

From AI pilots to real transformation: some lessons learned on turning ideas into impact

actuview

Alexandre Boumezoued, Principal, Milliman AI Solutions

OCTOBER 14, 2026

Common AI challenges

Technical

Organizations often struggle with fragmented data, legacy systems, governance requirements, and selecting the right technology approach.

Business

Many organizations know AI is important but struggle to identify high-value use cases and connect AI initiatives to measurable business outcomes.

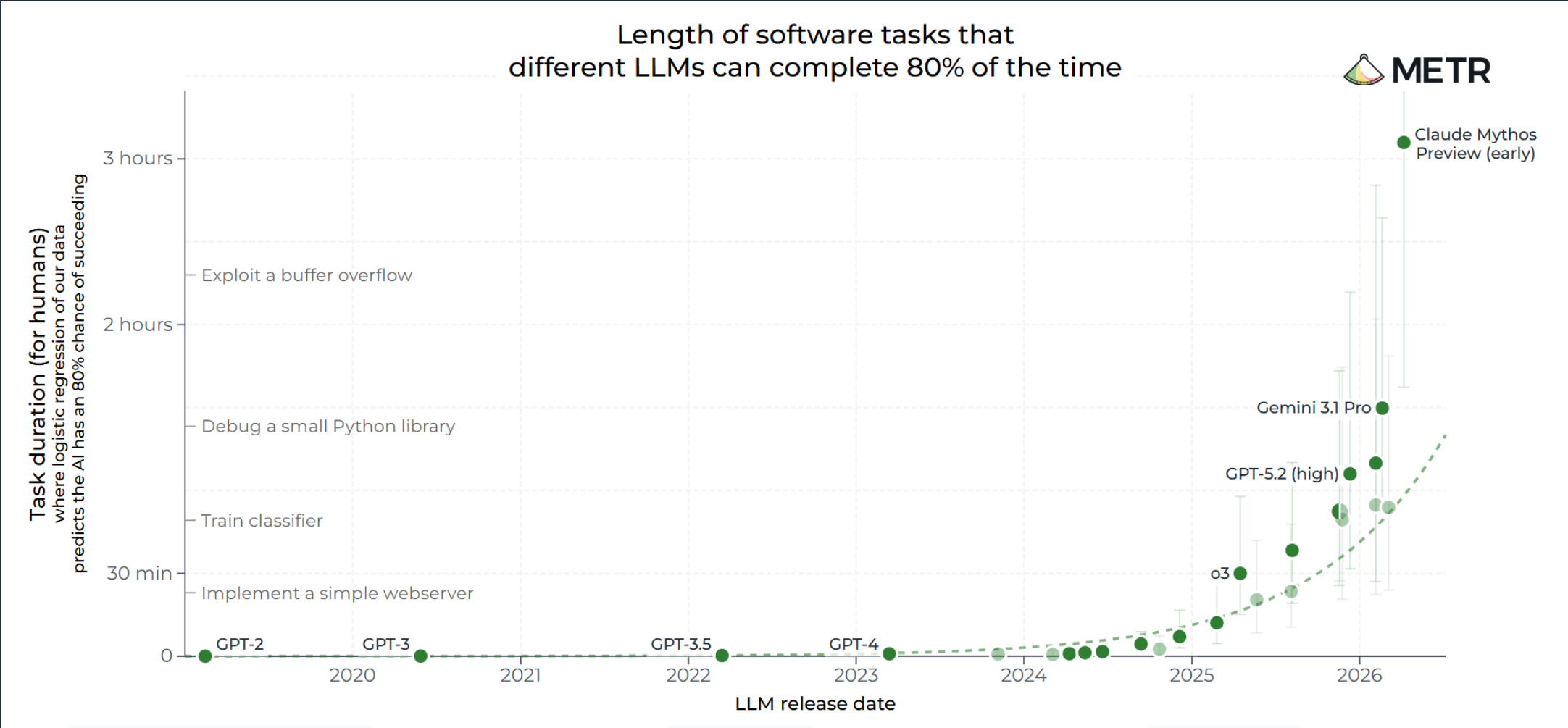
People

AI initiatives often stall because of limited stakeholder alignment, unclear ownership, resistance to change, or insufficient AI literacy.

1. Technical



Duration (for humans) of software tasks LLMs can complete 80% of the time



Measuring AI Ability to Complete Long Tasks - METR

Different models lead in different professions, yet even top-tier AI faces accuracy limits today

Corporate Lawyer		View more
Drafts and reviews contracts, conducts legal research, and advises clients on regulatory and transactional matters. Collaborates with partners on litigation, mergers and acquisitions, and compliance while managing heavy workloads across cases.		
	 GPT-6 Astra	73.4%
	 Fable 5.1	71.9%
	 Opus 5.5	71.2%
Experts from Latham & Watkins, Skadden, Cravath		
Management Consultant		View more
Analyzes industries, evaluates markets, and builds strategic or financial models to guide client decisions. Work often includes preparing presentations, drafting reports, and synthesizing research into actionable recommendations.		
	 Opus 5.5	80.0%
	 Sonnet 5.5	79.7%
	 Gemini 3.8 Flash	74.7%
Experts from McKinsey, BCG, Deloitte, Accenture, EY		
Investment Banking Analyst		View more
Builds financial models, values companies, and prepares pitch materials for potential deals. Responsibilities include conducting industry research, supporting transaction execution, and producing client-ready presentations under tight deadlines.		
	 Sonnet 5.5	77.2%
	 Gemini 3.7 Flash	71.3%
	 Opus 5.5	69.3%
Experts from Goldman Sachs, Morgan Stanley, JPMorgan, Barclays		

How to tailor a Large Language Model for actuarial tasks

A

Ground knowledge in actuarial information

Retrieval-augmented generation (RAG) improves the accuracy and reliability of Generative AI models by using information retrieved from specific and relevant data sources

B

Tailor reasoning capabilities for the actuarial domain

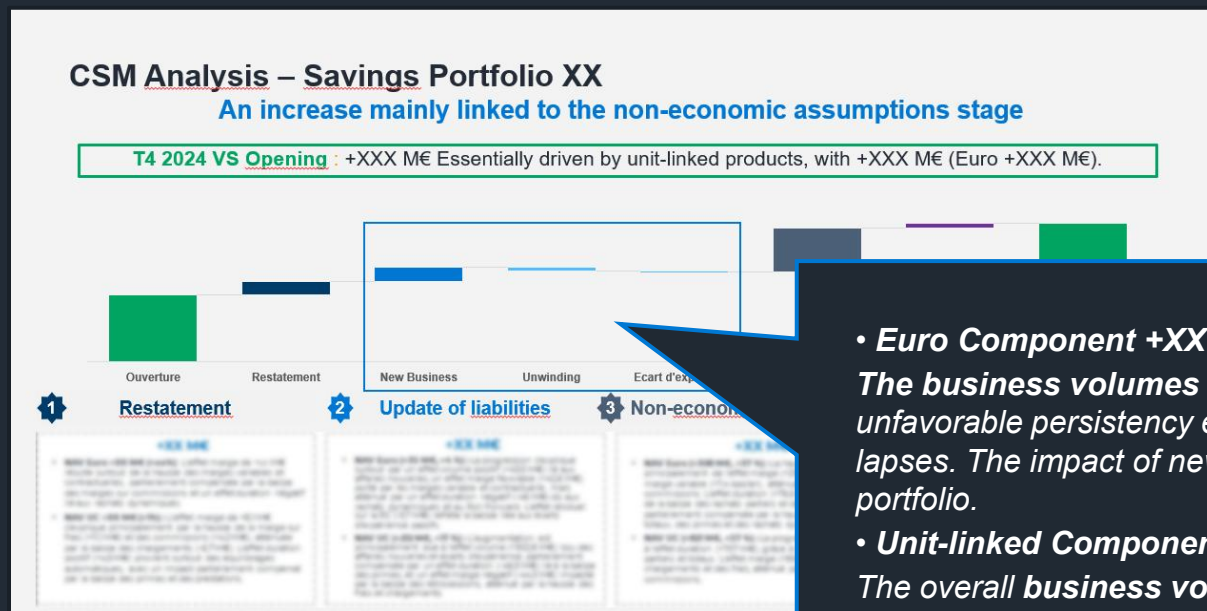
AI agents are autonomous systems that can develop action plans, make decisions, and perform actions to achieve specific goals

C

Leverage the ecosystem of actuarial tools

Model Context Protocol (MCP) has become a standard to make Generative AI models manage interactions with tools and external information

Case study – Automation of IFRS 17 CSM Analysis of Change



Operational Impact +xx M€

• Euro Component +XX M€ (+XX%)

The business volumes impact, slightly negative at XX million euros, is mainly explained by unfavorable persistency experience variances (-XX million euros in CSM), due to higher-than-expected lapses. The impact of new business (+XX million euros) partially offsets the run-off effect on the portfolio.

• Unit-linked Component +XX M€ (+XX%)

*The overall business volumes impact is positive at XX million euros, following very strong inflows (unit-linked mix rate up by xx bps). The additional variation of XX million euros is explained by a **profit margin impact** of XX million euros, linked to the decrease in the average commission rate observed after experience variances and new business (-xx bps).*

2. Business



“About 5% of AI pilot programs achieve rapid revenue acceleration; the vast majority stall, delivering little to no measurable impact on P&L. We observe a clear divide between success stories and stalled projects.”

—The GenAI Divide: State of AI in Business 2025 (July), published by MIT’s NANDA initiative

AI can deliver greater value beyond optimization



Optimization

Faster delivery and lower manual effort



Business Improvement

Better customer service, sales support and business decisions



Compliance

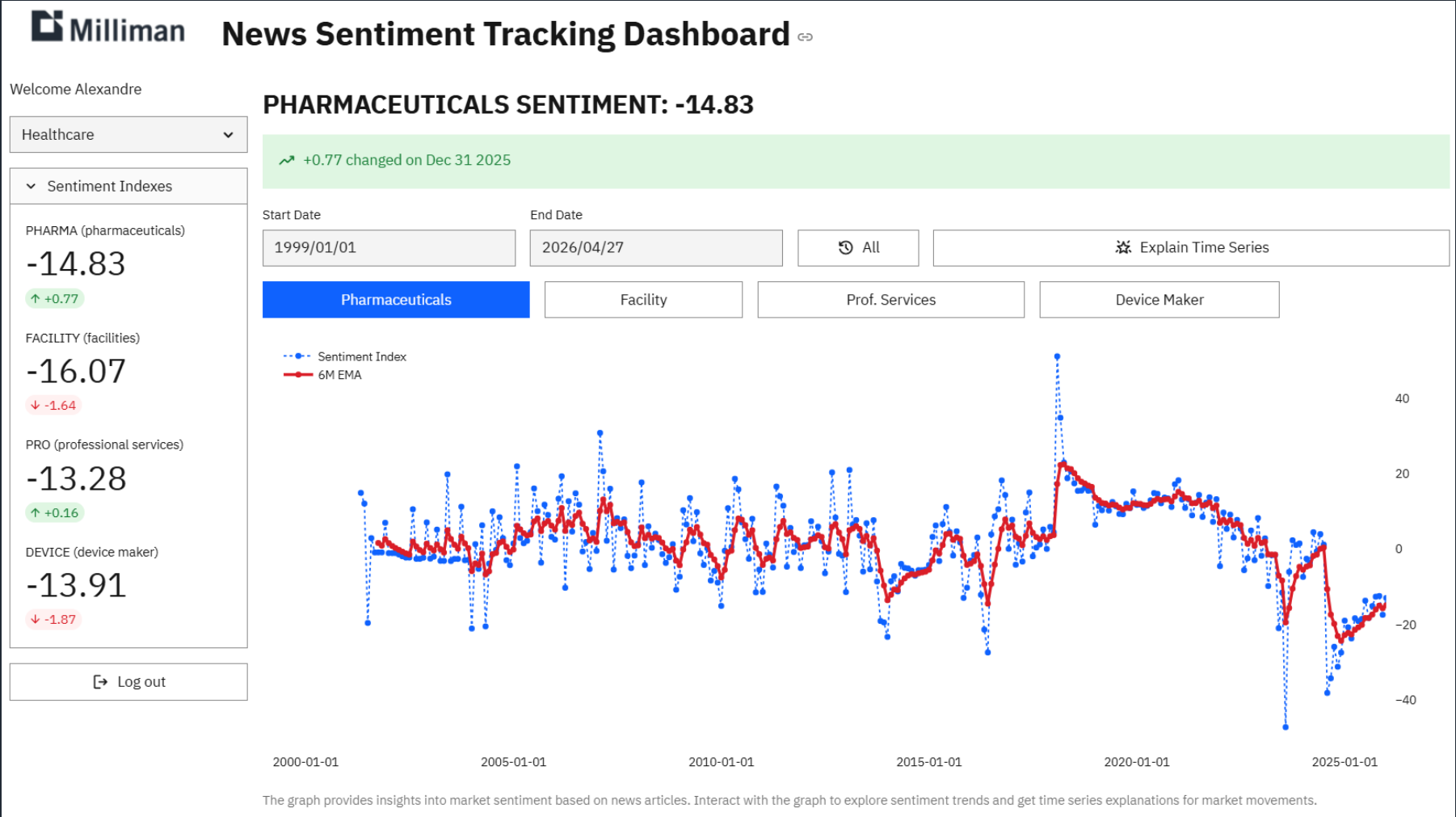
Stronger outcome quality, auditability and review effectiveness



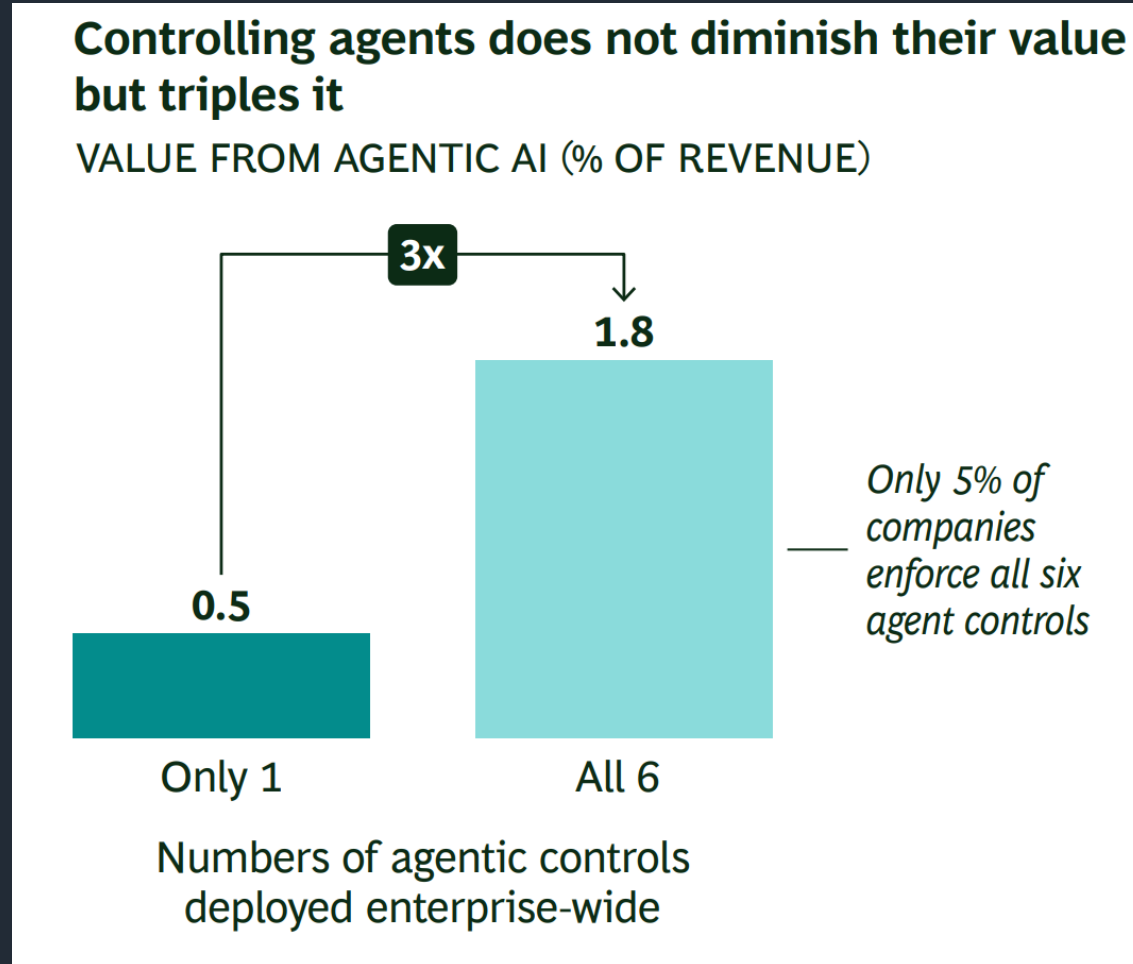
Transformation

Reinventing business model and human-agent interaction

Transformation: deriving tangible risk insights from news feeds



When agentic controls unlock value



—The Formula for Agentic AI Value, BCG report, September 30, 2026

3. People



“[Strategic workforce planning around humans and agents] is practiced by 55% of future-built companies versus 17% of laggards”

– The Formula for Agentic AI Value, BCG report, September 30, 2026

Multi-agentic system audit and validation by actuaries

Expert Review

Review batched runs and leave targeted feedback that becomes ground truth for validation.

TOTAL RUNS
20

STATISTICAL ACCURACY
5.0%

CORRECT
1

PARTIAL CORRECT
0

INCORRECT
0

Experiment Runs

20 results

Run %	Run Timestamp %	Input (Prompts) %	Output (Response)	Chain of Calls %	Important %	Tags %
d788...	11f5e	Sep 30, 04:27 PM	turn it into a management-style slide outline	## Management-style slide outline ## Slide 1 — Executive summary - ...	0 tools / 2 steps	No tags
69f92...	abba3	Sep 30, 04:25 PM	I approve this	Below is the consolidated response based only on your validated output...	0 tools / 2 steps	No tags
4b2a3...	1d3d6	Sep 30, 04:23 PM	run the Methodology Challenger	Methodology Challenger has completed. ## Specialist output ## 31 Cr...	2 tools / 8 steps	No tags
a92a8...	7d0cb	Sep 30, 04:23 PM	I approve this answer			
3d9fc...	ba8c0	Sep 30, 04:20 PM	run the Pattern Investigation			
41fc2...	c1149	Sep 30, 04:19 PM	I approve this answer			
54a68...	f6368	Sep 30, 04:18 PM	run the Movement Analysis			
e9a3d...	065c8	Sep 30, 04:18 PM	I approve the answer			
da544...	77034	Sep 30, 04:14 PM	can you perform the triangles diagnostic for projections			
2b647...	635e3	Sep 30, 01:05 PM	perform triangle diagnostic for projections 001 and 002			

10 runs per page

Expert Feedback Review

reserving-agents-proxied - merciful-ram-877

Select text from the Output Response below to add targeted feedback. Highlighted text indicates existing feedback.

METADATA

BATCH ID

VALIDATION STATUS

email: local

Initial-reserving-test

Pending

TAGS (0/5 MAX)

Type to add tags...

INPUT (PROMPT)

run the Pattern Investigation

OUTPUT (RESPONSE) — SELECT TEXT TO ADD FEEDBACK (MARKDOWN)

- Main issue is **material calendar-year effect**
- Latest immature AY shows a sharp drop in both projections

Stronger hypotheses

- Overall reserve increase is mainly mechanical from adding AY2023, not from broad

Why it fits

- AoM explicitly attributes **+1,081.08** to new business
- Assumption change is only **+28.85**
- Paid LDFs are very similar between 001 and 002

Efficient test

- Reconcile AY2023 reserve to exposure, premium, claim count, and paid-to-date versus prior immature years

- Favorable movement on AY2022 and AY2021 is most consistent with calendar-year payment acceleration or settlement timing

0 feedback notes

Cancel Save All Changes

Chain of Calls

Run 3d9fc80c-6887-4664-b338-dd7fbc18a0c0 - 6 tools / 12 steps. Agent nodes mark each agent (the coordinator and every specialist it delegates to) and own the steps that follow them. Tool calls and delegations show their arguments (long values shortened); recorded prompts and tool results are hidden.

Agent Tool call Execution step Declared parallel batch

Disclaimer

This presentation presents information of a general nature. It is not intended to guide or determine any specific individual situation and Milliman recommends that users of this presentation will seek explanation and/or amplification of any part of the presentation that they consider not to be clear.

Neither the presenter nor the presenter's employer shall have any responsibility or liability to any person or entity with respect to damages alleged to have been caused directly or indirectly by the content of this presentation. All persons who choose to rely in any way on the contents of this presentation do so entirely at their own risk.

The contents of this presentation are confidential and must not be modified, copied, quoted, distributed or shown to any other parties without Milliman's prior written consent.

Copyright © Milliman 2026. All rights reserved

