

Reziliența operațională digitală în lumea Inteligenței Artificiale

Managementul riscurilor AI și controale
bazate pe ISO/IEC 42001

Călin Mihai Rangu

Cyberthreats IBR, 15 octombrie 2026

AI și profilul rezilienței bancare

Se urmărește capacitatea de a detecta, limita și corecta efecte, împreună cu menținerea serviciului.

Declarația ESAs din 31 iulie 2026 leagă riscurile modelelor avansate de procesele și controalele DORA existente

Amenințări	AI accelerează atacurile și fraudele, inclusiv pentru băncile care nu o folosesc.
Vunerabilități	AI poate influența deciziile, datele și serviciile.
Apărare	AI susține detecția și răspunsul, cu propriile limite și dependențe.

Scenariul bancar de referință

Context	Un asistent AI extrage informații din documente pentru pregătirea unei plăți corporate.
Eveniment	Un document extern conține instrucțiuni ascunse care influențează interpretarea beneficiarului.
Vulnerabilitate	Fluxul preia recomandarea fără verificare independentă sau îi acordă agentului drepturi excesive.
Consecință	Date incorecte ajung în pregătirea operațiunii și pot influența plata dacă autorizarea nu le detectează.

Scenariu ipotetic. Impactul depinde de drepturile agentului și de controalele plății.

Utilizări AI și expuneri bancare

Utilizare	Beneficiu urmărit	Expunere principală
Asistent pentru angajați	Acces rapid la informații	Divulgare, informații incorecte
Chatbot pentru clienți	Răspuns și disponibilitate	Informare eronată, confidențialitate
Evaluarea bonității	Analiză și consecvență	Erori, discriminare, explicabilitate
AML și prevenirea fraudei	Prioritizarea alertelor	Cazuri omise, alarme false
AI în operațiunile de securitate	Analiza evenimentelor	Recomandări sau intervenții greșite
Agenți pentru operațiuni	Automatizarea activităților	Acțiuni neautorizate, propagarea erorilor

Evaluarea se face pentru fiecare utilizare concretă și pentru efectele ei în proces.

Clasificarea utilizărilor după impact și autonomie

Mod de utilizare	Efect posibil	Control proporțional
Redactare și sinteză	Influențează un document sau analiza	Surse verificate, reguli de date, revizuire
Recomandare într-un proces	Influențează decizia asupra clientului	Validare independentă, explicații, escaladare
Execuție prin instrumente / API	Modifică date sau inițiază acțiuni	Mandat limitat, autorizare distinctă, oprire
Utilizare într-un proces important	Poate afecta continuitatea serviciului	Alternativă testată, monitorizare, recuperare

Clasificarea internă a riscului și încadrarea juridică AI Act se documentează distinct.

Riscurile AI care afectează serviciile bancare

Grup de riscuri	Exemplu	Efect operațional
Date și confidențialitate	Dosare introduse într-un instrument neaprobat	Expunere de informații
Corectitudine și echitate	Analiză incorectă sau tratament neechitabil	Decizii afectate, remedierea cazurilor
Autonomie și responsabilitate	Agent care depășește mandatul	Acțiuni neautorizate
Shadow AI și competențe	Conturi personale utilizate în proces	Pierderea vizibilității și controlului
Furnizori și schimbări	Actualizare care degradează rezultatele	Întrerupere sau funcționare nesigură
Atacuri și trasabilitate	Manipularea documentelor consultate	Compromiterea fluxului, investigații dificile

Registrul AI se corelează cu riscul operațional, TIC, de model și de conformitate.

DORA, AI Act și ISO/IEC 42001

Reper	Rol în cadrul băncii
DORA	Riscul TIC, continuitate, incidente, testare și furnizori
AI Act	Obligații în funcție de sistem, scop și rolul organizației
GDPR și cadrul bancar	Protecția datelor, drepturile persoanelor și regulile procesului
ISO/IEC 42001	Responsabilități, evaluări, controale, dovezi și îmbunătățire pentru managementul AI

Calendar actualizat AI Act: anexa III, 2 decembrie 2027. Controalele pentru riscurile actuale continuă.

ISO/IEC 42001: sistemul de management AI

Componentă	Întrebarea de management	Rezultat verificabil
Context și leadership	Unde folosim AI și cine răspunde?	Domeniu, politică, roluri
Planificare și suport	Ce riscuri avem și ce resurse cer?	Evaluări, obiective, competențe
Operare	Cum controlăm utilizarea și schimbările?	Controale, aprobări, evidențe
Evaluare și îmbunătățire	Cum verificăm și corectăm?	Indicatori, audit, revizuire, măsuri

Risc identificat, control selectat, dovadă verificată, reevaluare.

ISO/IEC 42001: aplicabilitate și controale

Element	Rol în sistemul de management AI
Domeniu și structură	Standardul se aplică organizațiilor care dezvoltă, furnizează sau utilizează AI. Clauzele 4–10 formează structura sistemului de management.
Anexele A și B	Anexa A oferă controale de referință. Anexa B oferă ghidaj pentru implementare.
Selecția controalelor	Selecția pornește de la riscurile identificate.
Declarația de aplicabilitate (SoA)	Statement of Applicability documentează controalele necesare, justificările, implementarea și justificarea excluderilor din setul de referință.
Evaluarea impactului AI	Se corelează, după caz, cu DPIA și cu evaluarea drepturilor fundamentale, fără a presupune că aceste evaluări sunt identice.

Banca poate avea nevoie și de măsuri suplimentare.

Controale înainte de introducerea în utilizare

Punct de control	Referințe ISO	Dovadă
Scop, limite și responsabil desemnat	A.3.2; A.9.4	Fișa utilizării și mandat aprobat
Date, aplicații și dependențe	A.4.2–A.4.5	Inventar și arhitectură
Impact asupra clienților	A.5.2–A.5.4	Evaluare documentată și măsuri
Furnizor și responsabilități	A.10.2; A.10.3	Evaluare și prevederi contractuale
Testare și aprobare pentru producție	A.6.2.4; A.6.2.5	Rezultate ale testelor și aprobare

Introducerea în producție presupune rezultate demonstrate și acceptarea riscului rezidual.

Controale pe durata exploatării

Punct de control	Referințe ISO	Dovadă
Calitatea și proveniența datelor	A.7.4–A.7.6	Verificări și istoricul surselor
Utilizare în limitele aprobate	A.8.2; A.9.2; A.9.4	Instrucțiuni și verificarea respectării
Performanță și schimbări	A.6.2.4; A.6.2.6	Monitorizare și retestare
Trasabilitate	A.6.2.8	Jurnale protejate și validări
Raportare și comunicare a incidentelor	A.3.3; A.8.3; A.8.4	Dosar de incident și comunicări

Controalele tehnice acoperă datele, privilegiile și acțiunile executate.

Monitorizarea AI și controlul accesului

Control	Aplicare în exploatare
Monitorizare	Urmărește rezultatele, utilizarea, datele și integrările.
Schimbări și retestare	Modificarea modelului, a corpusului documentar, a prompturilor de sistem sau a drepturilor agentului poate schimba riscul și cere retestare.
Documente interne	Verificăm autorizarea și la extragerea informației, astfel încât utilizatorul să primească doar date la care are dreptul.
Agenți AI	Aplicăm privilegii minime, restricții de destinație și autorizări în sistemele care execută acțiunile.
Trasabilitate	Proiectăm trasabilitatea proporțional cu riscul, cu acces și retenție controlate, pentru a evita acumularea inutilă de date sensibile.

Filtrarea prompturilor este o măsură complementară.

Supravegherea umană în procesul bancar

- Persoana** Are competențe, mandat și timp pentru verificare.
 - Verificarea** Folosește surse independente și criterii clare de acceptare.
 - Intervenția** Poate respinge, modifica, escalada și opri utilizarea.
 - Dovada** Documentează verificarea, decizia și excepțiile relevante.
- Responsabilitatea, competența și posibilitatea de intervenție trebuie să funcționeze împreună.**

Riscul rezidual bazat pe dovezi

Etapă	Evaluare necesară	Dovadă
Control planificat	Beneficiu estimat, încă nedemonstrat	Plan de implementare
Control implementat	Proiectare adecvată și acoperire	Configurație, procedură, acces
Control testat	Funcționare în cazuri relevante	Teste și excepții
Risc rezidual	Probabilitate și impact reevaluate	Evaluare și acceptare autorizată

Exemplu de lucru: 20 redus la 9 la 55% eficacitate estimată. Reducerea cere validare.

Continuitatea serviciilor în trei scenarii AI

Scenariu	Răspuns operațional	Ce trebuie testat
AI indisponibilă	Activarea alternativei și gestionarea volumului restant	Capacitate, timp de activare, priorități
AI produce rezultate nesigure	Restricționare și verificarea operațiunilor afectate	Detectie, delimitare, corectare
AI sau sursele sunt compromise	Izolare și revenire la configurații / surse validate	Dovezi, integritate, condiții de reluare

Disponibilitatea, integritatea rezultatelor și capacitatea alternativei se testează separat.

Răspunsul la incident în scenariul plății

Moment	Acțiune	Rezultat urmărit
Limitare	Restricționarea integrării și a drepturilor afectate	Oprirea propagării
Investigație	Păstrarea documentelor, configurației și jurnalelor	Reconstruirea evenimentului
Delimitare	Identificarea operațiunilor influențate	Stabilirea impactului
Continuitate	Activarea fluxului alternativ și verificări independente	Menținerea serviciului
Reluare	Remediere, retestare și aprobare autorizată	Utilizare în condiții validate

Comunicarea și notificările se stabilesc în funcție de impact și de obligațiile aplicabile.

Guvernanță și responsabilitate

Rol	Responsabilitate
Conducere	Apetit la risc, resurse și limite semnificative
Prima linie / proprietar de proces	Utilizare, rezultate și implementarea măsurilor
IT și securitate	Controale tehnice în aria proprie
A doua linie	Metodologie, examinarea evaluărilor, monitorizare
DPO și juridic	Contribuție conform mandatului propriu
Audit intern	Asigurare independentă asupra cadrului și controalelor

Un responsabil principal pentru fiecare risc și separarea clară a funcțiilor de control.

Indicatori care susțin deciziile conducerii

Indicator propus	Decizie susținută
Utilizări active fără evaluare sau aprobare	Regularizare ori restricționare
Schimbări semnificative fără retestare	Limitarea utilizării până la validare
Erori cu impact asupra clienților	Remediarea procesului și a cazurilor
Utilizări sensibile fără trasabilitate	Consolidarea evidențelor
Procese importante fără alternativă testată	Investiții în continuitate
Dependențe comune și timp de activare a alternativei	Reducerea concentrării și îmbunătățirea răspunsului

Fiecare indicator are definiție, toleranță, responsabil și acțiune de escaladare.

Primele 90 de zile de implementare

Perioadă	Rezultate urmărite	Verificare
0–30 de zile	Inventar, responsabili, reguli de date și utilizare, identificarea Shadow AI	Acoperirea utilizărilor și expunerilor sensibile
31–60 de zile	Evaluarea cazurilor prioritare, controale, verificarea furnizorilor și testare	Dovezi și măsuri pentru riscurile importante
61–90 de zile	Exercițiu de incident, test de continuitate, indicatori și revizuire managerială	Capacitatea de răspuns și planul de remediere

Prioritate: impact asupra clienților, autonomie, date sensibile și criticitatea serviciului.

Întrebările conducerii despre AI

Știm unde folosim AI și ce poate face fiecare sistem?

Avem dovezi că funcționează controalele?

Cine poate limita sau opri utilizarea?

Am testat continuarea serviciului și corectarea efectelor?

Reziliența AI cere control asupra deciziilor și serviciilor pe întregul ciclu de utilizare.

Încheiere

Adoptarea AI cere o responsabilitate clară pentru efectele produse în procese.

Putem încuraja utilizarea atunci când definim scopul, verificăm rezultatele și limităm autonomia potrivit riscului.

ISO/IEC 42001 ajută organizația să gestioneze consecvent aceste condiții, iar DORA ne cere să menținem și să demonstrăm reziliența digitală în cadrul aplicabil.

Cele patru întrebări pot fi folosite în orice comitet care aprobă utilizări AI sau examinează incidente.

Anexele oferă o matrice a celor 18 riscuri din materialele inițiale și un catalog de 38 de controale adaptate pentru bănci.

Acestea sunt instrumente de lucru, nu evaluări ale unei instituții concrete.

Anexă: riscuri AI în procesele bancare (1/3)

ID	Risc	Scenariu bancar	Tratament propus
AI-R01	Divulgarea informațiilor confidențiale	Dosare de credit sau date despre conturi încărcate necontrolat	Clasificare, instrumente aprobate, acces minim, verificarea furnizorului
AI-R02	Prelucrarea necorespunzătoare a datelor personale	Date despre clienți utilizate fără evaluarea condițiilor de prelucrare	Minimizare, evaluări necesare, reguli de date și contracte
AI-R03	Halucinații și informații factuale eronate	Chatbotul comunică o condiție de produs inexistentă	Surse validate, teste, revizuire și limitarea utilizării
AI-R04	Automatizarea necontrolată a deciziilor	Rezultatul AI ajunge direct într-o decizie sau operațiune	Mandat, autorizare distinctă, supraveghere și escaladare
AI-R05	Bias sau tratament neechitabil	Rezultate diferite nejustificate în analiza clienților	Evaluarea impactului, testare și monitorizarea rezultatelor
AI-R06	Utilizare în afara scopului aprobat	Asistentul de sinteză este extins informal la aprobări	Limite documentate, aprobare pentru schimbarea scopului

Scenarii și tratamente orientative. Scorurile se stabilesc prin metodologia băncii.

Anexă: riscuri AI în procesele bancare (2/3)

ID	Risc	Scenariu bancar	Tratament propus
AI-R07	Aplicații AI neaprobate	Personalul folosește conturi personale în activitate	Inventar, alternative aprobate, reguli și măsuri tehnice
AI-R08	Dependență excesivă de outputul AI	Analistul preia concluzia fără verificarea surselor	Instruire, verificări independente și teste de supraveghere
AI-R09	Lipsa trasabilității	Banca nu poate identifica sursa sau versiunea unui rezultat	Jurnale, versiuni, surse și evidența validării
AI-R10	Dependență și control insuficient asupra furnizorului	Serviciul AI extern devine indisponibil	Evaluare, contract, monitorizare și alternativă testată
AI-R11	Schimbarea modelului sau degradarea performanței	Actualizarea schimbă rezultatele ori datele devin nereprezentative	Controlul schimbărilor, monitorizare și retestare
AI-R12	Prompt injection și manipularea contextului	Documentul extern influențează o recomandare operațională	Privilegii minime, validări în aplicații și teste adversariale

Scenarii și tratamente orientative. Scorurile se stabilesc prin metodologia băncii.

Anexă: riscuri AI în procesele bancare (3/3)

ID	Risc	Scenariu bancar	Tratament propus
AI-R13	Încălcarea drepturilor de proprietate intelectuală	Cod sau conținut reutilizat fără clarificarea drepturilor	Reguli IP, verificări și termeni contractuali
AI-R14	Neconformitate legală sau reglementară	Utilizarea începe fără încadrare și evaluarea obligațiilor	Maparea cerințelor, evaluare și reevaluare
AI-R15	Competențe și alfabetizare AI insuficiente	Utilizatorul nu recunoaște limitele sau utilizările interzise	Instruire pe roluri și verificarea înțelegerii
AI-R16	Responsabilități neclare	Businessul, IT și furnizorul presupun că altcineva aprobă	Responsabil principal, matrice de roluri și escaladare
AI-R17	Incident raportat sau tratat necorespunzător	Eroarea semnalată nu declanșează analiza operațiunilor afectate	Triere, investigație, notificări și acțiuni corective
AI-R18	Impact negativ neidentificat asupra persoanelor	Introducerea AI afectează clienți vulnerabili fără analiză	Evaluarea impactului, măsuri și mecanisme de contestare

Scenarii și tratamente orientative. Scorurile se stabilesc prin metodologia băncii.

Politici și organizare pentru AI

ID	Control, parafrizat	Aplicare bancară propusă	Dovadă tipică
A.2.2	Politica AI	Scopuri permise, restricții, aprobare și escaladare	Politică aprobată și comunicată
A.2.3	Corelarea politicilor	Aliniere cu riscul, securitatea, datele și achizițiile	Matrice de corelare
A.2.4	Revizuirea politicii	Revizuire periodică și la schimbări semnificative	Istoric și decizii de revizuire
A.3.2	Roluri și responsabilități	Responsabil de business, implementare, control și oprire	Matrice de roluri și mandate
A.3.3	Raportarea preocupărilor	Canal pentru erori, utilizări nepermise și incidente	Sesizări și analiză documentată

Selecția și justificarea controalelor se documentează în declarația de aplicabilitate.

Resurse pentru AI

ID	Control, parafrizat	Aplicare bancară propusă	Dovadă tipică
A.4.2	Documentarea resurselor	Inventar AI, date, utilizatori și dependențe	Registru și arhitectură
A.4.3	Resurse de date	Categorii de date admise și restricționate	Inventar și clasificare
A.4.4	Instrumente și platforme	Aplicații, servicii și versiuni relevante	Catalog aprobat
A.4.5	Sistem și calcul	Infrastructură, API, identități și dependențe cloud	Arhitectură și configurații
A.4.6	Competențe și persoane	Pregătire pe roluri, inclusiv pentru intervenție	Instruire și evaluări

Selecția și justificarea controalelor se documentează în declarația de aplicabilitate.

Evaluarea impactului pentru AI

ID	Control, parafrizat	Aplicare bancară propusă	Dovadă tipică
A.5.2	Proces de evaluare a impactului	Evaluare înaintea utilizării și la schimbări	Metodologie și evaluare
A.5.3	Documentarea impactului	Concluzii, ipoteze, măsuri și decizii	Dosar al utilizării AI
A.5.4	Impact asupra persoanelor	Drepturi, echitate, acces și contestare	Analiză și măsuri
A.5.5	Impact societal	Efecte mai largi, după natura și scara utilizării	Justificare și analiză

Selecția și justificarea controalelor se documentează în declarația de aplicabilitate.

Dezvoltare și cerințe pentru AI

ID	Control, parafrizat	Aplicare bancară propusă	Dovadă tipică
A.6.1.2	Obiective de dezvoltare responsabilă	Obiective pentru soluții dezvoltate sau configurate	Cerințe și criterii de acceptare
A.6.1.3	Proces de dezvoltare responsabilă	Ciclu controlat pentru aplicații, RAG și agenți	Procedură și înregistrări de proiect
A.6.2.2	Cerințe și specificații	Scop, date, securitate, performanță și limite	Specificații aprobate
A.6.2.3	Documentarea proiectării	Arhitectură, configurație și modificări	Documentație tehnică

Selecția și justificarea controalelor se documentează în declarația de aplicabilitate.

Validare și exploatare pentru AI

ID	Control, parafrizat	Aplicare bancară propusă	Dovadă tipică
A.6.2.4	Verificare și validare	Teste funcționale, negative și adversariale	Plan, rezultate și remedieri
A.6.2.5	Introducere în utilizare	Criterii și aprobări pentru producție	Dosar de lansare
A.6.2.6	Operare și monitorizare	Calitate, incidente, schimbări și utilizare	Rapoarte și indicatori
A.6.2.7	Documentație operațională	Configurație, limite și instrucțiuni actuale	Documentație și proceduri
A.6.2.8	Jurnalizarea evenimentelor	Evidențe pentru monitorizare și investigații	Jurnale și reguli de retenție

Selecția și justificarea controalelor se documentează în declarația de aplicabilitate.

Date pentru AI

ID	Control, parafrizat	Aplicare bancară propusă	Dovadă tipică
A.7.2	Date pentru dezvoltare	Guvernanță pentru antrenare sau îmbunătățire	Inventar și criterii de date
A.7.3	Colectarea și achiziția datelor	Sursă, drept de utilizare și condiții	Temei, licențe și acorduri
A.7.4	Calitatea datelor	Relevanță, actualitate și adecvare la scop	Criterii și verificări
A.7.5	Proveniența datelor	Origine și transformări relevante	Metadate și istoric
A.7.6	Pregătirea datelor	Curățare, etichetare și transformări controlate	Proceduri și validări

Selecția și justificarea controalelor se documentează în declarația de aplicabilitate.

Informarea părților interesate pentru AI

ID	Control, parafrizat	Aplicare bancară propusă	Dovadă tipică
A.8.2	Informarea utilizatorilor	Limite, reguli de date și verificarea rezultatelor	Instrucțiuni și instruire
A.8.3	Raportare externă	Identificarea obligațiilor și responsabilităților	Matrice de obligații
A.8.4	Comunicarea incidentelor	Informarea părților relevante potrivit impactului	Decizii și comunicări
A.8.5	Informarea părților interesate	Transparență proporțională cu utilizarea	Informări și mecanisme de contact

Selecția și justificarea controalelor se documentează în declarația de aplicabilitate.

Utilizare și relația cu terții pentru AI

ID	Control, parafrizat	Aplicare bancară propusă	Dovadă tipică
A.9.2	Procese de utilizare responsabilă	Reguli de aprobare, date, verificare și incidente	Procedură și evidențe
A.9.3	Obiective de utilizare responsabilă	Obiective și indicatori pentru utilizarea aprobată	Indicatori și toleranțe
A.9.4	Scop și limite	Utilizare intenționată și utilizări nepermise	Fișă și mandat
A.10.2	Responsabilități în lanț	Alocare între bancă, furnizori și parteneri	Contract și responsabilități
A.10.3	Furnizori	Evaluare și monitorizare a serviciului AI	Evaluare, contract și revizuirii
A.10.4	Relația cu clienții	Responsabilități pentru servicii AI oferite clienților	Termeni și informări

Selecția și justificarea controalelor se documentează în declarația de aplicabilitate.

Surse oficiale și repere de utilizare

ISO/IEC 42001:2023

Standardul de referință pentru sistemul de management AI

ISO, explicația sistemului de management

Guvernanță, aplicabilitate și certificare voluntară

ESAs, JC 2026 25, 31 iulie 2026

Riscuri TIC generate de modelele AI avansate

DORA, Regulamentul (UE) 2022/2554

Cadrul de reziliență operațională digitală

AI Act: surse și baza metodologică

Comisia Europeană, 27 iulie 2026

Intrarea în vigoare a AI Omnibus și calendarul actualizat

AI Act Service Desk, anexa III

Utilizări relevante pentru clasificarea sistemelor

AI Act, Regulamentul (UE) 2024/1689

Verificarea normelor aplicabile și a dispozițiilor tranzitorii

Materialele de lucru ale autorului

Registrul de 18 riscuri și catalogul de 38 de controale, adaptate pentru bănci

Vă mulțumesc!

Q&A

Lect.univ.dr.ing.ec. Călin M. Rangu

calin@rangu.ro

www.futureishere.eu

