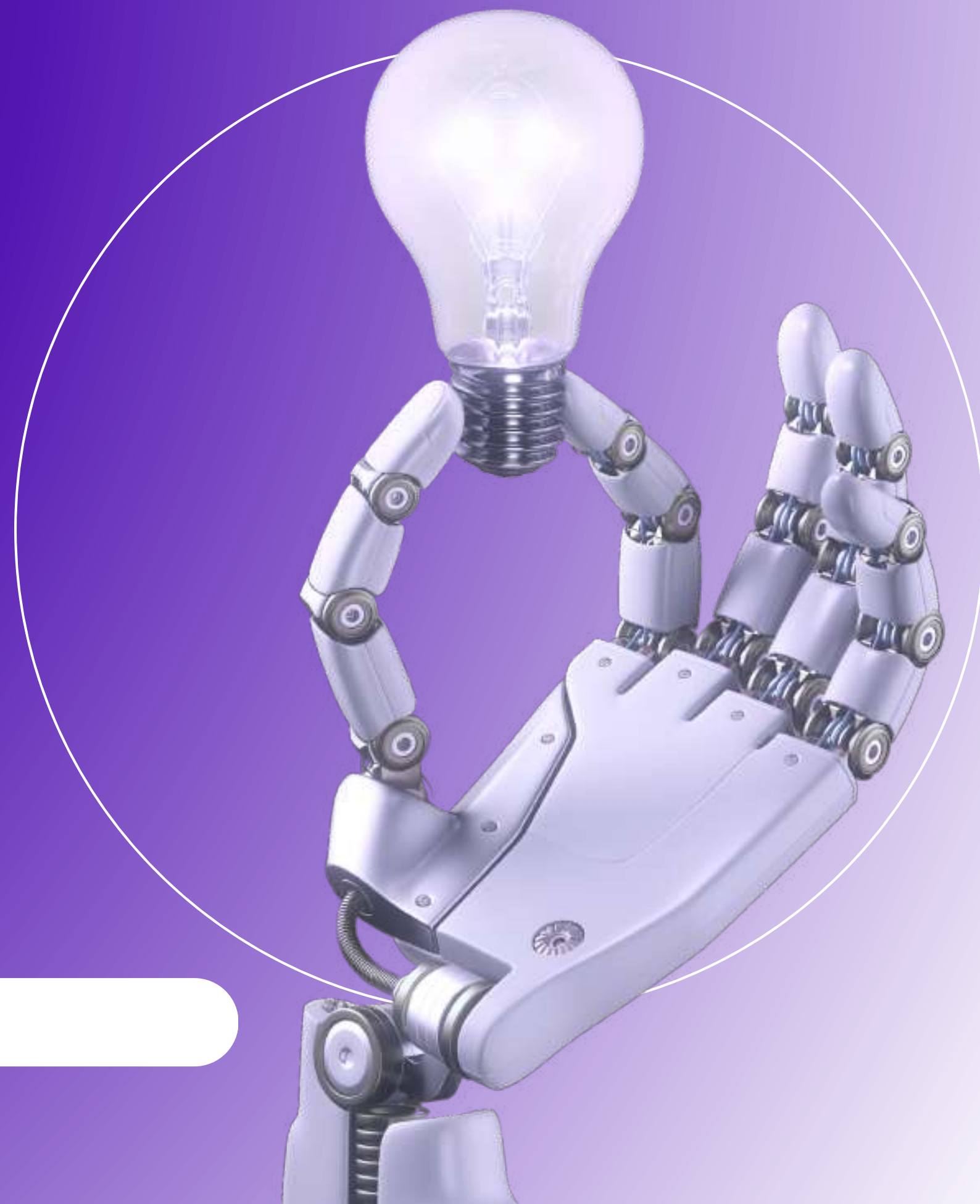

Artificial Intelligence and Supreme Audit Institutions

PROFESSOR DR. SEBASTIÃO HELVECIO RAMOS DE CASTRO

- Vice-President for Development and Public Policy, Instituto Rui Barbosa
- President, International Institute for Innovation, Public Policy and Solutions (3 Ipês)
- Data Analysis Specialist for External Control, Pedro Aleixo School of Accounts, Court of Auditors of Minas Gerais (TCE-MG)

PONTA DELGADA, 16 OCTOBER 2026





Disclaimer

- The Leviathan in motion: Hobbes (Leviathan, 1651) → the Shackled (Acemoglu & Robinson, The Narrow Corridor, 2019) → the Dileviathan (Asma Mhalla, Cyberpunk: le nouveau système totalitaire / Cyberpunk: The New Totalitarian System, Seuil, 2025)
- Artificial Intelligence and Byung-Chul Han (Non-Things, Undinge, 2021): AI is neither intelligent (inter-legere: to choose among options already given) nor artificial (it is software built by humans)



Hobbes's Leviathan (1651)



The Shackled (Acemoglu & Robinson, 2019)



The Dileviathan (Asma Mhalla, 2025)



*



State of the Art and Artificial Intelligence:

the reference proposal of the Dartmouth Summer Research Project (1956)

A PROPOSAL FOR THE
DARTMOUTH SUMMER RESEARCH PROJECT
ON ARTIFICIAL INTELLIGENCE

J. McCarthy, Dartmouth College
M. L. Minsky, Harvard University
N. Rochester, I. B. M. Corporation
C. E. Shannon, Bell Telephone Laboratories



"Moreover, in every period our men of science grow scarcer, and geniuses, like the highest peaks, arise only within mountain ranges. To produce a Galileo or a Newton, we need a legion of worthy investigators."

— Santiago Ramón y Cajal, Father of Neuroscience



To simulate is different from to understand:
AI reproduces behavior, it does not understand it



Sharm El-Sheikh Declaration

"Digitalization is a sustainable trend in the development of public administration. AI is regarded as an auxiliary tool capable of strengthening auditors' analytical functions through the processing of large volumes of data and the automation of routine operations. Auditors retain personal responsibility for audit outcomes, and the application of AI must comply with the principles of ethics, transparency, and data protection."

(Sharm El-Sheikh Declaration, XXV INCOSAI, Oct. 2025, own translation)

State of the Art.

A Conceptual Proposal for Artificial Intelligence Applied to External Control

"Artificial Intelligence is the set of computational systems capable of processing data, identifying patterns, generating analyses, and producing recommendations autonomously or in an assisted manner, subject to the following limitations:

- (i) subordinated to the constitutional mandate;
- (ii) modulated according to the specific professional standards of each activity — modulation that, in the Jurisdiction vertex, amounts to exclusion: the use of Artificial Intelligence is not admitted in the adjudication of accounts;
- (iii) permanently supervised by the auditor's independent professional judgment — supervision that operates precisely at the layer that simulation, by definition, does not reach: understanding;
- (iv) conditioned on the explainability of its methods and results;
- (v) accompanied by full human responsibility for the final product, given that a system that simulates without understanding cannot be a subject of accountability;
- (vi) restricted to institutionally sovereign computational environments, with a guaranteed isolated processing area, prohibiting the insertion of data under the custody of Courts of Audit into Large Language Models operated by third parties, and conditioned on the existence of a complete audit trail.

The human loop must be preserved under any and all circumstances as a constitutional, deontological, and epistemologically non-waivable requirement."

Explainability Principle and External Control

Internal Explainability:

Auditor → AI System

The auditor must be able to demonstrate how the system reached its result. If it cannot be explained, the report cannot be signed.

External Explainability:

Auditor → Auditee

The manager must understand the method in order to exercise an effective right to contest. Algorithmic opacity violates due process of law.

Democratic Explainability:

Auditor → Citizen

The citizen must understand the result in accessible language. Without this, external control produces formal accountability — not real accountability.

The Aphorism of Explainability



What is not understood cannot be explained.

What is not explained cannot be justified.

What is not justified cannot decide.

What does not decide with human responsibility is not external control — it is automation without accountability.

— Sebastião Helvecio

A Proposal to Update the INTOSAI Framework of Professional Pronouncements

INTOSAI-P 1
Lima Declaration

INTOSAI-P 10
Mexico Declaration on SAI
Independence

INTOSAI-P 12
The Value and Benefits of
SAIs

INTOSAI-P 20
Principles of Transparency
and Accountability

INTOSAI-P 50
Principles of Jurisdictional
Activities of SAIs

**A NEW PRINCIPLE IS
NEEDED:**
INTOSAIP-13. Explainability
(XAI)

The Five Justifications for the Explainability Principle (P-13)

1

WHEREAS ARTIFICIAL INTELLIGENCE SIMULATES INTELLIGENT BEHAVIOR WITHOUT CONSTITUTING UNDERSTANDING – A DISTINCTION ESTABLISHED SINCE THE VERY ORIGIN OF THE TERM AT DARTMOUTH (1956) – AN AUTONOMOUS PRINCIPLE IS REQUIRED TO DISCIPLINE HOW ITS METHODS AND RESULTS MUST BE RENDERED INTELLIGIBLE TO HUMAN JUDGMENT;

2

WHEREAS PRINCIPLE 12 (THE VALUE AND BENEFITS OF SAIS FOR CITIZENS) AND PRINCIPLE 20 (TRANSPARENCY AND ACCOUNTABILITY) PRESUPPOSE, WITHOUT EXPRESSLY STATING IT, THE INTELLIGIBILITY OF THE METHODS EMPLOYED BY THE SAI – A PRESUPPOSITION THAT THE INTRODUCTION OF AI SYSTEMS RENDERS, FOR THE FIRST TIME, SUBJECT TO RUPTURE;

The Five Justifications for the Explainability Principle (P-13)

3

WHEREAS ALGORITHMIC OPACITY COMPROMISES THE EXERCISE OF THE RIGHT TO CONTEST BY THE AUDITED MANAGER AND THE RIGHT TO DUE PROCESS OF LAW, EXPLAINABILITY CONSTITUTES A CONDITION OF PROCEDURAL LEGITIMACY, NOT A MERE TECHNICAL GOOD PRACTICE;

4

WHEREAS HUMAN RESPONSIBILITY FOR THE FINAL PRODUCT OF THE AUDIT, ALREADY ENSHRINED IN THE JURISDICTIONAL TRADITION OF SAIS (PRINCIPLE 50), CANNOT BE EXERCISED OVER RESULTS WHOSE ORIGIN THE AUDITOR IS NOT ABLE TO EXPLAIN;

The Five Justifications for the Explainability Principle (P-13)

5

WHEREAS SOCIETY'S TRUST IN SUPREME AUDIT INSTITUTIONS, THE ULTIMATE FOUNDATION OF THE LIMA DECLARATION (PRINCIPLE 1), DEPENDS ON THESE INSTITUTIONS' CAPACITY TO COMMUNICATE THEIR RESULTS IN LANGUAGE COMPREHENSIBLE TO THE CITIZEN, NOT ONLY TO THE SPECIALIST;

The World Is Transforming...

External Control Must Also Transform to Remain Relevant

Human intelligence created writing.



SUMÉRIA, 3.200 a.C.



ÍNDIA, séc. II a.C.



SÉCULO XXI

THE BINARY SYSTEM WAS BORN FROM HUMAN METER (PINGALA), NOT FROM SILICON VALLEY COMPUTING



Genius Also Transforms



LA CORONA, 1884



PERIODO AZUL, 1901



LES FEMMES D'ALGER (O.J. 1907)





The Auditor in the Age of Artificial Intelligence.

Hybrid Auditing

Institutional Quality the scope of External Control

Hybrid Auditing

Pillar 1.

Instrumental, Non-Substitutive Recognition

AI is recognized as a transformative tool for audit efficiency and quality, but always subordinated to ethical, responsible, and well-governed use.

Hybrid Auditing

Pillar 2.

The Hybrid Model as Methodological Axis

Intosai advocates a hybrid model that combines AI with professional judgment, strengthened governance frameworks, coordinated international efforts, capacity building, and oversight by governments.

Hybrid Auditing

Pillar 3.

Explicit Risk Management

The integration of AI into audit processes must be accompanied by risk management, particularly regarding algorithmic bias, data security, and data protection, with all Intosai bodies working in a coordinated manner.

Operational Requirements of the Hybrid Model

Governance and Normative Framework

Formal institutional policy on the use of AI in audits (no ad-hoc use by individual auditors);

————— →

Clear definition of which decisions may be AI-assisted and which remain exclusively human;

————— →

An internal body or function responsible for AI oversight.

————— →

Operational Requirements of the Hybrid Model

Professional Judgment as a Mandatory Filter

AI generates evidence or a hypothesis; the auditor certifies, validates, and assumes responsibility for the result;
_____ →

Prohibition of "automatic signature" of AI-generated results without substantive review;
_____ →

Mandatory training so that auditors know how to question — not merely accept — algorithmic results.
_____ →

Operational Requirements of the Hybrid Model

Technical Risk Management

Periodic testing and auditing of the AI systems themselves for algorithmic bias;

————— →

Security and data protection protocols (especially fiscal/sensitive State data);

————— →

Documentation of data and model provenance.

————— →

A blurred background image showing the back of an audience seated in a conference hall, looking towards a stage with a large screen. The image is partially obscured by a white vertical bar on the left and a purple circle on the right.

Capacity Building

The IDI recently reformulated LOTA as "Audit Tech & AI" in 2026, aligning the use of technology with the three ISSAI-based audit modalities (financial, performance, compliance) and preserving the dual role: using AI in auditing, and auditing the AI used by governments. This vision is incorporated into the IDI's 2024-2029 Strategic Framework.



Amplifying Risk Factors in the Use of AI in Administrative Decisions

INTRINSIC TO THE ALGORITHM'S DESIGN AND AUTONOMY

- Biased algorithms;
- Agentic intelligence.

CYBERATTACKS — EXTERNAL ADVERSARIAL VECTORS

- Prompt injection;
- Malicious swarms;
- Data poisoning;
- Supply chain compromise.



A Proposed Guideline for the Use of Prompts in Auditing





Goal, reach and connection with ISSAI 100

- The formulation of prompts is a professional competency not yet governed by the IFPP.
- This proposal guides the auditor through the planning, execution, reporting, and follow-up phases.
- The prompt is an act of auditing — subject to the same principles as ISSAI 100, the "mother standard."

The word "principles" is written in a cursive, chalk-like font on a dark, textured background that resembles a blackboard. The letters are white with some internal texture, giving them a hand-drawn appearance.

6 core principles

1. Subordination to the constitutional mandate.
2. Modulation by professional activity, full exclusion in Jurisdiction.
3. Skepticism and professional judgment as mandatory filter.
4. Communication-oriented explainability.
5. Full human responsibility.
6. Integrity of environment and evidence: hermetic grounded-generation sandbox.



The 7th Principle: Deliberate friction

Original doctrinal contribution — no precedent in external-control audit literature.

- Auditor's hypothesis recorded before consulting the AI;
- Written justification required to accept any AI-generated result; single-click approval prohibited;
- Periodic recalibration to prevent friction from becoming habitual.

Auditing Guidelines

Planning

AI supports preliminary surveying; the audit criterion is the auditor's exclusive act.

Execution

AI supports evidence gathering; sufficiency and appropriateness are the auditor's judgment.

Reporting

AI supports drafting; clarity, context, and accessibility remain the auditor's responsibility.

Follow-up

AI supports screening; certification of compliance is the auditor's exclusive act.



An invitation for the debate

- KSC, WGITA, and IDI show real institutional interest, yet no normative product exists on this specific gap.
- This draft is not a standard — it is a provocation subsidiary to ISSAI 100, submitted to the Intosai community's debate, not announced as a closed conclusion.

Hybrid Auditing

The Carbon Axis

feel, intuit, empathize, decide,
hold accountable

Carbon explains: How? Why?

The Silicon Axis

process, cross-reference, scale

Silicon points to: Where?
When?

Hybrid Auditing

“

To me, algorithms are silicon prostheses to see the invisible, and human oversight is the carbon arm that builds the solution.

Sebastião Helvecio

Hybrid Auditing

"In God we trust; the rest of us, we audit."

"We do not audit papers; we audit destinies."

"ISSAI is the compass; human dignity is our north."

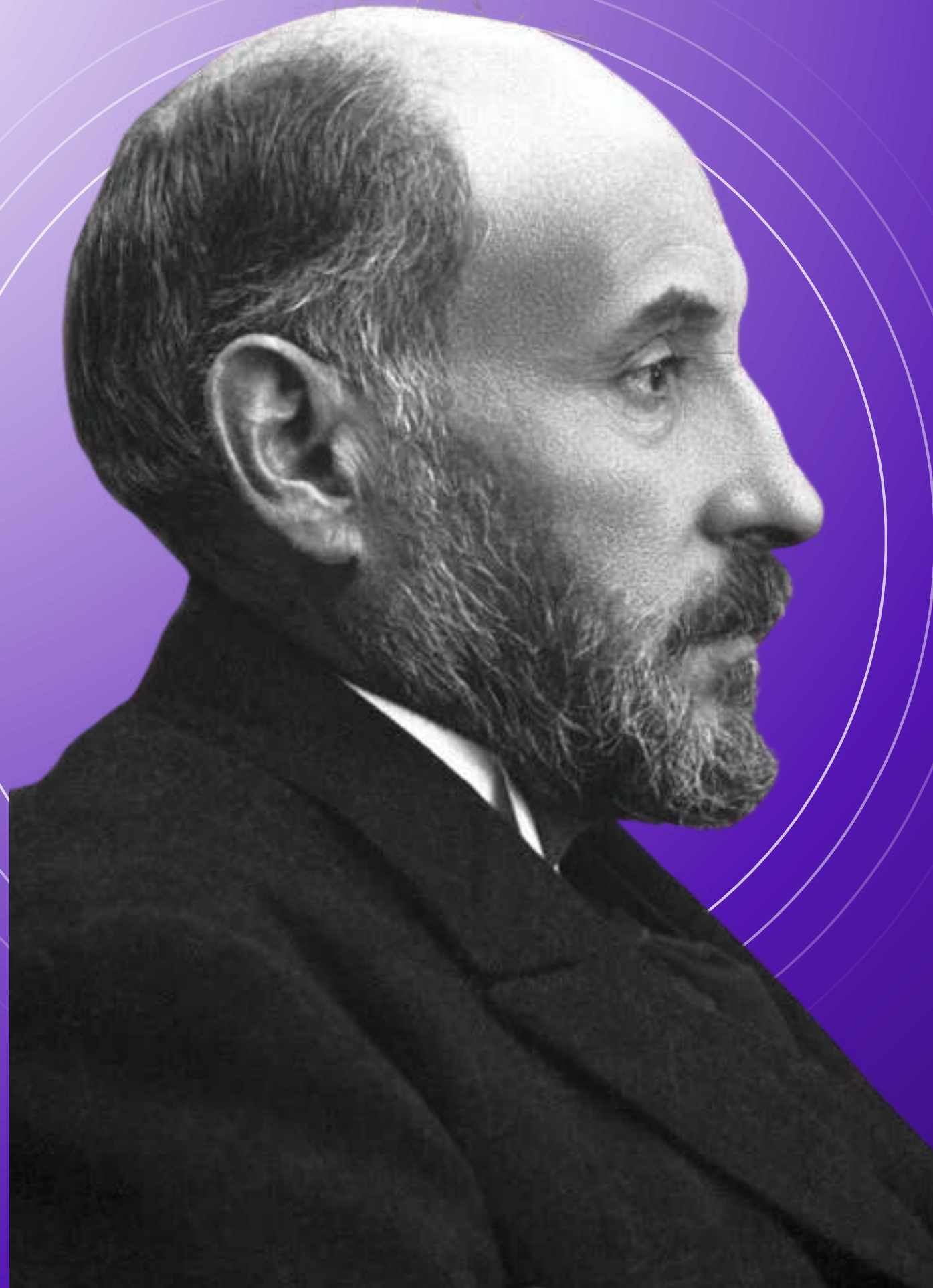
Hybrid Auditing:

The Decisive Stoichiometry

$$2 \text{ Ci} + (\text{C} \cdot \text{Si}) + \text{Nt} = \text{Public Value}$$

- Ci = Citizen
- C = Collective intelligence
- Si = Data processing
- Nt = Technical norms

Collective intelligence (C), guided by professional skepticism, is grounded in the Intosai NBASP criterion, especially Principle 12.



"Ideas do not last long.
Something must be done
with them."

— Santiago Ramón y Cajal,
Father of Neuroscience



WHEN WAS
THE LAST TIME
YOU DID
SOMETHING FOR
THE FIRST TIME?

Explore the New.
Live the Now.

Thank you.

Be happy!