to subsist in a compilation of data, the author of the compilation must also be a natural person (as opposed to a non-living “person” such as an incorporated body) who must first be identified. Otherwise, the work product cannot be deemed original for the purpose of copyright subsistence.⁴² This principle was alluded to although not expressly discussed in the *GYP Case*.⁴³ The principle was the subject of a different Court of Appeal decision in *Asia Pacific Publishing Pte Ltd v Pioneers & Leaders (Publishers) Pte Ltd*⁴⁴ (“*APP Case*”), where the court held that it is not necessary to name each and every human author to make out a claim for copyright protection but it has to be shown that the work product in question has been generated from human author(s) working alone or collaboratively, *ie*, the existence of such persons must be clearly established.⁴⁵

22 This principle has potentially wide-ranging implications given that “most of the produced and collected data in the digital economy, which are

41 *Global Yellow Pages Ltd v Promedia Directories Pte Ltd* [2017] 2 SLR 185 at [28].

42 *Asia Pacific Publishing Pte Ltd v Pioneers & Leaders (Publishers) Pte Ltd* [2011] 4 SLR 381 at [75].

43 See, *eg*, *Global Yellow Pages Ltd v Promedia Directories Pte Ltd* [2017] 2 SLR 185 at [24]: “for copyright to subsist in any literary work, there must be authorial creation that is causally connected with the engagement of the human intellect”, and [28]: “the sufficiency of intellectual effort will often be ... a question of ‘fact and degree’ ... What is common across the jurisdictions, however, is that each court seeks to characterise the human acts done and decisions made in reducing the universe of raw information into a work”.

44 [2011] 4 SLR 381.

45 *Asia Pacific Publishing Pte Ltd v Pioneers & Leaders (Publishers) Pte Ltd* [2011] 4 SLR 381 at [75].

so important for data analytics ... are machine-generated data”.⁴⁶ Further, large corporations that invest significant labour and financial resources to produce compilations often employ a high degree of automation and an army of workers in the entire production chain whose contribution lies in merely executing discrete tasks of data gathering, data verification, data entry and data presentation. It is not unusual for different steps in the production of these compilations to be outsourced or allocated to subcontractors or employees located in different countries. Even if the names of these individuals may be identified, it is unlikely that the human author requirement – which possesses the aforesaid element of creativity – would be satisfied.⁴⁷

23 To the extent that data compilations do not satisfy the human author requirement, copyright will not subsist in the same. Both this requirement and the creativity approach acknowledge the significant role of technology in producing compilation works and promote an important societal goal of ensuring public access to knowledge contained in data compilations.⁴⁸ Such material will be entirely free to access by other players in the data landscape, and this may then spur innovation and new business models.

E. “Thin” copyright in data compilations

24 The foregoing principles relate to copyright subsistence, *ie*, if a compilation of data is original and authored by human author(s), it would conceivably be eligible for copyright protection. However, even if copyright subsists, the protection conferred would be “thin”. The thinner the protection, the more substantial the copying must be before a finding of

46 Wolfgang Kerber, *A New (Intellectual) Property Right for Non-Personal Data? An Economic Analysis* at p 2 <https://www.uni-marburg.de/fb02/makro/forschung/magkspapers/paper_2016/37-2016_kerber.pdf> (accessed 14 February 2018).

47 David Tan, “Copyright in Compilations: Embarking on a Renewed Quest for the Human Author and the Creative Spark” (2013) 18 *Media and Arts Law Review* 151 at 160–161.

48 David Tan, “Copyright in Compilations: Embarking on a Renewed Quest for the Human Author and the Creative Spark” (2013) 18 *Media and Arts Law Review* 151 at 162.

copyright infringement will be made. The Court of Appeal's explanation in the *GYP Case* is worth setting out in full:⁴⁹

[T]he method of analysis is heavily shaped by what it is in a work that attracts copyright protection, especially where infringement and fair dealing are concerned. In other words, there is a nexus between the originality, skill and effort that goes into a work, and the substantiality of copying required to establish infringement. Thus, *although copyright may subsist in a work as a whole, there would be no infringement of such copyright unless one copies the work as a whole, or a substantial portion of the part of the work that attracts copyright protection in the first place ...* For example, the copyright that subsists in a *compilation of all the news articles* published by a certain agency in a certain month by virtue only of its *thematic arrangement* would not be infringed by another compilation of the same selection of articles arranged *chronologically* because, even though the same underlying material might have been copied, the arrangement, *in which copyright was found to subsist*, was not ... *the thinner the copyright protection, the more substantial the copying must be before a finding of infringement will be made.*[emphasis in bold italics added; emphasis in italics in original]

III. Reappraising the case for conferring stronger monopolies in data and data products

A. *Sui generis* right(s) in data and data products

25 As can be seen from the above principles, copyright law tends to lean in favour of allowing freedom of access to and use of data and data products. In the *GYP Case*, the Court of Appeal explained that “the commercial immorality underlying ... conduct [involving appropriation of data or facts that represents the fruit of an investment] ... is ... not within the purview of copyright law ... [and] is more properly the province of a *sui generis* database right, which has been recognised in the European Union but not in Singapore”.⁵⁰ Based on an incentive theory of intellectual property, the limited protection of data and data products under copyright law and lack of a *sui generis* database right may suggest that there is or will

49 *Global Yellow Pages Ltd v Promedia Directories Pte Ltd* [2017] 2 SLR 185 at [16].

50 *Global Yellow Pages Ltd v Promedia Directories Pte Ltd* [2017] 2 SLR 185 at [34].

be a lack of incentive to invest in and monetise data and data products, which in turn, will impede innovation in Singapore. The logic is that if subsequent compilers can take data from a compilation with impunity, and are thereby able to sell competing compilations at a lower price, there would be little incentive to invest in the creation of factual compilations in unchartered areas.⁵¹ Any advantages of lead time may be diminished in an industry where the buying public would be attracted to the later compilations because they are usually an updated version.⁵²

26 Such reasoning underpinned the introduction of Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases (“EU Database Directive”), which gives the maker of a database who shows that there has been qualitatively and/or quantitatively a substantial investment in obtaining, verifying or presenting the contents of the database, a right to prevent extraction and/or re-utilisation of the whole or of a substantial part, evaluated qualitatively and/or quantitatively, of the contents of that database.⁵³

27 Already, there are calls in the European Union (“EU”) to go beyond the EU Database Directive and create a “new data producer right with the objective of enhancing the tradability of non-personal or anonymised machine-generated data as an economic good”.⁵⁴ It is hoped that this would

51 Ng-Loy Wee Loon, “Copyright Protection for Traditional Compilations of Fact and Computerized Databases – Is Sweat Copyrightable?” [1995] Sing JLS 96 at 117.

52 Ng-Loy Wee Loon, “Copyright Protection for Traditional Compilations of Fact and Computerized Databases – Is Sweat Copyrightable?” [1995] Sing JLS 96 at 117.

53 Article 7(1) of the Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases. “Database” is defined in Art 1(2) as “a collection of independent works, data or other materials arranged in a systematic or methodical way and individually accessible by electronic or other means”, and “maker of a database” is defined in Recital 41 as the “person who takes the initiative and the risk of investing ... [and] excludes subcontractors”.

54 Commission Staff Working Document on the free flow of data and emerging issues of the European data economy, Accompanying the document Communication “Building a European Data Economy” (SWD/2017/02 final) at p 33 <<http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52017SC0002>> (accessed 14 February 2018).

facilitate and incentivise sharing of such data and protect investments and assets.⁵⁵ The proposal is in its preliminary stages. At the time of writing, the right has been envisaged as either a right *in rem* (*ie*, a property right enforceable against the world independent of contractual relations, which would mean a right to utilise the data and license its usage) or a set of purely defensive rights providing for the capacity to sue for misappropriation of data (similar to an action for breach of confidence).⁵⁶

B. Are monopolies in data and data products desirable or even appropriate?

28 Whether the introduction of such *sui generis* right(s) is the way to go for Singapore is a matter requiring careful consideration. For the present purpose, it suffices to observe that the appeal of the incentive theory (*ie*, lack of protection kills creativity) lies in the simplicity of its logic, which may not always hold true. In 2016, an economic study assessing the need for a new intellectual property right on data in the EU directly challenged the incentive theory.⁵⁷

In the discussion about data property so far, nobody has claimed that we have a general incentive problem in the digital economy as regards the collection, production, and analysis of data. To the contrary, the empirical fact of the massive and often simple production of huge amounts of data and their analysis seems to be one of the most important characteristics of Big Data and the digital economy. The amount of collected data is increasing exponentially, and it is widely expected that through the spreading of sensor technology and the ‘internet of things’ this trend will continue for the foreseeable future.

55 Commission Staff Working Document on the free flow of data and emerging issues of the European data economy, Accompanying the document Communication “Building a European Data Economy” (SWD/2017/02 final) at p 30.

56 Commission Staff Working Document on the free flow of data and emerging issues of the European data economy, Accompanying the document Communication “Building a European Data Economy” (SWD/2017/02 final) at pp 33–34.

57 Wolfgang Kerber, *A New (Intellectual) Property Right for Non-Personal Data? An Economic Analysis* at pp 8–9 <https://www.uni-marburg.de/fb02/makro/forschung/magkspapers/paper_2016/37-2016_kerber.pdf> (accessed 14 February 2018).

29 It may even be said that in the world of databases, *copying actually sparks innovation*. The US, where copying is allowed in the absence of a *sui generis* database right, has a much more vibrant database industry than Europe. In particular, as a result of the higher originality standard for compilation works in the US, the cost of creating new and better databases has been reduced. The European approach may have resulted in higher profits for individual database producers but it has not grown the overall industry, in part because it “chokes off the kind of beneficial tweaking and reworking that are so useful to innovation”.⁵⁸

30 It should therefore be considered whether the grant of monopolies in data and data products sits well with the “functional logic” of Big Data and the digital economy in the first place – to use data from many different sources, combine them, analyse them, derive new data, which again can be used for further analyses in combination with other data.⁵⁹ One of the characteristics of the data economy is that data can often be used for analyses in many different contexts and for solving many different problems. It has thus been claimed that in order to develop innovations in the digital economy, it is essential to have easy access to many different kinds of data and that all impediments to the free flow of data should be eliminated as far as possible.⁶⁰

31 Ultimately, the complexity of the contemporary data landscape should not be overlooked – whether in terms of the players in the landscape and relationship between one another;⁶¹ the range and nature of dealings with

58 David Tan, “Copyright in compilations: Embarking on a renewed quest for the human author and the creative spark” (2013) 18 *Media and Arts Law Review* 151 at 162–163.

59 Wolfgang Kerber, *A New (Intellectual) Property Right for Non-Personal Data? An Economic Analysis* at p 19 <https://www.uni-marburg.de/fb02/makro/forschung/magkspapers/paper_2016/37-2016_kerber.pdf> (accessed 14 February 2018).

60 Wolfgang Kerber, *A New (Intellectual) Property Right for Non-Personal Data? An Economic Analysis* at p 19 <https://www.uni-marburg.de/fb02/makro/forschung/magkspapers/paper_2016/37-2016_kerber.pdf> (accessed 14 February 2018).

61 In a study of the data landscape in Singapore conducted by KPMG Services Pte Ltd and commissioned by the Competition Commission of Singapore, the main players in the Singapore data landscape were identified as government

(continued on next page)

data and data products; or the benefits accompanying such dealings.⁶² This includes the emergence of “complex new value networks of firms”, where connected firms often contribute technically and economically to the production and processing of data. These complexities do not admit clear answers as to who should be granted monopoly rights and reap the economic value of data and data products: The data producer? The firm that is economically responsible for producing the data? The firm that can use the data most efficiently? Or the firm that can benefit the most from the data?⁶³ The very notion of monopoly rights may not be appropriate in these circumstances.

IV. Copyright exception for text and data mining

32 As a result of the increasing prominence and proliferation of data and data products, there is a burgeoning chorus of views that text and data

agencies, businesses, customers, data analytics solutions providers, data storage providers, and data aggregators and brokers, see *Data: Engine for Growth – Implications for Competition Law, Personal Data Protection, and Intellectual Property Rights* (16 August 2017) at paras 27–42 <<https://www.ccs.gov.sg/media-and-publications/publications/studies-research-papers/occasional-papers/data-engine-for-growth>> (accessed 14 February 2018).

62 For example, it has been suggested that the data economy measures the overall impact of the data market on the economy as a whole, and involves the generation, collection, storage, processing, distribution, analysis, elaboration, delivery and exploitation of data enabled by digital technologies. This includes not only the direct effects of the data market on the economy, but also the indirect and induced effects. The benefits extend beyond direct monetisation of data and data products, and can also translate into revenue increases, cost optimisation, operational efficiency, operational effectiveness and promotion of information and communication technology (“ICT”) adoption, entrepreneurship and new venture; see *Final Results of the European Data Market Study Measuring the Size and Trends of the EU Data Economy, Executive Summary* (1 February 2017) at pp 10 and 20–21 <<https://ec.europa.eu/digital-single-market/en/news/final-results-european-data-market-study-measuring-size-and-trends-eu-data-economy>> (accessed 25 February 2018).

63 Wolfgang Kerber, *A New (Intellectual) Property Right for Non-Personal Data? An Economic Analysis* at pp 6 and 15 <https://www.uni-marburg.de/fb02/makro/forschung/magkspapers/paper_2016/37-2016_kerber.pdf> (accessed 14 February 2018).

mining activities – which both create and use data and data products – should be conducted without risk of copyright infringement. Legitimising text and data mining activities is consistent with orthodox copyright principles, and as an alternative to granting monopolies in data and data products, may even achieve greater innovation and growth in the contemporary data-driven economy.

33 “Text and data mining” generally refers to the use of automated analytical techniques to analyse text and data in digital form in order to generate information and insights such as patterns, trends and correlations.⁶⁴ Such information and insights may not have been possible or feasible to obtain through manual effort. These techniques work by deriving information from machine-read material – copying large quantities of material, extracting the data, and thereafter analysing the data.⁶⁵ The actual process of analysis may involve making further copies of the material, including by storage in computers or on any medium by electronic means.

34 To the extent that such techniques involve copying or reproducing material that may be protected by copyright, text and data mining activities may give rise to liability for copyright infringement in the absence of a specific copyright exception. This can represent a barrier to innovation and economic opportunity – an independent review of intellectual property and growth in the UK in 2011 found that research scientists, including medical researchers, were being hampered from using computerised search and analysis techniques on data and text because the then-copyright law in the UK could possibly forbid or restrict such usage.⁶⁶

64 See, for example, the definitions of “text and data mining” in Art 2(2) of the Proposal for a Directive of the European Parliament and of the Council on copyright in the Digital Single Market <<https://ec.europa.eu/digital-single-market/en/news/proposal-directive-european-parliament-and-council-copyright-digital-single-market>> (accessed 15 February 2018); and Australian Law Reform Commission, *Copyright and the Digital Economy* (Discussion Paper 79, November 2013) at para 11.57.

65 (Hargreaves review) Supporting Document T – Text Mining and Data Analytics in *Call for Evidence* Responses at p 1 <<http://webarchive.nationalarchives.gov.uk/20140603125140/http://www.ipso.gov.uk/ipreview-doc-t.pdf>> (accessed 15 February 2018).

66 Ian Hargreaves, *Digital Opportunity: A Review of Intellectual Property and Growth* (May 2011) at para 5.10.

35 Given this, copyright legislation in the UK⁶⁷ and Japan⁶⁸ provide for a specific copyright exception for text and data mining. The exception has been proposed in Singapore,⁶⁹ as well as other jurisdictions such as Australia⁷⁰ and the EU.⁷¹

36 The exception is supported by “growing recognition” that text and data mining should not be infringement because it is “non-expressive use”.⁷² Non-expressive use leans on the fundamental principle that copyright law protects only the expression of facts, ideas and data. This point is aptly illustrated by the following example from the Australian Law Reform Commission:⁷³

Consider a computer algorithm employed to search through a text to obtain metadata, which discovers two facts about Moby Dick:

first, that the word “whale” appears 1119 times; second, that the word “dinosaur” appears 0 times. While a whale is certainly central to the expression contained in Moby Dick, this data is not. Rather, metadata of this sort ... is factual and non-expressive, and incapable of infringing the rights of copyright holders.

37 In other words, the text and data mining exception dovetails with the classic limits of copyright law. The fact-discovery purposes enabled by text and data mining should be permitted because such techniques do not seek

67 Copyright, Designs and Patents Act 1988 (c 48) (UK) s 29A (“Copies for text and data analysis for non-commercial research”).

68 Japan Copyright Law Arti 47^{septies} (“Reproduction, *etc* for information analysis”).

69 MinLaw and IPOS Public Consultation on Proposed Changes to Singapore’s Copyright Regime (23 August 2016), Proposal 9 at paras 3.61–3.64.

70 Australian Law Reform Commission, *Copyright and the Digital Economy* (Discussion Paper 79, November 2013) at paras 11.57–11.84.

71 Article 3 of the Proposal for a Directive of the European Parliament and of the Council on copyright in the Digital Single Market <<https://ec.europa.eu/digital-single-market/en/news/proposal-directive-european-parliament-and-council-copyright-digital-single-market>> (accessed 15 February 2018).

72 Australian Law Reform Commission, *Copyright and the Digital Economy* (Discussion Paper 79, November 2013) at para 11.60.

73 Australian Law Reform Commission, *Copyright and the Digital Economy* (Discussion Paper 79, November 2013) at para 11.60, citing M Jockers, M Sag & J Schultz, *Brief of Digital Humanities and Law Scholars as Amici Curiae in Authors Guild v Hathitrust* (2013) at p 18.

to make use of the expressions that are intended for copyright protection.⁷⁴ Uses that treat copyright works as mere data – rather than for their expressive value – do not compete with the original works and should not be treated as falling within the scope of the copyright owner’s rights.⁷⁵ As Professor Ian Hargreaves explained:⁷⁶

[Text and data mining] do[es] not directly trade on the underlying creative and expressive purpose of [copyright] work[s] ... This is not about overriding the aim of copyright – these uses do not compete with the normal exploitation of the work itself – indeed, they may facilitate it ... That these new uses happen to fall within the scope of copyright regulation is essentially a side effect of how copyright has been defined, rather than being directly relevant to what copyright is supposed to protect [*ie*, expressions].

38 The text and data mining exception may also be understood as an extension of the specific fair dealing exception for research (*eg*, the UK⁷⁷ and Australia⁷⁸) or a subset of fair use/general fair dealing (*eg*, the US⁷⁹ and Singapore⁸⁰). To the extent that the author’s original expression is not

74 See also, (Hargreaves review) Supporting Document T – Text Mining and Data Analytics in *Call for Evidence* Responses at p 1 <<http://webarchive.nationalarchives.gov.uk/20140603125140/http://www.ipo.gov.uk/ipreview-doc-t.pdf>> (accessed 15 February 2018): “There are strong arguments for amending the legal framework to provide a mechanism to allow for text and data mining to be undertaken without requiring permission from rights holders because the technique does not seek to make use of the expressions intended for protection.”

75 Australian Law Reform Commission, *Copyright and the Digital Economy* (Discussion Paper 79, November 2013) at para 11.62.

76 Ian Hargreaves, *Digital Opportunity: A Review of Intellectual Property and Growth* (May 2011) at para 5.24.

77 Copyright, Designs and Patents Act 1988 (c 48) (UK) ss 29 and 29A.

78 Australia Copyright Act 1968 s 40.

79 US Copyright Act 17 USC §107. For an analysis of how text and data mining may be understood as fair use within the meaning of §107, see Krista L Cox, “Research Libraries and New Technologies, Promoting Access to Information, Learning, and Innovation for Today and the Future” (2016) 13:1 I/S: J L & Pol’y for Info Soc’y 261 at 266–269.

80 Copyright Act (Cap 63, 2006 Rev Ed) ss 35 and 109. In *Global Yellow Pages Ltd v Promedia Directories Pte Ltd* [2017] 2 SLR 185, the Court of Appeal found that Promedia’s photocopying or scanning of the *Business Listings* into its temporary database amounted to fair dealing within s 35. The court’s

(continued on next page)

substituted, non-expressive uses in the form of text and data mining may be considered highly transformative: their “purpose and character” is such that they do not merely supersede the objects of the original creation.⁸¹

39 Copyright orthodoxy aside, the text and data mining exception is aligned with the “functional logic” of Big Data and the digital economy.⁸² At a commercial level, the ability to extract value is an increasingly important feature of the digital economy. Tellingly, there is “widespread acknowledgment” that text and data mining can benefit the research community and in doing so, encourage innovation;⁸³ it paves the way for novel discoveries, increased research output, early identification of problems,⁸⁴ and has the potential to grant “immense inferential power” to allow businesses, researchers and institutions to make proactive, knowledge-driven decisions.⁸⁵ Businesses can better understand and predict customers’ interests, focus efforts and resources on more profitable areas, and achieve better performance overall.⁸⁶ Text and data mining has also been known to increase the speed of processes and reduce transaction costs across a range of applications.⁸⁷

40 To reap such benefits and more, a text and data mining exception has been proposed as part of a comprehensive review of Singapore’s copyright

reasoning suggests that Promedia only made non-expressive use of the Business Listings (at [87]–[91]).

81 Australian Law Reform Commission, *Copyright and the Digital Economy* (Discussion Paper 79, November 2013) at para 11.61.

82 See para 30 above.

83 Recital 8 of the Proposal for a Directive of the European Parliament and of the Council on copyright in the Digital Single Market <<https://ec.europa.eu/digital-single-market/en/news/proposal-directive-european-parliament-and-council-copyright-digital-single-market>> (accessed 15 February 2018).

84 Australian Law Reform Commission, *Copyright and the Digital Economy* (Discussion Paper 79, November 2013) at para 11.58.

85 Australian Law Reform Commission, *Copyright and the Digital Economy* (Discussion Paper 79, November 2013) at para 11.59.

86 Australian Law Reform Commission, *Copyright and the Digital Economy* (Discussion Paper 79, November 2013) at para 11.59.

87 (Hargreaves review) Supporting Document T – Text Mining and Data Analytics in *Call for Evidence* Responses at p 5 <<http://webarchive.nationalarchives.gov.uk/20140603125140/http://www.ipso.gov.uk/ipreview-doc-t.pdf>> (accessed 15 February 2018).

regime.⁸⁸ The public consultation paper explained that text and data mining is “integral to Singapore’s Smart Nation initiative and allowing for such activities to operate freely would ... help to create and disseminate knowledge.”⁸⁹ The *Report of the Committee on the Future Economy*⁹⁰ similarly suggests that the exception assumes a strategic, national significance, as the CFE had advocated the use of data analytics to develop innovative solutions that serve the needs of and solve specific problems within various industries in Singapore.⁹¹

41 The proposed exception in Singapore will allow copying or reproduction of copyright works for the purpose of data analysis. Users of such works must have legitimate access to the works in the first place (such as through paid subscriptions or access that is not limited by any arrangement where content is accessible only after payment). While both commercial and non-commercial activities would be permitted under the proposed exception, the copying and reproduction must be for the purpose of data analysis to benefit from the exception; the exception is not intended to cover situations where the commercial benefit comes from the actual copies of the copyright works, instead of the data analysis.⁹²

V. Concluding remarks

42 The intrinsic balance in copyright law is aptly summed up by the Court of Appeal in the *GYP Case*:⁹³

88 MinLaw and IPOS Public Consultation on Proposed Changes to Singapore’s Copyright Regime (23 August 2016), Proposal 9 at paras 3.61–3.64.

89 MinLaw and IPOS Public Consultation on Proposed Changes to Singapore’s Copyright Regime (23 August 2016), Proposal 9 at paras 3.63.

90 February 2017.

91 Report of the Committee on the Future Economy (February 2017) at para 109a.

92 As proposed in para 3.64 of the MinLaw and IPOS Public Consultation on Proposed Changes to Singapore’s Copyright Regime (23 August 2016). Whether such an exception will be created, and if so, the details of the exception, is subject to the Government’s eventual recommendations taking into account stakeholder feedback from the consultation.

93 *Global Yellow Pages Ltd v Promedia Directories Pte Ltd* [2017] 2 SLR 185 at [74].

Copyright law is meant to promote creativity and innovation by granting exclusive rights to copyright holders, but there is equally a public interest in not allowing copyright law to hinder creativity and innovation.

43 Where data and data products are concerned, the balance in copyright law tends to lean in favour of allowing free access and use, given the fact-based nature of such material and high degree of automation often involved in its creation. A text and data mining exception may tilt the balance yet further towards allowing uninhibited access and use. Such developments would spur data-driven innovation and growth.

44 Of course, it is not a zero-sum game in that there is always a possibility that *sui generis* right(s) in data and data products can complement copyright law to achieve greater innovation overall. One should also consider other measures of protecting data and data products, such as the law of contract as observed in the *GYP Case*,⁹⁴ and depending on the circumstances, the law of confidence and use of technological protection measures.⁹⁵ Such discussions are, however, beyond the scope of this article. Nevertheless, like obsolete technologies, the case for granting monopolies in data and data products should be reappraised in the light of the realities and complexities of the contemporary data landscape.

94 *Global Yellow Pages Ltd v Promedia Directories Pte Ltd* [2017] 2 SLR 185 at [35] (citing the observations of *amicus curiae* Professor David Llewelyn).

95 For a discussion on these other modes of protection, see, for example, Tan Tee Jim SC, “New Law for Compilations and Databases in Singapore?” (2012) 24 SAclJ 745 at paras 113–123.