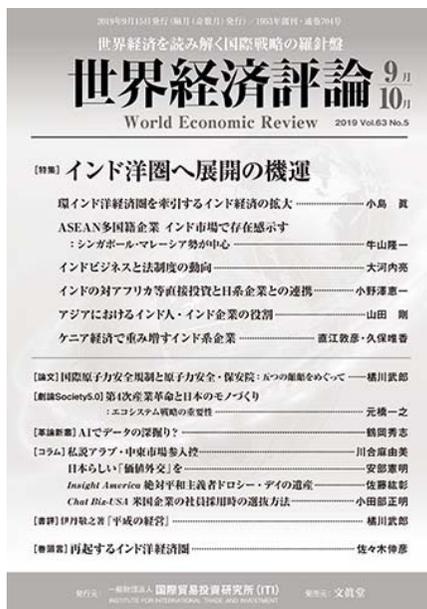


本論文は

# 世界経済評論 2019年9/10月号

(2019年9月発行)

掲載の記事です



世界経済評論

## 定期購読のご案内

年間購読料

1,320円×6冊=7,920円

6,600円

税込

17%

送料

無料

OFF



定期購読  
期間中

富士山マガジンサービス限定特典

※通巻682号以降

### デジタル版バックナンバー 読み放題!!



世界経済評論 定期購読



☎0120-223-223

[24時間・年中無休]

お支払い方法

Webでお申込みの場合はクレジットカード・銀行振込・コンビニ払いからお選びいただけます。  
お電話でお申込みの場合は銀行振込・コンビニ払いのみとなります。

Fujisan.co.jp

雑誌のオンライン書店



## AI でデータの深掘り？

信州大学先鋭材料研究所特任教授 鶴岡 秀志

2018年に続き、European Scientific Institute (ESI, 欧州研究機構) BioHealth Computing School の短期集中コース「Safe Nano Design, Summer School」の Faculty 一員として招聘された。ナノ材料応用について、大学院以上から選抜された学生向けの濃密なコースである。Faculty は講義だけではなく Coaching も行うため緊張の連続である。なお、ESI は欧州 CERN (欧州素粒子研究機構) の教育研究組織として位置づけられている。ジュネーブのフランス側 Archamps (アルシャン) 村の麦畑に囲まれた広大なキャンパス敷地を有していて、企業専用のオフィス棟も存在する官民共同の研究技術開発拠点である。この短期集中コースのテーマの1つが Bioinformatics である。

欧州共同体 EU では数年前から OECD と共同作業で欧州化学品登録規則 (REACH) に遺伝子影響情報と応用時の安全性予想を含めた高度統合システム Effectpedia として進化させる作業が行われている。最近の表現を使うと Bioinformatics を活用した化学品登録規則の新統合作業と言える。動物実験が制限されている欧州の安全性評価ハンデキャップを補い、これからのイノベーションを支える基盤技術として、膨大な既存データから新規化学物質の安全性予想を可能にしようとする意欲的な取り組みである。担当者の話では、京都大学大学院情報学研究科の Bioinformatics 研究成果を参考しているとのことである。我国の情報処理研究が世界有数の情報技術基盤構築のために使われていることは、この分野における我国の水準の高さを示す指標となるだろう。

昨今、データ分析が大流行で、統計はビジネス

だけではなくあらゆる分野で重要、イノベーションを創る打ち出の小槌、という言葉で満ち溢れている。しかし、記述統計学と推計統計学の分類がされず、本来数学である統計学が妙な話へと向かっている例が多々ある。統計熱力学は言うに及ばず、かなり高度な数学的知識を必要とする分野なのだが、用語の使い方が支離滅裂なコメンテーターによって「データは語る」「データを掘り下げる」的な言説が溢れている。しかし、データをいくら「深掘り」してもデータは何も語らないので、「データは語る」は騙りである。こんなコメンテーターや情報番組が幅を効かせている様では、せっかくの京都大学の先端研究も国内では間違って理解されてしまい、今後重要になる Bioinformatics の経済活動への貢献はお先真っ暗である。早晚、聞きかじりの知識で話を盛っている解説者を一掃しなければ、再生エネルギーや AI の様に本質とかけ離れた誤った理解が社会に浸透して、健全なイノベーションで社会が活性化していくことの妨げになりかねない。

情報処理というのは、事象を一旦数値や記号に置換えて、Variable (変数)、Function (関数)、Value (数値) の関係を演算・分析することである。その際、Parameter (媒介変数、母数) と Boundary Condition (境界条件) によって計算条件が変わる。メディアでは、やたらに「パラメーターは…」と発言して変数と媒介変数をごちゃまぜにしている話が多すぎるし、演算・分析を行う上で非常に重要な境界条件にはほとんど触れられていない。すなわち、芸能人的解説者によって情報処理に携わる専門家の考えることと世間一般の理解

が乖離してしまっている。この混乱の典型がNHKの「AIに聞いてみた」というシリーズ番組で、バケツの中に数値を全部放り込んでかき混ぜて、その中から「意外な結果が発見される」と放送している。因果関係が不確定なものをつなげれば「意外」なものが出現するのは当然なのだが、これを番組の売りにされては困る。情報処理は「風が吹けば桶屋が儲かる」を理論的に組み立てることで成り立っている。ディープ・ラーニングは樹状分析からなるので因果律から外れたら学問ではなくなってしまう。

「AIにきいてみた」みたいに、あらゆる事象（出来事）を闇鍋みたいな分析にかけてもおかしな結果しか出てこない。小学生でも習う簡単な2軸のグラフ（X軸、Y軸）のX軸に入れる数値は「等価」でなければならない。X軸を時間と決めたら距離を入れてはいけないのである。もちろん、多変数樹状分析、クラスター分析といった多変数分析手法もある。NHKが主張したいのはこのクラスター分析を用いて解析するというのだが、それでもデータセットの変数は「等価」という数学的条件をそれぞれの「決定木」で考慮しなければならない。人間の判断が入る「クレンジング」を詳らかにしなければ「AIにきいて」も訳が分からなくなるだけである。むしろ、経済人には馴染みのある日経平均株価の時間推移はクラスター分析の一つなのでこれを取り上げたほうが良かったのではないか。それぞれの銘柄の、価格の時間推移をまとめて一つの曲線で示す際、企業業績や社会情勢に影響を受ける個々の銘柄の変動はバラバラで統一性がない様に見えることがあるが、統計的手法でバラけ方を偏差として計算する。この株価変動グラフでは先達が「ローソク足」という一種の偏差を表す、判り易い便利な表示方法を発明してくれた。科学的論文でもグラフに偏差を表示することがあり、このローソク足に見た目が似ている。

この様に情報処理というのは、最低限、高校の

数Ⅱ程度の知識を前提として話をしなければ話がトンチンカンとなってしまいます。面白おかしく経済解説をしようとしている他局も芸能人に解説させる番組を持つので同罪である。もちろん、「AIによるInfomatics分析」というのは未だ限定的にしか可能でない。あくまでも、数値を行列として並べて関数による変換、多数値出力という行列計算を行なっているだけである。クラスター分析は時には群論を用いて処理することになるが、これも行列を応用した関数処理である。AIと称されるもののほとんどは高性能計算機による高度演算処理である。

ここまでの話でも、工学系出身でない方々にとって数学的な記述についていけないという諸兄も多いことと思う。テレビでは視聴率を維持しようと流行語の使いまわしや意外性の提示に終始して、結果的にいい加減な言説が流布してしまうのである。常々思うことは、専門家としてTVでコメントしている方々は、ビッグデータ分析に使われる数学を少しでも理解しているのかという疑問である。21世紀になって以来、プログラミングは数学を理解していなくても使えるものになってきた。昔はMNEMONICやFORTRANのように代数学を勉強しないと使えないコンピューター言語が主流であったが、C言語を経て、現在では人間の感覚に近い言語でプログラミングができるようになった。面倒な計算は、外部関数という形でブラックボックス化されているので「ボタンを押す」感覚でプログラムできる。しかし、今でもInfomaticsの研究者は、事象（出来事）を数学的記述に変換してアルゴリズムを設計し、プログラミングする方法を地道に追求している。解らないからと言って真の専門家を蔑ろにはしていない。

さて、欧州は数学者や化学者を束ねて化学品のデータベースを構築しているのが、そこから何を導ようとしているのか。答えは世界標準である。化学物質のデータベース（DB）を持つこ

とは国力と言えるので、ナノ材料の登録を切り口にして米国の保有する世界最大の化学物質 DB の Chemical Abstract (CA)\* に対抗しようとしていることは明らかである。米中の通商交渉では、今更手を出せない化学物質 DB の代わりに中国が通信を掌握しようとしているので覇権争いに至っているという見方もできる。Data Sheet は全ての製品の通商において必須で、その化学物質の正式名称、国際登録番号、基本物性は「全て」CA に記載されているものに準拠することが要求される。つまり、新規物質合成に成功しても CA に掲載されなければ通商に大きなハンデキャップを被ることになる。逆にいうと、米国は自国の利益に反する新規物質の登録を故意に遅らせる手段を持っていることになる。また、CA 利用は基本的に有償なので内容を確認するにも閲覧料を支払わなければならないが、昔の数十冊の CA 便覧本に比べて電子化された CA はずいぶんと高価なものになっている。欧州はナノ材料で米国に先行して化学品登録 REACH の DB を構築しているのだろう。ただし、彼らは AI 活用とは一言も行っていない。

ESI での議論の中で欧州のナノ技術目論見が示された。大量使用されている物質をナノ材料で置き換えて、エネルギー効率がよく環境影響の少ない代替物を生み出す腹づもりである。また、我国同様にインフラの老朽化に悩んでいるので、特に土木建築資材にナノ材料の活用を狙っている。また、中国の廃プラ輸入禁止により大きな影響を受けているので、ナノ材料を用いて多くの樹脂、特にエンブラをリサイクル可能な樹脂コンポジットに置き換えようとしている。

我国も技術立国を標榜するならば、官民を挙げてこの Bioinformatics の覇権争いに立ち向かわなくてはいけない。そろそろナンチャッテ科学技術番組を止めて、真面目番組で先端技術を多くの日本人に正しく伝え、国民の支持を得てイノベーションを進めなければならない。

(つるおか しゅうじ)

[注]

\* 21 世紀になり、米国 CA がドイツの世界最大の有機化合物 DB である Beilstein を獲得したので CA が統一 DB になった。

**世界でも例をみない諸外国の直接投資統計を調べる統計年鑑**  
**世界主要国の直接投資統計集 (2019 年版) II. 国別編—CD-ROM 版—**

※印刷イメージの PDF 版と EXCEL 形式のデータ編で構成。 発行：2019 年 6 月 / 価格：70,000 円  
 ※米ドル建に換算した数値データが利用できます。(自国通貨建と米ドル建の 2 種類の表で構成)

日本で唯一の直接投資統計の年鑑。1997 年以来毎年発行し 24 回目。

- ・日本企業の進出が多い国・地域を中心に、対内および対外直接投資統計を収録。収録国数：57 개국 (日本を含む)
- ・各国・地域の中央統計局、中央銀行、外国企業誘致促進機関等が作成する直接投資統計をもとに最新時点までの時系列データを掲載

【収録国・地域】アジア・太平洋地域 [中国、香港、韓国、台湾、フィリピン、タイ、シンガポール、ベトナム、ラオス、マレーシア、ミャンマー、インド、バングラデシュ、パキスタン、オーストラリア、ニュージーランド]

米州 [米国、カナダ、メキシコ、ブラジル、チリ、ペルー]

欧州 [英国、ドイツ、フランス、アイルランド、ベルギー、オランダ、ルクセンブルク、ノルウェー、スウェーデン、フィンランド、デンマーク、オーストリア、スイス、スペイン、ポルトガル、ポーランド、チェコ、ハンガリー、ルーマニア、ブルガリア、エストニア、ラトビア、リトアニア、クロアチア、スロベニア、キプロス、ギリシャ]

その他 [ロシア、イスラエル、南アフリカ、トルコ]

- ・見本 [http://www.iti.or.jp/report\\_90.pdf](http://www.iti.or.jp/report_90.pdf) をご参照ください。
- ・姉妹統計年鑑の「世界主要国の直接投資統計集 I. 概況編」を併用してお使いになると便利です。

※お問合せ、ご購入をご希望の方は下記までご連絡ください。

一般財団法人 国際貿易投資研究所 (ITI) TEL: 03(5148)2601 / FAX: 03(5148)2677  
 〒104-0045 東京都中央区築地 1 丁目 4 番 5 号 第 37 興和ビル 3 階 E-Mail: jimukyoku@iti.or.jp / URL: <http://www.iti.or.jp/>