Evaluating the Disruption of COVID-19 on AI Innovation using Patent Filings

Michelle Alexopoulos  
Department of Economics  
University of Toronto  
Toronto, Canada  
m.alexopoulos@utoronto.ca

Kelly Lyons  
Faculty of Information  
University of Toronto  
Toronto, Canada  
kelly.lyons@utoronto.ca

Kaushar Mahetaji  
Faculty of Information  
University of Toronto  
Toronto, Canada  
kaushar.mahetaji@mail.utoronto.ca

Keli Chiu  
Faculty of Information  
University of Toronto  
Toronto, Canada  
keli.chiu@mail.utoronto.ca

Abstract—Economists have long recognized that technological innovation is a key contributor to economic growth due to its impact on productivity. In this paper, we explore the impact of COVID-19 on innovation in artificial intelligence (AI) to better understand future effects on economic growth and productivity. Using patents as a measure of innovation and knowledge production, we analyze monthly patent application filing data from January 2015 to June 2021 to compare and assess trends. Past research has shown that growth in patents in the fields of AI have accelerated since 2012, with 6.5 times more annual filings occurring from 2006 to 2017. Here, we focus specifically on determining if the pandemic has had an impact on this acceleration in AI-related innovation. To accomplish this task we must confront the challenge in using up-to-date patent data for this kind of analysis due to the fact that there are considerable time lags associated with patent filing dates and their ultimate publication dates. In real-time situations such as COVID-19, it is, therefore, difficult to ascertain impact using the publicly available patenting data directly. In this paper, we propose a novel approach for examining existing and up-to-date publicly available patent filing data and use that method to gain new insights into the pandemic’s effects on AI-related innovation. Our findings suggest that the pandemic has had a slowing impact on the rate of innovation in these areas but that the downturn may be reversing.

Index Terms—artificial intelligence, innovation, patents, economic issues of technology, COVID-19

I. INTRODUCTION

There is little doubt that artificial intelligence (AI) and data science are having significant economic and societal impacts [1], [2], with many saying that AI will transform the very process of innovation itself [3]. Before the pandemic, predictions suggested that AI could contribute as much as $15.7T to the global economy in 2030 [4], in part because AI is creating new industries and lowering barriers to participation and access [5]. In this paper, we consider the potential impact of the pandemic on technological innovation and knowledge production in AI and its related areas in data science by analyzing trends in patent filing data over time. In addition to presenting a novel method for utilizing publicly available patent application data to provide an initial read on the current trends in innovation, understanding these trends can aid in forecasting productivity and economic growth in the years to come given the relationship between patenting and economic activity [6], [7].

Patents, both grants and filings, have long been used as a measure of innovation [8], [9]. The knowledge production function first proposed by Griliches [10] captures the notion that research and development (R&D) expenditures influence the production of knowledge and that the evolution of knowledge leads to changes in tangible and intangible outputs, some of which can be observed through patent applications, patents granted, and the introduction of new products to market. Recent analyses show that growth in patents in the fields of AI have accelerated since 2012 with the number of published patent applications in these fields growing from 8,515 in 2006 to 12,473 in 2011 to 55,660 in 2017 [11]. We are interested in understanding if the pandemic has had an impact on this acceleration in patent growth, and hence on the creation of knowledge in these areas. We focus on patent application counts since the time frame precludes a weighting by citations. In addition, since R&D expenditures, another common measure of innovation, are released only after a significant lag, they are unusable for a contemporaneous analysis. Journal articles published by date, an alternative measure, are typically available from article databases at a yearly frequency, making it impossible to examine monthly changes in innovation that occurred as the pandemic evolved.

There has been recent interest in understanding the economic and social impacts of the 2020-2021 coronavirus pandemic [12], [13]. There is ample reason to believe that COVID-19 has altered incentives to invest in R&D activities and the commercialization and adoption of AI-related innovations. For example, COVID-19 has created short-run challenges moving forward for certain applications because of factors such as: supply-chain interruptions for necessary hardware (e.g., challenges in the education space moving online); lack of funds due to plunging profits and share prices (See e.g., [14] and [15]); and, access to skilled workers due to travel restrictions (e.g., the inability to have skilled labour immigrate and poor/blocked internet access from other jurisdictions). For some organizations, COVID-19 appears to have created opportunities. Legislative lock-downs have driven sharp spikes in demand for products and services allowing for
remote work (See e.g., [16]) and online purchases (See e.g., [17]). Soaring share prices and record profits for some big tech companies have also created opportunities to invest more in their R&D activities or purchase smaller start-ups.

In this paper, we examine whether there has been an impact of the pandemic on AI-related innovation by analyzing trends in patent filing data from January 2015 to June 2021. In order to do this, we had to identify and implement a method for dealing with publication lags in patent filing data. Our novel method is described in Section II. Results of our analysis show that there has been a decrease in the rate of AI patent filings through the latter part of 2020 indicating a negative impact of the pandemic on AI innovation. However, our findings also suggest the effect may be turning around.

II. METHODOLOGY

In order to forecast the potential impact on patenting activity, and hence, the impact on knowledge creation, economic output and productivity, we require a method for utilizing the available and most recent patent filing data. For our purposes, we use the publicly available patent applications that have been published by the United States Patent and Trademark Office (USPTO). While it is well known that not all patent applications are granted, patent applications provide important insights into the output of various R&D activities. Here, we propose and apply a novel method to utilize the USPTO’s patent application data released weekly between January 2015-June 2021 for a detailed analysis of the pandemic’s effect on AI-related innovation. The timespan examined gives an opportunity to observe recent pre-pandemic trends in patent applications and better assess the potential impact of COVID-19 on patent filings. Moreover, given that approximately half of applications result in granted patents, any decrease identified in filings will help provide information on likely changes in the quantity of new AI-related products and processes available to markets in the next few years.

The largest challenge to utilizing the publicly existing data for an up-to-date analysis is existing publication lags. According to the USPTO’s site, the general rule is that each patent application will be published after 18 months from the earliest filing date. If all applications are published at the USPTO only after 18 months, it would be impossible to use the publicly available data for any real-time analysis. However, as the data illustrates, there are a number of important exceptions to the 18 month rule that lead to a substantial fraction of applications being published earlier in the timeline. This variation in publication lags was also noted in a post by Martin in 2015 [18]. The reasons for the variations, which are outlined in detail on the USPTO’s site, range from requests for early publication, to foreign filing dates. Of course, some patent applications are also published after the 18 month window and others are not published at all (e.g., provisional patents, patents that were filed with a non-publication request, design patents, patents with a national security implication, etc.). We argue, however, that so long as there are regular patterns in the fraction of non-published patents and filing publications lags across time periods, then changes in the number of observed patents filed and published within a timeframe can be used to provide some information about changes in patenting behavior. Our method for analyzing the most recent and up-to-date patent data (described in detail below) makes use of the fact that regular patterns do appear to exist.

To help illustrate the appropriateness of our method, we provide some simple data to support the idea that there is regularity in recent filing and publication patterns. First, we examined the ratio of the number of published applications by year of filing from 2015-2018 available in the USPTO’s AppFT database to the number of total patents filed by calendar year available in USPTO’s reports. The most recent dates (2019-2021) are excluded due to the fact that large portions of the applications will not have been published because of the variation in publication time lags. The ratio of the number of published applications to the number of total patents filed is very stable across the 4 years ranging from 59%-60.4% and is within the range seen when examining the same ratio for the ten year period from 2009-2018. This would indicate that the fraction of non-published applications remains fairly stable during the examined time period.

Next we examine the regularity of the publication date patterns for applications filed within each month. Specifically, we group applications by their month and year of filing, and then examine the number of each group’s filings that are published in each of the months afterwards through June 2021. For example, for all patents filed in January 2015, we count the number that are published zero months later (i.e., February 2015), one month later (i.e., February 2015), and so on. A comparisons across the years and across the months show that the trends in the publication dates are remarkably regular. Fig. 1 shows the number of patent filings (y-axis) published zero months, one month, up to 48 months (x-axis) after the filing date for each month (January 2015, February 2015, up to January 2021). There are clear spikes in the data that occur at months 4, 6 and 18, with smaller spikes occurring at 22 and 24 months published after the filing.

Given the regularities in the data, we propose two mechanisms for estimating the growth rates in number of patents filed using patterns of filing/publication lags from past years in conjunction with current data. Before explaining the details of these approaches, we begin by explaining our notation.

Let \( m_y \) be the actual number of patents filed in month \( m \) in year \( y \), and \( \hat{m}_y \) represents the estimated value of the number of patents filed in month \( m \) in year \( y \). As illustrated above, because applications are published over an extended period of time, we generally do not know the true value of \( m_y \) until

\[\text{average ratio for the 10 year period from 2009-2018 is 60.5\%}, \text{and the range is from 59\% to 62.1\%.}\]
years later (e.g., publications only approach 0 per month after about 48 months in Fig. 1). Therefore, in what follows we will define \( m_{xy} \) to be the number of patent filings published \( x \) months past the filing date of month \( m \) in year \( y \). This is the series that is shown in the graph in Fig. 1. Further, we will let \( \bar{x}(m, y) \) be the number of available months of data at the time of analysis; that is, the number of months since a filing date of month \( m \), year \( y \) and the fixed date June 2021, yielding the equation \( \bar{x}(m, y) = ((2021 - y) \times 12) + (6 - m) \). As a result \( m_{y}^{x} \) is available from the data whenever \( x \leq \bar{x}(m, y) \) for a given value of \( m \) and \( y \) such that \( \sum_{x=0}^{\bar{x}(m, y)} m_{y}^{x} \leq m_{y} \).

In practice, if the number of months since filing is less than \( \sim 48 \), then the observed number of published patents filed in the public database is less than the true value of \( m_{y} \) due to the available data being truncated. Moreover, without taking differences in the degree of truncation into account, it is impossible to accurately compare observed totals across months or years.

One option for comparison is to impose the same level of truncation across the filing dates to be compared. For example, we can compare the number of January 2020 applications published from January 2020-June 2021 to January 2019 applications published in the first 18 months (January 2019-June 2020) to the number of applications from January 2018 over an 18 month window (January 2018-June 2019) to provide some insights into predicted changes in patent filing behavior between the January filings across the years (See Fig. 2). Then, by modifying the length of the window to adjust for the data truncation (i.e., to account for the fact that there are fewer than 18 months of data available for more recent filings), we can get similar estimates of changes in growth rates for each month of interest. Of course, as time passes, and more applications are published, the error associated with the estimate in the growth rate diminishes, and in the limit, it would approach the error related only to the fact that some patent applications are never made public.

More formally, this amounts to calculating the sum of \( m_{y}^{x} \) for \( x \) over a fixed window of time which remains constant. For example, to compare filings for December of any given year to the December 2020 filing period, we use 6 months of available patent publishing data and to estimate filings for January of any given year to January 2020 filings, we use 17 months of available patent publishing data. That is, we use \( \bar{x}(m, 2020) = ((2021 - 2020) \times 12) + (6 - m) = 18 - m \) and

\[
\left( \frac{\sum_{x=0}^{\bar{x}(m, 2020)} m_{y}^{x}}{\sum_{x=0}^{\bar{x}(m, 2020)} m_{y}^{x} - 1} \right) \times 100
\]

(1)

to calculate estimated growth rate.

The second approach estimates monthly patent filings using the total number of patents published after a filing date of month \( m \) in year \( y \) to June 2021 and applies an adjustment to this number based on the 2015 patent publication patterns; that is, it is adjusted by the fraction of patents published \( x \) months after month \( m \) in the year 2015.

\[
\hat{m}_{y} = \frac{\sum_{x=0}^{\bar{x}(m, y)} m_{y}^{x}}{adj(\bar{x}(m, y))}
\]

(2)

where

\[
adj(\bar{x}(m, y)) = \frac{\sum_{x=0}^{\bar{x}(m, y)} m_{y}^{x}}{m_{y}^{2015}}
\]

(3)

and \( m_{y}^{2015} \) is the total number of patents filed in 2015. We utilize the 2015 numbers for this exercise since all months in the year have at least 48 months worth of published applications. The adjustments can be seen in Fig. 3.

In order to test the accuracy of our method and each of these mechanisms, we calculated the change in growth rate of all patent applications using filing month and past lag times (in months) for the period 2019-2020 and compared that with the USPTO’s reported application totals for the same period (including all publishable and non-publishable applications). According to the USPTO data, the total number of applications fell from 669,434 in 2019 to 646,244 in 2020, a decline by \( \sim 3.46\% \) over the year, while the total applications for utility patents (“patents for invention”) fell 3.9% in spite of the USPTO reporting no significant disruption to its operations and the launch of their COVID-19 Prioritized Examination
### III. Findings and Results

To begin, we illustrate the publication patterns for each month’s AI-related patent filings (See Fig. 4). Consistent with the findings for aggregate filings, there are regular patterns seen in the timing of publication lag after filing. Specifically there are large discernible spikes in publications 4, 6, and 18 months after the initial filing date. Notably, there appears to be far fewer AI-related applications that are published beyond the 18 month window than in the case of all patents. An examination of the patterns for the earlier years where the totals are less likely to be affected by truncation (i.e., 2015, 2016, 2017) shows that, while there is some variation depending on the month of filing, ~ 30% of the total number of AI-related patents filed are published within 6 months of filing, with the number rising to about 45-50% by 12 months after, and by the 18th month, generally 75-80% of the total number of filed patents have been published.

Fig. 4 shows the calculated number of patents published (y-axis) by time lag in months (x-axis) for each month, January to December. Each graph shows data for one of the months for years 2015 to 2020. The graphs for January and February also show data for 2021. Examining the graphs in Fig. 4, it is evident that, for the pre-COVID period (prior to March 2020) the number of applications in each publication month has rapidly increased from 2015-2019. This is clearly seen in each of the peak publication months. However, it is also apparent that while the rate of increase (from 2019 to 2020) appears highest in January 2020, growth stalled, and in a few months turned into negative growth after the onset of the pandemic (March 2020). However, the positive finding is that the available data for the most recent filings months suggest that there appears to be a return to at least the 2019 levels of patent filings in the early months of 2021. Given the fact that, as of June 2021, the publicly available data on filings for the years 2019, 2020, and 2021 are not directly comparable due to various degrees of truncation (e.g., there have only been 6 months since December 2020 filings), we report some statistics below to help determine the magnitude of the pandemic’s impact on AI-related patent filings. Table I presents estimates using both of our approaches. On the left side of Table I, we report a series of statistics to facilitate a year-over-year comparison by, in effect, imposing the same level of truncation for the years prior to 2020, as is currently present for each of the 2020 month’s filings. Column 2 in the table gives the number of months used to calculate the totals given by $\bar{m}(2020) = \frac{(2021 - 2020) \times 12 + (6 - m)}{12} = 18 - m$, so that the January 2020 filings examine the data from January 2020 to June 2021 (i.e., for 17 months after filing). Consistent with our first approach, we then calculate the number of applications published in the 17 months following the January 2017, 2018, and 2019 filings and report the corresponding year-over-year growth rates in columns 3-5 of Table I. The
exercise is then repeated for the remaining months (reported in subsequent rows of Table I) with the number of months included in the calculations changing to reflect the current levels (i.e., 2020 levels) of truncation associated with each month. If we compare the growth patterns of monthly AI-related patent filings between 2018 and 2019 using the fixed window method, we see that the number of patents published each month grew at an average rate of over 50%. The most growth in a given month, conditional on the fixed window, was over 65% between April 2018 and April 2019 and the smallest amount of growth was almost 33% between May 2018 and May 2019. Comparing data from 2019 and 2020 presents a much different story. Specifically, from January to July, the average growth rate for this filing data was just under 26%, with the rates of change falling thereafter. In August and November, the growth rates were negative, and a comparison of the September and October statistics show only single digit growth rates. Perhaps a positive sign is that growth from December 2019 to December 2020 moved back up to double digits (to 24.35%). Given that much of the R&D activities leading to a patent filing is performed months in advance, the patterns we see are consistent with larger effects on filings seen a few months after the onset of the pandemic.

Next, we utilize our second method, relying on the distribution for AI-related patents seen in 2015 for the adjustment for the baseline case. The estimates are reported in the last three columns of Table I. Overall, the results again indicate the presence of a slowdown in growth rates from the previous years occurring during the July to November period. Moreover, there is clearly a decline in the predicted year-over-year growth rate for 2020. Growth rates for the periods 2017-2018, and 2018-2019 are estimated to be ~35% while the growth rate for 2019-2020 is estimated to have fallen to 13.5% using the 2015 adjustment. The overall conclusion is similar to the suggested patterns using our fixed window approach. While the growth in AI-related patent filings is not negative during the 2020 year, as it has been for overall patent filings, the growth rate has severely slowed. Moreover, while it may
be too early to ascertain if the filing patterns in early 2021 have returned to normal given there have been less than 5 months’ time since January 2021, and even less for the following months, it is at least reassuring that December growth estimates are returning to more robust levels.

IV. Conclusion

In this paper we utilized USPTO publicly available monthly patent application filing data from January 2015 to June 2021 to assess the impact of COVID-19 on AI-related innovation. We focus on this area since many have identified AI as a general purpose technology that will significantly transform our economy and employment opportunities hence, understanding the current trends in AI-related innovation can aid in forecasting productivity, economic growth and employment in the wake of the pandemic. To identify whether COVID-19 has caused disruption in AI-related innovation, we developed a novel method which takes into account the fact that the publicly available information on recent filings is incomplete due to the fact that there are publication lags. Our findings suggest that the pandemic likely slowed the rate of AI innovation (as measured by patent filings) by approximately 20%. Much of this slowdown occurred during the second half of 2020, which is consistent with the view that COVID-19 impacted R&D activities in addition to AI adoption. The positive news is that the most recent data suggests that the downturn may be reversing. However, given there is still considerable noise in the estimates, more data will need to be examined in the following months to determine if a recovery is indeed underway.

Acknowledgment

The authors wish to thank P. Layarda, H. Pandya, A. Yang, and M. Diaz for their contributions.

REFERENCES


### TABLE I

<table>
<thead>
<tr>
<th>Month of Filing</th>
<th>Growth Rates Approach 1 (%)</th>
<th>Growth Rates Approach 2 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan</td>
<td>17</td>
<td>31.92</td>
</tr>
<tr>
<td>Feb</td>
<td>16</td>
<td>34.15</td>
</tr>
<tr>
<td>March</td>
<td>15</td>
<td>39.35</td>
</tr>
<tr>
<td>April</td>
<td>14</td>
<td>28.31</td>
</tr>
<tr>
<td>May</td>
<td>13</td>
<td>45.71</td>
</tr>
<tr>
<td>June</td>
<td>12</td>
<td>26.82</td>
</tr>
<tr>
<td>July</td>
<td>11</td>
<td>50.77</td>
</tr>
<tr>
<td>Aug</td>
<td>10</td>
<td>58.13</td>
</tr>
<tr>
<td>Sept</td>
<td>9</td>
<td>41.91</td>
</tr>
<tr>
<td>Oct</td>
<td>8</td>
<td>69.78</td>
</tr>
<tr>
<td>Nov</td>
<td>7</td>
<td>46.31</td>
</tr>
<tr>
<td>Dec</td>
<td>6</td>
<td>23.57</td>
</tr>
</tbody>
</table>

Total

35.0 | 34.5 | 13.5 |