

DAVLAT TASHKILOTLARIDA GENERATIV SUN'IY INTELLEKTDAN FOYDALANISHDA KIBERXAVFSIZLIKNI TA'MINLASH: XORIJ TAJRIBASI VA ZAMONAVIY YONDASHUVLAR

Abduraxmanov Zafar Batirovich

O'zbekiston Respublikasi Iqtisodiyot va moliya vazirligi huzuridagi

Bojxona qo'mitasining Bojxona instituti katta o'qituvchisi

Annotatsiya. Maqolada davlat tashkilotlarida generativ sun'iy intellektdan foydalanish bilan bog'liq kiberxavfsizlik muammolari tahlil qilinadi. AQSH, Buyuk Britaniya, Yevropa Ittifoqi va Singapurning rasmiy qo'llanmalari hamda amaliy yechimlari qiyoslanadi. Ma'lumotlar sizishi, zararli ko'rsatmalar, bilimlar bazasi zaifliklari va agentlarga ortiqcha vakolat berish xavflari tizimlashtiriladi. Tahlil asosida ma'lumot sezgirligi, bajariladigan amal oqibati va vakolat chegaralarini birgalikda hisobga oluvchi ko'p qatlamli himoya modeli taklif etiladi.

Kalit so'zlar: generativ sun'iy intellekt, davlat tashkiloti, kiberxavfsizlik, prompt injection, ma'lumotlar maxfiylik, RAG, minimal vakolat, inson nazorati, audit.

KIRISH

Generativ sun'iy intellekt (GenSI) davlat tashkilotlarida hujjat loyihalarini tayyorlash, katta hajmdagi matnlarni umumlashtirish, axborot izlash va murojaatlarga dastlabki javob shakllantirish uchun qo'llanishi mumkin. OECDning 2025-yilgi hisobotida davlat sektorida SI imkoniyatlaridan foydalanish ma'lumotlar boshqaruvi, xodimlar malakasi, infratuzilma va javobgarlik mexanizmlariga bog'liqligi ko'rsatiladi. Ushbu hisobot umumiy SI amaliyotini qamrab oladi; undagi barcha misollarni generativ SIga tenglashtirish to'g'ri emas [1].

Davlat tashkiloti uchun GenSI xavfsizligi faqat modelning to'g'ri javob berishi bilan belgilanmaydi. Xizmat hujjatlarining ruxsatsiz oshkor etilishi, axborot tizimlariga noto'g'ri buyruqlar uzatilishi yoki model natijasining tekshiruvsiz rasmiy jarayonga

kiritilishi ham hisobga olinishi kerak. NIST yondashuvida SI xatarlari butun hayotiy sikl davomida boshqariladi; generativ SI profili esa axborot xavfsizligi, maxfiylik va ishonchli ko‘ringan noto‘g‘ri javoblarni alohida ko‘rib chiqadi [2, 3].

Maqolaning maqsadi - xorijiy qo‘llanma va amaliyotlarni qiyoslash orqali davlat tashkilotlarida GenSIDan foydalanish uchun tekshiriladigan himoya talablarini tizimlashtirishdir. Asosiy savol quyidagicha qo‘yiladi: xodim, hujjat, model va tashqi axborot tizimi o‘rtasidagi aloqada xavfsizlik nazorati qayerda va qanday amalga oshirilishi lozim? Mualliflik taklifi mavjud himoya vositalarini yangi nom bilan takrorlashga emas, ularni vazifa, vakolat va audit dalillari bilan bog‘lashga qaratilgan.

MATERIALLAR VA USULLAR

Maqola maqsadli manba tanlash, hujjatlar kontent-tahlili, qiyosiy tahlil va konseptual modellashtirishga asoslangan. 2023–2026-yillarda e‘lon qilingan rasmiy hisobotlar, texnik qo‘llanmalar va ilmiy ishlar, shuningdek, 2026-yil sentabr oyi holatida tekshirilgan rasmiy xizmat sahifasi qo‘llanildi. Tahlil korpusi adabiyotlar ro‘yxatidagi 15 manbadan iborat. Manbalar bevosita GenSI xavfsizligiga yoki davlat sektoridagi qo‘llanish sharoitiga aloqadorligi, muallifi aniqligi va ochiq tekshirilishi mumkinligi asosida tanlandi.

Qiyoslashda besh mezon qo‘llanildi: ma‘lumotlarni himoyalash, kirish va amal vakolatlari, hayotiy sikl nazorati, inson javobgarligi hamda tekshiruv dalillarining mavjudligi. Rasmiy tavsiya, ilmiy ishda kuzatilgan xavf va amalda ishlayotgan xizmat haqidagi ma‘lumot bir-biridan ajratildi. Ushbu ish to‘liq tizimli sharh yoki meta-tahlil emas; uning natijalari tanlangan manbalar doirasidagi analitik umumlashtirish va konseptual takliflardan iborat. Empirik sinov o‘tkazilmagan.

NATIJALAR

Generativ SIga xos xavflar va ishonch chegaralari

OWASPning 2025-yilgi ro‘yxati ko‘rsatmalar orqali manipulyatsiya, sezgir axborotning oshkor bo‘lishi, ta‘minot zanjiri zaifliklari, natijani xavfsiz qayta ishlamaslik va ortiqcha agent vakolatlarini alohida xavf sifatida belgilaydi [4]. Bunda prompt injection

- modelga beriladigan yoki model o'qiydigan matn orqali uning belgilangan vazifasini o'zgartirishga urinishdir. Greshake va hammualliflar tashqi hujjat yoki veb-sahifaga joylashtirilgan ko'rsatma foydalanuvchining bevosita buyrug'isiz ham model xatti-harakatiga ta'sir ko'rsatishi mumkinligini namoyish etgan [5].

Modelga elektron pochta, fayl yoki ma'lumotlar bazasi bilan ishlash imkoniyati berilganda muammo javob matni bilan cheklanmaydi. Zhan va hammualliflarning InjecAgent ishida tashqi kontent orqali vositalarga ulangan agentlarni noto'g'ri harakatga yo'naltirish va ma'lumot chiqarishga undash xavflari o'rganilgan [6]. Bu natijalar barcha modellar uchun yagona xavf foizini bermaydi, biroq model ishlab chiqqan amalni alohida vakolat tekshiruvidan o'tkazish zarurligini asoslaydi.

1-jadval. Davlat tashkilotlaridagi GenSIning xavflari va nazorat yo'nalishlari

Xavf	Ehtimoliy oqibat	Nazorat yo'nalishi
Ma'lumot sizishi	Xizmat ma'lumotining oshkor bo'lishi	Ma'lumotni minimallashtirish; DLP; tasdiqlangan muhit
Bilvosita prompt injection	Hujjatdagi ko'rsatmaning buyruq sifatida talqin qilinishi	Tashqi kontentni ajratish; modeldan mustaqil vakolat nazorati
Ta'minot zanjiri buzilishi	Model yoki komponentning almashtirilishi	Kelib chiqish, versiya va yaxlitlik tekshiruvi
RAGda ruxsat buzilishi	Boshqa bo'lim hujjatining javobga qo'shilishi	Qidiruvdan oldingi hujjat darajasidagi ruxsat nazorati
Ortiqcha agent vakolati	Ruxsatsiz yozish, o'chirish yoki yuborish	Minimal vakolat; amalni tasdiqlash
Nazorat qilinmagan sarf	Xizmat uzilishi yoki xarajat oshishi	So'rov, vaqt va resurs limitlari

Ma'lumotlar sizishini faqat foydalanuvchi yuborgan prompt bilan cheklab bo'lmaydi: yuklangan fayl, qidiruv orqali olingan parcha va audit jurnali ham axborot oqimining qismidir. Model bergan noto'g'ri fakt esa har doim kiberhujum natijasi emas; u ishonchlilik muammosi ham bo'lishi mumkin. Shu sababli zararli aralashuvni aniqlash va javobning faktik to'g'riligini tekshirish bir-birini to'ldiradi, lekin bir xil vazifa hisoblanmaydi [3, 7].

Xorij tajribasining qiyosiy tahlili

AQSH: havflarni boshqarish va generativ SI profili. NIST AI RMF 1.0 to'rtta o'zaro bog'liq funktsiyani belgilaydi: GOVERN - boshqaruv va javobgarlik, MAP - kontekst va xatarlarni aniqlash, MEASURE - baholash, MANAGE - xatarlarni boshqarish. Bular bir marta bajariladigan ketma-ket tekshiruv emas, takrorlanadigan jarayondir. Framework ixtiyoriy foydalanish uchun ishlab chiqilgan; uni barcha davlat idoralariga majburiy sertifikat sifatida talqin qilish mumkin emas [2]. Generativ SI profili ushbu yondashuvni model natijalarining ishonchliligi, axborot xavfsizligi va uchinchi tomon bilan bog'liq xavflarga moslashtiradi [3]. Mazkur tajribadan olinadigan analitik xulosa - GenSI xizmatining egasi, ruxsat etilgan vazifasi va qoldiq xavfi hujjatlashtirilishi kerak.

Buyuk Britaniya: davlat xizmatchisi va tizim ishlab chiquvchi uchun alohida talablar. 2025-yilgi AI Playbook davlat tashkilotlarida SI vositalarini tanlash, xarid qilish va qo'llash masalalarini yoritadi. Unda shaxsiy ma'lumotlar, xavfsizlik, inson nazorati va tashkiliy javobgarlik birgalikda ko'rib chiqiladi [8]. NCSC va xalqaro hamkorlarning 2023-yilgi qo'llanmasi esa xavfsiz loyihalash, ishlab chiqish, joriy etish hamda ekspluatatsiya va texnik xizmat ko'rsatishni ajratadi [7]. Ikki hujjatni qiyoslash xodimning xavfsiz foydalanish qoidasi bilan dasturchining texnik himoya talabi bir-birini almashtirmasligini ko'rsatadi. Birinchisi foydalanish tartibini, ikkinchisi tizim arxitekturasini va hayotiy siklini qamrab oladi.

Yevropa Ittifoqi: umumiy, SIga xos va sohaviy himoya. ENISAning 2023-yilgi modeli uch qatlamdan iborat: bazaviy kiberxavfsizlik, SIga xos xavfsizlik va sohaviy xavfsizlik [9]. Bundan kelib chiqadigan texnik xulosa shuki, GenSI uchun maxsus filtr o'rnatilishi odatiy autentifikatsiya, yangilash va tarmoq himoyasi o'rnini bosmaydi.

Sohaviy qatlam esa xizmat jarayonining oqibatlarini hisobga olishga imkon beradi. ENISAning 2026-yil 7-iyuldagi hisobotida tezkor SI tahdidlariga qarshi operatsion salohiyatni rivojlantirish bo'yicha dastlabki tavsiyalar berilgan; tashkilot ularni to'liq va yakuniy tekshiruv ro'yxati deb ko'rsatmaydi [10]. Ushbu manbalar barcha YI tashkilotlarida bir xil tizim joriy qilinganini emas, texnik-uslubiy yondashuvlarni ifodalaydi.

Singapur: hayotiy sikl himoyasi, agent nazorati va amaliy xizmat. CSAning 2024-yilgi qo'llanmasi xavfsizlikni tizimni loyihalashdan boshlab hisobga olishni hamda an'anaviy va SIga xos tahdidlarni birgalikda boshqarishni ko'zda tutadi [11]. 2026-yil yanvarida taqdim etilgan agentli SI bo'yicha IMDA modeli tavsifida agent vakolatini chegaralash, muhim nuqtalarda inson tasdig'i, texnik nazoratlar va foydalanuvchi mas'uliyati ajratilgan [12]. Amaliy misol sifatida Pair davlat xizmatchilari uchun yozish, tahlil va boshqa xizmat vazifalarini qo'llab-quvvatlaydi. Rasmiy sahifada ayrim ma'lumot toifalari bilan ishlashga ruxsat hamda LLM provayderlari tomonidan jurnal yuritilmasligi ko'rsatilgan [13]. Bu xizmat tavsifi mustaqil xavfsizlik auditi xulosasi emas; Singapurdagi tasniflarni boshqa mamlakatning maxfiylik toifalariga avtomatik tenglashtirish ham mumkin emas.

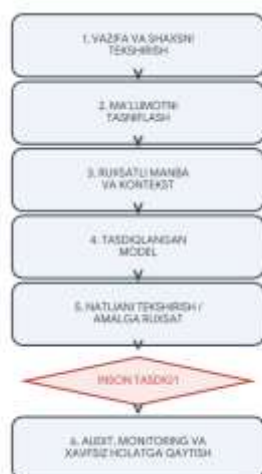
2-jadval. Xorijiy manbalarning qiyosiy mazmuni

Manba va hudud	Asosiy qamrov	Davlat tashkiloti uchun analitik xulosa
NIST, AQSH [2, 3]	Havf boshqaruvi va GenSIga xos xatarlar	Vazifa egasi, baholash va qoldiq xavfni qayd etish
AI Playbook va NCSC, Buyuk Britaniya [7, 8]	Foydalanish tartibi va xavfsiz hayotiy sikl	Xodim qoidasi hamda texnik talabni ajratish

Manba va hudud	Asosiy qamrov	Davlat tashkiloti uchun analitik xulosa
ENISA, Yevropa Ittifoqi [9, 10]	Bazaviy, SIga xos va sohaviy himoya	Maxsus SI nazoratlarini mavjud himoyaga qo‘shish
CSA, IMDA va Pair, Singapur [11–13]	Hayotiy sikl, agent vakolati va xizmat namunasi	Ma’lumot toifasi va amal vakolatini birgalikda cheklash

Davlat tashkiloti uchun ko‘p qatlamli himoya modeli

Qiyosiy tahlil asosida “**vazifa–ma’lumot–vakolat–dalil**” bog‘lanishiga asoslangan konseptual model taklif etiladi. Uning tayanch qoidasi: GenSI ma’lumotni qayta ishlashi mumkin, ammo o‘ziga yangi huquq bera olmaydi. Ruxsat haqidagi qaror model matnidan emas, tashkilotning autentifikatsiya, hujjatga kirish va amalni tasdiqlash mexanizmlaridan olinadi. Model uch omilni birgalikda ko‘rib chiqadi: ma’lumotning sezgirliigi, amalning oqibati va agentga berilgan vakolat ko‘lami.



1-rasm. GenSI xizmatini nazorat qilishning konseptual modeli

Birinchi qatlam - vazifa va javobgarlik. Har bir xizmat uchun maqsad, mas’ul bo‘linma, ruxsat etilgan amallar va to‘xtatish shartlari belgilanadi. Masalan, ochiq hujjatni qisqartirish uchun yaratilgan yordamchiga ichki tizimdagi yozuvlarni o‘zgartirish vazifasi

berilmaydi. Foydalanuvchi va xizmat hisobi alohida aniqlanib, kirish huquqi xizmat ehtiyojidan kelib chiqadi.

Ikkinchi qatlam - ma'lumotlarni himoyalash. Kiruvchi ma'lumot tasniflanadi va vazifa uchun zarur bo'lmagan qismlar chiqariladi. Ma'lumotlar sizishining oldini olish vositasi - DLP - qo'shimcha nazorat sifatida qo'llanadi. Biroq ismni o'chirishning o'zi to'liq anonimlashtirish hisoblanmaydi. Xizmatni baholashda ma'lumotni saqlash, qayta ishlash, provayder kirishi va modelni o'qitishda foydalanish shartlari alohida tekshiriladi; ularni barcha GenSI xizmatlari uchun bir xil deb taxmin qilish mumkin emas [8, 11].

Uchinchi qatlam - bilimlar bazasi va kontekst. RAG, ya'ni tashqi manbadan ma'lumot topib javob yaratishni boyitish usulida, hujjatga ruxsat model javob berganidan keyin emas, qidiruv natijasini modelga uzatishdan oldin tekshiriladi. Hujjatning manbasi, versiyasi va amal qilish holati saqlanadi. OWASP vektor bazalarida foydalanuvchi guruhlari orasidagi ma'lumot aralashuvi va ruxsatsiz foydalanishga qarshi mayda darajali nazoratlarni ko'rsatadi [14]. RAGning mavjudligi o'z-o'zidan maxfiylik yoki faktik aniqlik kafolati emas.

To'rtinchi qatlam - model va ulangan vositalar. Tasdiqlangan modelga faqat xizmat uchun zarur vositalar ulanadi. Model taklif qilgan buyruq ijroga uzatilishidan oldin mustaqil dasturiy nazoratdan o'tadi: amal turi, parametrlar, manzil va foydalanuvchi vakolati tekshiriladi. Hujjat o'qish huquqi uni o'chirish huquqini anglatmaydi. Inson tasdig'i kerak bo'lgan amallar va ruxsat etilgan xizmatlar oldindan belgilanadi [12].

Beshinchi qatlam - natija va inson nazorati. Matnli javob uchun manba va mazmun tekshiriladi; axborot tizimini o'zgartiradigan amal uchun esa ijrodan oldingi aniq tasdiq talab etiladi. Tasdiqlovchi xodim faqat "davom etish" tugmasini emas, qaysi hujjat, qaysi maydon yoki qaysi tashqi manzil bilan qanday amal bajarilishini ko'rishi kerak. Model matni rasmiy qaror yoki tayyor dasturiy buyruq sifatida avtomatik qabul qilinmaydi.

Oltinchi qatlam - audit va hodisaga javob. Xizmatda model versiyasi, murojaat vaqti, foydalanilgan manbalar, ruxsat tekshiruvlari va tasdiqlangan amallar qayd etiladi.

Jurnallarni saqlashda maxfiy matn va hisob ma'lumotlarining ortiqcha nusxalanishiga yo'l qo'yilmaydi. Shubhali holatda vosita ulanishini uzish, hisob vakolatini bekor qilish va xavfsiz holatga qaytish tartibi bo'ladi. Bu qatlam NCSCning ekspluatatsiya va monitoring yondashuviga mos keladi [7].

MUHOKAMA VA AMALIY TAVSIYALAR

Taklif etilgan modelning hissasi yangi himoya algoritmi yaratilganida emas, turli hujjatlardagi talablarni davlat tashkilotining ish jarayoniga bog'lashidadir. Modelda nazorat faqat prompt va yakuniy javobga emas, hujjatni olish hamda amalni bajarish oralig'idagi ruxsat chegaralariga ham joylashtiriladi. Shuning uchun hujjatda xavfsizlik qoidasi mavjudligi bilan uning dasturiy jihatdan majburiy bajarilishi farqlanadi.

Birinchi muhim cheklov - mahalliy server yoki yopiq tarmoqning o'zi xavfsizlikni kafolatlamaydi. U ayrim tashqi ma'lumot oqimlarini kamaytirishi mumkin, biroq noto'g'ri ruxsatlar, zaif komponentlar va ichki kontentdagi manipulyatsiya saqlanib qolishi mumkin. Xalqaro qo'shma tavsiyada tashqi ishlab chiquvchi yaratgan SI tizimini joriy etishda muhit konfiguratsiyasi, tizimni tekshirish va uzluksiz monitoring zarurligi ko'rsatilgan [15]. Demak, joylashtirish shakli hamda himoya sifati alohida baholanadi.

Ikkinchi cheklov - promptni filtrlashni yagona himoya vositasi qilishdir. Bilvosita ko'rsatma oddiy foydalanuvchi so'rovida emas, o'qilayotgan hujjatda bo'lishi mumkin [5]. Shu sababli filtrni ruxsat nazorati, bilimlar bazasi yaxlitligi va agent amallarini cheklash bilan birlashtirish taklif etiladi. Shu bilan birga, haddan tashqari bloklash xizmatning foydali vazifalarini ham to'xtatishi mumkin; xavfsizlik nazoratlari funksional yaroqlilik bilan birga baholanadi.

O'zbekiston davlat tashkilotlari uchun qo'llash tartibi. Dastlab GenSI xizmatlari, tashqi provayderlar va ulangan ma'lumot manbalarining inventari tuzilishi mumkin. Har bir xizmat uchun vazifa egasi, kiruvchi ma'lumot toifasi va ruxsat etilgan amallar qayd etiladi. Texnik ruxsatlar, ma'lumotlarni qayta ishlash shartlari va hodisaga javob tartibi aniqlanmaguncha xizmatni ichki hujjatlarga ulash taklif etilmaydi. Bu texnik-tashkiliy

taklif bo‘lib, muayyan tashkilotning amaldagi huquqiy majburiyatlarini aniqlash o‘rnini bosmaydi.

Joriy etishni ochiq hujjatlar bilan ishlaydigan, tashqi tizimga yozish huquqi bo‘lmagan yordamchidan boshlash mumkin. Keyingi bosqichda ichki qidiruv faqat hujjat darajasidagi ruxsatlar tekshirilganidan so‘ng qo‘shiladi. Tizimga yozish yoki hujjat yuborish imkoniyati esa alohida nazorat nuqtasi sifatida ko‘riladi. Bu ketma-ketlik umumiy muddat yoki majburiy bosqichlar sonini emas, vakolat kengaygani sari qo‘shimcha tekshiruv zarurligini ifodalaydi.

Qabul qilish mezonlari oldindan belgilanishi lozim. Masalan, vakolatsiz hujjatning qidiruv natijasiga tushmasligi, tasdiqsiz yozish amalining bloklanishi va bloklangan hodisaning auditda qayd etilishi tekshirilishi mumkin. Bunda haqiqiy maxfiy hujjatlar o‘rniga sun‘iy belgilangan test ma‘lumotlaridan foydalaniladi. Har bir tekshiruv uchun ssenariy, kutilgan natija, kuzatilgan natija, dalil va tuzatish chorasi qayd etiladi. Ushbu maqolada mazkur tekshiruvlar bajarilmagan; ular modelni keyinchalik baholash mezonlari sifatida taklif etilmoqda.

Xodimlarni tayyorlashda faqat prompt tuzish emas, ma‘lumotni tasniflash, model javobini tekshirish, ruxsatsiz vositani aniqlash va hodisa haqida xabar berish ko‘nikmalari ham qamrab olinishi kerak. Hujjatdagi zararli ko‘rsatmaga duch kelish misoli tushuntirilganda, maqsad hujum usullarini tarqatish emas, tashqi matnni avtomatik buyruq deb qabul qilmaslikni o‘rgatishdir. Xodim tayyorgarligi texnik cheklovlarni to‘ldiradi, ularning o‘rnini bosmaydi.

Tahlilning chegarasi shundaki, ochiq qo‘llanma va xizmat tavsifi tashkilot ichidagi barcha nazoratlarning ishlashini tasdiqlamaydi. Shuningdek, keltirilgan texnik tavsiyalar turli infratuzilmalarda bir xil natija berishi isbotlanmagan. Shu bois konseptual modelning samaradorligi haqida foizli yoki sababiy xulosa chiqarilmaydi; uning amaliy qiymati keyingi texnik audit va tekshiriladigan natijalar orqali aniqlanadi.

XULOSA

Xorijiy manbalar tahlili davlat tashkilotlarida GenSI xavfsizligini ta'minlashni alohida filtr yoki bitta model tanloviga qisqartirish mumkin emasligini ko'rsatadi. NIST havflarni boshqarish va javobgarlikni, Buyuk Britaniya qo'llanmalari foydalanish tartibi va hayotiy siklni, ENISA ko'p qatlamli himoyani, Singapur manbalari esa agent vakolatlari va boshqariladigan xizmat muhitini yoritadi [2, 7–12].

Maqolada taklif etilgan “vazifa–ma'lumot–vakolat–dalil” modeli ma'lumot sezgirligi, amal oqibati va vakolat ko'lamini birgalikda ko'rib chiqadi. Uning markaziy tamoyili - model ma'lumotni qayta ishlasa ham, ruxsat berish va javobgarlik mexanizmlarini almashtirmasligidir. O'zbekiston davlat tashkilotlari sharoitida mazkur yondashuv GenSI xizmatlariga qo'yiladigan texnik talablar, kirish nazorati va audit mezonlarini ishlab chiqish uchun konseptual asos bo'lishi mumkin. Modelning amaliy samaradorligi esa alohida tekshiruvni talab qiladi.

FOYDALANILGAN ADABIYOTLAR

1. OECD. (2025). Governing with Artificial Intelligence: The State of Play and Way Forward in Core Government Functions. OECD Publishing. DOI: 10.1787/795de142-en. [Manba](#)
2. NIST. (2023). Artificial Intelligence Risk Management Framework (AI RMF 1.0). NIST AI 100-1. DOI: 10.6028/NIST.AI.100-1.
3. Autio, C., Schwartz, R., Dunietz, J. va boshq. (2024). Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile. NIST AI 600-1. DOI: 10.6028/NIST.AI.600-1.
4. OWASP. (2025). Top 10 for Large Language Model Applications. 2025-yilgi nashr. OWASP Gen AI Security Project.
5. Greshake, K., Abdelnabi, S., Mishra, S., Endres, C., Holz, T., Fritz, M. (2023). Not what you've signed up for: Compromising Real-World LLM-Integrated Applications with Indirect Prompt Injection. arXiv:2302.12173, preprint.

6. Zhan, Q., Liang, Z., Ying, Z., Kang, D. (2024). InjecAgent: Benchmarking Indirect Prompt Injections in Tool-Integrated Large Language Model Agents. Findings of ACL 2024, 10471–10506. DOI: 10.18653/v1/2024.findings-acl.624.
7. NCSC, CISA va xalqaro hamkorlar. (2023). Guidelines for Secure AI System Development. 27 November.
8. Government Digital Service. (2025). Artificial Intelligence Playbook for the UK Government. 10 February. [Manba](#)
9. ENISA. (2023). Multilayer Framework for Good Cybersecurity Practices for AI. DOI: 10.2824/588830. [Manba](#)
10. ENISA. (2026). ENISA’s View on Cybersecurity in the Frontier AI Era. 7 July.
11. Cyber Security Agency of Singapore. (2024). Guidelines and Companion Guide on Securing AI Systems. 15 October.
12. Ministry of Digital Development and Information, Singapore. (2026). Singapore Launches New Model AI Governance Framework for Agentic AI. 22 January.
13. Open Government Products. (sanasiz). Pair. Davlat xizmatchilari uchun SI xizmatining rasmiy sahifasi. Murojaat sanasi: 29.09.2026.
14. OWASP. (2025). LLM08:2025 Vector and Embedding Weaknesses. OWASP Gen AI Security Project.
15. Canadian Centre for Cyber Security. (2024). Joint Advisory on Deploying AI Systems Securely. NSA, CISA va hamkorlar qo‘shma qo‘llanmasi haqidagi rasmiy material. 15 April.