

**UNITED STATES PATENT & TRADEMARK OFFICE**

**229 YEARS OF DATA**



**Emerging Technologies in USPTO Business Solutions**

May 25<sup>th</sup> 2018

# STATS @



- ~8500 Patent Examiners;
- Receive ~600k patent applications a year; have ~605,000 applications in inventory.
- ~1.2M patent applications in flight.
- ~2.5M granted patent that are enforceable.
- Issue ~305,000 patents a year with ~5 million “claims” that can be litigated.
  
- *Time to issue* first millionth Patent – ~75.8 years;
- *Time to issue* our last millionth patent – ~3.6 years.
  
- 875 Trademark employees (7% of USPTO)
- 579 examining attorneys (66% of Trademarks)
- Trademark Operations receive ~594,100 classes for registration in FY 2017.



# Uses within the USPTO

---

- I. Patent Enriched Citation Data
- II. Patent CPC Analytics - Trends
- III. Browser Based End Point Advanced Analytics for Patents
- IV. Trademark Image Classifier
- V. Research Projects at the USPTO

# I. Enriched Citation & DocDB

- Extracting the associated citations and related rejection type using emerging technologies.
- This allows us to provide this information to the IP5 Office in an internationally consumable format.
- DocDB (granular international family mapping with IP Office specific citations) advanced analytical analysis.

# Developer Hub for Patent Filing Data

The screenshot shows the USPTO Open Data Portal homepage. At the top, it says "Open Data Portal (beta)". Below the header, there's a navigation bar with links like "Home", "About Us", "Contact Us", "Feedback", "Help", "Privacy", "Terms", "Accessibility", "Site Map", "Glossary", "FAQ", "Help", "Privacy", "Terms", "Accessibility", "Site Map", "Glossary", "FAQ". The main content area features a large banner with the text "Become innovative." and four blue boxes with icons and text: "Data can be beautiful.", "Share with us.", "Explore our data.", and "Make something." Below this, there's a section titled "Application Programming Interfaces" with three cards: "PatentsView", "BDSS", and "Unified Event API".

This screenshot shows a data visualization page on the USPTO Open Data Portal. The title is "Utility Patent Application & Grant Count (2 visuals)". It features a line chart showing the trend of utility patent applications and grants over time. Below the chart, there's a section titled "Explore. Discover. Create." with three smaller visualizations: "Utility Patent Application & Grant Count (2 visuals)", "Apple Finances by Utility Patent Count", and "Utility Patents per State over Time". On the right side, there's a "Created something with our data?" section with a "Let us know" button.

This screenshot shows the "USPTO APIs" section of the USPTO Open Data Portal. It features a large blue header with the text "USPTO APIs" and a sub-header "See what new APIs The US Patent and Trademark Office has created with...". Below this, there's a "What is an API?" section with a "Let us know" button. At the bottom, there are three cards for "PatentsView", "BDSS", and "Unified Event API", each with a "View API" button.

This screenshot shows the "Our community" section of the USPTO Open Data Portal. It features a large blue header with the text "Our community" and a sub-header "The open data community is a collection of like-minded developers, entrepreneurs and business professionals who share ideas, stories and resources on USPTO data." Below this, there's a "Search for tags" section with a search bar and a "Filter by" dropdown. The main content area is a grid of featured stories, including "Mobile App for 2019 Government Services", "Using Data to Drive Your Entrepreneurship", "Models to Help Business Manager Change", "Consumer Finance Protection Board Update Site", "The Normal Entrepreneur", "Successful Grants", "Magic Emp - Coaster from east", "Improving Your Data Quality", and "VibeCity Atlanta".

<https://developer.uspto.gov/>

# https://developer.uspto.gov/api-catalog/

Open Data Portal <sup>beta</sup>

Home Explore Learn Build About USPTO open data



## USPTO Office Action Rejection API (beta)

v1 | Patents

This API allows for retrieval rejection data from 4.4 million Office actions mailed during the 2008 to mid-2017 period from USPTO examiners to the applicants of 2.2 million unique patent applications. It contains document level data including the type of actions taken on claims in the office action.

The United States Patent and Trademark Office (USPTO) Office Action Research Dataset for Patents contains detailed information derived from the Office actions issued by patent examiners to applicants during the patent examination process. The "Office action" is a written notification to the applicant of the examiner's decision on patentability. It generally discloses the reasons for any rejections, objections, or requirements and includes relevant information or references that the applicant may find useful for responding to the examiner and deciding whether to continue prosecuting the application.

This API allows for retrieval rejection data from 4.4 million Office actions mailed during the 2008 to mid-2017 period from USPTO examiners to the applicants of 2.2 million unique patent applications. It contains document level data including the type of

- API Syntax
- Patent Application Office Actions Research Dataset
- USPTO Patent Prosecution Research Data: Unlocking Office Action Traits Working Paper

Last updated: Mar 1st, 2018

Created something with our data?

Let us know

Share your work with the USPTO Open Data Community - use the button below to submit your work.

Share your work!

uspto

About Us Jobs Contact Us

Open Data Portal <sup>beta</sup>

Home Explore Learn Build About USPTO open data



## USPTO Office Action Citations API (beta)

v1 | Patents

This API allows for retrieval of citation data from 4.4 million Office actions mailed during the 2008 to mid-2017 period from USPTO examiners to the applicants of 2.2 million unique patent applications. It uses information derived from citations referenced on the Form PTO-892, Form PTO-1449, and text of Office actions.

The United States Patent and Trademark Office (USPTO) Office Action Research Dataset for Patents contains detailed information derived from the Office actions issued by patent examiners to applicants during the patent examination process. The "Office action" is a written notification to the applicant of the examiner's decision on patentability. It generally discloses the reasons for any rejections, objections, or requirements and includes relevant information or references that the applicant may find useful for responding to the examiner and deciding whether to continue prosecuting the application.

This API allows for retrieval of citation data from 4.4 million Office actions mailed during the 2008 to mid-2017 period from USPTO examiners to the applicants of 2.2 million unique patent applications. It uses information derived from citations referenced on the Form PTO-892, Form PTO-1449, and text of Office actions.

- API Syntax
- Patent Application Office Actions Research Dataset
- USPTO Patent Prosecution Research Data: Unlocking Office Action Traits Working Paper

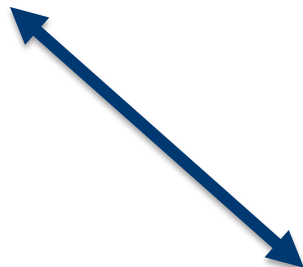
Last updated: Mar 1st, 2018

Created something with our data?

Let us know

Share your work with the USPTO Open Data Community - use the button below to submit your work.

Share your work!



# II. CPC Classifying Trends/Analytics

- **Data Sources**

- CPC Scheme Definition
  - Concatenated the definition to the root using XML parser
- Patents/PGPubs Specifications
  - Summary text, a mix of “Background of invention”, “Brief description of the invention”, etc.
- Patents/PGPubs Claims
  - Independent and dependent claims of patents

- **Machine Learning/AI Algorithms**

- Supervised/Unsupervised Machine Learning Algorithms
- Ensemble Algorithms

# Prototype Interface

The screenshot displays the 'CPC Explorer' application window. At the top, there is a navigation bar with 'Text' and 'Render' tabs, and an 'Analyze' button. The main content area shows a text input field containing a patent claim snippet, followed by a row of buttons for various CPC codes: H03M13/6362, H03M13/2957, H03M13/2792, H04L1/0071, H04L1/0065, H04L1/0066, H04L1/0057, and H04L1/0067. Below this is a 'CPC Codes' input field, a 'Reset' button, an 'App ID' input field, and a 'Classify' button. The 'Analysis' section is active, showing a table with columns for ID, Claim, CPC, Score, and Sentence. Two rows are visible, each with a '+' icon in the ID column. A search bar and a dropdown menu with options like JSON, XML, CSV, TXT, SQL, and MS-Excel are also present.

6. (Currently Amended) Awireless communication apparatus, comprising a processor and a non-transitory memory having processor-executable instructions stored thereon, wherein the processor-executable instructions, when executed by the processor, facilitate performance of the following: obtaining a system polar code by encoding data to be transferred, wherein the system polar code corresponds to a generator matrix; dividing the system polar code output into system bits and parity bits, wherein the system bits correspond to a unit matrix part in the generator matrix and the parity bits correspond to a check matrix part in the generator matrix; performing quadratic interleaving on the system bits to obtain a first group of interleaved bits, and quadratic interleaving on the parity bits to obtain a second group of interleaved bits, wherein a mapping function of the quadratic interleaving is such that a  $c(m)$ th bit before interleaving is mapped onto a  $c(m+1)(\text{mod } N)$ th bit after interleaving; and determining a rate-matched output sequence based on the first group of interleaved bits and the second group of interleaved bits.

H03M13/6362 H03M13/2957 H03M13/2792 H04L1/0071 H04L1/0065 H04L1/0066 H04L1/0057 H04L1/0067

CPC Codes   App ID

Analysis Summary Selections

Search

ID	Claim	CPC	Score	Sentence
+ 18	1	H03M13/2792	0.1368	1. (Currently Amended) A polar code rate matching method, comprising: obtaining, by a wireless communication apparatus, a encoding data to be transferred, wherein the system polar code corresponds to a generator matrix; dividing, by the wireless communication apparatus, the system polar code into system bits and parity bits, wherein the system bits correspond to a unit matrix part in the generator matrix and the parity bits correspond to a check matrix part in the generator matrix; performing, by the wireless communication apparatus, quadratic interleaving on the system bits to obtain a first group of interleaved bits, and quadratic interleaving on the parity bits to obtain a second group of interleaved bits, wherein a mapping function of the quadratic interleaving is $isc(m) \rightarrow c(m+1) \pmod{N}$
+ 23	1	H04L1/0071	0.0927	1. (Currently Amended) A polar code rate matching method, comprising: obtaining, by a wireless communication apparatus, a system polar code by encoding data to be transferred, wherein the system polar code corresponds to a generator matrix; dividing, by the wireless communication apparatus, the system polar code into system bits and parity bits, wherein the system bits correspond to a unit matrix part in the generator matrix and the parity bits correspond to a check matrix part in the generator matrix; performing, by the wireless communication apparatus, quadratic interleaving on the

Showing 1 to 10 of 14 rows  rows per page

- Classifies entered text (ie patent application)
- “Explains” why/how algorithm arrived at its conclusion
- Enables human validation and model training



# III. Browser based end point Advanced Analytics for Patents

The screenshot displays a patent analysis tool interface. On the left, a document viewer shows a patent abstract with highlighted terms: "access", "radio", "ap", "resource", "set", "point", "control", "resources", "aps", "controls", "without", "rrc", "parts", "part". Below this is a search results table with columns for Select, Result #, Viewed, and a grid of checkboxes. A "Concept Text" window is open in the foreground, showing a list of terms and queries. A blue arrow points from the "Concept Text" window to a patent document on the right, which includes a network diagram. A text box at the bottom right states: "AUTOMATICALLY GENERATE MEANINGFUL RELATED CONCEPTS OR SYNONYMS FROM ANY WEB-TOOL TO AUGMENT PE2E SEARCH STRINGS".

Document Viewer (Left):  
Text | Images  
KWIC Off | Settings | Find Within  
Highlight: access radio ap resource set point control resources aps controls without rrc parts part Highlights

Document Viewer (Right):  
Text | Images  
100% | Full  
Notes Off  
United States Patent Application Publication  
(19) Patent No.: US 2010/0000000 A1  
(12) Patent Application Publication  
(43) Pub. Date: 2010.11.2010  
(54) MEANS AND METHOD FOR CONTROLLING NETWORK ACCESS IN INTEGRATED COMMUNICATIONS NETWORKS  
(57) US 2010/0000000 A1  
(76) Inventor: Jonathan Sachs, Aschen (US)

Search Results (Bottom Left):  
Hit Terms  
Highlight: part control set without portion parts point limited  
L23: 29 results found. Currently displaying all results. Filtered by Family ID (17)

Concept Text (Center):  
Type | Query | Action  
Terms: "access", "network domain", "criterion", "authorization entity", "radio resource"  
types: Terms  
query: "access", "network domain", "criterion", "authorization entity", "radio resource"  
access network domain criterion authorization entity radio resource  
Speak CPC Codes Search  
+ Terms: "access", "radio", "radio resource", "message", "user"  
+ Terms: "criterion", "traffic", "parameter value", "load", "user"  
+ Terms: "criterion", "traffic", "information", "s.sub-2", "authorizes uts"  
+ Terms: "access", "network domain", "radio", "method", "function"

Network Diagram (Right):  
340 UT  
330 RNC A2 MRRM  
330 Core Network  
330 A1  
305 AP  
305 APC  
330 Access Network 2 (M-2)

Automatically Generate Meaningful Related Concepts or Synonyms from any Web-tool to Augment PE2E Search Strings

# Platform AI/BD Capabilities

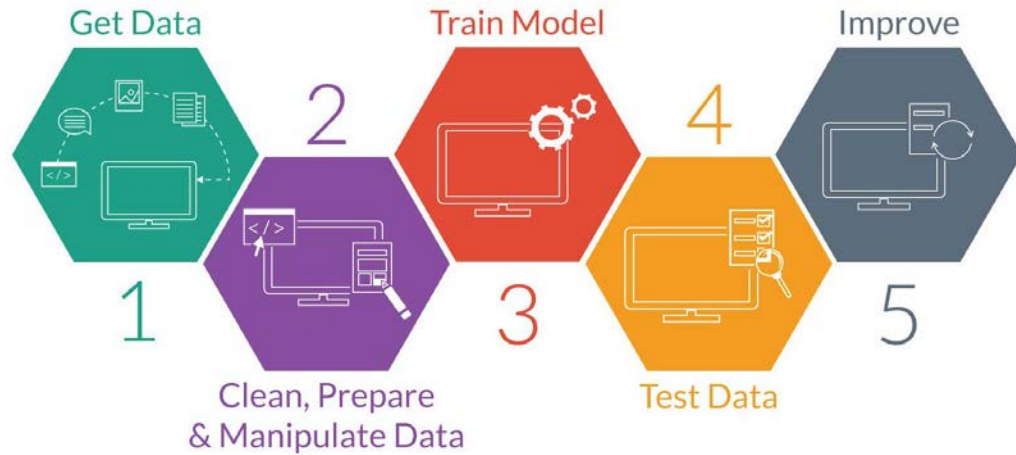
The screenshot displays a patent database interface with the following components:

- Search Bar:** Located at the top, containing the search term "MEANS AND METHOD FOR CONTROLLING NETWORK ACCESS IN INTEGRATED COMMUNICATIONS NETWORKS".
- Application Details:** A sidebar on the left showing information for Application No. 11/993,669, including the title "MEANS AND METHOD FOR CONTROLLING NETWORK ACCESS IN INTEGRATED COMMUNICATIONS NETWORKS" and examiner "LESNIEWSKI, Victor (2493)".
- Claims Tree:** A tree view on the left showing a hierarchy of claims, with claim 1 selected.
- Claim 1 Text:** A large text area displaying the text of claim 1, which describes a method for assisting the control of a User Terminal's (UT's) access to an access network domain. The text is partially highlighted in grey.
- Search Results:** A table below claim 1 listing search results. The table has columns for ID, Pub #, Abstract, Origin, Language, and Score.
 

ID	Pub #	Abstract	Origin	Language	Score
1	US2009182028A1	A method of controlling transmission rate by adjusting error correction packets and a communication apparatus using the same is provided. According to an embodiment of the present invention, the method of controlling transmission rate by adjusting error correction packets includes transmitting a first packet group received on a wireless radio link of data packets, and error correction packets, transmitting a second packet group composed of the data packets and the error correction packets in a ratio adjusted based on feedback information about the first packet group, and controlling the transmission rate based on the feedback information about the second packet group.	Google	en	0.22
2	EP2112427A1	The present disclosure relates to a method for controlling the transmission rate of a transmission channel. The method includes receiving a first packet group received on a wireless radio link of data packets, transmitting a second packet group composed of the data packets and the error correction packets in a ratio adjusted based on feedback information about the first packet group, and controlling the transmission rate based on the feedback information about the second packet group.	France	en	0.19
3	US2009182028A1	Methods and apparatus for adjusting error correction packets in a data network, a method provided for operating a network in a data network. The method includes computing a multi-set cost indicator and a network cost indicator that are associated with multiple and varied transmission of data. The method also includes determining that the network cost indicator is less than the network cost indicator, and generating a frame retransmission schedule that describes when the data will be retransmitted on the data network. The method also includes representing the transmission schedule, and multicasting the data according to the frame retransmission schedule.	Google	en	0.38
4	US2009182028A1	A network system comprises one or more servers to receive compressed video files for a group of users, each prior to capturing and/or receiving a digital picture image. The system also includes a server to receive the digital picture image, and to generate a data packet for each picture image. The system also includes a server to receive the data packet, and to generate a data packet for each picture image. The system also includes a server to receive the data packet, and to generate a data packet for each picture image.	Google	en	0.37
- Claim 10 Text:** A separate text box at the bottom of the interface showing the text of claim 10, which is partially highlighted in grey.
- Annotations:** A blue arrow points from the search results table to the text of claim 1. A red arrow points from the search results table to the text of claim 10.

**AGGREGATED "ONE-CLICK" PATENT, PRE-GRANT, FOREIGN AND NLP THROUGH ENSEMBLE AI RANKED RELEVANCY SEARCH RESULTS**

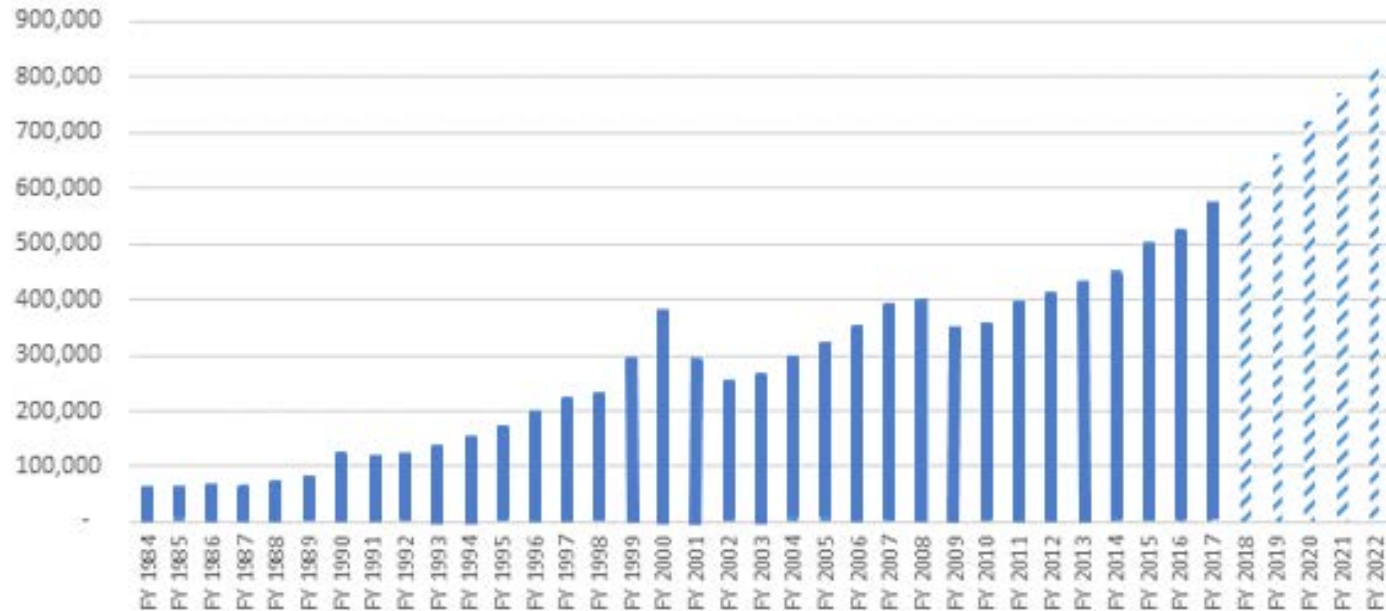
# IV. Trademark Machine Learning Framework



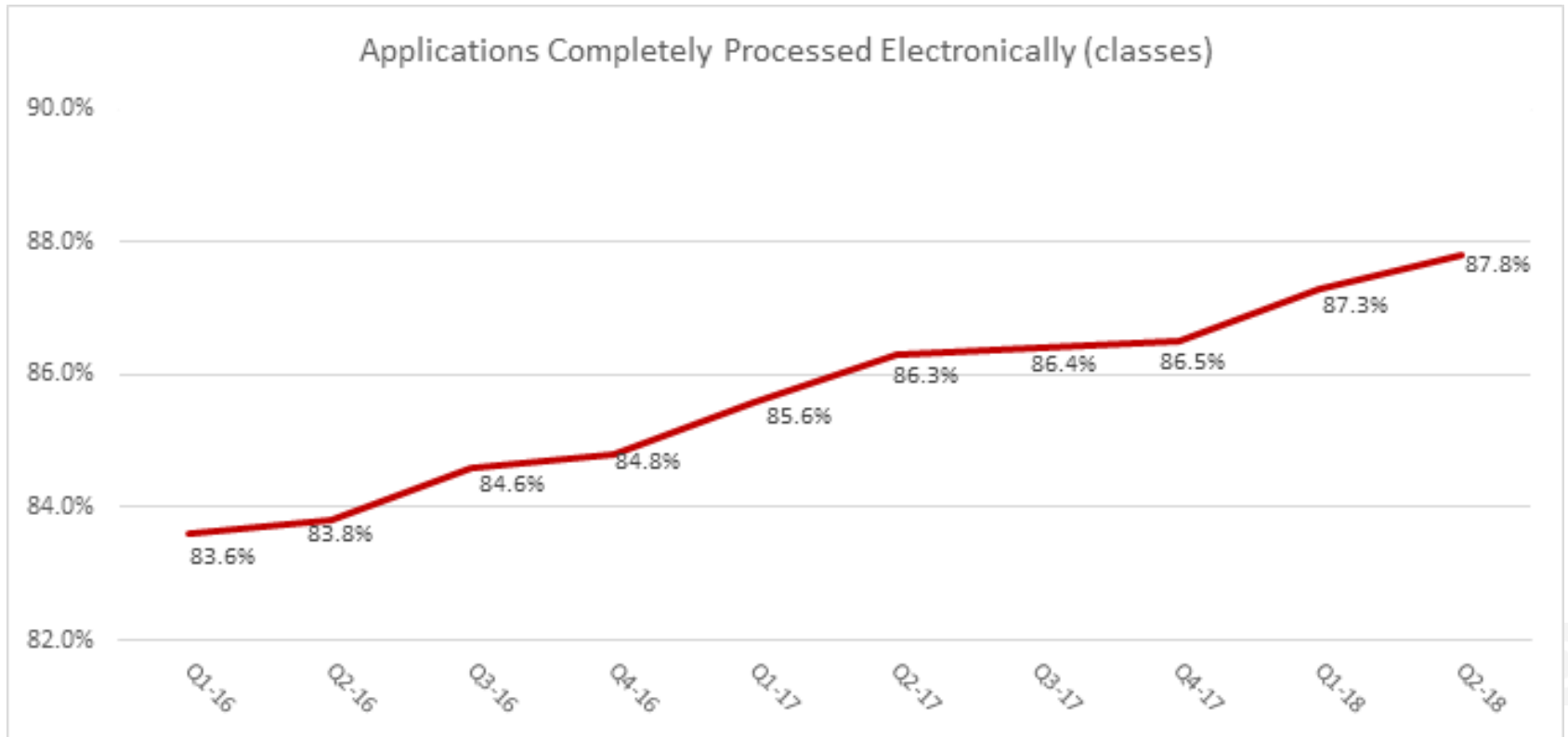
# New Application Filings

- FY16: 530,270 classes filed FY17: 594,107 classes filed
- FY18 up 8.7% compared to FY17. Expecting 646,000 classes this year

Applications for Registration - Classes

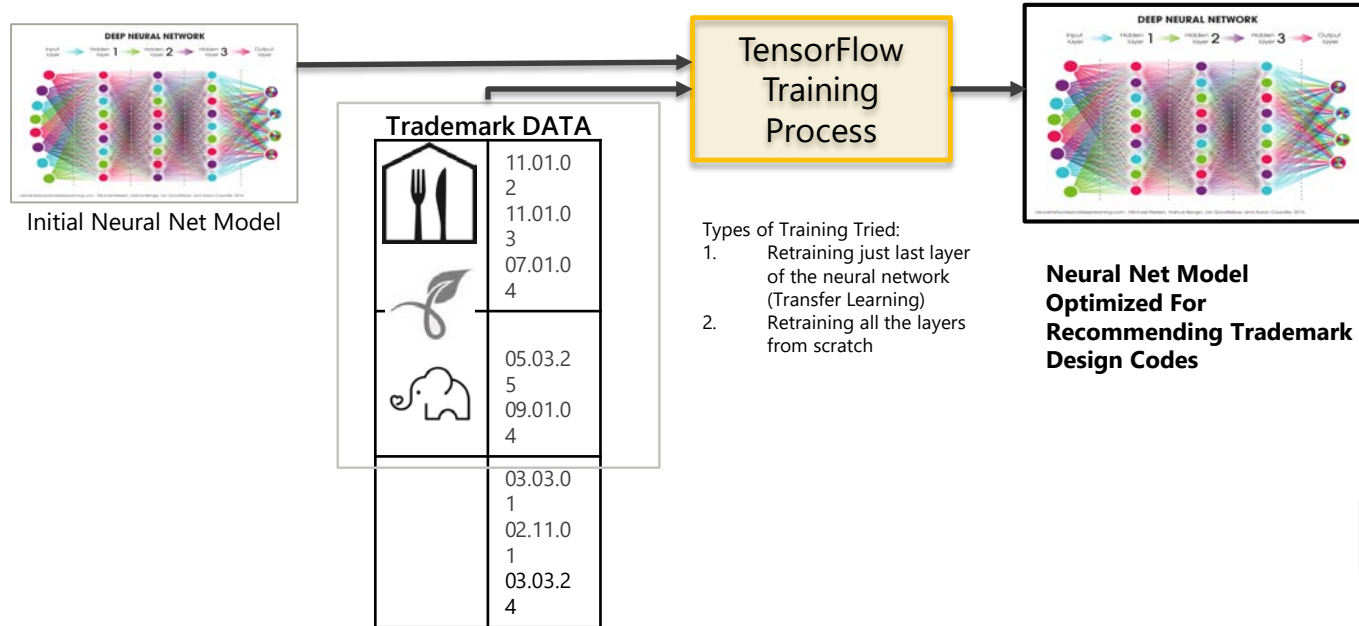


# Electronic TM Prosecution, End to End



# Trademark Design Code Suggestions

A six-digit numerical design search code is assigned to each design element of a trademark, such as a depiction of a star (01.01.03) or flower(05.05.25). Using years of images with corresponding examiner-annotated design codes, we are able to train deep learning systems that can predict design codes of a new trademark image.



# Design Code Based on Image

uspto

Trademark Ensemble Classification **Experimental**

Serial Number

Enter a serial number

or

Choose from the Gallery



TSDR Design Search Code(s)

- 01.01.10 - Three or more stars; Stars, three or more
- 01.01.13 - Stars - multiple stars with five points
- 02.03.01 - Busts of women facing forward; Heads of women facing forward; Portraits of women facing forward; Women - head, portrait or busts facing forward
- 02.03.24 - Women, stylized, including women depicted in caricature form
- 04.03.03 - Mermaids; Triton
- 26.01.12 - Circles with bars, bands and lines
- 26.01.17 - Concentric circles, two; Circles, two
- 26.01.21 - Circles that are totally or partially stacked.

Mark Description

The mark consists of the wording "STARBUCKS COFFEE" and design elements.comprising a siren (a two-tailed mermaid) wearing a crown with a star, two separate stars, and several circular creases.

Submit

**Design Codes On Record**

PREDICTED RESULTS

	DESIGN CODE	DESIGN CODE DESCRIPTION	INDICATORS	CONFIDENCE
<input checked="" type="checkbox"/>	04.03.03	Mermaids, Triton	A	<div style="width: 100%;"></div>
<input checked="" type="checkbox"/>	01.01.13	Stars - multiple stars with five points	A	<div style="width: 100%;"></div>
<input checked="" type="checkbox"/>	26.01.21	Circles that are totally or partially shaded.	A	<div style="width: 100%;"></div>
<input checked="" type="checkbox"/>	01.01.10	Stars, three or more, Three or more stars	A	<div style="width: 100%;"></div>
	26.01.18	Circles, three or more concentric, Concentric circles, three or more, Three or more concentric circles	A	<div style="width: 100%;"></div>
	24.11.02	Crowns open at the top	A	<div style="width: 100%;"></div>
<input checked="" type="checkbox"/>	02.03.24	Women, stylized, including women depicted in caricature form	A	<div style="width: 100%;"></div>
<input checked="" type="checkbox"/>	02.03.01	Busts of women facing forward, Heads of women facing forward, Portraits of women facing forward, Women - head, portrait or busts facing forward	A	<div style="width: 100%;"></div>
	24.11.01	Crowns closed at the top	A	<div style="width: 100%;"></div>
<input checked="" type="checkbox"/>	26.01.17	Circles, two concentric, Concentric circles, two, Two concentric circles	A	<div style="width: 100%;"></div>

**Neural Net Recommended Design Codes**

**Confidence scores**

MISSED PREDICTIONS

DESIGN CODE	DESIGN CODE DESCRIPTION	INDICATORS
26.01.12	Circles with bars, bands and lines	A

PREDICTIONS BY IMAGE MODEL

DESIGN CODE	DESIGN CODE DESCRIPTION
02.03.03	Princesses, Queens, Women wearing crowns or other symbols of royalty, including queens and princesses, Crowns (women wearing)
26.01.08	Circles having letters or numerals as a border, Circles having punctuation as a border, Letters, numerals or punctuation forming or bordering the perimeter of a circle
26.17.02	Bands, wavy, Bars, wavy, Lines, wavy, Wavy line(s), band(s) or bar(s)

PREDICTIONS BY TEXT MODEL

DESIGN CODE	DESIGN CODE DESCRIPTION
01.01.03	Star - a single star with five points
01.01.09	Stars, two, Two stars

# V. Research Projects at the USPTO

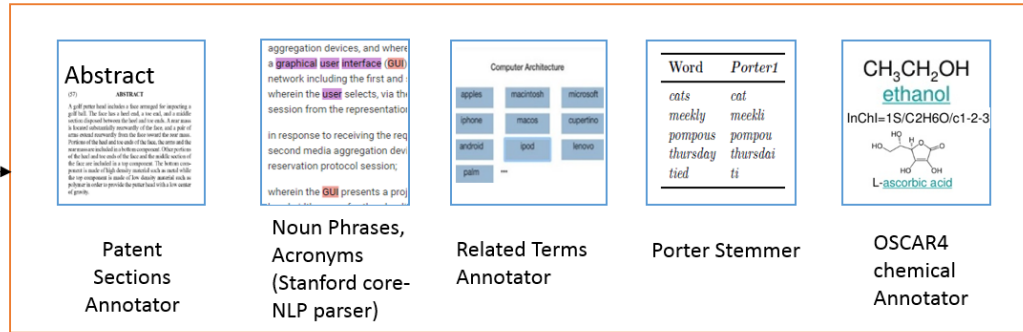
- AI Assisted Patent Searching
- AI based Patent Term Library Generator
- AI based Trademark Image Search
- Deep Machine Learning Chat Bots



# AI Patent Search Capabilities

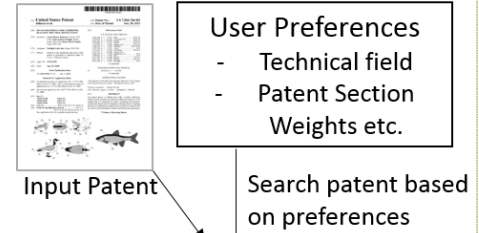
## AI-BASED PATENT PROCESSING

### INGESTION PIPELINE



US Patent & Pre-Grant publications as xml

## SEARCHING

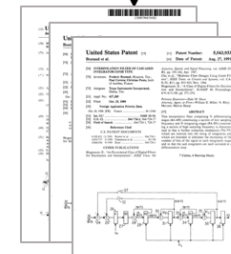


### SEARCH SYSTEM

Solr/Lucene

- fast, scalable, widely-used
- "more-like-this" option allows searching based on an entire document instead of key words
- supports multiple retrieval calculations such as TF/IDF or BM25

retrieves



Related Patents

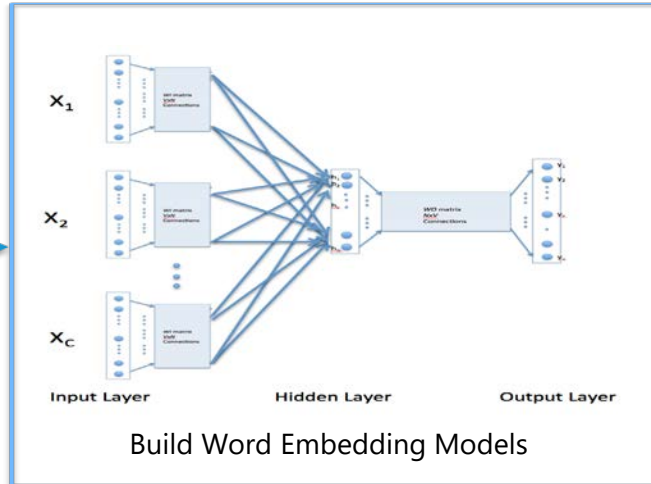
# Patent Synonyms Generator Tool

THE BELL SYSTEM  
TECHNICAL JOURNAL

IEEE 802 Standard documents

Corpus  
Patent Applications,  
Technical Journals,  
Technical Standards,...

Training



Generating  
Synonyms  
By Technical  
Fields

Videophones and Telephonic Communications; Audio Signals; Digital Audio Data Processing; Linguistics, Speech Processing and Audio Compression

media	storing	medium
removable	store	data
nontape	memory	computer-accessible
retrieval	persistent	...

Cable and Television

nonvolatile	storing	store
stored	medium	volatile
removable	memory	stores
cdrom	computer-readable	...

Process, Nucleic acid, Protein, Carbohydrate Chemistries and Diagnostics

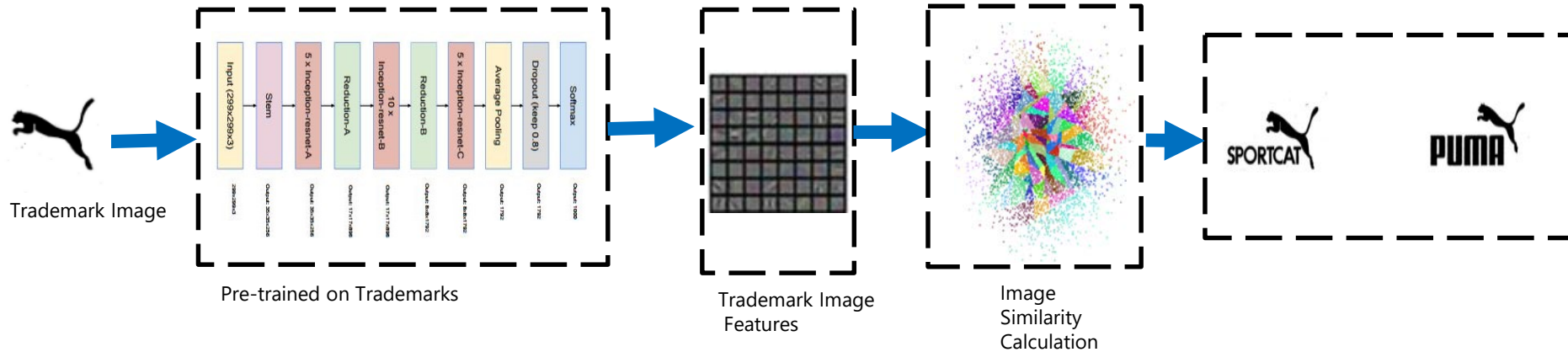
storage	storing	stored
cargo	destination	store
onboard	pipelines	tank
pipeline	offloading	...

Computer Architecture

disk	memory	storages
disks	ssd	nonvolatile
backup	storing	raid
volatile	...	

# Trademark Image Search

Searching for marks that look similar is an essential job function of a trademark examiner. Using neural networks we are able to retrieve and store features of mark images that we can then compare to other marks' features.



# TM Image Search Prototype using Machine Learning

uspto

Trademark Image Classification **Experimental**


Serial Number

or





**Choose from the Gallery**

or

**Upload your own**  
Please note, we only accept JPG documents.


**Input Image** → 

DESIGN CODE SUGGESTIONS    **REVERSE IMAGE SEARCH** ← **Uses Machine Learning**

			
87083516	87072116	86969706	79193066

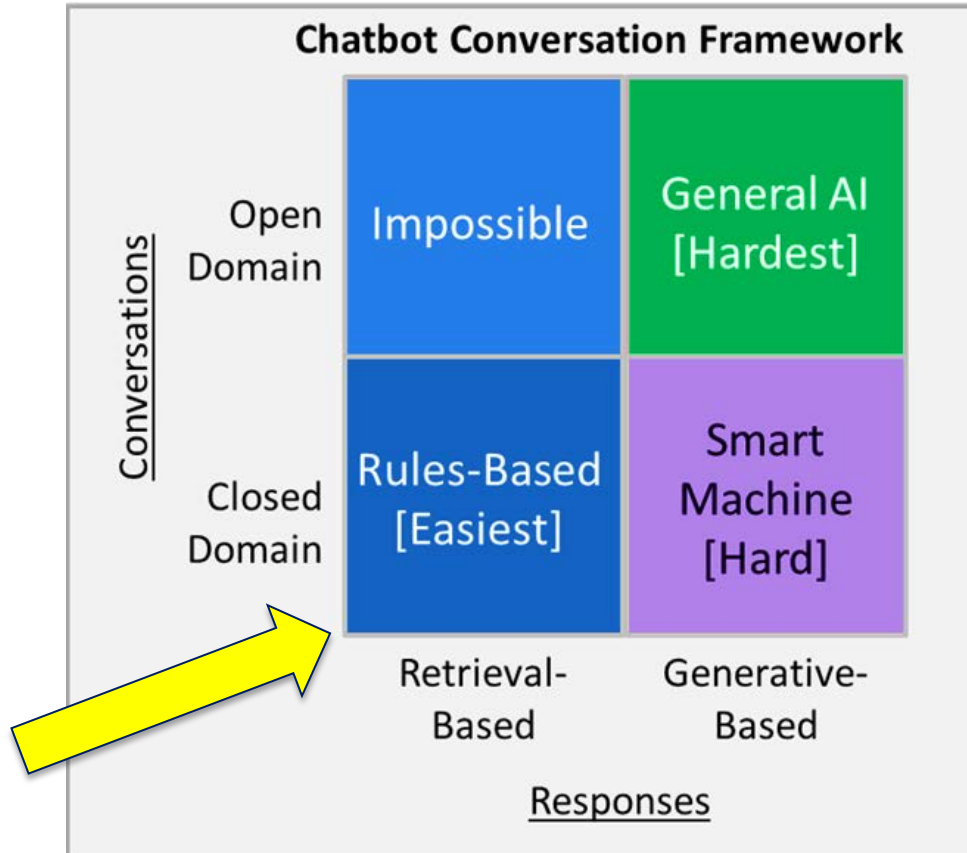
TSDR Design Search Code(s)

- 24.11.02 - Crowns open at the top

**Results of Matching Images** → 

Choose Model

# V. Leveraging Deep Learning for Chatbot Service



# Prototype Chatbot Service

User: what's 101

---

Wolcott: It is type of rejection based on whether or not the claimed subject matter is subject matter eligible 35 U.S.C. §101 "Whoever invents or discovers any new and useful process, machine, manufacture, composition of matter, or any new and useful improvement thereof, may obtain a patent therefor...."

---

User: how about 102

---

Wolcott: It is type of rejection based on whether or not the claimed subject matter is novel.

---

User: finally, what's 103

---

Wolcott: It is type of rejection based on whether or not the claimed subject matter is obvious. A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

---

User: rules to go final

---

Wolcott: Section 2271 of the MPEP - Final Action: Before a final action is in order, a clear issue should be developed between the examiner and the patent owner. To bring the prosecution to a speedy conclusion and at the same time deal justly with the patent owner and the public, the examiner will twice provide the patent owner with such information and references as may be useful in defining the position of the Office as to unpatentability before the action is made final. In making the final rejection, all outstanding grounds of rejection of record should be carefully reviewed and any grounds or rejection relied on should be reiterated. The grounds of rejection must (in the final rejection) be clearly developed to such an extent that the patent owner may readily judge the advisability of an appeal.

