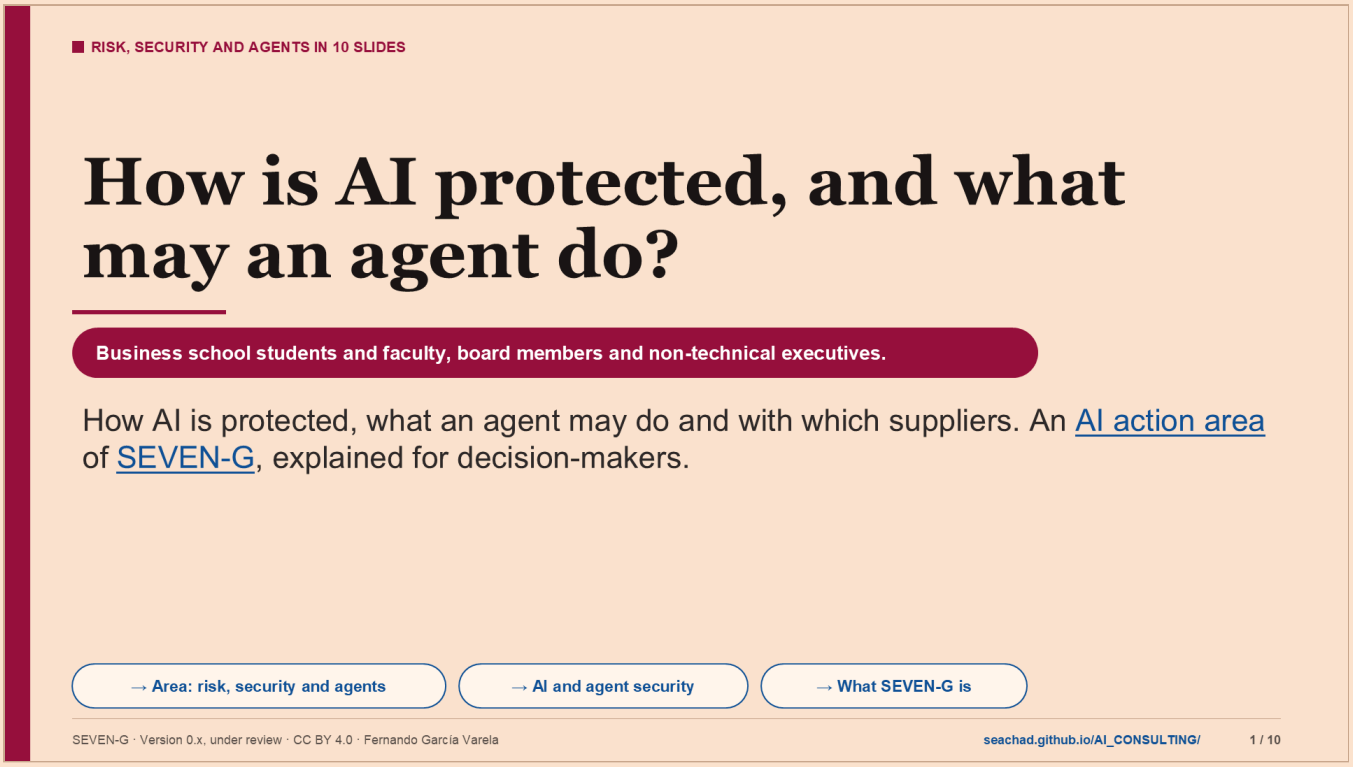


Notes

This presentation explains the Risk, security and agents action area of SEVEN-G, the framework for the governance, value and measurement of AI. It answers three questions: how AI is protected, what an agent may do and which suppliers the company works with.

What it is for. Until recently, a company's AI answered questions or summarised documents. Now agents are appearing: systems that execute actions with tools and permissions. They pay, write to customers, change data or touch production systems. When a system acts, an error or an attack is no longer a bad answer: it becomes an action with effects. That is why AI security is a matter for management and the board, not only for the technical team.

What you will find in the ten slides:

- Why it matters: what AI changes in the company's risk.
- The key concepts, with the glossary definitions (document 02).
- How SEVEN-G treats it: autonomy levels A0 to A3, what is required at each one and who decides.
- At which gate of the lifecycle each thing is checked and with which templates.
- A fictitious case, frequent mistakes, questions for management or the classroom, the relationship with other areas and how to start.

Why it matters. The basic rule of the area is simple: the more autonomy an agent has, the more controls it needs. And the controls are decided before production, not after the first incident.

Question for the classroom: how many systems in your organisation already execute actions on their own, and who knows about it?

TERMS ON THIS SLIDE

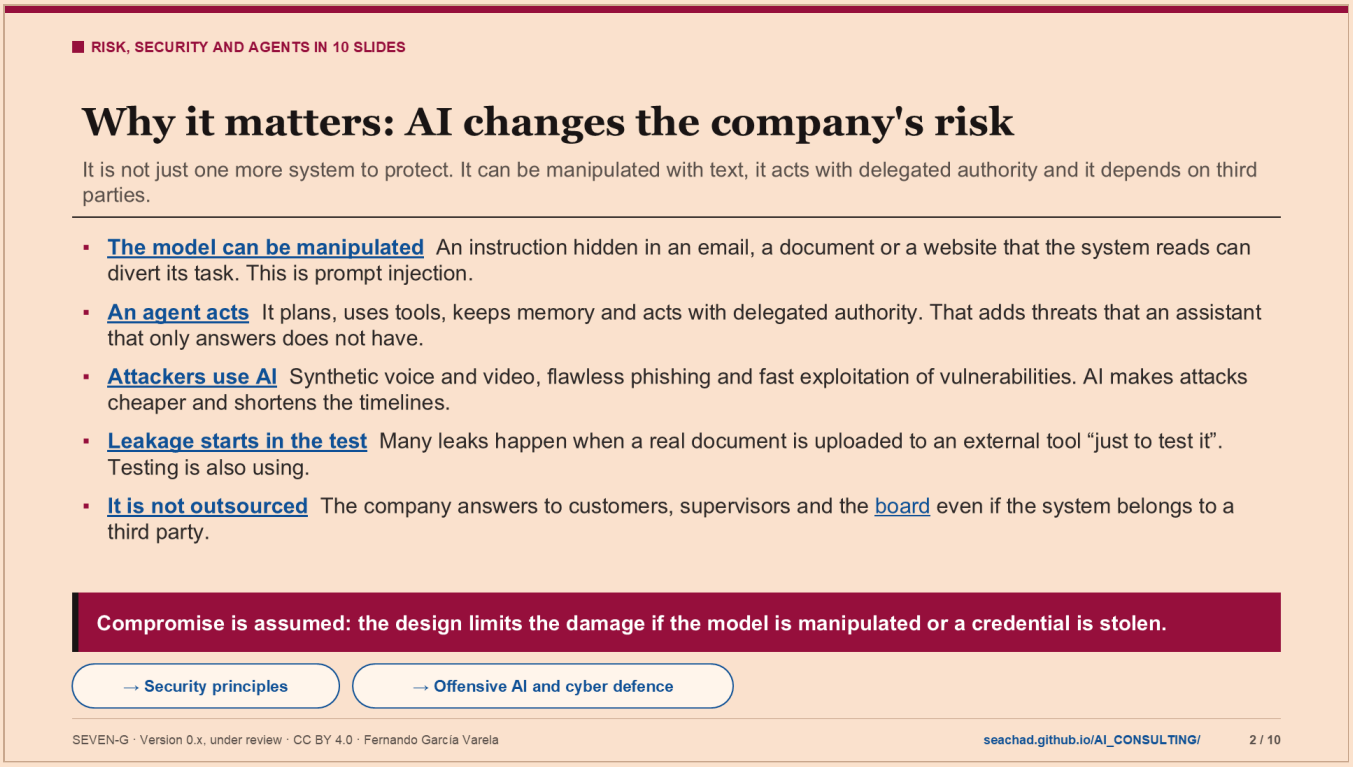
- **AI action areas**: Twelve entry points into the framework by topic: strategy, use cases, value, governance, organisation, maturity, risk, security, standards, board, data and operations, and implementation. Each area

answers a question and leads to its documents, templates, tools and a page of its own. It lets you reach a subject without going through the whole method. The link opens the areas on the SEVEN-G entry page. https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/entrada/index.html#areas

- **SEVEN-G**: SEVEN-G (Seven-phase Enterprise Value & Governance) is an open framework for a company to implement, govern and measure AI, and to know whether it is transforming or only gaining efficiency. It treats each use case as an investment with decision gates and gives the board figures it can trust. Management, the AI office, risk and the board use it. The link opens its entry page, with the story, the questions and the examples. https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/entrada/index.html

SOURCES AND LINKS

- **Area: risk, security and agents** https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/areas/seguridad.html
- **Document 35 · AI and agent security: purpose and scope** https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/35_SEVEN-G_Seguridad_de_IA_y_agentes.html#purpose-and-scope
- **Document 36 · Third parties and AI suppliers** https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/36_SEVEN-G_Terceros_y_proveedores_de_IA.html
- **AI and agent security** https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/35_SEVEN-G_Seguridad_de_IA_y_agentes.html



Notes

AI is not just one more system to protect with the usual rules. It changes the company's risk in five ways.

- The model can be manipulated with text. Prompt injection introduces orders that alter the intended behaviour. It is direct if the user writes it and indirect if it is hidden in what the system reads: documents, emails, websites or tool responses.
- An agent acts. New threats appear: that it pursues a goal different from the one entrusted to it, that it uses more permissions than its task needs or that it induces a person to approve something inappropriate.
- Attackers also use AI. Most attacks are not new, but AI reduces their cost, makes them more convincing and shortens the timelines. One example is CEO fraud with a fake video call that orders an urgent payment.
- Leakage starts in the test. That is why SEVEN-G says that testing is also using: no external tool is used with internal, personal or confidential data without first assessing the supplier.
- Responsibility is not outsourced. Even if the model belongs to a third party, the company is accountable.

The SEVEN-G response starts with eight principles. The most important one for an executive: the model is not a security boundary. The limits that matter (permissions, amounts, recipients and data) are enforced outside the model. And compromise is assumed: the design starts from the idea that one day the model will be manipulated or a credential will be stolen.

Why it matters. If protection depends on the model obeying, a well-written text is enough to bypass it.

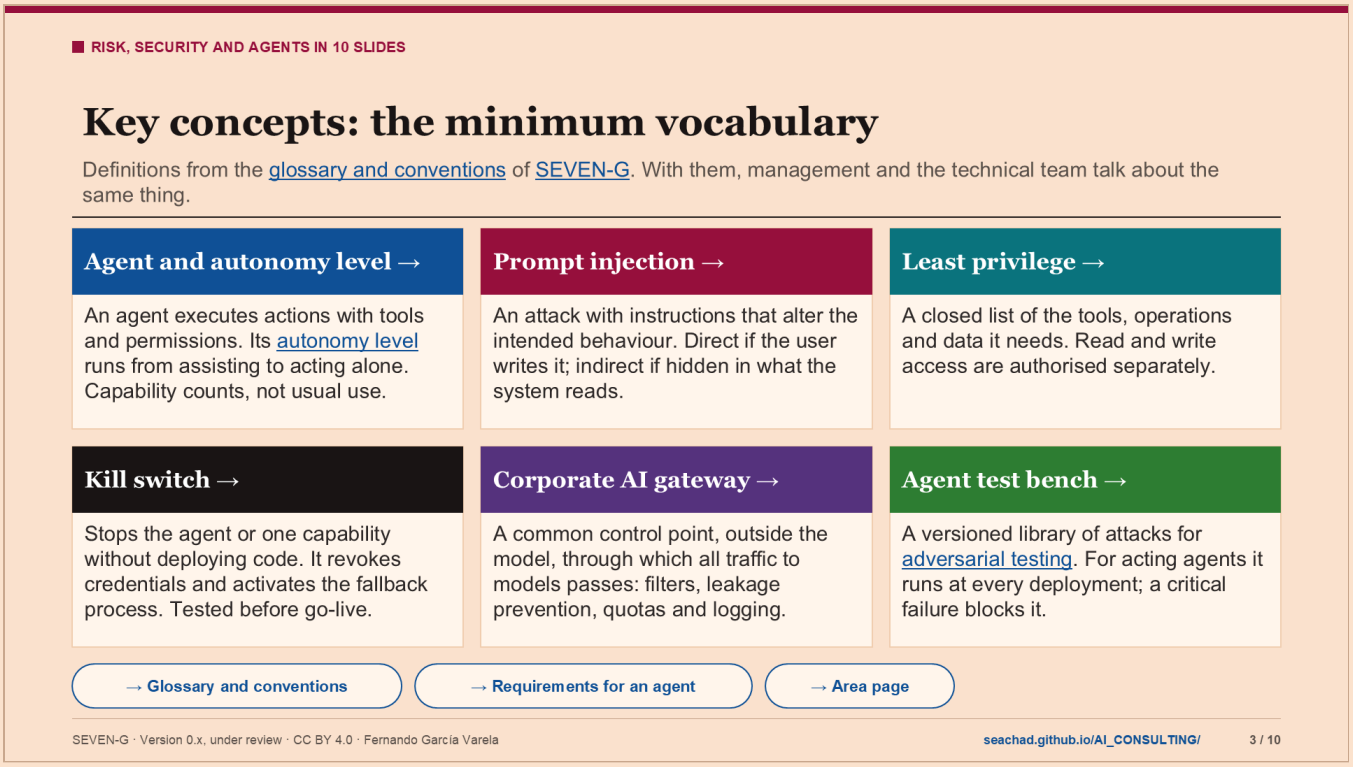
Question for the classroom: does any payment order in your organisation depend only on recognising a voice or a face?

TERMS ON THIS SLIDE

- **Board of directors**: The company's highest body. In SEVEN-G it approves the AI thesis and risk appetite (C2), Transform bets and their scaling, and oversees AI every quarter. AI is a strategic asset: its big decisions are not only for technology. The link opens its role in the governance model (document 30). https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/30_SEVEN-G_Modelo_de_gobierno.html#3-1-board-of-directors

SOURCES AND LINKS

- **Document 35 §2 · Security principles** https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/35_SEVEN-G_Seguridad_de_IA_y_agentes.html#security-principles
- **Document 35 §3 · Threats to AI systems** https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/35_SEVEN-G_Seguridad_de_IA_y_agentes.html#threats-to-ai-systems
- **Document 36 §2 · Principles (testing is also using)** https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/36_SEVEN-G_Terceros_y_proveedores_de_IA.html#principles
- **OWASP Top 10 for LLM Applications (2025)** https://genai.owasp.org/resource/owasp-top-10-for-llm-applications-2025/
- **CCN-CERT BP/36 · Good practices against offensive AI** https://www.ccn-cert.cni.es/es/informes/informes-de-buenas-practicas-bp/7469-ccn-cert-bp-36-buenas-practicas-ia-ofensiva/file.html
- **Security principles** https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/35_SEVEN-G_Seguridad_de_IA_y_agentes.html#security-principles
- **Offensive AI and cyber defence** https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/35_SEVEN-G_Seguridad_de_IA_y_agentes.html#exposure-to-offensive-ai-and-use-of-ai-in-cyber-defence



Notes

Six concepts are enough to follow a conversation about AI security. All of them are defined in the SEVEN-G glossary (document 02).

- Agent and autonomy level. An agent is an AI system that executes actions on other systems, data or people, or prepares them for human validation, by means of tools and permissions. The autonomy level measures how much it acts without prior human validation of each action. The level assigned is the highest of the actions the system can execute, not that of its usual use. An agent that almost never pays, but can pay, is classified as if it paid.
- Prompt injection. It is the typical attack on generative AI. Its controls are demonstrated with adversarial testing, not with statements.
- Least privilege. The agent has only what its task strictly needs, and the permissions are justified in the security design (P18).
- Kill switch. When activated, it revokes credentials, blocks new actions, preserves logs and activates the fallback process. The AI operations owner and security can activate it.
- Corporate AI gateway (control SEG-26). It applies the same measured configuration to every use case. It fails closed for sensitive actions and degraded for queries.
- Corporate agent test bench (control SEG-27). Its scenarios come from the threat model and from the company's own incidents, and they cover all channels: text, documents, email, web, memory, messages between agents and synthetic voice.

A seventh concept completes the area: the AI supplier, the third party that supplies models, platforms or software with AI. It is managed with three requirement levels, from N1 to N3 (document 36).

Why it matters. Without a common vocabulary, “the agent is secure” means nothing. With it, management can ask: what level does it have? who can stop it? does it go through the gateway?

TERMS ON THIS SLIDE

- **Glossary and conventions** (documento 02): The official SEVEN-G glossary: each term with its definition, its Spanish equivalent and the document that develops it, plus the scales and codes used across the framework. It ensures everyone calls the same thing by the same name. The link opens the document. https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/02_SEVEN-G_Glosario_y_convenciones.html
- **SEVEN-G**: SEVEN-G (Seven-phase Enterprise Value & Governance) is an open framework for a company to implement, govern and measure AI, and to know whether it is transforming or only gaining efficiency. It treats each use case as an investment with decision gates and gives the board figures it can trust. Management, the AI office, risk and the board use it. The link opens its entry page, with the story, the questions and the examples. https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/entrada/index.html
- **Autonomy levels** (A0–A3): A scale of how much an AI system acts on its own. A0: only informs; A1: proposes and a person validates each action; A2: acts within limits and a person supervises; A3: acts alone in a defined scope. The more autonomy, the more controls: an A2 or A3 agent with effects on third parties is Enterprise. The link opens the scale in document 35. https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/35_SEVEN-G_Seguridad_de_IA_y_agentes.html#autonomy-levels
- **Adversarial testing** (P53): Tests in which a team tries to trick or break an AI system as an attacker would: prompt injection, data leakage, abuse of an agent's tools. A critical scenario in which the attacker succeeds prevents deployment. They are planned and reported with template P53. The link opens the template. https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/plantillas/P53_SEVEN-G_Plan_e_informe_de_pruebas_adversarias.html

SOURCES AND LINKS

- **Glossary (document 02) · Autonomy of systems and agents** https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/02_SEVEN-G_Glosario_y_convenciones.html#4-11-autonomy-of-systems-and-agents
- **Document 35 §4 · Essential requirements for an agent** https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/35_SEVEN-G_Seguridad_de_IA_y_agentes.html#essential-requirements-for-an-agent
- **Document 35 §6 · AI security controls (SEG)** https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/35_SEVEN-G_Seguridad_de_IA_y_agentes.html#catalogue-of-ai-security-controls-seg
- **Document 36 §4 · Requirement levels N1–N3** https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/36_SEVEN-G_Terceros_y_proveedores_de_IA.html#requirement-levels-n1-n3
- **Requirements for an agent** https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/35_SEVEN-G_Seguridad_de_IA_y_agentes.html#essential-requirements-for-an-agent
- **Area page** https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/areas/seguridad.html



RISK, SECURITY AND AGENTS IN 10 SLIDES

How SEVEN-G treats it: controls by autonomy

What the system can do is what counts. A critical control not designed blocks the design gate; not tested, it blocks go-live.

Level	Role of the person	What is required	Who decides
A0 · Assistance	Decides and executes.	Least privilege, secure credentials and external content treated as data.	Proposed in the feasibility and risk phase and set in the solution design phase.
A1 · Recommendation	Validates each action before it is executed.	Also: own identity, action log, adversarial testing and human validation of every action.	Recorded in the AI system inventory. Moving up a level is a relevant change.
A2 · Supervised action	Supervises, can interrupt and reviews after the fact.	Also: intent control, limits, a tested kill switch and the test bench at every deployment.	With effects on third parties, money, personal data or production: Enterprise intensity.
A3 · Autonomous action	Sets limits, supervises aggregates and can stop it.	Everything above, with more frequent tests and reviews (kill switch, every quarter).	The AI committee, at the design and go-live gates, states why supervision is not enough.

Proceed with conditions is not permitted for a critical control. Disabling one in production is a critical nonconformity.

→ Autonomy levels | → Critical controls | → Enterprise criteria

SEVEN-G · Version 0.x, under review · CC BY 4.0 · Fernando García Varela seachad.github.io/AI_CONSULTING/ 4 / 10

Notes

The table summarises the central rule of the area: the more autonomy, the more controls. Each row adds requirements to the previous one.

- A0 · Assistance. The system informs, summarises or generates content; the person decides and executes. Least privilege, credential management and treating all external content as data are mandatory. If it accesses corporate systems with its own credentials, its own identity and approved tools are also mandatory.
- A1 · Recommendation. The system proposes a specific action and the person validates each one before it is executed. Its own identity, traceability between action and intent, the action log and adversarial testing are added.
- A2 · Supervised action. The system executes actions within defined limits. All applicable agent controls are required: a decision point outside the model that checks each action, amount and volume limits, a kill switch tested every six months and the test bench at every deployment.
- A3 · Autonomous action. The system executes sequences without individual review. The kill switch is tested every quarter and permissions are reviewed every quarter.

Who decides and when. The level is proposed in phase 3 (feasibility), set in phase 4 (design) and recorded in the inventory. An A2 or A3 system that affects third parties, money, personal data or production meets the criterion of “agents with the ability to act” and goes with Enterprise intensity. An A3 system needs a decision by the AI Committee at the design (G4) and go-live (G5) gates. After a serious incident related to autonomy, the system moves down a level or is stopped until the analysis.

Why it matters. This rule turns “trusting AI” into a concrete decision that is approved, supervised and audited.

Question for the classroom: why does SEVEN-G classify by what the system can do and not by what it usually does?

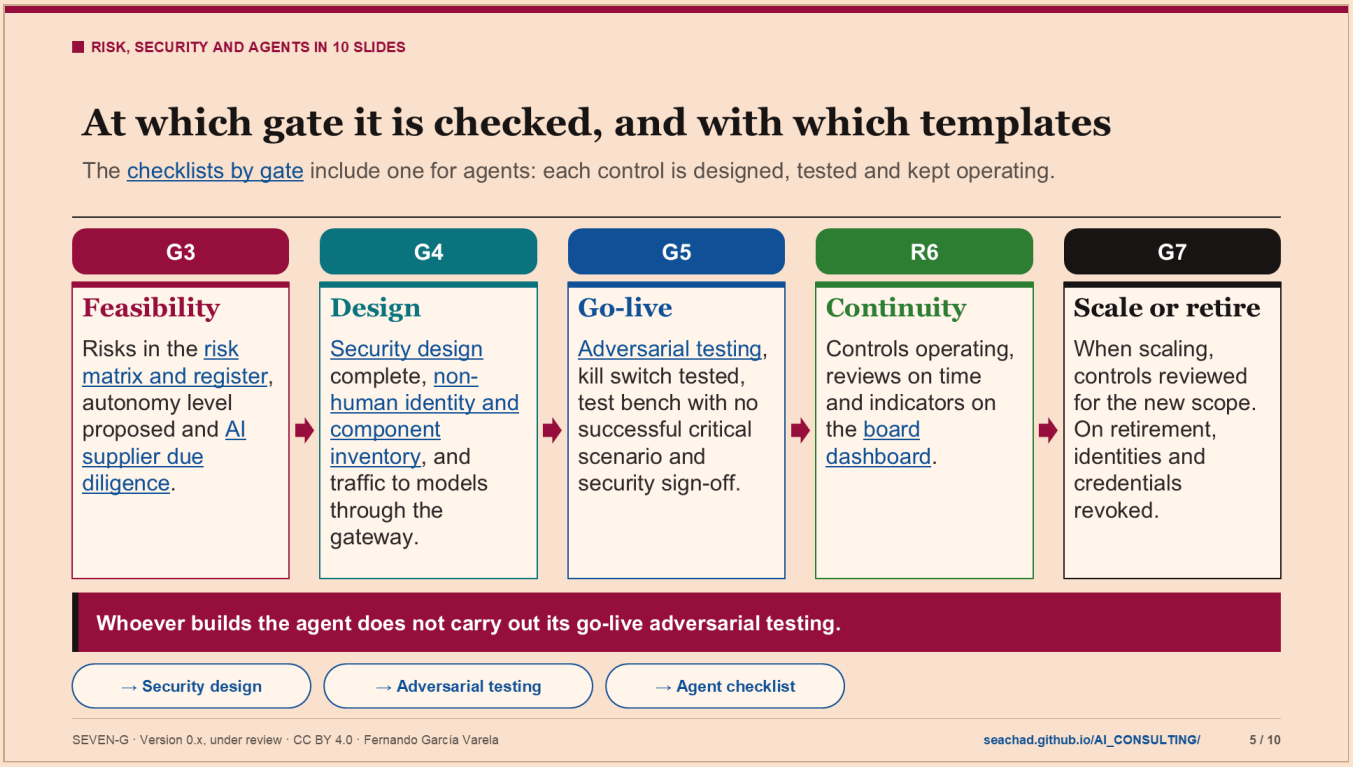
TERMS ON THIS SLIDE

- **Design gate** (G4): The gate that approves the full design before building: architecture, data, security, human oversight, rollback and adoption. It checks that the controls are planned from the start. The link opens its criteria in document 21. https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/21_SEVEN-G_Criterios_de_gate_y_auditoria.html#6-5-g4-design
- **Go-live gate** (G5): The gate that authorises using the AI system with real users, customers or data. It requires validation and pilot passed, a tested rollback plan and multi-level sign-off (business, technology, risk). Nothing goes into production without this sign-off. The link opens its criteria in document 21. https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/21_SEVEN-G_Criterios_de_gate_y_auditoria.html#6-6-g5-go-live
- **Feasibility and risk** (fase 3): The phase in which a limited test checks whether the use case is technically and economically feasible, it is classified under regulation and its risks are recorded. It ends at the feasibility gate (G3), the main stop gate before the big spend. It is where stopping in time saves the most money. The link opens the phase manual in document 20. https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/20_SEVEN-G_Manuales_de_fase.html#phase-3-feasibility-and-risk
- **Solution design** (fase 4): The phase in which the full solution is designed: architecture, data, security, human oversight, rollback plan and adoption plan. It ends at the design gate (G4). It ensures controls are designed before building, not added later when they cost more. The link opens the phase manual in document 20. https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/20_SEVEN-G_Manuales_de_fase.html#phase-4-solution-design
- **AI system inventory** (T02): A list of all the company's AI systems: its own, third-party AI embedded in processes, the assistants employees use and any unauthorised use detected. Each system has an owner, a regulatory classification, an intensity and an autonomy level. It comes first, because you cannot govern what you do not know. It is a view of the use case register: open the link and choose «Inventory» (fictitious data). https://seachad.github.io/AI_CONSULTING/SEVEN-G/herramientas/T01_registro_iniciativas/registro.html
- **Agent test bench** (SEG-27): A common, versioned library of simulated attacks, drawn from the threat model and the company's own incidents, that tests agents and assistants across all their channels, tools and connectors. For agents that act it runs at every deployment, and if a critical scenario succeeds the deployment does not go ahead. The link opens the control in document 35. https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/35_SEVEN-G_Seguridad_de_IA_y_agentes.html#cod-seg-27
- **AI committee**: A management body that decides the portfolio and the gates at its level, allocates budget and escalates to the board what belongs there. It brings together business, technology, finance and risk, and follows rules of procedure (P38). It matters because it turns scattered decisions into a single portfolio. The link opens its role in the governance model (document 30). https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/30_SEVEN-G_Modelo_de_gobierno.html#3-3-ai-committee
- **Nonconformity** (P50): A breach of a framework rule or a control, even if it caused no harm (for example, a use case in production without sign-off). It is classed as minor, major or critical, recorded (P50), its root cause analysed (P52) and corrected on time. It lets governance learn from its failures. The link opens the process in document 37. https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/37_SEVEN-G_No_conformidades_e_incidentes.html#nonconformity-process

SOURCES AND LINKS

- **Document 35 §5 · Autonomy levels** https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/35_SEVEN-G_Seguridad_de_IA_y_agentes.html#autonomy-levels
- **Document 35 §5.3 · Minimum controls per level** https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/35_SEVEN-G_Seguridad_de_IA_y_agentes.html#5-3-minimum-controls-per-level
- **Document 35 §7 · Agent controls (AG)** https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/35_SEVEN-G_Seguridad_de_IA_y_agentes.html#catalogue-of-agent-controls-ag
- **Document 01 §9.2 · Enterprise criteria** https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/01_SEVEN-G_Metodologia_fundacional.html#9-2-enterprise-criteria
- **Autonomy levels** https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/35_SEVEN-G_Seguridad_de_IA_y_agentes.html#autonomy-levels
- **Critical controls** https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/35_SEVEN-G_Seguridad_de_IA_y_agentes.html#5-4-critical-controls

- **Enterprise criteria** https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/01_SEVEN-G_Metodologia_fundacional.html#9-2-enterprise-criteria



Notes

In SEVEN-G each use case goes through phases separated by decision gates. Security is checked at five of them, always with evidence.

- Feasibility gate (G3). Generative AI and security risks are assessed in the risk matrix (P12). The autonomy level is proposed. If there is a supplier, its security is assessed with the due diligence questionnaire (P55).
- Design gate (G4). The security design (P18) records the threat model, the autonomy level, the identities, the permission matrix, intent control, the limits, the sensitive actions, the kill switch and the test plan. The agents' identities are entered, with their human owner, in the inventory of non-human identities (P54). Traffic to models goes through the corporate gateway.
- Go-live gate (G5). Adversarial testing (red teaming) is carried out and documented in P53. The kill switch is tested. The test bench cannot have any successful critical scenario. Information security signs within the multi-level sign-off.
- Continuity review (R6). It is checked that the controls are still operating and that campaigns and reviews are on time. The indicators reach the board dashboard (T17).
- Scale or retire gate (G7). When scaling, controls are reviewed for the new scope. On retirement, no orphaned identities or credentials may remain.

Who does what. The AI Technical Owner designs and implements the controls. Information security reviews them. The AI Risk Owner assesses the residual risk. The AI Auditor verifies the evidence with the LV-AG checklist. The agent security assessment (T10) is applied with the security design (P18). The AI security profile (P72) gives the overall view in the corporate cycle.

Why it matters. Separating whoever builds from whoever tests prevents anyone from approving their own work.

TERMS ON THIS SLIDE

- **Checklists by gate** (documento 22): Practical checklists to prepare and review each gate and some cross-cutting topics (evidence, agents). They turn the criteria of document 21 into yes-or-no questions answered before the session. The link opens the document. https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/22_SEVEN-G_Listas_de_verificacion_por_gate.html
- **Risk matrix and register** (T06): A likelihood-by-impact matrix (5 × 5) and register of each use case's risks: inherent and residual risk, controls and their effectiveness, response, contingency and who accepts the risk. It is mandatory from phase 3. It is filled in with template P12 or in the «Risks» view of the use case register, which also downloads an Excel file for the use case with proposed mitigation plans. Fictitious data. https://seachad.github.io/AI_CONSULTING/SEVEN-G/herramientas/T01_registro_iniciativas/registro.html
- **AI supplier due diligence** (P55): A twelve-block questionnaire to assess an AI supplier before contracting: security, data, models, continuity, compliance, subcontractors and exit. A nonconforming result in the critical blocks prevents using it with production data. The link opens the template. https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/plantillas/P55_SEVEN-G_Cuestionario_de_diligencia_debida_a_proveedores_de_IA.html
- **Security design** (P18): A template to design the solution's security before building it, agents included: threats, controls, identities, permissions, logs and planned tests. It is reviewed at the design gate. Adding security afterwards costs more and protects less. The link opens the template. https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/plantillas/P18_SEVEN-G_Diseno_de_seguridad.html
- **Non-human identity and component inventory** (P54): A template listing the identities used by AI agents and services (accounts, keys, permissions) and the software components of each solution. An agent with too many permissions is an open door; this list makes it possible to review and remove them. The link opens the template. https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/plantillas/P54_SEVEN-G_Inventario_de_identidades_no_humanas_y_componentes.html
- **Adversarial testing** (P53): Tests in which a team tries to trick or break an AI system as an attacker would: prompt injection, data leakage, abuse of an agent's tools. A critical scenario in which the attacker succeeds prevents deployment. They are planned and reported with template P53. The link opens the template. https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/plantillas/P53_SEVEN-G_Plan_e_informe_de_pruebas_adversarias.html
- **Board dashboard** (T17): A supervision dashboard generated from the use case register. It shows the board cost, validated versus declared value, the funnel, risks, incidents, adoption, maturity and what holds back scaling, with a «?» help on each card. It is for deciding, not just for informing. The link opens the example dashboard with fictitious data; the interface is in Spanish and the help texts are in both languages. https://seachad.github.io/AI_CONSULTING/SEVEN-G/herramientas/T17_panel_consejo/ejemplo/salida/t01_Dashboard_Casos_Uso_IA_v8.html

SOURCES AND LINKS

- **Document 35 §11 · Integration with the lifecycle** https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/35_SEVEN-G_Seguridad_de_IA_y_agentes.html#integration-with-the-lifecycle
- **Document 35 §12 · Associated tools and templates** https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/35_SEVEN-G_Seguridad_de_IA_y_agentes.html#associated-tools-and-templates
- **Template P72 · AI security profile** https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/plantillas/P72_SEVEN-G_Perfil_de_seguridad_de_IA_CSF_y_Cyber_AI_Profile.html
- **Board dashboard (T17)** https://seachad.github.io/AI_CONSULTING/SEVEN-G/herramientas/T17_panel_consejo/index.html
- **Agent checklist** https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/22_SEVEN-G_Listas_de_verificacion_por_gate.html#lv-ag-agents



Notes

The case is fictitious, like all the cases in document 92 (application cases): companies, people and figures are illustrative. A logistics operator wants to open a new service: booking on-demand transport capacity for its customers with an agent that buys loads on freight exchanges. It is a Transform bet, with Enterprise intensity.

How the area is applied, step by step:

- Risk appetite. At the direction stage (C2) the board sets how far it goes: agents with supervised action (A2), with a financial limit per transaction approved by the committee, and no A3 in transactions involving third-party money. Security starts at the board, before the agent exists.
- Autonomy level. The agent acts alone up to €2,500 per load. Above that, a person decides. It is an action limit enforced outside the model, not an instruction the model could forget.
- Design gate (G4). The AI Committee checks that the limit per load, the action log and the kill switch are designed.
- Go-live gate (G5). A pilot with random assignment (520 loads with the agent and 510 in the control group) gives a margin improvement of 0.6 points, below the 1.0 set as the stop criterion. Errors: 0.8 %. The committee stops the use case. The team asked for another iteration with a lower threshold and the board decided not to relax the criteria.

What it teaches. Security controls do not make a use case good: they let it work without harm if its value is demonstrated. Here the value was not demonstrated, and the company stopped with €165,000 invested and a recorded lesson. The image shows the lifecycle, with its phases and gates.

Question for the classroom: if the pilot had reached the margin, what would have had to be tested before putting the agent into production?

TERMS ON THIS SLIDE

- **Board of directors**: The company's highest body. In SEVEN-G it approves the AI thesis and risk appetite (C2), Transform bets and their scaling, and oversees AI every quarter. AI is a strategic asset: its big decisions are not only for technology. The link opens its role in the governance model (document 30). https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/30_SEVEN-G_Modelo_de_gobierno.html#3-1-board-of-directors
- **Design gate** (G4): The gate that approves the full design before building: architecture, data, security, human oversight, rollback and adoption. It checks that the controls are planned from the start. The link opens its criteria in document 21. https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/21_SEVEN-G_Criterios_de_gate_y_auditoria.html#6-5-g4-design
- **AI committee**: A management body that decides the portfolio and the gates at its level, allocates budget and escalates to the board what belongs there. It brings together business, technology, finance and risk, and follows rules of procedure (P38). It matters because it turns scattered decisions into a single portfolio. The link opens its role in the governance model (document 30). https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/30_SEVEN-G_Modelo_de_gobierno.html#3-3-ai-committee

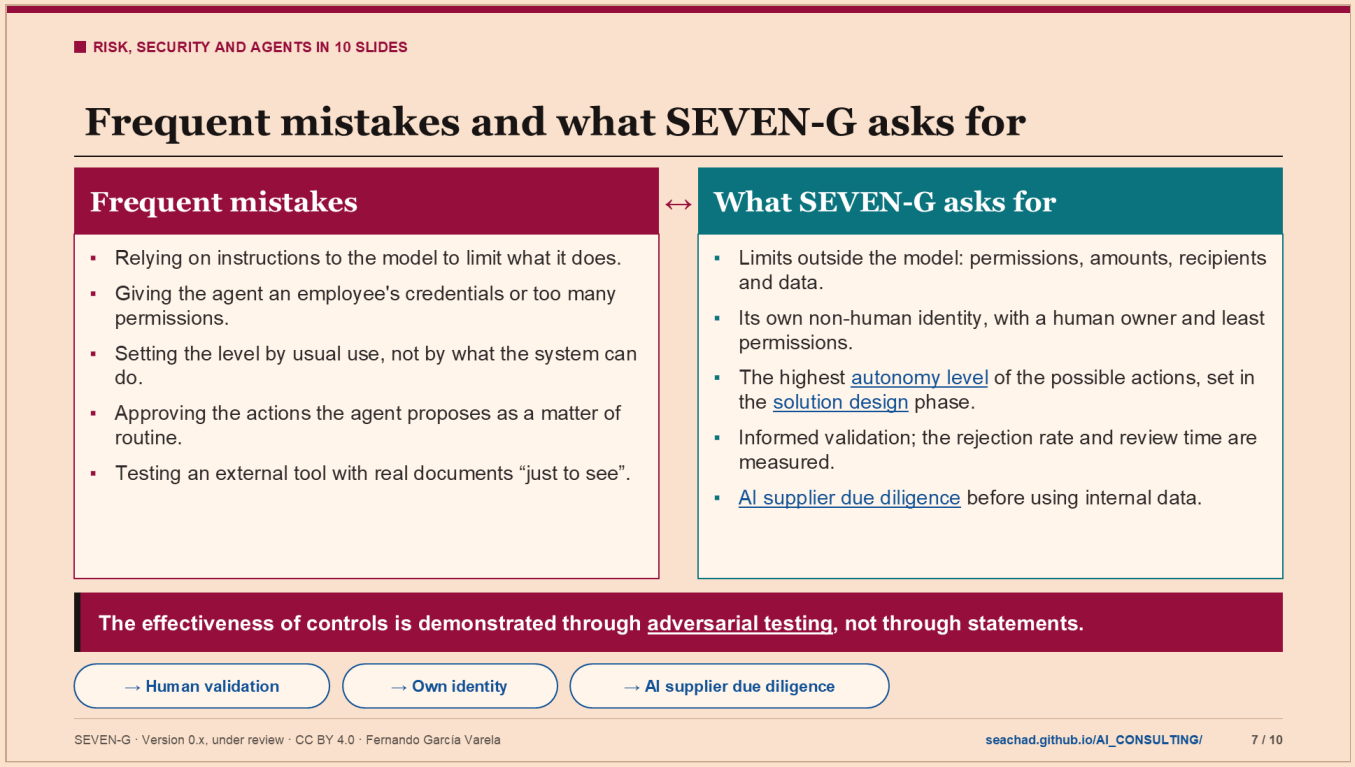
LINKS IN THE CHART

- Proposed · phases 0 and 1: https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/00_SEVEN-G_Que_es_y_para_que_sirve.html#4-3-use-case-level-lifecycle-with-decision-gates
- Value hypothesis · phase 2: https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/00_SEVEN-G_Que_es_y_para_que_sirve.html#4-3-use-case-level-lifecycle-with-decision-gates
- Proof of concept · phase 3: https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/00_SEVEN-G_Que_es_y_para_que_sirve.html#4-3-use-case-level-lifecycle-with-decision-gates
- In development · phases 4 and 5: https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/00_SEVEN-G_Que_es_y_para_que_sirve.html#4-3-use-case-level-lifecycle-with-decision-gates
- In use · phases 6 and 7 · or disengaged: https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/00_SEVEN-G_Que_es_y_para_que_sirve.html#4-3-use-case-level-lifecycle-with-decision-gates
- Phase 0 · Context and constraints: https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/20_SEVEN-G_Manuales_de_fase.html#phase-0-context-and-constraints
- Phase 1 · Opportunity discovery: https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/20_SEVEN-G_Manuales_de_fase.html#phase-1-opportunity-discovery
- Phase 2 · Value hypothesis: https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/20_SEVEN-G_Manuales_de_fase.html#phase-2-value-hypothesis
- Phase 3 · Feasibility and risk: https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/20_SEVEN-G_Manuales_de_fase.html#phase-3-feasibility-and-risk
- Phase 4 · Solution design: https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/20_SEVEN-G_Manuales_de_fase.html#phase-4-solution-design
- Phase 5 · Delivery and validation: https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/20_SEVEN-G_Manuales_de_fase.html#phase-5-delivery-and-validation
- Phase 6 · Operation and governance: https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/20_SEVEN-G_Manuales_de_fase.html#phase-6-operation-and-governance
- Phase 7 · Evolution or retirement: https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/20_SEVEN-G_Manuales_de_fase.html#phase-7-evolution-or-retirement
- G0 · Authorisation: https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/21_SEVEN-G_Criterios_de_gate_y_auditoria.html#6-1-g0-authorisation
- Checklist LV-G0: https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/22_SEVEN-G_Listas_de_verificacion_por_gate.html#lv-g0-authorisation
- G1 · Opportunity: https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/21_SEVEN-G_Criterios_de_gate_y_auditoria.html#6-2-g1-opportunity

- Checklist LV-G1: https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/22_SEVEN-G_Listas_de_verificacion_por_gate.html#lv-g1-opportunity
- G2 · Hypothesis: https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/21_SEVEN-G_Criterios_de_gate_y_auditoria.html#6-3-g2-hypothesis
- Checklist LV-G2: https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/22_SEVEN-G_Listas_de_verificacion_por_gate.html#lv-g2-hypothesis
- G3 · Feasibility: https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/21_SEVEN-G_Criterios_de_gate_y_auditoria.html#6-4-g3-feasibility
- Checklist LV-G3: https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/22_SEVEN-G_Listas_de_verificacion_por_gate.html#lv-g3-feasibility
- G4 · Design: https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/21_SEVEN-G_Criterios_de_gate_y_auditoria.html#6-5-g4-design
- Checklist LV-G4: https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/22_SEVEN-G_Listas_de_verificacion_por_gate.html#lv-g4-design
- G5 · Go-live: https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/21_SEVEN-G_Criterios_de_gate_y_auditoria.html#6-6-g5-go-live
- Checklist LV-G5: https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/22_SEVEN-G_Listas_de_verificacion_por_gate.html#lv-g5-go-live
- Multi-level sign-off: https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/21_SEVEN-G_Criterios_de_gate_y_auditoria.html#multi-level-go-live-sign-off-g5
- R6 · Continuity: https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/21_SEVEN-G_Criterios_de_gate_y_auditoria.html#6-7-r6-continuity-review
- Checklist LV-R6: https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/22_SEVEN-G_Listas_de_verificacion_por_gate.html#lv-r6-continuity-review
- G7 · Scale or retire: https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/21_SEVEN-G_Criterios_de_gate_y_auditoria.html#6-8-g7-scale-or-retire
- Checklist LV-G7: https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/22_SEVEN-G_Listas_de_verificacion_por_gate.html#lv-g7-scale-or-retire

SOURCES AND LINKS

- **Document 92 · Case 2, logistics operator L (fictitious)** https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/92_SEVEN-G_Casos_de_aplicacion.html#case-2-logistics-operator-l-fictitious
- **Document 35 §4.5 · Action limits** https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/35_SEVEN-G_Seguridad_de_IA_y_agentes.html#4-5-action-limits
- **Document 35 §4.7 · Kill switch** https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/35_SEVEN-G_Seguridad_de_IA_y_agentes.html#4-7-kill-switch
- **Case 2 · Logistics operator** https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/92_SEVEN-G_Casos_de_aplicacion.html#case-2-logistics-operator-l-fictitious
- **Action limits** https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/35_SEVEN-G_Seguridad_de_IA_y_agentes.html#4-5-action-limits



Notes

Five frequent mistakes, each with the SEVEN-G rule that prevents it.

- Relying on instructions to the model. Telling the model “do not make payments of more than a thousand euros” helps, but an attacker can bypass it with another instruction. SEVEN-G enforces limits outside the model, with an independent decision point that blocks whatever does not fit.
- An employee's credentials or too many permissions. If the agent uses a person's account, nobody knows who did what. SEVEN-G requires its own identity, registered with its human owner, and a closed list of permissions. When it acts on behalf of a user, it never has more permissions than that user.
- Setting the level by usual use. A system that almost always proposes, but can execute, is an acting system. It is classified by what it can do.
- Routine approval. If a person approves a hundred actions a day without reading them, human validation does not protect. SEVEN-G asks for informed validation: the person sees which action is proposed, with which data, why and with what effects, and can reject it at no cost to themselves. The rejection rate and review time are measured to detect routine approval.
- Testing with real documents. No external tool is used with internal, personal or confidential data without the supplier assessment proportionate to its level. A test with synthetic or public data does not need it.

Why it matters. All these mistakes have something in common: they give a feeling of control that nobody has checked. SEVEN-G asks for it to be demonstrated with adversarial testing.

Question for the classroom: which of these five mistakes do you think is most frequent in your organisation?

TERMS ON THIS SLIDE

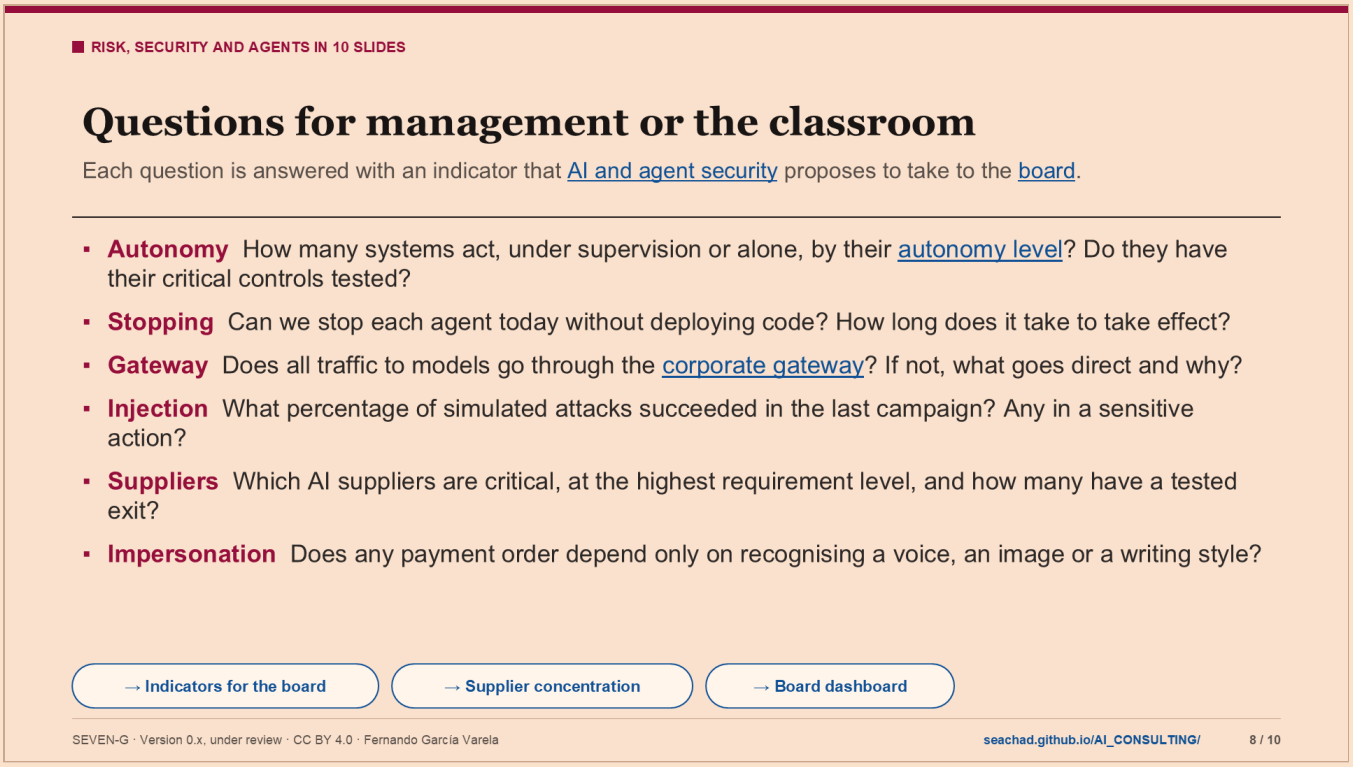
- **Autonomy levels** (A0–A3): A scale of how much an AI system acts on its own. A0: only informs; A1: proposes and a person validates each action; A2: acts within limits and a person supervises; A3: acts alone in a defined scope. The more autonomy, the more controls: an A2 or A3 agent with effects on third parties is Enterprise. The

link opens the scale in document 35. https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/35_SEVEN-G_Seguridad_de_IA_y_agentes.html#autonomy-levels

- **Solution design** (fase 4): The phase in which the full solution is designed: architecture, data, security, human oversight, rollback plan and adoption plan. It ends at the design gate (G4). It ensures controls are designed before building, not added later when they cost more. The link opens the phase manual in document 20. https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/20_SEVEN-G_Manuales_de_fase.html#phase-4-solution-design
- **AI supplier due diligence** (P55): A twelve-block questionnaire to assess an AI supplier before contracting: security, data, models, continuity, compliance, subcontractors and exit. A nonconforming result in the critical blocks prevents using it with production data. The link opens the template. https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/plantillas/P55_SEVEN-G_Cuestionario_de_diligencia_debida_a_proveedores_de_IA.html
- **Adversarial testing** (P53): Tests in which a team tries to trick or break an AI system as an attacker would: prompt injection, data leakage, abuse of an agent's tools. A critical scenario in which the attacker succeeds prevents deployment. They are planned and reported with template P53. The link opens the template. https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/plantillas/P53_SEVEN-G_Plan_e_informe_de_pruebas_adversarias.html

SOURCES AND LINKS

- **Document 35 §2 · Security principles** https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/35_SEVEN-G_Seguridad_de_IA_y_agentes.html#security-principles
- **Document 35 §4.4 · Intent-based access control** https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/35_SEVEN-G_Seguridad_de_IA_y_agentes.html#4-4-intent-based-access-control
- **Document 35 §4.6 · Human validation of sensitive actions** https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/35_SEVEN-G_Seguridad_de_IA_y_agentes.html#4-6-human-validation-of-sensitive-actions
- **Document 36 §2 · Principles (testing is also using)** https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/36_SEVEN-G_Terceros_y_proveedores_de_IA.html#principles
- **Human validation** https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/35_SEVEN-G_Seguridad_de_IA_y_agentes.html#4-6-human-validation-of-sensitive-actions
- **Own identity** https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/35_SEVEN-G_Seguridad_de_IA_y_agentes.html#4-1-own-identity



Notes

These six questions are useful for a board session, a management committee or a classroom case. Each one is backed by an indicator from document 35 (section 10) or document 36 (section 7), with its expected reading.

- Autonomy. Indicators: systems by autonomy level and critical agent controls. Reading: A2 and A3 systems may grow, but with 100 % of their critical controls tested and operating. Any deviation is explained.
- Stopping. Indicator: kill switch test. Reading: 100 % of A2 and A3 systems with a current test and a measured time to take effect.
- Gateway. Indicator: traffic through the corporate gateway. Reading: 100 %; direct traffic is explained and regularised.
- Injection. Indicator: success of prompt injection in the last adversarial testing campaign. Reading: a downward trend and zero in sensitive actions.
- Suppliers. Concentration indicators: systems per supplier, value that depends on each one and critical suppliers without a tested exit. For N3 suppliers, the exit is tested at least once a year or before renewal.
- Impersonation. Governance measure: no payment order depends only on recognising a voice, an image or a writing style. At least one annual drill of fraud with synthetic impersonation is run, involving finance and management.

How to use them in the classroom. Ask students to choose a company they know and answer with what they know. The usual finding is that several answers are “we don't know”. That finding is already the first result.

Why it matters. The board does not need to understand how a model works. It needs to ask the right questions and require a figure behind each answer. In SEVEN-G those indicators reach the board dashboard (T17).

TERMS ON THIS SLIDE

- **AI and agent security** (documento 35): The threats specific to AI (prompt injection, data leakage, poisoning, agent abuse) and their controls: autonomy levels, a corporate gateway, an agent test bench, non-human identities and the use of AI in cyber defence. The link opens the document. https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/35_SEVEN-G_Seguridad_de_IA_y_agentes.html
- **Board of directors**: The company's highest body. In SEVEN-G it approves the AI thesis and risk appetite (C2), Transform bets and their scaling, and oversees AI every quarter. AI is a strategic asset: its big decisions are not only for technology. The link opens its role in the governance model (document 30). https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/30_SEVEN-G_Modelo_de_gobierno.html#3-1-board-of-directors
- **Autonomy levels** (A0–A3): A scale of how much an AI system acts on its own. A0: only informs; A1: proposes and a person validates each action; A2: acts within limits and a person supervises; A3: acts alone in a defined scope. The more autonomy, the more controls: an A2 or A3 agent with effects on third parties is Enterprise. The link opens the scale in document 35. https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/35_SEVEN-G_Seguridad_de_IA_y_agentes.html#autonomy-levels
- **Corporate AI gateway** (SEG-26): A common control point, outside the model, through which all traffic to AI models passes, in-house or third-party: input, output and context filters, leakage prevention, quotas, approved models and logging; for agents, control of each action against an authorised intent. It has the same measured configuration for every use case. The link opens the control in document 35. https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/35_SEVEN-G_Seguridad_de_IA_y_agentes.html#cod-seg-26

SOURCES AND LINKS

- **Document 35 §10 · Indicators for the board** https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/35_SEVEN-G_Seguridad_de_IA_y_agentes.html#indicators-for-the-board
- **Document 35 §9.2 · Governance measures against offensive AI** https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/35_SEVEN-G_Seguridad_de_IA_y_agentes.html#9-2-governance-measures
- **Document 36 §7 · Technological dependency and concentration** https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/36_SEVEN-G_Terceros_y_proveedores_de_IA.html#technological-dependency-and-concentration
- **Indicators for the board** https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/35_SEVEN-G_Seguridad_de_IA_y_agentes.html#indicators-for-the-board
- **Supplier concentration** https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/36_SEVEN-G_Terceros_y_proveedores_de_IA.html#7-2-concentration
- **Board dashboard** https://seachad.github.io/AI_CONSULTING/SEVEN-G/herramientas/T17_panel_consejo/index.html



Notes

The area hosts two sub-areas, which keep their own page and their own slides: "Risk and regulation" and "Standards". The first two cards are those sub-areas.

AI security does not stand alone. It relies on other SEVEN-G areas, on the rules for building the solution and on SPHERES.

- Risk and regulation. Security threats have their own typical risks (generative AI and security) that are assessed with the probability by impact scale of document 33. The EU AI Act requires high-risk systems to have an appropriate level of accuracy, robustness and cybersecurity, including resilience against manipulation of data and models.
- Standards. Each SEVEN-G security control states the NIST CSF 2.0 function it contributes to: govern, identify, protect, detect, respond and recover. With that, the AI security profile (P72) is built. NIST's AI profile, the Cyber AI Profile, is still a preliminary draft: it is used only as guidance and never underpins a gate criterion.
- Data, build and operations. The operations manual (document 52) runs monitoring, the kill switch and continuous evaluations, and measures guardrail effectiveness. AI security incidents follow the nonconformities and incidents circuit (document 37).
- Security when building. SEVEN-G document 53 requires security, secrets and dependency analysis on AI-generated code before go-live (G5); in Enterprise, at every change. Dependencies proposed by the AI are checked before they are used. Agents that build operate with least privilege and without production credentials, and their autonomy is classified with the A0–A3 scale of document 35. A critical or high finding blocks G5, unless the risk is formally accepted.
- SPHERES. It is the explanatory layer of the spheres. In sphere 07 (Decision) it uses the A0–A3 scale to decide how much is delegated in each type of decision. It reminds us that the autonomy level is not the same as the ambition level.

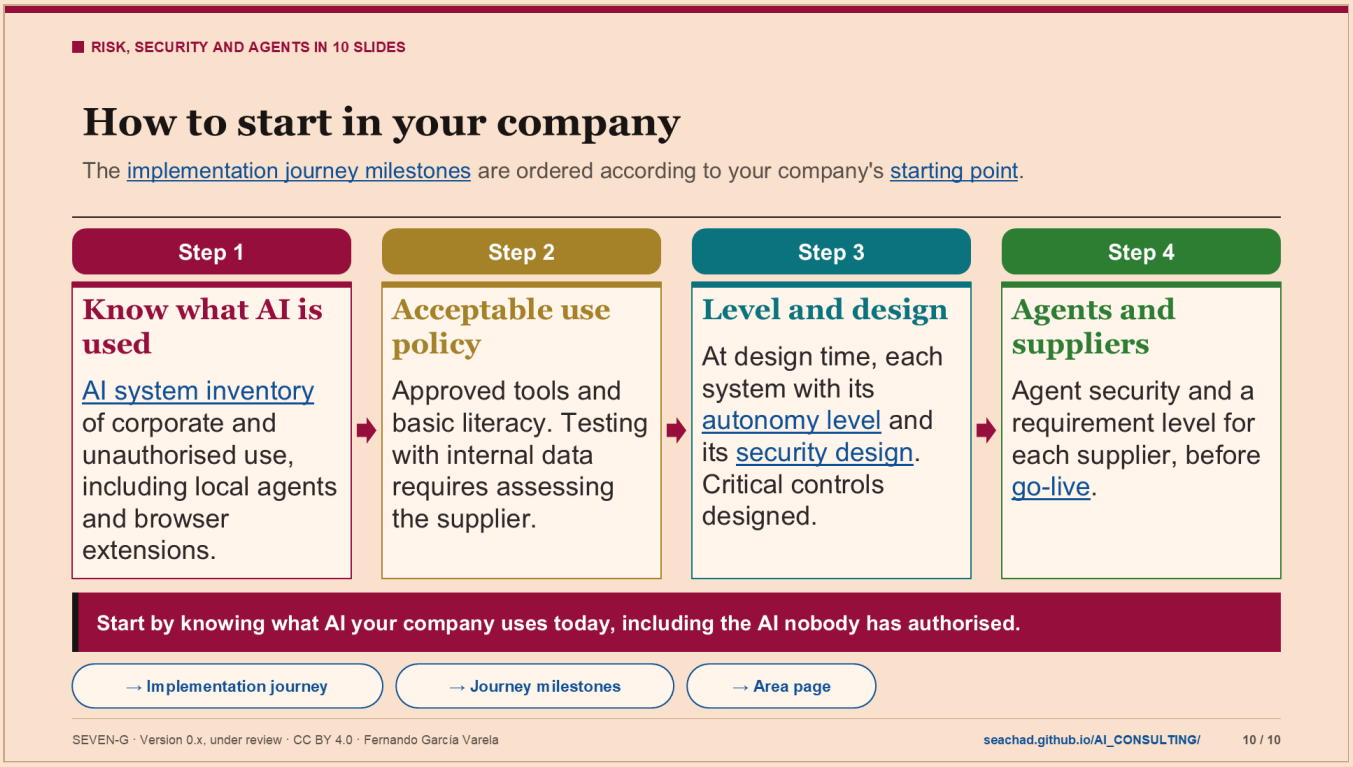
Why it matters. The same vocabulary in SEVEN-G and SPHERES prevents the board, the product team and the development team from talking about different things.

TERMS ON THIS SLIDE

- **AI risk methodology** (documento 33): How AI risks are identified, assessed and accepted: likelihood and impact scales, the 5 × 5 matrix, ten categories, a catalogue of typical risks and who may accept each level. The risk register is mandatory from the feasibility phase. The link opens the matrix in document 33. https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/33_SEVEN-G_Metodologia_de_riesgos_de_IA.html#4-3-risk-level-and-5-5-matrix
- **AI security profile (NIST CSF)** (P72): A template for companies that use the NIST cybersecurity framework: current and target profile of AI security, subcategory by subcategory, with the gap and the actions. The NIST AI profile it relies on is still a draft. It is conditional: only if the company uses that framework. The link opens the template. https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/plantillas/P72_SEVEN-G_Perfil_de_seguridad_de_IA_CSF_y_Cyber_AI_Profile.html
- **AI incident and its severity** (S1–S4): An event in which an AI system causes or may cause harm: serious error, bias, data leak, improper action by an agent. It is classed from S1 (the most serious) to S4. S1 and S2 reach the board by the fast track and may require notifying authorities (P51). The link opens the severity scale in document 37. https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/37_SEVEN-G_No_conformidades_e_incidentes.html#4-2-severity
- **Go-live gate** (G5): The gate that authorises using the AI system with real users, customers or data. It requires validation and pilot passed, a tested rollback plan and multi-level sign-off (business, technology, risk). Nothing goes into production without this sign-off. The link opens its criteria in document 21. https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/21_SEVEN-G_Criterios_de_gate_y_auditoria.html#6-6-g5-go-live
- **SPHERES**: A methodology that explains where AI can create value in a company and with what ambition, through three questions: where do we play?, with what ambition? and what do we ask? It is the explanatory layer of the spheres and ambition levels; the operating rules are in SEVEN-G. It supports conversations with management and the board. The link opens its presentation document. https://seachad.github.io/AI_CONSULTING/SPHERES/html/en/00_SPHERES_Que_es_y_para_que_sirve.html#what-spheres-is
- **Autonomy levels** (A0–A3): A scale of how much an AI system acts on its own. A0: only informs; A1: proposes and a person validates each action; A2: acts within limits and a person supervises; A3: acts alone in a defined scope. The more autonomy, the more controls: an A2 or A3 agent with effects on third parties is Enterprise. The link opens the scale in document 35. https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/35_SEVEN-G_Seguridad_de_IA_y_agentes.html#autonomy-levels

SOURCES AND LINKS

- **Document 34 §5.3 · NIST CSF 2.0 and Cyber AI Profile (draft)** https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/34_SEVEN-G_Mapeo_regulatorio.html#5-3-nist-csf-2-0-and-cyber-ai-profile
- **Document 53, section 7.4 · Security** https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/53_SEVEN-G_Construccion_de_soluciones_con_IA.html#7-4-security
- **SPHERES 03 · The four autonomy levels** https://seachad.github.io/AI_CONSULTING/SPHERES/html/en/03_SPHERES_Esferas_habilitadoras.html#the-four-autonomy-levels
- **The NIST Cybersecurity Framework (CSF) 2.0** https://csrc.nist.gov/pubs/cswp/29/the-nist-cybersecurity-framework-csf-20/final
- **NIST IR 8596 · Cyber AI Profile (initial preliminary draft)** https://csrc.nist.gov/pubs/ir/8596/iprd
- **Regulation (EU) 2024/1689 (EU AI Act)** https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng
- **Building solutions with AI** https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/53_SEVEN-G_Construccion_de_soluciones_con_IA.html
- **What SPHERES is** https://seachad.github.io/AI_CONSULTING/SPHERES/html/en/00_SPHERES_Que_es_y_para_que_sirve.html
- **Area page** https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/areas/seguridad.html



Notes

There is no need to set everything up at once. The SEVEN-G implementation journey (document 96) orders the milestones according to each company's starting point. For this area, four steps mark the way.

- Know what AI is used (milestone HI-02). The AI Office and security carry it out. It includes unauthorised use: external assistants with company data, local agents with credentials on devices and AI browser extensions. Everything is inventoried and authorised like any other tool. If there is evidence of unauthorised use, or it is not known, this milestone moves to top priority.
- Acceptable use policy (milestone HI-03). It sets which tools are approved and includes basic training. It applies the principle that testing is also using.
- Level and design (phase 4 of each use case). Each system receives its autonomy level and its security design (P18). Critical controls are designed before the design gate (G4).
- Agents and suppliers (milestone HI-16). Security, procurement and risk lead it: the agent requirements of document 35 and the N1, N2 or N3 requirements for each supplier according to document 36. It must be ready before go-live (G5).

The implementation journey tool (T23) classifies the company's starting point with a questionnaire and shows the prioritised journey. It works in the browser with fictitious example data.

Why it matters. What is not inventoried cannot be protected. Starting with the inventory and the policy is cheap and gives the board a first honest answer: what AI we use and who controls it.

Question for the classroom: if you had to take an inventory of your organisation's AI tomorrow, who would you ask first?

TERMS ON THIS SLIDE

- **Implementation journey milestones** (HI-01–HI-22): Twenty-two steps to implement SEVEN-G in five stages, from naming a sponsor and inventorying AI to the annual review and the declaration. Each milestone says what to do, who does it, with which template or tool and which maturity question substantiates it; its priority

depends on the starting point. The link opens the milestones in document 96. https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/96_SEVEN-G_Puntos_de_partida_y_recorrido_de_implantacion.html#the-journey-milestones

- **Starting points** (PP-A–PP-F): Six types of company by their starting point with AI: zero point, user of third-party AI, automator with rules or RPA, in-house AI in production, many pilots and at scale. The type changes the order, emphasis and calendar of implementation, never the destination or the rules. The link opens the six types in document 96. https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/96_SEVEN-G_Puntos_de_partida_y_recorrido_de_implantacion.html#2-1-the-six-archetypes
- **AI system inventory** (T02): A list of all the company's AI systems: its own, third-party AI embedded in processes, the assistants employees use and any unauthorised use detected. Each system has an owner, a regulatory classification, an intensity and an autonomy level. It comes first, because you cannot govern what you do not know. It is a view of the use case register: open the link and choose «Inventory» (fictitious data). https://seachad.github.io/AI_CONSULTING/SEVEN-G/herramientas/T01_registro_iniciativas/registro.html
- **Autonomy levels** (A0–A3): A scale of how much an AI system acts on its own. A0: only informs; A1: proposes and a person validates each action; A2: acts within limits and a person supervises; A3: acts alone in a defined scope. The more autonomy, the more controls: an A2 or A3 agent with effects on third parties is Enterprise. The link opens the scale in document 35. https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/35_SEVEN-G_Seguridad_de_IA_y_agentes.html#autonomy-levels
- **Security design** (P18): A template to design the solution's security before building it, agents included: threats, controls, identities, permissions, logs and planned tests. It is reviewed at the design gate. Adding security afterwards costs more and protects less. The link opens the template. https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/plantillas/P18_SEVEN-G_Diseno_de_seguridad.html
- **Go-live gate** (G5): The gate that authorises using the AI system with real users, customers or data. It requires validation and pilot passed, a tested rollback plan and multi-level sign-off (business, technology, risk). Nothing goes into production without this sign-off. The link opens its criteria in document 21. https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/21_SEVEN-G_Criterios_de_gate_y_auditoria.html#6-6-g5-go-live

SOURCES AND LINKS

- **Document 31 §4.7 · Personal agents and automations** https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/31_SEVEN-G_Politica_corporativa_y_uso_aceptable.html#4-7-personal-agents-and-automations
- **Document 35 §6 · Control of corporate AI use (SEG-20)** https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/35_SEVEN-G_Seguridad_de_IA_y_agentes.html#catalogue-of-ai-security-controls-seg
- **Tool T23 · Implementation journey** https://seachad.github.io/AI_CONSULTING/SEVEN-G/herramientas/T23_recorrido_implantacion/recorrido.html
- **Implementation journey** https://seachad.github.io/AI_CONSULTING/SEVEN-G/herramientas/T23_recorrido_implantacion/recorrido.html
- **Journey milestones** https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/96_SEVEN-G_Puntos_de_partida_y_recorrido_de_implantacion.html#4-1-what-each-milestone-is-who-does-it-and-how-it-is-substantiated
- **Area page** https://seachad.github.io/AI_CONSULTING/SEVEN-G/html/en/areas/seguridad.html

LEGAL NOTICE

Legal notice: a generic, free methodological aid provided "as is" for information purposes; it is not legal advice and has not been legally reviewed for any case. Each organisation adapts it and is responsible for its legal review and its regulatory compliance; it may not reflect recent regulatory changes. To the fullest extent permitted by law, the author accepts no liability whatsoever for the effects of its application. It does not grant certification of any kind. Details in document 93.

CC BY 4.0 (content) · MIT (code) · Fernando García Varela · seachad.github.io/AI_CONSULTING/