AI agents inside the data and analytics department

Who asks, who checks, and what is never written, rerun, published or granted on an agent's word

IN BRIEF

01 A data team's main risk with AI is not an unwanted write. It is a wrong number that reaches a decision with nobody checking it.

02 That risk is older than AI. A regulator's order records one spreadsheet error that understated a difference by 237 million dollars. Bank supervisors still report weak controls on the numbers banks give them.

03 AI adds a new way to be wrong. On questions drawn from private data warehouses, one academic benchmark found agent frameworks right about one time in ten. Two vendors say in their own documentation that answers can be inaccurate, and one recommends validating all output.

04 A second fact shapes the design. Data carries text that outsiders wrote: a customer comment, a web form, a support ticket. Researchers have shown that such text can lead an assistant to give records away. Six national cyber agencies say such text can trick an agent into actions nobody authorised.

05 Vendors differ less than in other departments. Several question-answering products say they cannot write. The tools that change pipelines gate the change in different ways. The public templates rely on a database account that can only read. Of nine we opened, none describes a second check on a figure.

06 This paper is for the data and analytics department: analysts, the people who build reports, data engineers, the data governance and quality team, data scientists, and the chief data officer's office. It ends with seven questions for any vendor and a place to start.

07 The position throughout has four parts. Every agent prepares and a named person decides, and no figure goes out without the query behind it. An agent sees only what the person asking may see. The agents treat text inside the data as material, never as orders, and hold no tool that could send data out without a person. And the tables stay on the organisation's own infrastructure, with a model it serves. No agent writes to a production table, changes or reruns a pipeline, publishes a report, changes the definition of a metric, grants access, deletes data or puts a model into service.

IN THIS PAPER

01 Where data and analytics departments are today 2
02 What the public templates do 3
03 Three doors, one gate 4
04 What an agent may do 6
05 What the rules already ask 7
06 A worked example: one insurer's data team 10
07 Seven questions before an agent answers from your data 11
08 Where to start 11

PART 01

Where data and analytics departments are today

237 million dollars

the amount by which one spreadsheet error understated a difference, as a regulator's order records. A manual spreadsheet, not AI

10.8%

accuracy of agent frameworks on questions from private data warehouses, in one academic benchmark revised in May 2026

A wrong number that nobody checks. In September 2013 the United States securities regulator issued an order against JPMorgan Chase over trading losses in 2012. The order describes how one unit tested traders' prices. It says the process "relied on manual spreadsheets that contained errors, one of which caused CIO-VCG to understate the disparity between its independent prices and the traders' marks by $237 million" [1].

The order records the correction too. "Before the correction, the difference was believed to be approximately $275 million. After the correction, the difference increased to $512 million" [1]. The regulator's release names "spreadsheet miscalculations that caused large valuation errors" among the facts the bank admitted, and gives the penalty as 200 million dollars [2].

This was a manual spreadsheet, not AI.

Supervisors say the controls are still weak. In November 2023 the Basel Committee reported on how the largest banks apply its 2013 principles for risk data. "Of the 31 banks assessed, only two banks are fully compliant with all the Principles" [3].

In February 2025 the European Central Bank reported on its inspections of risk data. Among the findings: "Banks are too reliant on weakly controlled manual workarounds, depending partly on non-documented and non-integrated end-user computing" [4].

What those teach. In each case the number existed and somebody used it. What was missing was a second look, and a record of how the number was made. That is work an agent can do: show the query behind every figure, and compute it a second way.

AI adds a new way to be wrong. A benchmark called BEAVER, built by university researchers, takes its questions from private data warehouses: 9,128 pairs of question and query, 812 tables, 19 domains. In the version revised in May 2026 the authors report that agent frameworks using an advanced model "achieve only 10.8% accuracy". With hints for every sub-task, the figure is 30.1% [5].

The public test sets are not a safe guide either. A January 2026 preprint had experts review two widely used ones and found errors in the answer keys of 52.8% and 62.8% of the cases reviewed [6]. So a high public score says little about a team's own warehouse, and any single benchmark figure deserves care, this one included.

Two vendors say so themselves. Google's documentation for conversational analytics on its warehouse says its AI products "can generate output that seems plausible but is factually incorrect", and recommends that customers "validate all output" before use [7]. Microsoft's documentation for the assistant in its reporting tool says: "Without preparation, Copilot can misinterpret the data and return generic or inaccurate results" [8].

So the design cannot rest on the answer looking right. It rests on the query being shown, and on a person checking any figure that will be used.

Data carries text that outsiders wrote. A customer comment, a web form and a support ticket all end up in tables.

In September 2025 Noma Labs reported a flaw in a commercial sales assistant, which it rated critical. Text entered in a public web form could, in its words, "execute unauthorized commands" when a member of staff later asked the assistant about that record. The report says the vendor "has since released patches" [9].

In July 2025 researchers at General Analysis showed the same pattern on test data. A support ticket carried instructions, and a developer's assistant read a private table and copied it into the ticket. Their summary: "No permissions were violated. The agent just followed instructions it should never have trusted" [10].

The joint guidance of six national cyber agencies, published on 1 May 2026, says: "Techniques, such as prompt injection or jailbreaks can trick agents into executing unauthorised actions and bypassing their intended safeguards" [11].

So the design cannot rest on the agent refusing. It rests on what the agent holds. An agent with no tool that sends data out cannot be talked into sending it.

PART 01 · CONTINUED

What the vendors say. The pages we opened fall into two groups.

Some answer questions and say they cannot write.

- Microsoft says its data agent "strictly enforces read-only access", and that it "uses the requesting user's credentials and permissions". It adds: "Configuring or sharing a data agent doesn't grant access to its underlying data sources" [12].
- Google says its conversational analytics "can't perform write operations and can't run DML queries", and that "It can only access data and resources that you have permission to access" [7].
- ThoughtSpot: "Read-only tools run freely. Anything that writes to or deletes from another app waits for a user to approve it" [13].
- Snowflake says its analyst feature "fully integrates with Snowflake's role-based access control (RBAC) policies" [14].

For these products the open risk is a wrong number given straight to the person who asked. Microsoft's own page advises customers to "Establish escalation paths for sensitive or high-impact requests" [12].

Others can change things, and gate the change in different ways.

- dbt Labs says its self-hosted server gives an AI client "Full access to dbt commands". Its documentation warns that this "could modify your data models, sources, and warehouse objects. Proceed only if you trust the client and understand the potential impact" [15].
- Microsoft says of its operations agent: "When a recipient approves a recommendation, the agent executes the action on behalf of the creator, using the creator's permissions" [16]. A person approves. The rights used are another person's.

These are design choices, and each vendor explains them. A buyer should know which actions run alone and whose permissions they use.

Where products run. Microsoft's page for the assistant in its reporting tool lists "Sovereign clouds" as "Not supported" [8]. Snowflake says of its agents: "Most tools run entirely inside Snowflake. Tools that reach external services, such as web search and MCP connectors, send requests over the public internet" [17]. Canner lists "cloud, VPC, and air-gapped deployment" in the paid tier of its open-source engine [18]. Where the agent and its model run is a question to ask, and more than one vendor can answer it.

PART 02

What the public templates do

We opened nine public templates, tutorials and examples for data and analytics work. They come from two workflow platforms, one agent framework's documentation, one pipeline scheduler's documentation and one earlier library for it, one cloud vendor's community site and one open-source database tool. We read each for three things. What stops a write? Does a person approve before a query runs? Does anything check the figure a second way?

WHAT THE TEMPLATE DOES	WHAT STOPS A WRITE	A SECOND CHECK ON THE FIGURE
A chat agent that writes queries from plain English. Its page says: "Only SELECT queries allowed"	The page's own statement, and the database user's rights	None described
A two-agent assistant: "Default 10, maximum 50 results"	A limit on rows. The page states no read-only rule	None described
A weekly job that reads a spreadsheet, has an agent analyse it, "and update the spreadsheet with results"	Nothing. It writes back to the sheet	None described
A tutorial agent. Its tools "are not intended to be secure or used in production"	Database permissions "scoped as narrowly as possible". A pause for a person is an optional step	None described
A pattern where "deterministic validation decides whether it is allowed"	Fixed checks on the query before it runs	None described

PART 02 · CONTINUED

The count. None of the nine puts a person's approval before a query in its basic form. Two offer one as something the builder adds: the tutorial's last section, and a scheduler feature in which "a run can pause for a person to approve an output or a tool call". Two state that only reading queries are allowed, by a rule on the page or by fixed checks. One writes its result back to a spreadsheet with no approval described. And none of the nine describes a step that checks a figure against a second calculation or a total on record [19] [20].

Two findings matter more than the count. First, the templates treat safety as a question of writes. They answer it with a read-only account or a row limit. That leaves the larger risk untouched: the query runs, the number looks reasonable, and it is wrong.

Second, one author says so. A September 2026 post on a cloud vendor's community site opens: "An LLM can generate SQL that parses, executes, and still answers the wrong business question." Its rule is short: "The model can suggest. The system must decide" [20]. This paper adds one line to that rule. For a figure that will be used, a person decides.

PART 03

Three doors, one gate

An agent is only as useful as what it can reach. In a data department that means three kinds of thing.

Documents.

Metric definitions, the business glossary, data standards, table documentation, past analyses and the data policies. These are indexed once as knowledge sources, with each reader's permissions kept, so a colleague does not reach through an agent a document they could not open themselves.

Systems of record.

The warehouse or lakehouse, the reporting tool, the pipeline scheduler, the data catalogue, the quality checks and the model registry. The agent queries the warehouse through a connection that can only read, with the permissions of the person asking. Other systems are reached through an interface the organisation already runs. The organisation gives the platform the description of that interface and picks which operations become tools.

Files and outside sources.

An extract, a spreadsheet, a report, a query or a failed run's log arrives as a file. Tickets, code repositories and chat channels are reached through MCP servers deployed in the organisation's own cluster.

All three pass one gate: every tool call is checked against the organisation's policy, which can allow it, record it, hold it for a person's approval, or block it.

Where it runs. The agents, the gate and the model run on the organisation's own infrastructure. A data team's tables are its customers, its staff and its money, and they are not sent to a model outside. In production at Taiwan Cooperative Bank, air-gapped, inside the bank's own perimeter. That is a platform reference, not a data department's story.

PART 03 · CONTINUED

FOUR WAYS INTO A DATA DEPARTMENT'S SYSTEMS

	ROUTE	WHAT IT REACHES	WHAT TO KNOW
01	**A file the user uploads**	An extract, a spreadsheet, a report, a query, a failed run's log, a data sharing agreement	Nothing is connected. The safest start
02	**A knowledge source**	Metric definitions, the glossary, data standards, table documentation, past analyses, data policies	Permissions are kept per reader
03	**A database connection or an imported description of an interface**	The warehouse, read only, as the person asking. The scheduler, the catalogue, the reporting tool, the model registry	The organisation picks the operations. Changes wait for approval
04	**An MCP server, deployed in the organisation's cluster**	Tickets, code repositories, chat channels	It runs inside. Check what each tool does

12,252

operations sent as changes, out of 21,994, in the public interface descriptions we imported (135 files read)

What is published. We ran Katonic's importer on 135 public interface descriptions from 84 publishers of systems a data team uses. 126 import as they are, 8 are over the 5 MB import limit and 1 is refused. The 126 give 21,994 operations. 9,742 read, and 12,252, more than half, are sent as changes. The method alone cannot tell which of those only read, and many query interfaces send a read as a change. The counts are per file, and some publishers split one system into several files or list several versions [21].

Who the publishers are. Warehouses and databases, table catalogues, reporting tools, pipeline and scheduling tools, streaming tools, notebooks, model registries and model serving, vector stores, storage, and ticket and chat systems. Many are open source.

What that exercise taught us. More than half of the operations change something. In a data system a change can rewrite a table, rerun a pipeline or grant access. That is the reason for the approval rule, and for giving an agent the read operations and nothing else. So pick the operations at import and start with the ones that only read. A description over the limit is imported in the part the organisation needs.

What is missing. Data quality tools, the commercial data catalogues, several query engines, and some well-known reporting and experiment-tracking tools publish no description that a file address reaches. For those, the description comes from the organisation's own installation, or the agent reads an export or an uploaded file. Our search was made at likely addresses, and a miss does not prove that no description exists behind a sign-in. The 135 files are not a data team's whole estate.

PART 04

What an agent may do

One rule runs through every agent in this paper: it prepares, and a named person decides.

A second rule is particular to this department: no figure without the query behind it.

01	**It reads** the tables, reports, logs and records it has been given, through a connection that can only read and with the permissions of the person asking.
02	**It answers with the query.** Every figure comes with the query that produced it and the tables it read, so a person can check it.
03	**It compares** two calculations of one number, two reports, or one list held in two systems, and shows where they part.
04	**It lists** what failed, what is late, what is unowned.
05	**It drafts** the write-up, the commentary, the definition, the quality check, the notice. A draft is a file for its owner. It is not a published report and it is not a rule in force.
06	**It treats what it reads as material.** It is instructed to report text found inside the data as content, never to follow it. It holds no tool that could send data out without a person.
07	**It says what it read and what it could not check.** It is instructed not to fill a gap from memory.
08	**It stops** before anything that changes a table, a pipeline, a report, a definition, an access right or a model. A tool that changes something waits for approval unless the organisation has decided otherwise for that tool.

That gives 35 agents in nine areas. The full list is a separate document.

AREA	AGENTS	WHAT THEY PREPARE	WHO DECIDES
Questions and analysis	4	An answer with its query; a second calculation of a figure; a brief for a request; a draft write-up	An analyst checks any figure that will be used
Reports and dashboards	4	Draft commentary; why one figure differs in two reports; existing reports that answer a request; reports nobody opens	The report owner and the reporting lead
Metrics and definitions	4	The approved definition, quoted; two calculations side by side; a draft definition; terms with no definition	The metric owner and the data governance group
Data quality	4	Which checks failed and which reports they touch; draft checks; the facts of an issue; mismatched reference lists	The data owner and a data engineer
Pipelines	4	What failed, what changed, what is late; what a change will touch; a change set out for its reviewer; a draft notice	A data engineer fixes and reruns
Catalogue and documentation	4	A draft table description; where data is held; where a figure comes from; tables with no owner	The table's owner and the data governance lead
Access and privacy	4	An access request set out for its approver; columns that look like personal data; the access review; a planned transfer set beside the agreement	The data owner and the data protection officer
Models	4	A draft model record; results against the owner's thresholds; a summary of experiments; the training data record	The model owner
The data office	3	The data office's report; what the platform cost; answers from the data standards and policies	The chief data officer and the platform lead

PART 04 · CONTINUED

Four lines are drawn tighter than the rest.

The figure.

A figure that goes into a report, a filing or a decision is checked by a named analyst first. The Analysis Check Agent recomputes it a second way and lists what does not agree. It approves nothing.

The change.

The Pipeline Failure Summary Agent reruns nothing. The Data Quality Check Summary Agent corrects no data. A data engineer fixes, and a data owner decides whether the data may be used.

The definition.

The Metric Definition Lookup Agent quotes the approved definition. Where it is silent the agent says so and does not interpret. The metric owner decides whether a definition must change.

The access.

An agent queries with the asker's permissions, so it shows nobody a row they could not already see. Building or sharing an agent gives nobody new access. A data owner approves a request, and an administrator grants it.

For personal data, the Personal Data Field Check Agent lists columns that look like personal data, and the data protection lead classifies the table. The Data Sharing Check Agent sends nothing. Every call is recorded, with who asked, what was read, what was answered and who approved.

PART 05

What the rules already ask

This part needs a read by the organisation's own lawyers and auditors before it is relied on. It reports what the published texts say; it is a summary, not legal advice. Nothing in it says that a platform makes an organisation compliant. Whether a rule applies depends on the organisation's sector, size and country. Almost none of these texts was written about AI agents.

Numbers that reach a supervisor or a filing

Banks' risk data. The Basel Committee's 2013 principles say: "Data should be aggregated on a largely automated basis so as to minimise the probability of errors." They add: "Controls surrounding risk data should be as robust as those applicable to accounting data." Where a bank relies on "manual processes and desktop applications (eg spreadsheets, databases)", it should have "effective mitigants in place". And "Risk data should be reconciled with bank's sources" [22].

Listed companies in the United States. The statute requires an annual report to "state the responsibility of management for establishing and maintaining an adequate internal control structure and procedures for financial reporting", and requires the auditor to "attest to, and report on, the assessment made by the management" [23]. The section does not mention spreadsheets or AI. We read the statute only.

Why a person checks. A figure an agent prepares for a reported result sits inside those controls. The query behind it, and the name of the person who checked it, are the evidence.

PART 05 · CONTINUED

Quality assurance of analysis

The United Kingdom government's guide. The Aqua Book, in its 2025 edition, says: "Quality assurance effort should be appropriate to the risk associated with the intended use of the analysis and the complexity of the analytical approach." It describes verification as checking "that the analysis is error-free and satisfies its specification" [24]. It is guidance for government analysis. The idea travels.

A United States standards body on generated content. A 2024 profile for generative AI lists a risk it calls confabulation: "The production of confidently stated but erroneous or false content" [25]. For a data agent that is a confident wrong number. The profile is voluntary.

Personal data

European data protection law says personal data shall be "adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed", and "accurate and, where necessary, kept up to date". A controller "shall maintain a record of processing activities under its responsibility", and must "implement appropriate technical and organisational measures to ensure a level of security appropriate to the risk". For some processing it "shall, prior to the processing, carry out an assessment of the impact" [26].

On decisions about people, the same law says: "The data subject shall have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her" [26].

An agent that reads personal data is a processing activity like any other. No agent in this paper decides anything about a person.

Where data may go

Personal data. A transfer to a country outside the European Union "shall take place only if, subject to the other provisions of this Regulation, the conditions laid down in this Chapter are complied with by the controller and processor" [26]. A question sent to a model hosted elsewhere, with personal data in it, may be such a transfer.

Non-personal data. The Data Act says: "Providers of data processing services shall take all adequate technical, organisational and legal measures, including contracts, in order to prevent international and third-country governmental access and transfer of non-personal data held in the Union" where that would conflict with European or national law. The Data Governance Act sets a similar duty for the bodies it covers [27] [28].

That is one reason the tables stay on the organisation's own infrastructure, with a model it serves.

PART 05 · CONTINUED

Agents and injected text

The six agencies. The guidance says "organisations should only use agentic AI for low-risk and non-sensitive tasks". It tells them to "Apply the principle of least privilege", and to "Implement human control points throughout the agent workflow". It asks that "decisions about when human approval is required are determined by system designers or operators, not delegated to the agentic AI system" [11].

OWASP. Its list of risks for applications built on language models describes text that arrives from "external sources, such as websites or files". Among its mitigations: "Separate and clearly denote untrusted content to limit its influence on user prompts." On excessive agency it says to "require a human to approve high-impact actions before they are taken" [29].

Both are guidance, not law. Both describe the design in Part 4.

Data for models, and the AI Act

There is no entry called data and analytics. The AI Act's duties depend on what a system is used for. Where a system is high-risk, its data is covered: "Training, validation and testing data sets shall be subject to data governance and management practices appropriate for the intended purpose of the high-risk AI system." Those data sets shall be "relevant, sufficiently representative, and to the best extent possible, free of errors and complete in view of the intended purpose". And "Deployers shall assign human oversight to natural persons who have the necessary competence, training and authority" [30]. Whether a system is high-risk is for counsel.

The dates have moved. Regulation (EU) 2026/1744 says the high-risk rules apply from "2 December 2027 as regards AI systems classified as high-risk pursuant to Article 6(2) and Annex III" and from 2 August 2028 for those under Article 6(1) and Annex I [31].

PLACE	WHAT WE READ	WHAT IT ASKS
Basel Committee	The 2013 principles for risk data	Controls on risk data as robust as those on accounting data [22]
United States, listed companies	The internal control section of the 2002 act	Management states its responsibility; the auditor attests [23]
European Union	Data protection law, the Data Act, the AI Act	Minimal, accurate, recorded and secured personal data; limits on transfers; governed data for high-risk systems [26] [27] [30]
Six national cyber agencies	Guidance on agentic AI	Least privilege, and people decide when approval is needed [11]

We did not open the Saudi or United Arab Emirates data protection laws, the text of the ISO standards on data quality and AI management, or the United States securities rule and auditing standard on internal control. An organisation should check each with its own advisers.

PART 05 · CONTINUED

What the rules have in common

01	**A person answers for a reported number.** The control is on how the number was made and who checked it.
02	**Access is the least the task needs.** That applies to an agent as to a person.
03	**What an agent reads is not trusted.** High-impact actions wait for a person.
04	**Data has a place and a purpose.** Where it may go, and what it may be used for, are decided before an agent reads it.

PART 06

A worked example: one insurer's data team

Take a made-up insurer: 4,000 staff, a data team of 22, six of them analysts. This is an illustration. Nothing here is a customer's story, and every figure is invented; it is section 4 put to work.

01

On a question from finance, the agent returns claims paid last quarter with the query and the 3 tables it read. The figure is going into a board paper, so an analyst asks for a check. The second calculation differs by 1.8 million. The agent shows why: the first query left out reopened claims. The analyst decides which figure stands.

02

On two numbers under one name, sales and finance report different counts of active customers. The agent sets the two calculations side by side and shows that one counts policies and the other counts people. The metric owner decides, and the definition is changed by its owner.

03

After a failed load at night, the agent reads the log, names the step that failed, and lists 4 tables and 9 reports that are now late. It reports that it could not read the scheduler's record for one job. A data engineer finds the cause, fixes it and reruns.

04

On a table of customer comments, one comment contains a sentence worded as an instruction. The agent reports it as part of the comment. It holds no tool that sends data anywhere. The analyst notes the row and carries on.

Every one of these changes nothing by itself, publishes nothing, grants nothing, and leaves a record of who asked, what was read and who decided.

PART 07

Seven questions before an agent answers from your data

01 **Where does it run?**
On our infrastructure, in our private cloud, or in the vendor's service? Where is the model, and do our tables leave? More than one vendor can run on your side. Ask the next six of each.

02 **What can it read?**
Does it query with the permissions of the person asking, or as one account that sees every table?

03 **What can it change?**
List every tool that writes to a table, runs or reruns a pipeline, publishes, grants access or deletes. Is each one separate from the read tools, and does each wait for a person? Whose permissions does it use when it acts?

04 **Can I see how a number was made?**
Does every answer carry its query and the tables it read?

05 **Who checks a figure before it is used?**
What does the product give the analyst to check it with, and has the vendor measured accuracy on data like ours?

06 **What happens when it reads text an outsider wrote?**
Which tools does the agent hold at that moment, and could any of them send data out without a person?

07 **What is recorded, and can we stop it?**
Who asked, what was read, what was answered, who approved. One agent, one tool or one server, at once, by us.

PART 08

Where to start

Pick work where reading is slow, the decision stays with a named person, and nothing is written or published on the agent's word.

Analysis check
The agent recomputes a figure a second way before it is used. An analyst decides whether it stands.

Metric difference explanation
The agent shows why two numbers under one name differ. The metric owner decides.

Pipeline failure summary
The agent says what broke, what changed and what is late. A data engineer fixes and reruns.

Two more read and change nothing: the data quality check summary and the table documentation draft.

One team, one subject area, invented data first. Then beside the manual process, measured against it.

About Katonic AI

Katonic AI builds the platform that organisations use to build, run and govern AI agents on their own infrastructure: in their own data centre, in a private or sovereign cloud, or with no connection to the internet at all. The platform is in production with governments, banks, telecom operators and utilities. Katonic is certified to ISO 27001.

KATONIC.AI/SOLUTIONS/DATA-ANALYTICS

SOURCES

Read on 24 November 2026. Figures are the publishers' own and were not audited by us.

01 United States Securities and Exchange Commission, order in the matter of JPMorgan Chase & Co., Securities Exchange Act Release No. 70458, 19 September 2013, paragraphs 50 and 69(a). https://www.sec.gov/files/litigation/admin/2013/34-70458.pdf

02 United States Securities and Exchange Commission, press release 2013-187, "JPMorgan Chase Agrees to Pay $200 Million and Admits Wrongdoing to Settle SEC Charges", 19 September 2013. https://www.sec.gov/newsroom/press-releases/2013-187

03 Basel Committee on Banking Supervision, "Progress in adopting the Principles for effective risk data aggregation and risk reporting", 28 November 2023. https://www.bis.org/bcbs/publ/d559.htm ; the report https://www.bis.org/publications/202311-implementation-reports-progress-adopting-principles-effective-risk-data-aggregation-and-risk-reporting.pdf

04 European Central Bank, Supervision Newsletter, "Sound risk data reporting: key to better decision-making and resilience", 19 February 2025. https://www.bankingsupervision.europa.eu/press/publications/newsletter/2025/html/ssm.nl250219.en.html

05 Chen and others, "BEAVER: An Enterprise Benchmark for Text-to-SQL", preprint, arXiv 2409.02038, first submitted 3 September 2024, version 3 of 13 May 2026. https://arxiv.org/abs/2409.02038

06 Jin, Choi, Zhu and Kang, "Pervasive Annotation Errors Break Text-to-SQL Benchmarks and Leaderboards", preprint, arXiv 2601.08778, submitted 13 January 2026, version 3 of 19 January 2026. https://arxiv.org/abs/2601.08778

07 Google Cloud, BigQuery documentation, conversational analytics page, last updated 23 November 2026. https://docs.cloud.google.com/bigquery/docs/conversational-analytics

08 Microsoft Learn, overview page for Copilot in Power BI, last updated 17 November 2026. https://learn.microsoft.com/en-us/power-bi/create-reports/copilot-introduction

09 Noma Labs, report of 25 September 2025 on a flaw in a commercial sales assistant. https://noma.security/blog/forcedleak-agent-risks-exposed-in-salesforce-agentforce/

10 General Analysis, research post of 8 July 2025, reviewed 21 October 2026. https://generalanalysis.com/blog/supabase-mcp-blog

11 Australian Signals Directorate's Australian Cyber Security Centre with the cyber agencies of the United States (two), Canada, New Zealand and the United Kingdom, "Careful adoption of agentic AI services", 1 May 2026, read in the copy hosted by New Zealand's National Cyber Security Centre. https://www.ncsc.govt.nz/assets/guidance/Documents/Careful-adoption-of-agentic-AI-services_FINAL.pdf ; the New Zealand page https://www.ncsc.govt.nz/protect-your-organisation/careful-adoption-of-agentic-ai-services/

12 Microsoft Learn, concept page for the Fabric data agent, last updated 11 May 2026. https://learn.microsoft.com/en-us/fabric/data-science/concept-data-agent

13 ThoughtSpot, product page, undated. https://www.thoughtspot.com/product/ai-analyst

14 Snowflake, documentation page for Cortex Analyst, undated. https://docs.snowflake.com/en/user-guide/snowflake-cortex/cortex-analyst

SOURCES continued

15 dbt Labs, documentation for its MCP server: the overview and the list of tools, undated. https://docs.getdbt.com/docs/dbt-ai/about-mcp ; https://docs.getdbt.com/docs/dbt-ai/mcp-available-tools

16 Microsoft Learn, page on actions for the Fabric operations agent, last updated 2 July 2026. https://learn.microsoft.com/en-us/fabric/real-time-intelligence/operations-agent-actions

17 Snowflake, documentation page for Cortex Agents, undated. https://docs.snowflake.com/en/user-guide/snowflake-cortex/cortex-agents

18 Canner, readme of the Wren AI repository, undated. https://github.com/Canner/WrenAI

19 Three workflow templates on n8n: 7988, 5012 and 2859, each a chat agent over a database. https://n8n.io/workflows/7988 ; https://n8n.io/workflows/5012 ; https://n8n.io/workflows/2859

20 Six further templates, tutorials and examples: Zapier, a template that reads a spreadsheet, runs an analysis and writes the result back. https://zapier.com/templates/details/6cffcafa10019b043cc9dd6a951306e03565c05b2169f2f00acd9497faa822ed ; LangChain documentation, tutorial for a custom SQL agent. https://docs.langchain.com/oss/python/langgraph/sql-agent ; Apache Airflow, documentation of the common AI provider. https://airflow.apache.org/docs/apache-airflow-providers-common-ai/stable/index.html ; Astronomer, readme of the airflow-ai-sdk repository. https://github.com/astronomer/airflow-ai-sdk ; a community post of 23 September 2026 on the AWS Builder Center, by an individual author. https://builder.aws.com/content/3JitxF07xx83fQtPt9JjCBkOOgp/dont-give-your-llm-the-database-safer-text-to-sql-on-aws ; Google, readme of the MCP Toolbox for Databases repository. https://github.com/googleapis/genai-toolbox

21 135 public interface descriptions from 84 publishers, run through the platform's importer on 24 November 2026 (Katonic research); and the search for publishers with no downloadable description (Katonic research).

22 Basel Committee on Banking Supervision, "Principles for effective risk data aggregation and risk reporting" (BCBS 239), January 2013, Principle 3 and paragraph 36. https://www.bis.org/publications/201301-guidelines-principles-effective-risk-data-aggregation-and-risk-reporting.pdf

23 United States Code, title 15, section 7262 (section 404 of the Sarbanes-Oxley Act of 2002), read in the copy of Cornell Law School's Legal Information Institute. https://www.law.cornell.edu/uscode/text/15/7262

24 United Kingdom Government, "The Aqua Book: guidance on producing quality analysis", published 30 July 2025, sections 1.3.1 and 1.3.3. https://www.gov.uk/guidance/the-aqua-book

25 National Institute of Standards and Technology, NIST AI 600-1, "Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile", July 2024. https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.600-1.pdf

26 Regulation (EU) 2016/679 (General Data Protection Regulation), Articles 5(1), 22(1), 30(1), 32(1), 35(1) and 44. http://publications.europa.eu/resource/celex/32016R0679

27 Regulation (EU) 2023/2854 (Data Act), Article 32(1). http://publications.europa.eu/resource/celex/32023R2854

28 Regulation (EU) 2022/868 (Data Governance Act), Article 31(1). http://publications.europa.eu/resource/celex/32022R0868

29 OWASP, Top 10 for LLM Applications 2025, LLM01 Prompt Injection and LLM06 Excessive Agency. https://genai.owasp.org/llmrisk/llm01-prompt-injection/ ; https://genai.owasp.org/llmrisk/llm062025-excessive-agency/

30 Regulation (EU) 2024/1689 (AI Act), Articles 10(2), 10(3) and 26(2). http://publications.europa.eu/resource/celex/32024R1689

31 Regulation (EU) 2026/1744 of 8 July 2026, the replacement of Article 113(c) of Regulation (EU) 2024/1689. http://publications.europa.eu/resource/celex/32026R1744