あすへの考

生成AI 新たな「人類の脅威」

編集委員室 大塚隆一

人間が書いたような文章を作ることができる生成AI（人工知能）の利用が急拡大している。一方で「人類や文明の存続を脅かす」などの警鐘も相次ぐ。なぜ、それほど恐れるべきなのか。核兵器や気候変動など他の脅威と何が違うのか。それらとの比較で何が見えてくるのか。

パンドラの箱

「人類存亡の脅威」と聞くと、何を思い浮かべるだろうか。

自然がもたらす破局的な脅威には小惑星の衝突や超巨大火山の噴火などがあるが、ここでは人間の活動、特に技術や産業の発展で生じた脅威を取り上げたい。

具体的には「気候変動」「核兵器」「遺伝子の改変」「人工知能（AI）」の四つだ。程度の差はあるが、どれも人類の存続を揺るがすリスクをはらむ。それらを年代順に並べたのが下の表である。

このうち「気候変動」と「核兵器」は「エネルギー」に関わる。

18世紀後半からの産業革命で化石燃料の採掘・利用が広がり、温暖化につながった。気温上昇が続けば、地球の環境は激変する。

20世紀に入ると、人類は原子核のエネルギーを発見し、核爆弾を製造した。今や世界の核は文明を破壊しつくす潜在力を持つ。

一方、「遺伝子の改変」は「命」の謎解きに挑む生命科学が生んだリスクだ。特に近年、遺伝子を自在に操作できるゲノム編集が登場したことで懸念が強まった。

この技術を人の生殖細胞や受精卵に使い、遺伝子を望み通りに変えた「デザイナーベビー」を誕生させるとどうなるか。改変の影響は子々孫々まで残る。専門家は人類の多様性や進化に未知の問題を生じさせかねないと危惧する。

もちろん科学技術や産業の発展は多大な恵みをもたらしてきた。

数次の産業革命で私たちの暮らしは豊かになった。核エネルギーは原子力という新しい電源を生んだ。核医学にも期待が集まる。

遺伝子の研究は難病の治療や新薬の開発、作物の品種改良などでめざましい成果を上げてきた。

ただ人類は、恵みと引き換えに、扱いを誤れば自らの生存を危うくするリスクを背負った。

いわば、災いが詰まった「パンドラの箱」を開けてしまった。

おおつか・りゅういち 科学部、国際部、ジュネーブ、ニューヨーク、ワシントン支局を経て記者生活43年目。時代の大きな流れや構図を伝える記事が目標。

言語乗っ取り 人心操る懸念

民主主義の破壊

「人工知能」のうち、いま話題の「生成AI」は「知」の分野の驚くべき成果だ。代表格の対話型AI「チャットGPT」は人間のような巧みさで「言語」を操る。

それゆえ、脅威にもなりうる。こちらの「パンドラの箱」はどんな災いをもたらすのか。

政府のAI戦略会議は先月、懸念されるリスクとして、偽情報の氾濫、犯罪の巧妙化、著作権の侵害など7項目を挙げた。

一方、イスラエルの歴史家ユヴァル・ノア・ハラリ氏は英誌エコノミストで、「言語」が人類の文明を築いてきたことを考えれば、生成AI問題はもっと「大きな構図」で捉えるべきだと論じた。

「民主主義は対話であり、対話は言語による。AIが言語を乗っ取れば、有意義な対話、すなわち民主主義は破壊されかねない」

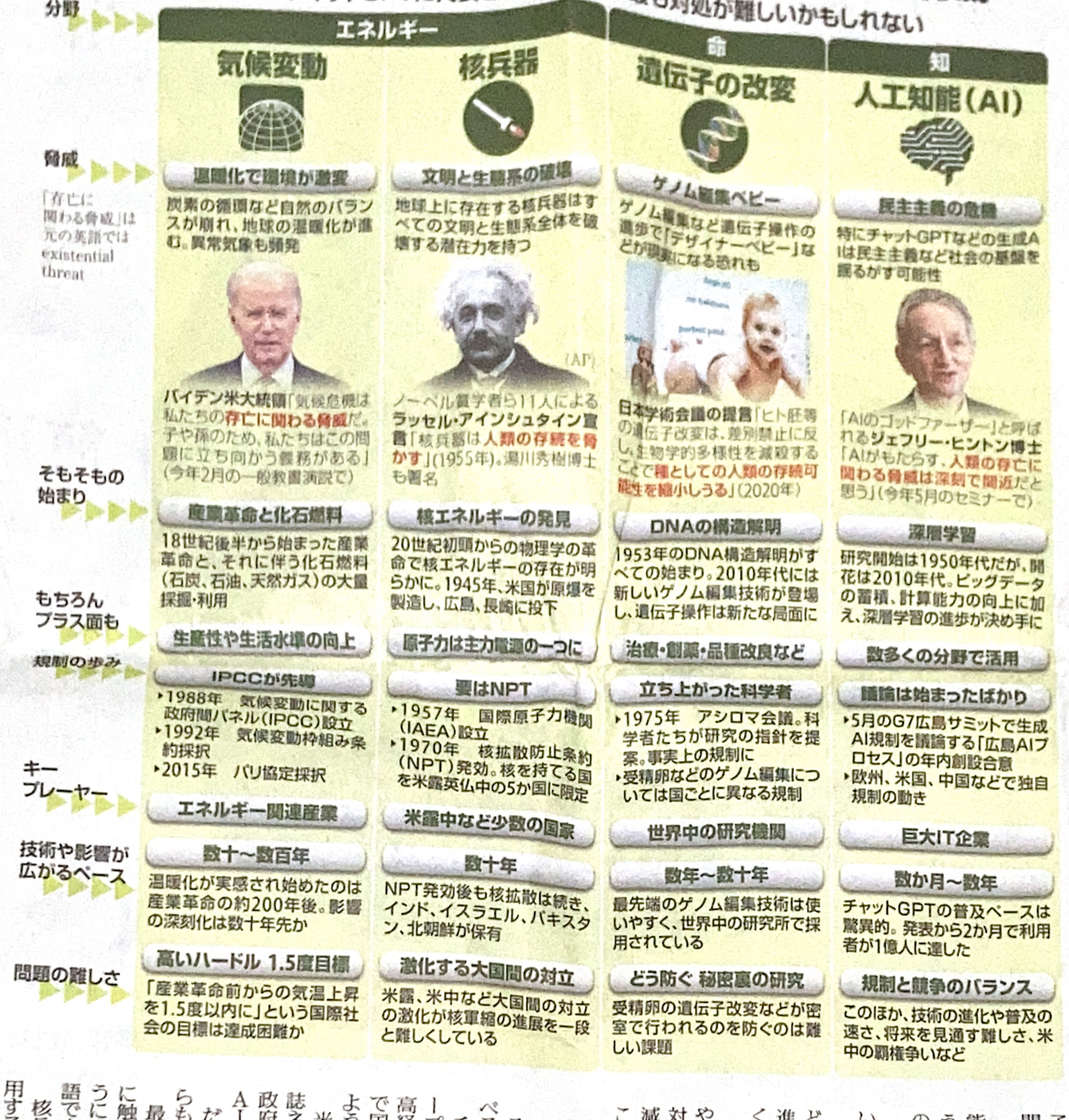
さらに「核兵器は文明を物理的に破壊できる」が、生成AIは「私たちの精神世界と社会」を滅ぼす「新しい大量破壊兵器」になりうる、とまで指摘した。

もちろんハラリ氏も、生成AIが社会の抱える様々な課題の解決に役立つ可能性は認めている。

しかし今は、その能力を見極め、規制を優先すべき時だと訴える。

人類の存続や文明を揺るがしかねない4つの脅威

チャットGPTに代表される生成AIは最も対処が難しいかもしれない



分野	エネルギー	エネルギー	命	知
	気候変動	核兵器	遺伝子の改変	人工知能（AI）
脅威 「存亡に関わる脅威」は元の英語ではexistential threat	**温暖化で環境が激変** 炭素の循環など自然のバランスが崩れ、地球の温暖化が進む。異常気象も頻発 バイデン米大統領「気候危機は私たちの存亡に関わる脅威だ。子や孫のため、私たちはこの問題に立ち向かう義務がある」（今年2月の一般教書演説で）	**文明と生態系の破壊** 地球上に存在する核兵器はすべての文明と生態系全体を破壊する潜在力を持つ （AP） ノーベル賞学者ら11人によるラッセル・アインシュタイン宣言「核兵器は人類の存続を脅かす」（1955年）。湯川秀樹博士も署名	**ゲノム編集ベビー** ゲノム編集など遺伝子操作の進歩で「デザイナーベビー」などが現実になる恐れも 日本学術会議の提言「ヒト胚等の遺伝子改変は、差別禁止に反し、生物学的多様性を減殺することで種としての人類の存続可能性を縮小しうる」（2020年）	**民主主義の危機** 特にチャットGPTなどの生成AIは民主主義など社会の基盤を揺るがす可能性 「AIのゴッドファーザー」と呼ばれるジェフリー・ヒントン博士「AIがもたらす、人類の存亡に関わる脅威は深刻で間近だと思う」（今年5月のセミナーで）
そもそもの始まり	**産業革命と化石燃料** 18世紀後半から始まった産業革命と、それに伴う化石燃料（石炭、石油、天然ガス）の大量採掘・利用	**核エネルギーの発見** 20世紀初頭からの物理学の革命で核エネルギーの存在が明らかに。1945年、米国が原爆を製造し、広島、長崎に投下	**DNAの構造解明** 1953年のDNA構造解明がすべての始まり。2010年代には新しいゲノム編集技術が登場し、遺伝子操作は新たな局面に	**深層学習** 研究開始は1950年代だが、開花は2010年代。ビッグデータの蓄積、計算能力の向上に加え、深層学習の進歩が決め手に
もちろんプラス面も	生産性や生活水準の向上	原子力は主力電源の一つに	治療・創薬・品種改良など	数多くの分野で活用
規制の歩み	**IPCCが先導** ▸1988年 気候変動に関する政府間パネル（IPCC）設立 ▸1992年 気候変動枠組み条約採択 ▸2015年 パリ協定採択	**要はNPT** ▸1957年 国際原子力機関（IAEA）設立 ▸1970年 核拡散防止条約（NPT）発効。核を持てる国を米露英仏中の5か国に限定	**立ち上がった科学者** ▸1975年 アシロマ会議。科学者たちが研究の指針を提案。事実上の規制に ▸受精卵などのゲノム編集については国ごとに異なる規制	**議論は始まったばかり** ▸5月のG7広島サミットで生成AI規制を議論する「広島AIプロセス」の年内創設合意 ▸欧州、米国、中国などで独自規制の動き
キープレーヤー	エネルギー関連産業	米露中など少数の国家	世界中の研究機関	巨大IT企業
技術や影響が広がるペース	**数十～数百年** 温暖化が実感され始めたのは産業革命の約200年後。影響の深刻化は数十年先か	**数十年** NPT発効後も核拡散は続き、インド、イスラエル、パキスタン、北朝鮮が保有	**数年～数十年** 最先端のゲノム編集技術は使いやすく、世界中の研究所で採用されている	**数か月～数年** チャットGPTの普及ペースは驚異的。発表から2か月で利用者が1億人に達した
問題の難しさ	**高いハードル 1.5度目標** 「産業革命前からの気温上昇を1.5度以内に」という国際社会の目標は達成困難か	**激化する大国間の対立** 米露、米中など大国間の対立の激化が核軍縮の進展を一段と難しくしている	**どう防ぐ 秘密裏の研究** 受精卵の遺伝子改変などが密室で行われるのを防ぐのは難しい課題	**規制と競争のバランス** このほか、技術の進化や普及の速さ、将来を見通す難しさ、米中の覇権争いなど

進歩の速さ 規制追いつかず

複雑で厄介

同様の懸念は生成AIの開発者たちも抱いている。その一人、ジェフリー・ヒントン博士は米国で先月開かれたセミナーで、チャットGPTの予測以上の賢さに驚き、怖さも感じると語った。

「最新のモデルは常識的な推論ができる。IQ（知能指数）は80か90くらいか。古今の小説とマキャベリの人心掌握術から学び、もっと賢くなったら、私たちを巧みに操れるようになるだろう」

歴史を顧みれば「核」や「遺伝子」でも研究開発に関わった科学者たちは真っ先に警鐘を発した。

核エネルギー発見に貢献したアインシュタイン博士は他のノーベル賞受賞者らと核兵器の脅威を警告する歴史的な宣言を出した。

DNAの構造解明で知られるワトソン博士は遺伝子操作の自主規制を話し合った有名なアシロマ会議の呼びかけ人の一人だった。

だが、生成AIが他の脅威と共通するのはこの点くらいだ。表が示すように異なる点の方が多い。

このため、生成AIの規制は一段と複雑で厄介なものになる。

第一に、技術の進歩や影響の波及のスピードが桁違いだ。だから規制が追いつかない。

チャットGPTの利用者は昨年11月の公開から5日で100万人、2か月で1億人に達した。

気候変動は産業革命の幕開けから温暖化が実感されるまで200年以上、核兵器は5か国への拡散に約20年、北朝鮮を含む9か国に増えるまでに約60年かかった。どちらも規制について議論するそれなりの時間的余裕はあった。

第二に、関わるプレーヤーが質も数も他の脅威とは異なる。

まず、主役は巨大IT企業だ。このことは営利企業が人類の命運に関わるかもしれない技術の根幹を握ることを意味する。

また、生成AIは文章の作成がとても楽になる。スマホでも使えるようになってきた。悪用しようと思えば、誰でも可能だろう。

核を持つのは少数の国家、遺伝子の改変も扱えるのは限られた専門家であるのとはワケが違う。

第三に、生成AIを含む人工知能は競争力や軍事力の源泉になりうる。従って規制の議論には米中の覇権争いなども絡んでくる。

第四に、何をどう規制すればよいのか、不明な点が多い。

そもそも生成AIがなぜこれほどの能力を得たのか、今後どんな進歩が予想されるか、専門家もよく分かっていないからだ。

この点は課題が明確な「気候」や「核」と大きく異なる。温暖化対策には二酸化炭素などの排出を減らせばいい。核軍縮もなすべきことははっきりしている。

難しいバランス

それでも他の脅威の経験から学べる点はないだろうか。

チャットGPTを開発した米オープンAIのサム・アルトマン最高経営責任者（CEO）は米議会で国際原子力機関（IAEA）のような組織が必要だと証言した。

米国の科学者グループは英科学誌ネイチャーで気候変動に関する政府間パネル（IPCC）の生成AI版の創設を提案した。

だが様々な利害が交錯し、どちらも実現は簡単ではなさそうだ。

最後に、もう一つの重要な論点に触れたい。ハラリ氏らが言うように、生成AIで特に怖いのは言語で人心を操作しうる点だ。

核兵器や遺伝子改変の脅威は悪用する人間がいて初めて現実のものになる。生成AIはそんな人間を生み出す恐れがある。ヒントン博士の言葉を借りれば「AIはレバーを引けなくても、人に引かせることができるかもしれない」。

その意味で、生成AIは深刻さのレベルが一段高いとも言える。

繰り返しになるが、生成AIは多大な恩恵をもたらす可能性を秘める。一方で最悪の場合、文明の基盤を揺るがしかねない。

どんな規制が有効なのか。開発や活用とのバランスはどうすべきか。技術で悪用は防げるのか。人類は核や気候変動とは異質の、複雑な難題と向き合うことになる。

先週の出来事

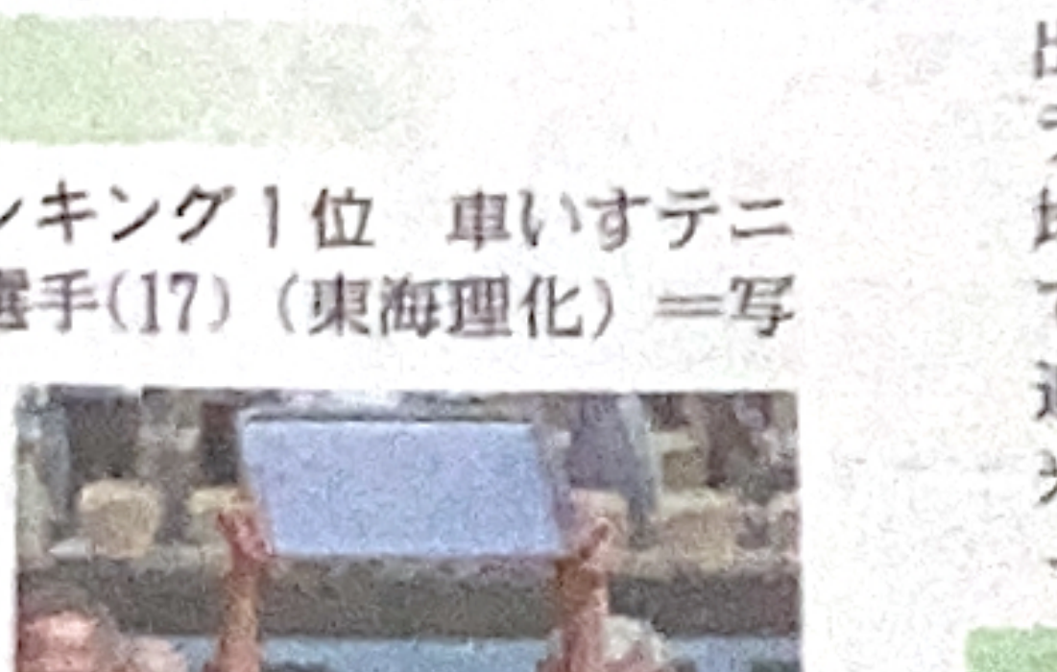
6月11日～17日

11日 日

▷「参院のドン」青木幹雄さん死去

12日 月

▷**小田選手が世界ランキング1位** 車いすテニス男子の小田凱人選手(17)（東海理化）＝写真、AP＝が、国際テニス連盟の12日付の世界ランキングで1位となった。11日に閉幕した四大大会、全仏オープンの男子シングルス決勝で、アルフィー・ヒューエット選手（英）を破り、史上最年少で初優勝。それまで1位だったヒューエット選手をポイントで上回った。



13日 火

▷**トランプ氏、37件無罪主張** 機密文書を持ち出した連邦法違反などで起訴された米国のトランプ前大統領がフロリダ州マイアミの連邦地方裁判所に出廷し、計37件の起訴事実全てで無罪を主張した。米大統領経験者が連邦法違反で起訴されたのは史上初。機密文書には米国や外国の核戦力を含む軍事力の記述もあった。

14日 水

▷**陸自訓練中銃撃、3人死傷** 陸上自衛隊「日野基本射撃場」（岐阜市）で自衛官候補生の男(18)が訓練中、指導役の男性自衛官3人に小銃を発砲し、1曹(52)と3曹(25)が死亡、別の3曹も負傷した。岐阜県警は男を殺人未遂容疑で現行犯逮捕した。15日には、容疑を殺人に切り替え、岐阜地検に送検した。

▷**米利上げ見送り** 米連邦準備制度理事会（FRB）は連邦公開市場委員会（FOMC）で、政策金利の引き上げを11会合ぶりに見送った。急ピッチで進めた利上げによる経済への影響を見極める。一方、高水準が続くインフレ（物価上昇）抑制のため、年内に計0.5%の追加利上げを行うことを示唆した。

15日 木

▷産総研の技術漏えい容疑 中国人研究員逮捕

▷首相、今国会での解散見送り表明

▷北ミサイル2発 日本のEEZ内に落下

16日 金

▷防衛財源法・LGBT法が成立

▷杉下茂さんが12日に死去したことが明らかに

▷北別府学さん死去

海外ニュースは現地時間

今週の予定

6月18日～24日

18日 日	▷ブリンケン米国務長官訪中（～19日）
19日 月	▷衆参両院が旧優生保護法に基づく強制不妊の被害実態をまとめた調査報告書の全文を公表
20日 火	▷独中政府間協議（ベルリン） ▷サッカー・キリンチャレンジ杯 日本―ペルー（大阪・パナソニックスタジアム吹田）
21日 水	▷通常国会会期末 ▷ウクライナ復興会議（～22日、ロンドン）
22日 木	▷新たな国際的開発資金取り決めのための首脳会合（～23日、パリ）
23日 金	▷天皇、皇后両陛下がインドネシア公式訪問から帰国 ▷沖縄慰霊の日 ▷5月の全国消費者物価指数を発表
24日 土	▷先進7か国（G7）男女共同参画・女性活躍担当相会合（～25日、栃木県日光市）