

PART FIVE

# ***Access***



# Copyright Reform and Fact-Based Works

Teresa Scassa\*

## A. INTRODUCTION

Information is a hot commodity in today's economy. In recent years, there has been a dramatic growth in the number of websites, databases, tools and applications that use data from a variety of public and private sources to offer innovative information-based services to a wide range of users. In some cases, the providers of these information tools are traditional disseminators of information. In others, they are established businesses that have developed new information tools and products. In many cases, the innovators are upstarts — small companies or individuals that see opportunities for new and useful applications. The dissemination of information and the development of new information tools are not limited to commercial enterprises. There is much free content from sources ranging from academic to purely amateur.

Property rights are an important element of control and are the foundation for financing, licensing, transacting and other activities within the marketplace. It is not surprising, therefore, that the issue of rights in

---

\* I am grateful to Ariel Katz for his comments and suggestions on a presentation of some of the content of this paper, to Elizabeth Judge for sharing her thoughts on these issues, and to Michael Geist and an anonymous reviewer for their insightful comments. I am also appreciative of the research assistance of Julien Bourgeois. The support of the Canada Research Chairs Program and the GEOIDE Network is gratefully acknowledged.

facts and information is re-emerging within this flourishing information landscape. Copyright law sets relatively clear boundaries with respect to the protection of works that are the vehicles for information. Text-based accounts are literary works, and the reproduction of all or a substantial part of such works is infringing. Videos, photographs and maps that are vehicles for information are likewise also works that fall within the traditional copyright categories. Yet the law remains uncertain when dealing with issues of rights in the facts or information that is represented in these works. Further, the greater the proportion of data in relation to the “work,” the greater the uncertainty regarding the scope of protection. Where the work is merely a compilation of data, the copyright protection is, in the words of the United States Supreme Court, “thin.”<sup>1</sup>

Bill C-32<sup>2</sup> does little to address the uncertainties surrounding copyright in facts and in fact-based works. This is perhaps not surprising for several reasons. The uncertainty as to scope is widespread. International copyright treaties exhibit ambivalence around the issue. For example, Article 5 of the *WIPO Copyright Treaty* provides:

Compilations of data or other material, in any form, which by reason of the selection or arrangement of their contents constitute intellectual creations, are protected as such. This protection does not extend to the data or the material itself *and is without prejudice to any copyright subsisting in the data* or material contained in the compilation (emphasis added).<sup>3</sup>

Similar language is used in the *Agreement on Trade-Related Aspects of Intellectual Property Rights* (TRIPs).<sup>4</sup> The language suggests that while copyright in a compilation will only be in the original selection or arrangement of the work, there may still be separate copyrights in the “data or material contained in the compilation.” In Europe, database protection was achieved through *sui generis* legislation specifically tailored to this type

- 
- 1 *Feist Publications, Inc. v. Rural Telephone Service Co.*, 499 U.S. 340, <http://caselaw.lp.findlaw.com/scripts/getcase.pl?court=US&vol=499&invol=340>, 111 S.Ct. 1282 at 349 (1991) [*Feist* cited to U.S.] .
  - 2 Bill C-32, An Act to amend the Copyright Act, 3d Sess., 40th Parl., 2010, [www2.parl.gc.ca/Sites/LOP/LEGISINFO/index.asp?Language=E&query=7026&Session=23&List=toc](http://www2.parl.gc.ca/Sites/LOP/LEGISINFO/index.asp?Language=E&query=7026&Session=23&List=toc) [Bill C-32].
  - 3 *WIPO Copyright Treaty*, 20 December 1996, 36 I.L.M. 65 at Art. 5 (entered into force 6 March 2002), [www.wipo.int/treaties/en/ip/wct/trtdocs\\_wo033.html](http://www.wipo.int/treaties/en/ip/wct/trtdocs_wo033.html), [WCT].
  - 4 *Agreement Establishing the World Trade Organization*, Annex 1 C: Agreement on Trade-Related Aspects of Intellectual Property Rights, 15 April 1994, 1869 U.N.T.S. 299 at Art. 10(2), [www.wto.org/english/tratop\\_e/trips\\_e/t\\_agmo\\_e.htm](http://www.wto.org/english/tratop_e/trips_e/t_agmo_e.htm).

of information asset.<sup>5</sup> Although a similar *sui generis* regime has been considered in the US and in Canada, there has been little appetite to legislate in this area.<sup>6</sup> The status quo, therefore, has represented the government approach to protecting data. The status quo, according to the case law, is that there is no copyright protection for facts themselves but copyright protection for original expressions of those facts. In our dynamic and evolving information economy, the *status quo* is showing signs of strain.

Bill C-32 represents a set of copyright reforms in the works since the last major reforms in 1997.<sup>7</sup> Aspects of the Bill are clearly designed to address Canada's international obligations with respect to copyright on the internet,<sup>8</sup> and the preamble to the Bill speaks of the need to foster innovation in a knowledge economy.<sup>9</sup> Indeed, copyright law is seen by the govern-

---

5 *Sui generis* legislation in the EU member states was required by virtue of EC, European Database Directive 96/9/EC of the European Parliament and of the Council of the European Union of 11 March 1996 on the legal protection of databases, [1996] O.J. L 77/20, <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31996L0009:EN:HTML> [European Database Directive].

6 In 1998, the federal government commissioned a study on the protection of facts and information in Canada. See Robert Howell, prepared for Industry Canada and Canadian Heritage, Database Protection and Canadian Laws (Ottawa: Industry Canada, 1998), <http://dsp-psd.pwgsc.gc.ca/Collection/C2-370-1998E.pdf>. No further steps were taken. In the US bills to create a *sui generis* regime for the protection of databases were introduced around the same time, and failed. See: US, Bill H.R. 3531, *Database Investment and Intellectual Property Antipiracy Act*, 104th Cong., 1996, <http://thomas.loc.gov/cgi-bin/query/z?c104:h.r.3531>; and US, Bill H.R. 2652, *Collection of Information Antipiracy Act*, 105th Cong., 1997, <http://thomas.loc.gov/cgi-bin/query/D?c105:2:/temp/~c1050UfJy1::>. See also U. Copyright Office, *Report on Legal Protection for Databases* (1997), <http://www.copyright.gov/reports/dbase.html>.

7 Copyright reform to modernize the *Copyright Act*, R.S.C. 1985, c. C-42, was conceived of as a three-phase process. The first phase was implemented in 1988. Phase II was initiated with the first Bill C-32, introduced in 1997. Reform has been stalled since that time. See: Canada, Heritage Canada, *A Framework for Copyright Reform* (2002), / [www.ic.gc.ca/eic/site/crp-prda.nsf/eng/rp01101.html](http://www.ic.gc.ca/eic/site/crp-prda.nsf/eng/rp01101.html).

8 Canada has signed, but has not yet implemented two key international treaties dealing with copyright and the internet. These are the *WIPO Copyright Treaty*, above note 3, and the *WIPO Performances and Phonograms Treaty*, 20 December 1996, 36 I.L.M. 76 (entered into force 20 May 2002), [www.wipo.int/treaties/en/ip/wppt/trt-docs\\_woo34.html](http://www.wipo.int/treaties/en/ip/wppt/trt-docs_woo34.html) [WPPT].

9 Bill C-32, above note 2, at Preamble. In particular, the Preamble identifies copyright law as "an important marketplace framework law," and acknowledges its impact on "many sectors of the knowledge economy." The Preamble concludes by noting: "Canada's ability to participate in a knowledge economy driven by innovation and network connectivity is fostered by encouraging the use of digital technologies for research and education."

ment as a pillar in Canada's digital economy strategy.<sup>10</sup> Rather than tackle the issue of the growing commercial importance of fact-based works, digital information tools and resources, Bill C-32 is silent in certain key areas. In addition, some measures in the Bill may pose a barrier to access and innovation in relation to fact-based works. In an innovation economy, clarity around the status and use of data in protected works is crucial.<sup>11</sup>

This chapter will provide an overview of the current state of the law in relation to the protection of fact-based works in Canadian copyright law. The analysis begins with a discussion of the current state of the law using contemporary examples. It next considers the fault lines appearing within the established doctrines. As this is a book about the current state of copyright reform in Canada, the broader question will be whether Bill C-32 does anything to change the settled law or expectations in relation to facts and fact-based works, and whether there are missed opportunities to address an issue crucial to the digital economy and information society.

## B. THE NEW WORLD OF FACT-BASED WORKS

The landmark decision in *Feist Publications, Inc. v. Rural Telephone Service Company, Inc.*<sup>12</sup> in the United States, and its Canadian counterpart, *Tele-Direct (Publications) Inc. v. American Business Information Inc.*,<sup>13</sup> came at a time when an increase in computing power and declining costs of computer memory had not only made fact-based compilations more commercially versatile and valuable, they were increasingly easy to reproduce and modify. Fast-forward almost twenty years, and computing power is dramatically greater, memory is far less expensive and the tools for reproduction and manipulation of data are more sophisticated. Perhaps the most significant change of all is that computing power, memory and tools are no longer the exclusive preserve of major corporate interests. The Web 2.0 revolution has put the power to harvest, control and manipulate

10 Canada, Industry Canada, *Improving Canada's Digital Advantage: Strategies for Sustainable Prosperity* (Ottawa: Public Works, 2010) at 14, [http://de-en.gc.ca/wp-content/uploads/2010/05/Consultation\\_Paper.pdf](http://de-en.gc.ca/wp-content/uploads/2010/05/Consultation_Paper.pdf).

11 See, for example: Christine Galbraith, "A Panoptic Approach to Information Policy: Utilizing a More Balanced Theory of Property in Order to Ensure the Existence of a Prodigious Public Domain" (2007) 15 J. Intell. Prop. L. 1, [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=1607959](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1607959).

12 *Feist*, above note 1.

13 *Tele-Direct (Publications) Inc. v. American Business Information Inc.* (1996), 74 C.P.R. (3d) 72 (F.C.T.D.), *aff'd* [1998] 2 F.C. 22, (1997), 76 C.P.R. (3d) 296 (C.A.), leave to appeal to S.C.C. denied, [1998] 1 S.C.R. xv [*Tele-Direct* cited to S.C.R.].

data — and even more importantly, to disseminate it — in the hands of ordinary individuals.<sup>14</sup> The result is an environment that will generate new conflicts over rights in data and in fact-based works.

An added element that is important to appreciating this new context is the shifting role of government with respect to key collections of data. Government has long been the primary generator of certain types of data. These include geospatial data, as well as data about natural and physical resources, aeronautical data, climate data and a vast range of demographic data. In addition, government is a source of vast stores of information about citizens' interactions with government at all levels. Governments around the world, including Canada, have sought to develop policies and infrastructures for managing and disseminating geospatial data, now recognized as a building block for research and innovation within a knowledge society.<sup>15</sup> At the same time, access to information processes are increasingly relied upon to gain access to a wealth of government information. In keeping with the nature of the Web 2.0 revolution, this information can be used and disseminated in a variety of ways by a wide range of actors.<sup>16</sup> Public records are also being mined for the information they contain, and form the basis for a vast array of commercial and non-commercial initiatives.<sup>17</sup> Yet governments are not the only source of data — data are also

---

14 Christopher C. Miller, "A Beast in the Field: The Google Maps Mashup as GIS/2" (2006) 41 *Cartographica* 187, <http://utpjournals.metapress.com/content/jolo53012262n779/fulltext.pdf>.

15 Canada's spatial data infrastructure (SDI) is spearheaded by Geoconnections: [www.geoconnections.org/en/index.html](http://www.geoconnections.org/en/index.html). It includes portals for free access to collections of government geospatial data. See Geobase, [www.geobase.ca](http://www.geobase.ca) (providing geospatial data from all levels of government); and GeoGratis, <http://geogratis.cgdi.gc.ca> (providing free geospatial data from the Earth Sciences Sector of Natural Resources Canada). In the United States, resources are made available through the National Spatial Data Infrastructure, [www.fgdc.gov/nsdi/nsdi.html](http://www.fgdc.gov/nsdi/nsdi.html). For a clearinghouse of international governmental geospatial data initiatives, see Global Spatial Data Infrastructure Association, [www.gsdi.org/SDILinks](http://www.gsdi.org/SDILinks).

16 The online version of the *Toronto Star* used to feature a weekly column, "Map of the Week," which plotted a wide variety of information about Toronto onto a map of the city. In many instances, the information came from access to information requests. See "Toronto Star Map of the Week" *Toronto Star* (10 July 2010), <http://thestar.blogs.com/maps/>.

17 For example, *USA Today* created a map of mortgage foreclosures in the city of Denver using information from public registries. The map allows users to navigate around a Google map of part of the city that shows in red all houses foreclosed upon, and provides the precise address and amount of money owing at the time of foreclosure; see Brad Heath & Ron Coddington, "Denver Foreclosures: One Hard Hit Neighborhood

generated by researchers and private enterprises.<sup>18</sup> Ordinary individuals are the sources of a wealth of sought-after data; personal information has become a major commodity.<sup>19</sup> Individuals can also be the source of other kinds of information. Increasingly, corporations that create information-based products are involving users in the “crowdsourcing” of this information or in its correction, revision and modification.<sup>20</sup>

Within this environment we have also witnessed a significant movement towards the democratization of information. Open source and open access movements seek to guarantee broad, open access to all manner of copyright-protected works.<sup>21</sup> Other movements have sought to decentral-

---

at a Glance” *USA Today* (2010), [www.usatoday.com/news/graphics/foreclosure\\_map/foreclose.htm](http://www.usatoday.com/news/graphics/foreclosure_map/foreclose.htm).

- 18 A recent dispute between Century 21 and Rogers, Inc. in Canada involves the scraping of real estate information from real estate agents’ web sites in order to populate Rogers’ Zoocasa house hunting site. The matter has not yet gone to trial. See Gary Marr, “Century 21 Canada does battle with Rogers” *Financial Post* (7 September 2009), [www.financialpost.com/story.html?id=1969611](http://www.financialpost.com/story.html?id=1969611).
- 19 Personal information is widely sought to create increasingly detailed consumer profiles for marketing and other purposes. See, for example, Perri 6, “The personal information economy: trends and prospects for consumers” in Susanne Lace, *The Glass Consumer: Life in a Surveillance Society* (Bristol: The Policy Press, 2005) at 17. Recently in the US the popular social networking site Facebook used copyright arguments to attempt to stop PowerVentures Inc. from using scraping technology to mine its site for personal information posted by its users. See *Facebook, Inc. v. Power Ventures, Inc.*, 2009 WL 1299698, [http://scholar.google.ca/scholar\\_case?case=196488893558356568&hl=en&as\\_sdt=2002&as\\_vis=1](http://scholar.google.ca/scholar_case?case=196488893558356568&hl=en&as_sdt=2002&as_vis=1), 91 U.S.P.Q. 2d 1430 (N.D.Cal.) [*Facebook, Inc.*].
- 20 Crowdsourcing is a term used to describe an open invitation to the public to contribute to the development of a product (including information based-products) or to the solving of a problem. For example, the news network Al Jazeera used crowd-sourced information to create a map titled “War on Gaza.” Users are invited to contribute information, photographs and opinions on events in Gaza. See <http://labs.aljazeera.net/warongaza>. Openstreetmap.org offers a venue for crowd-sourced mapping around the world, [openstreetmap.org](http://openstreetmap.org). Even Google is getting into crowd-sourcing. It now encourages users to edit, update and modify its Google Maps, [www.google.com/help/maps/edit](http://www.google.com/help/maps/edit).
- 21 The Creative Commons movement is a leading example. For Creative Commons in Canada, see, <http://creativecommons.ca>. In the realm of software, see, Open Source Initiative, [www.opensource.org](http://www.opensource.org). In 2009, the University of Ottawa became the first Canadian university to join the Compact for Open Publishing Equity, when it announced a new open access initiative to make scholarly research freely accessible to a wide audience: University of Ottawa, News Release, “University of Ottawa among North American leaders as it launches open access program” (8 December 2009), [www.media.uottawa.ca/mediaroom/news-details\\_1824.html](http://www.media.uottawa.ca/mediaroom/news-details_1824.html). See also, Open Access uOttawa, [www.oa.uottawa.ca/index.jsp?language=en](http://www.oa.uottawa.ca/index.jsp?language=en).



ize the production and dissemination of information.<sup>22</sup> In Canada, the pushback of ordinary Canadians against a hard-line creators' rights approach to copyright law resulted in an unprecedented set of grassroots consultations on copyright law,<sup>23</sup> which ultimately led to the current Bill C-32. It is fair to say that the public tolerance for laws that unduly limit access to content, or constrain uses that have traditionally been permitted or tolerated, is greatly diminished.

It is in this environment, then, that it becomes necessary to reconsider the state of copyright law in relation to facts and fact-based works.

### C. THE STATE OF COPYRIGHT LAW IN RELATION TO FACTS

It is a basic principle of copyright law that there is no copyright in facts, only in the original expression of facts.<sup>24</sup> Facts can be expressed in a number of different ways across a range of copyrighted "works." A photograph, for example, may be considered an original expression of the facts visible within it; the photograph itself is an artistic work.<sup>25</sup> Similarly, a map is essentially an original expression of certain geographical "facts," and maps are also artistic works.<sup>26</sup> A biography or an historical account would be examples of literary works that are expressions of fact, and a documentary film can also be an expression of facts in the form of a cinematographic work. In each of these examples, facts are central to the works, but the works themselves involve significant expressive activity. Extracting the bare facts from the expressive content is less troublesome because the core value of the work lies in its expression.

22 Wikipedia is a classic example of an attempt to crowd source information on a massive scale: [www.wikipedia.org](http://www.wikipedia.org). See also, Cass R. Sunstein, *Infotopia: How Many Minds Produce Knowledge* (Oxford: Oxford University Press, 2006). Of course, Wikipedia is not without its critics. See, e.g., Roger A. Longhorn & Michael Blakemore, *Geographic Information: Value, Pricing, Production, and Consumption* (Boca Raton, FL: CRC Press, 2008) at 86–87.

23 Industry Canada, News Release, "Government of Canada Launches National Consultations on Copyright Modernization" (20 July 2009), [/www.ic.gc.ca/eic/site/ic1.nsf/eng/04840.html](http://www.ic.gc.ca/eic/site/ic1.nsf/eng/04840.html) [Consultation]. The vast majority of submissions to this consultation process were made by individuals.

24 Miriam Bitton, "Feist, facts and functions: historical perspective", in Robert F. Brauneis, ed., *Intellectual Property Protection of Fact-based Works: Copyright and Its Alternatives* (Cheltenham, U.K.: Edward Elgar, 2009) 3 at 16.

25 *Copyright Act*, R.S.C. 1985, c. C-42, <http://laws.justice.gc.ca/en/C-42/index.html>, s. 2, definition of "artistic work."

26 *Ibid.*

More challenging issues arise when one is dealing with facts expressed in more basic ways — ways in which the factual content outweighs the expressive activity (and the core value of the work resides in the facts). For example, directories, lists, tables of data and databases are all works in which the factual content outweighs the expression. Such works are compilations of facts, and while compilations are considered to be copyright-protectable works,<sup>27</sup> the expression of a compilation lies essentially in the original selection or arrangement of the contents, and it is only this selection or arrangement that is protected.<sup>28</sup> In a factual compilation, it is the selection of the facts or their arrangement in the work that can be protected, but not the underlying facts.<sup>29</sup>

Not only must there be a selection or arrangement of facts to give rise to copyright protection in a compilation of facts, that selection or arrangement must be “original.” In Canada and the US, originality has been defined to mean something other than labour or investment. Thus, following the *Feist* decision in the US, originality is considered to be a “minimal level” or “spark” of creativity.<sup>30</sup> In Canada, the Supreme Court of Canada in *CCH Canadian Ltd. v. Law Society of Upper Canada*<sup>31</sup> ruled that originality results from an exercise of “skill and judgment.”<sup>32</sup> Both courts exclude labour or investment alone as bases for originality. Moreover, telephone directories have been found in both countries to lack originality — in spite of the fact that they are products of significant labour and investment — because

---

27 *Ibid.* at s. 2, definition of “every original literary, dramatic, musical and artistic work.”

28 *Robertson v. Thomson Corp.*, 2006 SCC 43, [www.canlii.org/en/ca/scc/doc/2006/2006scc43/2006scc43.html](http://www.canlii.org/en/ca/scc/doc/2006/2006scc43/2006scc43.html), [2006] 2 S.C.R. 363, at paras. 37–38 [*Robertson* cited to S.C.R.].

29 Per Abella J. *et al.*, dissenting in part in *Robertson*, *ibid.* at para. 100.; *Tele-Direct*, above note 13.

30 *Feist*, above note 1, at 358–59.

31 *CCH Canadian Ltd. v. Law Society of Upper Canada*, 2004 SCC 13, [www.canlii.org/en/ca/scc/doc/2004/2004scc13/2004scc13.html](http://www.canlii.org/en/ca/scc/doc/2004/2004scc13/2004scc13.html), [2004] 1 S.C.R. 339 [*CCH Canadian* cited to S.C.R.].

32 It has been suggested that the US and Canadian standards of originality are not significantly different. See, e.g., Teresa Scassa, “Recalibrating Copyright Law?: A Comment on the Supreme Court of Canada’s Decision in *CCH Canadian Ltd. v. Law Society of Upper Canada*” (2004) 3(2) C.J.L.T. 89, [http://cjlt.dal.ca/vol3\\_no2/pdftarticles/scassa.pdf](http://cjlt.dal.ca/vol3_no2/pdftarticles/scassa.pdf), and Daniel Gervais, “Canadian Copyright Law Post-*CCH*” (2004) 18 I.P.J. 131. Note as well that the High Court of Australia in *Telstra Corporation Limited v. Phone Directories Company Pty Ltd.*, [2010] FCA 44, [www.austlii.edu.au/au/cases/cth/FCA/2010/44.html](http://www.austlii.edu.au/au/cases/cth/FCA/2010/44.html) [*Telstra* #2], uses both the Canadian and US standards to refer to essentially the same threshold (at para. 344).

they are factual compilations which combine a “whole universe” selection (all subscribers who have not specifically asked to be excluded) with an obvious arrangement (alphabetical).<sup>33</sup> Neither the selection nor the arrangement is “original” in the sense required. Since the facts themselves cannot be copyrightable subject matter, and there is no originality in the expression of those facts, the works are not protectable by copyright.

In contrast, telephone directories in Australia were ruled protectable by copyright. The court’s decision defined originality to include the expenditure of significant labour or capital.<sup>34</sup> However, in *Telstra Corporation Limited v. Phone Directories Company Pty Ltd*,<sup>35</sup> a more recent case, the High Court of Australia was careful to distinguish the prior decision<sup>36</sup> by noting that in the earlier case the parties had conceded that the phone directories were authored. In their view, it was impossible to find authorship in the electronic telephone directory databases at issue before them.<sup>37</sup> The court ruled that “[a]uthorship and originality are correlatives,”<sup>38</sup> and found that it was necessary first to identify authors and then to assess their contributions to the work. In a highly automated process, it would be difficult to identify any particular “authorial” contribution to the selection or arrangement.

In Australia, as in Canada or the US, the data contained in a compilation of facts are theoretically not capable of protection under copyright law; protection depends upon an original selection or arrangement of the data. Once such an original selection or arrangement is found (and the threshold may not be particularly high),<sup>39</sup> then the work is protected by copyright. This means that the reproduction of the work as a whole or a

33 *Feist*, above note 1; *Tele-Direct*, above note 13.

34 *Telstra Corporation Ltd. v. Desktop Marketing Systems Pty Ltd.* [2001] FCA 612, [www.austlii.edu.au/cgi-bin/sinodisp/au/cases/cth/FCA/2001/612.html](http://www.austlii.edu.au/cgi-bin/sinodisp/au/cases/cth/FCA/2001/612.html) [*Telstra #1*].

35 *Telstra #2*, above, note 32.

36 *Ibid.* at paras. 52, 134, 157.

37 *Ibid.* at paras. 90–91, 333.

38 *Ibid.* at para. 344.

39 For example, in *Key Publications, Inc. v. Chinatown Today Publishing Enterprises, Inc.*, 945 F.2d 509 (2d Cir. 1991), <http://openjurist.org/945/f2d/509/key-publications-inc-v-chinatown-today-publishing-enterprises-inc>, the court found sufficient originality in the selection and arrangement of entries in a yellow pages directory of Chinese-American businesses. See generally, David E. Shipley, “Thin But Not Anorexic: Copyright Protection for Compilations and Other Fact Works” (2007) 15 J. Intell. Prop. L. 91, [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=1076789](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1076789). But note that in both *IceTV Pty Limited v. Nine Network Australia Pty Limited* [2009] HCA 14 [*IceTV*] and *Telstra #2*, above note 32, two Australian cases in the Web 2.0 era, the High Court found, in one case a lack of originality in the expression of the

substantial part of the work is not permissible without licence. Since the expressive content in a compilation of facts is the selection or arrangement of the data, the substantial reproduction that infringes copyright must be of the expressive content and not the underlying facts. Thus, the question becomes: what amounts to a substantial taking of the selection or arrangement of the facts? One approach is to ask how many facts, or which facts, amount to a substantial reproduction of the selection or arrangement expressed in the compilation. In *IceTV Pty Limited v. Nine Network Australia Pty Limited*,<sup>40</sup> by contrast, the court considered whether there was copying of the original authorial contribution, and since there was no originality in the expression of the facts themselves, there could be no substantial reproduction.<sup>41</sup>

Although the basic principles around copyright and fact-based works are fairly settled, there can be a great deal of divergence in how the principles are applied. Since the decisions in *Feist* in the US and *Tele-Direct* and *CCH Canadian* in Canada, there has been considerable uncertainty as to the scope of protection available for fact-based works in these countries — particularly regarding compilations of fact.<sup>42</sup> For example, in *B & S Publications Inc. v. Max-Contacts Inc.*,<sup>43</sup> the plaintiff complained that the defendant had taken and used the facts in its publication, the *Oil and Gas Index*, for the purposes of producing a competing publication. The plaintiff's index listed explorers and producers in the oil and gas industry in Alberta and supplemented these details with additional factual informa-

---

data and in the other, a lack of authorship and originality in the product of an automated data process.

40 *IceTV*, above, note 39.

41 *Ibid.* at para 42.

42 This uncertainty is evident in *GeoConnections*: Canada, Natural Resources Canada, The Dissemination of Government Geographic Data in Canada: Guide to Best Practices, Version 2 (Ottawa: Natural Resources Canada, 2008), [www.geoconnections.org/publications/Best\\_practices\\_guide/Guide\\_to\\_Best\\_Practices\\_Summer\\_2008\\_Final\\_EN.pdf](http://www.geoconnections.org/publications/Best_practices_guide/Guide_to_Best_Practices_Summer_2008_Final_EN.pdf). At page 58, after acknowledging that there is no copyright in data, the Guide refers to "copyrighted digital data." For a discussion of the difficulties with the licensing of data evidenced in this Guide, see Elizabeth F. Judge & Teresa Scassa, "Intellectual Property and the Licensing of Canadian Government Geospatial Data: An Examination of Geoconnections' Recommendations for Best Practices and Template Licences," (2010) 54:3 *Canadian Geographer*, available at [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=1567482](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1567482). See also William M. Landes & Richard A. Posner, *The Economic Structure of Intellectual Property Law* (Cambridge, MA: Belknap Press, 2003) at 104 (noting difficulty in protecting large electronic databases using the *Feist* analysis).

43 [2001] A.J. No. 143.

tion. Although Hutchison J. found that the design layout and format of the work was “original and unique,” and thus copyrightable, he also held that the defendant’s copying of the data violated the plaintiff’s copyright — not in the selection and arrangement of the data — but in the data itself. He noted:

I am of the opinion that the data relating to the exploration and production companies as researched and presented by the plaintiff is capable of and by itself of being copyrighted, owing in part to the criteria used to select the names by the plaintiff, the research done on its currency and accuracy, and its categorization.<sup>44</sup>

The decision suggests that some courts might view facts as protectable where they are considered sufficiently ‘original’ in their own right. This view admits that a fact might be the product of an exercise of skill and judgment and some facts are more than just representations of the world around us.

The uncertainty in the law around the protection of fact-based works has been met, from time to time, with industry calls for *sui generis* legislation, similar to what exists in the EU.<sup>45</sup> This would give separate protection for databases, since they in particular are most vulnerable under this approach.<sup>46</sup> In a realm of virtually unlimited computing power, whole universe sets of relevant data are very attractive but display little original selection. As it is the searcher who uses a search engine to extract the data relevant to them, it is also difficult to identify any original arrangement

---

44 *Ibid.* at para. 44. The case was decided prior to *CCH Canadian*, above note 30, and thus prior to the clear exclusion of labour alone as a basis for finding originality. Nevertheless, Hutchison J. does not appear to be basing his reasons on the sheer labour involved in compiling the data. What he describes arguably involves an exercise in skill and judgment in arriving at the data itself.

45 European Database Directive, above note 5.

46 It is by no means clear that *sui generis* legislation is the preferred approach. For a sample of perspectives on the pros and cons of such an approach, see Amy C. Sullivan, “When the Creative Is the Enemy of the True: Database Protection in the U.S. and Abroad” (2001) 29 A.I.P.L.A.Q.J. 317; Charles C. Huse, “Database Protection in Theory and Practice: Three Recent Cases” (2005) 20 Berkeley Tech. L.J. 23; Jacqueline Lipton, “Balancing Private Rights and Public Policies: Reconceptualizing Property in Databases” (2003) 18 Berkeley Tech. L.J. 773, [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=471885](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=471885); Yijun Tian, “Reform of Existing Database Legislation and Future Database Legislation Strategies: Towards a Better Balance in the Database Law” (2005) 31 Rutgers Computer & Tech. L.J. 347, [www.accessmylibrary.com/article-1G1-139431582/reform-existing-database-legislation.html](http://www.accessmylibrary.com/article-1G1-139431582/reform-existing-database-legislation.html).

within the overall compilation.<sup>47</sup> North American copyright law offers little real protection for these types of databases.<sup>48</sup> In general, the level of protection for compilations of fact is uncertain and unpredictable.<sup>49</sup>

#### D. WHY FACTS ARE NOT PROTECTED BY COPYRIGHT LAW

The rationale for not protecting facts in copyright law has typically been rooted in the view that facts are not original. As Justice O'Connor stated in *Feist*, facts "do not owe their origin to an act of authorship."<sup>50</sup> She describes census takers, for example, as "copying" facts from the world around them. Since facts are not original, no one can claim authorship in a fact, and any claim to authorship can only reside in the original expression of a fact. Where a fact is only capable of expression in a very limited number of ways, the doctrine of merger may also be relevant to prevent monopolies where the expression of the fact and the fact itself have "merged."<sup>51</sup>

---

47 Landes & Posner, above note 42 at 104. Of course, it can be argued that a searchable database is actually the result of a complex architecture and a great deal of skill and judgment. In *Telstra #2*, above note 32, however, this made it more difficult to establish copyright in the database as it was impossible to disentangle the myriad contributions to the database in order to identify actual authorship (at para 87).

48 In the United States, the "hot news" doctrine that originated in *International News Service v. Associated Press*, 248 U.S. 215, <http://supreme.justia.com/us/248/215/case.html>, 39 S.Ct. 68 (1918) has proven useful in protecting some fact-based works. The hot news doctrine is not copyright law — it emerges from unfair competition law more generally, and it applies in situations where someone has compiled information at some expenditure of capital, labour or time, and the information has a certain commercial value. A competitor who appropriates this information for commercial gain can be found liable for a species of misappropriation. Because this is a commercial tort and not a property right, the information is only protected for the period of time in which it has commercial value, and the doctrine only applies in the context of unfair competition. A non-commercial user of the information would not be restrained from using the same information. The hot news doctrine has not been precluded in Canada, nor has it been expressly adopted or applied.

49 Ian Masser, *Governments and Geographic Information* (London: Taylor & Francis, 1998) at 82; Shipley, above note 39.

50 *Feist*, above note 1 at 347.

51 The classic "merger" case in the US is *Baker v. Selden*, 101 U.S. 99, [http://scholar.google.ca/scholar\\_case?case=16308210976883953911&hl=en&as\\_sdt=2&as\\_vis=1&oi=scholar](http://scholar.google.ca/scholar_case?case=16308210976883953911&hl=en&as_sdt=2&as_vis=1&oi=scholar), 25 L.Ed. 841 (1880). Just as one cannot have a monopoly over a fact or idea, there is no monopoly in the expression of the fact or idea where there is only one way (or a very limited number of ways) to express the fact or idea. In such cases, there is said to be a "merger" between the fact/idea and its expression. In Canada, the merger doctrine has been considered and approved in *Delrina Corp. v. Triolet Systems, Inc.* (2002),

A better argument for not protecting facts is found in more recent Australian case law. In *IceTV*, for example, the High Court expressly indicated that public policy reasons justified not extending copyright protection to facts. The court stated that copyright “does not confer a monopoly on facts or information because to do so would impede the reading public’s access to and use of facts and information.”<sup>52</sup> This public policy-based rationale dictates a much stricter approach to copyright in fact-based works. In Australia, courts must now consider the extent of human authorship in compilations of fact, and will protect only original authorial expression in the selection or arrangement of the facts.

The problem with a rationale for not protecting facts that is not based on public policy, but that is founded instead on the character of facts as not original, is that it relies on a very particular concept of fact as a form of observable truth. Yet not all facts are equal, and some things considered to be facts may actually not be the result of simple observation, but of a significant exercise of skill and judgment. Thus, one of the fissures in this area of the law is the potential for what can be called “original” facts.<sup>53</sup> A lack of clear articulation of the law in relation to facts can lead to courts finding, as occurred in *B & S Publications*,<sup>54</sup> that certain facts are themselves original.

## E. “ORIGINAL” FACTS

An “original fact” is one which itself displays some of the attributes of authorship. Originality in the Canadian context, as noted above, requires an exercise of skill and judgement and that the work itself not be copied. There are many types of facts that might actually qualify as original under

---

58 O.R. (3d) 339, at para. 52, [www.canlii.org/en/on/onca/doc/2002/2002canlii11389/2002canlii11389.html](http://www.canlii.org/en/on/onca/doc/2002/2002canlii11389/2002canlii11389.html), 17 C.P.R. (4th) 289 (C.A.) where the court stated: “The merger notion is a natural corollary of the idea/expression distinction which, as I have said, is fundamental in copyright law in Canada, England and the United States. Clearly, if there is only one or a very limited number of ways to achieve a particular result in a computer program, to hold that that way or ways are protectable by copyright could give the copyright holder a monopoly on the idea or function itself.”

52 *IceTV*, above note 39 at para 28.

53 See Teresa Scassa, “Original Facts: Skill, Judgment and the Public Domain” (2006) 51 McGill L.J. 253, [http://lawjournal.mcgill.ca/documents/1224868339\\_Scassa.pdf](http://lawjournal.mcgill.ca/documents/1224868339_Scassa.pdf); Justin Hughes, “Created Facts and the Flawed Ontology of Copyright Law” (2007) 83 Notre Dame L. Rev. 43, [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=1012071](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1012071).

54 Above note 43.



this standard.<sup>55</sup> For example, much scientific fact is not simply copied from the world around us. A scientific fact may begin with a hypothesis. The scientist may develop a methodology to test this hypothesis, and may run many experiments as a result of this methodology. The results of the experiments may be expressed as data, but it is at least arguable that this data, too, is authored, as different tests following different methodologies may produce other results in other labs. The data is in part the product of an original conception and execution. Where tests repeatedly confirm certain outcomes, the hypothesis may be considered proven and may eventually come to represent scientific fact. Latour and Woolgar characterize the resultant scientific facts as “the set of statements considered too costly to modify [that] constitute what is referred to as reality.”<sup>56</sup> They go on to describe scientific activity as being “a fierce fight to construct reality.”<sup>57</sup> In the same vein, Justin Hughes writes about “social facts” as being facts that arise from “human agreement.”<sup>58</sup>

To offer another example, maps have long been considered to represent facts—the map itself is an expression of fact that is clearly protectable under copyright law, and there is much originality in the expression. Indeed, there is so much originality in the expression of the facts in a map that some of these facts bear only a general approximation to reality.<sup>59</sup> The art of mapmaking is not a precise reproduction of geographic reality. As Wood and Fels put it, “The map is not a picture. It is an argument.”<sup>60</sup> Notwithstanding this, while courts have treated maps as artistic works as a whole, they nevertheless consider them to be representations of fact. In *R. v. Allen*, for example, the court stated:

- 
- 55 Green describes the idea that facts are not copyrightable because they are not independently created as a “fallacy”: Michael Steven Green, “Two fallacies about copyrighting factual compilations” in Robert F. Brauneis, ed., *Intellectual Property Protection of Fact-based Works: Copyright and Its Alternatives* (Cheltenham, UK: Edward Elgar, 2009) 109 at 109–10.
- 56 Bruno Latour & Steve Woolgar, *Laboratory Life: The Construction of Scientific Facts* (Princeton, N.J.: Princeton University Press, 1986) at 243.
- 57 *Ibid.*
- 58 Huse, above note 46 at 59.
- 59 Mark Monmonier, *How to Lie with Maps*, 2d ed. (Chicago: University of Chicago Press, 1996). Monmonier observes at 1: “To portray meaningful relationships for a complex, three-dimensional world on a flat sheet of paper or a video screen, a map must distort reality.”
- 60 Denis Wood & John Fels, *The Natures of Maps: Cartographic Constructions of the Natural World* (Chicago: University of Chicago Press, 2008) at xvi.



In the world of map making, roads exist. Drawing a road on paper to show where that road exists in relation to other roads cannot create a subsisting copyright. It is the manner of compiling and the way that information is presented that creates originality and artistry that qualifies as a subsisting copyright.<sup>61</sup>

This statement reflects the general difficulties in negotiating facts and their expression in copyright law.<sup>62</sup> While the roads represented in the map do exist, and their existence or location cannot be the subject matter of copyright, the statement that “[d]rawing a road on paper to show where that road exists in relation to other roads cannot create a subsisting copyright” conflates the existence of the road with its expression in a map. The representation of a road on a map is nothing other than the expression of that “fact,” and a different mapmaker might make different choices in expressing its location and details. Thus, while the “fact” may be the road itself, the line on a map is not a fact, but rather an expression of the fact. The longer or more winding the road, the less likely there will be merger between the fact and its expression.

Perhaps the most extreme example of original facts arises in the context of popular culture, where the popularity of certain works leads to products that explore their internal “facts.” In *Castle Rock Entertainment v. Carol Publishing Group*,<sup>63</sup> the owners of the rights in the popular television series *Seinfeld* successfully sued the publisher of *The SAT* (*Seinfeld* Aptitude Test). *The SAT* consisted of a series of trivia questions based on the characters and events in the *Seinfeld* series. The Second Circuit Court of Appeals agreed with the court below that the creators’ copyright was infringed and rejected arguments that the trivia questions merely reproduced facts derived from the series. Citing *Feist*, the Court stated: “Unlike the facts in a phone book, which ‘do not owe their origin to an act of authorship’ . . . , each ‘fact’ tested by *The SAT* is in reality fictitious expression created by *Seinfeld*’s authors.”<sup>64</sup> The court went on to distinguish between “true” facts (for example, the identity of the actors who play the characters) and fictionalized facts (those drawn from events in the series).

---

61 *R. v. Allen*, 2006 ABPC 11 at para. 11 5, [www.canlii.org/en/ab/abpc/doc/2006/2006abpc115/2006abpc115.html](http://www.canlii.org/en/ab/abpc/doc/2006/2006abpc115/2006abpc115.html), 399 A.R. 245.

62 See Green, above note 55 at 110–14.

63 *Castle Rock Entertainment, Inc. v. Carol Publishing Group, Inc.*, 150 F.3d 132, [www.law.cornell.edu/copyright/cases/150\\_F3d\\_132.htm](http://www.law.cornell.edu/copyright/cases/150_F3d_132.htm) (2d Cir. 1998).

64 *Ibid.* at 139.

In *Warner Bros. Entertainment Inc. v. RDR Books*,<sup>65</sup> a case which followed *Castle Rock*, the court took a similar view of a Lexicon based on the series of Harry Potter novels. The court stated: “Even if expression is or can be used in its ‘factual capacity,’ it does not follow that expression thereby takes on the status of fact and loses its copyrightability.”<sup>66</sup> In these examples, fact becomes a relative term, perhaps more acutely so because of the iconic cultural status of the works in question.

The fictional “fact” cases represent the far end of a spectrum of facts ranging from the wholly observed to the wholly authored. The difficulty in applying copyright doctrines in this area is most acute with respect to those facts that occupy the middle range of the spectrum. Illustrations can be drawn from contemporary disputes around rights in “facts.”

## F. THE FACTS ABOUT PUBLIC TRANSIT

A flurry of disputes has recently erupted in the United States over public transit data. These disputes raise interesting questions about the nature of “facts” and their protection under copyright law. In some cases, the disputes have arisen in contexts where public transit authorities have contracted with a company called NextBus.<sup>67</sup> NextBus uses proprietary algorithms to crunch data from transit timetables with data harvested from GPS systems installed on buses and that communicate with electronic readers at set points along routes. The result is prediction data—predictions about when the next bus is likely to arrive at any given stop. The information is made available to transit riders, and the goal is to improve their public transit experience. In a number of American cities that have launched NextBus services, entrepreneurial individuals have created iPhone applications which harvest and make the prediction data accessible to iPhone users in a format more convenient to them than the web-based NextBus interface. NextBus claims copyright in its prediction data;<sup>68</sup> the

---

65 *Warner Bros. Entertainment Inc. v. RDR Books*, 575 F. Supp. 2d 513, [http://scholar.google.ca/scholar\\_case?case=13852164224811081270&hl=en&as\\_sdt=2&as\\_vis=1&oi=scholar](http://scholar.google.ca/scholar_case?case=13852164224811081270&hl=en&as_sdt=2&as_vis=1&oi=scholar), (S.D.N.Y. 2008).

66 *Ibid.* at 536.

67 NextBus, [www.nextbus.com](http://www.nextbus.com). NextBus is a subsidiary of Grey Island Systems, Inc., [www.interfleet.com/index.asp](http://www.interfleet.com/index.asp).

68 The NextBus licence agreement contains the following: “Nextbus predictions and other information are copyrighted. You agree not to resell our service or use your account access to provide data from our service to any other user or to publish the data in any way.” NextBus Mobile Terms and Conditions, <http://nextbusmobile.com/s/terms.webui>.

application developers insist that the information is “fact” and in the public domain.<sup>69</sup> At least one transit company has muddied the waters by asserting rights in the underlying transit timetable data.<sup>70</sup> The disputes that have arisen over the rights in the underlying data offer a good illustration not only of the uncertainties of copyright law in this area, but also of the dynamic and likely contentious area into which we are moving regarding data, innovation and copyright.

There are two main sets of data that are relevant in the transit data disputes. The first is the transit timetable data. In its printed form, it is easy to say that the route, stop and time information arranged into a coherent timetable is a work in which copyright subsists. The situation is less clear if that data is stripped from the schedules offered by the transit company and made available through an iPhone app, as has already occurred in many North American cities. In such circumstances, some transit authorities have responded by asserting rights in the schedule data and issuing takedown notices.<sup>71</sup> A quick glance at licence agreements suggests that some transit authorities assert rights in their data. For example, although the San Francisco Municipal Transportation Agency (SFMTA) has chosen to provide free public access to their transit data, they do so by way

---

69 Anthony Ha, “Apple kills Routesy app, my iPhone gets less useful” *Social Beat* (27 June 2009), <http://social.venturebeat.com/2009/06/27/apple-kills-routesy-app-my-iphone-gets-less-useful>; Rafe Needleman, “Who Owns Transit Data” *CNet News* (24 August 2009), [http://news.cnet.com/8301-19882\\_3-10315749-250.html](http://news.cnet.com/8301-19882_3-10315749-250.html).

70 This was the case in New York City. See Dan Oshiro, “NY Transit Authority Cites Schedules as Copyrighted Material” *ReadWriteWeb* (20 August 2009), [www.readwriteweb.com/archives/ny\\_transportation\\_authority\\_cites\\_schedules\\_as\\_cop.php](http://www.readwriteweb.com/archives/ny_transportation_authority_cites_schedules_as_cop.php). In Washington DC, the Washington Metropolitan Area Transit Authority (WMATA), in response to strong demand, made its transit data publicly available. See Washington Metropolitan Area Transit Authority, News Release, “Metro Makes Schedule and Route Data Available Via Web Site” (20 March 2009), [www.wmata.com/about\\_metro/news/PressReleaseDetail.cfm?ReleaseID=2506](http://www.wmata.com/about_metro/news/PressReleaseDetail.cfm?ReleaseID=2506). However, it also commissioned a study to take “a look at intellectual property, such as schedule data to determine whether there are revenue opportunities in the future.” *Ibid.*

71 In New York City, for example, a software developer created an iPhone application called StationStops that provided commuters with schedule information for New York’s largest commuter rail system. In the summer of 2009, the New York Metro Transportation Authority (NYMTA) contacted the developer claiming ownership of the information in the schedule and demanding payment of a \$5000 license fee. When the developer refused to pay, the NYMTA sent a takedown notice to Apple requesting that the application be banned from iPhone. Apple complied with the takedown notice. The dispute has since been resolved. See Bryan Chaffin, “StationStop Gets Ticket to Ride Again” *The MacObserver* (8 October 2009), [www.macobserver.com/tmo/article/stationstop\\_gets\\_ticket\\_to\\_ride\\_again](http://www.macobserver.com/tmo/article/stationstop_gets_ticket_to_ride_again).

of a licence agreement, which is premised on underlying rights in the data. The licence states that the SFMTA “retains full title and ownership and all rights and interest in the Data.”<sup>72</sup> A similar term is present in the Edmonton Transit Licence Agreement. It reads: “Edmonton Transit retains all right, title, and interest in the Data, and any intellectual property rights embodied in the Data, including any copyright.”<sup>73</sup> The contract terms are puzzling, since if the data are “facts” they are in the public domain. Yet it can be argued that transit data actually represent not observed fact, but rather authored “fact.” If bus #102 is expected to depart stop #5 at 6:05 a.m., this is because the transit authority has developed their schedule accordingly, and they retain the discretion to change or adjust the times and routes. Indeed, the timetable as a whole is like a careful choreography, with buses scheduled so as to provide the necessary levels of service and the desired connections, depending on the peaks and ebbs of demand throughout the day. Bus arrival and departure times are less “facts” than they are points in the choreography.

If it can be argued that transit timetable data are authored facts, what then can be said of prediction data? The prediction data used by NextBus are generated by algorithms that combine timetable data with GPS data to produce arrival time predictions. On the one hand, one could argue that the results are not “facts” and therefore may have a sufficient degree of authorship to be protected by copyright law. Another company writing its own algorithms might arrive at different predictions, and of course, predictions are merely well-informed guesses as to likely outcomes. Yet counter-arguments are also possible. Prediction data of this kind are not dissimilar to some forms of scientific fact. Indeed, one could argue that predictions are closer to observed fact than transit timetable data. While timetable data must originate with the authors of the timetable, prediction data are based on known and observed information. The issue is not resolved.<sup>74</sup>

---

72 SFMTA Transit Data Licence Agreement and Download, updated to December 17, 2009, [www.sfmta.com/cms/asite/transitdata.htm](http://www.sfmta.com/cms/asite/transitdata.htm).

73 Edmonton Transit’s Google Transit Feed Specification Data Terms of Use, [www.edmonton.ca/transportation/ets/ets-data-for-developers.aspx](http://www.edmonton.ca/transportation/ets/ets-data-for-developers.aspx). The Edmonton Transit System makes its transit data available under licence from its website to encourage the private sector development of useful applications for transit users. This data is the basic route and arrival time information. The data is licensed at no cost for non-commercial users, but the ETS retains the right to charge a licence fee to the developers of commercial applications.

74 Note that the Weather Network asserts copyright in the “information” available on its site: “Without the expressed written consent of PMI, no information or material

## G. FACTS AND INFORMATION: THE ROOT OF THE PROBLEM

It is possible that the root of the copyright problem lies in the distinction between facts and information. This distinction is one that is well-recognized in the geospatial information field. Geospatial data is typically defined as the raw data recording geographical points, such as coordinates represented by longitude and latitude. These data are classic “observed” facts. Information, by contrast, consists of facts placed in some form of context. It is this contextualization of fact that results in *information* and that may consist of the elusive element of authorship. While there may be no copyright in facts, there will be copyright in information if the element of authorship is readily apparent. The authorship in a map or a biography is evident enough. In a collection or compilation of data, it is perhaps more difficult to discern. But data expressed in a compilation is still contextualized in some way—the compilation tells us something about something—and in this sense it is information.

A classic rationale for not protecting facts (i.e., data) in copyright law is that they are not authored. Information as contextualized fact, by contrast, is the result of authorship. Information can therefore, at least theoretically, be protected by copyright law. Yet not all information will be protected. The rationale for not protecting some forms of information is different from the rationale for not protecting facts. Where information—contextualized fact—is not protected, it is most likely because there has been a merger of expression and fact. In other words, the information conveys the underlying facts in such a way that it is difficult or impossible to separate out the fact from its expression.

The challenge of the information society is to recognize the extent to which facts are constantly being transformed into information, to recognize the difficulties in separating the information from the underlying fact, and to decide what to do about recognizing, protecting and rewarding the authorship of information where warranted. This challenge will only become more difficult with time. The disputes in the US over bus timetables and prediction data offer just one illustration of some of the problems and conflicts that can arise in a context where marked technological advances have blurred the distinctions between users and creators,

---

from this Web site may be copied, reproduced, republished, uploaded, posted, transmitted, or distributed in any way whatsoever. . .”; The Weather Network, [www.theweathernetwork.com/copyright](http://www.theweathernetwork.com/copyright).

created new — even if tiny and highly specialized — markets for innovative information products, and rapidly expanded the ways in which data can be used and disseminated.

The transit data context offers a neat encapsulation of some of the issues and tensions in this area. Data compiled by a public body are of relevance to innovators in the private sector who are both traditional and non-traditional players. The public has an interest in useful and innovative data-based products, and in an open and competitive marketplace. The public body has certain responsibilities to the public as well as its own needs to recover or reduce costs. At the centre of it all are sets of data and the ability to control or limit access to them through the vehicle of copyright law. The transit data disputes highlight the unanswered questions in copyright law regarding facts and information — questions that are likely to increase in importance over time. The concern is that this level of uncertainty, and the disparity in power between established and upstart innovators or users, may lead to a brake on innovation around fact-based works and applications.

## H. BILL C-32 AND FACT-BASED WORKS

Bill C-32 offers nothing that is particularly addressed to the protection or use of fact-based works. This is not entirely surprising, as the copyright reform process had not identified this area as one needing attention. Nevertheless, the legislative choices reflected in Bill C-32 will inevitably have an impact on such works, and this section will consider some of these effects. The most important of these effects have to do with access to and use of factual content. As noted above, facts fall outside the scope of copyright for public policy reasons. In the words of the Australian High Court, copyright grants no monopoly on facts “because to do so would impede the reading public’s access to and use of information.”<sup>75</sup> To the extent, then, that Bill C-32 limits access to and use of facts and information, it is highly problematic.

Perhaps the most important impact of Bill C-32 on fact-based works may come from the proposed amendments to the *Copyright Act* that would make it actionable to circumvent technological protection measures. It was inevitable that the Bill should address this issue, as one of the driving forces for copyright reform was the need to implement the WIPO treat-

---

75 *IceTV*, above note 39 at para 42.

ies.<sup>76</sup> The WCT requires that anti-circumvention measures be present in the copyright legislation of ratifying states.<sup>77</sup> What was not inevitable was the form that such measures would take. An earlier Canadian copyright bill,<sup>78</sup> which died on the order paper, had made anti-circumvention actionable only when it was done for purposes that would infringe copyright.

Bill C-32 takes a different approach: circumvention of a technological protection measure that controls access to a work is actionable, regardless of the purpose of the circumvention.<sup>79</sup> The only exceptions to this general rule are limited, and include law enforcement and national security,<sup>80</sup> encryption research,<sup>81</sup> the need to make computer programs interoperable,<sup>82</sup> providing access to persons with disabilities,<sup>83</sup> and protecting personal information.<sup>84</sup> These exceptions operate only within limited parameters. Circumvention to gain access to data not protected by copyright — data in the public domain — would not be permissible if that data were expressed in a compilation that was itself a work, even if all that was protectable by copyright law was the original selection or arrangement of the data and not the underlying facts themselves. The “thin” copyright for factual compilations — a protection that is thin for legitimate public policy reasons — is rendered “thick” by the addition of anti-circumvention technology. It has already been pointed out that anti-circumvention measures may prevent legitimate fair dealing with protected works.<sup>85</sup> In the case of fact-based works, the problem is amplified where the underlying data is not itself

---

76 WCT, above note 3; WPPT, above note 8.

77 WCT, *ibid.* at Art. 11.

78 Bill C-60, *An Act to amend the Copyright Act*, 1st Sess., 38th Parl., 2004–2005, s. 34.02, [www.parl.gc.ca/PDF/38/1/parlbus/chambus/house/bills/government/C-60\\_1.pdf](http://www.parl.gc.ca/PDF/38/1/parlbus/chambus/house/bills/government/C-60_1.pdf). For a discussion of the anti-circumvention features of this bill, see Michael Geist, “Anti-circumvention Legislation and Competition Policy” in M. Geist, ed., *In the Public Interest: The Future of Canadian Copyright Law* (Toronto: Irwin Law, 2005) at 211.

79 Bill C-32, above note 2, s. 41.1.

80 *Ibid.*, s. 41.11.

81 *Ibid.*, s. 41.13.

82 *Ibid.*, s. 41.12.

83 *Ibid.*, s. 41.16.

84 *Ibid.*, s. 41.14.

85 Ian Kerr, Alana Maurushat, & Christian S. Tacit, “Technological Protection Measures: Tilting at the Copyright Windmill” (2002–2003) 34 *Ottawa L. Rev.* 6, [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=793504](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=793504) at 47–48.



copyrightable. This could serve as a significant brake on innovation and competition,<sup>86</sup> but also on the free exchange of ideas and information.

Although the anti-circumvention provisions in Bill C-32 will have implications for fact-based works, there are other issues that are also important. Crown copyright,<sup>87</sup> which provides for ownership by the Crown of works created under its direction, remains highly problematic. Because some of the most important and useful collections of data about Canada — its territory, resources, and citizens — are created by government departments or agencies, claims to copyright in these collections of data can pose a barrier to access.

Bill C-32 does not address Crown copyright. Indeed, Crown copyright was not on the table, nor was it addressed in previous rounds of copyright reform. This is notwithstanding the fact that Crown copyright was raised as an issue in the submissions of a number of participants in the consultation process.<sup>88</sup> Crown copyright is an area where there seems to be no government appetite for reform — perhaps not surprisingly. Crown copyright serves government both by permitting it to develop revenue streams based on the sale of its works,<sup>89</sup> and by permitting it to exercise control over the dissemination of its works.<sup>90</sup> This element of control is perhaps most significant in the context of access to information. While copyright is not typically a barrier to an access request for documents in the hands of government, Crown copyright can be used to prevent the publication

---

86 See Geist, above note 78 at 235, discussing the impact of a similar provision in the US *Digital Millennium Copyright Act*, Pub. L. No. 105-304, 112 Stat. 2860 (1998), [www.copyright.gov/title17/](http://www.copyright.gov/title17/).

87 *Copyright Act*, above note 7, s. 12.

88 Eight organizations representing users of copyright works raised Crown copyright as an issue, as did four non-profit organizations. The issue was also raised in the submissions of numerous individuals. See *Consultation*, above note 23.

89 Stanbury argues that this rationale for Crown copyright is inappropriate as the Crown requires no economic incentive to produce the materials it does: W.T. Stanbury, "Aspects of Public Policy Regarding Crown Copyright in the Digital Age" (1996) 10 I.P.J. 131 at 138, <http://www.lexum.com/conf/dac/en/stanbury/stanbury.html>. Longhorn & Blakemore, above note 22 at 76–78, offer a critique of government's role in commercializing geographic information.

90 For a detailed discussion of Crown Copyright, see Elizabeth F. Judge, "Crown Copyright and Copyright Reform in Canada", in M. Geist, ed., *In the Public Interest: The Future of Canadian Copyright Law* (Toronto: Irwin Law, 2005) at 551. Although control is sometimes asserted as necessary to ensure quality, it does not necessarily serve this purpose. See, e.g., David Vaver, "Copyright and the State in Canada and the United States" (1996) 10 I.P.J. 187 at 200, [www.lexum.com/conf/dac/en/vaver/vaver.html](http://www.lexum.com/conf/dac/en/vaver/vaver.html), and Judge & Scassa, above note 42.



and further dissemination of these documents.<sup>91</sup> Even outside the access to information context, there are numerous examples of governments asserting copyright in fact-based works, as well as in the underlying data.<sup>92</sup> The control exercised by government — or the potential for the exercise of this control — might pose a significant threat to the free flow of information and be a barrier to innovation.<sup>93</sup> While the federal government and some provincial governments are moving towards the dissemination of some government data under relatively open licenses, these licences are vulnerable to changes in policy direction.<sup>94</sup> In Bill C-32, Crown copyright remains an unchanged feature of our copyright law; one that is of real significance in the area of fact-based works.

It was inevitable that Bill C-32 would deal with the issue of the liability of internet service providers (ISPs) or providers of information location

91 See Judge, above note 90 at 572; Teresa Scassa, "Table Scraps or a Full Course Meal? The Public Domain in Canadian Copyright Law", in *Intellectual Property at the Edge: New Approaches to IP in a Transsystemic World, Proceedings of the Meredith Lectures*, (Cowansville, Québec: Editions Yvon Blais, 2007) 347 at 369.

92 See Statistics Canada, "Copyright/Permission to Reproduce" (18 January 2010), [www.statcan.gc.ca/reference/copyright-droit-auteur-eng.htm](http://www.statcan.gc.ca/reference/copyright-droit-auteur-eng.htm); Geoconnections, above note 42.

93 See, for example, Jacques Frémont, "Normative State Information, Democracy and Crown Copyright" (1996) 11 I.P.J. 19 at 31. In the UK, where data is protected under a *sui generis* database protection regime, the Royal Mail recently faced criticism over its refusal to make postal code data freely available. The Royal Mail derives revenues of £1.3 million a year from licensing this data, but the cost of licences is prohibitive for some who would otherwise use the data in a variety of innovative applications. See, for example, the notice of shutdown by JobCentre ProPlus: [www.jobcentre-proplus.com](http://www.jobcentre-proplus.com), and Ernest Marples Blog, "Ordnance Survey to release postcode data?" (10 December 2009), <http://ernestmarples.com/blog>; and Open Rights Group, [www.openrightsgroup.org/press/press-releases/royal-mail-stop-job-search](http://www.openrightsgroup.org/press/press-releases/royal-mail-stop-job-search).

94 For examples of licence templates for unrestricted use of federal government data, see Geoconnections, above note 42. For a municipal government example, consider the City of Toronto's open data initiative, [www.toronto.ca/open/index.htm](http://www.toronto.ca/open/index.htm). Licence terms and conditions can be found at: [www.toronto.ca/open/terms.htm](http://www.toronto.ca/open/terms.htm). For a discussion of the vulnerability of licensed information to a change in government policy, see Teresa Scassa, "The Best Things in Law are Free: Towards Quality Free Public Access to Primary Legal Materials in Canada" (2000) 23 Dal. L.J. 301. Of course, one can argue that control can be used to ensure the quality and reliability of downstream uses of data and can even ensure that the downstream uses are not exclusive or proprietary. For example, the City of Toronto's open data licence requires licensees to acknowledge that they acquire no proprietary interests in the licensed data or data sets. See City of Toronto, "Terms of Use for our Datasets," [www.toronto.ca/open/terms.htm#licence](http://www.toronto.ca/open/terms.htm#licence). Of course, this does not avoid the problem of the scope or impact of claims to copyright in an original selection or arrangement of the source data.

tools. In the US, a “notice and takedown” approach has been adopted, requiring ISPs to remove postings or files when they are given notice that the material infringes copyright. Notice and takedown has been criticized as posing risks to freedom of expression.<sup>95</sup> It has also raised concerns about eliminating fair use of copyright protected works. If a mere claim of infringement is enough to force the takedown of a posted work, notice and takedown can significantly chill the dissemination of works that are legitimate and non-infringing. With fact-based works, where there are major issues about the subsistence and scope of copyright, notice and takedown can be used to squelch competition and to stifle the dissemination of information. In the context of the transit data example above, notice and takedown measures were relied upon by NextBus or transit authorities to stop the sale of iPhone applications that made use of transit and prediction data, even though the application creators maintained they had taken only facts in the public domain.

Bill C-32 reflects a different approach to the problem of infringing materials on the internet by adopting a “notice and notice”<sup>96</sup> system that will apply to ISPs or providers of information location tools. Under notice and notice, a copyright owner who believes that their rights in a work are being infringed, where those works are posted or communicated over the internet, may give notice of claimed infringement to the relevant ISP or information location tool provider. The notice must be sufficiently detailed to identify the work in question, the identity of the rights holder, and the interest of the rights holder in the work.<sup>97</sup> The notice must also give details of the claimed infringement. An ISP in receipt of such a notice must forward it to the party who has allegedly committed the infringing acts and must then keep records that will permit the identity of the person to whom the relevant IP address belongs to be determined.<sup>98</sup> A rights holder who then decides to take legal action against the alleged infringer will be able to use legal means to discover this identity. While notice and notice is generally preferable to notice and takedown, in the context of fact-based works, it is particularly important as the copyright claims in such works are often uncertain, tenuous or very limited in scope. Notice and notice favours access to fact-based works by ensuring that they cannot so easily

---

95 Sheryl N. Hamilton, “Made in Canada: A Unique Approach to Internet Service Provider Liability and Copyright Infringement”, in M. Geist, ed., *In the Public Interest: The Future of Canadian Copyright Law* (Toronto: Irwin Law, 2005) 285 at 300.

96 Bill C-32, above note 2, ss. 41.25-41.27.

97 *Ibid.*, s. 41.25(2).

98 *Ibid.*, s. 41.26.

be forced from circulation by parties asserting uncertain copyrights. By adopting notice and notice, Bill C-32 reflects an important element of balance between rights holders and users/innovators.

The exception for non-commercial user-generated content in Bill C-32 is also important in the context of fact-based works. This is because many users mine fact-based works to create non-commercial works that rework these facts, either on their own, or in a mashup with other facts.<sup>99</sup> If Bill C-32 is passed into law, a new section 29.21 would permit individuals to “use an existing work or other subject-matter . . . which has been published or otherwise made available to the public, in the creation of a new work or other subject-matter in which copyright subsists.” The exception would also extend to the dissemination of such a work as long as the dissemination is solely for non-commercial purposes, the source and author of the existing work are mentioned, and where the use of or dissemination of the new work “does not have a substantial adverse effect, financial or otherwise, on the exploitation or potential exploitation of the existing work.”<sup>100</sup> This includes circumstances where the new work cannot be a substitute for the existing one.<sup>101</sup>

The exception is not without its limits. If one were to apply this provision to the transit data example, iPhone application developers who used NextBus data in their iPhone apps would be unable to rely on the user-generated content exception for two reasons: their works are commercial (this alone clearly takes the use out of the scope of the exception); and their apps function essentially as a substitute for the NextBus service. Of course, this analysis presumes that any claim to copyright in the data can be substantiated—a matter that is still an open issue. However, the user-generated content exception might be useful if information were extracted from an information-based product or service, and presented in a non-commercial context. Thus, even where copyright is asserted (with or without justification) in compilations of data, the user-generated content exception could provide a level of security to users who mine the data in order to create free, publicly accessible, non-commercial works. The potential benefits to the

---

99 To illustrate, the website “ProgrammableWeb” offers a listing of mashups created using the Google Maps Application Programming Interface (API). Mashups are created by taking the Google Maps interface and combining it with data from other sources. The list of mashups is enormous and represents an extremely wide variety of information. See [www.programmableweb.com/api/google-maps/mashups](http://www.programmableweb.com/api/google-maps/mashups) [*Programmable Web*].

100 *Bill C-32*, above note 2, s. 29.21(1)(d).

101 *Ibid.*

public in the area of fact-based works are significant, particularly because there appears to be a fairly active group of individuals willing to compile and disseminate information at no cost to the public.<sup>102</sup> Of course, licence terms and anti-circumvention measures could both significantly limit the scope of the user-generated content exception.<sup>103</sup> In addition, the exception would not apply if the user-generated works are considered to meet the rather open-ended criterion of having “a substantial adverse effect” on the actual or potential exploitation of the source work.

## I. CONCLUSION

Although data and fact-based works are surging in importance in the digital economy, and although new issues are raised by the tools now in the hands of creators—large and small—to fashion new information-based applications, the state of copyright law in relation to facts remains stagnant. For some, the uncertainty over the scope of protection for fact-based works may be a barrier to investing in innovation in this area; for others, the uncertainty over potential liability for using facts contained in other compilations or data sets may also be a barrier to innovation. Although the existing copyright principles can be used to resolve cases as they come before the courts, the principles are such that it will be difficult to give real shape to the law in this area. The uncertain distinction between fact and expression (or fact and information), and the fact-specific nature of any inquiry into originality of a selection or arrangement, or substantial taking from this selection or arrangement, give little hope of clear guidance in this area. Bill C-32 offers no solutions to the core issues. Reforms, if they are to come, will have to wait until some future date.

Government is a major producer of data and a prime data source for many private sector and research activities. The same uncertain rules regarding copyright in facts play out where the Crown is the owner of copyright, with the added implications of the Crown being in a position to censor, restrict or otherwise control public data. Crown copyright thus raises public policy issues in relation to ownership and control of data and its impact on innovation, as well as in relation to ownership and control and their impact on democracy, accountability and freedom of expression.

---

102 For examples, see *Programmable Web*, above note 99.

103 Note that Statistics Canada, which asserts copyright in its data, currently uses a licence to create its own non-commercial user-generated content exception: Statistics Canada, above note 92.

Crown copyright remains far from the legislative agenda, even though it was an issue frequently raised in the consultation process.

Although Bill C-32 does not address fact-based works specifically, a number of its provisions will be relevant to the use and exploitation of such works. Anti-circumvention provisions will effectively extend the copyright protection of a selection or arrangement of data to the data themselves when there is a layer of technology around the compilation restricting access and use. While some might view this as a necessary means of protecting investments in valuable sets of data, it is a clumsy and inadequate tool for addressing the necessary balance between exclusive rights and access, between ownership and the public domain. A notice and notice scheme to manage ISP liability is also important. Experiences in the US have shown how notice and takedown can be used to quickly remove data-based works from circulation or sale over the internet. Where the issues about the scope of any copyright protection in fact-based works and about the protectable elements in any compilation of facts are so uncertain, claims to copyright are difficult to verify or assess. When combined with notice and takedown, this would have an unnecessarily chilling effect.