

Cipher Classifiers – Your view of the technology space

20th June 2023

Presenters



Nigel Swycher

CEO
LexisNexis Cipher



Sebastian Muller-Borges

Head of Cipher Certified
LexisNexis Cipher



Sarah Newton

Head of Sales, North
America
LexisNexis Cipher



William Mansfield

Head of Consulting and
Customer Success
LexisNexis Intellectual
Property Solutions

Contents

01 Why Patent Classification

02 How does Patent Classification work

03 How do customers use Classifiers

04 Patent Classification in PatentSight

05 Q&A



Delivering best in class classification and analysis



Powerful AI Technology Classification

Industry leading technology classification through AI. Simplifying and enabling ongoing and high quality patent technology searching



Intelligent Patent Analytics Platform

Business-focused analytics engine designed specifically for patent data. Highly processes and cleaned data enabling quick routes to insights.

Why patent classification?

Why PatentSight + Cipher?



More complex & interconnected technology

More and more players are stepping out of their traditional markets to leverage their technology in other domains, for example digitation.



Growing number of patent filings

As emerging nations develop and more actively adopt a robust patent system more patents are filed, like China. Even in more developed nations filings increase year on year



Increasing demands from the business

The business is becoming more and more knowledgeable about what insights can be derived from IP. Organizations like PDG have also worked hard to develop this.



No significant growth in IP head count

IP team head count is not growing inline with these increased demands and complexity

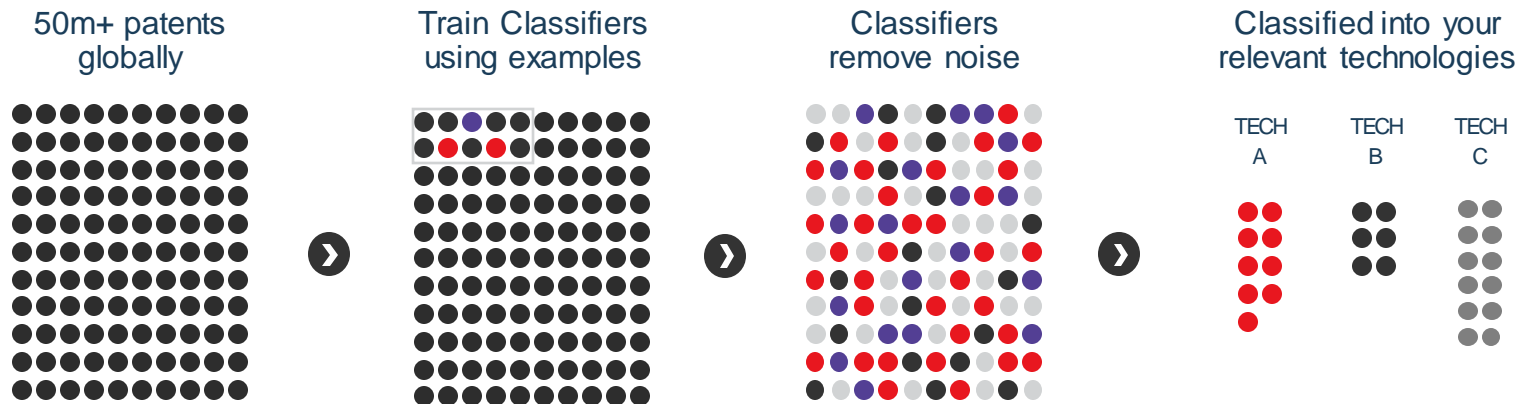


Need to do more with the same

IP teams need to be able to do more and delivery high quality results with the same or even fewer resources. IP tools need to support teams to realize this.

Cipher Classification

Classifiers are **supervised machine learning algorithms** trained with positive and negative patent examples that allow you to find all patents relating to specific technologies



Strategic Patent Intelligence



Benchmarking

Compare a patent portfolio to other owners through a technology lens



Budget Management

Justify patent budgets to CFOs and others to communicate the impact of an investment



Portfolio Optimisation

Ensure the right portfolio to meet strategic patenting objectives



Competitive Intelligence

Understand who's doing what by automating patent to technology mapping



Technology Trends

Understand, monitor and react to the latest technology trends



Due Diligence

Automate manual reviews for efficient execution of M&A and licensing transactions



Inbound Assertion

Be prepared with evidence to create a fast and effective threat assessment



Monetisation

Identify opportunities to create value through licensing or sale of patent assets



Cross-licensing

Combine patent and revenue data to determine rational licensing outcomes



Risk Mitigation

Understand, quantify and communicate patent portfolio risk

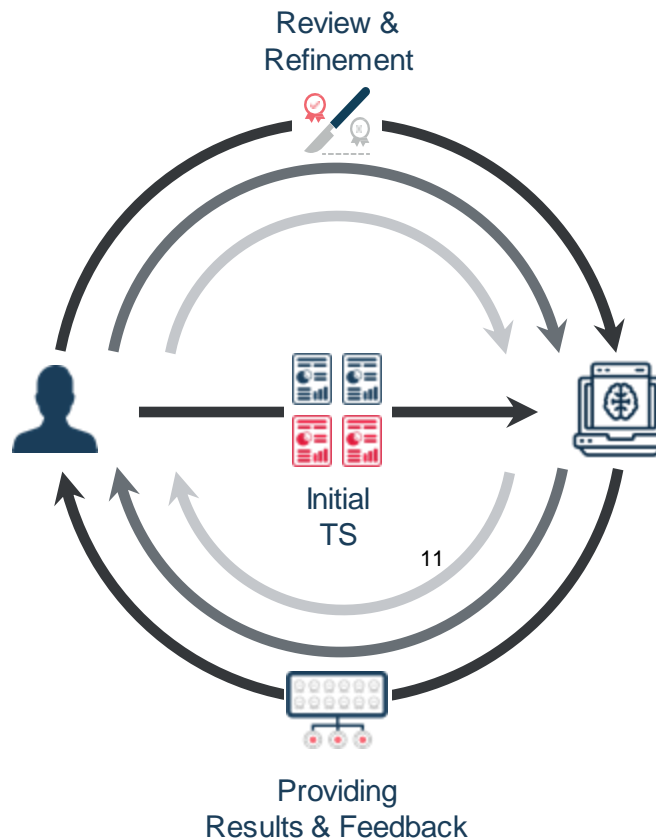
Loyal customers in both Europe and US

| | | | |
|-----------------------|-------------|-------------------------------------------------|----------------------------------------------|
| Automotive | DAIMLER | NISSAN | tu simple |
| Consumer | BAT | Firmenich <small>for good, naturally</small> | Unilever |
| Finance | AON | BARCLAYS | iD dealroom.co |
| Industrials | BAE SYSTEMS | Rolls-Royce | SCHOTT <small>glass made of ideas</small> |
| Semiconductors | arm | infineon | NVIDIA |
| Technology | Meta | DOLBY | Google |
| Professional Services | Cravath | LATHAM & WATKINS <small>LLP</small> | RICHARDSON OLIVER INSIGHTS |

How do Cipher Classifiers work?

Building a Classifier

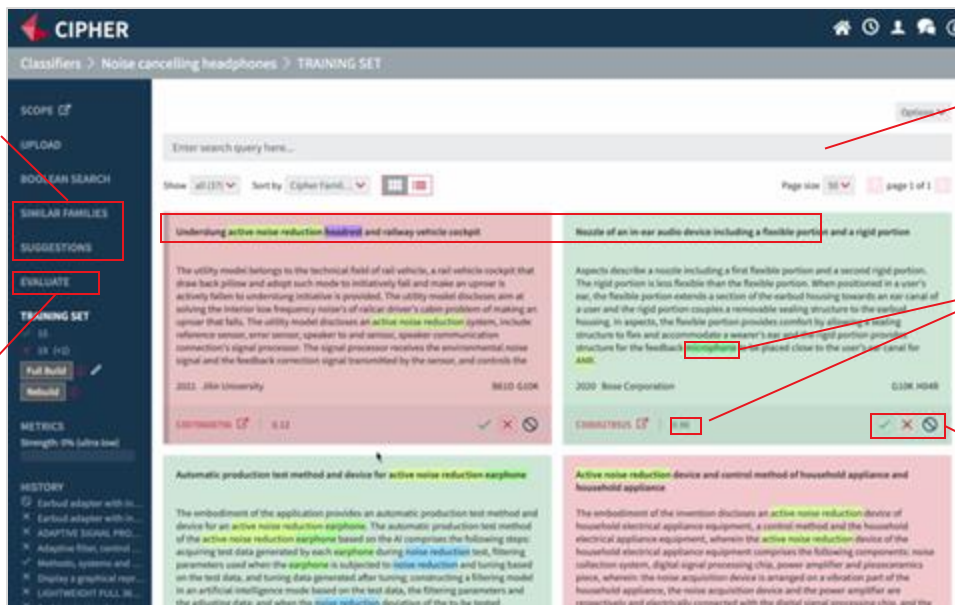
- 1 Training the Classifier with examples**
User provides positive and negative families to start training the classifier
- 2 Classifier provides feedback**
Each family receives a relevance score. Engine provides results and suggestions
- 3 Refining results**
User reviews and refines results and feeds those into the classifier training set
- 4 Iterations until good performance**
User reviews and refines results and feeds those into the classifier training set



Our Classifier Builder enables Cipher Certified

Smart Features: Suggestions & Similar Families analyse meta data to suggest additional patents that improve and diversify the training set

Tools to test performance and coverage before it is added to client's taxonomy



Text searching enables user to find other examples of patents to add to the training set

Phrase highlighting and classifier confidence scores assist user understanding of the ML algorithm

Example patents can be marked **positive** or **negative** to train the ML classifier

Classifier scope notes

Human curated scope notes support transparency and control - guiding the human **not** the machine

The screenshot displays the CIPHER interface for a classifier named "Voice Recognition - 3 Eval". The interface is divided into several sections:

- SCOPE**: Shows the classifier name and a search bar.
- SCOPE**: A detailed description of the classifier's scope, including a definition of voice recognition and a note that it excludes speech recognition.
- Build notes**: A note explaining that the classifier applies analyses of a person's voice to verify their identity.
- RECOGNITION**: Lists related terms such as speaker recognition, speech recognition, speaker verification, and speaker authentication.
- USE CASE/APPLICATION**: Lists applications like verification and identification, and notes that the classifier is text-dependent and independent.
- VOICE**: Lists voice-related terms like utterance, voice, and vocal print/profile/characteristic.
- TECH**: Lists technical details such as frequency estimation, hidden Markov models, Gaussian mixture models, pattern matching algorithms, neural networks, matrix representation, and vector quantization.

On the right side, there is a "Highlighting on" section with a list of phrases to highlight, including "voice, speaker, vocal, spoken, acoustic, vocal print, voice print, utterance, w", "G06Q 20/40145, G10L 2025/906", "G10L 17/00", "passphrase, acoustic features, person, user, unique, sound-groove", "authent, verif, access, analyse, ID, biomet, signature, correlat", "recogni", "G10L 17", "identity, identi", "GAUSSIAN MIXTURE MODEL, GMM", "sound wave, spectrum, voice characteristic", and "speech, speech to text, phrase key, phrase, word, Voice command, text-to-".

Classifier training data - getting started

Building classifiers requires high quality training data - Boolean search can deliver an effective starting point

CIPHER
Classifiers > Voice Recognition - 3 Eval > BOOLEAN SEARCH

SCOPE

UPLOAD

BOOLEAN SEARCH

SIMILAR FAMILIES

SUGGESTIONS

EVALUATE

SETS
Train 81 (+5 -2)
Test 95 (+1 -4)
Build

See guide [History](#) [Random USA grants](#)

{voice OR speaker OR vocal OR spoken OR acoustic} AND {authent!* OR identi!* OR verif!* OR access OR ID OR biomet*} fetch 500 by releva

Top 500 results (from 266,564 total)

Show all (500) Sort by Score (desce... Lines Single Columns default Page size 50 page 1 of

| Cipher Family ID | Owners | Title | Priorit... | Sc... |
|------------------|------------------------------------|------------------------------------------------------------------------------------------------------------|------------|-------|
| C0052185220 | Beihang University | A kind of method for recognizing sound-groove based on Android | 2017-09-11 | 0.99 |
| C0054308252 | | A kind of teaching Work attendance method based on voice recognition | 2018-06-13 | 0.99 |
| C0011441806 | Shengle Information Technolog... | Voice-print authentication system having voice-print password picture prompting function and realizatio... | 2010-09-25 | 0.99 |
| C0014330647 | Guangdong Univ of Foreign Sta... | A kind of speaker's method for quickly identifying and system based on model growth cluster | 2015-09-06 | 0.99 |
| C005018610 | Iflytek Co., Ltd. | Method and system for identifying speaker | 2011-08-18 | 0.99 |
| C0010243624 | Beijing Yuanjian Technologies C... | Dual-factor identity authentication method and system based on Application on Voiceprint Recognition a... | 2015-04-14 | 0.99 |

Developing the training data - Similar Families

Our advanced **Similar Families** algorithm supports exploration of both general and specific areas

The screenshot displays the LexisNexis 'Similar Families' interface. The breadcrumb trail at the top reads 'Classifiers > Voice Recognition - 3 Eval > SIMILAR FAMILIES'. A left-hand navigation menu includes options like SCOPE, UPLOAD, BOOLEAN SEARCH, SIMILAR FAMILIES, SUGGESTIONS, EVALUATE, and SETS. The main area features a search form with 'Family IDs' (containing 'C0011441607') and a 'Text' input field. Below the form are controls for 'Number of results' (set to 100) and 'Text type' (set to General), with a red 'Search' button. A 'See guide' link and an 'Options' dropdown are also present. A filter input field is labeled 'Enter keywords to filter results'. Below this, there are controls for 'Show' (all 100), 'Sort by' (Relevance), and 'Page size' (50). A summary line states 'This selection contains 100 families'. A table of results is shown with columns for Cipher Family ID, Owners, Title, Priority, and Score. Two results are visible, both from 'Shengle Information Technolog...'. The first result has a score of 0.99 and a date of 2010-09-25. The second result has a score of 0.98 and the same date.

| Cipher Family ID | Owners | Title | Priorit... | Sc... |
|------------------|----------------------------------|-------------------------------------------------------------------------------------------------------------|------------|-------|
| C0011441606 | Shengle Information Technolog... | Voice print authentication system having voice print password picture prompting function and realization... | 2010-09-25 | 0.99 |
| C0011435799 | Shengle Information Technolog... | Voiceprint authentication system and implementation method thereof | 2010-09-25 | 0.98 |

Developing the training data - Suggestions

The **Suggestions** functionality is Cipher prompting the classifier builder to provide the most useful information to the system

Classifiers > Voice Recognition - 3 Eval > SUGGESTIONS

SCOPE
UPLOAD
BOOLEAN SEARCH
SIMILAR FAMILIES
SUGGESTIONS
EVALUATE

SETS
Train Test
✓ 81 (+1 -3)
✗ 90 (+1 -4)
Build

METRICS
CV Test
prec 0.82 N/A
recall 0.82 N/A
f1 0.82 N/A
Strength: 82% (high)

Quantity: 10 Strategy: Automatic

Fetch more suggestions

A mix of all available suggestion strategies

9 suggestions (using strategy Automatic)

Show: all (9) Sort by: Score (desc...)

Page size: 50 page 1 of 1

Similar Citations

Dynamic cipher generation method for electronic equipment for **fingerprints** and **voice recognition**

The invention discloses a dynamic cipher generation method for electronic equipment for **fingerprints** and **voice recognition**. The dynamic cipher generation method comprises the following steps: acquiring a **ingerprint** image of a **user** and extracting **ingerprint** features; performing matching the **ingerprint** features and a **ingerprint** template; when the **ingerprint** features are matched successfully, generating and displaying **voice verification** content; acquiring the **voice** message of the **voice verification** content read by the **user** and extracting the **voice verification** content and **voiceprint** features from the **voice** message; when the extracted **voice verification** content is matched with the generated **voice verification** content,

2014 Wuxi Biokee **Biometrics** Technology Co Ltd G06F

C0002874033 0.95

Similar Phrases to Positives

Divide the method for account management **user** on TV using sound groove **recognition** technology in e

The invention discloses a kind of method for dividing account management **user** on TV using sound groove **recognition** technology in e, the present invention be **voiceprint identification** module, Android operation system support under realize, when smart television is started shooting, start **voiceprint identification** module, pass through **voice print** database collection and data analysis, match the account of **user**, further according to each item data in **user's** account to television image pattern, acoustic pattern, hobby program, using hobby and other use habits to TV carry out parameters setting, so as to meet the individual demand of different **user**.

2016 Sichuan Changhong Electrical Appliance Co.,... H04N

C0022112183 0.75

Evaluation - testing and refining the Classifier

Cipher has many tools to assist with the **Evaluation** of a Classifier

The screenshot shows the CIPHER interface for evaluating a classifier. The breadcrumb trail is 'Classifiers > Voice Recognition - 3 Eval > EVALUATE'. The left sidebar contains navigation options: SCOPE, UPLOAD, BOOLEAN SEARCH, SIMILAR FAMILIES, SUGGESTIONS, and EVALUATE. Under EVALUATE, there are 'SETS' for 'Train' (81 (+5 -2)) and 'Test' (80 (+1 -4)). The main area is titled 'Create evaluation report' and includes options for 'Random' (20,000, USA grants), 'Landscape' (All patents), and 'Custom'. Below this is a 'Previous' section with a list of reports, including one for 'Google, Amazon.com'.

| | Google | Amazon.com | TOTAL |
|----------------------------|--------|------------|-------|
| Voice Recognition - 3 Eval | 49 | 13 | 62 |
| TOTAL | 49 | 13 | 62 |

This identifies Google and Amazon patents responding to the draft classifier for evaluation

Training data - a more transparent approach

+ive and -ive **training sets** can be reviewed (and corrected) at any time

The screenshot displays a patent database interface with two panels. The top panel shows a list of patents related to voice authentication, with a sidebar on the left for navigation. The bottom panel shows a similar list of patents, but with a different set of suggestions and metrics.

Top Panel:

- SIMILAR FAMILIES**
- SUGGESTIONS**
- EVALUATE**
- SETS**
 - Train: 81 (+5 -2)
 - Test: 90 (+1 -4)
 - Build
- METRICS**
 - CV: 0.82
 - Test: 0.78

| Cipher Family ID | Owners | Title |
|------------------|----------------------------------|------------------------------------------------------------------------------------------------------------|
| C0011441807 | Shengle Information Technolog... | Voice-print authentication system having voice-print password voice prompting function and realization ... |
| C0059127580 | Cirrus Logic Inc. | SPEAKER IDENTIFICATION |
| C0043800073 | Alibaba Group Holding Limited | Method and system for distinguishing humans from machines |
| C0034191190 | NTT | Voice authentication system |
| C0024959506 | Tencent Holding Ltd | Sound-groove model training method, method for recognizing sound-groove and device |
| C0046437459 | | Identity authentication system and method |
| C0028063176 | Baidu Inc | Method and apparatus for authenticating user |
| C0008023484 | Baidu Inc | Vocal print login method and device based on artificial intelligence |

Bottom Panel:

- SUGGESTIONS**
- EVALUATE**
- SETS**
 - Train: 81 (+5 -2)
 - Test: 90 (+1 -4)
 - Build

| | | |
|-------------|-------------------------------------|------------------------------------------------------------------------------------------------------------|
| C0060145779 | Shanghai Transsion Information... | Handle method, terminal and the computer storage medium of incoming call requests. Method, terminal ... |
| C0046779436 | Intel Corporation | METHOD AND APPARATUS FOR ESTABLISHING DEVICE CONNECTIONS |
| C0008063862 | IBM | System and method for processing personal telephony recorder commands |
| C0004243492 | Luzhou Wzwood Planning Ct | Filtering network data management system based on <u>signature</u> verification |
| C0062191279 | Korea Advanced Institute for Sci... | SENTENCE SELECTION DEVICE FOR <u>SPEECH</u> SYNTHESIS TRAINING TO BUILD A <u>SPEECH</u> SYNTHESIZER BAS... |
| C0036917488 | Ricoh Co. Ltd. | Inkjet recording device and method of releasing cover lock of inkjet recording device |

Taxonomies and classifiers

Cipher enables set of **classifiers** to be delivered in structured **taxonomies**

CIPHER

SELECT CLASSIFIERS TO APPLY Done

Filter all available

[SHOW ONLY SELECTED] [SELECT] [DESELECT]

FinTech classifiers
[expand] [collapse]

- FinTech
 - Authentication and Authorization
 - Biometric
 - Biometrics (broad) ⓘ
 - Facial Recognition ⓘ
 - Fingerprint Scanning ⓘ
 - Iris Recognition ⓘ
 - Voice Recognition ⓘ
 - Non-Biometric
 - Open Banking
 - Trusted Execution Environment (TEE) ⓘ
 - Verification and Know your customer (KYC) ⓘ
 - Banking Services and Products
 - Computing Infrastructure
 - Customer Services
 - Payment Technologies

Uncategorised

- Voice Recognition - 3 Eval ⓘ

Voice recognition (also called voice authentication, voice identification, voice ID) is a type of user authentication that uses voiceprint or voice profile biometrics and relies on the unique vocal characteristics of individuals. Not to be confused with Speech Recognition (technology used in speech-to-text applications and virtual assistants).

Benefits of Classification

Classifiers are **supervised machine learning algorithms** trained with positive and negative patent examples that allow you to find all patents relating to specific technologies.



Trained by Human Powered by ML

Humans provide context and direction & ML connects data points.



Smart Features

Smart features help build better training sets and therefore create better results



Transparency

Full transparency in results allows the review of all patents by relevance.



Full Control

The Classifier Builder gives full control over what is in- and excluded.

Testing the accuracy of machine learning (ML) algorithms

Cipher has tested its ML algorithms using test data generated independently by a third party, achieving 96% accuracy

Test methodology

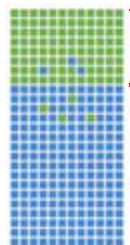
Cipher has tested its ML algorithm by:

- Taking a test data set generated independently by a third party (Patinformatics)
- Cipher's classifiers were trained on a portion of the data
- The accuracy was tested against the manually curated test set

Note: The test process is described in detail in Cipher's paper "Construction and evaluation of gold standards for patent classification", published in World Patent Information.

Test results

From 300 patents, Cipher's algorithm achieved 96% accuracy



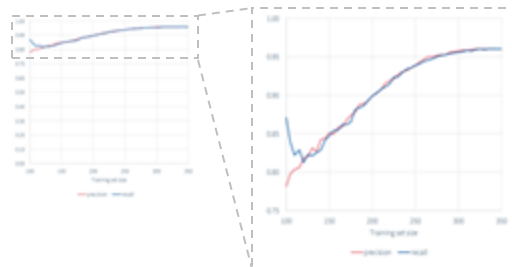
Predicted Positives

93 true positives
3 false positives

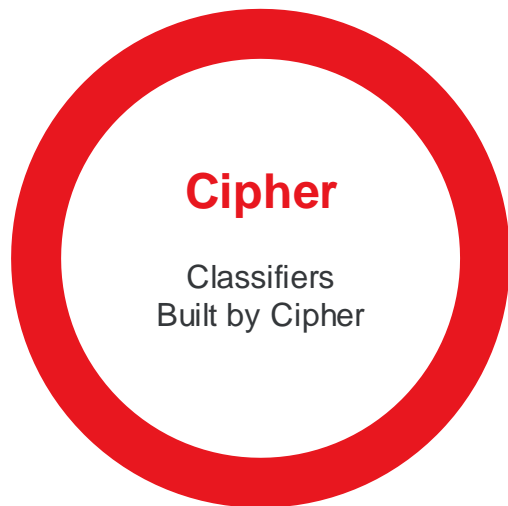
Predicted Negatives

200 true negatives
4 false negatives

Test data shows that Cipher's ML reaches a stabilised level of accuracy from a relatively small amount of data



Cipher Classifiers - two options



Client-defined Classifiers
built by the Cipher Team



Training program and support
to create inhouse Capabilities
for Classifier building

How are Cipher Customers using Cipher Classifiers?

The impact on our clients:

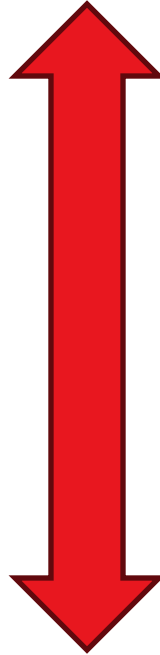
Top down success as a leader and bottom up success as a team

IP at the heart of the business

Top Down:

Cipher enables our clients to enhance the productivity and efficiency of their teams.

Classification automates recurring tasks such as competitor monitoring, so their time can be used on high value tasks



Bottom Up:

Portfolio analytics and robust data give our clients a seat at the table and an ability to justify their decisions, budget requests and a means to change the narrative around the portfolio being a cost centre

Automation of processes

“From cost cutting request to budget increase”

Challenge:

Our client was asked to cut a portfolio by 10% to save on costs. There was little understanding as to the business value the portfolio provided.

Insight:

Competitor benchmarking and risk models in Cipher showed that making these cuts would leave the business open to multi-million dollar assertion risks.

Results:

The GC changed tac and budgets were in fact increased as the process had identified pre-existing gaps in the portfolio.

“Four days per month saved by automation”

Challenge:

Our Client had a team member who cumulatively spent at least one day week running competitor searches, reading through new publications and sorting them.

Solution:

Cipher’s classifiers were implemented and the team member instead received an instant, automated, weekly digest of all new publications within their own business area.

The team were able to review, tag and comment on the patents within Cipher before then sharing notable patents to R&D colleagues, all seamlessly within Cipher.

Results:

That team member was then able to use their time to deliver more analytics to their business, including competitor filing strategy analytics and M&A scouting.

What is the Journey?

Cipher + PatentSight integration plan



Now

Leverage PatentSight's powerful analytics using Cipher ML classifiers

Cipher Certified available using Cipher's Classifier Builder



Mid-term

Additional features and integration – based on customer feedback



**Long-term:
Full integration**

Your own custom technology using Cipher industry leading classification engine directly into PatentSight



Questions?



Thank You

Will Mansfield

Head of Consulting and Customer Success
william.mansfield@lexisnexis.com