

# 知財情報検索と解析の将来展望

## — 拡張知財情報科学者 (Enhanced & Augmented IP Information Scientist) に 3 年以内になるためにはどうするか? —

Future perspective for Intellectual Property Information Retrieval and Analysis: What should you do to become an Enhanced & Augmented IP Information Scientist within 3 years ?



一般財団法人情報科学技術協会 パテントドキュメンテーション部会主査

### 桐山 勉

2004 年から INFOSTA-SIG-パテントドキュメンテーション部会のコアパースン。2006 年 2 月に帝人知的財産センターを定年退職し、2006 年 4 月から Japio 特許情報研究所の客員研究員を務める。2011 年～2016 年に IPI-Award Selection Board (国際特許情報賞選考委員) を務めた。2014 年より World Patent Information 専門誌の EAB 編集員も務める。2013 年 4 月より、はやぶさ国際特許事務所の顧問も務める。知財情報解析の専門家でもある。

## 1 はじめに

この三年間で TV<sup>1)</sup>、新聞、雑誌、広告などの巷の一般的な会話の中でも頻りに AI (人工知能) という言葉が出現するようになって来た。筆者の定点観測地点の対象として IPI-ConfEx<sup>2)</sup>、East meets West<sup>3)</sup>、PIUG 年次大会<sup>4)</sup>、II-SVD<sup>5)</sup> (または ICIC 会議)、PIAC in China<sup>6)</sup>、特許情報フェア&カンファレンス<sup>7)</sup>、EPO-PIC<sup>8)</sup> などの国際会議に注目し、この三年間にそこで発表されたプレゼン資料を集めて読み漁った。それに専門雑誌の World Patent Information 誌<sup>9)</sup> と日々の日経新聞情報とアジア特許情報研究会の仲間との情報交換などが筆者の主な情報源である。

昨年の特許情報フェア&カンファレンス以降<sup>10)</sup> の 10 か月の間に筆者に強く影響したのは次の 7 項目である。

- A) **新人 AI 研究者の育成** : 8 社からの 9 億円支援で東大に AI の寄付講座<sup>11)</sup>。2021 年 5 月までの 5 年計画。3 省連携による次世代の人工知能技術の研究開発と人材確保。高校生に対する AI 技術の数学的な考え方の教育も既に行われ始めた<sup>12)</sup>。米国にはシンギュラリティ大学<sup>13)</sup> という名の大学も既に設立されて加速度的な変化に備えるための教育がなされていると日経新聞にも掲載されるこの頃である。
- B) 大手 IT 企業の**機械学習・深層学習のフレームワークの公開**、および米 Google 社が無料で AI 研究者にスパコン開放<sup>14)</sup>。
- C) 欧州特許庁の TAPAS 構想 (**T**oward **A**utomated **P**rior **A**rt **S**earch) の研究<sup>15)</sup>。2013 年に averbis

社と共同研究を開始。

- D) 米国 USPTO における Sigma-PLUS<sup>16)</sup> という Prior Search の研究。IP.com と IEEE との支援を受けての試行研究。
- E) Oxford 大学のオズボン准教授と野村総研の共同研究<sup>17)</sup> をうけて、その後に **Advanced Attorney** の出現と活躍。
- F) インド系企業の Srinivasan Parthiban 氏の II-SDV<sup>5)</sup> および PIUG2017 年次大会<sup>4)</sup> における発表プレゼン。
- G) **コンピューター言語 Python** は、ユビキタスに誰でも例えば有料 Web セミナーを通じて受講できる時代である。

情報科学技術協会の月刊誌 7 月号に筆者らの総論記事<sup>18)</sup> が掲載される機会を得た。その執筆準備作業が AI に関する諸々の情報の蒐集と勉強に極めて役立った。

## 2 拡張 IP 情報科学者になるためには

筆者が主唱する造語、**拡張 IP 情報科学者**とは **Enhanced & Augmented IP Information Scientist** のことである。**IP 情報科学者**という概念は、2008 年に (当時の特許検索競技大会委員長の) 中出良治氏と欧州 PDG を訪問した際に当時の PDG 会長の Minoo Phillip 女史から教わった考え方である。単なる特許情報の依頼代行調査をするのではなく知財情報と非特許情報も含めて必要情報をすべて集め、理解し、更に科学者らしく情報解析し、依頼元にその結果を報告できる高度な情報科学者

という考え方である。

**Enhanced Human (拡張人間)** という語彙は AI 記事の中で頻繁に使われている。平たく言えば、AI 技術を身に付けて AI 技術を活用できる能力を持った人間(AI 技術で補強された能力を持った人間)を意味する言葉である。

Augmented は **Augmented Reality (拡張現実)** というバーチャルリアリティの関係で良く使われ始めた語彙である。また高度医療のお蔭で、義足と義腕で通常の人と全く変わらぬ活動をしている人達も多くなってきた。筆者も心臓病を患いリハビリ治療中であるが、そのリハビリ仲間には人工心臓の人も内蔵心臓ペースメーカーを埋め込んだ人も多種多様であり、高度医療の恩恵を受けた **Augmented Human (拡張人間)** とお付き合いをするように変わった。だから、筆者の経験から思いついた新造語が「**Enhanced & Augmented IP Information Scientist (拡張知財情報科学者)**」である。

福澤諭吉翁のエピソードを知って、自己啓発のための中期 3 年～5 年の計画を思いついた。諭吉は 24 歳の時(1859 年)に横浜の外国人居留区を見物に行った。蘭方医・緒方洪庵の私塾「適塾」でオランダ語を学んだ福澤は横浜の外国人居留区では蘭語が通じず、英語が話されていたことに挫折した。福澤は大阪の適塾に戻った翌日から英語の独学に切り替えたエピソードである。また、「慶応義塾信義」には三カ月(初歩書)、六カ月(地理書または窮理書)、六カ月(歴史書)の合計 1 年 3 か月の修行が先ず必要と書かれていると教わった。筆者も、上記のエピソードを知った後で、友人の安藤氏に強く勧められて、人工知能のコンピューター言語の Python を学ぶ決心をした。

筆者は持論として、三つの目(①鳥の目、②魚の目、③虫の目)の他に、**第 4 の目**、所謂、**予測の目・シミュレーションの目**が必要と考えている。それをいつも One-Sheet 形式でイメージアップすることが実践行動には必要と考え、先ず、その One-Sheet を作成した(図 1 参照)。三段階で把握するために、現状と 3～5 年後の近未来予測と約 10 年後の遠未来イメージ推測を強制発想にて描いた。すると、副題の「拡張知財情報科学者(Enhanced & Augmented IP Information Scientist)に 3 年以内になるためにはどうするか?」が自分自身の脳裏に臍げに浮かんできた。コンピューター言語の Python を学ぶ一歩を踏み出してから更に徐々にではあるがこの One-Sheet が具体的に映り始めて来た。

### 3 必要と推測する「5 つの Level」(私案)

AI 技術のことを調べていると自動運転技術の技術開発の目標の 5 つの level に遭遇する。それにちなんで、筆者の独断にて拡張 IP 情報科学者になるための途中経過を考慮しつつ、5 つの level の私案を思いついた。人間には願望が見つかったとしてもそれを実現するまでには相当の修行時期が必要となり、日々の訓練が必須となる。頭だけでなく身体で覚えて慣れること(習慣化)が特に大事である。そのためには努力と覚悟が必須である。

筆者は、3 年以内に少なくとも表 1 の Level-3 に到達したいと願望している。好ましくは、その Level-4 に到達したいと願望している。そして AI 技術を活用した特許情報総合検索システムの開発に貢献できる **拡張 IP 情報科学者** になりたいと夢と強い願望を持っている。

現在(傾向・兆候) ◀	3～5年後予測(2020～2022年) ▶	10年後イメージ推測(2027年後) ▶
<ul style="list-style-type: none"> <li>・囲碁(AlphaGo)、将棋(PONANZA)◀</li> <li>・電動車椅子(cf. WHILL)◀</li> <li>・自動運転車(Level 3)◀</li> <li>・高度医療(義足・義腕、心臓治療)◀</li> <li>・Advanced Searcher◀</li> <li>・Advanced Attorney◀</li> <li>・TAPAS(EPO)◀</li> <li>・Sigma-PLUS/APES(USPTO)◀</li> <li>・オープン・リソース(主要 IT 企業 etc)◀</li> <li>・AI 研究者養成(9 億円寄付講座)/サーバー解放(Google)◀</li> </ul>	<ul style="list-style-type: none"> <li>・開発終了→他分野へ AI 技術応用研究◀</li> <li>・電動車椅子 with スマホ(音声ガイド、GPS、他)◀</li> <li>・自動運転車(Level 4 &amp; 5)◀</li> <li>・高度医療 Augmented Human (I)◀</li> <li>・<b>Enhanced &amp; Augmented IP Information Scientist with AI</b>◀</li> <li>・進化した TAPAS(EPO)、審査官連携◀</li> <li>・進化した Sigma-PLUS/APES(USPTO)、審査官連携◀</li> <li>・オープン AI エコシステム(<b>オープンソフト活用 AI サービス</b>)◀</li> <li>・Web セミナー日常化/サーバー解放の常識化(通信 G5)◀</li> </ul>	<ul style="list-style-type: none"> <li>・他分野へ AI 技術応用普及・常識化◀</li> <li>・小型無人タクシー型自動車、リニア新幹線◀</li> <li>・自動運転車(Level 5)、ドローンの応用化普及◀</li> <li>・高度医療 Augmented Human(II)◀</li> <li>・Enhanced &amp; Augmented IP Information Scientist through AI◀</li> <li>・主要特許庁/WIPO-IEEE 連携の消滅化◀</li> <li>・審査業務の大幅改善&amp;変質 through AI◀</li> <li>・<b>PreSingularity 真近かに迫る</b>◀</li> <li>・新次世代通信&amp;Exa Scale Server 普及?◀</li> </ul>

図 1 筆者の 3 ステップ予測のイメージ図

表 1 筆者が推測する5つのレベル (私案)

level	各レベル内容の状態	モデル人物
Level-0	AI技術に無関心 AI技術には全くの素人	2015年までの筆者
Level-1	AI技術に興味・関心を持つ 文献・特許の蒐集・収集	2016年4月時点の筆者 2017年1月時点の筆者
Level-2	AI技術の専門用語を理解する 深層学習の仕組み・言語を理解する AI技術を使ったAPIを活用している	2017年12月時点の筆者目標 2018年8月時点の筆者目標 2016年の昭和電工のT氏
Level-3	AI技術の深層学習のプログラムを書く。 自分のPCにて 特許情報の教師ありを解析	2016年時点の友人A氏 <b>2018年8月時点の筆者目標</b>
Level-4	AI技術に関して4人で議論ができるサーチャー ①コーディネータ、②Deep Learning技術者、 ③Platform-Edgeモデル設計SE、④拡張IP科学者(橋渡し役)	2020年8月時点の筆者目標
Level-5	大人の人工知能の改善ができる。 大手IT企業が提供するAIの無料システムが触れる深層学習技術を持っている。	2020年以降の筆者目標 2015年時点のIBMの知人 (Stephen Boyer氏)

Abraham Lincoln曰く、「誰かが素晴らしい成功を収めたということは、他の人にも同じことができるという証明である」と。この格言を信じて実践行動を開始したい。

## 4 コンピューター言語 Python を学び始める

大手IT企業はフレームワークを公開している。それらはどのようなコンピューター言語を使っているのだろうか。質問を友人に送り、教えて頂いたのが図2である。

また、図2の中に Srinnivasan Parthiban 氏が「Pythonが多用されている」ことを何度も説明したプレゼン・スライドを埋め込んでみた。

No.	IT企業名	公開フレームワーク	CPU言語	OS言語	備考1
1	Google	TensorFlow	Python, C++	Ubuntu, Linux, Mac OS X	日経記事 2017.5.19 6 AI研究者にスノコン解放(無料)
2	Facebook	Torch	Lua	Ubuntu, Mac OS	
3	Microsoft	CNTK	Python, C++	Linux, Windows	
4	Amazon	DSSTNE			
5	PEN 日本	Chainet	Python	Ubuntu, CentOS, Linux, Mac OS	
6	FRONTEO	未公開			
7	NTTデータ 数理システム	未公開	Python		
8	日本IBM	Watson		Linux	
9	XLPAT(パナ)	未公開			
10	IRD 国産特許 事務所	未公開			
11	トヨタ(ト)	未公開			
12	ソニー	公開			日経記事 2017.6.27

図2 Pythonは広く活用されているという結果

筆者は、アジア特許情報研究会 知財情報解析研究グループにも所属している。そのリーダーの安藤氏に強く勧められたことがある。

「今は誰でも機械学習の深層学習を自分のPCで学べて触れる時代ですよ!」「コンピューター雑誌「Interface」(2017年6月号)に掲載されたJupyter-Notebookを利用して、有料WebセミナーUdemyの中で、公開されたPython(Version3.5)をインストールしながら公開オープン・リソース・ライブラリーをインポートして自分のPC上で演習する119

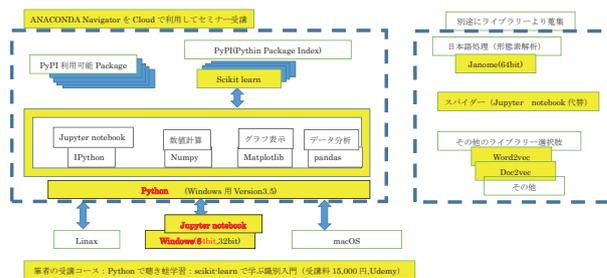


図3 筆者が学んだPythonと各種ライブラリー (Interface誌6月号のP68の図1を参照して作成)

コースから成るセミナーがある」と。安藤氏を信じているが「本当に素人の筆者でも着いて行けるのか」という不安で一杯であった。

しかも15000円の有料である。(1週間以内に申込みれば92% offの1200円で申し込めるとのこと。)清水の舞台から飛び降りるがごとく大きな覚悟が筆者には必要であった。一旦、登録購入したものの119コースを修了するまで、解からないことだらけで最初の十数時間が苦痛の何物でもなかった。しかし、全119コースを30時間以上の忍耐後に終了し、現在随意に受講できる第2回目の自己学習を50時間以上かけて行っている最中である。

この有料Webセミナーを受講して解ったことがある。

- A) 確かにコピキタスに誰でも自分のPCにて機械学習の深層学習をPython言語で触れる時代であることが素人の筆者にも確認できた。
- B) 機械学習→深層学習のPython言語に対する拒否アレルギーが無くなった。
- C) **数学的科学手法のVector法**の次元数に限界が無く、また、**次元削減法**とか**二値化識別法**とか**深層学習の最適法**などの客観的手法が既に公開ライブラリーの中に沢山存在することが判明した。その結果、同義語・類推語などを過去の経験から見つけるベテラン・サーチャーの域にはAIは到底達成できないだろうという思いは「**筆者の偏見に過ぎない**」と実感できた。所謂、数学と知財情報サーチャーの仕事が繋がった。
- D) AIの強みは、構築されたアルゴリズムとシステムにより自己学習が人間の1000倍以上の学習経験時間を短時間にこなしてしまう**超スピード機能**であると実感できた。まさに、囲碁と将棋のAI電王戦にて、「AIが人間の経験時間の1000年以上の自己対戦研究ができる」という特徴とよく似ている。

## 5 【知財分析+IP ランドスケープ】の時代

2017年7月17日の日経新聞に「知財分析 経営の中枢に」「IP ランドスケープに注目が集まる」という記事が掲載され、知財業界で話題になっている。筆者が尊敬する三井物産戦略研究所の山内明先生がその約1か月前に日経新聞社からインタビューを受けられて、その結果の内容が詳細に日経新聞記事になったものである。今年の Japio YEAR BOOK 2017 には著名な山内明先生の論文「特許マーケティングを中心にした IP ランドスケープ実践に役立つ知財情報戦略」が収録されると聞いているので、是非とも精読して参照して頂きたい。

また、今年の特許情報フェア&カンファレンスに来日し講演される予定の Anthony Trippe 氏は Patinformatics.LLC の創業者であり、現在ではイリノイ工科大学 (IIT) の IP Management and Markets の准教授でもある。彼は 2010-2012 年には米国 PIUG の会長を務めた重鎮である。彼は、早くから IP ランドスケープに注目し、WIPO の委託を受けて Guideline for IP Landscape を作成した人物でもある。彼も今年の Japio YEAR BOOK に「二値化識別法の有用性」について論述すると聞いている。是非英語文を精読して貰いたい。

筆者が尊敬するお二人とも知財情報を解析することの重要性を早くから指摘され、IP ランドスケープを実践してきた国内と海外の IP ランドスケープ活用の牽引者（この業界にて率先垂範する機関車的な存在の尊敬すべき重鎮）である。

この様に注目される IP ランドスケープを実践するために、筆者も努力を重ねている。筆者が常用する特許検索システムは Shareresearch と JapioGPG/FX である。それらの検索結果をダウンロードしてレイテック社の PAT-LIST-GLS にて通常の特許分析を日常茶飯事的に行っている。そして、それらの俯瞰可視化と IP ランドスケープ化を色々な手法で行っている。最近では Grid 図（モザイク図）と引用被引用ネットワーク図と Word Cloud 図を併用することが多い。

更に筆者は新しいシステムを試行するために、できる限り交渉して試行実践ができないか世界の知人人脈を通じて努力を惜しまない。筆者は既に帝人を退職し 10 年以上も経つので自分のやりたいこと・やりたい研究を継

続するには知人の皆様のご支援をお願いする以外には手段がない。

## 6 事例研究( I )—【Semantic Search +AI】

この三年間で色々な検索システムが裏で AI が稼働して改良されている。定点観測地点として海外の国際会議に毎年継続して参加しているとグローバル特許情報業界における大きな潮流が見えて来る。その一つが、【Semantic Search +AI】および【Citation Search +AI】という考え方である。その潮流を確認するための事例研究として、特許検索競技大会 2016 の化学・医薬分野の問 2 の問題の解答を色々なシステムにより試行してみるのが良いと考えている。その試行の結果から、何か新しい知見が得られないかを個人的に研究している。最終的な報告は INFOPRO シンポジウム 2017 にて報告するつもりであるが、その中間結果をこの場にて一部だけ記述する。

筆者らは、「情報の科学と技術」の 7 月号に執筆した総論<sup>18)</sup>において、「特許検索競技大会において AI も参加が許容され、AI が 5 年以内に Gold 賞を受賞する日が訪れるのでないか」と予測した。だからこそ、その特許検索競技大会の事例にて研究を具体的に行った。

### 【Shareresearch<sup>19)</sup>による概念検索テスト】

筆者の常用検索システムの予備検索として概念検索も多用していた。しかし、「用いる単語の重み付けを詳細に変えた概念検索」をする習慣が抜けていた。筆者の死角であった。この死角を徹底的に研究し、Shareresearch の効果を最初に見つけたのは友人の安藤氏であった。その驚くべき知見を追試してみたら、筆者も図 5 の結果を得た。

Shareresearch の第一段階の概念検索（通常多用しているもの）に次の文を入力した。

【入力文】熱可塑性樹脂フィルム基材層 (A 層)、酸化ケイ素蒸着層 (B 層)、ポリビニルアルコール系樹脂と粘土鉱物を含む塗膜層 (C 層) が他の層を介して又は介さずにこの順に積層されてなることを特徴とするガスバリア性包装用フィルム。

ここで、驚くべきことは、入力文において「粘土鉱物」



図4 Shareresearchの単語重み付の指定画面



図5 Shareresearchの単語重み付結果の上位画面

の類似語である「モンモリロナイト」は単語重み付け指定の画面（図4）には全く表面にでていない。しかし、その結果の図5では**第一位と第四位と第七位に正解が表示された**。（友人の安藤氏の知見と正確には全く同じではないが、上位に該当3件が含まれていることは、検索時期が多少1か月遅れても、筆者のPCでも再現できた。）

特許検索競技大会の問2は、「正解特許の内の異なる5社の特許を選んで答えよ」という指定なので、その5件の6割の3件を上位10件以内に表示させる「**概念検索+ AI**」の威力に筆者も驚かされた。

（当然のことながら、ベテランの特許検索サーチャーが作る模範検索式をSharesearchで追試して作成すれば、28件（筆者メモ控えS2913）がヒットし、正解は5件が含まれていたことも追試検証できた。）

#### 【XLPAT<sup>20</sup>】による Search 検証

前述の表1に記した**Level-2**に到達するためには、AIを組み込まれた特許検索システムを使う経験がどうしても必要であった。そのために、特許情報フェア&カンファレンスに出展していたXLPAT-lab社のブースを訪れ、詳細に説明を聞き、別会場の企業プレゼンにも参加した。その後、直に英語メールを送り、Trialを強く希望する旨を何度もお願いしたが、「日本の製造業企

業のTrial申し込みを優先するから」と筆者は拒絶された。しかし、今年の7月20日にNGB社で開催されたセミナーには受講者特典として10日間のTrialが付与されていた。岐阜県大垣市から25,000円の交通費を自己負担して参加した。更に、「セミナー受講者でアンケートに答えた場合には、更に5日間のTrialが追加される」と説明があったので、それを申し込んだ。昨年の特許情報フェアから10カ月も待ち続けたTrailが幸運にもできた。早速、田辺氏の論文「ビッグデータ時代における特許情報調査への人工知能の活用」の3.2節に記載のXLPATの記述内容を検証した。その論文の図3、図4、図5、及び図6に記載の内容も全て検証確認ができた。

次に、特許検索競技大会2016の化学・医薬分野の問2をXLPATを用いて解くことに挑戦した。

データベースはLexisNexis社のものを共用しているので問2の問題を英文化することから始めた。

【検索概念の英語文（筆者の英訳）】 A **thermoplastic resin film** base layer (layer A), a **silicon oxide vapor-deposited** layer (layer B), a coating film layer (C layer) containing a **polyvinyl alcohol resin** and a **clay** mineral in this order wherein the **gas barrier** laminate film is laminated.

残念ながら、XLPATにはSemantic Search 英文を入力して検索する機能が見つけれなかった。そこで、複雑な近接演算式を考えて、Search機能において日本特許、公関係、フルテキスト検索対象で検索を行った。

【実際に入力した複雑な近接演算式】  
((“Silicon oxide” OR SiO) W/n30 (deposition OR CVD)) AND ((“polyvinyl alcohol” OR PVA) W/n20 (clay OR montmorillonite)) AND (gas W/n2 barrier) AND packaging

非常に速い応答スピードで66件がヒットした。（図6参照）

更に、驚くことは絞り込み検索機能（+add field）があり、Title+Abstract+Claimにて絞り込むと6件がヒットした。（図7）

<input type="checkbox"/>	S.No	Number/Title/Abstract
<input checked="" type="checkbox"/>	1	JP2001287319A - Gas Barrier Film <a href="#">Show/Hide Abstract</a>
<input checked="" type="checkbox"/>	2	JP2011218805A - Gas Barrier Laminated Film <a href="#">Show/Hide Abstract</a>
<input checked="" type="checkbox"/>	3	JP11165369A - Gas Barrier Film <a href="#">Show/Hide Abstract</a>
<input type="checkbox"/>	4	JP2007098646A - Gas Barrier Film <a href="#">Show/Hide Abstract</a>
<input checked="" type="checkbox"/>	5	JP2013006283A - Laminated Film <a href="#">Show/Hide Abstract</a>
<input type="checkbox"/>	6	JP2011218803A - Gas Barrier Laminated Film <a href="#">Show/Hide Abstract</a>
<input checked="" type="checkbox"/>	7	JP2009006508A - Laminated Film <a href="#">Show/Hide Abstract</a>
<input type="checkbox"/>	8	JP2011201280A - Laminate Gas Barrier Laminated Film And Package

図6 XLPATにおける66件ヒットの上位画面 (No.1、No.2、No.3、No.5、No.7が正解であった。)

<input type="checkbox"/>	S.No	Number/Title/Abstract
<input checked="" type="checkbox"/>	1	JP11310680A - Gas Barrier Film And Package Using The Same <a href="#">Show/Hide Abstract</a>
<input type="checkbox"/>	2	JP2001283814A - Gas Barrier Laminate Excellent In Stiffness Strength And Impact Resistance, Having Shading Properties <a href="#">Show/Hide Abstract</a>
<input checked="" type="checkbox"/>	3	JP2000062108A - Gas Barrier Film <a href="#">Show/Hide Abstract</a>
<input checked="" type="checkbox"/>	4	JP2009006508A - Laminated Film <a href="#">Show/Hide Abstract</a>
<input type="checkbox"/>	5	JP2000202937A - Vapor Deposition Film Laminated <b>PACKAGING</b> Material <a href="#">Show/Hide Abstract</a>
<input checked="" type="checkbox"/>	6	JP2003025483A - High Gas Barrier Laminate <a href="#">Show/Hide Abstract</a>

図7 XLPATにおける+add field 絞り込み検索結果 (No.1、No.3、No.4、No.6が正解であった。)

【絞り込み検索式】

((“Silicon oxide” OR SiO) W/n30 (deposition OR CVD)) AND (gas W/n2 barrier) in Title, Abstract, and Claim.

ただし、この結果で注意したいことは、もしも「clay mineral」の類似語である「montmorillonite」を複雑な近接演算式に組み込まなかった場合の検索の実証確認結果である。

残念ながら「montmorillonite」を組み込まなかった場合にはヒット22件が得られたが、正解は1件も含まれなかった。

だから、問2の問題の難しさは、「粘土鉱物」の類似語をウィキペディアで調べれば気づくかも知れないが、

専門家でないとは気づかないことだと筆者は判断した。

(註：しかし、Shareresearchの第2段階の単語重み付け概念検索の結果では、裏側で動くAIがそれを見つけて来ている観点に凄さがあると筆者は判断している。)

【InnovationQ Plus<sup>21)</sup>による概念検索テスト】

筆者がイタリアのローマ市で3月上旬に開催されたIPI-ConfEx2017に参加した時、IP.com社とIEEE協会の共同展示ブースが設けられていた。そこで、「米国USPTOのSTICにてSigma-PLUSを試行研究しているが、それは貴社らのものか？」という内容を英語で質問したら「Yes, その通り」という回答が返ってきた。そこで、特許情報フェア&カンファレンス2016にてUSPTOのThomas A Beach氏の英語スライドの4枚を見せながら、「このスライドの図と同じ画面をここで確認したいので、見せて欲しい」と約30分間も粘り交渉した結果、展示ブースにてデモ説明するお二人では説明できないと言い出した。そこで、「日本にて日本語で私にデモ説明できる人を紹介して欲しい。」と要望しながら名刺を交換した。すると5月になってIP.comのJim Durkin氏からメールが入り「電話で話したい」と言ってきた。その後7月になり、丸善雄松堂(株)から筆者の携帯スマホに突然に電話が入り「IEEE協会からの依頼事項でInnovationQ Plusを説明させて欲しい」と連絡が入った。幸運の女神の降臨であった。

8月10日にこちらの要望の「特許検索競技大会2016の化学・医薬分野の問2を使いデモをして欲しい」とお願いした。しかし、近接演算のコマンドが正確に詳細に判らず、困った。Trialの許可を漸く頂き、8月23日にSign Inに成功した。

早速、InnovationQ Plusにも、英語の質問文を投げた。

InnovationQ Plusの凄いところは、質問文にmontmorilloniteが無くても、Shareresearchの結果と同様にTop 8位とTop 17位に正解が出現する点である。その他の試行結果の詳細は、INFOPROシンポジウムにて報告したい。

Semantic Search式をclay mineral such as montmorilloniteに修正すると、上位30件の中に正解が

## 7 事例研究(Ⅱ)―【Citation Search +AI】

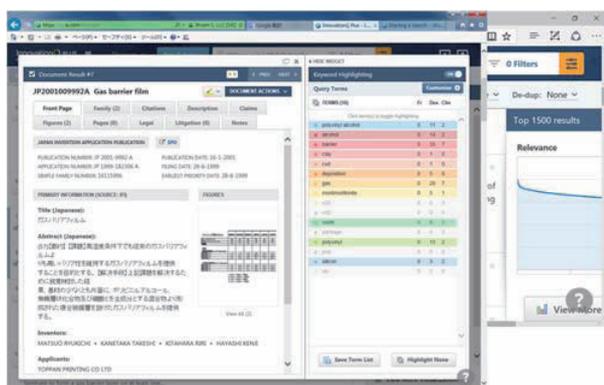


図8 InnovationQ Plus の Top8 位の結果図

7件も増加する結果になった。短期間の試行期間であるので筆者が詳細に機能を把握できていない可能性がある。InnovationQ Plusは、「Semantic Search + AI」を中心的な想定活用法として位置付けて、高度な Boolean 近接演算式の作成を想定していない。なるほど、米国 USPTO が注目して試行検討をするはずと直感した。

以上の三種類の Trial により表1の Level-2 に漸く辿り付けた。

### 【同義語・類似語を見つける工夫】

欧州特許庁は word2vec を TAPAS に応用していると聞く<sup>15)</sup>。米国 USPTO の Sigma-PLUS では OpenNLP と WordNet を活用しているらしい<sup>16)</sup>。友人の安藤氏は Word2vec と doc2vec を使い分けて、自分の PC にて専門家として研究している。米国 PIUG2016 と 2017 年次大会<sup>4)</sup>では、elementary IP 社の Sumeet Sandhu 氏 (女性 CEO) が、二段階 (Macro Analytics, Micro Analytics) にて Search の keywords (同義語・類似語) を見つける Tool を開発して upgrade 中と発表している。(筆者はスライドの第6枚目に注目している。)

以上、色々なプレゼン資料をダウンロードして精読し、解からない部分を周囲の知人・友人に相談し、できるだけ正確に理解するように努力している。いつも知人・友人に感謝のし通しである。

Ambercrite.com の副社長とは米国 PIUG2013 年次大会にて質問してからは知人となった。Doris Spielthener さん (女性社長) とは 2015 年に Web-Ex 会議を数回も行った間柄のために、最近市販され始めた Ambercrite Ai<sup>22)</sup> も幸運にも特別に試行することができた。二人が出願している特許を Shareresearch にて検索すると、多数がヒットする。その代表的な特許群は次のものである。

AU2013332274B, AU2015246129A, WO2015059491A, EP2909756A, 特許 06063053, US2015/0220535

それらの特許番号を 1 件ずつ Ambercrite Ai に入力して前 3 世代 + 本件 + 後 3 世代 = 合計 7 世代の引用被引用のネットワーク関係を調べている。

Ambercrite Ai には残念ながら【Semantic Search + AI】の文章表現の入力画面が無いので、あくまでも特許番号の入力が中心になる。そこで、特許検索競技大会 2016 の化学・医薬分野の問 2 の正解特許番号群を色々入力して調べた。

事例研究 (I) のような結果は得られないが、図 9 のような引用被引用の 7 世代間のネットワーク図が得られ、Boolean 検索とも、「Semantic Search + AI」検索とも異なった観点からの検索手法として、筆者は大変に重宝している。

Ambercrite Ai の結果のネットワーク表示は、Amberscope と類似の表現方法である。Ambercrite Ai の Similarity 一覧表も Amberscope と同様な表現

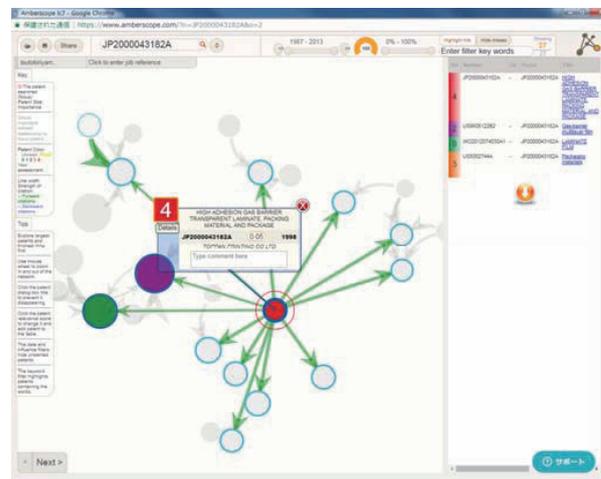


図9 特許検索競技大会 2016 の化学・医薬分野の正解特許 (特開 2000-043182) のネットワーク図

ではあるが、upgrade しているために個々の詳細な項目が改善されている。筆者がチェックした特許番号が容易に EXCEL にダウンロードできるので、有用な手段として使っている。

筆者は、IP.com-IEEE 協会が共同開発した InnovationQ Plus に興味をもっている。Jim Durkin 氏が示した特許群 (US6038561, US8539000, US8548951) を精読するだけでなく、Ambercite Ai にて 7 世代間に亘る関連特許も調べた。図 10 に US8548,951 の引用被引用のネットワーク図 (NW 図) を示す。また図 11 にそのネットワーク図の Similarity 結果のデータ画面も示す。この中には、Aureka の Themescape の開発企業の Augirin 特許も含まれ、広範囲に関係特許群を含めて把握することが容易にできた。

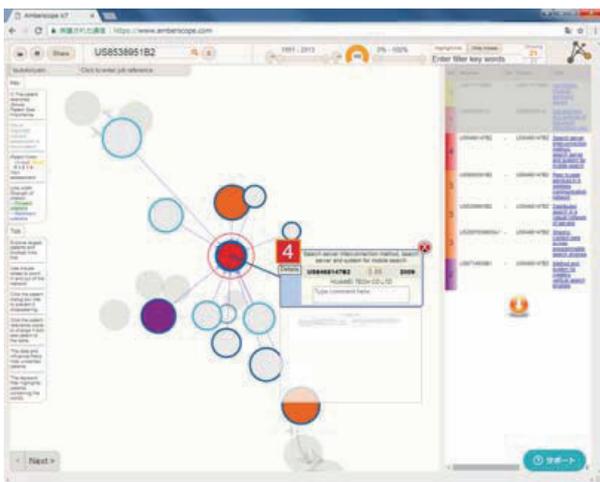


図 10 IP.com 社の Jim Durkin 氏が指摘する特許 (US6038561B2-US8468147B2) の NW 図

Patent No.	Title	IPC Class	Priority	Appropriation
US61201482	Label search	G06F17/30884 G	2003-12-29	Yes 0.83
US67149381	Method and app	G06F17/30884 Y1	2003-07-31	Yes 2.42
US79424782	Blending metho	G06F17/30553 G	2007-01-24	Yes 1.56
US79322882	Patented ass	G06F17/30884 G	2007-02-16	Yes 1.57
US70962080	System and me	G06F17/30884	2005-09-20	Yes 0.46
US82759982	Search method	G06F17/30881 G	2004-09-28	Yes 0.46
US10688882	Peer-to-peer s	G06D3/00 H04L	2004-07-06	Yes 0.68
US200700286	Sharing combi	G06F17/30887	2005-09-10	Yes 5.9
US24088882	Distributed ass	G06F17/30884 G	2007-05-02	Yes 3.07
US20070394C	System and me	G06F17/30 H04L2	2005-09-23	Yes 0.67
US20042282C	Method of data	G06F17/30 H04L2	2004-09-16	Yes 0.82
US7163882	Aggregating co	G06F17/30837 G	2005-09-12	No 16.77
US6893282	Peer to peer se	H04L29/08822 H	2004-04-13	No 7.86
US200700286	Generating and	G06F17/30884 G	2005-09-28	No 9.05
US200200050	Method, syste	H04L29/08 H04L	2004-10-02	No 5.58
US69347282	Method and app	G06F17/30887 Y1	2003-03-04	No 4.26
US6880982	Method and app	G06F17/30553 G	2003-01-10	No 2.81
US6882782	Baseline metho	G06F17/30 G06F2	2003-09-30	No
US6405782	Distributed net	G06F17/30884 G	1999-12-06	No

図 11 US8468147B2 の Similarity 結果

筆者が Ambercite Ai を特に有効活用する場合は、情報源である日経新聞記事に関する特許を先ず Sharesearch にて概念検索と Boolean 検索で調べてから、その代表的な特許群の引用被引用特許群を調べて、周辺特許群と過去特許群／最近特許群を網羅的に収集する際に活用している。

筆者は日経新聞記事を読む際に自分の注目したい技術テーマ (Private SDI theme) を 10 個ぐらい持ちながら大綱を張り、待ち伏せ作戦にて SDI 情報を集めている。例えば、(イ)AI 技術関連、(ロ)自動運転技術、(ハ)ドローン技術、(ニ)空飛ぶ自動車技術、(ホ)情報解析 IT 技術などである。

大きな綺麗な結晶は最初に小さな結晶核と長い成長時間が必要である様に、切り抜き新聞記事群(A)と Sharesearch で検索した手探り特許群(B)が最初の結晶核と位置付けている。それを継続的に数カ月間 → 1 年間 → 3 年間と続けることにより、Japio YEAR BOOK 原稿のネタと INFOPRO シンポジウム発表のための研究ネタが徐々に雪だるま的に大きく膨らみ、質的にも徐々に良好な Intelligence に改良されていくものと信じている。それを習慣行動にしている。71 歳の高齢になり視力と記憶力がすこぶる低下した状態において、若い IP 情報科学者 (良好な視力と抜群の記憶力を備えた若者たち) と上手くお付き合いをするために、筆者の弱点を補完する「習慣実践行動+AI 活用」が必要と信じている。だからこそ、AI を学ばざるを得ない。

## 8 結論

筆者の情報源と Intelligence への変換の手法と AI の第三のブームの図などを One-Sheet に纏めた。(図 12)

一気に拡張 IP 情報科学者には成れず、成長過程と訓練努力と時間が必要になる。自己啓発をスキップすることはできない。夢と願望を実現するためには、紙に描いた夢と願望を書斎の壁に貼り、日々眺めている。超具体的に日々一歩ずつ亀の如く登り続けるのが最善の方法と信じている。

筆者は Japio YEAR BOOK に執筆することと

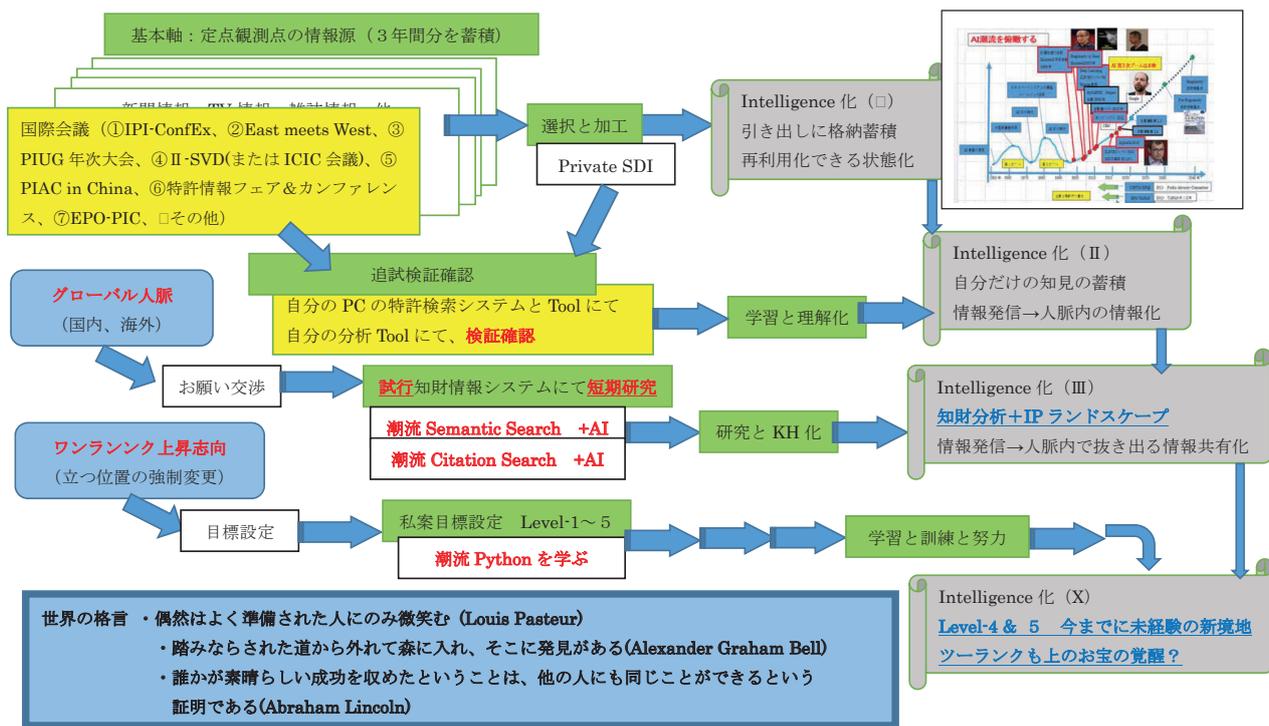


図 12 結論の俯瞰可視化 One-Sheet 図

INFOPRO シンポジウムに発表することを自分自身に強制している。最近では海外の IPI-ConfEx と米国 PIUG 年次大会における発表を自分自身に強制的に仕向けている。

米国 PIUG の正式の登録メンバーに日本人が少ないことを懸念している。中国の場合には PIUG-China-branch があり既に 1 万人以上が Jhon Wu 会長の下に組織化されている。今年の 9 月には、友人の Cynthia Barcelon Yang さん (元 PIUG 会長経験者) が PIUG からの派遣団の団長となり北京の PIAC in China 国際会議に行くご連絡が入った。

本記事の執筆内容は筆者の私見であり、所属機関の INFOSTA-SIG-パテントドキュメンテーション部会、またはアジア特許情報研究開の知財情報解析研究グループの公式見解ではないことを、ここにてお断りしておく。

## 9 おわりに

「世界を知り世界と共に生きる」をモットーに 48 歳から 71 歳までこの業界にて活動してきた。特許情報に関する 7 つの国際会議を通じて構築した Linkedin の人脈ネットワークの知人・友人達に心から感謝する。50 人以上の顔写真も収集している。筆者はそれらの中に知

財情報分野の 10 人の師を持っている。

今回この原稿を執筆するに際して、下記の支援者に謝辞を述べる。

XLPAT システムの創業者のお二人の Jitin Talwar 氏と Sandeep Singh Kohli 氏に試行の機会を与えて頂いたことに感謝する。

IEEE 協会 (日本、米国) 及び IP.com 社と、その仲介をして頂いた丸善雄松堂 (株) に、InnovationQ Plus の試行をさせて頂いたことに感謝する。

また、Ambercrite Ai の試行をさせて頂いた Doris Spielthener 氏と Mike Lloyd 氏に感謝する。

また、Shareresearch を使わせて頂いた日立製作所および日立技術情報サービス社にこの場を借りて御礼申し上げます。

特許分析ツールの PAT-LIST-GLS を常時使わせて頂いているレイテック社にこの場を借りて御礼申し上げます。

筆者は 71 歳の退職 OB なので、大手企業のサーチャーのように高額な有料の検索システムを日常的には利用できない。この謝辞にて述べた皆様のご理解とご支援とご協力の上で筆者の研究環境が初めて整う。長い間の信頼関係に基づき、今年もご協力ご支援を頂けたことに深く感謝する。

また、筆者が所属するアジア特許情報研究会の知財情

報解析研究グループの安藤リーダーをはじめとする皆様には微々細に亘りご指導ご支援を頂き、御礼申し上げます。

最後に、国内の専門家との討論だけでなく世界の専門家との討論も是非行われるように皆さまにお薦めする。

今後とも、「知財情報を通して企業の R&D 効率の向上を目指す」ことを筆者のライフワークとしたい。

## 参考文献

(web 参照日は全て、2016 年 8 月 27 日である)

- 1) NHK スペシャル「驚異の人工知能 天使か悪魔か」：2017 年 6 月 25 日、TV 放送
- 2) 欧州国際会議 IPI-ConfEx : <http://www.ipi-confex.com/>
- 3) East meets West : <http://www.epo.org/learning-events/events/conferences/emw2017.html>
- 4) 米国国際会議 PIUG 年次大会 2017 : <http://www.piug.org/an17meeting/>
- 5) II-SDV 2017 : <https://www.slideshare.net/Haxel/iisdv-2017-the-international-information-conference-on-search-data-mining-and-visualization>
- 6) PIAC in China : <http://www.piac-china.com/>
- 7) 特許情報フェア&カンファレンス : <http://www.pifc.jp/2016/>
- 8) EPO-PIC : <http://www.epo.org/learning-events/events/conferences/pi-conference.html>
- 9) World Patent Information 誌 : <http://www.sciencedirect.com/science/journal/01722190?sdc=1>
- 10) 筆者の JapioYEARBOOK2016 記事 : [http://www.japio.or.jp/00yearbook/files/2016book/16\\_2\\_02.pdf](http://www.japio.or.jp/00yearbook/files/2016book/16_2_02.pdf)
- 11) 東大に AI の寄付講座 9 億円基金 : [https://www.nikkei.com/article/DGXLASDZ30HYN\\_Q6A530C1TJC000/](https://www.nikkei.com/article/DGXLASDZ30HYN_Q6A530C1TJC000/)
- 12) 日経プラス 10 「AI 人材を育てろ！ 企業エンジニアが高校出前授業：2017 年 8 月 21 日の 22 時から TV 放送。
- 13) 米国シンギュラリティ大学 : [https://en.wikipedia.org/wiki/File:Singularity\\_University\\_Logo\\_2016.png](https://en.wikipedia.org/wiki/File:Singularity_University_Logo_2016.png)
- 14) AI 研究者にスパコン開放 米グーグル、無料で : [https://www.nikkei.com/article/DGKKASDC18H2O\\_Y7A510C1T11000/](https://www.nikkei.com/article/DGKKASDC18H2O_Y7A510C1T11000/)
- 15) TAPAS at EPO : <https://www.slideshare.net/Haxel/iisdv-2017-towards-semantic-search-at-the-european-patent-office>
- 16) USPTO の Sigma-PLUS by Thomas A Beach 氏 : <http://www.japio.or.jp/fair/files/2016/201609uspto.pdf>  
スライド 22-25 枚目
- 17) Oxford 大学のオズボーン准教授と野村総研との共同研究 : [https://www.nri.com/jp/news/2015/151202\\_1.aspx](https://www.nri.com/jp/news/2015/151202_1.aspx)
- 18) 特許情報と人工知能：情報の科学と技術、Vol.167, 2017, No.7, P340-349, 桐山勉、安藤俊幸
- 19) 日立総合特：許情報システム Shareresearch : <http://www.hitachi.co.jp/Prod/comp/app/tokkyo/sr/>
- 20) XLPAT システム : <https://www.xlpat.com/>
- 21) InnovationQ Plus : <http://innovationqplus.ieee.org/>
- 22) Ampercite Ai : <https://www.ampercite.com/>