

Visual Mining Studio、Visual R Platform、Deep Learner、Big Data ModuleはAlkanoとしてリニューアルしました。

開本 亮 様

公益財団法人 京都高度技術研究所／
特許庁委託事業 知財戦略デザイナー

学術論文と特許をマッピングする

知的財産の保護と活用において多方面で活躍中の開本亮様。大学の学術論文と特許をマッピングし、特許取得の可能性のある論文の抽出や、取得している特許がどの学問分野に関連があるかを一望できる、新発想のシステムを開発している。NTTデータ数理システムの**Visual Mining Studio** (以下、**VMS**)、**Deep Learner**、**Text Mining Studio** (以下、**TMS**) も利用して開発されているこのシステムについて、話をうかがった。



公益財団法人 京都高度技術研究所
インキュベーションマネージャー
特許庁委託事業 知財戦略デザイナー
博士(工学)・法学修士・MBA・弁理士
開本 亮 様

Interview

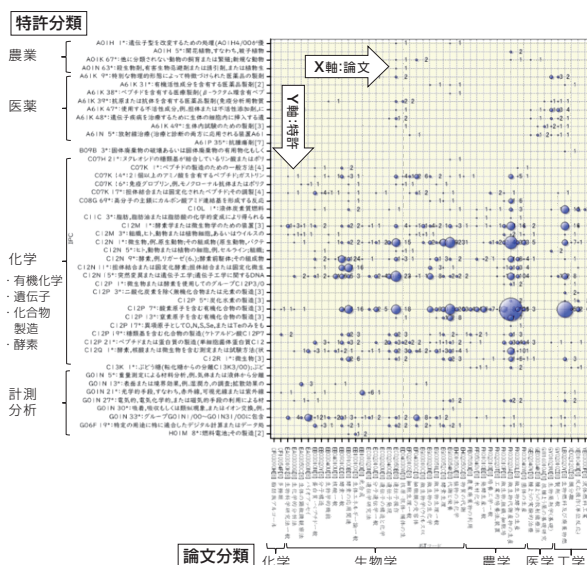
この論文はどういう特許が取得できそうか、を解析

大学の論文と特許の関係性を研究されているとお聞きしました。

開本 大学にはたくさんの知財が眠ったままで、非常にもったいない状況です。私は神戸大学で知的財産部門の部門長を務めていたことがあるのですが、そのときから痛感しています。大学の知財といえば学術論文です。神戸大学の場合、20年間で10万件の論文が発表されていますが、そこからどのくらいの特許が出願されているかというと、論文100本に対して特許1件程度、100対1の割合でしかありません。この比率の差をもっと縮めたい。大学の知財を特許というカタチで評価して社会に広く役立てていただき、そこから得られるリターンによって、大学の研究開発を行うというサイクルを回していきたい。そのためにまず、大学に眠っている論文の定量化と可視化をするシステムの開発を目指したのです。

論文をどのように定量化・可視化するのですか。

開本 最初に手掛けたのが、論文から特許へリンク付け、マッピングするシステムの構築です。論文にはもともと学術上の位置づけを示す論文分類が付与されていますが、特許分類は付与されていません。そこで共同研究者である中央大学理工学部教授の難波英嗣先生が作成されたAIを使って論文にも特許分類を付与し、特許上の位置づけを示すようにしました。そして、その結果を、論文分類は横軸、特許分類は縦軸としてマッピング表を作り、どの論文がどのような特許取得の可能性があるか分かるようにしています。この結果を参考にすることで、特許出願を増やして100対1だった論文と特許の比率を、100対10ぐらいには近づけることができるかなと思っています。また、このマッピング表は論文と特許の関連性の強さも評価しており、特許取得の可能性が高い分類を見たり、取得できそうな特許の領域を広く概観したりすることもできます。さらに論文と特許のほかに、時間の座標軸も加えた3Dでの表示もでき、どの学問分野の論文が増えたか、減ったかといった変遷を時系列で表示することもできます。



論文と特許の二次元マトリックスによるマッピング例

PROFILE

ひらきもと あきら
開本 亮 様

1981年、株式会社島津製作所航空産機事業本部入社。1984年、同社基盤技術研究所主任研究員等を歴任し、2007年、同社知的財産部専門部長。2011年、京都国際特許事務所副所長、2014年、神戸大学学術・産業イノベーション創造本部教授。2006年、弁理士登録。研究・イノベーション学会、日本弁理士会、日本知財学会、京都技術士会に所属し、特許出願から審判、訴訟、ライセンス等、知的財産の領域で活躍中。

[裏面に続く]

システム開発は、次の段階に入っているそうですね。

開本 特許から論文をリンク付けする逆パターンのシステムも開発中です。特許にはもともと特許分類が付与されていますが、論文分類は付与されていません。そこでAIを使って特許にも論文分類を付与しようとしています。そのAIには、難波先生のシステムに加えて、NTTデータ数理システムの**VMS**、**Deep Learner**、**TMS**なども活用しています。現段階は、大学が取得している特許を対象に開発を進めています。これができれば、特許がどの学問分野に関連性があるか分かります。また、企業にとっては産学連携の際の指針としても使えるでしょう。加えて、先の論文→特許のマップと重ね合わせることで、論文はたくさん書かれているのに特許出願がされていない学問分野を明らかにすることもできます。論文発表も特許出願もあまり進んでいない領域を見つけられれば、研究方針の策定などにも使えるでしょう。また、このマップがあれば、大学の全体的な研究状況が把握できるので、大学の研究者は同僚の研究状況が見やすくなり、自分の研究と同僚の研究を組み合わせたらどんなことができそうか、といったアイデア出しにも活用できると思います。

論文と特許を、どのように具体的にリンク付けするのですか。

開本 論文→特許の場合、ひとつの方法として、まず論文に対して日本で公開されている全特許、数百万件を母集団として類似する特許を見つけ、それらの特許の特許分類を類似度で重みをつけて集計することで、論文に対する特許分類を計算し付与できます。前述したように、対象とする論文データの集合に対し、付与された特許分類をクロス集計し、マッピング表にしています。マッピングにあたって、論文は国立研究開発法人科学技術振興機構によるJST分類を、特許は国際特許分類（IPC分類）を基準としています。JSTは学問体系で、またIPCは現実的な目的や利用シーンでそれぞれ分類され、項目立てもその内容も異なっているため、相互にリンクさせるために非常に苦労しました。また論文→特許の場合、両者が100対1の関係性のため「参考となる答えが非常に少ない」といった状況ですので、両者の関連づけが難しく、最良の結果となるよう論文の言語処理を難波先生によるAIで処理して、**VMS**によりクロス集計を行っています。

JST分類								IPC分類				
第1階層	第2階層	第3階層	第4階層	第1階層	第2階層	第3階層	第4階層	第1階層	第2階層	第3階層	第4階層	
A: 科学技術一般	5	35	64	M: 原子力	4	18	64	A: 農学・遠学等	16	84	1106	
B: 物理学	12	76	446	N: 電気工学	4	39	155	B: 地理・運輸	37	168	1985	
C: 化学	7	39	391	P: 熱機関	4	14	36	C: 化学・冶金	21	87	1321	
D: 宇宙・地球	6	29	132	Q: 機械工学	10	52	152	D: 織造・紙	9	39	350	
E: 生物科学	11	57	209	R: 建設工学	4	33	161	E: 固定構築物	8	31	323	
F: 農林水産	10	50	230	S: 環境	3	12	61	F: 機械・照明	18	97	1064	
G: 医学・薬学	25	143	363	T: 交通	5	5	23	G: 物理	14	81	696	
H: 工学一般	5	21	50	U: 資源	1	11	79	H: 電気	6	51	546	
I: システム・制御	3	9	30	W: 金属	5	18	114					
J: 情報工学	5	29	61	X: 化学工学	5	12	44					
K: 経営工学	2	14	41	Y: 化学工業	12	50	238					
L: エネルギー	5	8	19	Z: その他	2	20	25					

JST分類とIPC分類の比較

4階層・計3188にのぼる分類数

5階層・計約8万にのぼる分類数

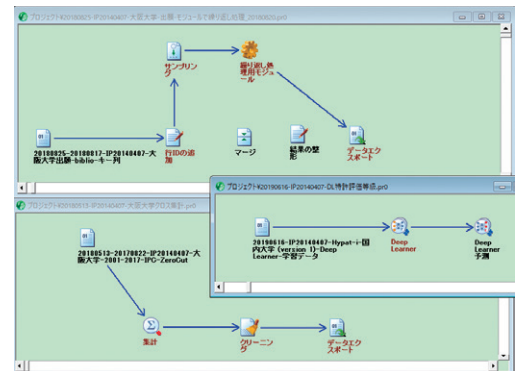
一方で、特許→論文の比率は逆で、「参考となる答えが非常に多い」という状況ですので、**TMS**で言語解析した結果を**Deep Learner**で解析し、**VMS**で統計処理することで最適な結果が得られると予測しています。現在、そのシミュレーションをさまざまに行っているところです。

当社の製品はどのようにお使いでしょうか。

開本 NTTデータ数理システムを知ったのはいまから6年前、その際に**TMS**や**VMS**などのツールを導入しました。当時、数理科学系の解析ソフトウェアは海外製が主流の中、国内企業でよくやっているし、製品の開発力もしっかりしていてこれから日本のデファクトになっていくのでは、と予感したのです。その3年後、今回の研究開発が科学研究費助成事業に採択されて本格化し、同社のツールを使う頻度も一気に高まりました。特に**VMS**はデータの预处理やクロス集計に欠かせません。同様の作業はExcelでも可能ですが、データのソートや他のデータとの照合、アウトプットまですべて手作業となります。プロジェクトで1つでも作業を怠るとデータが再現できなくなります。その点、**VMS**はインプットとアウトプットを決めたら、その結果は何度繰り返してもブレがなく、非常に効率的です。また、集計の際もどのデータをどこまで出すか簡単に設定でき、とても便利です。

このシステムはまだまだ成長するそうですね。

開本 実は、企業で開発部門の責任者をしていたとき、自社の特許が某企業から侵害されたことがありました。私は弁理士の資格を取り知財について勉強し、相手の会社に対抗する陣頭指揮を行いました。その際に痛感したのが、知財を守り活用するには特許しかないということで、それが今回の開発につながっています。大学の研究機関、企業、そして社会のエコシステムを回していくためにも、この開発をぜひとも成功させたいと思っています。今後は、この開発で得た解析手法をさらに進化させ、新規論文の発表時期やその内容を予測し、新たな特許取得や企業とのコラボレーションにつなげたりできるようなシステムの開発も視野に入れていきます。



VMSプロジェクト画面例