Signals From the Future

Mapping AI Patent Landscapes

Lee Allen Kuczewski
M.S. Data Visualization Candidate, Parsons School of Design

Submitted in partial fulfillment of the requirements for the degree of Master of Science in Data Visualization at Parsons School of Design | May 2021

Acknowledgements

As with most endeavors, we can never fully claim them as our own. This thesis is no exception. I'd like to thank Daniel Sauter and Aaron Hill for their generous teachings, guidance, and dedication to the Data Visualization program and building of the generous community at Parsons. Let your dedication to teaching stand as an example for the future of higher education.

To my classmates, who inspired and amazed me time and time again in their dedication and talents. Zhibang Jiang, Inhye Lee, and Seungyu Paik (Soonk), I'm indebted to your long conversations and for teaching me to see things that weren't there. To Jessie Shefrin and Nahum Smith for your friendships, and illuminating generosity to always lend a hand. The team at Lens.org, the USPTO, and to Stephen Metts for his guidance in GIS and mapping. To my family and friends both far and near, those with us and departed, especially to those who are no longer here.

To my dearest wife, Cléa -- you'll always be on a line all of your own. You are truly the blessing of my life.

Abstract

This thesis visualizes AI-related patent filings by geographic region in the United States between 2016 - 2020 to better understand the landscape and asymmetries existing throughout regional clusters.

Inventions are signals from the future, as much as they are from the past. They provide us with a barometer for measuring the patterns of economic activity, policy direction, and competitive moat-building by international companies and governments across the world. These visual explorations will serve to connect policy makers, investors, entrepreneurs and academic commercialization offices with opportunities to make better decisions, bring ideas to market, and grow the economy sustainably within the United States.

This research addresses the actors within the patent ecosystem by examining who its participants are, and importantly by connecting filings to those who are being assigned the patents, more specifically those who take ownership of the intellectual property. Additional county level statistics are examined such as minority populations by county and economic income to better understand the relationships between the number of machine learning patents assigned within a region and the communities, which compose those regions.

Table of Contents

1 Introduction

'The future is already here – it's just not evenly distributed.'
— **William Gibson**

On July 29, 2019, two patents were filed with the United States Patent and Trademark Office and later with worldwide patent agencies. One of the inventions was a light beacon that flashed in a novel manner to attract attention (US16/524,350), and the second, a container based on fractal geometry (US16/524,532).[1] At first glance it was just another patent application by an inventor; but on closer inspection the inventor listed on the patent is named "DABUS". DABUS stands for "device for the autonomous bootstrapping of unified sentience", and is in fact not a human.[2] The representative listed on the patent is Stephen L. Thaler, a physicist and inventor with deep experience in producing disruptive and novel ideas. His original patent for DABUS, and the number of inventions autonomously produced by it creates new questions related to "inventorship" by non-humans. With the possibility of naming an AI machine as an inventor, a wave of legal proceedings across the world has begun unfolding; and international patent organizations are beginning to question the extent to which an AI machine can be claimed as an inventor on patents.

Artificially intelligent machines are transforming our lives. In past decades the idea of machines gaining the characteristics of creativity and human-like intelligence were the realm of science fiction writers; today they are much more real. Daniel Kahneman, the renowned psychologist and economist states, "I do not think that there is very much that we can do that computers will not eventually be programmed to do".[3]

Practitioners within the artificial general intelligence (AGI) community have no agreed upon definition for artificial general intelligence. Some define AGI as self-referencing "synthetic intelligence", while others expand on learning; such as the "adaptation to the environment using insufficient resources", as described by researcher Pei Wang.[4] Thaler's invention, which has been brought to a halt by the patent and trademark office and legal proceedings, could be argued to have solved at least a component in the evolution of intelligent agents regarding the creation of inventions, which moves the debate about inventorship, autonomy for machines, and legal frameworks in new directions. Many of Thaler's prior patents, including the Creativity Machine™ (US 163423423) and others, set the stage for autonomous inventing by artificially intelligent machines.[5]

---

1 Thaler, Stephen. "In the United States Bankruptcy Court for the Eastern District of Virginia Alexandria Division." *New York Times (1923-Current File),* Oct 14, 2002.

2 Thaler, Stephen. "In the United States Bankruptcy Court for the Eastern District of Virginia Alexandria Division." *New York Times (1923-Current File),* Oct 14, 2002.

3 Agrawal, Ajay, Joshua Gans, and Avi Golfarb. "The Economics of Artificial Intelligence: An Agenda." NBER Reporter 2019, (2019): 43

4 Goertzel, Ben. "Artificial General Intelligence: Concept, State of the Art, and Future Prospects." *Journal of Artificial General Intelligence* 5, no. 1 (2014): 1-48.

5 Thaler and L. Stephen. *Electro-Optical Device and Method for Identifying and Inducing Topological States Formed among Interconnecting Neural Modules*. Vol. 61924624. United States:

In early 2021, A judge in the UK upheld a refusal for DABUS to be named as an inventor for the reason that an inventor needs to be a "natural person".[6] The AI cannot own or transfer rights to a patent. Ownership by an AI machine is at the core of the ruling, and the court suggests that because the machine is not considered a legal person, it can't apply or transfer rights that do not exist. We will undoubtedly see more of these legal proceedings, ethical questions of agency and autonomy, and defining the relationships which AI machines will continue to play in the years ahead.

Inventions are signals from the future, as much as they are from the past. They provide a barometer for measuring the patterns of economic activity, policy direction, and competitive moat-building by international companies across the world. An instance of this last point can be seen in the "smartphone" patent wars of 2011, where an estimated 250,000 patents covered the smartphone category.[7] From the liquid crystal displays to speech signal processing the labyrinth of aggregated entities related to smartphone production stretched to 314,490 during this time.[8] Increasingly, we are dependent on these technologies when participating in something as mundane as making a phone call, or as a product of remote communication during the global COVID-19 pandemic, the dreaded video call. These technologies, which both facilitate and monitor our communication are being driven by novel patents in AI and machine learning.

The goal of this thesis is to visualize AI-related patent filings by geographic region in the United States between 2016-2020 to better understand the landscape and asymmetries existing throughout regional county areas. I will address the innovation ecosystem by examining who its participants are, and importantly by connecting filings to those companies who are being assigned the patents, more specifically those who take ownership of the intellectual property. Other county level indicators such as minority population and economic income will be analyzed to understand the relationships between machine learning patents and the communities, which they exist within.

6 [2020] WLR(D) 526, [2020] EWHC 2412 (Pat), [2020] Bus LR 2146 (United Kingdom)

7 RPX Corporation S-1 1 Ds1.Htm FORM S-1.". https://www.sec.gov/Archives/edgar/data/1509432/000119312511240287/ds1.htm.

8 Team, Research, Joel R. Reidenberg, Stanley D. , Nikki Waxberg Chair, N. Cameron Russell, Maxim Price, and Anand Mohan. *Patents and Small Participants in the Smartphone Industry* 2015.

2 Background

‘If it’s a good idea, go ahead and do it.

It is often easier to ask for forgiveness than to ask for permission.’

— Grace Hopper

Stories of intelligent machines have been embedded within our shared narratives and histories for thousands of years.[9] From the “self-pumping bellows” and “self-opening gates” within Homer’s Iliad[10], to the author of *The Wonderful Wizard of Oz,* Frank Baum’s mechanical man named “Tiktok” in 1907, whose “Extra-Responsive, Thought-Creating, Perfect-Talking Mechanical Man … Thinks, Speaks, Acts, and Does Everything but Live”.[11]

Though the imaginary and mathematical foundations for what is now called “artificial intelligence” began long before 1956, the popular genesis story is that a proposal authored by John McCarthy, Marvin Minsky, Nathaniel Rochester, and Claude Shannon was the first to use the term “Artificial Intelligence”.[12] A summer research project at Dartmouth University brought together a group of researchers and academics, studying automatic computing, language-based programming, neuron nets, self-improvement, randomness and creativity, as well as theoretical brain models.

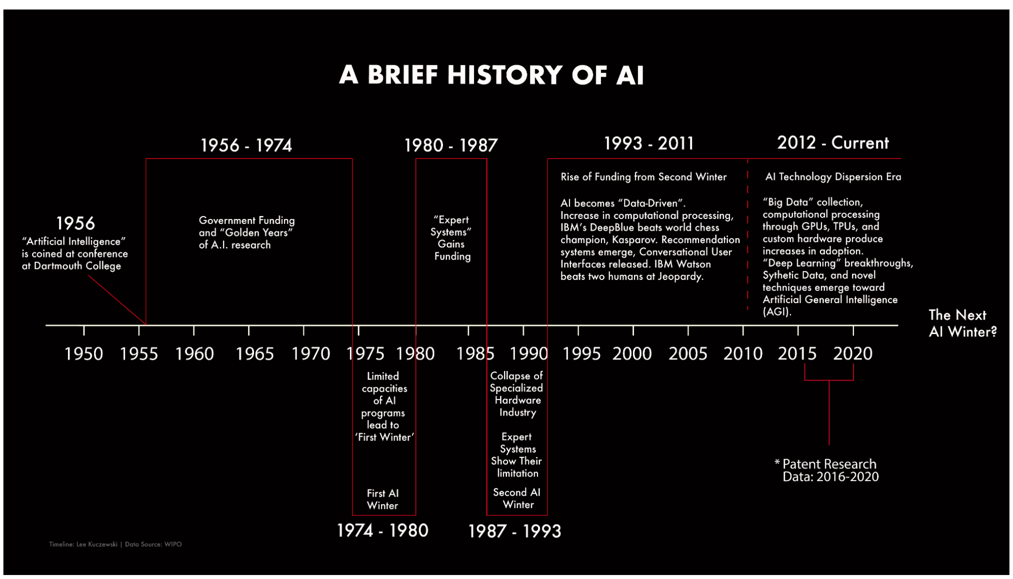


Figure 1. A Brief History of AI

---

9 Liveley, Genevieve and Sam Thomas. *Homer's Intelligent Machines AI in Antiquity*.

10 Buchanan, Bruce G. “A (very) brief history of artificial intelligence.” AI Magazine, vol. 26, no. 4, Winter 2005, p. 53+.

11 Liveley, Genevieve and Sam Thomas. *Homer's Intelligent Machines AI in Antiquity*.

12 McCarthy, J., M. L. Minsky, N. Rochester, I. B. M. Corporation, and C. E. Shannon. "A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence."

The decades following the emergence of AI were affected by a number of boom and bust cycles related to funding the research. Approximately 18 years of government funding fueled the first "Golden Years" of AI research, and in 1974 the community and its researchers experienced what would be termed the first "AI winter".[13] Until 1993, the academic research cycle followed alongside the funding opportunities, with research fueled by innovations in "Expert Systems", which emulated the decision making ability of a human expert. In 1987, the specialized hardware industry collapsed; AI began to show its limitations in practical use cases, and the research and capital resources vanished once again; the second "AI winter" lasted until 1993. As computational processing power increased, the theoretical foundations of AI were able to be realized. Specifically, as in the case with machine learning, enormous datasets and specialized hardware, such as Graphics Processing Units (GPUs) and Tensor Processing Units (TPUs) continued to accelerate research and the application of that research. In May of 2021, we are experiencing a prolonged upward cycle once again, and there is only speculation as to when the next winter may be upon us.

Patent History

Though there remains a debate as to when the idea of patents first emerged, the Venetian Act of 1474 represents one of the first enactments of streamlining the patent process, which lasted successfully for over 300 years. The system helped to propel Venice in transforming itself from a "nation of sailors to being a nation of artisans and engineers, and ultimately the center of technological development in Europe".[14] It was only in the last 230 years that the United States declared its first Patent Act into law. Lasting three years, from 1790 to 1793, it became the foundational system for future acts to expand upon. The patent system was overhauled in 1836 when a new Patent Act was written by Sen. John Ruggles; creating a novel framework for recording, numbering, and classifying patents. Over the next 155 years until 1991, 5 million patents were issued, and it took an astonishing 27 additional years for the next 5 million patents to become issued, bringing the aggregated total to over 10 million patents issued.[15]

As artificial intelligence research was in its "Golden Years", patents containing AI made up approximately 9% technology subclasses, and by 2018 it reached 42%. The number of yearly AI applications grew by more than 100% from 30,000 per year to over 60,000, annually from 2002 to 2018, and active AI inventor-patentees grew from 1% in 1976 to 25% in 2018.[16] The functional applications within AI patents have been dispersing, as new research develops and new techniques emerge from practice. Areas including knowledge representation and reasoning, speech processing, predictive analytics, distributed AI, Natural

---

13 WIPO Technology Trends 2019.

14 Fusco, Stefania. "Lessons from the Past: The Venetian Republic's Tailoring of Patent Protection to the Characteristics of the Invention." *Northwestern Journal of Technology and Intellectual Property* 17, no. 3 (2020): 301-348.

15 USPTO, Patents Through History, https://10millionpatents.uspto.gov/

16 USPTO, Inventing AI: Tracing the diffusion of artificial intelligence with U.S. patents

language processing, robotics, computer vision, control methods, and planning and scheduling, make up the bulk of the current classification.[17]

2.1 The Landscape and Actors

Interactions within the patent landscape rely on connections well beyond the patent office. A number of actors such as educational institutions and commercialization offices, entrepreneurs, venture capitalists and startup studios, independent research centers, government policy makers, private individuals, corporations and the list goes on. Each of these actors plays an important role in the process of creating value and bringing these inventions to the market.

The Entrepreneur

Both the entrepreneur and the firms formed by them play a critical role in patent filings. Over the centuries the term "entrepreneur" has shifted in meaning. Stemming from the latin phrase "inter prehender", meaning "seize with the hand", to transition in the 16th Century to a more broad meaning of "a person who undertakes something". Even later with 18th Century economic theory as "one who behaves actively, or one who acts", the word has undergone a number of transformations throughout the centuries.[18] In this sense, it's the entrepreneur who provides the coordination and movement to bring together the resources necessary to commercialize an idea within a market. The more recent emergence of platform technologies has provided new methods for the entrepreneur to play an active connector role in matchmaking seekers with solvers.

The Investor

Funding for AI startups has also continued to increase. From $7.5 billion in 2016 to $33 billion in 2020, resources are moving to enterprises which are "AI-first".[19] The investors come in various shapes and sizes. Whether the focus of the investments are indiscriminately about gaining profits, or for the social impact, investors are mostly always looking for a return on their resources provided. Without funding sources, entrepreneurs, inventors, and their institutions can only bring an idea so far. In order to scale a product, commercialize an idea, or build a team, investors become an integral part of the innovation ecosystem. Angel investors, venture capitalists, foundations, family offices, and alternative asset funds often play a linchpin role when a product aligns with a market.

The Policy Maker

Opportunities occur when policy and funding align. The U.S. government has committed to making artificial intelligence a priority again, with the National AI Initiative Act becoming law in 2021. Policy can

---

[17] WIPO technology trends 2019

[18] Boutillier, Sophie and Dimitri Uzunidis. *The Entrepreneur : The Economic Function of Free Enterprise*. London: John Wiley & Sons, Incorporated, 2016. http://ebookcentral.proquest.com/lib/newschool/detail.action?docID=4792658.

19 CB Insights, Research Report: Enterprise AI Trends To Watch in 2021.

either accelerate the momentum behind research, or bring it to a halt. The recent release of AI.gov by the White House, seems to reintegrate AI as a strategic initiative where funding will be allocated. Policy making will continue to impact the types and techniques within machine learning research.

Academic Institutions and Commercialization Offices

Universities are bridges to industry, and the partnerships between them are vital to the innovation lifecycle. In addition to training and educating student learners, universities and colleges must have an obligation to society to remain innovative in their research. A number of high-impact companies have emerged from the fabric of academia and research because of the generous intellectual property licensing agreements coming from these institutions. Commercialization offices and technology transfer departments have long been facilitators for creating products from research. As the work of Jaffe, Acs, Audretsch, and Feldmann suggests, the knowledge from the university lab "spills over" and contributes to the "generation of commercial innovations by private enterprises".[20]

[20] Acs, Zoltan J. and David B. Audretsch. *Handbook of Entrepreneurship Research: An Interdisciplinary Survey and Introduction* New York, NY: Springer New York, 2010.

3 Geographies & Patent Assignees

'Everything is related to everything else, but near things are more related than distant things.'
-Waldo Tobler, The First Law of Geography

Geography has important implications for what we know about navigating the world. As the American-Swiss geographer Waldo Tobler wrote in 1970 in his First Law of Geography, "Everything is related to everything else, but near things are more related than distant things". This relatedness by measuring proximity is critical. Since Tobler's writing, the world has experienced unprecedented acceleration related to global network effects of innovation.

Regional Studies, looks at this problem through the lens of "regional clusters''. The field of Regional Studies, or Area Studies, continues to be influenced by the work of Michael Porter, distinguished professor, author, and researcher at Harvard's Institute for Strategy and Competitiveness. Porter has written extensively for decades on the idea of geographical clusters as represented by a "*geographical proximate group of interconnected companies and associated institutions in a particular field, linked by commonalities and externalities.*"[21] Porter's work on clusters has influenced economic and political strategies for entire countries, and also provides a formidable theory on how resources aggregate within certain areas to provide regional advantages.

AI Patent Geographies Within the U.S.

Prior to any attempt at clustering, the research uncovered 141 counties as active in filing machine learning specific patents. This number illustrates that only 4.5% of all U.S. counties had active assignees in our sample, who had a machine learning patent granted between 2016-2020. Further analysis shows the increased growth from year-to-year for filing of 31.85%, 45.51%, 151.74%, and 72.55% from 2016 -2020.

Year	Number of Patents	Percent Change
2016	135	
2017	178	31.85%
2018	259	45.51%
2019	652	151.74%
2020	1125	72.55%

Figure 2. AI Patents Statistics 2016-2020

[21] Porter, Michael E. On Competition. Boston, MA: Boston, MA : Harvard Business School Publishing, 1998.

The figures below were created to better understand where the low and high activity levels were within these U.S. counties. Future analysis showed that 10 regional clusters accounted for 86% of all the machine learning patents filed in our sample. The highest level of concentration (38.6%) was found in a west coast "super-region" in Santa Clara, CA and the surrounding areas of San Mateo, San Francisco, and Alameda.

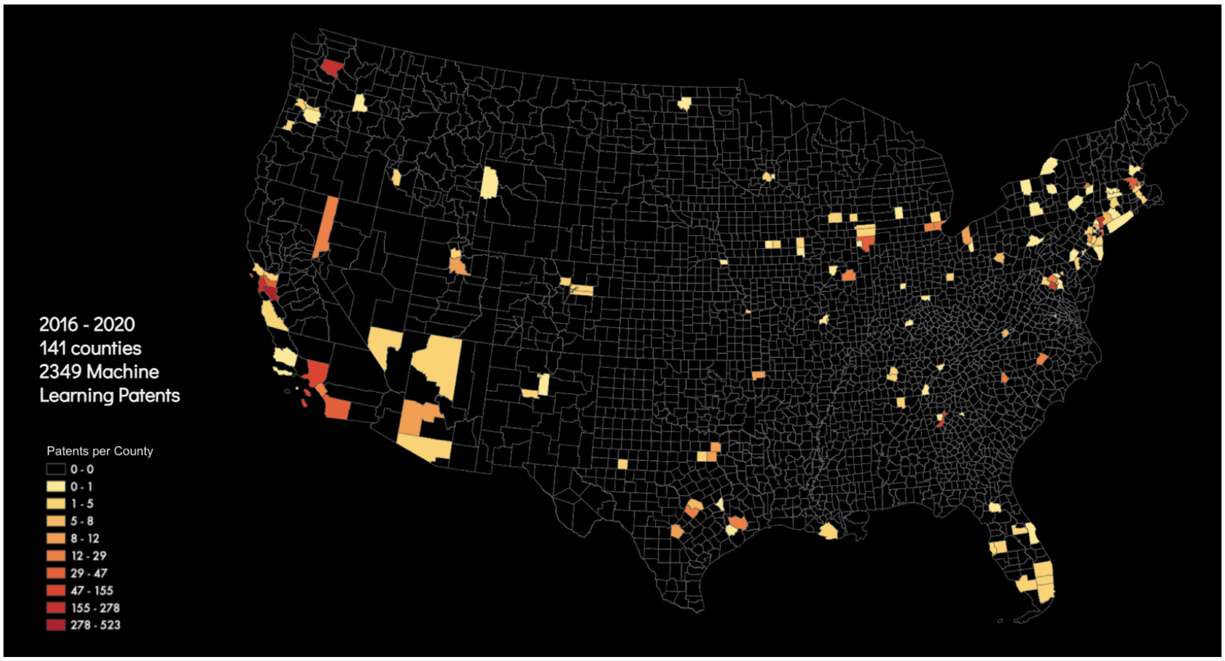


Figure 3. Lee Kuczewski | QGIS analysis 2021, Thematic Map | Patents per County in the U.S.

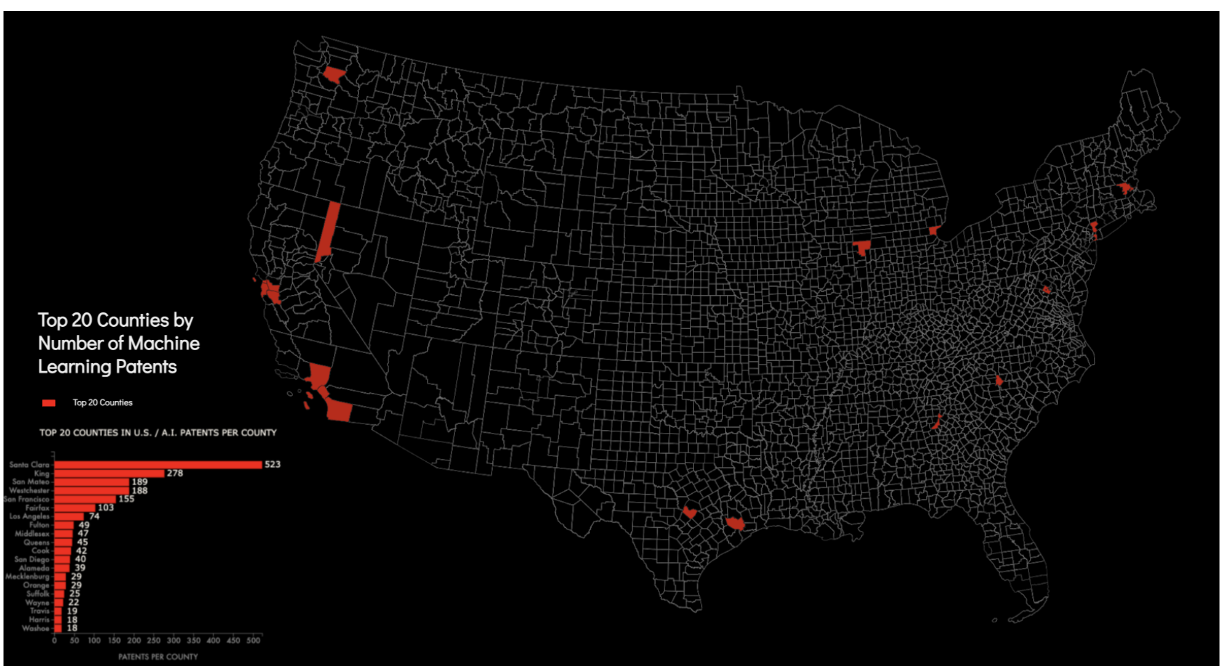
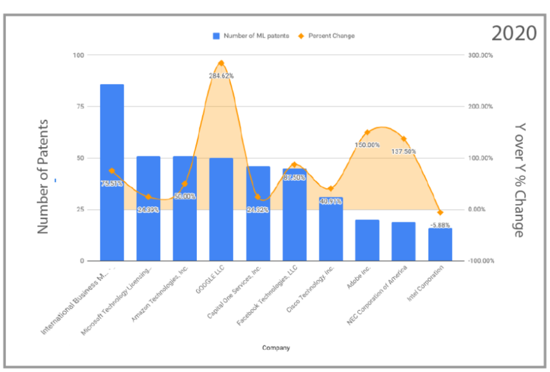


Figure 4. Lee Kuczewski | QGIS analysis 2021, Thematic Map | Top 20 Counties by Number of Machine Learning Patents.

3.1 US. County Assignees

Within the U.S., companies appear to be driving the machine learning patents. The growth rates of companies from year-over-year are considered as a measure of activity by year. Here we see enormous growth year-over-year (in orange) at the individual company level for machine learning patent assignments. The charts illustrate the top ten companies in each year from 2016-2020. The blue bars indicate the number of patents assigned on the left Y axis, and the orange area indicated by the right Y axis is the percentage change in number of patents granted to each company on the X axis.



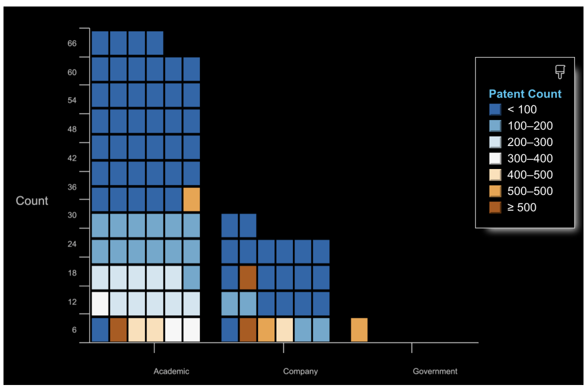
From 2016 to 2020 there was significant growth in the number of machine learning filings per company. One example is the 700% growth rate of Splunk Inc. in 2017, and the 1750% increase year-over-year of Capital One Services, Inc. in 2019. Many companies saw a triple-digit increase year-over-year in the number of machine learning patents granted.

Figure 5. U.S. County-level assignees by company, 2016 - 2020.

3.2 China and U.S. Assignees

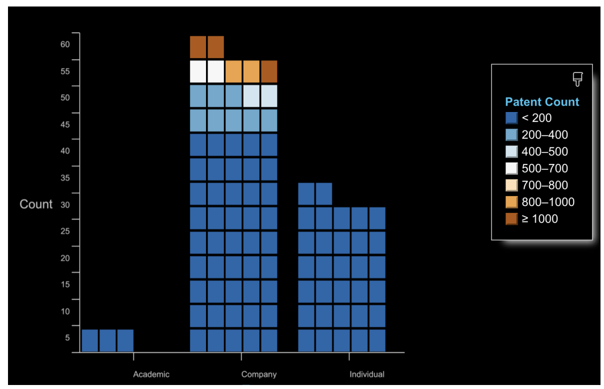
An example of regionalism related to the AI specific patent filings can also be seen in computer vision patent filings between the U.S. and China. Further analysis from a dataset from Georgetown's Center for Security and Emerging Technologies (CSET), revealed the average patent counts by entity type, the total sum of patents by organizational type, and sum of organizations by type for computer vision specific filings in the United States and China. The outcome of this analysis seems to suggest a difference in who is assigned the intellectual property. Academic institutions appear to outnumber company assignments in China, whereas in the U.S. the majority of computer vision patents are held by companies and some individuals.

CHINA



China: analysis using VS code/ Sanddance.

UNITED STATES



United States: Analysis using VS code/ Sanddance.

Computer vision specific patents China:

Average of Patent counts by entity type

Academic	158.3571429
Company	167.8076923
Government	457
Grand Total	**163.9690722**

Sum of Organizations by Type

Academic	70
Company	26
Government	1
Grand Total	**97**

Sum of Computer Vision Patent Count

Academic	11085
Company	4363
Government	457
Grand Total	**15905**

Computer vision specific patents United States:

Average of Patent counts by entity type

Academic	107.6666667
Company	218.7419355
Individual	71
Grand Total	**166.5670103**

Sum of Organizations by Type

Academic	3
Company	62
Individual	32
Grand Total	**97**

Sum of Computer Vision Patent Count

Academic	323
Company	13562
Individual	2272
Grand Total	**16157**

Figure 6. International Assignees, CSET, expanded Dataset.

4 Data & Methodology

4.1 Data Sets and API

Patent data can be quite complex as there is no single categorization schema for every country. A new classification system called the Cooperative Patent Classification (CPC) was created in 2010 between the United States Patent and Trademark Office (USPTO) and the European Patent Office (EPO), to jointly establish a robust and efficient schema. This study used data returned from the United States Patent and Trademark Office's PatentsView API.[22] Search queries related to the CPC scheme could have extended to over 60 CPC categories and 75 search-term categories (Appendix 1,2, 3). Due to the limitations of the API and time allotted for this research, I opted for a more direct approach related to returning all patents from 2016-2020 where "machine learning" was explicitly mentioned within the patent abstract.

```
https://api.patentsview.org/patents/query?q={"_and":[{"_gte":{"patent_date":"2016-01-01"}},{"inventor_country":"US"},
{"_text_phrase":{"patent_abstract":"machine learning"}}]}&f=
["wipo_field_title","patent_number","patent_title","patent_date","inventor_last_name",  "inventor_first_name",
"inventor_city", "inventor_county","inventor_state", "inventor_country","inventor_county_fips", "inventor_state_fips",
"inventor_longitude", "inventor_latitude","assignee_first_name", "assignee_last_name","assignee_organization",
"assignee_city", "assignee_county","assignee_total_num_patents", "assignee_county_fips",
"assignee_state_fips","assignee_country", "assignee_lastknown_latitude", "assignee_lastknown_longitude",
"assignee_latitude", "assignee_longitude"]&o={"per_page":2675}
```

Figure 6. Search Query, PatentsView API, USPTO

The goal was to understand both the locations and assignees of the machine learning specific patents, and we therefore returned both the latitude and longitude information, in addition to the Assignee and Inventor County FIPS codes. County FIPS codes, or the Federal Information Processing Standards are 5 digit codes which are given to counties or their equivalents in every state. This information can be joined to other tables, such as income data, race data, and other metrics in the geoprocessing stages.

Design patents, which are different from utility patents were excluded from our search terms, as we focused less on style and more on functionality. Design patents cover such things as the shape of the rounded corners on the iPhone or other strict design forms. These patents are not within the scope of the project.

For the purpose of this research I was concerned with the assignees on record, those who took ownership of the patent. In the U.S. these tended to be companies, unlike in places such as China, where the vast majority of patents are granted to research institutes and government sponsored universities.
Cleaning the data, and "flattening" the JSON structure to CSV, tabular format was the one priorities before beginning the geoprocessing. This was a fairly manual process and went through several iterations as

[22] USPTO, PatentsView: https://patentsview.org/apis/api-endpoints

we parsed unnecessary keys. As the query returned a large volume of individual inventor names, we had to manually parse these out. In many cases employees of these companies are inventors of the patents, but they "assign" their ownership rights to a company per their employment contracts; therefore, "assigning" all the ownership rights to the company at the time of invention.


```
{
    "patents": [
        {
            "patent_number": "10001775",
            "patent_title": "Machine learning systems and techniques to optimize teleoperation and/or planner
                decisions",
            "patent_date": "2018-06-19",
            "inventors": [
                {
                    "inventor_last_name": "Rege",
                    "inventor_first_name": "Ashutosh Gajanan",
                    "inventor_city": "San Jose",
                    "inventor_county": "Santa Clara",
                    "inventor_state": "CA",
                    "inventor_country": "US",
                    "inventor_county_fips": "6085",
                    "inventor_state_fips": "6",
                    "inventor_longitude": "-121.831",
                    "inventor_latitude": "37.2969",
                    "inventor_key_id": "287955"
                },
```


Inventor and Assignee JSON returned


```
            "assignees": [
                {
                    "assignee_first_name": null,
                    "assignee_last_name": null,
                    "assignee_organization": "Zoox, Inc.",
                    "assignee_city": "Foster City",
                    "assignee_county": "San Mateo",
                    "assignee_total_num_patents": "143",
                    "assignee_county_fips": "6081",
                    "assignee_state_fips": "6",
                    "assignee_country": "US",
                    "assignee_lastknown_latitude": "37.5698",
                    "assignee_lastknown_longitude": "-122.226",
                    "assignee_latitude": "37.5698",
                    "assignee_longitude": "-122.226",
                    "assignee_key_id": "326711"
                }
            ],
```


Figure 7. JSON Sample from USPTO query.

Using QGIS, a geographic information system, I imported these assignee coordinates as a Comma Delimited Text file, and created a point for every latitude and longitude coming back from our API call specific to the assignee data. The initial import showed 2349 patents, which were assigned across 141 counties. To get a basic sense of the shape of the data, pivot tables, custom charts, and interactive graphs were created and designed based on year, county, and the organizations. The following figure (Figure 8), is one sample of this process, though there are dozens more.

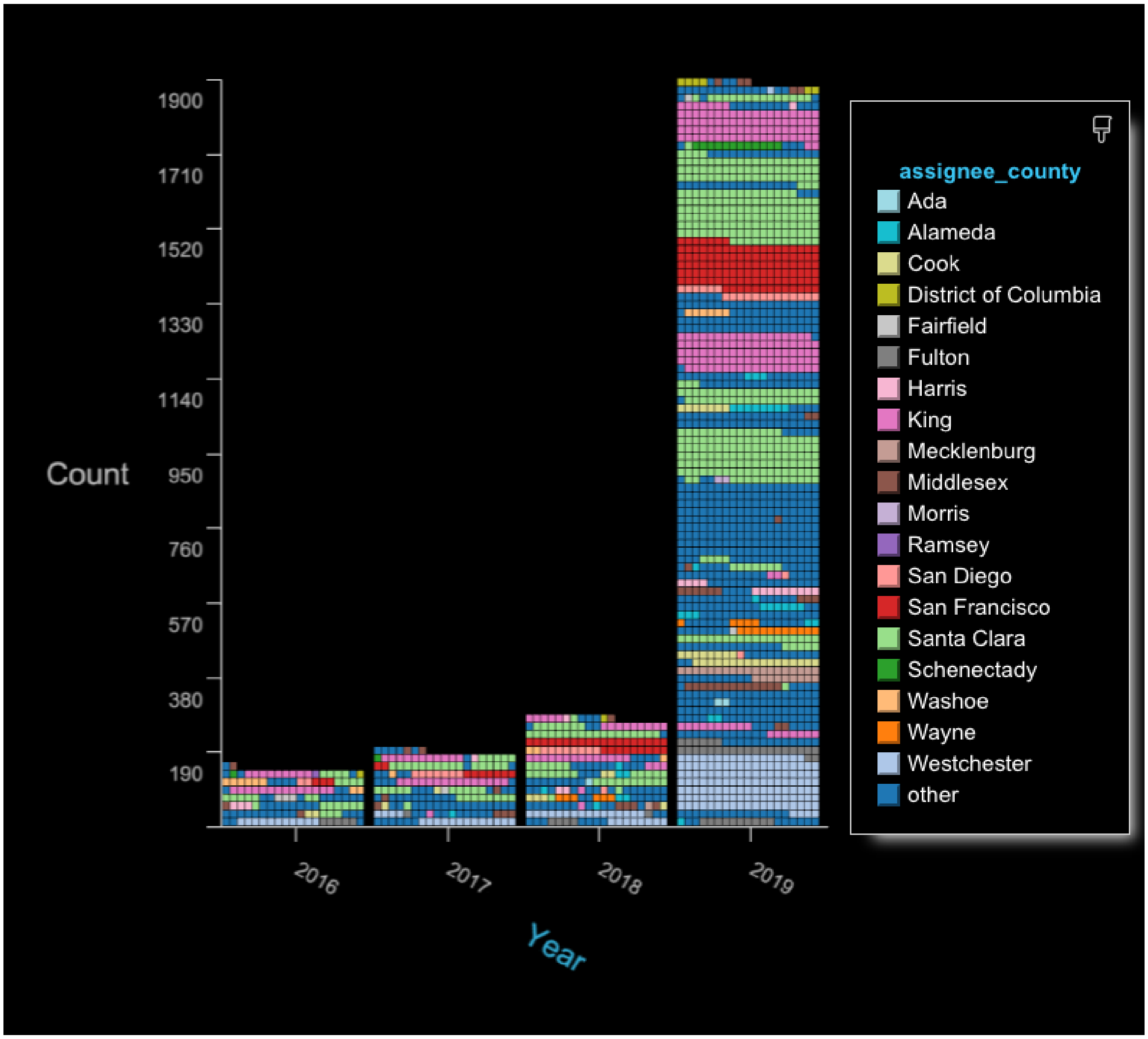


Figure 8. VS Code / Sanddance Regional

After creating an attribute table of the patent data in QGIS, additional data sources were imported to understand the regional characteristics of the locations. Personal income per county region was included from the U.S. Bureau of Economic Analysis (BEA), and consisted of 2019 person income across 2,964 counties.[23] Per capita income was computed using Census Bureau mid-year population estimates. Estimates reflected population estimates available as of March 2020. The income data was joined at the county level in QGIS through a join command on County Fips codes, after cleaning and matching the "countyfips" to account for the additional "0". Structured Query Language (SQL) was used to change the data type and place an additional column using a text type. I then checked the length, and converted both 3 and 4 digits to 5 digits with beginning "0"s to match on county Fips codes.

[23] U.S Bureau of Economic Analysis: https://www.bea.gov/news/2020/personal-income-county-and-metropolitan-area-2019

Another asset created in the process of data cleaning from our Dataset was an interesting pattern around the seasonality of the patent grants, with a greater concentration of grants happening in the months of September through December, and less activity happening in January and February. These findings require further investigation.

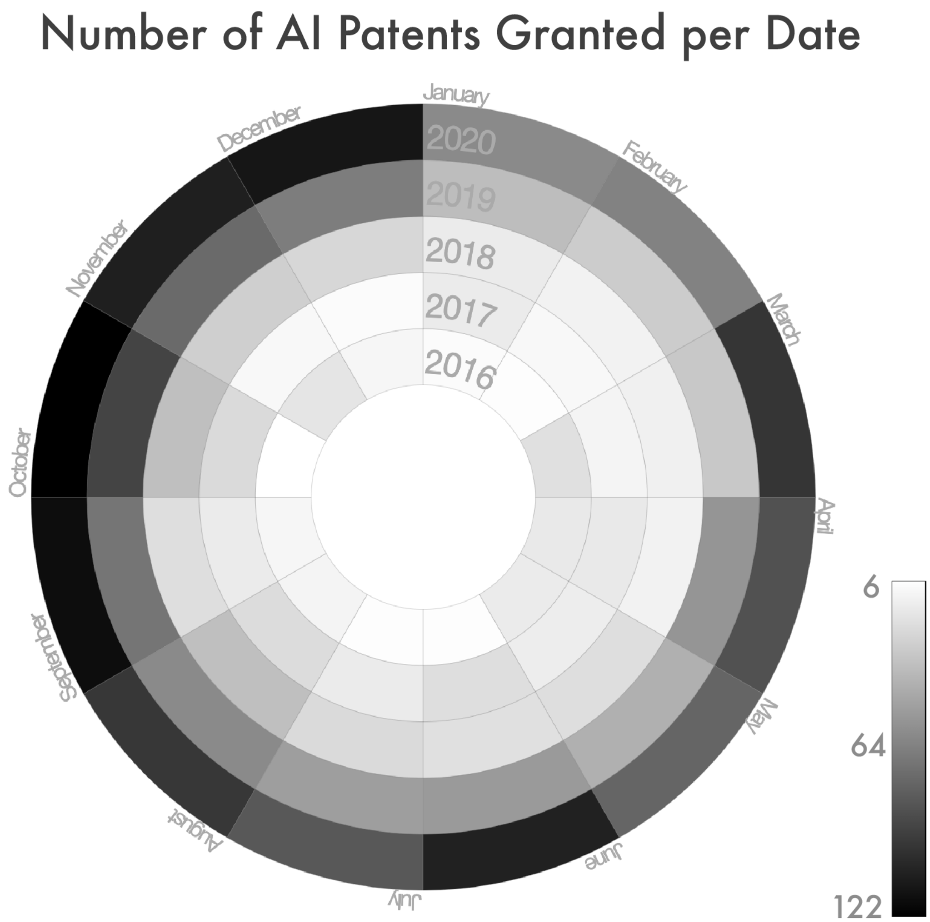


Figure 9. D3.js / QGIS Seasonality Finding

4.1 GIS Analysis and Methods

Using a technique called Spatial Autocorrelation, “clusters of counties” were created and ranked based on the concentration of patents files. A maximum of 523 machine learning patents which were assigned in Santa Clara, and a minimum of 0 in the vast majority of counties. Overall, we found 141 counties under consideration for our analysis, and looked at 10 regional clusters which brought together a number of the counties. We used a statistical calculation called Moran’s I (Figure 10) to get a better sense of how these clusters relate to one another, and determined the Z score for each county (Figure 12). Machine Learning Patent Density per county with a range of 0 (in black) to 523 (dark red). The dark blue within the figure represents a low spatial correlation, and the dark red shows a high spatial correlation per county region. A K-nearest neighbor algorithm was also performed with the results shown in (Figure 11).

The Moran’s I statistic for spatial autocorrelation is given as:

$$I = \frac{n}{S_0} \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{i,j} z_i z_j}{\sum_{i=1}^{n} z_i^2} \quad (1)$$

where z_i is the deviation of an attribute for feature i from its mean $(x_i - \bar{X})$, $w_{i,j}$ is the spatial weight between feature i and j, n is equal to the total number of features, and S_0 is the aggregate of all the spatial weights:

$$S_0 = \sum_{i=1}^{n} \sum_{j=1}^{n} w_{i,j} \quad (2)$$

The z_I-score for the statistic is computed as:

$$z_I = \frac{I - \mathrm{E}[I]}{\sqrt{\mathrm{V}[I]}} \quad (3)$$

where:

$$\mathrm{E}[I] = -1/(n-1) \quad (4)$$

$$\mathrm{V}[I] = \mathrm{E}[I^2] - \mathrm{E}[I]^2 \quad (5)$$

Figure 10. Moran’s I, ESRI ArcGIS[24]

[24] ESRI ArcGIS :https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-statistics/h-how-spatial-autocorrelation-moran-s-i-spatial-st.htm

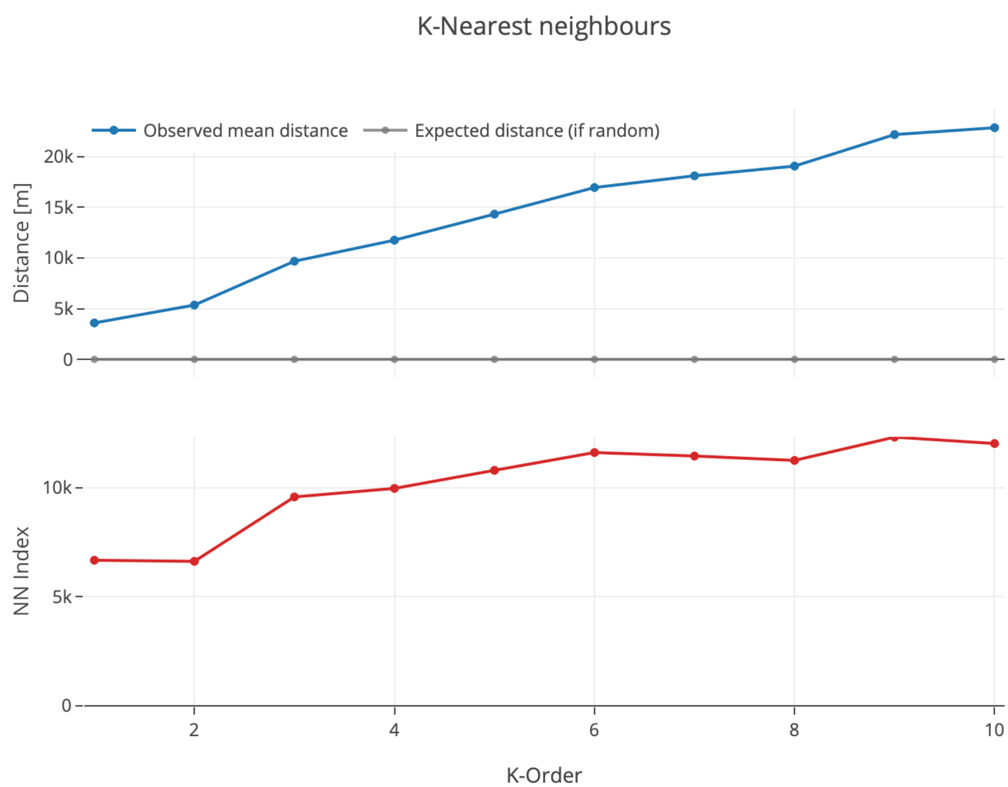


Figure 11. K-Nearest Neighbor Analysis.

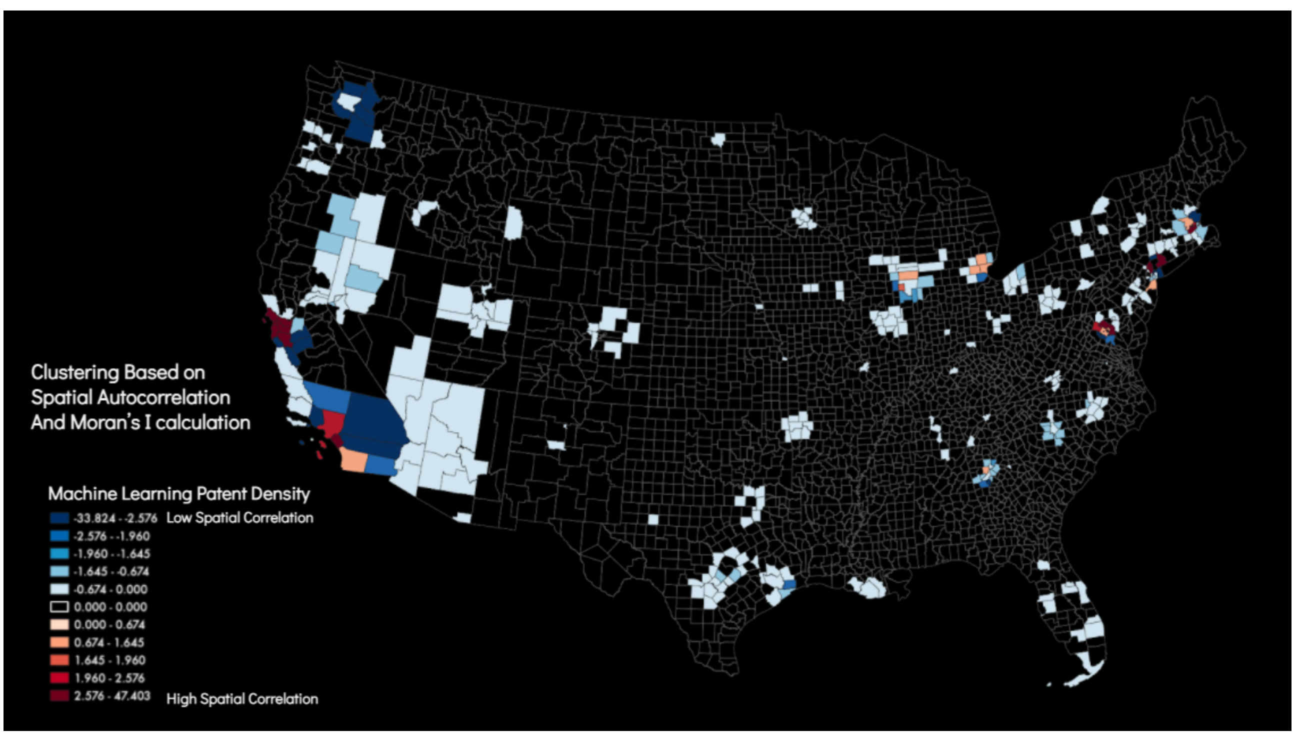


Figure 12. QGIS analysis 2021 | Moran's I Mapping, Lee Kuczewski

Personal Income

Personal Income by county region for 2019 was used in our calculation. We joined the County FIPS codes and tested that with a Pearson's r correlation (Figure 13) to better understand if income was connected with high or low patent counts per county. We found a positive weak correlation in our sample between numbers of patents per county and personal income expressed as +0.36.

$$r = \frac{n \cdot \sum XY - \sum X \cdot \sum Y}{\sqrt{\left[n \sum X^2 - (\sum X)^2\right] \cdot \left[n \sum Y^2 - (\sum Y)^2\right]}}$$

$$r = \frac{141 \cdot 227952703 - 2366 \cdot 9456550}{\sqrt{\left[141 \cdot 480500 - 2366^2\right] \cdot \left[141 \cdot 719887851418 - 9456550^2\right]}} \approx 0.3565$$

Figure 13. Pearson's r | Personal Income Correlation

Counties with minority populations greater than 50%

This mapping addresses the question of where counties with minority populations greater than 50% and patent counts greater than zero overlapped. There were 204 counties with minority populations greater than 50%.

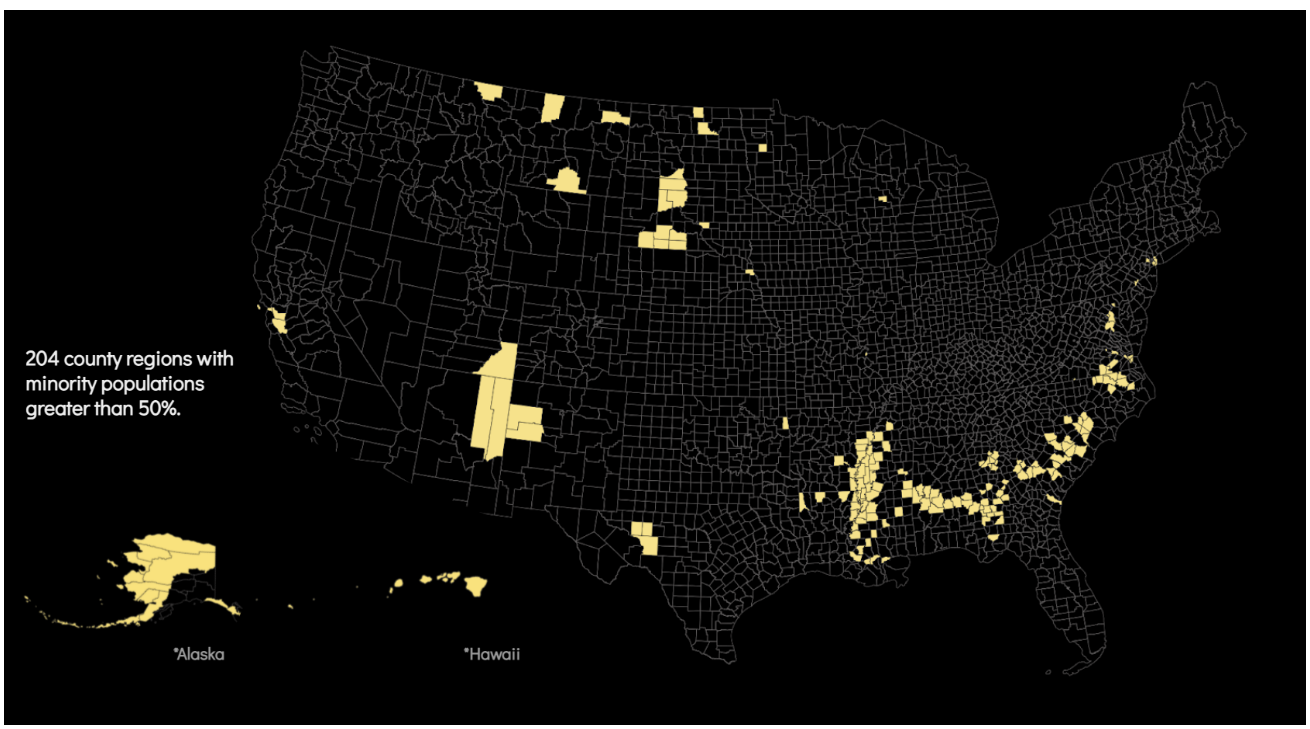


Figure 14. QGIS analysis 2021 | Minority population greater than 50%, Lee Kuczewski

Of the 204 counties with a minority population over 50%, we see that only 9.2% of the total counties filing machine learning patents have a minority population over 50%. Interestingly, the top three counties with a greater than 50% minority population accounted for 35.7% of all the machine learning patents in our study.

Figure 15. QGIS analysis 2021 | Minority population greater than 50% and greater than zero ML patents, Lee Kuczewski

5 Conclusion

"Any sufficiently advanced technology is indistinguishable from magic."

-Arthur C. Clarke

When the novelist and "cyberspace" coiner, William Gibson stated "The future is already here -- it's just not evenly distributed", he more than likely wasn't thinking about the effects of machine learning patent landscapes. His comment was prescient in many ways. First, it brings to focus the unequal distribution of knowledge, and secondly, the effects of that unequal distribution.

Signals From The Future is both concerned with the past as well as what we collectively do about our shared futures. Specifically, this research is about the decisions we make today and how we face the challenges and opportunities presented by AI technologies. We have an obligation as citizens to better understand the asymmetries and regional impact which these inventions are having on the greater society. Though further research is needed to understand additional connections and correlations between the complex network of the various actors within the patent landscape, this paper and visualization represents a snapshot of where we might begin to better understand our place in a quickly changing landscape.

This research paper opened with the possibility of an AI inventing patents, and the resulting effects that brought on questions about inventing, and importantly about being human. We all have a duty to understand where the impact is happening, and who to hold accountable if these impacts are at the expense of the communities they espouse to serve. We live in an age where the fiction from decades ago is no longer fiction, it's now our chance to maintain a critical perspective and to ensure we aren't leaving our fellow humans behind.

6 Appendix

Appendix 1. CPC Code Classification Schema from IPO, UK.

Bio-inspired approaches	ant-colony	bee-colony	differential*-evol* algorithm*
	evolution* algorithm*	evolution* comput*	fire-fly
	genetic program*	genetic* algorithm	memetic algorithm*
	particle-swarm*	swarm behav*	swarm intell*
Classification and regression trees	ensemble-learn*	fuzzy-c	kernel learn*
	k-means	multi* label* classif*	random-forest*
	spectral cluster*	vector-machine*	
Neural networks	adversar* network*	auto-encod*	back-propagat*
	back-propogat*	deep-belief network*	deep-learn*
	neural-network	neural-turing	neuro-morph comput*
	transfer-learn*		
Rule learning	association rule	q-learn*	
Supervised learning	collaborat* filter*	ensemble-learn*	factorisation machin*
	factorization machin*	feature engineer*	kernel learn*
	policy-gradient method	q-learn*	random-forest*
	recommender system*	reinforc* learn*	supervised learn*
Unsupervised learning	adversar* network*	auto-encod*	unsupervised learn*
Fuzzy logic	fuzzy environment*	fuzzy logic*	fuzzy number*
	fuzzy set*	fuzzy system*	fuzzy-c
	k-means		

Appendix 1. Artificial Intelligence -- An overview of AI patenting, IPO, United Kingdom ISBN 978-1-910790-61-8

Appendix 2. CPC Code Classification Schema from IPO, UK.

Table 2: Areas of the CPC scheme used by the search strategy				
B60G2600/1878	G05B13/0275	G06N3/004	G10L25/30	H04N21/4666
B60G2600/1879	G05B13/028	G06N3/02	G11B20/10518	H04Q2213/054
E21B2041/0028	G05B13/0285	G06N3/12	H01J2237/30427	H04Q2213/13343
F02D41/1405	G05B13/029	G06N5	H02P21/0014	H04Q2213/343
F03D7/046	G05B13/0295	G06N7	H02P23/0018	H04R25/507
F05B2270/707	G05B2219/33002	G06N20	H03H2017/0208	Y10S128/924
F05B2270/709	G05D1/0088	G06N99/005	H03H2222/04	Y10S128/925
F05D2270/707	G06F11/1476	G06T3/4046	H04L25/0254	Y10S706
F05D2270/709	G06F11/2257	G06T9/002	H04L25/03165	
F16H2061/0081	G06F11/2263	G06T2207/20081	H04L41/16	
F16H2061/0084	G06F15/18	G06T2207/20084	H04L45/08	
G01N29/4481	G06F17/16	G08B29/186	H04L2012/5686	

Table 2: Areas of the CPC scheme used by the search strategy				
A61B5/7267	G01N33/0034	G06F19/24	G10H2250/151	H04L2025/03464
B29C66/965	G01N2201/1296	G06F19/707	G10H2250/311	H04N21/4662
B29C2945/76979	G01S7/417	G06F2207/4824	G10K2210/3024	H04N21/4663
B60G2600/1876	G05B13/027	G06K7/1482	G10K2210/3038	H04N21/4665

Appendix 2. Artificial Intelligence -- An overview of AI patenting, IPO, United Kingdom ISBN 978-1-910790-61-8

Appendix 3. CPC Code Classification Schema from IPO, UK.

Table 3: Keywords used in the search strategy			
ant-colony	factorization machin*	high-dimensional* feature*	particle-swarm*
bee-colony	factorisation machin*	high-dimensional* input*	pattern-recogni*
fire-fly	feature engineer*	k-means	policy-gradient method
adversar* network*	feature extract*	kernel learn*	q-learn*
artificial*-intelligen*	feature select*	latent-variable*	random-forest*
association rule	fuzzy-c	link* predict*	recommender system*
auto-encod*	fuzzy environment*	machine intelligen*	reinforc* learn*
autonom* comput*	fuzzy logic*	machine learn*	sentiment* analy*
back-propagat*	fuzzy number*	map-reduce	sparse represent*
back-propogat*	fuzzy set*	memetic algorithm*	sparse*-code*
cognitiv* comput*	fuzzy system*	multi* label* classif*	spectral cluster*
collaborat* filter*	gaussian mixture model	multi*-objective* algorithm*	stochastic*-gradient*
deep-belief network*	gaussian process*	multi*-objective* optim*	*supervis* learn*
deep-learn*	genetic program*	natural-gradient	support-vector machine
differential*-evol* algorithm*	genetic* algorithm	neural-turing	swarm behav*
dimensional*-reduc*	high-dimensional* data	*neural-network*	swarm intell*
ensemble-learn*	high-dimensional* model*	neuro-morph comput*	transfer-learn*
evolution* algorithm*	high-dimensional* space*	non-negative matri* factor*	variation*-infer*
evolution* comput*	high-dimensional* system*	object-recogni*	vector-machine*

G10L, G06F17/2*	Speech recognition, Natural Language Processing
G06T	Image processing
G06K	Recognition of data
G06F17/50	Computer-Aided Design

Appendix 3. Artificial Intelligence -- An overview of AI patenting, IPO, United Kingdom ISBN 978-1-910790-61-8

7 References

Abood, Aaron and Dave Feltenberger. "Automated Patent Landscaping." *Artificial Intelligence and Law* 26, no. 2 (2018): 103-125. doi:10.1007/s10506-018-9222-4. http://link.springer.com/10.1007/s10506-018-9222-4.

Acs, Zoltan J. and David B. Audretsch. *Handbook of Entrepreneurship Research: An Interdisciplinary Survey and Introduction* New York, NY: Springer New York, 2010a.

Agrawal, Ajay, Joshua Gans, and Avi Golfarb. "The Economics of Artificial Intelligence: An Agenda."

Buchanan, Bruce G. "A (very) Brief History of Artificial Intelligence." 2005, 53+, https://link-gale-com.libproxy.newschool.edu/apps/doc/A145633110/AONE?u=nysl_me_newsch&sid=AONE&xid=206ddf27.

Delgado, Mercedes, Michael E. Porter, and Scott Stern. "Clusters and Entrepreneurship." *Journal of Economic Geography* 10, no. 4 (2010a): 495-518. http://www.jstor.org.libproxy.newschool.edu/stable/26161411.

Echeverri-Carroll, Elsie and Maryann P. Feldman. "Chasing Entrepreneurial Firms." *Industry and Innovation* 26, no. 5 (2019): 479-507. doi:10.1080/13662716.2018.1475220.

Feldman, Maryann P. "The Entrepreneurial Event Revisited"

Feldman, Maryann P., Serden Ozcan, and Toke Reichstein. "Falling Not Far from the Tree: Entrepreneurs and Organizational Heritage." *Organization Science* 30, no. 2 (2019): 337-360. doi:10.1287/orsc.2018.1222.

Fusco, Stefania. "Lessons from the Past: The Venetian Republic's Tailoring of Patent Protection to the Characteristics of the Invention." *SSRN Electronic Journal* (2019). doi:10.2139/ssrn.3331687.

Fusco, Stefania. *Number 3 (2020) Northwestern Journal of Technology and Intellectual Property*. Vol. 17.

Goertzel, Ben. "Artificial General Intelligence: Concept, State of the Art, and Future Prospects." *Journal of Artificial General Intelligence* 5, no. 1 (2014): 1-48. doi:http://dx.doi.org.libproxy.newschool.edu/10.2478/jagi-2014-0001..

Gompers, Paul A. and Joshua Lerner. *The Venture Capital Cycle*. Cambridge Mass.: MIT Press, 1999a. http://www.loc.gov/catdir/toc/fy034/99013957.html.

Huggins, Robert and Hiro Izushi. *Competition, Competitive Advantage, and Clusters: The Ideas of Michael Porter* Oxford: Oxford University Press, 2011a

Jelonek, Dorota, Ilona Pawełoszek, Cezary Stępniak, and Tomasz Turek. "Spatial Tools for Supporting Regional E-Entrepreneurship." *Procedia Computer Science* 65, (2015a): 988-995. doi:10.1016/j.procs.2015.09.061. https://www.sciencedirect.com/science/article/pii/S1877050915028914.

Kim, Jaeyoung, Janghyeok Yoon, Eunjeong Park, and Sungchul Choi. *Patent Document Clustering with Deep Embeddings* 2018. https://www.researchgate.net/profile/Sungchul_Choi2/publication/325251122_Patent_Document_Clustering_with_Deep_Embeddings/links/5b002920a6fdccf9e4f556da/Patent-Document-Clustering-with-Deep-Embeddings.pdf.

Kochan, Thomas A. and Lee Dyer. *Shaping the Future of Work: A Handbook for Action and a New Social Contract*. Abingdon, Oxon ; New York, NY: Routledge, 2021a.

Kulp, Patrick. "Generative AI Patent Filings Grow 500%: BRANDS ARE TESTING THE POTENTIAL BEHIND THIS EMERGING TECH." *Adweek (2003),* Oct 21, 2019, 7.

Liveley, Genevieve and Sam Thomas. *Homer's Intelligent Machines AI in Antiquity*.

Lyons, Thomas S., Roger E. Hamlin, and Amanda Hamlin. *Using Entrepreneurship and Social Innovation to Mitigate Wealth Inequality*. 1st ed. Boston: Boston: DEG Press, 2018a.

Maggitti, Patrick G., Ken G. Smith, and Riitta Katila. "The Complex Search Process of Invention." *Research Policy* 42, no. 1 (2013): 90-100. doi:10.1016/j.respol.2012.04.020. https://linkinghub.elsevier.com/retrieve/pii/S0048733312001370.

McCarthy, J., M. L. Minsky, N. Rochester, I. B. M. Corporation, and C. E. Shannon. "A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence." (: 13.

Plummer, Lawrence A. "Spatial Dependence in Entrepreneurship Research: Challenges and Methods." *Organizational Research Methods* 13, no. 1 (2010a): 146-175. doi:10.1177/1094428109334199.

Porter, Michael E. *The Competitive Advantage of Nations : With a New Introduction*. New York: New York : Free Press, 1998.

Porter, Michael E. "Location, Competition, and Economic Development: Local Clusters in a Global Economy." *Economic Development Quarterly* 14, no. 1 (2000a): 15-34. doi:10.1177/089124240001400105.

Pratheeban, M. "Patent Data Clustering: A Measuring Unit for Innovators." 1, no. 1 (: 8. http://www.iaeme.com/MasterAdmin/UploadFolder/PATENT%20DATA%20CLUSTERING-2.pdf.

Shane, Scott. *A General Theory of Entrepreneurship: The Individual-Opportunity Nexus*. Cheltenham, UK ; Northampton, MA, USA: E. Elgar, 2003a.

Strumsky, Deborah and José Lobo. "Identifying the Sources of Technological Novelty in the Process of Invention." *Research Policy* 44, no. 8 (2015a): 1445-1461. doi:10.1016/j.respol.2015.05.008. https://linkinghub.elsevier.com/retrieve/pii/S0048733315000840

Strumsky, Deborah, José Lobo, and Joseph A. Tainter. "Complexity and the Productivity of Innovation." *Systems Research and Behavioral Science* 27, no. 5 (2010a): 496-509. doi:10.1002/sres.1057. http://doi.wiley.com/10.1002/sres.1057

Team, Research, Joel R. Reidenberg, Stanley D. , Nikki Waxberg Chair, N. Cameron Russell, Maxim Price, and Anand Mohan. *Patents and Small Participants in the Smartphone Industry* 2015.

Thaler, Stephen. "In the United States Bankruptcy Court for the Eastern District of Virginia Alexandria Division." *New York Times (1923-Current File),* Oct 14, 2002. https://search.proquest.com/docview/92217070.

Thaler and L. Stephen. *Electro-Optical Device and Method for Identifying and Inducing Topological States Formed among Interconnecting Neural Modules*. Vol. 61924624. United States:.

Waters, Nigel. "Tobler's First Law of Geography." (2017). doi:10.1002/9781118786352.wbieg1011.

[2020] WLR(D) 526, [2020] EWHC 2412 (Pat), [2020] Bus LR 2146 (United Kingdom a)

CB Insights, Enterprise AI Trends To Watch in 2021. https://www.cbinsights.com/research/report/enterprise-ai-trends-2021/

"ClusterMapping.US Data API Documentation | U.S. Cluster Mapping." . http://clustermapping.us/content/clustermappingus-data-api-documentation

"The Competitive Advantage of Nations." *Harvard Business Review,* 1990, https://hbr.org/1990/03/the-competitive-advantage-of-nations

"Economic Research." . https://www.uspto.gov/ip-policy/economic-research

Financing Innovation in the United States, 1870 to Present, edited by Lamoreaux, Naomi R., Kenneth L. Sokoloff. Cambridge: MIT Press, 2007a. http://ebookcentral.proquest.com/lib/newschool/detail.action?docID=3338539.

"IP in China." . https://www.uspto.gov/ip-policy/ip-china https://www.uspto.gov/ip-policy/ip-china.

"The Lens - Free & Open Patent and Scholarly Search." . https://www.lens.org/lens https://www.lens.org/.

McCarthy Et Al. - A PROPOSAL FOR THE DARTMOUTH SUMMER RESEARCH PROJE g.

"Patft » Page 1 of 1." http://patft.uspto.gov/ http://patft.uspto.gov/.

Regional Data k. https://nvca.org/research/regional-data/ https://nvca.org/research/regional-data/.

"RPX Corporation S-1 1 Ds1.Htm FORM S-1." . https://www.sec.gov/Archives/edgar/data/1509432/000119312511240287/ds1.htm.

"Social Progress Imperative." . https://www.socialprogress.org/ https://www.socialprogress.org/.

"Supplementary Material - Inventing AI: Tracing the Diffusion of Artificial Intelligence with U.S. Patents." (m): 55. https://www.uspto.gov/sites/default/files/documents/OCE-DH-AI.pdf.

Tobler's First Law of Geography 2019. https://en.wikipedia.org/w/index.php?title=Tobler%27s_first_law_of_geography&oldid=929891532

"U.S. VC Median Deal Size by Stage 2020." . https://www-statista-com.libproxy.newschool.edu/statistics/829108/vc-median-deal-size-usa-by-stage/

"Why do Entrepreneurs Matter?" . https://www.kauffman.org/ecosystem-playbook-draft-3/entrepreneurs/ .