EBSA WEBINAR · PART 2

AI governance in the laboratory

12 practical points for responsible use



Stefan Wagener, PhD
Biorisk Management International (BRMI)
bioriskinternational.com · biorisk.ai

What is AI governance?

AI governance is the set of decisions, roles and checks that keep AI use purposeful, safe and accountable, from the first choice of tool until it is withdrawn.



What may it do?

Approved tasks, clear limits, rules for our data

Points 1 · 4 · 6



Who is responsible?

Named owner, reviewer and approver, with trained users

Points 2 · 10



Can we trust it?

Risk assessment, supplier evidence and our own tests

Points 3 · 5 · 7



Are we still in control?

Human approval, access control, monitoring, power to stop

Points 8 · 9 · 11 · 12



Not a ban, and not just an IT task. Governance is how the lab says yes to AI, safely.

Govern AI like a new employee



Job description

↓

Purpose and limits



Supervisor

↓

Responsibility



Reference checks

↓

Supplier evidence



Probation

↓

Controlled introduction



Same discipline, new colleague. No independent work until it has earned it.

Our example: *Biorisk-AI*

1 Configure

Agent, procedures, facility, people, incidents

2 AI analysis

Likelihood, consequence and uncertainty

3 Report

Hazards, scores, controls, citations

Biorisk-AI©

In beta testing with laboratories

biorisk.ai

Decision support, not decisions: the platform makes no final determinations or approvals.

12 governance points, in three phases

BEFORE CHOOSING

1. Purpose and limits
2. Responsibility
3. Risk assessment
4. Applicable rules

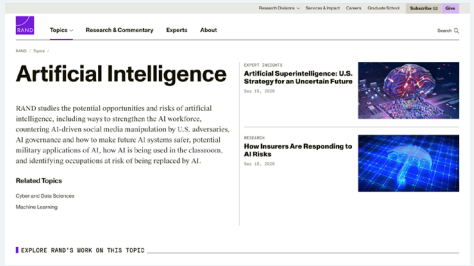
BEFORE ROUTINE USE

5. Supplier evidence
6. Data protection
7. Local testing
8. Human approval

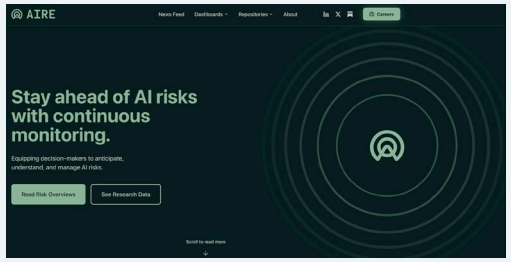
THROUGHOUT USE

9. Security and access
10. Procedures and skills
11. Controlled introduction
12. Monitoring and withdrawal

Tip: use external reports to ask better questions



RAND



AI Risk Explorer



AIxBio, Johns Hopkins

Tools and risks are moving fast: stay up to date.

EXTERNAL EVIDENCE

What was tested?

What limits?

Does it apply to our tool?

LOCAL DECISION

Check the workflow

Test the safeguards

Record the decision

1 Purpose and limits

BEFORE CHOOSING

RISK · BIORISK-AI EXAMPLE

Assessments drafted for review start being treated as approved.

KEY MITIGATION

Restrict use to approved tasks and decision limits.

EVIDENCE TO KEEP

Approved-use statement, task permissions, records of exceptions.



New employee: a job description and clear limits on duties.

2 Responsibility

BEFORE CHOOSING

RISK · BIORISK-AI EXAMPLE

Nobody owns the lab's Biorisk-AI use, checks its reports or can pause it.

KEY MITIGATION

An approved AI-use policy, and named people who own, review and approve.

EVIDENCE TO KEEP

AI-use policy, signed role assignments, dated review, approval, pause and restart decisions.

THREE ROLES

Tool owner
Keeps it controlled · can pause use

Qualified reviewer
Checks assessments · can reject advice

Approver
Accepts remaining risk · authorises uses

New employee: a supervisor who checks the work and can stop unsafe activity.

3 Risk assessment

BEFORE CHOOSING

RISK · BIORISK-AI EXAMPLE

Correct scenario inputs, but the likelihood, consequence or uncertainty score is unsuitable.

KEY MITIGATION

Implement and verify controls, such as qualified review, before use.

EVIDENCE TO KEEP

Risk assessment, control test results, approval of the remaining risk.



New employee: supervision and safety controls in place before independent work.

3 Rate each scenario, keep uncertainty separate

	Minor	Moderate	Severe
Likely	3	6	9
Possible	2	4	6
Unlikely	1	2	3

LIKELIHOOD × CONSEQUENCE = REVIEW PRIORITY



Priority for review, not approval



State how sure you are



Unreviewed is not uncertain

4 Applicable rules

BEFORE CHOOSING

RISK · BIORISK-AI EXAMPLE

Entering personnel or incident data breaches privacy or institutional rules.

KEY MITIGATION

Meet every applicable requirement before approving use.

EVIDENCE TO KEEP

Requirements checklist, completed actions, agreements, approval.

CHECK · ASK · KEEP

Safety and quality
Which rules apply? · Requirements checklist

Privacy and contracts
Personnel and incident data? · Supplier terms

Approval
Who confirms output? · Named decision

New employee: verify qualifications, training and confidentiality duties.

5 Supplier evidence

BEFORE ROUTINE USE

RISK · BIORISK-AI EXAMPLE

The curated evidence base may not cover your agents, procedures or country.

KEY MITIGATION

Restrict use where supplier evidence is inadequate.

EVIDENCE TO KEEP

Supplier evidence review, unresolved limits, approved-use restrictions.

ASK THE SUPPLIER

- **Model and version**
- **Curated sources (200+)**
- **Validation results**
- **Update schedule**
- **Data handling**

New employee: check references and test claimed skills in practice.

6 Data protection

BEFORE ROUTINE USE

RISK · BIORISK-AI EXAMPLE

Security-sensitive agent, facility and incident details are processed where the lab cannot control them.

KEY MITIGATION

Match where the tool runs to your most sensitive input; enforce storage, reuse and deletion rules.

EVIDENCE TO KEEP

Hosting and data-flow record, data-handling agreement, configuration record, deletion test results.

WHERE DOES IT RUN?

Local (on-premises)
Data stays on site · lab patches and backs up

Private cloud
Dedicated, named region · access by contract

Shared cloud (SaaS)
Supplier hosts · sent on to an AI provider?

New employee: works only in approved places, with access only to needed records.

7 Local testing

BEFORE ROUTINE USE

RISK · BIORISK-AI EXAMPLE

Reports for your own agents and procedures differ from expert-reviewed assessments.

KEY MITIGATION

Keep any task that fails local tests out of routine use.

EVIDENCE TO KEEP

Test plan and criteria, results, corrections, retest approval.

TEST THE ANSWER AND THE SAFEGUARDS

Assessment quality
Is the advice justified? · vs expert review

Information protection
Is access restricted? · permission tests

Human review
Are serious errors caught? · reviewer exercise

New employee: observed work on known-result samples against agreed criteria.

7 Red teaming: try to make it fail

Red teaming is a planned exercise in which a small team deliberately tries to make *Biorisk-AI* give wrong, unsafe or leaked output, using safe test records, before real users meet those failures.

Tricky inputs

Incomplete or contradictory scenario data

Pass: it flags the gaps

Access and leaks

Try to see another user's data or reports

Pass: access is refused

Weak citations

Check every source behind a suggested control

Pass: sources support it

Out-of-scope requests

Ask for tasks outside the approved use

Pass: it stays in scope

Plan, attack, record, fix, retest. Repeat before routine use and after every major update.

8 Human approval

BEFORE ROUTINE USE

RISK · BIORISK-AI EXAMPLE

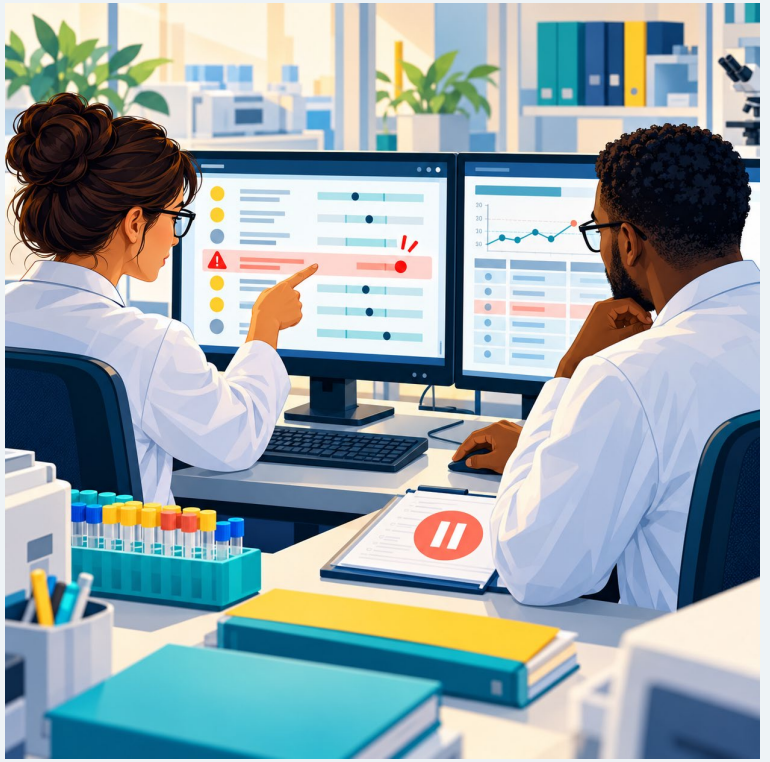
A Biorisk-AI report is used as the final assessment without qualified review.

KEY MITIGATION

A qualified person approves every output before use. No agentic use.

EVIDENCE TO KEEP

Reviewed assessment, corrections, reviewer identity, dated approval.



New employee: a supervisor checks important decisions until independence is earned.

9 Security and access

THROUGHOUT USE

RISK · BIORISK-AI EXAMPLE

Assessment reports are edited or approved by the wrong people, or versions are lost.

KEY MITIGATION

Control revisions; keep approved assessments available.

EVIDENCE TO KEEP

Change and approval history, edit permissions, recovery test.

GIVE ONLY THE ACCESS NEEDED

Laboratory staff
Use approved assessments

Reviewer
Review and approval

Administrator
Approved system settings

New employee: access matched to duties, and removed when it ends.

10 Procedures and skills



RISK · BIORISK-AI EXAMPLE

A reviewer accepts a suggested control without checking that its citation supports it.

KEY MITIGATION

Train staff and require demonstrated competence before use.

EVIDENCE TO KEEP

Training record, completed exercise, authorisation to use the tool.

PRACTICE EXERCISE

"The suggested control sounds right, but the cited source does not support it."

- Explain the problem
- Check the source
- Report and record

New employee: supervised practice and a competence check before independent work.

11 Controlled introduction

THROUGHOUT USE

RISK · BIORISK-AI EXAMPLE

Every lab adopts it at once instead of starting with a defined beta-test pilot.

KEY MITIGATION

Start with a supervised pilot; expand only on evidence.

EVIDENCE TO KEEP

Pilot plan, results, issue resolutions, approval stating limits.



New employee: supervised duties first; review performance before expanding them.

12 Monitoring and withdrawal

THROUGHOUT USE

RISK · BIORISK-AI EXAMPLE

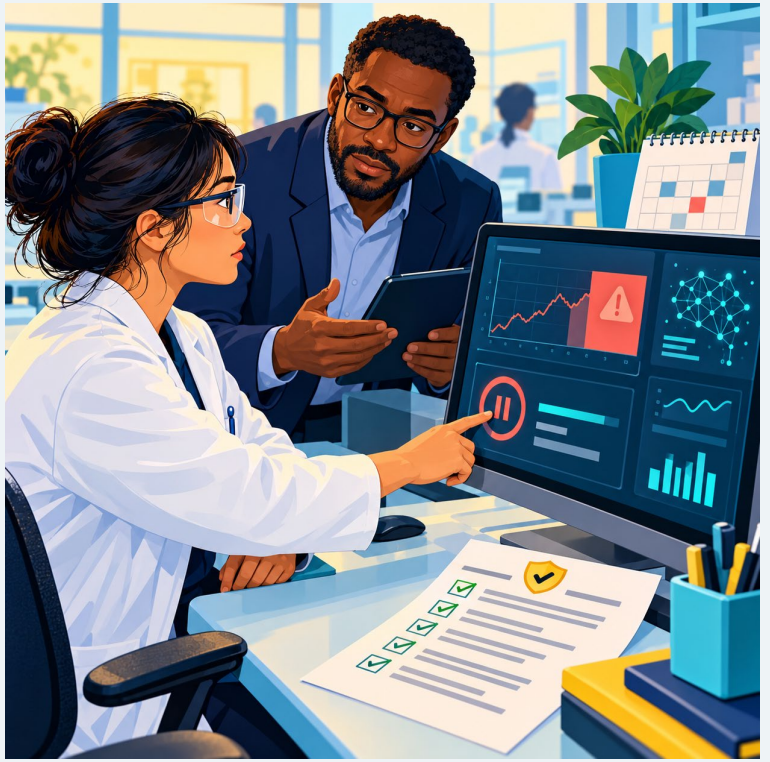
An update to the model or the evidence database changes scores without anyone noticing.

KEY MITIGATION

Retest after updates; report AI errors as near-misses; be ready to stop.

EVIDENCE TO KEEP

Version and change record, retest results, AI error and near-miss reports, restart approval.



New employee: reassess competence and access when duties change.

The 12 points align with major AI frameworks

ISO/IEC 42001

AI management system

Policy and roles	2
Impact assessment	3
Testing and monitoring	7, 12
Suppliers	5

NIST AI RMF

Risk management framework

Govern	2, 4, 10
Map	1, 3, 5
Measure	7
Manage	8, 11, 12

EU AI Act

Deployer duties

AI literacy (now)	10
Use per instructions	1
Human oversight	8
Monitoring and logs	12

ISO 35001

Your biorisk system

Risk assessment	3
Competence	10
Suppliers	5
Management review	12

Govern AI inside your biorisk management system, not beside it.

Make a clear decision

Approve a defined use

Relevant tests passed

Safeguards checked

Limits and remaining risk recorded

Pause and resolve

Important failures remain

Evidence can't support it

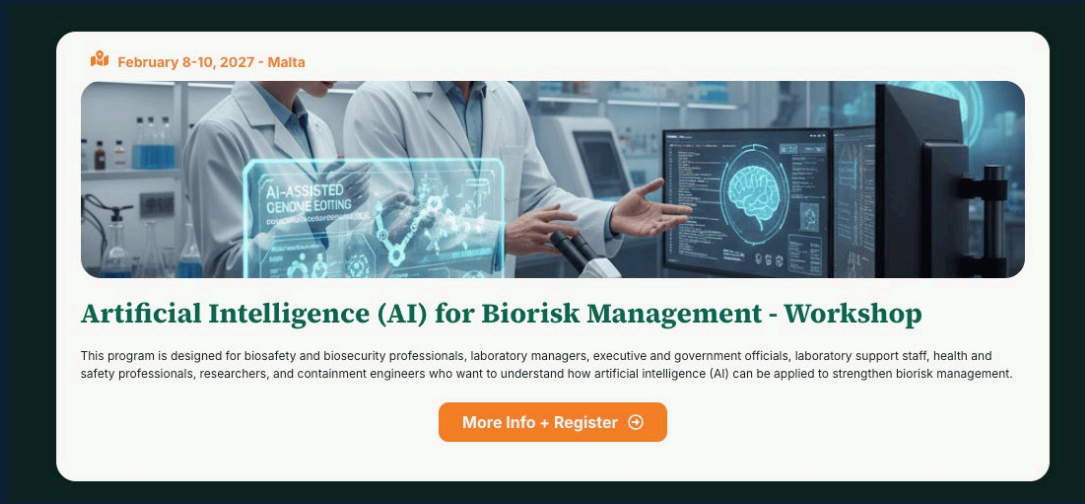
A required protection is missing

After approval: keep monitoring, and retest after important changes.

Next steps

- Work with biorisk AI tools
- Practice and develop governance
- AI cyber(bio)security
- Hands-on with GenAI

3-DAY HANDS-ON WORKSHOP



AI for Biorisk Management

Malta · 8–10 February 2027

Thank you

Questions and discussion

Stefan Wagener, PhD
stefan@bioriskinternational.com
Biorisk Management International (BRMI)

“We are moving biorisk management to the next level”