



米国防長官、ネブラスカ州の空軍基地を訪問(日付：2020年2月20日)[写真：AP/アフロ]

研究レポート

新興技術と核抑止関係

2021-03-30

戸崎洋史（日本国際問題研究所軍縮・科学技術センター 主任研究員）

「安全保障と新興技術」研究会 第8号

「研究レポート」は、日本国際問題研究所に設置された研究会参加者により執筆され、研究会での発表内容や時事問題等について、タイムリーに発信するものです。「研究レポート」は、執筆者の見解を表明したものです。なお、各研究会は、「研究レポート」とは別途、研究テーマ全般についてとりまとめた「研究報告書」を公表する予定です。

新興技術と核兵器システム

核態勢・抑止関係に関する議論において、新興技術がもたらし得る含意は重要な論点の一つとなってきた。核兵器システムへの新興技術の導入に関する核保有国の具体的な目的、構想、計画、開発状況は必ずしも明らかではないが、特に議論の焦点となっているのは、情報収集・警戒監視・偵察(ISR)や核指揮・統制・通信(NC3)への人工知能(AI)や量子技術などの導入である。

早期警戒、脅威の探知、状況認識、攻撃・被害結果の判定などを行うISRでは、たとえば量子センシングなどによりリモートセンシング技術が発展すれば、敵対国の攻撃能力に対する探知能力が向上し、その使用前の対処可能性を高め得る。クラウドコンピューティング、高速・大容量データ通信、AIなどの活用は、膨大な情報の効率的な集約と迅速な解析・分析を可能にする。

また、核兵器の使用に係る意思決定を司るNC3に先端的なAIを導入することで、核兵器の使用が想定される高い緊張状況、さらにはすでに核戦争下にあるという複雑な状況において、核兵器の使用の是非、使用する場合の核攻撃の規模、攻撃目標、使用する核兵器の割り当てなどに関する選択肢を指揮官に迅速、適切・正確かつ効率的に提供すると期待されている。上述のような

厳しい状況でのAIによる支援は、認知バイアスやヒューマンエラー、あるいは誤報などによる核兵器の不要・不適切な使用の可能性を低減するとも考えられている。さらに、量子通信の実用化は、核戦争のような苛烈な環境下でも確実な通信を可能にすると考えられる。

新興技術は主として、核兵器システムの「神経系」の能力を高めるものであり、核兵器の他を凌駕する圧倒的な爆発威力がもたらす「核革命」を覆すようなゲームチェンジャーになるわけではない。しかしながら、それは抑止関係への影響が軽微だということの意味するわけではない。

抑止関係の安定化？

たとえば、核兵器システムへの新興技術の上述のような導入は、危機においても先制核攻撃の誘因が生じにくい状態である「危機安定性(crisis stability)」、および戦略核戦力などの量的・質的増強の誘因が抑えられた状態である「軍拡競争に係る安定性(arms race stability)」からなり、核戦争が勃発する公算の低い状況を意味する「戦略的安定(strategic stability)」に寄与するとの見方がある。戦略的安定の維持には、抑止を試みる側(抑止国)が、敵対国(被抑止国)の先制核攻撃を受けても懲罰的な核攻撃(報復攻撃)を確実に実行する能力(懲罰的抑止力、deterrence by punishment)を保持することが鍵を握るとされてきた。ISR能力の強化によって、抑止国が持つ報復能力への敵対国による攻撃を早期に探知できれば、報復能力が破壊される前に対処する可能性を高めることができる。また、NC3の向上によって、情報の過多・煩雑や時間の切迫といった状況のなかでも、抑止国が報復攻撃を敢行する確実性を高め得る。こうして報復能力の残存性(survivability)および信頼性が高まれば、敵対国は抑止国に対する先制核攻撃の遂行を強く抑制されることになる。同時に、抑止国は、懲罰的抑止力が無効化される懸念から先制核攻撃を遂行したいとの誘因も抑制される。

また、敵対国の核攻撃に対して報復能力を自動的に発動するシステムの構築は、それが技術的・政治的に適切に作動する限りにおいて、核兵器使用の権限を持つ指導者が敵対国の攻撃目標になったとしても報復能力を発動できるとの抑止国に安心感を与えるものとなり得る。あるいは、AIの活用によって高度なシミュレーションや演習の実施が可能になり、それらを通じて抑止国は自国の報復能力に対する自信を持つことができる。

新興技術の導入によって報復能力の非脆弱性を高めることができれば、抑止国・被抑止国の双方にともに先制核攻撃の誘因が生じにくい状態、すなわち危機安定性の維持・向上が期待し得る。また、報復能力の非脆弱性が担保されれば、報復攻撃に必要な規模以上の核戦力を保持する必要性が低減されるため、軍拡競争に係る安定性も保たれることになる。

抑止関係の不安定化？

他方で、新興技術の核兵器システムへの導入は、抑止関係を不安定化させるとの見方もある。

たとえば、抑止国が、強化されたISRとNC3(あるいは通常戦力用または核・通常両用のC3)によって、敵対国の核戦力を高いレベルで弱体化させ、あるいは無効化できるような効果的な攻撃能力(損害限定能力)を備えれば、その抑止国は敵対国が核戦力を使用する前に先制攻撃を敢行したいと考えるかもしれない。これに対して、そうした状況に直面する敵対国も、その核戦力が無力化される前に使用したいとの誘因を強めるかもしれない。とくに、先端的なAIを導入したNC3によって意思決定プロセスが高速化すれば、それだけ抑止国と敵対国の間で軍事力の早期使用に係る競争も加速化することになる。

また、抑止国による効果的な損害限定能力の獲得は、報復能力の弱体化・無効化を懸念する敵対国に、その量的・質的拡充を迫るものとなり得る。抑止国がこれに対抗して損害限定能力の一層の獲得を図るといったスパイラルに陥れば、軍拡競争に係る安定性は損なわれる。

しかも、敵対国の攻撃に対する損害限定能力は「自国を守る」兵器であり、技術的に可能であればその獲得・強化に制動をかけることは容易ではない。冷戦期に米ソが相互確証破壊(MAD)状況を受忍し、戦略的安定性を目的として二国間の核軍備管理を追求した重要な要因の一つは、他方の核戦力を無効化できるほどの効果的な損害限定能力を技術的に保持し得ないという現実であった。

現時点で、核兵器システムへの新興技術の導入が、技術的に高い損害限定能力の達成をもたらすか否かは分からないが、通常兵器システムに導入される新興技術は核兵器システムへの導入も検討されるであろうし、ISRやC3については核・通常両用のシステムが構築される可能性もある。仮にそれらが損害限定能力の強化にさほど寄与しない場合でも、敵対国はその報復能力が大

大きく損なわれる可能性を念頭に核戦略や運用・調達計画などを立案するであろう。そのことが、結果として戦略的安定性を低下させることにつながり得るとも指摘されている。

意図せざる核兵器使用のリスク

戦略的安定性への含意以上に注視されているのが、新興技術の核兵器システムへの導入が、誤解、誤認、誤算あるいは事故などによる意図せざる核兵器使用のリスクを増幅させるとの懸念である。

第一に、たとえば武力紛争では、敵対国の軍事行動を阻害すべく核兵器用、あるいは核・非核両用のISRやC3を攻撃する強い誘因が生じるであろうが、特に敵対国の核・非核両用のシステムに対して、その通常攻撃に対する損害限定を企図して攻撃を敢行したにもかかわらず、敵対国は核兵器システムへの攻撃だと判断して核報復攻撃で対応するかもしれない。

第二に、意思決定の自動化が進む場合、自動化バイアス、すなわち高ストレス下でAIが下した判断や選択肢を、仮にそれが不完全、不正確あるいは不確実であったとしても、まさにAIが下したとの理由で「最適解」だと信頼し、選択するリスクも指摘されている。もとより、AIが示す判断や選択肢をどこまで信頼できるかという問題もある。AIの信頼性はデータの質、トレーニングの有効性、操作のパラメータに大きく依存しており、特に1945年8月の原爆投下以来、75年以上にわたって実戦での使用がない核兵器について、それらが十分に信頼できるレベルで組み込めるか、どのレベルであれば信頼し得るかを判断する基準は容易には設定できない。

第三に、新興技術を導入した核兵器システムの複雑性・不確実性や、アルゴリズムのミス、さらには予見されていないリスクの顕在化により、予期せぬ重大な事故によって核兵器が使用される可能性もある。特に、新興技術導入の競争のなかで、信頼性の低い技術の早期採用、あるいは潜在的・顕在的なリスク要因を軽視、過小評価した導入がなされれば、意図せざる核兵器使用のリスクを一段と高めかねない。

第四に、抑止関係は敵対国間の認識の相互作用によって規定されるが、NC3のAI化が進めば、彼我の人と人との間だけでなく、人とAIの間、およびAIとAIの間の認識をどのように考えるかという問題も生じよう。多数国間の抑止関係と同様に、AIという新たな要素が加わることで、認識の相互作用がより複雑化し、それだけ誤解、誤認、誤算の可能性は高まる。AI化されたNC3の段階的な移行・発展によって新旧システムが自国・敵対国に混在することで、人間と機械の認識の相互作用を一層不安定化させかねない。

第五に、NC3やISRのデジタル化が一段と進むことで、サイバー攻撃に対する脆弱性が高まりかねない。偽情報の提供、通信の混乱化、コミュニケーション・チャネルの妨害・破壊などは、状況認識の正確性を低下させよう。また、サプライチェーンを通じて、あるいはその他の方法で、核兵器の有効性を損なう可能性のある方法で、核兵器に欠陥や悪意のあるコードを導入し、核兵器システムの誤作動や不正な制御をもたらす可能性もある。しかもそうしたサイバー攻撃は、武力紛争時だけでなく平時からなされ得るとすれば、グレーゾーン状況から核のエスカレーションの可能性を考えなければならず、核兵器使用の新たなリスクを生じさせている。

おわりに

新興技術が核抑止関係にいかなる含意をもたらすか、少なくとも現状では明確な方向性を示すことは難しい。ただ、戦略的安定性の低下や、意図せざる核兵器使用といった重大なリスクがあり、これをいかにして抑制するかが、核抑止問題における重要な課題の一つとなることは明らかである。そこでは、通常兵器と核兵器の規模と洗練度、技術導入の速度、地理的・地政学的緊張、技術的対称性・非対称性、戦略的関係の状況と成熟度など多様な要素が作用すると論じられている(Boulanin, et.al.)。そうした複雑な課題についてリスクを抑制・低減するためには、核保有国による新興技術の慎重な導入を期待するだけでなく、国・専門家が多角的な見地からこの問題を議論し、そこから信頼醸成、透明性、およびリスク低減のための措置、あるいは規範の構築などといった取り組みを、リスクが顕在化する前に導き出すことが求められる。

参考文献

James M. Acton, "Escalation through Entanglement: How the Vulnerability of Command-and-Control Systems Raises the Risks of an Inadvertent Nuclear War," *International Security*, Vol. 43, No. 1 (Summer 2018).
Vincent Boulanin, Lora Saalman, Petr Topychkanov, Fei Su and Moa Peldán Carlsson, *Artificial Intelligence, Strategic Stability and Nuclear Risk* (SIPRI, June 2020).

Mark Fitzpatrick, "Artificial Intelligence and Nuclear Command and Control," *Survival*, Vol. 61, No. 3 (June-July 2019).

Michael C. Horowitz, Paul Scharre, and Alexander Velez-Green, "A Stable Nuclear Future? The Impact of Autonomous Systems and Artificial Intelligence," Cornell University, December 2019.

James Johnson, "The AI-cyber nexus: implications for military escalation, deterrence and strategic stability," *Journal of Cyber Policy*, September 2019.

James Johnson, "Delegating strategic decision-making to machines: Dr. Strangelove Redux?" *Journal of Strategic Studies*, (2020).

Michael T. Klare, "'Skynet' Revisited: The Dangerous Allure of Nuclear Command Automation," *Arms Control Today*, Vol. 50, No. 3 (April 2020).

Keir A. Lieber and Daryl G. Press, "The New Era of Counterforce: Technological Change and the Future of Nuclear Deterrence," *International Security*, Vol. 41, No. 4 (Spring 2017).

Kelley M. Saylor, "Artificial Intelligence and National Security," CRS Report, R45178, November 21, 2019.

Yuna Huh Wong, et.al., *Deterrence in the Age of Thinking Machines* (RAND, 2020).