



## 研究レポート

# 新興技術の規制可能性：軍備管理の視点からの論点整理

2021-03-30

秋山信将（一橋大学大学院教授／日本国際問題研究所客員研究員）

### 「安全保障と新興技術」研究会 第7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。

### はじめに

安全保障における新興技術のインパクトは様々な面から評価されているが、軍備管理の分野でも新興技術をどう扱うかは大きな争点となりつつある。軍備管理は、主として対立する国家間の関係の安定を図るため、力の均衡などを考慮しつつ軍備の規制や制限を行うことによって、危機のエスカレーションや軍拡競争といった安全保障上のリスクを管理することを目的とする。また、同時に軍備管理の枠組みに参加する国々は常に軍事戦略上(あるいは時に政治戦略上)の競争優位の確立をめぐる競争を繰り広げるが、その競争の枠組みを設定する。

従来の米口間の核軍備管理は、防御兵器における軍拡を弾道弾迎撃ミサイル制限条約(ABM条約)で制限し、攻撃兵器の領域は、核弾頭の運搬手段を射程距離によって「戦略」、「中距離」と区切ったうえで、戦略レベルで相互確証破壊の状態にあることについて合意することにより、そこに「戦略的安定性」があることを相互に了解するという形をとっていた。本稿は、新興技術がこのような従来の軍備管理を通じた安定性のあり方にどのような影響を与えるのかを考察するうえで重要な論点を整理することを目的とする。軍備管理は安全保障政策の一環であり、本来であれば抑止態勢のあり方および国家安全保障戦略と一体的に考えるべきであるが、本稿では、軍備管理の領域に絞って論じることとする。

## 軍備管理における新興技術の影響を評価するうえでの論点

軍備管理における新興技術の影響を考察する場合、最大の論点は、軍備管理が提供する安定性の実効性の担保ということになろう。それは、第一に、戦略的関係において何をもち「安定」とみなすのか、その「安定」のフォーミュラを見出すことができるか、第二に、遵守の担保(検証制度におけるチートの可能性、遵守の報酬と不遵守に対する制裁)、すなわち軍備管理の提供する安定性が持続的であり得るかどうかである。

なお、軍備管理を通じた安定性は、「危機の安定性」と「軍拡競争の安定性」から成っており、そのいずれの観点からも検討する必要がある。また、これまで軍備管理を論じる際に使われてきた概念や用語について、新興技術がもたらす現象にそのまま援用が可能かどうかについても再検討が必要となろう。

現時点において評価は定まっていないものの、新興技術がキネティックな攻撃力の向上に貢献する部分に関しては、規制の対象(スコープ)の設定や規制の方法(検証を含む)の設定次第では従来の軍備管理・抑止概念の延長線上で論じることができそうである。(なお、軍備管理の枠組み設定の妥当性に関しては、すでに、新興技術以外の問題、すなわち大国間の戦略環境の変化、すなわち軍備管理の枠外にいる中国の影響力の拡大やロシアによる中距離ミサイル全廃条約(INF条約)違反による中距離ミサイル能力保有などによって大幅な減退が認められる。)

勿論、軍事理論上、安定性の解をどう見出すことができるのか、またその解について政治的に合意ができるのか、その見通しを論じるためには、これらの新興技術の能力についてさらなる評価を続ける必要があるだろう。それよりもおそらく不透明性が高く、また軍備管理のモダリティの解の中に取り込むことが難しいのは、AI、リモートセンシング、ビッグデータなどによってサポートされる意思決定の部分での影響である。

結論を先取りしていえば、軍備管理での新興技術の主たるインパクトは、主として既存の軍備管理(もしくは抑止関係)の中ですでに存在しているリスクを増幅させる効果であろう。たとえば、軍備管理におけるエスカレーション・ラダーの想定妥当性、「戦略的曖昧性」の計算の不確実性、相手の意図の不可知性、といった面で影響が大きいように見える。もし、こうした従来の概念に継続的な活用を困難にするような大きな変化を新興技術が起すのであれば、既存の軍備管理の枠組みや制度の設計思想の中でこれらの問題を捕捉することは困難であり、そのような技術の規制(ルール化)の行方は不透明なものとなる。

## 「危機の安定性」の側面への影響

キネティックな打撃力の面で見れば、例えば、極超音速滑空弾(HGV)は、その敏捷性と機動性・操縦性という特性は打撃力の向上につながりえる。このHGVが、大国間の戦略的安定を損なうとすれば、どのような状況だろうか。シェリングによれば、戦略的安定性は、双方の戦略的核戦力が先制的な対兵力攻撃に対して脆弱である場合にのみ危機の不安定性が急激に悪化し、「奇襲攻撃に対する相互の恐怖」が生じることで壊れるとされる。しかし、現時点において、このHGVによる中口の戦力投射能力は、戦域レベルの打撃力の変化にとどまり米国の戦略核の脆弱性を高めるには至っていないとされる。

ただし、攻撃のターゲットになりえるような通常戦力と核戦力の配備場所が数百キロほどしか離れていない場合、HGVの特性ゆえに、どちらが標的となっているかを明確に判別する前に反撃・報復の意思決定が求められることになる。北東アジアにおいては、米国にとっては、戦域レベルの紛争で、戦略的に死活的ではないとしても、中国にとっては戦略レベルの死活的利益がかかっているという非対称性が存在する。その場合、戦略レベルの核能力では劣位にある中国にとって、攻撃の対象が判明する前に反撃の意思決定を下す必要があるというのは、意図せざるエスカレーションのリスクが高い状況であるということになる。なお中口は通常戦力と核戦力を同一基地に配備している(co-location)とされるため、従来からこの問題は指摘されてきたが、HGVが考慮に入ることによって、数百キロの分離ではこのco-locationの問題への解決策とはなりえないということになる。

また、キネティックな打撃力ではないが、攻撃能力という観点からすれば、サイバー攻撃による核戦争時指揮統制通信(NC3)や早期警報システムへの攻撃は、危機の安定性におけるリスク要因となることは言うまでもない。

意思決定の面で見ると、AIにサポートされたインテリジェント・システムでは、モニタリング(戦場・現場の情報収集)→分析→最適解→命令→作戦実施という兵力運用のサイクルが加速化されることが想定される。このような意思決定までの時間の短縮に伴うリスクとしては、事態が急速に変化していく中で、エスカレーション・コントロールにおける判断への心理的な圧力が増す可能性である。とりわけ、能力で劣位にある側の心理としては、戦闘において敗北を受け入れるかいなか、すなわち「use-it-or-lose-it(核)を使うのか、敗北を受け入れるのか)」という究極の選択の圧力を受けやすくなる。

また、AIなど新興技術による状況把握能力の向上は、いわゆる「戦場の霧(Fog of War)」を晴らすことが期待されるが、一方で、意思決定をAIに依存するようになった場合に懸念されるのは、意思決定プロセスにおいて、意思決定者から、情報処理過程が不可視化、あるいは不可知化されることである。思考経路の不可視化・不可知化は、意思決定者が自らの決定に対する自信を欠き、その結果として一貫した合理的決定(と意思決定者が信じる決定)をできなくする恐れがある。もちろん、こうした心理状態は、AIその他の技術が介在しない状況においても起こりえるが、逆に言えば、技術によって解決が可能な問題ではないということになる。

そのほか、想定しうる「不安定性」のリスクとしては、AIやビッグデータ、リモートセンシング技術等を活用したサポート・システムの非対称性により生じるものがある。つまり、敵対する相手がAI等を活用したシステムによって状況把握や意思決定の支援を受けていると疑心暗鬼な状態に陥った場合、(時間的)劣勢に対する強迫観念から、AIサポート・システムの導入を急ぐあまり、不完全なアルゴリズムなど潜在的にバグを抱えたシステムを運用してしまうリスクや、相手の状況把握能力を過剰に見積もった場合の心理的圧力などが、判断を誤らせる可能性がある。

## 「軍拡競争の安定性」の側面への影響

軍拡競争の安定性への影響という側面でもいくつかの可能性が想定しうるが、一例をあげるとすると、情報収集・警戒監視・偵察(ISR)能力の向上が軍拡意欲に与えるインパクトである。核保有国同士の相互抑止の関係は必ずしも対称的な能力を持つ必要はなく、中国と米国の間では、米国は中国に対して公的には脆弱性を認めない優越の側にあるとし、他方中国は米国に対する一定程度の報復能力を保持することで十分とする最小限抑止を想定してきた。しかし、最小限抑止を採用する側の前提としては、相手のISR能力には限界があり、配備されたすべての核戦力(とりわけ移動型)を捕捉することは不可能で、それ故に報復の力を確保できるというものであった。しかし、ISRの能力が飛躍的に向上し、ほとんどの核戦力の所在が捕捉されるといったような事態になれば、最小限抑止は成立せず、より強靱な核戦力の構築に向かう、すなわち軍拡意欲を掻き立てる可能性もある。

## 軍備管理領域の概念の変化

軍備管理・抑止の領域における既存の概念や用語が、新しい技術環境下での「戦略的安定性」の議論で妥当性を持ちえるかという点も、検討が必要であろう。とりわけ、サイバー攻撃におけるエスカレーション、抑止、透明性、(サイバー軍備管理における)検証などは、物理的な打撃力を前提とした概念とは異なるであろう。あるいは無人兵器による戦闘においてもエスカレーション・ラダーを上るとしたらどのような状態になった場合が想定されるのか、また、サイバー攻撃に対する、あるいは無人兵器による攻撃に対する報復において、戦時国際法にある比例原則(proportionality)を適用するとして、例えば、重要インフラに対するサイバー攻撃による損害をどう見積もり、それに均衡する報復はサイバー攻撃以外の打撃力による報復によっても成立するかなどについて明確な考え方は存在しない。昨年、米軍の偵察ドローンをイランが撃墜したことに對し、米国がサイバー攻撃によって報復したと発表した。果たしてこれが比例原則に照らして妥当なのか、またこれが先例となって比例原則の考えが方向づけられるのか不明であるが、安定的な戦略関係の構築のためには今後国際的な議論を深めていく必要がある。

これらは、軍備管理の枠組みが提供する、戦略関係における安定性のための方程式を確立するうえで必要な概念の整理ということになる。

## まとめに代えて

新興技術の戦略関係への影響は、新START条約の後継条約の議論の中でも、ミサイル防衛と並び重要な論点となってくるであろう。とりわけ、戦略核と非核戦略兵器、戦略レベルと非戦略レベル、そして新ドメイン(サイバー)の境界線があいまいになる作用が新興技術にはあり、この特徴は新条約が規制の対象に含める兵器体系、そして規制の枠組みの提供する安定性への信頼度を規定することになるであろう。後継条約のありそうな形としては、A)HGV、ミサイル防衛など非核能力も含め、包括的に戦略兵器を規制する条約、B)従来の戦略核+INFレンジのミサイルで条約本体を構成し、HGVを含む非核戦略兵器とミサイル防衛は追加議定書などで改めて規定する、C)従来通り兵器体系を区分し、個別のカテゴリーごとに条約を策定、という3つの方法が考えられる。

これらのシナリオを合意の容易さで並べれば、C)、B)、A)という順番になるであろうが、戦略的安定性の実質を担保する実効性の順はその逆となろう。軍備管理とは、安全保障政策の一環ではあるが、同時に関係国間の政治的な関係を規定する枠組みでもあり、それは、合意内容については極めて政治的な判断になることを意味する。とすれば、問題は、米中ロ三か国の間に、合意に向けて実質の伴った協議をしていく政治的な意思があるのかどうか、という身もふたもない話になってしまう。

加えて、AIの活用をどこまで認めるかなど、NC3への新興技術の実装が意思決定に与える影響や、NC3の脆弱性の戦略的安定性に対する重要性の評価が関係国の間で合意されれば、そこも軍備管理の枠組み内で対処すべき項目になるかもしれない。ただ、サイバーやAIの実装については、実効性を持つ(あるいは検証可能な)規制の方法を見出す困難さ、透明性確保の困難さを考えると、軍備管理条約の一部としてこれらを考慮することは少なくとも現時点においては極めて難しそうである。従って、今できることは、各国のサイバー攻撃のドクトリンの情報交換や報復における比例原則等に関するお互いの考え方を交換する等の信頼醸成や、AIの導入については、例えば、導入にあたっての行動規範や、意思決定への関与における人間の思考能力(速度)を超越しないように規制をかけるセーフガードのメカニズムの導入は可能か、といった角度から、双方の懸念に関し相互理解を深めるところから始めることが必要であろう。

## 【参考文献】

Nobumasa Akiyama, "AI Nuclear Winter or AI that Saves Humanity? AI and Nuclear Deterrence," Joachim von Braun et al., eds., *Robotics, AI, and Humanity: Science, Ethics, and Policy*, Cham Switzerland: Springer, 2021, pp. 161-170.

Andrew F. Krepinevich Jr., "The Eroding Balance of Terror," *Foreign Affairs*, January/February 2019, <https://www.foreignaffairs.com/articles/2018-12-11/eroding-balance-terror>.

Rafael Loss and Joseph Johnson, "Will Artificial Intelligence Imperil Nuclear Deterrence?" *War on the Rocks*, September 19, 2019, <https://warontherocks.com/2019/09/will-artificial-intelligence-imperil-nuclear-deterrence/>.



Robert Peters, Justin Anderson and Harrison Menke, "Deterrence in the 21st Century: Integrating Nuclear and Conventional Force," *Strategic Studies Quarterly*, Vol. 12, No. 4(Winter 2018), pp. 15-43.

T. C. Schelling, *The Reciprocal Fear of Surprise Attack*, P-1342, Santa Monica, CA: RAND, 1958.