在人工智能（AI）时代如何守护日本的主权——将日语的特殊性转化为“武器”，发展日本独自的人工智能

超越“肌肉”：在 AI 时代，网络空间与情报体系本身就是主权国家的“头脑”和“神经系统”。
AI generated by Gemini

北村滋（前国家安全保障局长）

直到冷战时期，国家安全保障的重心主要置于军事力量与外交力量这两大支柱之上。以核武器为代表的武器体系的数量与质量，以及同盟关系和经济合作网络的广度，无疑一直是衡量国家生存能力与国际影响力的重要指标。隔着“铁幕”相互对峙的美苏两大阵营，正是在不断扩张核力量与同盟体系的竞争之中，在这种脆弱的势均力敌的平衡上勉强维持了和平。

然而，21 世纪已经过去四分之一的时间，国家实力的构成正在发生明显变化。决定国家实力的要素，已经不再局限于传统意义上的军事力量或可直观观察的外交活动范围，而是扩展到数据与网络的安全性、关键产业供应链的韧性，以及与本国或盟国合作掌握包括人工智能（AI）在内的先进技术的能力。

换言之，如果说护卫舰、战斗机和坦克是国家的“肌肉”，那么与之相对应的“神经系统”和“头脑”便是网络空间与 AI。无论肌肉多么强健，如果神经系统被切断、判断功能陷入瘫痪，国家也将无法行动。当代安全保障的核心，正是在于如何守护并重塑这一“神经系统”与“头脑”。

正如明治时期的日本以富国强兵与文明开化为双轮[1]，同时整备军事力量与官僚制度一样，对于21世纪的日本而言，如果不锻造网络与AI这一新的“神经系统”与“头脑”，就无法谈论作为主权国家应有的形态。

国家“神经系统”的脆弱性

在俄罗斯入侵乌克兰的过程中，一个具有象征意义的现象是：在军事入侵正式开始之前，针对乌克兰发电站和政府机构的网络攻击已经连续发生，通信与行政服务因此受到严重影响。在入侵军越过国境之前，对方国家的行政手续已出现迟滞，民众也难以获得准确的信息。针对电网的恶意软件攻击导致一度停电，铁路和金融系统也受到波及。与此同时，对卫星通信系统的干扰不仅扰乱了战场上的指挥体系，也影响到了民间企业与普通市民。在战争状态下，“混乱”已经不再只通过轰炸的闪光呈现出来，而同样可能以屏幕冻结与密集错误代码的形式出现。

2017年席卷全球的勒索软件“NotPetya”，通过乌克兰企业使用的一款会计软件更新机制传播，最终导致欧美港口运营公司、物流企业和制药企业的系统瘫痪。这一事件象征性地表明：网络攻击可以跨越国境，通过供应链使整个世界经济循环陷入停滞。

同样在2017年，勒索软件“WannaCry”重创英国国家医疗服务体系（NHS），导致医院预约系统发生故障、诊疗记录无法查阅，一些医院甚至被迫暂停接收急诊患者。这表明网络攻击已经不再只是企业经营损失的问题，而可能直接危及人的生命。在病房里挤满患者的背景下，看不见的攻击者却在束缚医疗现场的判断能力——这种异样的现实，我们必须正视。

鉴于这些情况，日本尝试将单个医院纳入《经济安全保障推进法》所规定的特定社会基础设施运营者范围，可以说是一项重要进展[2]。

美国也接连发生重大事件。2021年的“Colonial Pipeline”攻击导致美国东海岸的汽油供应网络部分停摆，人们涌向加油站的画面迅速传遍世界。虽然实施攻击的是犯罪组织，但其技术能力与目标选择方式，与国家级战术有着惊人的相似性。

2020年的“SolarWinds事件”中，植入该公司IT管理软件Orion更新程序中的恶意代码，悄然侵入美国国土安全部、财政部以及盟国政府机构和企业系统。系统后台的管理软件，仿佛成为外部势力的隐秘通道。

---

[1] 富国强兵是明治政府提出的国家口号。19世纪后半叶，在亚洲诸国相继沦为欧美列强殖民地的背景下，日本为了维持自身独立，迫切需要获得“能够与西方匹敌的经济实力与军事力量”，这一口号正是这种国家战略的体现。另一方面，文明开化则是一场积极吸收西方文明、推动社会制度与生活方式全面革新的文化运动。它并非单纯的社会风潮，而是通过引入自由、平等、科学教育等现代价值观，使日本在国际社会中被承认为“近代国家”的重要过程。

[2] 特定重要基础设施运营者：根据《经济安全保障推进法》（ESPA），日本政府在电力、燃气、通信、金融等15个领域指定“特定重要基础设施运营者”。这些企业在安装关键设备时必须接受政府的事前审查，以防范网络攻击及供应链风险。

围绕石油的攻防也延伸到了网络空间。2012 年，沙特阿拉伯国有石油公司沙特阿美遭到名为“Shamoon”的擦除型恶意软件攻击（即通过彻底删除硬盘数据使其无法恢复的病毒），超过三万台计算机遭到毁灭性破坏。据称，攻击还将企业标志替换为骷髅图像并删除硬盘内容。自 1990 年海湾战争以来，围绕石油权益的斗争，已经从战场上的轰炸转移到了恶意代码的世界。

朝鲜则事实上把网络空间作为一种获取外汇的手段。从令人联想到日本 TBS 电视剧《VIVANT》[3]（2023）开场情节的孟加拉国中央银行巨额非法转账事件，到针对加密资产交易所的黑客攻击导致资金流失，都是典型案例。2014 年，美国电影公司索尼影业也因一部电影而遭到大规模网络攻击，导致信息泄露和业务中断。

在国际社会通过制裁限制金融交易框架的情况下，一些国家利用匿名性极高的数字空间开展国家行为。无论古今中外，“海盗行为”被部分国家默许并成为充实国库的手段的现象并不罕见。

中国针对台湾的网络行动则更进一步。除了入侵台湾政府网站之外，还通过日常基础设施——如便利店显示屏或铁路信息系统——显示政治信息。这种方式将心理战与信息战结合在一起，形成一种新的攻击模式。

2007 年，爱沙尼亚在迁移苏联战争纪念碑时，遭遇大规模网络攻击，政府网站、银行系统和媒体一度瘫痪。被认为有俄罗斯黑客参与的这一事件，被视为围绕历史认知问题的外交争议在网络空间演变为攻击行为的先例。

台湾和波罗的海国家的经验表明，网络攻击正在演变为直接影响舆论的手段。而随着 AI 技术的发展，攻击者可以以前所未有的速度识别系统漏洞，这也使局势变得更加严峻。

“末端”才是左右安全保障的关键

这种国际趋势，对日本而言绝非隔岸观火。针对朝日集团控股公司（Asahi Group Holdings）、ASKUL 等企业的勒索软件攻击已经表明，一旦企业活动中断，经济与社会功能的瘫痪便会随之而来。

针对防卫相关企业、半导体制造商、大学与研究机构、通信运营商等机构的定向攻击持续不断，一些地方政府的信息系统也曾遭到攻击，导致行政服务暂停。地方政府职员日常工作所使用的终端设备，以及外包的系统运维结构，往往成为攻击的入口。这种结构清楚地说明：日本信息管理体系的“末端”反而可能决定国家整体安全。

[3] TBS 电视台周日剧场《VIVANT》（2023 年播出）：该剧讲述一名卷入巨额误汇款事件的商社职员，实际上是隶属于自卫队秘密情报组织“别班”的成员，并在调查国际恐怖组织之谜的过程中展开一系列行动的故事。

“多样化的主体支撑着国家功能，而这些主体同时也可能成为攻击国家的入口”——这一结构已经成为国际社会的“常识”。

第二次世界大战时期，攻击敌国军需工厂和港口、切断其兵站补给是一种典型战略。而在当代，这一战略形态已经发生变化：攻击目标不再局限于军事设施，而是延伸至构成战略物资供应链一环的中小企业服务器、云端业务系统，以及支撑地方政府窗口服务的网络设备等日常基础设施。

基于这种结构变化，日本于 2022 年制定了《经济安全保障推进法》，建立了重要物资稳定供应、关键基础设施服务稳定提供、以及先进关键技术开发支持等制度框架[4]。不仅半导体、蓄电池、关键矿产等实体资源，通信基础设施、云服务等“看不见的基础设施”也成为安全保障的重要关注对象。

进一步地，2025 年日本完善了“主动网络防御法制”，使国家从过去以事故发生后的应对为前提的被动防御，转向能够在攻击前阶段进行探测、阻止与无害化处理的主动应对。同时，新设立了负责协调各省厅权限与责任的国家网络安全统括室，从而在国家层面建立了综合协调与指挥控制的基础。

这些措施背后的基本认识是：网络空间已经不再只是技术部门处理的专业领域，而是关系国家决策基础本身的安全保障问题。

冷战时期，围绕核威慑的讨论曾关注“如何防止因错误认知而导致的升级”。而在今天，一个新的问题被摆到台面上：当信息系统陷入瘫痪时，国家还能在多大程度上保持正常判断能力？这种认识转变，成为日本安全保障政策的重要转折点，也将成为未来政策发展的前提。

当然，对国家中枢的威胁并不止于网络攻击。围绕 AI 的讨论，也已经超越了单纯的产业竞争或生产效率问题，深化为一个根本性课题：如何确保国家的信息主权。

AI 之间的战争

在生成式 AI 被逐步引入政策制定、行政事务、危机管理、金融监管乃至国防领域的当下，国家已经无法再把 AI 视为“方便的工具”。一些 AI 模型已经开始被用于起草行政文件、制作法律草案初稿、检索判例以及整理审查标准。然而，只要其判断过程仍是黑箱，责任归属便不可避免地变得模糊。

欧盟通过《AI 法案》对高风险领域的 AI 使用实施严格监管，同时法国正在推动 Mistral，德国则推进 Aleph Alpha 等本国主导的基础模型。

在英国，AI 被用于医疗影像诊断，以提高癌症早期发现率；与此同时，关于诊断错误责任归属的问题也引发了广泛讨论。

[4] ESPA（2022 年）确立了四项制度支柱：1）强化重要物资供应链；2）保障关键基础设施安全；3）支持先进关键技术研发；4）保护敏感专利（专利非公开制度）

爱沙尼亚作为电子政府的先行者，已经将纳税申报、公司注册乃至法院记录管理全面数字化，并尝试通过 AI 进一步实现行政服务自动化。在那里，作为国家“手脚”的行政事务，正逐渐被算法所取代。

在亚洲，韩国将 NAVER 的大规模语言模型 HyperCLOVA X 引入公共部门，用于行政文件撰写和居民咨询的初步回应。

新加坡在“智慧国家”构想下，将 AI 应用于城市管理、交通控制与医疗服务，在提升市民生活便利性的同时，也面临如何监管背后海量个人数据的问题。

中国则通过监控摄像网络、面部识别技术与大数据分析构建社会信用评分体系，并将其用于治安维护与行政管理。

这些例子表明：AI 不仅是治理工具，也正在成为重新定义“主权”的基础技术。

在美国，国防部的“Maven 计划（Project Maven）”具有象征意义。AI 通过分析无人机与卫星传回的海量影像数据，协助识别目标并辅助战场态势判断。随着这类系统不断发展，战场上最关键的判断——敌我识别——将越来越依赖算法。

在乌克兰战场上，据称 AI 已经被用于整合无人机影像和战场信息，并辅助确定炮兵射击优先级。AI 已不再只是军官的辅助参谋，未来甚至可能逐渐承担部分“决策者”角色，这其中蕴含着潜在风险。

问题也已经延伸到普通民众的日常生活。目前，日本钓鱼诈骗和恶意重定向攻击正急剧增加。日本网络犯罪对策中心和消费者厅已经发现大量伪装成正规品牌的假购物网站。此外，日本信息处理推进机构也警告称，一些网站通过伪装“下一页”等普通按钮，将用户引导至仿造正规界面的恶意页面。

普通用户几乎难以区分这些链接的真伪，而 AI 的使用正使这种伪装界面以惊人的速度增加。乍看之下，这似乎只是网络犯罪问题，但同样的技术手段也可以用于窃取高度敏感信息，进而引发涉及国家安全的事件。

当然，如果 AI 被正确使用，许多简单的编码错误有可能大幅减少。由代码错误引发的网络漏洞以及由此产生的犯罪问题，也有可能通过 AI 技术得到改善。

另一方面，攻击者也在利用 AI 迅速提升发现系统漏洞的能力。因此，几乎所有信息系统和应用程序的开发，都逐渐演变为 AI 与 AI 之间的战场。

例如，两年前 CrowdStrike 公司事件中，该公司软件中的漏洞在 Windows 操作系统深层触发，导致航空公司系统、银行系统等众多社会基础设施受到影响，服务一度中断。这一事件再次说明：轻易开放操作系统核心层依然存在巨大风险。

在这样的背景下，全球各国对“主权 AI”这一概念的关注迅速上升。所谓主权 AI，是指根据特定国家或地区的法律与监管框架开发和运营的 AI 系统。

正如前文所述，法国、德国、韩国正在推进本国主导的基础模型，以减少对美国平台的过度依赖。印度也在战略性地培育能够适应多语言环境的 AI 基础体系。在中东地区，则正在开发专门针对阿拉伯语的 AI 模型，以构建对文化与宗教语境更加敏感的对话系统。

各国都在逐渐形成共识：能够准确理解本国语言、文化与法律制度的AI，正是国家主权的核心所在。

对日本而言，最优选择是什么？

那么，日本应当选择怎样的发展方向？

日本在安全保障与技术两方面都以日美同盟为基础，而美国云计算与 AI 企业所拥有的技术优势也极为明显。在这种现实条件下，如果无视这一格局，试图像欧洲部分国家那样追求“完全自立型”的AI主权，显然并不现实。

但另一方面，如果国家核心功能长期完全依赖外部主体，也会带来主权国家难以接受的风险。冷战时期，各国围绕通信电缆与卫星通信线路的控制权展开激烈竞争；而在当今时代，这一角色已经被云平台与AI基础设施的控制权所取代。

在这一前提下，日本应当采取的方向其实十分清晰。

第一，在以日美同盟为前提的协同框架内，确保日本自身的政策裁量空间。在这一过程中，不仅要考虑国家层面，也必须兼顾产业界。换言之，在日本或美国具有技术优势的领域，不应轻易迫使日本企业或盟国美国企业过度开放技术与知识产权。否则就可能为潜在竞争国家提供可乘之机，最终受损的不仅是日本与美国企业，也包括国家安全本身。例如《智能手机软件竞争促进法》便是一个典型案例，应当以安全为最优先原则，采取审慎而克制的执行方式。

第二，在行政、司法、医疗、教育、危机管理等直接关系国家判断能力的领域，至少应建立具有一定自主性的技术基础设施。

第三，聚焦日本长期积累优势的领域——如日语环境、高龄化社会、自然灾害应对、机器人技术、材料科学与危机管理——通过“选择与集中”的战略，培育具有高度不可替代性的主权技术。而在这些领域中，日语这一语言环境的特殊性尤为关键。

将日语的特殊性转化为战略优势

日语的意义理解高度依赖语境，并大量使用敬语和委婉表达。因此，如果只是简单套用通用的国际模型，在行政文件、法律制度、医疗记录等对精度与责任要求极高的领域，很难保证足够的可靠性。能够准确理解并处理日语行政文书、司法记录和医疗信息的AI，本身就是支撑主权国家决策能力的“主权技术”。如果这一关键能力完全依赖外国商业模型，从长期来看，等同于将国家判断能力的一部分外包给外部主体。

同样地，在机器人技术、材料科学、医学研究、危机管理、卫星与地理空间信息等日本长期保持国际竞争力的专业领域推进AI开发，也具有同样的重要意义。关键并不在于规模或算力的绝对优势，而在于通过不可替代的专业优势，为国家争取更大的战略自主空间。

战后日本曾在汽车、家电、半导体等产业领域，通过在细分领域做到“即使是小众领域也做到世界第一”，从而提升了在国际谈判中的话语权。AI 时代的主权，也应当通过在特定领域形成不可替代的存在感来加以支撑。

然而，无论制度与技术如何完善，仅凭这些仍不足以完成国家安全保障。国家运作的主体始终是人，而人的脆弱性往往是最大的风险来源。与外部网络攻击或情报窃取相比，内部人员的信息泄露、不当访问、长期使用陈旧系统、以及对信息管理缺乏意识等问题，往往带来更为严重的损害。国际社会的多起案例已经反复证明这一点。例如，斯诺登事件中美国国家安全局（NSA）的大量机密信息外泄；美国军方机密文件被带出至个人终端；以及欧洲多国国防相关信息的外泄事件等，都表明内部安全漏洞往往能够从根本上动摇国家安全体系。

为此，日本于 2024 年制定了《安全审查制度法》，并与此前的《特定秘密保护法》相结合，建立了对国家机密与敏感信息访问管理的统一框架。该法律不仅适用于政府机构，也将民间企业纳入适用范围，明确国家需要保护的信息与技术，并为相关访问权限提供法律依据与透明机制。通过这些制度，日本在人员安全管理方面已建立起一定的结构性基础。在国家与企业界限日益模糊的经济安全保障时代，仅将信息保护局限于国家机构内部已不可能，包含民间部门在内的“全社会型安全保障”将成为不可或缺的前提。

“数字地缘政治”的最前线

如果把视野扩大到东南亚，就会发现：中美双方数字基础设施并存的结构，正逐渐逼近其极限。 过去，该地区许多国家通过如下方式维持某种平衡：在通信和监控系统方面采用中国设备，在云服务和金融服务方面利用美国企业的技术。然而，如今云计算与 AI 已经进入足以直接规定国家功能本身的阶段，若仍认为这种“两边下注”的格局能够长期维持，未免过于乐观。国家数据究竟应放在哪个国家的基础设施之上？一旦发生危机，哪一方的云平台能够持续可靠运行？出现故障时又该由谁承担责任？ 这些问题无一不触及国家主权的核心。

新加坡正将涉及国家核心功能的信息集中到政府主管的云平台之上；印尼则在与美国企业合作建设国家数据中心的同时，转向更加重视物理基础设施安全保障的方向。越南和菲律宾也在推进对中国制造基础设施依赖度以及云服务合同的重新审视。

如今，东南亚正逐渐成为数字领域地缘政治选择的最前线。正如冷战时期的中东曾是围绕石油展开地缘政治竞争的焦点，21 世纪的东南亚则正逐渐成为围绕数字、数据与 AI 展开的新地缘政治竞争焦点。

当然，把东南亚各国理解为被迫在中美之间作出非此即彼的全面倾斜，也并不准确。更现实的战略应当是：在涉及国家安全核心的中枢领域，向值得信赖的伙伴——也就是同盟国或准同盟国——靠拢；而在商业领域，则保持多样性，并推动本国技术能力的提升。 这种兼顾不同层级的战略，才真正具有可行性。

与此同时，在扩大对特定“关切国家”的监管时，也必须充分注意，不能因此不当地限制盟国企业的活动，结果反而削弱本国的产业基础。如何把握真正必要且有效的规制方式，也将成为今后日本外交、安全保障与经济政策中的一个重要课题。

以上，我们考察了网络安全、AI、人员安全保障、地区秩序等看似不同的领域，但贯穿其中的核心问题其实最终汇聚为一个：国家如何维持自身的判断能力与行动自由。

“主权”，就是“决策的自主性”

网络攻击会从外部夺取主导权，对 AI 基础设施的过度依赖会使决策过程变得不透明，内部不正当行为、信息泄露以及软件开发中的简单失误，会从内部侵蚀包括产业基础在内的国家根基，而地区秩序的变化，又会加大外部力量对国家在数字服务基础设施选择上的压力。面对这种复合型危机，国家不能停留在零碎、被动、逐次应对的层面，而必须从多维视角出发，构建一体化的韧性体系。

2025 年的日本，通过实施安全审查制度法、完善主动网络防御法制、设立国家网络安全统括室、并推进及修订经济安全保障推进法等一系列举措，稳步整备了可谓国家安全保障“神经系统”的基础。然而，真正为制度与组织这些“器皿”注入生命的，是国家意志与社会共识。经济安全保障并不只是产业政策的延长线，而应被定位为：在同盟关系的基础上，为将国家决策的自由与责任传递给下一代而展开的一项“主权再设计”工程。

在一个不确定性常态化的时代，“主权”的核心就在于决策的自主性。如何守护国家的判断基础，又如何重新锻造这一基础——对这一课题的回应，将决定今后日本的前进方向。

经作者及出版社许可，转载自《文艺春秋》2026 年 2 月号，第 126—135 页。

北村滋（前国家安全保障局长）

北村 Economic Security 合同会社代表。1956 年生于东京。毕业于东京大学法学部。1980 年进入日本警察厅。曾赴法国国立行政学院（ENA）留学，历任警察厅警备局警备课长、外事课长、内阁总理大臣秘书官（第一次安倍内阁）、警察厅长官官房总括审议官等职。2011 年在野田内阁中出任内阁情报官；在第二次至第四次安倍内阁中继续留任该职。2019 年就任国家安全保障局长，菅义伟内阁时期亦继续留任，至 2021 年 7 月卸任。