

March 2016

## Public Transit Data Through an Intellectual Property Lens: Lessons About Open Data

Teresa Scassa  
*University of Ottawa*

Follow this and additional works at: <https://ir.lawnet.fordham.edu/ulj>

 Part of the [Energy and Utilities Law Commons](#), [Intellectual Property Law Commons](#), [Internet Law Commons](#), and the [Transportation Law Commons](#)

---

### Recommended Citation

Teresa Scassa, *Public Transit Data Through an Intellectual Property Lens: Lessons About Open Data*, 41 Fordham Urb. L.J. 1759 (2014). Available at: <https://ir.lawnet.fordham.edu/ulj/vol41/iss5/8>

This Article is brought to you for free and open access by FLASH: The Fordham Law Archive of Scholarship and History. It has been accepted for inclusion in Fordham Urban Law Journal by an authorized editor of FLASH: The Fordham Law Archive of Scholarship and History. For more information, please contact [tmelnick@law.fordham.edu](mailto:tmelnick@law.fordham.edu).

# PUBLIC TRANSIT DATA THROUGH AN INTELLECTUAL PROPERTY LENS: LESSONS ABOUT OPEN DATA

*Teresa Scassa\**

## ABSTRACT

This Article examines some of the challenges presented by the transition from ‘closed’ to open data within the municipal context, using municipal transit data as a case study. The particular lens through which this Article examines these challenges is intellectual property law. In a ‘closed data’ system, intellectual property law is an important means by which legal control over data is asserted by governments and their agencies. In an ‘open data’ context, the freedom to use and distribute content is a freedom from IP constraints. The evolution of approaches to open municipal transit data offers some interesting examples of the role played by intellectual property at every stage in the evolution of open municipal transit data, and it highlights not just the relationship between municipalities and their residents, but also the complex relationships between municipalities, residents, and private-sector service providers.

## TABLE OF CONTENTS

Introduction .....	1760
I. A Typology of Transit Data .....	1763
II. Delivering Transit Data.....	1767

---

\*Canada Research Chair in Information Law, University of Ottawa. This Article was initially presented as part of the Smart Law for Smart Cities symposium hosted by the Urban Law Journal of Fordham University School of Law, in February 2014. Many thanks to Alexandra Diebel for her excellent assistance with the research for this Article and for her comments on drafts. Thanks also to Robert Giggey, Tracey Lauriault, Patrick Levesque, Graham Reynolds, and Charles Sanders. The support of the Social Sciences and Humanities Research Council of Canada for the Geothink project (Geothink.ca), and the support of the Canada Research Chairs Program are both gratefully acknowledged.

III. The Economic Value of Transit Data .....	1773
IV. Asserting Intellectual Property Rights in Transit Data .....	1775
A. Copyright in Transit Data .....	1779
B. Transit Data Intellectual Property Skirmishes .....	1789
1. Transit Maps .....	1789
2. Static Transit Data .....	1791
3. Real-Time Transit Data.....	1798
4. Legal Interoperability and Clear Licensing .....	1803
5. Patent Claims .....	1805
Conclusion.....	1808

### INTRODUCTION

The open government data movement<sup>1</sup> is gathering steam both at the domestic and international levels, and, to some extent, it has more recently been folded into the broader embrace of open government.<sup>2</sup> The promises of open data are many, and they include transparency and accountability,<sup>3</sup> improved efficiency in the delivery of services and in planning activities,<sup>4</sup> greater citizen engagement,<sup>5</sup> better uptake of government services,<sup>6</sup> and the stimulation of innovation and

1. For background on this movement, see *The Annotated 8 Principles of Open Government Data*, OPENGOVDATA, <http://opengovdata.org/> (last visited Nov. 9, 2014).

2. President Obama's Memorandum on Transparency and Open Government set the stage for domestic developments in the United States. See *Memorandum on Transparency and Open Government*, 74 Fed. Reg. 4685 (Jan. 21, 2009), available at [http://www.whitehouse.gov/the\\_press\\_office/TransparencyandOpenGovernment/](http://www.whitehouse.gov/the_press_office/TransparencyandOpenGovernment/). At the international level, the Open Government Partnership (OGP) represents a growing international coalition of governments and other stakeholders dedicated to advancing open government. See OPEN GOV'T PARTNERSHIP, <http://www.opengovpartnership.org> (last visited Nov. 9, 2014). States that are members of the OGP commit to establishing and meeting goals within the framework established by that organization. See *Open Government Declaration*, OPEN GOV'T PARTNERSHIP, <http://www.opengovpartnership.org/about/open-government-declaration> (last visited Nov. 9, 2014); see also OECD, TOWARDS RECOVERY AND PARTNERSHIP WITH CITIZENS: THE CALL FOR INNOVATIVE AND OPEN GOVERNMENT (2010), available at <http://www.oecd.org/governance/ministerial/46342001.pdf>.

3. Patrice McDermott, *Building Open Government*, 27 GOV'T INFO. Q. 401, 403 (2010); see also Katleen Janssen, *Open Government Data and the Right to Information: Opportunities and Obstacles*, J. COMMUNITY INFORMATICS (2012), <http://ci-journal.net/index.php/ciej/article/view/952/954>.

4. Janssen, *supra* note 3; Barbara Ubaldi, *Open Government Data: Towards Empirical Analysis of Open Government Data Initiatives* (OECD Working Papers on Pub. Governance, No. 22, 2013), available at <http://dx.doi.org/10.1787/5k46bj4f03s7-en>.

5. McDermott, *supra* note 3, at 403–04; Janssen, *supra* note 3, at § 2.1.

6. One of the arguments for open data in the transit sector is that the development of apps that make transit information more easily accessible will

economic growth.<sup>7</sup> The challenges posed by open data are also significant. Perhaps one of the most difficult is managing the change in institutional culture from closed to open data.<sup>8</sup> Wrapped up in the institutional cultural shift are issues of information control, concerns over quality and liability, difficulties in negotiating relationships with private sector suppliers around open data, inexperience and lack of resources, and concerns over lost opportunities for revenue generation.<sup>9</sup> At the same time, individuals are becoming increasingly vocal about their desire to have access to government data, and increasingly engaged in finding uses and applications for this data.<sup>10</sup>

This Article examines some of the challenges presented by the transition from closed to open data within the municipal context, and uses municipal public transit data as a case study. The choice of a municipal data case study is driven by the fact that there has been very strong and early interest in municipal data. This data relates to the communities in which people live, and to the services upon which they rely most directly in their daily lives. Within the broad category of municipal data, transit data is particularly interesting because developers have been keen to access it as open data, and there has been more experience with its use as a result. Francisca Rojas describes transit data as “one of the earliest and arguably most successful cases of open data adoption in the U.S.”<sup>11</sup> Municipal

---

increase ridership. A study of the adoption of a real-time transit data system in Chicago found that the provision of real-time transit data (in this case both from official transit authority sources and through apps made by independent developers) did lead to a slight increase in ridership. See Lei Tang & Piyushimita Thakuriah, *Ridership Effects of Real-Time Bus Information System: A Case Study in the City of Chicago*, 22 *TRANSP. RES. PART C: EMERGING TECH.* 146 (2012), available at <http://dx.doi.org/10.1016/j.trc.2012.01.001>.

7. Janssen, *supra* note 3, at § 2.1.

8. FRANCISCA M. ROJAS, *TRANSIT TRANSPARENCY: EFFECTIVE DISCLOSURE THROUGH OPEN DATA* 22 (2012), available at [http://www.transparencypolicy.net/assets/FINAL\\_UTC\\_TransitTransparency\\_8%2028%202012.pdf](http://www.transparencypolicy.net/assets/FINAL_UTC_TransitTransparency_8%2028%202012.pdf); see also, EUROPEAN PUB. SECTOR INFO. PLATFORM, *OPEN TRANSPORT DATA MANIFESTO* (2012), available at <http://www.scribd.com/doc/111890372/Helsinki-Open-Transport-Data-Manifesto>.

9. ROJAS, *supra* note 8, at 22.

10. Ubaldi, *supra* note 4, at 19–20, notes the growing role of private actors in making use of government data. In the transit context in particular, when checked on May 8, 2012, the Google Group for transit developers had 795 members. See *Transit Developers*, GOOGLE GROUPS, <https://groups.google.com/forum/#!aboutgroup/transit-developers> (last visited May 8, 2012).

11. ROJAS, *supra* note 8, at 15; see also *Memorandum on Transparency and Open Government*, *supra* note 2 (dating back to January 2009). In the United Kingdom, the open data portal, [data.gov.uk](http://data.gov.uk), launched in January of 2010. Kevin Anderson, *Tim Berners-Lee Launches UK Public Data Website*, *GUARDIAN* (Jan. 21, 2010),

transit data also represents a category of data in which there is both strong public interest<sup>12</sup> and potential economic value.<sup>13</sup> There has been a great deal of resident<sup>14</sup> mobilization in favor of open transit data. There has also been a significant degree of uptake in the development of apps related to transit data in those municipalities which have made it open.<sup>15</sup>

The particular lens through which this Article examines the struggles over municipal public transit data is intellectual property law.<sup>16</sup> In a “closed data” system, intellectual property law is the means by which legal control over data is asserted by governments and their agencies. The classic definition of “open” in relation to both data and content provides that “[a] piece of data or content is open if anyone is free to use, reuse, and redistribute it—subject only,

---

<http://www.theguardian.com/technology/blog/2010/jan/21/timbernerslee-government-data>. The Open Government Partnership launched on September 20, 2011. See OPEN GOV'T PARTNERSHIP, *supra* note 2. As will be seen in the discussion below, demand for open transit data in various U.S. municipalities dates back to the mid-2000's.

12. See DANIEL DIETRICH, STATE OF PLAY: RE-USE OF TRANSPORT DATA 4 (2012), available at [http://www.epsiplatform.eu/sites/default/files/TopicReport\\_TransportData.pdf](http://www.epsiplatform.eu/sites/default/files/TopicReport_TransportData.pdf) (stating “transport data is highly relevant to citizens’ everyday lives, whether they are using private or public transport, or a mixture of both,” and in doing so, highlighting the importance of such data for both industry and business).

13. City-Go-Round is a grant-funded organization that has as its mission making public transit more convenient. See *About City-Go-Round*, CITY-GO-ROUND, <http://www.citygoround.org/about/> (last visited Nov. 9, 2014). It has a website that acts as a clearinghouse for transit data apps that are built upon open transit data. *Id.* The proliferation of these apps speaks not just to the level of public interest in this data, but also to its potential to support innovation.

14. It is quite common in the literature around open government and open data to see the term “citizen” used to refer to individuals within the relevant community (whether it is national, state-level, or municipal). However, “citizen” is a term loaded with legal connotations relating to status within a country. I use the less loaded term “resident” to refer to those living within a municipality, as this category may include many non-citizens, from recent immigrants to students studying on visas.

15. City-Go-Round maintains a list of transit agencies that provide open transit data—and a list of those that do not. See *All Transit Agencies*, CITY-GO-ROUND, <http://www.citygoround.org/agencies/> (last visited Nov. 9, 2014). As of November 9, 2014, of a total of 1026 transit agencies considered worldwide, City-Go-Round listed 291 of those as having open data, with 735 remaining closed. *Id.*

16. The focus of this Article is predominantly on copyright and patent law issues, with a brief mention of the relevance of database rights in the European context. It is worth noting that other related IP issues may arise. For example, municipal transit authorities have asserted trademark rights in logos, symbols, and marks, and open data licenses generally restrict the use of trademarks in relation to downstream products created using the licensed data. See *infra* note 76. A discussion of trademark issues, while interesting, is beyond the scope of this Article.

at most, to the requirement to attribute and/or share-alike.”<sup>17</sup> The freedom to use and distribute content is inherently a freedom from IP constraints. In this way, IP rights, and their effective waiver, are intrinsic to open data. The transition to open data, therefore, necessarily requires a relaxing of this form of proprietary control. This shift from closed to open is often difficult for governments. It is more than a simple policy change; it may require a broader change to institutional culture.<sup>18</sup> The evolution of approaches to open municipal transit data offers some interesting examples of the role played by intellectual property at every stage, and it highlights not just the relationship between municipalities and their residents, but also the complex relationships between municipalities and private sector service providers.

This discussion of open transit data through the lens of IP law unfolds as follows: Part I of this Article offers a discussion of the different types of transit data, with a particular focus on transit maps, static transit data, and real-time GPS data. Part II considers how transit data is delivered to the public, and how the modes of delivery have changed with emerging technologies. The economic value of transit data is also affected by technology, and Part III of this Article discusses this issue. Part IV considers how, and in what circumstances, intellectual property rights have been asserted in transit data, and examines both the scope of copyright in transit data and the different legal skirmishes over IP rights—including patent troll activities—in relation to the use of such data. The Article concludes with a discussion of the lessons to be drawn from this study of claims to IP rights in municipal transit data.

### I. A TYPOLOGY OF TRANSIT DATA

There are many different types of data generated through the operation of public transportation systems, and the variety of such

---

17. *The Open Definition*, OPEN DEFINITION, <http://opendefinition.org/od/> (last visited Nov. 9, 2014).

18. For example, many governments adopted a cost-recovery model in relation to licensing data. See, e.g., METROGIS DATA PRODUCERS WORK GROUP, *Making Public GIS Data Free and Open: Benefits and Challenges*, in METROGIS: FREE & OPEN ACCESS TO DATA RESEARCH & REFERENCE DOCUMENTS 3 (2013), available at [http://metrogis.org/MetroGIS/media/gis-documents/publications/MetroGIS\\_014\\_FreeAndOpenDataResearch.pdf](http://metrogis.org/MetroGIS/media/gis-documents/publications/MetroGIS_014_FreeAndOpenDataResearch.pdf). Brett Goldstein writes of “fear” as being a major barrier to municipal government adoption of open data. See Brett Goldstein, *Open Data in Chicago: Game On*, in BEYOND TRANSPARENCY: OPEN DATA AND THE FUTURE OF CIVIC INNOVATION 20–22 (Brett Goldstein & Lauren Dyson eds., 2013).

data continues to grow with technology.<sup>19</sup> While ‘transit data’ as a general category could include data about any transportation system, the focus of this Article is on municipal public transit data. This Article will consider three specific categories of data that have been of interest to transit users: route maps, static transit data, and real-time GPS data. Nevertheless, it is important to keep in mind that there are many other kinds of transit data and the transit data context is in a constant state of evolution. As technology evolves, so too do the kinds and volume of data that is collected. For example, data from smart payment systems represents a relatively new category of transit data and one in which there may soon be a rather broad interest.<sup>20</sup>

A map is not data so much as it is a particular way in which data is presented. Prior to the development of Web 2.0,<sup>21</sup> transit maps were relatively simple things; they provided visual representations of transit routes, an overview of the transit network, and the location of stops and connection points. While the plotted information might change or evolve over time, such maps were not generally subject to rapid change. These types of transit maps offered transit users a way to visualize the system on which they were travelling. While these more static, non-interactive maps are still used, they co-exist with

---

19. For example, transit data, considered broadly, could include not just transit system data, but also data about traffic flows, accidents, infrastructure, ridership, expenditures, and so on.

20. Smart cards are chip-enabled cards that allow for fare payment from an electronic purse established by a transit user and associated with the user’s card. *See Smart Card Primer*, SMART CARD ALLIANCE, <http://www.smartcardalliance.org/smart-cards-intro-primer/> (last visited Nov. 9, 2014). A transit user swipes or taps the card on a card reader installed on a public transit vehicle in order to effect payment of the fare. *See id.* For a discussion of potential uses of smart card data, see BRUNO AGARD ET AL., CENTRE INTERUNIVERSITAIRE DE RECHERCHE SUR LES RÉSEAUX D’ENTREPRISE, LA LOGISTIQUE ET LE TRANSPORT, MINING PUBLIC TRANSPORT USER BEHAVIOUR FROM SMART CARD DATA (2007), *available at* <https://www.cirrelt.ca/DocumentsTravail/CIRRELT-2007-42.pdf>; Marie-Pier Pelletier et al., *Smart Card Data Use in Public Transit: A Literature Review*, 19 *TRANSP. RES. PART C: EMERGING TECHNOLOGIES* 557 (2011).

21. The term “Web 2.0” is generally used to refer to a vision of the World Wide Web as an interactive platform that features user-generated content, networked collaboration, and social networking. *See* Tim O’Reilly, *What Is Web 2.0: Design Patterns and Business Models for the Next Generation of Software*, O’REILLY MEDIA (Sept. 30, 2005), <http://oreilly.com/web2/archive/what-is-web-20.html>. Rather than the delivery of static information to the public via websites, the concept of “Web 2.0” sees web users contributing to the creation and dissemination of content. *See id.*

more contemporary maps that are interactive, and that may embed a great deal of additional information.<sup>22</sup>

Schedule data, also known as static transit data,<sup>23</sup> offers a different layer of information linked to transit routes and stops. In essence, it is timetable data. This kind of data is the result of a significant planning exercise. In drawing up transit timetables, consideration must be given to which routes are in high demand, daily and seasonal traffic patterns, demographic considerations, limitations in terms of available vehicles and their load capacity, and so on. For larger municipalities, this process would be carried out using software tools supplied by third party vendors.<sup>24</sup> The resulting transit timetables may be subject to adjustment or periodic review or reworking of the transit system.

Real-time transit data are defined as “data that are being collected at the same time as they are being generated and that may be disseminated immediately.”<sup>25</sup> Such data changes rapidly and are generated as part of the day-to-day operations of a transit authority.<sup>26</sup> They are generated when a GPS unit installed on a transit vehicle communicates information to a server at regular intervals. This information may include the geographic coordinates of the vehicle, the vehicle and route identifiers, and the time at which the coordinates were recorded.<sup>27</sup> These types of data are useful in planning, assessing performance on particular routes, and evaluating the overall operation of the transit system. They have also proven to be useful to transit users who want to know whether a particular bus is likely to arrive late or early at a given stop.

---

22. Some of this data may not be transit data, or, viewed alternatively, the transit data may be embedded in a multi-purpose map which provides information about local businesses, public institutions, bike paths, and so on.

23. The term “static transit data” generally contrasts timetable data with real-time GPS data, which is discussed later in this Article. Schedule data is considered “static” in comparison, because it does not change from day to day (or from minute to minute). Changes tend to be at predictable intervals (for example, the introduction of a summer schedule).

24. ROJAS, *supra* note 8, at 22.

25. Larry W. Thomas, *Legal Arrangements for Use and Control of Real-Time Data*, 37 LEGAL RES. DIG. 3, 3 (2011), available at [http://onlinepubs.trb.org/onlinepubs/tcrp/tcrp\\_lrd\\_37.pdf](http://onlinepubs.trb.org/onlinepubs/tcrp/tcrp_lrd_37.pdf) (noting that the data that can be collected using real-time GPS systems goes beyond location information and may include other details such as the speed of the vehicle, braking, and the opening and closing of doors).

26. Rojas notes that real-time transit data served many operational needs including system management, monitoring and adjusting performance, and facilitating the location of vehicles in emergency situations. ROJAS, *supra* note 8, at 19.

27. *Id.*



While schedule data can be downloaded in bulk, real-time transit data is constantly changing. As a result, an Application Programming Interface (API) is required to make proper use of these data. An API is a communication protocol between an app and a data source.<sup>28</sup> For example, an API might be designed to retrieve subsets of a transit authority's real-time GPS data at set intervals. This data can then be transformed/reformatted and published on a website or made available through an app. While outside developers could create their own APIs for real-time transit databases made open by municipalities, the preferable approach is for the municipality to provide and maintain an API for developers.<sup>29</sup> This way, if the municipality makes changes to its real-time GPS database, the app developers do not necessarily need to change their applications.

Real-time data are more complex than static transit data for a number of reasons. The actual volume of data is much higher: each GPS-equipped vehicle operating within the transit system communicates a live stream of data throughout its period of operation.<sup>30</sup> While the data may follow similar patterns over time, this is not the same on a repeating basis. The progress of vehicles may be affected by weather, traffic, accidents, or other factors.<sup>31</sup> As a result, for any given vehicle on any given route, the real-time transit data may vary—often significantly—from both the static timetable data and from one day to the next. The real value of these data (from the perspective of the transit user) is the ability to correlate these data with particular points along transit routes—in other words, to be able to predict with some degree of accuracy when the vehicle will approach any given stop.<sup>32</sup> Thus, algorithms can be applied to real-time transit data in order to produce this predictive data.<sup>33</sup>

---

28. *Definition of API*, PC MAG. ENCYCLOPEDIA, <http://www.pcmag.com/encyclopedia/term/37856/api> (last visited Nov. 9, 2014).

29. A single API developed by a transit authority will standardize the data requests made by developers and is thus more efficient and more manageable. *See* ROJAS, *supra* note 8, at 58.

30. For a discussion of the functioning of automatic vehicle location (AVL) technologies, see PETER G. FURTH ET AL., USING ARCHIVED AVL-APC DATA TO IMPROVE TRANSIT PERFORMANCE AND MANAGEMENT 25 (2006), *available at* [http://onlinepubs.trb.org/onlinepubs/tcrp/tcrp\\_rpt\\_113.pdf](http://onlinepubs.trb.org/onlinepubs/tcrp/tcrp_rpt_113.pdf).

31. Transit vehicles may arrive at stops either earlier or later than their scheduled times for a variety of reasons that are often unpredictable and that may change from one day to the next. A snowstorm, for example, may significantly affect the on-time performance of buses; traffic jams, electrical outages, accidents, or other events may also cause off-schedule arrival or departure times.

32. *See* ROJAS, *supra* note 8, at 39 (suggesting that real-time data may ultimately displace timetable data). In other words, customers will no longer think in terms of

These three categories of transit data—maps, static transit data, and real-time GPS data—map onto both the evolution of available data due to technology (from very simple to more complex and real-time), and the evolution of technology in the hands of transit users (from paper-and-ink to mobile and interactive devices). This, in turn, has driven changes regarding whether and how data has been made available to those users. In other words, the demand for open transit data is driven in part by the technology in the hands of transit system users, and in part by the increasing relevance of the available data. Thus, it is important to keep in mind that there has been an evolution in the nature and volume of available data as well as a marked progression in the technologies available to gather, process, and share this information. There is every reason to expect that both the nature and volume of data and the related technologies will continue to change and expand. In this way, principles and practices developed in relation to earlier (and simpler) categories of data may have a significant impact in shaping how later, more complex types of data are managed.

## II. DELIVERING TRANSIT DATA

This section considers the manner in which transit data is delivered to the public, and how these modes of delivery have been affected by emerging technologies. Municipalities that operate public transportation systems generate a significant body of data related to the operations of that system.<sup>34</sup> Much of that data is of a kind that is directly of interest to those who use the public transit system. In the case of system maps and transit timetables, the sharing of this data with transit users is an essential part of rendering the system useable. Although real-time transit data may have initially been collected for internal system management purposes, it quickly became of interest

---

when a bus is *supposed* to arrive at a given stop; rather, they will simply focus on when the next bus is predicted to do so.

33. Note that real-time transit data can also be used to ensure greater transparency and accountability of transit authorities. For example, developers could develop apps that use this data to assess the overall performance of the system and to provide an outside assessment of its operations. *See id.*

34. Such data includes timetable and route information. However, it also includes information gathered by transit authorities regarding customer satisfaction, incidents and safety concerns, ridership levels, payment, on-time numbers, total trips, and data gathered from the use of real-time GPS systems. For a synthesis of some of the routinely collected transit data, see AM. PUB. TRANSP. ASS'N, 2013 PUBLIC TRANSPORTATION FACT BOOK (2013), *available at* <http://www.apta.com/resources/statistics/Documents/FactBook/2013-APTA-Fact-Book.pdf>.

to transit users and app developers.<sup>35</sup> Other data that may be collected by transit agencies may be of lesser value to ordinary citizens,<sup>36</sup> but nonetheless may be of great interest to researchers, urban planners, developers, community groups, and private sector companies engaged in data profiling.<sup>37</sup> This may include data about transit ridership, patterns of transit use, and so on. As more transit systems move towards smart card technologies for fare payment,<sup>38</sup> the quality of this data becomes more fine-grained, and thus of greater potential interest.<sup>39</sup> What is important to consider here is that ‘transit data’ is not a closed category; it continues to expand in both volume and detail.<sup>40</sup>

Public transit authorities have traditionally published maps, schedules, and timetables in a variety of formats for their ridership. Prior to digital technologies and the internet, this information was chiefly made available in the form of maps or schedule data displayed on signs in vehicles or at stops, or distributed to transit users in paper

---

35. ROJAS, *supra* note 8, at 7.

36. It should be noted that, as was the case with real-time data, the full potential for such data may be latent.

37. For example, data from electronic ticketing systems for public transit have already been used by researchers to study commuting patterns, with a view to improving city planning. See *Oyster Gives Up Pearls*, UCL ENGINEERING, <http://www.engineering.ucl.ac.uk/blog/projects/oyster-gives-up-pearls/> (last visited Nov. 9, 2014). An interesting example of citizen use of transit data in combination with demographic data is the story of an Ottawa man who used this data to show that the planned location for a light rail stop in Ottawa was not as convenient for potential riders as another location. See *Citizens Fact-Check Transit Claims, Open Data Initiative Means Tech-Savvy Residents Can See Info First-Hand*, OTTAWA COMMUNITY NEWS, Aug. 2, 2012, available at <http://www.ottawacommunitynews.com/news-story/3964959-citizens-fact-check-transit-claims-open-deata-initiative-means-tech-savvy-residents-can-see-info-fir/>.

38. See, e.g., SMART CARD ALLIANCE, A GUIDE TO PREPAID CARDS FOR TRANSIT AGENCIES (2011), available at [http://www.smartcardalliance.org/resources/pdf/Prepaid\\_Cards\\_for\\_Transit\\_Agencies\\_20110212.pdf](http://www.smartcardalliance.org/resources/pdf/Prepaid_Cards_for_Transit_Agencies_20110212.pdf).

39. One issue not addressed in this Article is where the line will be drawn in terms of the nature and quality of data that is made available to the public under open data programs. This is clearly an important issue in the context of Smart Cities, as large volumes of detailed, high quality data about any manner of services will be at the core of Smart Cities. Although the sharing of transit-user data raises interesting issues, a discussion of these is beyond the scope of this Article.

40. The expansion of categories of transit data is evident from the types of data tracked by the American Public Transportation Association in its annual Public Transportation Fact Books. See AM. PUB. TRANSP. ASS’N, *supra* note 34. The kind of information collected may be directly related to technologies, as certain technologies produce entirely new types of data. This is the case with real-time GPS systems, for example, which permit the collection of detailed information regarding the location of vehicles at specific points in time. See, e.g., FURTHET AL., *supra* note 30.

format.<sup>41</sup> As the World Wide Web evolved, both maps and static transit data were displayed on transit company websites.<sup>42</sup> Trip planners—online interfaces that provide users with timetable information and connections to take them from their point of departure to their destination—were also eventually provided by many transit agencies.<sup>43</sup> Today, static transit data is still available on paper, on websites, and through trip planners.<sup>44</sup> Increasingly, system users access transit data by telephone, via text messaging systems, and through mobile transit apps.<sup>45</sup>

While transit authorities were once exclusive sources of transit information, the digital era has seen a variety of new players emerge as providers of information, either independently or in partnership with transit authorities. Google was a relatively early entrant into the contemporary field of provision of transit data.<sup>46</sup> It saw an opportunity to enhance its Google Maps service by incorporating route and schedule information into the “directions” feature of Google Maps, with some transit stops and schedule data automatically overlaid on the maps themselves.<sup>47</sup> Google began to negotiate agreements with transit authorities for access to schedule data.<sup>48</sup> As part of this process, Google, in conjunction with Portland’s

---

41. ROJAS, *supra* note 8, at 16.

42. *See id.* at 7.

43. *Id.* at 16.

44. *Id.* at 16–19.

45. Text messaging transit information services are common. *See, e.g., How-to Guide: Bus Tracker by Text*, CHI. TRANSIT AUTHORITY, [http://www.transitchicago.com/riding\\_cta/how\\_to\\_guides/bustrackertext.aspx](http://www.transitchicago.com/riding_cta/how_to_guides/bustrackertext.aspx) (last visited Nov. 9, 2014). Telephone-based systems may also be available. *See, e.g., TripFinder*, ENGHOUSE TRANSP., <http://www.enghousetransportation.com/products/public-transportation/trip-planner.php> (last visited Nov. 9, 2014). Mobile transit apps are increasingly ubiquitous. *See, e.g., CITY-GO-ROUND*, <http://www.citygoround.org/> (last visited Nov. 9, 2014) (providing a public transit app-finding tool).

46. Google began working on this issue in 2005 in conjunction with Portland’s Trimet Transit Agency. *See* Matthew Roth, *How Google and Portland’s TriMet Set the Standard for Open Transit Data*, STREETS BLOG SF (Jan. 5, 2010), <http://sf.streetsblog.org/2010/01/05/how-google-and-portlands-trimet-set-the-standard-for-open-transit-data/>.

47. For a list of transit authorities that share data with Google Maps from around the world, see *Cities Covered*, GOOGLE MAPS TRANSIT, <http://www.google.com/landing/transit/cities/index.html> (last visited Nov. 9, 2014). Of course, not all of these share real-time data, but all of them share at least static transit data. *See* ROJAS, *supra* note 8, at 24 (describing Google Maps Transit data as “a free alternative to the expensive online trip planners in which larger agencies had invested”).

48. Not all transit authorities were initially keen to provide their data to a large private corporation such as Google without evidence of a clear benefit. *See* LAUREN PESSOA ET AL., ENABLING TRANSIT SOLUTIONS: A CASE FOR OPEN DATA 18 (2011). Note as well that Google did not become an open transit data repository as a result of

TriMet transit agency, developed a standard format for organizing static transit data, known as the Google Transit Feed Specification.<sup>49</sup> After working cooperatively with developers, Google made its transit feed specification an open standard, and as a result, its name was changed to the General Transit Feed Specification (GTFS).<sup>50</sup> When transit authorities began to collect real-time data, Google also developed a real-time standard<sup>51</sup> and sought to have this data incorporated into its maps as well.<sup>52</sup>

---

this incorporation of transit data into its maps. See EDWARD L. HILLSMAN & SEAN J. BARBEAU, ENABLING COST-EFFECTIVE MULTIMODAL TRIP PLANNERS THROUGH OPEN TRANSIT DATA 66 (2011), available at <http://www.nctr.usf.edu/wp-content/uploads/2011/06/77926.pdf>. Application developers for transit data could not acquire this data from Google, and still had to source them directly from the relevant transit authority. See *id.* Thus if the transit authority had an agreement with Google, but did not otherwise provide open transit data, this would be a barrier to application development. See *id.* Hillsman and Barbeau also note that transit agencies are dependent upon Google to upload their data, and note that wait times may sometimes be significant, especially as the demand by transit agencies to have their data included grows. See *id.* at 3.

49. ROJAS, *supra* note 8, at 23. Google's interest in publishing municipal transit data has been credited with pushing transit authorities to structure their data according to the General Transit Feed Specification (GTFS), thus making it more useful to developers when it was ultimately released as open data. See ROJAS, *supra* note 8, at 8. Note that Microsoft's Bing Maps also now uses transit data in the GTFS format. See Brian Hendricks, *Bing Maps Gets Transit Directions*, BING BLOGS (Sept. 16, 2010), [http://www.bing.com/blogs/site\\_blogs/b/maps/archive/2010/09/16/bing-maps-gets-transit-directions.aspx](http://www.bing.com/blogs/site_blogs/b/maps/archive/2010/09/16/bing-maps-gets-transit-directions.aspx). For a description of the GTFS, see HILLSMAN & BARBEAU, *supra* note 48, at 6–7.

50. PESSOA ET AL., *supra* note 48, at 7; Bibiana McHugh, *Pioneering Open Data Standards: The GTFS Story*, in BEYOND TRANSPARENCY: OPEN DATA AND THE FUTURE OF CIVIC INNOVATION 125, 132 (Brett Goldstein & Lauren Dyson eds., 2013) (noting that this name change proved important to a broad range of parties, including transit agencies “who were worried about losing control of their data” to a private sector corporation). The acronym GTFS is also used to refer to the renamed standard. Indeed, the decision to replace ‘Google’ with ‘General’ was made in order to preserve the acronym. See Joe Hughes, *General Transit Feed Spec Changes, Proposal: Remove ‘Google’ from the Name of GTFS*, Google Groups (Oct. 19, 2009), [https://groups.google.com/forum/#!topic/gtfs-changes/ob\\_7MIOvOxU](https://groups.google.com/forum/#!topic/gtfs-changes/ob_7MIOvOxU). Note that there is a transit developers group on Google Groups for those interested in sharing information related to open transit data. See *Transit Developers*, GOOGLE GROUPS, *supra* note 10. The Open Knowledge Foundation has also created a working group to discuss Open Transport initiatives. See OPEN KNOWLEDGE FOUND., OPEN TRANSPORT WORKING GROUP, <http://transport.okfn.org> (last visited Nov. 9, 2014).

51. This is the General Transit Feed Specification—Real Time. See *infra* notes 222–23 and accompanying text.

52. Google Maps encourages transit authorities to supply real-time transit data to its maps. See *Live Transit Updates*, GOOGLE MAPS CONTENT PROVIDERS, <http://maps.google.ca/help/maps/mapcontent/transit/live-updates.html> (last visited Nov. 9, 2014).

The explosive rise in the use of mobile devices also gave rise to a vibrant app developer community, and, in this way, developers became a new source of repackaged public transit data.<sup>53</sup> This development of mobile apps was driven largely by transit users who, in the mid-2000's, began to seek out transit data in order to create apps that would make these data available to them in more useful formats.<sup>54</sup> This demand for transit data was a strong driver of municipal open data initiatives.<sup>55</sup> As transit authorities began to deploy GPS tracking systems on their vehicles, developer demand also grew for access to this data in order to create apps that would provide users with better estimates of the actual (as opposed to scheduled) time of arrival of transit vehicles at particular stops.<sup>56</sup> The fact that many transit authorities had already organized their data according to the GTFS (and later its real-time counterpart—"GTFS realtime") gave developers access to data in a readily usable format.<sup>57</sup>

The transition to real-time transit data introduced another set of players in delivering transit data. Many transit authorities entered into contracts with private sector companies to supply GPS units for vehicles as well as the hardware and software necessary to collect and process this data.<sup>58</sup> The software packages also offered predictive data regarding the expected time of arrival at individual stops.<sup>59</sup>

---

53. See generally ROJAS, *supra* note 8 (describing the evolution of the developer community).

54. *Id.* at 43.

55. *Id.* This same relationship between demand for transit data and the growing availability of open transit data is also observed in the European context. See DIETRICH, *supra* note 12, at 7.

56. Real-time GPS tracking systems were initially used by transit authorities for internal purposes that included system-wide monitoring and performance evaluation. See, e.g., ROJAS, *supra* note 8, at 7.

57. As noted above, the GTFS was initially developed by Google in conjunction with Portland's TriMet transit authority. PESSOA ET AL., *supra* note 48, at 7. For an account of this collaboration, see McHugh, *supra* note 50; Michael Perkins, *Metro Refuses to Participate in Google Transit*, GREATER GREATER WASHINGTON (Dec. 13, 2008), <http://greatergreaterwashington.org/post/1495/metro-refuses-to-participate-in-google-transit/>. Open data may also be made available in file formats that allow easy reuse of the data. The standards discussed above provide a means of organizing transit data. Making data open in reusable file formats does not prevent a developer from using a particular standard to organize the data.

58. ROJAS, *supra* note 8, at 19.

59. NextBus, the leading North American company in this field, claims to provide its services to over 135 transit agencies. See *NextBus Real-Time Passenger Information Solutions*, NEXTBUS, <http://cts.cubic.com/en-us/solutions/realtimepassengerinformation/nextbus,inc.aspx> (last visited Nov. 9, 2014). Other real-time GPS related services are provided by companies such as RouteMatch. See *Traveler Information Systems*, ROUTEMATCH SOFTWARE, <http://routematch.com/>

Some of these companies—such as NextBus—partnered with municipalities to provide this data to transit users on electronic sign boards or through other notification systems under the control of the transit authority.<sup>60</sup> Demand for access to this data by app developers was quick and strong, and many municipalities began to make it available as open data.<sup>61</sup>

With the growth and popularization of both mobile devices, shared software, and the proliferation of amateur software development skills, transit riders and app developers have become more active, not just as consumers of transit data, but also as users of this data. They have sought ways in which they can bend the data to their own purposes, and ways in which they can share the resultant apps with the broader public.<sup>62</sup> One of the possible benefits of open transit data is that it has the potential to stimulate the development of apps that go beyond what a transit authority might otherwise be prepared or capable of supporting, either technically or financially.<sup>63</sup> For example, developers may have reasons to develop apps that are of use to transit users with different types of disabilities,<sup>64</sup> or apps in

---

solutions/public-transit-fixed-route/traveler-information-systems/ (last visited Nov. 9, 2014).

60. See *Detailed Product Information*, NEXTBUS, <http://cts.cubic.com/solutions/real-timepassengerinformation/nextbus,inc/hownextbusworks/detailedproductinformation.aspx> (last visited Nov. 9, 2014).

61. City-Go-Round maintains a list of transit agencies that provide open data (and a list of those that do not). See *All Transit Agencies*, CITY-GO-ROUND, *supra* note 15.

62. See *Transit Developers*, GOOGLE GROUPS, *supra* note 10 (providing a large and active group of transit app developers on Google Groups). See also CITY-GO-ROUND, *supra* note 45 (dedicating a website to making the use of public transit more convenient and supporting both open transit data, and the development of transit apps).

63. Pessoa et al. note that app developers can provide information to transit users that crosses different agency and regional boundaries—something that might be beyond the scope of a single transit authority's mandate. PESSOA ET AL., *supra* note 48, at 4. The authors note that with open data, developers can develop information tools for all manner of devices—something that a transit authority may not have the resources to do. *Id.* at 4–5. Developers have also created apps, which combine different features, for example, a feature that would emit a sound designed to wake a snoozing commuter as their vehicle approaches their stop. See ROJAS, *supra* note 8, at 33.

64. ROJAS, *supra* note 8, at 33; see also PESSOA ET AL., *supra* note 48, at 5 (identifying the task of making transit data available to people with a broad range of perceptual disabilities as one of the challenges for transit authorities); Aaron Steinfeld et al., *Mobile Transit Rider Information Via Universal Design and Crowdsourcing*, 2217 TRANSP. RES. REC.: J. TRANSP. RES. BOARD 95 (2011), available at [http://www.cs.cmu.edu/~astein/pub/Steinfeld\\_TRB11\\_final.pdf](http://www.cs.cmu.edu/~astein/pub/Steinfeld_TRB11_final.pdf).

minority languages spoken within particular urban communities.<sup>65</sup> Today, many transit authorities rely upon user-generated apps to meet the needs of their transit users,<sup>66</sup> marking a milestone for municipal open data and a new era of partnership between transit authorities, developers and transit users.

The new technologies that made it possible to deliver static and real-time transit data to transit system users dramatically increased the accessibility of this information.<sup>67</sup> The proliferation of parties with interests in using transit data has, not surprisingly, also generated conflicts with respect to ownership and control of the data.<sup>68</sup>

### III. THE ECONOMIC VALUE OF TRANSIT DATA

Transit system maps and timetable data has little or no commercial value on its own. In order to make their systems attractive and useful to transit users, transit authorities must provide this information, and they have traditionally provided it free of charge.<sup>69</sup> Thus, at the point at which the earliest app developers were seeking these datasets, the sets held no commercial value to the transit authorities. Notwithstanding the early resistance to their release, their wider sharing and dissemination was arguably of real benefit.<sup>70</sup>

To the extent that app developers were able to develop useful apps that provided transit users with more convenient access to transit data, the data had a downstream commercial value in that it might lead to a commercial demand for such apps. In reality, many, if not most, public transit apps are free, or available for a very low cost.<sup>71</sup>

---

65. Sara M. Kaufman, *Getting Started with Open Data: A Guide for Transportation Agencies*, RUDIN CENTER FOR TRANSP. POL'Y & MGMT. 3 (May 1, 2012), <http://wagner.nyu.edu/files/rudincenter/opendata.pdf>.

66. The website City-Go-Round, <http://www.citygoround.org>, keeps a list of transit apps that make use of open data provided by transit authorities.

67. In this context, "accessibility" means primarily ease of access by transit riders via different mobile devices.

68. These conflicts are the subject of Part V of this Article.

69. ROJAS, *supra* note 8, at 16 (providing examples of the media traditionally used by transit authorities to provide information to their customers. These include paper schedules and maps, as well as similar information provided from websites.). Such information remains freely available today. *See, e.g., Schedules*, CHI. TRANSIT AUTHORITY, <http://www.transitchicago.com/schedules/> (last visited Nov. 9, 2014); *Schedules*, MTA.INFO, <http://www.mta.info/schedules> (last visited Nov. 9, 2014).

70. *See, e.g.,* McHugh, *supra* note 50.

71. City-Go-Round's Apps page provides links to a broad range of transit apps that are either free or available for download for anywhere from \$0.99 to a few dollars. *See Transit App Gallery*, CITY-GO-ROUND, <http://www.citygoround.org/apps/> (last visited Nov. 9, 2014).



As demand grew for both static and real-time transit data, some public transit authorities also explored the potential to derive a revenue stream from the use or licensing of such datasets.<sup>72</sup> Real-time transit data, in particular, has greater economic potential.<sup>73</sup> This data is “value-added,” in the sense that it is not essential to the transit rider’s use of the system, but may nonetheless be highly desirable.

In this respect, some transit authorities have contemplated ways in which value-added transit data might be delivered in conjunction with advertising content as another way of generating revenue. For example, advertising space can be sold in relation to real-time GPS data displayed on digital noticeboards in subway stations.<sup>74</sup> Advertising content might also be delivered through “official” apps or online trip planners.<sup>75</sup>

More indirect economic benefit can be derived from transit data where a company incorporates this data with other services it provides in order to enhance the usefulness or attractiveness of its services.<sup>76</sup> Certainly this may be an important motivation for Google, which now regularly incorporates both real-time and static transit data into its Google Maps service.<sup>77</sup> Transit data might also be

---

72. Washington’s Metropolitan Area Transit Authority (WMATA) was also concerned about the loss of potential to monetize its data, and commissioned a study to explore its commercial potential. *See* ROJAS, *supra* note 8, at 66.

73. The greater economic potential for this data lies in the fact that it is not essential to a transit rider’s commute in the same way that timetable data is. There is therefore no need to provide it to users free of charge, and there may be room to exploit its value-added nature.

74. *See, e.g.*, BRUCE SHALLER, TRANSP. RES. BOARD, TRANSIT ADVERTISING SALES AGREEMENTS: A SYNTHESIS OF TRANSIT PRACTICE 36 (2004), *available at* [http://onlinepubs.trb.org/onlinepubs/tcrp/tcrp\\_syn\\_51.pdf](http://onlinepubs.trb.org/onlinepubs/tcrp/tcrp_syn_51.pdf) (providing an example of the Greater Cleveland Regional Transit Authority installing electronic signs featuring real-time GPS data in order to generate revenue from advertising sales); *see also* *Traveler Information Systems*, ROUTEMATCH SOFTWARE, *supra* note 59 (offering real-time GPS data services to transit authorities with the potential, among other things, to increase their advertising revenues); *TSO Public Transportation*, TSO MOBILE, <http://www.tsmobile.com/TSO-Public-Transportation.html> (last visited Nov. 9, 2014) (touting the advertising potential of such digital noticeboards).

75. For example, WMATA was reported to have derived advertising revenue for each visit to its online trip planner. *See* ROJAS, *supra* note 8, at 66.

76. Google Maps encourages transit authorities to supply their data for incorporation into its maps. *See* *Google Transit Partner Program*, GOOGLE MAPS CONTENT PROVIDERS, <http://maps.google.com/help/maps/mapcontent/transit/> (last visited Nov. 9, 2014). While this is pitched as a benefit to transit authorities, it also enhances the usefulness of Google Maps.

77. *Id.* Note as well that transit data is also available through Microsoft’s Bing Maps service. *See* *Bing Transit Partners*, BING MAPS, <http://www.bing.com/maps/TransitPartners.aspx> (last visited Nov. 9, 2014).

incorporated into house-hunting tools, online news sites, and any number of other services in order to enrich or enhance the sites.

While municipalities might arguably benefit from licensing their data for a fee to others who plan to derive direct or indirect economic benefit from the exploitation of this data, others have argued that the core service of a municipal transit authority is to provide transit services.<sup>78</sup> From this perspective, if disseminating the data freely and as widely as possible makes the use of public transit more attractive to urban residents, the benefits of increased ridership will greatly outweigh any revenue from licensing.

#### IV. ASSERTING INTELLECTUAL PROPERTY RIGHTS IN TRANSIT DATA

This Part considers the extent to which intellectual property rights are asserted in transit data, and the circumstances in which such claims are made. It also examines the scope and extent of copyright in transit data and in compilations of such data. In addition, this Part takes into account other related IP issues, including the licensing of such data for use by developers, and the more recent assertion of patent claims in relation to real-time GPS data applications.

The transit data landscape described above is one in which transit authorities, through their operations, generate significant volumes of data that are relevant and of interest to a broad range of actors. It is perhaps not surprising that, as interest grew among different actors in repurposing this data, as opposed to just passively consuming it, many transit authorities reacted by asserting IP rights in their transit data or in the formats in which it was expressed.<sup>79</sup> This initial response has relaxed considerably over time. Today, some open data license agreements or terms of use seem to abandon any claim to copyright in transit data,<sup>80</sup> although others still assert rights to transit data—both real-time and static.<sup>81</sup>

---

78. See McHugh, *supra* note 50, at 129–31.

79. North American transit authorities that assert rights in their data make claims to copyright in data, schedules, and maps. See, e.g., Flegenheimer, *supra* note 78. In Europe, where *sui generis* database rights are recognized, database rights have been asserted in some cases. See DIETRICH, *supra* note 12.

80. See, e.g., *MTA License Information*, MTA.INFO, <http://web.mta.info/developers/license.html> (last visited Nov. 9, 2014) (“If you only wish to access and use the MTA’s data, this *does not* require a license.”) (emphasis in original).

81. See, e.g., *Terms of Use*, TRIMET DEVELOPER RESOURCES, [http://developer.trimet.org/terms\\_of\\_use.shtml](http://developer.trimet.org/terms_of_use.shtml) (last visited Nov. 9, 2014) (“All materials published on the Site, including, but not limited to, trademarks, service marks, maps, schedules,

Another phenomenon that is worth noting here is that as transit data evolved—alongside the related technologies and modes of expression—it became more questionable whether this data even could be protected by copyright.<sup>82</sup> In other words, the more static and concrete the expression of the data, the easier it was to recognize copyrightable expression (even if the available protection was relatively ‘thin’). Complex and digitized data sets and consistently evolving real-time data are inherently more difficult to categorize as works in which copyright subsists. Certainly, it becomes much more difficult to conceptualize the organization of data within a database as reflecting a particular arrangement. It is also more difficult to identify authorship in complex, non-finite collections of data. Finally, where the compilation as a whole is not copied (for example, in the case of real-time transit data) but rather just selected live-streamed data, it becomes more difficult to argue that something other than facts is being taken. Thus, the evolution of both the nature of the data and its mode of expression is relevant to the underlying intellectual property issues.

In the discussion of intellectual property issues that follows, it is important to keep in mind the range of factors that motivated transit authorities to initially refuse to share data. These motivations were often mixed.<sup>83</sup> A key concern was that of control over quality.<sup>84</sup> Transit authorities used intellectual property rights as a means of maintaining control over the dissemination of information regarding

---

arrival information, fare information, photographs and illustrations (collectively, the “Content”), are the property of TriMet unless otherwise indicated.”).

82. Thomas, *supra* note 25, at 4 (concluding that in a commissioned study on the intellectual property rights in real-time GPS transit data, it would be difficult to argue that copyright subsists in this type of data).

83. As will be seen in the discussion that follows, these motivations included concerns over quality control and authoritativeness, cost recovery, and profit-making.

84. PESSOA ET AL., *supra* note 48, at 19; ROJAS, *supra* note 8, at 60. Quality concerns may reflect different but related preoccupations. One is that the public may be misled or harmed by flawed data. The other is that the transit authority’s reputation may suffer if flawed information is disseminated. It is interesting to note that quality control is a key justification for Crown (state) copyright in those commonwealth jurisdictions that still assert government copyright over data. See Elizabeth F. Judge, *Crown Copyright and Copyright Reform in Canada*, in *IN THE PUBLIC INTEREST: THE FUTURE OF CANADIAN COPYRIGHT LAW* 551 (Michael Geist ed., 2005) (discussing in detail Crown copyright). Although control is sometimes asserted as necessary to ensure quality, it does not necessarily serve this purpose. See, e.g., 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*, 54 *CANADIAN GEOGRAPHER* 366, 366–74 (2010); David Vaver, *Copyright and the State in Canada and the United States*, 10 *INT’L PROP. J.* 187, 200 (1996).

their systems. In doing so, they demonstrated a reluctance to allow other non-authorized, non-official sources to communicate the same information, notwithstanding arguments that this information was in the public domain. In any event, these concerns over quality may well have been misplaced. A study by Rojas found that complaints about transit apps have generally been routed to the developers, who often have feedback processes in place to help identify bugs and to improve their apps.<sup>85</sup> Transit riders did not generally attribute the source of problems with third-party apps to the transit authority's data.<sup>86</sup> Control, of course, may be over more than just quality. Transit authorities may have concerns about liability for harm caused by faulty reuse of the data, or may be concerned about the impact on their reputations if data is used in ways that are embarrassing or scandalous.<sup>87</sup>

Concerns regarding the need to control the quality of “authoritative” information are not necessarily a good justification for refusing to make data open. For example, if app developers are determined to use certain types of data, in the absence of open data, they may use other methods to obtain the data that may make less reliable the data that is ultimately provided to the public.<sup>88</sup> Further, any problems with the quality of apps using open data, or with the quality of the data (which may trace back to its source) might be easily resolved and might not negatively impact the municipal data source.<sup>89</sup> In this respect, although quality control may be a motivating factor for asserting intellectual property rights, concerns over quality with open data may be misplaced. The growing comfort with less control is part of the institutional/cultural shift required for the evolution of open data.

---

85. ROJAS, *supra* note 8, at 33.

86. *Id.*

87. PESSOA ET AL., *supra* note 48, at 17. Liability as a risk of open data is also addressed as a challenge in MetroGIS. See METROGIS DATA PRODUCERS WORK GROUP, *supra* note 18, at 8; see also ROJAS, *supra* note 8, at 10. Embarrassment at chronic late arrivals might also result from greater transparency with respect to real-time GPS transit data. See, e.g., ROJAS, *supra* note 8, at 10, 38.

88. Developers might seek the data under access to information regimes and then scan it, or they might scrape data from transit system websites. See discussion *infra* Part IV.B.2.

89. It should be noted that control and integrity issues are also related to issues around the use of the trademarked names and logos of transit authorities in conjunction with the transit data. See PESSOA ET AL., *supra* note 48, at 20. Open licenses tend to place limits on the use of trademarks in order to ensure there is no implication that the transit authority is the source of the app that makes use of the data. Although these issues around trademarks are part of the broader transit data/intellectual property picture, they are beyond the scope of this Article.

Another motivation for transit authorities to restrict access to and use of transit data was a concern about lost potential for revenue generation.<sup>90</sup> With technologies for information dissemination in a state of rapid evolution, transit authorities were reluctant to take steps that would make it difficult for them to extract future revenue streams from the use of their data. For example, even as the Washington Metropolitan Area Transit Authority (WMATA) was making its static transit data open to developers, it commissioned a \$500,000 study to assess the commercial value of this data.<sup>91</sup> This value might be found in royalties charged for use of the data or from the sale of advertisements on electronic signage communicating information about the predicted arrival time of the next subway train.<sup>92</sup> Thus, another motivation to resist open data was the need of cash-strapped municipal agencies to balance their budgets or to recover the costs of their operations.<sup>93</sup> This too is likely to be a factor motivating resistance to opening other categories of government data. What is interesting to note in the transit data context is that open data policies evolved in the context of the less valuable static data, and many were in place as the more valuable real-time transit data became available.<sup>94</sup> An interesting question, therefore, is whether the open data movement would have had as much traction as it did, had the data initially sought been more obviously of commercial interest.

A final consideration—and one that will be discussed in further detail later in this Article—relates to the often complex relationships between governments and private sector companies from whom services are procured. The area of procurement is, indeed, a thorny

---

90. See ROJAS, *supra* note 8, at 22, 66.

91. Tim Jones, *Who Controls Data about Public Transportation?*, ELECTRONIC FRONTIER FOUND. (Sept. 10, 2009), <https://www EFF.ORG/deeplinks/2009/09/who-controls-data-about-public-transportation>; see ROJAS, *supra* note 8, at 66 (discussing Washington's approach to its transit data).

92. See *supra* notes 72–73 and accompanying text (providing more information on advertising programs).

93. See PESSOA ET AL., *supra* note 48, at 21 (identifying the high cost of information delivery as one of the challenges facing transit authorities). The report suggests that open transit data may actually reduce information delivery costs for transit authorities. *Id.*; see also Kaufman, *supra* note 65, at 2. At the same time, however, it is important to remember that producing regular and reliable open data also has its costs, whether it is done in house or through outsourcing. See PESSOA ET AL., *supra* note 48, at 16.

94. See McHugh, *supra* note 50 (providing an account of the development of the GTFS standard for static transit data); ROJAS, *supra* note 8, at 20–22 (stating these developments led to a greater push for open transit data).

one for open data generally.<sup>95</sup> Intellectual property considerations can become a barrier to open data where a department or agency has not been sufficiently attentive to the location of intellectual property rights (such as they may be) in any data generated through the supply of services by a private-sector company.<sup>96</sup> Issues may also arise where a municipality that contracts for data-related services permits the supplier to use its own proprietary software or standard for organizing or processing the data.<sup>97</sup> Doing so may complicate arguments around rights to access and to use the data. The use of proprietary standards may also make the data less fully useful than it would have been if it were made available in an open standard.<sup>98</sup>

### A. Copyright in Transit Data

One of the interesting issues with municipal transit data—and with open data more generally—is that the very concept of “openness” implies a lifting of intellectual property barriers to access and use. “Open,” after all, is the same term used in “open source” and “open access.”<sup>99</sup> Unlike computer software, or print-based works made available under “open access,” however, data is not protected under copyright law.<sup>100</sup> It is only protected to the extent that it is part of an

---

95. See, e.g., Steve Spiker, *Oakland and the Search for the Open City*, in *BEYOND TRANSPARENCY: OPEN DATA AND THE FUTURE OF CIVIC INNOVATION* 105, 120–21 (Brett Goldstein & Lauren Dyson eds., 2013); see also DIETRICH, *supra* note 12, at 4 (determining whether transit data can be made open “depends on the specific legal framework and on specific contracts between government authorities and the private companies undertaking the transport services”).

96. PESSOA ET AL., *supra* note 48, at 17; ROJAS *supra* note 8, at 22, 60. Rojas suggests that one of the reasons why Portland’s TriMet agency was an early innovator with open transit data was that it was a smaller agency and had therefore not outsourced the preparation of its static transit data. ROJAS *supra* note 8, 24. Of course, it may also be that a department or agency of government could deliberately choose to locate ‘ownership’ of any data with the private-sector company in order precisely to avoid access and transparency.

97. See *infra* notes 206–09 and accompanying text (discussing this issue in the context of the Bay Area Rapid Transit Authority (BART) and NextBus).

98. See discussion *infra* Part IV.B.3.

99. Tim G. Davies & Zainab Ashraf Bawa, *The Promises and Perils of Open Government Data (OGD)*, J. COMMUNITY INFORMATICS (2012), <http://ci-journal.net/index.php/ciej/article/view/929/926>. Davies and Bawa note that the open source movement emerged in part as a form of resistance to the extension of intellectual property rights to software. See *id.* They observe that “[i]n many settings, specific notions of ‘open’ are primarily articulated in opposition to some ‘closed’ sets of arrangements that are being challenged.” *Id.*

100. That copyright protects neither facts nor ideas is a basic principle of copyright law. This is discussed in more detail in this subsection. See *infra* notes 115–19 and accompanying text.

original compilation of data that is, itself, sufficiently original to qualify for protection.<sup>101</sup> By referring to it as “open data,” therefore, there is a semantic parallel between open data and other open works, even though the two are not equivalent. It is fundamentally important then, to take into consideration the rather weak footing from the outset, for claims to copyright in data.

Under the U.S. Copyright Act of 1976,<sup>102</sup> the federal government has no claim to copyright in the works it generates (with some exceptions),<sup>103</sup> and this would include compilations of facts. Thus, at the federal level in the United States, public domain is the default rule. This is in contrast with state and municipal governments, where the existence of intellectual property rights in works is the default.<sup>104</sup> Even those governments that have made data “open” under open data licenses have generally done so through the use of open licenses which are based on the starting premise that the government is the owner of copyright in the licensed work.<sup>105</sup> An open license does not mean that intellectual property rights do not exist; quite the contrary,

---

101. *Feist Publ'ns, Inc. v. Rural Tel. Serv. Co.*, 499 U.S. 340 (1991). This is discussed in greater detail below.

102. 17 U.S.C. §§ 101–1332 (2012).

103. *Id.* § 105 (“Copyright protection under this title is not available for *any work of the United States Government*, but the United States Government is not precluded from receiving and holding copyrights transferred to it by assignment, bequest, or otherwise.”) (emphasis added).

104. Thomas, *supra* note 25, at 4, (“[T]he majority rule appears to be that, unless prohibited by state law, state and local agencies may seek copyright protection for their works.”).

105. See *Open Government Licence*, NAT'L ARCHIVES, <http://www.nationalarchives.gov.uk/doc/open-government-licence/version/2/> (last visited Nov. 9, 2014) (beginning with an assertion of copyright and database rights in the license material). The terms of service for the Open New York data portal assert both specific and non-specific intellectual property rights in the data and documents made available through the portal. See *generally Terms of Service*, STATE OF N.Y., <https://data.ny.gov/> (last visited Nov. 9, 2014). At the municipal level, the City of Providence places its open data site under the same terms of service for its general city website, and those terms assert broad intellectual property rights. See *generally Terms of Use*, CITY OF PROVIDENCE, <https://www.providenceri.com/terms-of-use> (last visited Nov. 9, 2014). The Data Policies document for the San Francisco Open Data site indicates that intellectual property rights, in particular data sets, may be asserted on a case-by-case basis. See *generally Data Policies*, CITY & COUNTY OF S.F., <http://www6.sfgov.org/index.aspx?page=220> (last visited Nov. 9, 2014). This is also the case with the City of Chicago's open data portal. See *generally Data Terms of Use*, CITY OF CHI., [http://www.cityofchicago.org/city/en/narr/foia/data\\_disclaimer.html](http://www.cityofchicago.org/city/en/narr/foia/data_disclaimer.html) (last visited Nov. 9, 2014). The Open Data Portal for the City of Los Angeles is licensed under a Creative Commons Licence which is premised upon copyright in the licensed works. See *Control Panel*, SOCRATA, INC., <https://controllerdata.lacity.org/> (last visited Nov. 9, 2014).

open licenses often explicitly license works protected under copyright.<sup>106</sup>

Although state and municipal governments may claim copyright in their works, these works are nonetheless created in the public interest, using taxpayers' money. In the United States, the default rule of public domain for federal works may create a broader public sense of entitlement to all works created for public purposes using taxpayer dollars, notwithstanding the level of government that has generated the works.<sup>107</sup> Although, as we have seen, copyright can be asserted by governments to serve public purposes (such as cost recovery or quality control), it may be that when it comes to already weak copyright claims in data, the public status of a rights holder might weigh against a finding of even a thin copyright protection.<sup>108</sup> It should be noted as well that the incentive provided by the copyright monopoly is not needed to encourage transit authorities to produce schedule data or maps; these are generated as part of normal

---

106. It is quite common for open licenses to be premised upon the existence of copyright in the work being licensed. *See, e.g., About the Licenses*, CREATIVE COMMONS, <http://www.creativecommons.org> (last visited Nov. 9, 2014). The U.K. government open data license is also premised upon copyright and database rights in the licensed information. *See Open Government Licence, supra* note 105. The assertion of even weak property rights gives a broader level of control with respect to licensed data. While one might use public domain data available under the terms of a license, the license will only bind the parties to the contract. A third party is free to take and make use of the non-proprietary data. By contrast, a data set that is protected by copyright can be provided under license terms that limit aspects of its use; while a third party may not be bound by those contractual terms, they are still not free to make use of the intellectual property of another. To do so legally, they would have to seek out their own license.

107. This view may be becoming increasingly prevalent. *See* Tim O'Reilly, *Government as a Platform*, in *OPEN GOVERNMENT: COLLABORATION, TRANSPARENCY, AND PARTICIPATION IN PRACTICE* 11, 12 (Daniel Lathrop & Laurel Ruma eds., 2010) ("There is a new compact on the horizon: information produced by and on behalf of citizens is the lifeblood of the economy and the nation; government has a responsibility to treat that information as a national asset.").

108. There are other considerations which may weigh against a court finding copyright in transit data. *See* *N.Y. Mercantile Exch. Inc. v. IntercontinentalExchange, Inc.*, 497 F.3d 109 (2d Cir. 2007) (reacting negatively to a claim by the plaintiff to copyright in its data on a number of grounds). It is worth noting that the court also observed that because of the nature of its business, N.Y. Mercantile Exch. needed no copyright incentive to create the data since it not only needed to generate this data as part of its operations, but it was also required by law to do so. *See id.* at 118. In the case of transit data, it would be similarly possible to argue that transit authorities do not need copyright incentives to create transit timetables because their operations require them to do so. Such factors may carry some weight in a court's consideration of the merits of any copyright claim.



operations and are published to facilitate use of their systems by the public.<sup>109</sup>

Although case law in the United States has treated maps as compilations of data, in the past, revisions to the Copyright Act of 1976 categorized maps as “[p]ictorial and graphic works of authorship.”<sup>110</sup> Nevertheless, it remains the case that maps are generally considered to be fact-based works, and copyright in maps will not extend to the facts they represent.<sup>111</sup> However, to the extent that the graphic representation of the facts represents an original expression, that particular expression of the facts may be protected by copyright law.<sup>112</sup> The more basic the representation of the mapped “facts,” the greater the risk that the merger doctrine will apply to find that the authorial expression in the map has merged with the facts or ideas represented by the map. Transit maps tend to be fairly simple representations of routes, and the simpler the representation, the more likely a finding that the idea and expression has merged.<sup>113</sup> The copyright monopoly in a subway map, for example, would extend to the original elements of the graphic representation, but anyone wishing to convey subway stop information in map form would still be free to do so, so long as they do not copy the original design elements of the transit authority’s map.<sup>114</sup>

Since the decision of the U.S. Supreme Court in *Feist Publications, Inc. v. Rural Telephone Service Co.*,<sup>115</sup> it has been resoundingly clear that: (1) there is no copyright in facts; and (2) copyright in any

---

109. See *supra* note 69 and accompanying text.

110. 17 U.S.C. § 101 (2012) (“Pictorial, graphic, and sculptural works’ include two-dimensional and three-dimensional works of fine, graphic, and applied art, photographs, prints, and art reproductions, maps, globes, charts, diagrams, models, and technical drawings, including architectural plans.”).

111. See *Kern River Gas Transmission Corp. v. Coastal Corp.*, 899 F.2d 1458 (5th Cir. 1990) (holding that Kern River’s map markings indicating their proposed locations for a pipeline were not protected because such protection would grant them a monopoly over their proposed location ideas, “a foreclosure of competition that Congress could not have intended to sanction through copyright law”).

112. See *Mason v. Montgomery Data, Inc.*, 967 F.2d 135 (5th Cir. 1992); David B. Wolf, *New Landscape in the Copyright Protection for Maps: Mason v. Montgomery Data, Inc.*, 40 J. COPYRIGHT SOC’Y U.S.A. 401 (1992).

113. See, e.g., Katleen Janssen & Joseph Dumortier, *The Protection of Maps and Spatial Databases in Europe and the United States by Copyright and the Sui Generis Right*, 24 J. MARSHALL J. COMPUTER & INFO. L. 195, 210 (2006).

114. In one case where an app developer was challenged over his use of a subway map, he was reported to have decided to simply design his own version of the map rather than copy the authority’s map. See Wendy Seltzer, *The New Threat: Subway Map Sharing*, COPYRIGHT (Sept. 26, 2005), [http://copyright.corante.com/archives/2005/09/26/the\\_new\\_threat\\_subway\\_map\\_sharing.php](http://copyright.corante.com/archives/2005/09/26/the_new_threat_subway_map_sharing.php) (last visited Nov. 9, 2014).

115. 499 U.S. 340 (1991).

compilation of facts is “thin.”<sup>116</sup> Facts are regarded as being part of the public domain, and therefore available to all.<sup>117</sup> Different theories can justify this inability to monopolize facts. In *Feist*, Justice O’Connor suggested that it was because facts were incapable of originality; they were copied from the world around us.<sup>118</sup> On this view, facts are not capable of authorship and thus no one may acquire an intellectual property monopoly in facts simply by recording them. According to Justice O’Connor, “facts do not owe their origin to an act of authorship. The distinction is one between creation and discovery.”<sup>119</sup>

From a public policy perspective, it would make no sense to permit authors to monopolize facts because to do so would hamper the diffusion and dissemination of knowledge, and the creation of new works.<sup>120</sup> Monopolies over facts would also be economically inefficient.<sup>121</sup> While “sweat of the brow” or “industrious collection”<sup>122</sup> doctrines allowed for a more robust protection for factual compilations, they also required competitors to waste resources on re-gathering facts from scratch, rather than expending their efforts on finding new and innovative ways to present those facts to the public or to apply those facts to new and useful solutions to problems.

---

116. *Id.* at 349.

117. *Id.* at 348.

118. *Id.* at 347.

119. *Id.* at 347. The argument that facts are “discovered” and not created has been criticized. *See, e.g.*, Wendy J. Gordon, *Reality as Artifact: From Feist to Fair Use*, 55 L. & CONTEMP. PROBS. 93 (1992). Even scientific ‘facts’ are sometimes later disproved. In this sense, some facts are closer to theories. Of course, theories are close to ideas, and ideas are also not subject to copyright protection. *See id.* In some cases, courts have declined to give copyright protection to ‘facts’ that are little more than speculations. *See, e.g.*, *Nash v. CBS Inc.*, 899 F.2d 1537 (7th Cir. 1990) (declining to find that a movie based on a book’s theory that John Dillinger was alive and living on the West Coast, infringed copyright in the book); *see also* *Hoehling v. Universal City Studios Inc.*, 618 F.2d 972, 978 (2d Cir. 1980) (“To avoid a chilling effect on authors who contemplate tackling an historical issue or event, broad latitude must be granted to subsequent authors who make use of historical subject matter, including theories or plots.”).

120. *See, e.g.*, Miriam Bitton, *Feist, Facts and Functions: Historical Perspective*, in INTELLECTUAL PROPERTY PROTECTION OF FACT-BASED WORKS: COPYRIGHT AND ITS ALTERNATIVES 3, 16–17 (Robert F. Brauneis ed., 2009).

121. *See* *Harper & Row, Publishers, Inc. v. Nation Enters.*, 471 U.S. 539, 556–57 (Brennan, J., dissenting).

122. *Feist*, 499 U.S. at 352 (1991) (explaining that for these theories the “underlying notion was that copyright was a reward for the hard work that went into compiling facts”). The Court in *Feist* unanimously rejected “sweat of the brow” as a basis for copyright protection. *Id.* at 354.

While copyright will not protect facts, it will protect an original expression of those facts.<sup>123</sup> In a compilation of facts, the original expression will lie in the selection or arrangement of those facts.<sup>124</sup> Anyone asserting copyright in a compilation of facts must therefore be prepared to demonstrate that they have achieved an original selection or arrangement. In *Feist*, the U.S. Supreme Court found that there was no originality in the selection or arrangement of facts in a telephone directory.<sup>125</sup> The selection was dictated by the nature of the telephone directory (a compilation of all names, addresses and phone numbers of subscribers who had not opted out of listing) and, as the court noted, there “is nothing remotely creative about arranging names alphabetically in a white pages directory. It is an age-old practice, firmly rooted in tradition and so commonplace that it has come to be expected as a matter of course.”<sup>126</sup> Since there was no possible claim to copyright in the facts, and no original selection or arrangement of facts, the directories could not be protected by copyright law.<sup>127</sup>

Since *Feist*, other courts have grappled with what constitutes an original selection or arrangement of facts.<sup>128</sup> In some instances, judges have been prepared to find either an original selection or an original arrangement, or both. For example, while a basic telephone directory is not capable of copyright protection, a specialized directory—one that is a compilation of information about certain types of businesses, arranged thematically, for instance—might represent an original work of authorship as a result of an author’s choices as to what businesses to include, and what categories or themes to use in order to organize the information.<sup>129</sup> In *Matthew Bender & Co., Inc. v. West Publishing Co.*, the Court suggested the following guidelines: “creativity in selection and arrangement,

---

123. *Id.* at 345; see also Bitton, *supra* note 120.

124. *Feist*, 499 U.S. at 348.

125. *Id.* at 362–63.

126. *Id.*

127. *Id.* at 364.

128. See Michael Steven Green, *Two Fallacies About Copyrighting Factual Compilations*, in INTELLECTUAL PROPERTY PROTECTION OF FACT-BASED WORKS: COPYRIGHT AND ITS ALTERNATIVES 109 (Robert F. Brauneis ed., 2009).

129. See *Key Publ’ns, Inc. v. Chinatown Today Publ’g Enters., Inc.*, 945 F.2d 509 (2d Cir. 1991); see generally David E. Shipley, *Thin But Not Anorexic: Copyright Protection for Compilations and Other Fact Works*, 15 J. INTELL. PROP. L. 91 (2007). Similarly, a compilation that established prices for used cars was found to have sufficient originality because of the degree of judgment required to arrive at the prices. See *CCC Info. Servs. v. Maclean Hunter Mkt. Reports*, 44 F.3d 61 (2d Cir 1994).

therefore, is a function of (i) the total number of options available, (ii) external factors that limit the viability of certain options and render others non-creative, and (iii) prior uses that render certain selections ‘garden variety.’”<sup>130</sup> However, the underlying data or information is still part of the public domain. Another person might use that data in their own work, so long as they did not copy a substantial part of either the selection or the arrangement.

Because the idea of “selection” requires some conscious choice by the author, the requirement for an original selection of data may not be met when a compilation of data is a “whole universe” selection. Applying this reasoning to transit data, a data set consisting of all bus stop times in a timetable for a given route, or all bus stop times for all routes, does not result from an original selection. Many open data sets, including municipal transit data, will not reflect an original selection simply because they contain all the available data.<sup>131</sup> Such data sets might still be considered capable of being protected by copyright if they reflected an original *arrangement* of the data.<sup>132</sup> However, timetable data organized chronologically does not reflect an original arrangement, nor does data organized according to a standard specification, such as the GTFS.<sup>133</sup> Data within a database similarly may not reflect an original arrangement.<sup>134</sup>

Real-time data pose their own particular copyright challenge, since this type of data are gathered in real-time, and, as a result, do not represent a static or finite collection. Although such data may be stored in a database for later analysis (as, for example, to assess the overall performance of the transit system), such data, generated and used in real-time, may simply be too ephemeral to constitute a

---

130. *Matthew Bender & Co. v. West Publishing Co.*, 158 F.3d 674, 682–83 (2d Cir. 1998).

131. *See, e.g.*, Judge & Scassa, *supra* note 84.

132. *See Feist*, 499 U.S. at 352.

133. The chronological organization of timetable data is an arrangement that lacks originality in the same way that the alphabetical organization of telephone directory information does. *See id.* at 362–63. Where data is organized according to an external standard, it will similarly lack an original arrangement since the arrangement is due to the standard rather than to any creative spark from the “author” of the compilation.

134. *See, e.g.*, Green, *supra* note 125, at 115 (suggesting that under the *Feist* approach, “an electronic database that simply stores facts in a raw form, allowing them to be searched and organized by the consumer” may not reflect any “arrangement” that would qualify for copyright protection); Thomas, *supra* note 25, at 5. For a case that did find copyright to subsist in the *structure* of a database (though not in the underlying data), see *Assessment Techs. of Wis., LLC v. WIREdata, Inc.*, 350 F.3d 640 (7th Cir. 2003).

“work” in which copyright would subsist.<sup>135</sup> In a comprehensive study of legal rights in real-time data, Larry Thomas concludes that such data are not copyrightable.<sup>136</sup>

There has been some litigation with respect to so-called “original facts”<sup>137</sup> or “creative facts”<sup>138</sup>—things that are only facts because they have been generated by human creativity. For example, the rights holders in the hugely popular *Seinfeld* series were successful in asserting their copyright in the series against a company that had created a *Seinfeld* trivia game.<sup>139</sup> Although the game involved questions and answers regarding events that occurred in different episodes of the television series (the “facts” of the show), the trivia game was nonetheless considered to be a derivative work that exploited the original work.<sup>140</sup> A similar result was reached in a case dealing with a Lexicon built around J.K. Rowling’s *Harry Potter* series of novels.<sup>141</sup> Although characters, events, creatures, and places in the novels were “facts” of a sort, they were only facts within the context of the overall creative work.<sup>142</sup>

This recognition of the ability to assert copyright over “creative” facts has spilled over into the data context in cases where plaintiffs have asserted that the “facts” within a given compilation actually owe their origin to the plaintiffs, and are therefore capable of copyright protection in and of themselves.<sup>143</sup> For example, in *New York*

---

135. See Thomas, *supra* note 25, at 6.

136. *Id.*

137. See Teresa Scassa, *Original Facts: Skill, Judgment and the Public Domain*, 51 MCGILL L.J. 253 (2006); see also Alan L. Durham, *Speaking of the World: Fact, Opinion and the Originality Standard of Copyright*, in INTELLECTUAL PROPERTY PROTECTION OF FACT-BASED WORKS: COPYRIGHT AND ITS ALTERNATIVES 133 (Robert F. Brauneis ed., 2009) (discussing other borderline “facts” such as conjecture, opinions, and predictions).

138. See Justin Hughes, *Created Facts and the Flawed Ontology of Copyright Law*, 83 NOTRE DAME L. REV. 43 (2007) [hereinafter Hughes, *Flawed Ontology*]; Justin Hughes, *Created Facts and Their Awkward Place in Copyright Law*, in INTELLECTUAL PROPERTY PROTECTION OF FACT-BASED WORKS: COPYRIGHT AND ITS ALTERNATIVES 186 (Robert F. Brauneis ed., 2009) [hereinafter Hughes, *Awkward Place*].

139. *Castle Rock Entm’t, Inc. v. Carol Publ’g Grp., Inc.*, 150 F.3d 132 (2d Cir. 1998).

140. *Id.* at 139.

141. *Warner Bros. Entm’t Inc. v. RDR Books*, 575 F. Supp. 2d 513, 536 (S.D.N.Y. 2008).

142. See *id.*

143. See, e.g., *N.Y. Mercantile Exch., Inc. v. IntercontinentalExchange, Inc.*, 497 F.3d 109 (2d Cir. 2006); *BanxCorp v. Costco Wholesale Corp.*, 978 F. Supp. 2d 280 (S.D.N.Y. 2013); *RBC Nice Bearings, Inc. v. Peer Bearing Co.*, 676 F. Supp. 2d 9, 21 (D.Conn. 2009).

*Mercantile Exchange, Inc. v. IntercontinentalExchange, Inc.*,<sup>144</sup> the plaintiffs asserted copyright in the real-time data they generated which could be used to fix values for futures contracts for natural gas and crude oil.<sup>145</sup> These values would shift and change according to a web of different factors.<sup>146</sup> The plaintiffs argued that the data they generated were essentially original, and should be protected.<sup>147</sup> The Second Circuit Court of Appeals declined to decide this point, calling it a “close question,” and ruled instead against the plaintiffs on other grounds.<sup>148</sup> These grounds themselves are important—the court was prepared to apply the merger doctrine to the generated data.<sup>149</sup> In other words, even if the facts were considered “original” expressions that could be protected by copyright, where an idea and its expression are so closely merged that there is no other reasonable way to express the idea, there will be no copyright monopoly.<sup>150</sup> The court was therefore of the view that the plaintiffs would have had to show that their calculations of values for the contracts would be sufficiently different from any other reasonable calculations.<sup>151</sup> Since the goal of such calculations is to arrive at an accurate figure (and one based on more or less the same inputs), merger would be a significant problem in this area.<sup>152</sup>

Arguments about computer-generated data raise other copyright issues as well. Even if it is possible to argue that the data generated by the application of an algorithm to input data is itself original, there is also a potential argument that the “authorship,” which is so

---

144. *N.Y. Mercantile Exch., Inc.*, 497 F.3d at 112.

145. *See id.* at 110–13; *see also* Thomas, *supra* note 25, at 10 (acknowledging that it could be argued that predictions based on real-time transit data are not discoveries of existing facts, but skeptical that such data could be protected by copyright).

146. *N.Y. Mercantile Exch., Inc.*, 497 F.3d at 110–11.

147. *Id.* at 113.

148. *Id.* at 114 (“While the line between creation and discovery is often clear-cut, we recognize that it is a difficult line to draw in this case.”). Nevertheless, the court was prepared to accept that in some cases, where relatively little input data was available, the settlement prices were closer to predictions, which in turn “appears closer to creation.” *Id.* at 116.

149. *Id.*

150. *Id.* at 116–17 (“[E]ven expression is not protected in those instances where there is only one or so few ways of expressing an idea that protection of the expression would effectively accord protection to the idea itself.”) (quoting *Kregos v. Associated Press*, 937 F.2d 700, 705 (2d Cir. 1991)).

151. *Id.* at 115 (“[T]here is one proper settlement price; other seemingly-accurate prices are mistakes which actually overvalue or undervalue the futures contract.”).

152. *Id.* at 118.

essential to copyright protection, is lacking.<sup>153</sup> In an Australian case involving telephone directories compiled using highly automated processes, the High Court of Australia observed that “[a]uthorship and originality are correlatives.”<sup>154</sup> In other words, a work without an author could not be original. The Court ruled that in order to find copyright to subsist in a work, it is necessary first to identify the author or authors and then to assess the extent of their contributions to the work.<sup>155</sup> In the case of a work created by a highly automated process, it might be difficult to identify any particular “authorial” contribution to the selection or arrangement.<sup>156</sup> If this were the case, copyright protection would be unavailable. Similar principles would likely apply in the U.S. context.<sup>157</sup> Essentially, originality in a compilation of data requires some spark of creativity or authorial judgment. A process by which selection or arrangement of data occurs according to pre-set parameters may lack the necessary authorship.<sup>158</sup>

This brief review of copyright principles in relation to facts reveals that, under copyright law, facts themselves are not capable of copyright protection. Only an original selection or arrangement of facts can be protected, but it is an open question whether any given selection or arrangement will qualify for copyright protection. While a map, as a visual representation of facts, may be protected by copyright (even though the underlying facts remain in the public domain)<sup>159</sup> the expressive dimension is significantly diminished in a compilation of facts. Further, automated processes for collection or arrangement may remove authorship from any resulting compilation, and thus undermine the potential for any claim over copyright in the

---

153. Thomas, *supra* note 25, at 7 (identifying a lack of authorship as a particular problem in the case of real-time data, which are generated by the interaction of hardware and software).

154. *Telstra Corp. Ltd. v Phone Directories Co.* (2010) 264 ALR 617, ¶ 344 (Austl.).

155. *Id.* at ¶ 28.

156. In many instances, even static transit timetable data is produced via an automated process using the proprietary systems of third-party companies. *See, e.g., ROJAS, supra* note 8, at 23.

157. *E.g.* Thomas, *supra* note 25, at 7 (suggesting that real-time transit data lack the element of authorship necessary to support copyright).

158. *Id.* at 5.

159. The artistic or graphic component of a map reflects the original expression that can be protected by copyright. However, maps are based upon geographical facts such as the location of roads, watercourses, or other features. Copyright in a map does not give the mapmaker a monopoly over the facts represented in the map. *See Feist Publ'ns, Inc. v. Rural Tel. Serv. Co.*, 499 U.S. 340 (1991).

result.<sup>160</sup> While some scholars have argued that certain types of data are capable of being “original” in their own right,<sup>161</sup> these arguments have their greatest success in the context of facts derived from works that are fictional in their entirety.<sup>162</sup> With respect to data generated in order to help predict or understand phenomena in the real world, these arguments are much weaker, and run up against other barriers to copyright protection such as the doctrine of merger.

## B. Transit Data Intellectual Property Skirmishes

Notwithstanding the rather weak copyright claims in relation to transit data, such claims have been regularly asserted. In this section, we consider some of these claims, the stated reasons for making them, and their role in the evolution towards open transit data.

### 1. Transit Maps

The first skirmishes over transit data were in relation to non-interactive transit maps. Such maps are typically static. A map of a subway system for example, would only need updating if a new subway station were added (something that would be a relatively rare occurrence) or if a subway stop were renamed (also relatively rare). Bus route maps might change more frequently, but even so, the changes would be relatively few, and generally at easily predictable intervals (for example, seasonal changes).

In 2005, at the very outset of a new era in mobile digital technology, New York City’s Metropolitan Transportation Authority (MTA) sent a cease and desist letter<sup>163</sup> to William Bright, a developer who had copied the subway map from the MTA website, and adapted it for viewing on the iPod.<sup>164</sup> Bright’s goal was to allow users to access

---

160. *Telstra Corp. Ltd. v Phone Directories Co.* (2010) 264 ALR 617, ¶ 338 (Austl.).

161. See generally, e.g., Durham, *supra* note 137; Hughes, *Flawed Ontology*, *supra* note 138; Hughes, *Awkward Place*, *supra* note 138; Scassa, *supra* note 137.

162. See generally, e.g., *Castle Rock Entm’t, Inc. v. Carol Publ’g Grp., Inc.*, 150 F.3d 132 (2d Cir. 1998); *Warner Bros. Entm’t Inc. v. RDR Books*, 575 F. Supp. 2d 513 (S.D.N.Y. 2008).

163. See *MTA Wants to Dislocate iPod Subway Maps*, CHILLING EFFECTS (Sept. 14, 2005), <https://www.chillingeffects.org/copyright/notice.cgi?NoticeID=2369> (providing a copy of the cease and desist letter that the MTA sent to iPodSubwayMaps.com).

164. See Mike Musgrove, *Lost Underground? Check your iPod*, WASH. POST (Oct. 4, 2005), <http://www.washingtonpost.com/wp-dyn/content/article/2005/10/03/AR2005100301581.html>. The first iPod capable of showing videos and photos was released in 2005. See Apple Press Info, *iPod and iTunes Timeline*, APPLE, <http://www.apple.com/pr/products/ipodhistory/> (last visited Nov. 13, 2014). Mobile



subway information while on the go.<sup>165</sup> A similar cease and desist letter was sent to Bright by the Bay Area Rapid Transit authority (BART) in the same year, also in relation to a subway map designed for use on the iPod.<sup>166</sup>

Both cease and desist letters asserted copyright in the maps.<sup>167</sup> While the MTA letter was relatively terse, the BART letter was more specific about its justification for asserting its rights. The letter stated: “BART is concerned that the unauthorized copying of content from the BART website will mislead consumers by providing inaccurate information with the implication that it is official BART information.”<sup>168</sup> The asserted concern in this case was not about interference with any revenue streams; rather, it was about the accuracy of the posted information. This element of control—the perceived need to maintain control over information flowing to the public in order to ensure that only accurate information is disseminated—is a recurring theme in the context of government data more generally.<sup>169</sup>

---

phone use was not widespread in North America in 2005, although the BlackBerry device was rapidly gaining in popularity. See *BlackBerry Timeline: A Look Back at the Tech Company's History*, GLOBAL NEWS (Sept. 24, 2013), <http://globalnews.ca/news/860689/blackberry-timeline-a-look-back-at-the-tech-companys-history/>. The first generation iPhone was not launched until 2007. *iPod and iTunes Timeline*, *supra* note 164.

165. Bright created a repository of transit system maps viewable on the iPod. See Musgrove, *supra* note 164.

166. See *BART Wants Rapid Takedown of iPod Subway Maps*, CHILLING EFFECTS (Sept. 1, 2005), <https://www.chillingeffects.org/copyright/notice.cgi?NoticeID=2368> (providing a copy of the cease and desist letter that BART sent to iPodSubwayMaps.com).

167. See *id.*; *MTA Wants to Dislocate iPod Subway Maps*, *supra* note 163.

168. *BART Wants Rapid Takedown of iPod Subway Maps*, *supra* note 166.

169. For example, one reason governments have been slow to make use of volunteered geographic information has been a concern over the quality of such data and their impact on their authoritative data sets. See, e.g., DIETRICH, *supra* note 12; Michael F. Goodchild & Linna Li, *Assuring the Quality of Volunteered Geographic Information*, 1 *Spatial Statistics* 110 (2012); Peter A. Johnson & Renee E. Sieber, *Situating the Adoption of VGI by Government*, in *GROUNDING GEOGRAPHIC KNOWLEDGE: VOLUNTEERED GEOGRAPHIC INFORMATION (VGI) IN THEORY AND PRACTICE* 65 (Daniel Sui et al. eds., 2013). Copyright battles over transit maps are, of course, not unique to the United States. See, e.g., DIETRICH, *supra* note 12, at 7 (reporting that a startup company in France was forced to remove a map of the Paris subway system from a mobile app after the system operator asserted its copyright in the map). More recently, in 2010, New York's MTA was reported to have sent out a cease and desist letter over the use of the subway map on a dress. See Flegenheimer, *supra* note 78. In this latter instance, the issue appears more to be one related to the commercialization of trademarks, logos, and symbols of the MTA than it is a matter of exercising a monopoly over the represented information. See *id.* Note that trademark issues are latent in the BART letter, as well. *BART Wants Rapid*

What can be drawn from these early skirmishes over transit data is that public transit information was carefully guarded by municipal transit authorities that saw their mandate as including oversight to ensure that they were the sole authorized source of information about transit operations.<sup>170</sup> What is also clear is that intellectual property rights—in these cases copyright<sup>171</sup>—were the means used to assert that control.

## 2. *Static Transit Data*

The next generation of battles over transit data was related to static transit data—essentially, timetable information. Prior to the development of mobile technologies, transit timetable information was primarily made available to transit users through paper schedules, or through online versions of those paper schedules.<sup>172</sup> Some innovation was happening in the form of web-based trip planners that would allow users to go online to garner information about specific trips they wished to take.<sup>173</sup> Transit authorities developed phone and text message services that would allow users to contact the transit authority to get specific schedule information about, for example, a particular bus stop.<sup>174</sup> With the emergence of the iPhone and other smart mobile devices, two phenomena coincided. First, there was a much broader demand from the public for information that could be easily accessed from a smart phone.

---

*Takedown of iPod Subway Maps*, *supra* note 166. The concern over inaccurate information was combined with the concern that this information wrongfully attributed BART as the source. *Id.*

170. As noted earlier, in those jurisdictions with Crown copyright, this copyright is sometimes asserted as a means of control over quality or authoritativeness. See Judge, *supra* note 84; see also Teresa Scassa, *The Best Things in Law are Free: Towards Quality Free Public Access to Primary Legal Materials in Canada*, 23 DALHOUSIE L.J. 301, 321–22 (2000).

171. BART also asserted trademark rights. See *BART Wants Rapid Takedown of iPod Subway Maps*, *supra* note 166.

172. PESSOA ET AL., *supra* note 48, at 10. New York City's MTA licenses data while restricting the use of its trademarks. See *Developer Resources*, MTA.INFO, <http://web.mta.info/developers/> (last visited Nov. 13, 2014).

173. See, e.g., *Rider Tools*, MASS. BAY TRANSP. AUTHORITY, [http://www.mbta.com/rider\\_tools/trip\\_planner/](http://www.mbta.com/rider_tools/trip_planner/) (last visited Nov. 9, 2014); *Trip Planner*, MTA.INFO, [http://tripplanner.mta.info/MyTrip/ui\\_web/customplanner/tripplanner.aspx](http://tripplanner.mta.info/MyTrip/ui_web/customplanner/tripplanner.aspx) (last visited Nov. 9, 2014).

174. For example, in May 2011, New Jersey's NJ Transit announced a new text messaging service that would provide train schedule information. See *NJ Transit Train Schedules Now Available Via Text Message*, N.J. TRANSIT (May 11, 2011), [http://www.njtransit.com/tm/tm\\_servlet.srv?hdnPageAction=PressReleaseTo&PRESS\\_RELEASE\\_ID=2673](http://www.njtransit.com/tm/tm_servlet.srv?hdnPageAction=PressReleaseTo&PRESS_RELEASE_ID=2673).

This was not just whole timetables that could be viewed through a tiny browser window. Rather, the public sought apps that would permit the user to select the desired information and to access it in a useful format.<sup>175</sup> The second phenomenon was the encouragement, initially by Apple, but later by other companies such as Google, of the development of apps by the community of users of mobile devices.<sup>176</sup> Users were encouraged not only by the availability of tools and information for app development, but also by a platform from which (conforming) apps could be distributed freely or for a price, as provided by Apple's App Store (and later the Android Market).<sup>177</sup>

Because static transit data generally do not change with great frequency, and because changes come at predictable intervals, this type of data presents relatively few challenges for application developers. Early difficulties with the use of static transit data were chiefly the result of the assertion by some transit authorities of copyright in their timetables and in the underlying data.<sup>178</sup> Transit authorities resisted the sharing of their timetable data in different ways. One way was simply to deny access to these data in reusable formats. Unlike subway maps, timetable data in a large municipality would be voluminous, and subject to periodic change. If it were not made available in a digital reusable format, reducing the data to such a format would be a time consuming barrier for many developers.<sup>179</sup> In the early days of transit app development, some developers chose to obtain transit data either through access to information requests or

---

175. PESSOA ET AL., *supra* note 48, at 10.

176. Kim W. Tracey, *Mobile Application Development Experiences On Apple's iOS and Android OS*, IEEE Potentials, July–Aug. 2012, at 30, 30–31, available at <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6248786>. Tracey suggests that the provision of easy-to-access app marketplaces by both Apple and Google and support by both companies for app development were factors in the proliferation of apps for these platforms. *Id.*

177. *Id.*

178. See Martin B. Cassidy, *MTA Clash Over Train Schedules on iPhone*, STAMFORD ADVOC. (Aug. 31, 2009), <http://www.stamfordadvocate.com/news/article/Blogger-MTA-clash-over-train-schedules-on-iPhone-1691.php> (describing how New York's MTA sent a cease-and-desist letter to a developer who created an app featuring commuter rail timetables). Washington's WMATA encumbered the use of certain schedule and route data with restrictive terms of use, even when the agency made said data available for public download in 2009. ROJAS, *supra* note 8, at 66.

179. See ROJAS, *supra* note 8, at 20–22 (discussing the growing demand by developers for transit data in reusable formats, which could be used much more quickly and efficiently, and are less likely to contain the errors that might be present in data that was scraped from websites or scanned from paper documents).

by scraping the data from transit authority websites.<sup>180</sup> Ironically, as stated earlier, such practices were much more likely to introduce errors and inaccuracies—the very things the municipalities sought to avoid by refusing to release the data as open data.<sup>181</sup>

Some municipalities that were initially unwilling to share their data with app developers under open licenses nonetheless chose to work with Google to have their data embedded in Google Maps.<sup>182</sup> As transit authorities began to adopt the GTFS, app developers also started to adopt and use this standard for their own app projects.<sup>183</sup> It should be noted that the choice of an open standard for data, as opposed to a closed or proprietary standard, is another piece of the open data and IP puzzle. When those municipalities who used the GTFS because of their relationship with Google later decided to make their data open to developers, the fact that they had organized their data according to a commonly used and non-proprietary standard was an important benefit.<sup>184</sup>

---

180. *See, e.g.*, PESSOA ET AL., *supra* note 48, at 8; ROJAS, *supra* note 8, at 20–21. Data scraping has also been used in Europe in cases where open data are not available from transit authorities. *See* DIETRICH, *supra* note 12, at 14. The Supreme Court of British Columbia defines scraping as a “form of indexing that looks for specific information located in known positions on selected web pages with known layouts. Used particularly to build specialized websites combining information from other websites.” *Century 21 Canada Limited Partnership v. Rogers Communications Inc.*, 2011 BCSC 1196, ¶ 10 (Can.).

181. PESSOA ET AL., *supra* note 48, at 8, 19.

182. Not all municipalities were initially willing to embed their data with Google Maps. *See, e.g.*, McHugh, *supra* note 50, at 127; Perkins, *supra* note 57. However, this initial reluctance was overcome as the public interest in Google Transit became evident. *See* McHugh, *supra* note 50, at 127. McHugh writes, “[t]he biggest advantage of being part of the GTFS standard for agencies is that their information appears in a global set of search products that are easy to use and visited by millions and millions of people every day.” *Id.* at 130. McHugh notes as well that users of Google Transit are familiar with the interface as they move from city to city, and can also find other relevant information while using Google Maps. *Id.* In a 2011 report, more than 125 U.S. transit agencies had incorporated their transit data into Google Maps. HILLSMAN & BARBEAU, *supra* note 48, at 2.

183. *See* ROJAS, *supra* note 8, at 24–25.

184. In 2011, a study found that almost twenty-five percent of transit authorities in the United States used the GTFS to publish their static timetable data. PESSOA ET AL., *supra* note 48, at 7. City-Go-Round equates open transit data with data made available in GTFS format. CITY-GO-ROUND, *supra* note 45. It maintains a list of agencies that make their data available in this format. *Id.* This list is in turn derived from the GTFS Data Exchange, <http://www.gtfs-data-exchange.com/>, which is a site designed to facilitate the sharing and retrieval of transit data in GTFS format by app developers and municipalities. *See* HILLSMAN & BARBEAU, *supra* note 48, at 66. The GTFS exchange acts as a kind of open transit data repository and lists data in GTFS format from more than 700 transit authorities world-wide. *Id.* at v.

As transit data became open, developers in cities across the United States were motivated to develop apps for static transit data. In some cases, these motivations were very personal—the developer sought to create something that would be of use to him or her personally, or to a friend or loved one who was a transit user.<sup>185</sup> In other cases, the goals were more mixed, ranging from commitment to developing useful apps, to personal skill development and even commercialization.<sup>186</sup> It should be noted that not all apps that use transit data will be exclusively public transit apps. For example, the open source OpenTripPlanner<sup>187</sup> integrates transit data with trip planning information for both pedestrians and cyclists.<sup>188</sup> Another app, Walk Score,<sup>189</sup> helps apartment hunters to make choices based on the availability of public transit.<sup>190</sup> In a sense, Google's use of transit data is a precursor to these types of apps. The incorporation of transit data into Google Maps was not so much about the creation of a transit data information tool as it was about creating a rich, multi-level information tool. In this way, transit data become interwoven with other information about particular urban areas. This shifts the focus from transit data as a proprietary data set under the control of transit authorities, with a primary relevance to public transit, to transit data as one of a number of interlinked data sets that are part of a broader narrative about life in a given municipality.

In a modern, high volume, urban transit system, transit timetables are produced with the aid of computer software, although transit authorities must provide the necessary parameters for the timetables, including data regarding frequency, peak hours, and so on.<sup>191</sup> One

---

185. See, e.g., ROJAS, *supra* note 8, at 20–21. Factors motivating developers to create transit apps were multiple and profit was generally the least motivating factor. *Id.* at 32.

186. See Birgitta Bervall-Kåreborn & Debra Howcroft, *Mobile Applications Development on Apple and Google Platforms*, 29 COMM. ASS'N FOR INFO. SYSTEMS 566, 574–75 & tbl.4 (2011), available at <http://aisel.aisnet.org/cais/vol29/iss1/30/> (discussing motivations of app developers); see also ROJAS, *supra* note 8, at 32 (discussing motivations in the particular context of transit data app developers).

187. *Multimodal Trip Planning & Analysis*, OPENTRIPPLANNER, <http://www.opentripplanner.org/> (last visited Nov. 9, 2014).

188. *Id.* See generally HILLSMAN & BARBEAU, *supra* note 48 (studying multi-modal transit data).

189. *Apartments and Rentals—Find a Walkable Place to Live*, WALK SCORE, <http://www.walkscore.com/apartments/> (last visited Nov. 9, 2014).

190. See *id.*

191. See Christopher MacKechnie, *Designing Bus Routes and Schedules Part IV: Writing the Bus Schedule*, ABOUT.COM, [http://publictransport.about.com/od/Transit\\_Planning/a/Designing-Bus-Routes-And-Schedules-Part-Iv-Writing-The-Bus-Schedule.htm](http://publictransport.about.com/od/Transit_Planning/a/Designing-Bus-Routes-And-Schedules-Part-Iv-Writing-The-Bus-Schedule.htm) (last visited Nov. 9, 2014).

could argue that the schedules are an expression of this tremendous intellectual effort (admittedly performed substantially with the aid of computers) and are thus copyright protected works. On this view, the schedules do not record observations regarding the behavior of buses (observable facts) but rather reflect a series of choices by the transit authority regarding how different bus departure times and frequencies should be organized so as to meet the needs of the traveling public. Yet, this argument would be difficult to sustain from a legal point of view. In the first place, the planning effort necessary to set a transit schedule is now largely automated, and new schedules are generated by tweaking certain parameters of existing schedules.<sup>192</sup> It is not clear that there is a sufficient exercise of authorship, nor is it clear that there is even an identifiable author for copyright protection to be available.

There is also a distinction between the planning exercise and the printed timetables. Once drawn up, transit timetables reflect a series of transit facts: they tell us when certain buses or trains are scheduled to appear at particular stops. Thus, the schedules are, in essence, a collection of stop times; that is, a compilation of facts. This view separates the planning exercise necessary to operate a transit system (which does not require the incentive of copyright protection in order to occur) from the publication of the timetables, which provide information about those operations. Copyright protection for transit timetables is completely unrelated to their production—they would be generated regardless of whether copyright was available. There is simply no rationale to support copyright protection in these circumstances.<sup>193</sup>

While it is still possible that claims to copyright in bus timetables as a whole would be recognized by courts, the protection for these compilations would necessarily be “thin” in the sense discussed above, and would not extend to the underlying data. Furthermore, the circumstances in which a municipality would draw any advantage from enforcing such copyrights are unclear. Bus timetables are functional works whose goal is to provide information that makes it

---

192. *Id.*

193. *Cf.* Feist Publ'ns, Inc. v. Rural Tel. Serv. Co., 499 U.S. 340 (1991). The works at issue were telephone directories, which were generated by the plaintiff telephone company as an essential part of its operations. *Id.* at 342. Justice O'Connor observed that the statutory requirement to publish subscriber information supported the view that the selection of published facts was not original. *Id.* at 363. Although not necessarily dictated by law, the information in a bus timetable is dictated by necessity—the timetable cannot serve its function if it does not provide bus stop arrival times.

easier for those seeking to use public transit to do so effectively. A broad dissemination of this information is in the interest of the transit authority, which wants to encourage ridership.

Nevertheless, municipalities have not hesitated to claim copyright in their transit schedules.<sup>194</sup> Echoing its approach in the earlier era of protection of transit maps, the New York MTA asserted this type of claim in 2009, threatening an app developer with legal action if he did not license the data he was using for \$5000 plus royalties.<sup>195</sup> When the developer did not comply with the demand, the MTA issued a takedown notice to the Apple App Store, which complied and removed the app.<sup>196</sup> Conflicts over schedule data arose in other jurisdictions as well,<sup>197</sup> and while many transit authorities have since made these data available as open data, there are still a large number of municipalities that do not.<sup>198</sup> As with map data, one consideration has been the perceived need to control the flow of information about

---

194. Even so, there would still be no copyright in the underlying data, and it would not be copyright infringement for a developer to take that data and make it available to the public in a different form.

195. See Jones, *supra* note 91; Benjamin Kabak, *MTA Struggling in an Age of Open Information*, 2ND AVE. SAGAS (Sept. 14, 2009), <http://secondavenuesagas.com/2009/09/14/mta-struggling-in-an-age-of-open-information>; see also PESSOA ET AL., *supra* note 48, at 18–19.

196. See Jones, *supra* note 91; Kabak, *supra* note 195. The notice and takedown system is a powerful weapon in the hands of copyright owners. It can become a disproportionately powerful weapon in cases where the underlying IP claims are very weak, such as in the case of transit data. See, e.g., DIETRICH, *supra* note 12, at 15 (narrating Apple's takedown of a Belgian transit data app after the transit authority complained that its rights in the data were infringed); Robert McMillan, *San Francisco Misses the NextBus*, PC WORLD (July 3, 2009), <http://www.pcworld.com/article/167856/article.html> (reporting Apple's takedown of another transit data app over claims to copyright in real-time transit data). Notice and takedown has been criticized for its drastic impact on the circulation of works in which copyright is disputed or which may amount to fair use. See, e.g., Matthew Schonauer, *Let the Babies Dance: Strengthening Fair Use and Stifling Abuse in DMCA Notice and Takedown Procedures*, 7 I/S: J.L. & POL'Y FOR INFO. SOC'Y 135, 152 (2011).

197. For example, in Australia, Sydney's RailCorp was alleged to have threatened a law suit against a developer who had created an app based on transit timetables. See Asher Moses, *CityRail puts brakes on iPhone timetable app*, THE AGE (Mar. 5, 2009), <http://www.theage.com.au/news/digital-life/mobiles--handhelds/articles/bureaucrats-put-brakes-on-iphone-app/2009/03/05/1235842537210.html>. In Europe, for example, the National Belgian Railway Co. (NBRC) asserted database rights in its schedule data against the developers of a mobile app that made use of the data. See DIETRICH, *supra* note 12, at 10. Public outcry eventually led the NBRC to back down. See *id.*

198. City-Go-Round maintains a list of transit agencies that make their data open and those that do not. See *All Transit Agencies*, CITY-GO-ROUND, *supra* note 15. Information from their website indicated that of 1026 public transit agencies considered world-wide, 291 had open data and the remainder did not. See *id.*

the system, in order to ensure accuracy.<sup>199</sup> Nevertheless, in some cases, transit authorities have also sought royalties for the use of their timetable data.<sup>200</sup>

In spite of initial resistance, some transit authorities took the lead in exploring the potential of sharing static transit data as open data. Portland's TriMet transit authority led the way in making static transit data open to developers as early as 2006.<sup>201</sup> The Portland experience was instrumental in that it was largely positive and highly popular.<sup>202</sup> Much of its success was attributed to an open and engaged attitude from within the transit authority, as well as the hard work it took to develop a strong and positive relationship between the transit authority and the broader community of developers.<sup>203</sup> Portland served as a role model for other transit agencies that followed. Other early adopters were transit authorities in Boston and Chicago.<sup>204</sup> By 2010, municipalities such as Washington D.C. and New York City were beginning the process of making transit data open.<sup>205</sup> By this time, the transit data landscape had again changed, and the emerging transit data issues related to real-time GPS data.

---

199. In her guide for transit authorities considering a move to open data, Kaufman identifies quality control as a common concern. *See* Kaufman, *supra* note 64, at 10; *see also* Moses, *supra* note 192. Kaufman's response is to note that the market will take care of apps that use faulty data, as users will simply move to a better quality app. Kaufman, *supra* note 65, at 10.

200. *See, e.g.,* Jones, *supra* note 91. The impact on a government agency's ability to engage in cost-recovery is recognized as a barrier to open data. *See* MetroGIS Data Producers Work Group, *Making Public GIS Data Free and Open: Benefits and Challenges*, in METROGIS: FREE & OPEN ACCESS TO DATA RESEARCH & REFERENCE DOCUMENTS 5 (Randy Knippel & Geoff Maas eds., 2103), available at [http://metrogis.org/MetroGIS/media/gis-documents/publications/MetroGIS\\_014\\_FreeAndOpenDataResearch.pdf](http://metrogis.org/MetroGIS/media/gis-documents/publications/MetroGIS_014_FreeAndOpenDataResearch.pdf). Washington's WMATA was initially reluctant to share its data, and it explored options to derive revenue streams from these data. *See* ROJAS, *supra* note 8, at 66. Even when it did respond to vocal demand from developers, it was strongly criticized for setting terms of use that were very restrictive, and that left open the possibility that license fees might later be demanded. *See id.* at 67. Greater Greater Washington recorded some of the criticisms of the WMATA's approach. *See, e.g.,* Perkins, *supra* note 57; Michael Perkins, *Google Transit: What's in Metro's Terms?*, GREATER GREATER WASHINGTON (July 24, 2009), <http://greatergreaterwashington.org/post/3012/google-transit-whats-in-metros-terms/>.

201. *See* ROJAS, *supra* note 8, at 55. Portland had made its transit data available through Google Maps in December of 2005. *See id.*

202. *See id.* at 55–58 (narrating the Portland experience); *see generally* McHugh, *supra* note 50.

203. *See* ROJAS, *supra* note 8, at 55–58; McHugh, *supra* note 50, at 128–29.

204. *See* ROJAS, *supra* note 8, at 59–65.

205. *See id.* at 66–71.



### 3. Real-Time Transit Data

Real-time transit data is a stream of data that is communicated from GPS units placed on transit vehicles to a central hub.<sup>206</sup> The information is communicated from the vehicle to the hub at regular intervals.<sup>207</sup> In order for the data to be useful, they must be formatted according to a certain set of parameters which are determined by the software that is used for this purpose.<sup>208</sup> As will be discussed, different choices are available—but essentially they come down to the use either of a proprietary or an open data standard. The use of a proprietary data standard introduces another layer of intellectual property considerations. The discussion below examines the evolution of real-time transit data, the related data standards, and the associated intellectual property issues.

In the mid-to-late 2000's, transit authorities began to install GPS technology on transit vehicles for a variety of operational purposes.<sup>209</sup> Knowing where vehicles were located in real time could enhance management of the system, particularly where accidents, traffic congestion, bad weather, and other events could disrupt planned operations. The data gathered by real-time GPS systems would also be useful for overall planning purposes, as it might give a clearer indication of patterns, problem areas, and performance issues. Yet it quickly became clear that real-time transit data could have customer-oriented uses as well.<sup>210</sup> Such data could be used to communicate information about predicted vehicle arrival times to customers.<sup>211</sup> This information might allow customers to improve their experiences, for example, by allowing them to catch buses that they might otherwise have thought they had missed, or to take shelter a bit longer before heading out to wait for a bus in a rainstorm. Studies also showed that, whether or not the information was of practical

---

206. The purpose for the data could include fleet management, performance management, safety and security considerations, and even the monitoring of drivers. The Transportation Research Board gives a detailed account of the functioning of real-time GPS tracking (also known as Automatic Vehicle Location or AVL). See generally FURTH ET AL., *supra* note 30.

207. *Id.* at 17–19.

208. For a detailed discussion of how real-time GPS systems are designed and deployed, and how their functionality is determined, see *id.* at 14–20.

209. See FURTH ET AL. (discussing the use of this technology for operational purposes).

210. ROJAS, *supra* note 8, at 19–20 (noting that while AVL systems were initially introduced for internal transit authority purposes, many transit agencies recognized that this information could be used to provide the public with more accurate arrival time information).

211. *Id.*

assistance to riders, it improved their general experiences, as customers were apparently less frustrated with wait times if they had some clear sense of when the vehicle might actually arrive.<sup>212</sup>

The advent of real-time transit data presented transit authorities with some interesting challenges. There was certainly a cost to installing the necessary equipment on the buses and to acquiring the software and hardware necessary to run the system. Many transit authorities sought the services of private sector companies to provide the necessary equipment and to gather, process, and format the data.<sup>213</sup> Where this occurred, it introduced new intellectual property challenges. In the first place, the private sector supplier was providing the equipment and software necessary to gather and process the data, and, in the early days at least, some of these companies made claims to copyright in the predictive data they produced.<sup>214</sup> In such circumstances, while the municipality might be licensed to use the data, it would not be in a position to make it open municipal data, and thereby available for use by developers. For example, in 2009, Routsey, an app developed after San Francisco's BART transit authority made its real-time data available as open data, was derailed when NextBus asserted that it had exclusive rights to that data.<sup>215</sup> The transit authority resisted these claims, and the issue was later resolved so as to permit the use of the data as open data.<sup>216</sup> Issues of "ownership" of resulting data can be negotiated in any procurement contract, yet, in the early days, it may not have been as obvious what the implications might be if control over the licensing and use of the data was left with the private sector company.

---

212. See, e.g., *id.* at 41–42. Tang & Thakuria found evidence of some slight increase in ridership following the roll-out of real-time transit information in Chicago. Tang & Thakuria, *supra* note 6.

213. NextBus, one of the leading companies in this field, boasts over 135 transit agencies as clients. See *NextBus Real-Time Passenger Information Solutions*, *supra* note 59.

214. For example, NextBus once asserted copyright claims to real-time data from San Francisco's transit system. See McMillan, *supra* note 191; see also Teresa Scassa, *Copyright Reform and Fact-Based Works*, in FROM "RADICAL EXTREMISM" TO "BALANCED COPYRIGHT": CANADIAN COPYRIGHT AND THE DIGITAL AGENDA 571, 586–88 (Michael Geist ed., 2010).

215. See McMillan, *supra* note 196; see also Eve Batey & Matt Baume, *Does a Private Company Own Your Muni Arrival Times?*, S.F. APPEAL (June 25, 2009), <http://sfappeal.com/2009/06/who-owns-sfmta-arrival-data/>.

216. See Thomas, *supra* note 25, at 22, for a reproduction of the clause negotiated between NextBus and the San Francisco Municipal Transportation Agency (SFMTA) regarding the ownership of real-time GPS data. The clause clearly provides that the SFMTA is the owner of this data, and that it is entitled to make it publicly available. *Id.*

Another issue, of course, was whether the data (assuming that it was capable of being protected under copyright law) was something that municipalities wanted to share, even if they were legally entitled to do so. Early plans to share real-time data with transit users involved, for example, electronic signage in subway stations for which advertising space could be sold.<sup>217</sup> Other possibilities existed for commercializing the data, and some transit authorities, always cognizant of the bottom line, were reluctant to simply give away data that might have a commercial value.<sup>218</sup>

In addition to issues around revenue generation, concerns over quality control also arose in relation to the use of real-time transit data. However, these quality issues were not exclusively focused on the real-time data. In one example, New Jersey's NJ Transit blocked access to its real-time data feed by a developer it accused of providing inaccurate data.<sup>219</sup> According to one source, NJ Transit's response to this incident was to emphasize that it was the best and most reliable source of transit information about its operations.<sup>220</sup> Data quality in this context could have a real impact on rider experience. For example, if bus arrival predictions are inaccurate, customers may be even more disgruntled than they would be with a bus that is simply late for its scheduled arrival. Further, from a reputational point of view, predictive data that shows a consistent marked divergence from the scheduled times could have a negative reputational effect on a transit authority. However, this latter "quality" issue might have less to do with the quality of the real-time data and more to do with the efficient operation of the transit system. Thus, the assertion of IP rights over data for quality concerns, in this context, might be motivated less by the desire to protect the public and more by a desire to blur transparency. The Chicago Transit Authority, for example, chose not to release its real-time transit data as open data until after it had improved its on-time performance within the system.<sup>221</sup>

---

217. *See supra* note 74.

218. For example, it was reported that the City of Ottawa's Transit Commission contemplated reversing a decision to make real-time GPS data publicly available, after city staff produced estimates of potential advertising revenues should the data be kept closed. *See* Jessica Smith, *OC Transpo Recommends Reneging on Open Data Promise*, METRO (Jan. 18, 2012), <http://metronews.ca/news/ottawa/39143/oc-transpo-recommends-reneging-on-open-data-promise/>.

219. *See* Dana Hawkins-Simons, *Clever Commute in Dispute with NJ Transit Over App*, BARISTANET (Aug. 7, 2012), <http://baristanet.com/2012/08/clever-commute-in-dispute-with-nj-transit-over-app/>.

220. *Id.*

221. ROJAS, *supra* note 8, at 38.

The nature of real-time transit data made standards—open or proprietary—much more of an issue for the use of this type of transit data. Google was quick to recognize the public interest there would be in real-time transit data, and it worked with developers and transit authorities early on to arrive at the GTFS-real-time standard.<sup>222</sup> Data prepared according to this open standard could easily be incorporated into the Google Maps interface. As an open standard, it could also be used in the development of apps that would present riders with real-time transit information via their mobile devices. Other real-time data standards exist. The Service Interface for Real Time Information (SIRI) is the dominant standard in Europe,<sup>223</sup> although it has not caught on to any great extent in North America. Private sector suppliers of GPS-real-time data services to transit authorities may also offer their own proprietary standards.<sup>224</sup> Data prepared according to these standards cannot be incorporated into Google Maps without conversion. Moreover, although apps can be written that make use of these data (assuming they are made available as open data), the use of a closed standard is more limiting in terms of development options and possibilities for interoperability of apps from one transit system to another.<sup>225</sup>

As is typically the case with this type of technology, users will eventually gravitate towards particular standards and other less popular standards either fall by the wayside or have key features absorbed by the more dominant standards. For the purposes of interoperability, which is discussed below,<sup>226</sup> a single common (and open) standard is the ideal end result.<sup>227</sup> In North America the GTFS for static data is now dominant, while the real-time context is still

---

222. See *What is GTFS-Realtime?*, GOOGLE DEVELOPERS, <https://developers.google.com/transit/gtfs-realtime/> (last visited Nov. 9, 2014).

223. See SERVICE INTERFACE FOR REAL TIME INFORMATION, <http://user47094.vs.easily.co.uk/siri/> (last visited Nov. 9, 2014).

224. For example, NextBus uses its own proprietary standard for real-time GPS data. See Kurt Raschke, *Open Standards for AVL and Other Real-Time Transit Data*, KURT RASCHKE (Apr. 12, 2011), <https://kurtraschke.com/2011/04/open-standards-for-avl>. Rojas recommends against the adoption of proprietary standards. ROJAS, *supra* note 8, at 46.

225. See Kaufman, *supra* note 64, at 2 (finding that in 2012 only four U.S. transit authorities used GTFS-realtime). This is, of course, an area that is rapidly evolving and changing. Since the publication of the Kaufman paper, for example, the New York City subway system has adopted GTFS-realtime. See *Subway Real-Time Data Feeds*, MTA.INFO, <http://datamine.mta.info> (last visited Nov. 9, 2014).

226. See discussion *infra* Part IV.B.4.

227. See ROJAS, *supra* note 8, at 44.

evolving.<sup>228</sup> There are several advantages to open standards. As they are shared by a community of users, those users may provide feedback and request changes to the standard either to improve it or to have it adapted to their circumstances.<sup>229</sup> Where a transit authority makes data available in a shared standard, it is possible to quickly adapt existing apps to the newly released data, making the development of useful apps for transit users much quicker.<sup>230</sup> Data formatted according to open standards will also be more easily combined with other data to produce innovative new information tools.<sup>231</sup>

With the evolution of real-time transit data, one continues to see the same concerns over quality control and cost recovery surface for transit authorities.<sup>232</sup> The quality issues, however, become more complex as the line blurs between the quality of the data and what it might reveal about the efficiency of the transit service being provided. Cost recovery concerns shift as well, since the real-time data have more commercial potential for municipalities than do static transit data. Nonetheless, those municipalities that had made static transit data open seem also to be committed to making the more complex real-time data open as well.<sup>233</sup> Real-time data introduced an additional dimension not present with more static forms of data. This was due to the presence of private sector corporations that entered into contracts with municipalities to provide the service of collecting and processing the real time data. In the early days at least, issues arose over rights—as between the contracting municipalities and the

---

228. See PESSOA ET AL., *supra* note 48, at 8. City-Go-Round maintains and updates a list of municipalities offering open transit data, *All Transit Agencies*, CITY-GO-ROUND, *supra* note 15, as well as a list of available apps, *Transit App Gallery*, CITY-GO-ROUND, *supra* note 71.

229. ROJAS, *supra* note 8, at 46.

230. PESSOA ET AL., *supra* note 48, at 8. It should be noted that some apps are open source, allowing for very rapid adaptation to newly available data. See, e.g., OPENTRIPPLANNER, *supra* note 187. Other apps may be proprietary, but are nonetheless available for use with multiple transit authorities. See, e.g., HOPSTOP, <https://www.hopstop.com/> (last visited Nov. 9, 2014). HopStop is adapted for multiple transit authorities in North America and in Europe. See *id.*

231. ROJAS, *supra* note 8, at 46; see HILLSMAN & BARBEAU, *supra* note 48, at 10–11 (underlining the importance of using open standards in order to create trip planners that span different transit authorities or multiple modes of transit); see also Kurt Raschke, *Why ‘They’re Not on NextBus’ Isn’t the Problem it Sounds Like*, KURT RASCHKE (Jan. 11, 2014), <https://kurtraschke.com/tag/siri>.

232. See ROJAS, *supra* note 8, at 63–68 (exemplifying the case studies of the Chicago CTA and the Washington WMATA).

233. See *All Transit Agencies*, CITY-GO-ROUND, *supra* note 15 (listing municipal transit authorities that provide open transit data).

private sector companies—in the data generated through these systems. This rendered data “ownership” issues more complex. In addition, the use of closed proprietary standards by private sector companies had an impact on downstream uses of the data even in those cases where municipalities choose to make the data open to developers.

#### 4. *Legal Interoperability and Clear Licensing*

Even in those cases where a transit authority chooses to make its data available as open data, issues may still arise in relation to the manner in which the data are licensed. The open data movement has generated a healthy volume of discussion regarding the idea of an open data license, and in fact, many different open licenses are available.<sup>234</sup> In some cases, governments or transit authorities have chosen to adopt a license from the Creative Commons suite of licenses.<sup>235</sup>

There is also an Open Database Licence that has been crafted to deal with the particular circumstances of licensing data.<sup>236</sup> Some governments have drafted their own open data licenses, which contain terms specific to the particular realities of government data.<sup>237</sup>

---

234. Open licenses tend to have different degrees of openness. *See, e.g.*, About the Licenses, Creative Commons, <http://creativecommons.org/licenses/> (last visited Nov. 9, 2014). Some have few or no restrictions, e.g. (such as the CC-BY license which requires only attribution), while others require a commitment to “share alike,” or permit only non-commercial downstream uses of the licensed content. *See id.* Problematic open licenses can be a barrier to the use of open data. *See, e.g.*, DIETRICH, *supra* note 12, at 14.

235. For example, transit data made available in Australia through Translink is available under a Creative Commons license. *See Open Data*, TRANSLINK, <http://translink.com.au/news-and-updates/open-data> (last visited Nov. 9, 2014). The Government of Australia has also adopted the Creative Commons (CC) Attribution 3.0 Australia license for its open data. *See Attribution 3.0 Australia*, CREATIVE COMMONS, <https://creativecommons.org/licenses/by/3.0/au/deed.en> (last visited Nov. 9, 2014). For a list of governments that have adopted CC licenses for use in various contexts, see *Government use of Creative Commons*, CREATIVE COMMONS, [http://wiki.creativecommons.org/Government\\_use\\_of\\_Creative\\_Commons](http://wiki.creativecommons.org/Government_use_of_Creative_Commons) (last visited Nov. 9, 2014).

236. These circumstances are distinct from other contexts in large part because the European Union has created a sui generis database right that offers protection to databases and the data they contain, independently of copyright law. *See, e.g.*, PAUL MILLER ET AL., OPEN DATA COMMONS, A LICENSE FOR OPEN DATA (2008), available at <http://events.linkeddata.org/ldow2008/papers/08-miller-styles-open-data-commons.pdf>. The Open Database License is maintained by Open Data Commons, and can be found at <http://opendatacommons.org/licenses/odbl/>.

237. The United Kingdom has led the way with its Open Government Licence (OGL). *See Open Government Licence*, *supra* note 105. Note that this license was created to address the lacunae in the Creative Commons licenses when it came to

In some cases, transit authorities have drafted licenses that are specific to their particular data sets.<sup>238</sup>

To some extent, an open license permits one to set aside the issues regarding whether there is any copyright in the compilation of data being licensed. This is so, even though many open licenses are still premised upon claims to copyright in the work that is the subject of the license.<sup>239</sup> Since the work is made available under an open license there is no real incentive to challenge claims to copyright in the data. Nevertheless, both the underlying claim to intellectual property rights and the license terms and conditions very much form a part of the context in which the data are made available.

The lack of general consensus over an appropriate open license for transit data (or for any open data) creates a context in which “legal interoperability” can become an issue.<sup>240</sup> Legal interoperability is a term used to describe the compatibility of different licenses in cases where multiple data sets are “mashed up” or combined together in a single product.<sup>241</sup> An app developer, for example, who sought to

---

addressing the European database rights. See KENT MEHWORT, CREATIVE COMMONS LICENCES: OPTIONS FOR CANADIAN OPEN DATA PROVIDERS 3 (2012), available at <https://www.cippic.ca/sites/default/files/Creative%20Commons%20Licenses%20-%20Options%20for%20Canadian%20Open%20Data%20Providers.pdf>. Other jurisdictions have adapted and adopted versions of the UK OGL. See, e.g., *Open Government Licence—Canada*, GOV'T CAN., <http://data.gc.ca/eng/open-government-licence-canada> (last visited Nov. 9, 2014).

238. Note that there was much criticism of Washington's WMATA's first open data license, the terms of which were considered to be unduly onerous, and thus discouraging to developers. See ROJAS, *supra* note 8, at 66; Michael Perkins, *It's Here! Metro Posts Transit Data Online*, GREATER GREATER WASHINGTON (Mar. 23, 2009), <http://greatergreaterwashington.org/post/1845/its-here-metro-posts-transit-data-online/>.

239. This is certainly the case with the CC licenses, as it is with the many of the transit-data open licenses. See, e.g., *License Agreement and Terms of Use*, REGIONAL TRANSP. DISTRICT, cl. 7, [http://www.rtd-denver.com/License\\_Agreement/License\\_Agreement.pdf](http://www.rtd-denver.com/License_Agreement/License_Agreement.pdf) (last visited Nov. 9, 2014); *MCTS Google Transit Feed Terms of Use*, MILWAUKEE COUNTY TRANSIT SYS., cl. 9, <http://kamino.mcts.org/gtfs/> (last visited Nov. 9, 2014); *Terms of Use*, OAHU TRANSIT SERVS., INC., cl. 5, <http://www.thebus.org/transitfeed/terms-of-use.asp> (last visited Nov. 9, 2014); *WMATA Developer License Agreement*, WASHINGTON METROPOLITAN AREA TRANSIT AUTHORITY, cl. 7, [http://www.wmata.com/rider\\_tools/license\\_agreement.cfm](http://www.wmata.com/rider_tools/license_agreement.cfm) (last visited Nov. 9, 2014).

240. See Federico Morando, *Legal Interoperability: Making Open (Government) Data Compatible with Businesses and Communities*, 4 ITALIAN J. LIBR. & INFO. SCI. 441 (2013); Harlan J. Onsrud, *Legal Interoperability in Support of Spatially Enabling Society*, in SPATIALLY ENABLING SOCIETY: RESEARCH, EMERGING TRENDS AND CRITICAL ASSESSMENT 163 (Abbas Rajabifard et al. eds., 2010), available at [http://works.bepress.com/harlan\\_onsrud/1](http://works.bepress.com/harlan_onsrud/1) (discussing the problems of legal interoperability).

241. See, e.g., MEHWORT, *supra* note 231, at 18–20; Onsrud, *supra* note 235, at 163.

combine data from transit authorities in adjacent municipalities, or to combine transit data with other open data (for example, relating to parks or recreation activities, or to embed it in data from OpenStreetMap) would need to ensure that the terms and conditions of the licenses under which the different data sets were made available were compatible.<sup>242</sup> If the terms of one data set barred commercial use, while the others did not, the sets could not be combined in an app that was to be made available commercially. In some cases, licenses may be drafted in such a way that it may not be entirely clear what, if any, restrictions apply, or what their actual scope may be.<sup>243</sup> As Daniel Dietrich notes, unclear license terms “create high transaction costs and [are] a burden a non-legal expert is unlikely to undertake.”<sup>244</sup> In this way, uncertainty as to license terms or a lack of clear interoperability can be an impediment to the use of open data.<sup>245</sup> Thus, even where claims to intellectual property rights are not themselves a barrier to the use of municipal transit data, the license under which the data is made available to the public may have an impact on uptake and use of the data.

### 5. Patent Claims

Not all of the issues arising with respect to the use of transit data have involved copyright claims. In fact, even as most of the major U.S. transit authorities moved to open data and backed away from legal challenges to third party use of these data, another set of IP claims emerged to cause disruption and alarm in this area. These claims are to patent rights in computer code used in the apps that collect and deliver real-time data. These claims have been brought against both transit authorities and app developers.

Non-practicing entities (NPEs)—often colloquially referred to as patent trolls—are companies that hold a portfolio of patents—usually software-related.<sup>246</sup> Their chief line of business is enforcing those patents against other companies that they claim have used them in

---

242. See DIETRICH, *supra* note 12, at 14.

243. See MEHWORT, *supra* note 237, at 20–22.

244. DIETRICH, *supra* note 12, at 14.

245. An online interoperability tool has been created by the CIPPIC Licensing Information Project for Open Licenses (CLIPOL). This tool attempts to offer non-expert users a means of assessing the legal interoperability of different open licenses. See CLIPOL, <http://www.clipol.org> (last visited Nov. 13, 2014).

246. See James Bessen et al., *The Private and Social Costs of Patent Trolls*, REGULATION, Winter 2011-2012, at 26, available at <http://object.cato.org/sites/cato.org/files/serials/files/regulation/2012/5/v34n4-1.pdf> (discussing NPEs).



their own operations.<sup>247</sup> In the transit context, one particular company, ArrivalStar, has asserted patent rights over code that is used in vehicle tracking systems.<sup>248</sup> One source reported that in the space of a year, ArrivalStar had filed more than one hundred lawsuits against companies—including transit authorities—that allegedly made use of the subject matter of the patent.<sup>249</sup> Typically, the transit authority or developer will receive a cease and desist letter and, rather than incurring the costs of litigation, will either stop using the technology in question or will pay a license fee for its use.<sup>250</sup> In the case of ArrivalStar, reports suggested that a significant number of municipal transit authorities in the United States who received such letters settled with ArrivalStar for sums ranging between \$50,000 and \$75,000.<sup>251</sup> While many of the suits filed have been against municipal transit agencies that deploy real-time GPS bus tracking systems, suits have also been brought against individual app developers.<sup>252</sup> In cases where individual app developers lack the resources to push back, and where the app they created generates little revenue, their choice might be to withdraw their app rather than pay license fees.<sup>253</sup> The NextBus system is also operated under license from ArrivalStar.<sup>254</sup>

---

247. *Id.*

248. Although ArrivalStar has primarily been the party asserting patent rights in the real-time GPS context in the United States, similar patent rights have also been advanced in Canada by other related entities. *See, e.g.,* Ariel Katz, *Troll Alert: Dovden Investments/ArrivalStar*, ARIEL KATZ ON INTELL. PROP. COMPETITION INNOVATION & OTHER ISSUES (June 27, 2013), <http://arielkatz.org/archives/2823>; “Patent Troll” Withdraws Lawsuit Against Bus App Developer, CBC NEWS (Sept. 18, 2013), <http://www.cbc.ca/news/canada/ottawa/patent-troll-withdraws-lawsuit-against-bus-app-developer-1.1858462>.

249. Joe Mullin, *A New Target for Tech Patent Trolls: Cash-Strapped American Cities*, ARS TECHNICA (Mar. 15, 2012), <http://arstechnica.com/tech-policy/2012/03/a-new-low-for-patent-trolls-targeting-cash-strapped-cities/>.

250. According to the American Public Transportation Association (APTA), “at least 11” public transit authorities had chosen to settle suits brought by ArrivalStar, rather than to face the costs of litigation. Press Release, Am. Pub. Transp. Ass’n, APTA Announces Settlement with ArrivalStar: Frivolous Patent Infringement Claims Against APTA Members Will Stop (Aug. 21, 2013), *available at* [http://www.apta.com/mediacenter/pressreleases/2013/Pages/130821\\_Patent.aspx](http://www.apta.com/mediacenter/pressreleases/2013/Pages/130821_Patent.aspx).

251. Mullin, *supra* note 249; *see also* ROJAS, *supra* note 8, at 22.

252. *See, e.g.,* “Patent Troll” Withdraws Lawsuit Against Bus App Developer, *supra* note 248.

253. For example, one developer who received a cease and desist letter relating to his transit app was asked for a \$10,000 license fee in relation to an app that grossed about \$200 per month. *See id.*

254. *About NextBus Predictions*, NEXTBUS, <http://www.nextbus.com/help/about/NextBusHelp.shtml> (last visited Nov. 9, 2014).

The patent law suits have added a new dimension to the development and use of open transit data. Transit authorities settling law suits may have to include in the terms of settlement a provision that insulates from lawsuits developers licensed under open data licenses to use the real-time GPS data, thus protecting the emerging app development sector. The impact of so many lawsuits filed against public agencies essentially funded by taxpayer money has also attracted attention. In 2012, the Electronic Frontier Foundation (EFF) in conjunction with the Samuelson Law, Technology and Public Policy Clinic at Berkeley Law School filed a request for re-examination of the ArrivalStar patent with the U.S. Patent and Trademark Office.<sup>255</sup> This ultimately led to a significant reduction in the scope of the patent.<sup>256</sup> In addition, the American Public Transportation Association (APTA) filed a lawsuit against ArrivalStar, challenging the validity of its patent.<sup>257</sup> The matter was settled shortly afterwards, with ArrivalStar agreeing to cease suing public transit authorities.<sup>258</sup> The APTA also asked the Federal Trade Commission to launch an investigation into the company's practices.<sup>259</sup> The United States government has also announced

---

255. Adi Kamdar, *30+ Examples of Prior Art to Help Combat ArrivalStar's Patent*, ELECTRONIC FRONTIER FOUND., (Dec. 12, 2012), <https://www.eff.org/deeplinks/2012/12/thirty-examples-prior-art-combat-arrivalstars-patent>.

256. Julie Samuels, *EFF Throttles Notorious Patent Used to Threaten Public Transit Systems*, ELECTRONIC FRONTIER FOUND. (June 26, 2013), <https://www.eff.org/press/releases/eff-throttles-notorious-patent-used-threaten-public-transit-systems>; see also Julie Samuels, *ArrivalStar: How to NOT Make Friends and Influence People*, ELECTRONIC FRONTIER FOUND. (June 27, 2013), <https://www.eff.org/deeplinks/2013/06/arrivalstar-how-not-make-friends-and-influence>; *ArrivalStar patent decision from USPTO*, ELECTRONIC FRONTIER FOUND., <https://www.eff.org/document/arrivalstar-patent-decision-uspto> (last visited Nov. 9, 2014) (providing the USPTO decision).

257. See Complaint, Am. Pub. Transp. Ass'n v. Arrivalstar S.A., No. 13 Civ. 4375 (S.D.N.Y. Jun. 25, 2013), available at <http://www.apta.com/mediacenter/Documents/APTA%20v%20ArrivalStar%20-%20Complaint%20%28STAMPED%29.pdf>; see also Joe Mullin, *Patent Troll that Sues Public Transit Systems Gets Hauled into Court*, ARS TECHNICA (June 27, 2013), <http://arstechnica.com/tech-policy/2013/06/patent-troll-that-sues-public-transit-systems-gets-hauled-into-court/>.

258. Alex Lawson, *ArrivalStar Calls Off Dogs in Public Transit Patent Suits*, LAW360 (Aug. 22, 2013), <http://www.law360.com/articles/466953/arrivalstar-calls-off-dogs-in-public-transit-patent-suits>; see also Press Release, Am. Pub. Transp. Ass'n, *supra* note 250.

259. See Mullin, *supra* note 257. The FTC has, in fact, launched a major investigation into patent trolls in general. See Edward Wyatt, *F.T.C. Is Said to Plan Inquiry of Frivolous Patent Law Suits*, N.Y. TIMES, June 19, 2013, available at [http://www.nytimes.com/2013/06/20/business/ftc-is-said-to-plan-inquiry-of-frivolous-patent-lawsuits.html?hpw&pagewanted=all&\\_r=1&](http://www.nytimes.com/2013/06/20/business/ftc-is-said-to-plan-inquiry-of-frivolous-patent-lawsuits.html?hpw&pagewanted=all&_r=1&).

measures to curb patent troll behavior.<sup>260</sup> This strong, multi-party response is interesting as it suggests that, at least in high value areas, undue interference with the use of public sector data will generate significant push-back. This may be particularly the case in those jurisdictions where there are organizations with sufficient funds and the mandate to act in the public interest. Where these resources do not exist, patent claims may well limit or stifle the development of apps. To the extent that a rich open data ecosystem stimulates the use of open data, patent trolling behavior can pose a significant problem.

### CONCLUSION

This Article has examined the evolution of open data in the context of municipal transit data, and through an IP lens. As is the case with much municipal data, transit data are not generated as a result of incentives provided by the copyright system; they are generated as a necessary by-product of the operation of a transit system. Perhaps this explains why, at least in the early days, municipal transit authorities appear to have asserted IP rights in their transit data predominantly out of concerns over quality control. As data became more complex and offered more commercial potential, concerns over cost recovery (never entirely absent) became more important. Nevertheless, the evolution in the complexity and commercial value of transit data also came after the first successes of open data. Once launched on an open data trajectory, it may be more difficult for municipalities to retreat from a policy of open data.

The initial IP battles in this arena were between closed and open data. In other words, they involved struggles to compel transit authorities to make their data open to developers. Once data were made open, other intellectual property issues emerged. These included the use of proprietary or non-proprietary data standards, and concerns over the interoperability of data licenses. These issues persist today.

The growing complexity of the data being generated by transit systems reflects the ongoing evolution of technology. The technological evolution often required transit authorities to contract out for the provision of new data-related systems such as the GPS

---

260. Press Release, The White House, FACT SHEET—Executive Actions: Answering the President's Call to Strengthen Our Patent System and Foster Innovation (Feb. 20, 2014), <http://www.whitehouse.gov/the-press-office/2014/02/20/fact-sheet-executive-actions-answering-president-s-call-strengthen-our-p>.

tracking of vehicles. The contracting out of services led to new IP issues—ones that will be particularly important in the evolving Smart Cities<sup>261</sup> context. These issues relate to how rights in data are managed as between governments and their agencies and private sector service providers. Where a private sector company provides the hardware and software to collect and process data in relation to the operation of municipal services, it is necessary to consider—and to negotiate—to whom those data belong, and to whom rights of access should be given.

The open transit data narrative also illustrates the dynamics wrought by technological change that put sophisticated digital and mobile technologies in the hands of independent developers and ordinary transit users. The drive for open data is fueled in large part by a demand for access to resources perceived as public that are ripe for exploration and exploitation by an ever-broadening range of actors. In this context, IP has predominantly acted as a barrier: a barrier to access, a barrier to interoperability, and a barrier to exploitation.<sup>262</sup>

Another feature of this landscape is one that is shared in many other contexts as well. Although there is no copyright in data and copyright in compilations of data is ‘thin,’ the assertion of weak or non-existent claims to copyright by economically stronger parties against those with few resources to litigate tends to achieve effects similar to the exercise of strong IP rights. In this respect, the transit data context is rife with claims to IP rights—whether they are claims made by municipalities, third party service providers or patent trolls—that have questionable legal foundation, but yet that have the power to shape and determine relations. An interesting element within this environment, therefore, is the extent to which public pressure—reflected in newspaper articles, outraged blog posts, and social media—have an impact in building resistance to dubious claims exercised in a context of disparate economic and bargaining power. In the case of patent trolls, the transit data context also sees publicly-

---

261. ‘Smart Cities’ is a term that is broadly inclusive. It has been defined as encompassing “almost any form of technology-based innovation in the planning, development, and operation of cities.” Colin Harrison & Ian Abbott Donnelly, *A Theory of Smart Cities*, PROCEEDINGS OF THE 55<sup>TH</sup> ANNUAL MEETING OF THE ISSS (2011), available at <http://journals.iss.org/index.php/proceedings55th/article/viewFile/1703/572>; see also, ANTHONY M. TOWNSEND, SMART CITIES: BIG DATA, CIVIC HACKERS, AND THE QUEST FOR A NEW UTOPIA (2013).

262. This is not to say that IP rights have not played some role in incentivizing the technological innovation that has driven the development of new and more complex forms of transit data—although the extent of that role remains uncertain.

funded transit authorities uniting to fend off patent infringement claims, and civil society organizations also becoming involved in these battles.

The opening of transit data has, in many cases, led to the development of a wide range of apps, and to the incorporation of the data into other useful tools for urban dwellers and for travelers. What is created goes well beyond what a municipal transit authority would have the resources or mandate to develop, and this alone has been a strong argument for making this kind of data open. Because, in many respects, copyright claims in compilations of data are so weak, openness in this context is more about the willingness of governments, as a matter of policy and practice, to make data sets available to the public in reusable formats. Once this is done, IP issues do not fade away—rather they manifest themselves in new ways, particularly as the landscape changes with many new players and new types of data.