

MSIS 549 HW2 — Compliance Checklist

Assignment: Agentic AI for Real-World Impact **Student:** Taashi Manyanga **Course:** MSIS 549 B — AI and GenAI for Business Applications **Instructor:** Prof. Léonard Boussieux **Date:** February 2026 **Total Points Available:** 15 + 2 bonus

Deliverables Checklist

Required Deliverable	Status	Location
System Artifact (shareable link)	✓ COMPLETE	https://github.com/taashim-eng/linkedin-thought-leadership-agent
Tutorial Write-up (PDF)	✓ COMPLETE	TUTORIAL_WRITEUP.pdf
Benchmark Appendix	✓ COMPLETE	benchmark/BENCHMARK_APPENDIX.md
[Optional] Demo Recording	⚠ NOT SUBMITTED	Not recorded — see note below

Demo Note: A live test run was executed on 2026-02-11 with the topic "IT Moving from AI Chatbot to Dev Tool." Week 1 was published live to LinkedIn (URN: urn:li:share:7430329696413843458). A screen recording of this session was not captured. If submitting, recommend recording a walkthrough of the GitHub repo + archive file as a substitute.

Grading Rubric — Self-Assessment

Category 1: Problem Selection (1 pt)

Requirement: Workflow is genuinely meaningful to you/your context with clear justification

Check	Evidence	Status
Real workflow (not fictional)	LinkedIn thought leadership is explicitly listed in the assignment examples ("Drafting and iterating LinkedIn posts")	✓
Meaningful to personal context	Section 1 of Tutorial explains the personal pain point: leaders have expertise but can't maintain consistent LinkedIn presence due to overhead of planning a cohesive series	✓
Can be run at least twice on real inputs	Run 1: SQL Query Performance (Q1 2026 series). Run 2: Data Cleanliness. Run 3: AI Chatbot to Dev Tool (live test, Feb 2026)	✓
Clear justification for agentic approach	Tutorial §1.2: explains why single-prompt LLM fails (generic, disconnected) vs. multi-stage pipeline (authentic, cohesive)	✓

Self-Score: Excellent (1/1)

Category 2: System Design (4 pts)

Requirement: Distinct agents/nodes/skills with clear roles, logical flow, and thoughtful integration

Check	Evidence	Status
Minimum 3 distinct components	9 skills, each with a single distinct responsibility (verified by Skill 0-8 files)	✓

Clear roles — no overlap	Skills are separated by concern: interview (1), planning (2), drafting (3), voice (4), format (5), evaluate (6), archive (7), publish (8)	✓
Logical flow	Sequential pipeline with explicit phase gating documented in Skill 0 and Tutorial §2.3	✓
Thoughtful integration	HITL checkpoints with routing logic (voice issue → Skill 4, structure → Skill 3). State management documented.	✓
Architecture diagram	diagrams/architecture_infographic.html — A4 landscape reference architecture	✓
Pipeline flow diagram	ASCII flow diagram in Tutorial §2.3 + Design Document §5	✓

Self-Score: Excellent (4/4)

Category 3: Implementation Quality (3 pts)

Requirement: Runs end-to-end reliably, outputs are usable and well-formatted. Prompts are well-structured with at least one meaningful iteration and honest critique.

Check	Evidence	Status
Runs end-to-end	3 complete runs documented (SQL, Data Cleanliness, AI Chatbot to Dev Tool)	✓
Usable, well-formatted outputs	outputs/Q1_2026_SQL_Performance_Series.md, outputs/test_run_AI_Dev_Tool_Series.md — 6 ready-to-post LinkedIn posts per run	✓
Prompts well-structured	Each skill has: name, purpose, inputs, outputs, step-by-step workflow, constraints, failure modes, examples	✓
At least one meaningful prompt iteration	Tutorial §4.1: Voice Refiner iterated 3 times (v1: "sound professional" → v2: named patterns → v3: detection table). Voice score jumped from 3.0 to 4.7	✓
Honest critique of prompt quality	Tutorial §4.3: explicit strengths and weaknesses. Identifies hard-coded voice profile and vague input dependency as weaknesses	✓

Self-Score: Excellent (3/3)

Category 4: Real Usage + Iteration (2 pts)

Requirement: Clear evidence of 2+ real runs with documented changes and reflections

Check	Evidence	Status
2+ real runs	Run 1 (SQL), Run 2 (Data Cleanliness), Run 3 (AI Chatbot, live published)	✓
Inputs documented	Topic, audience, anecdote, CTA, tone documented for each run in archives and Tutorial §5	✓
Outputs documented	Full 6-post series outputs saved in outputs/ directory	✓
Changes after first run	Tutorial §5.1: added "In the era of AI" to banned list; added system diagram suggestion to Skill 5; revised Week 5 draft	✓
Changes after second run	Tutorial §5.2: added CTA pushback prompt to Skill 1; noted Week 5 vision weakness for future improvement	✓
Whether system helped	Tutorial §7 "Would I keep using this?": Yes — saves 4-5 hours per series, more cohesive than scattered sessions	✓

Where it failed/slowed	Tutorial §7 "What Didn't Work": Week 5 generic, voice profile one-size-fits-all, no multi-format output	✓
------------------------	---	---

Self-Score: Excellent (2/2)

Category 5: Tutorial Writing Overall Quality (3 pts)

Requirement: Clear, well-structured, genuinely replicable — covers problem statement, building process, prompts, and benchmarking

Check	Evidence	Status
Replicable by a new reader	Tutorial §8 "How to Replicate": prerequisites, 9 numbered steps, MCP setup instructions	✓
Problem statement	Tutorial §1: pain point, current status quo, why agentic, tech stack	✓
Building process documented	Tutorial §3: tools & timeline table, 3 key bottlenecks, decisions made	✓
Prompts documented	Tutorial §4: exact prompts shown for Skills 2 and 4, before/after iterations	✓
Benchmarking	Tutorial §6: methodology, summary table, 4 key revelations	✓
Figures and links	Architecture infographic, ASCII flow, benchmarking table, GitHub link	✓
Honest discussion of bottlenecks	Tutorial §3.2: 3 explicit bottlenecks with root causes and fixes	✓
Submitted as PDF	TUTORIAL_WRITEUP.pdf generated via md-to-pdf	✓

Self-Score: Excellent (3/3)

Category 6: Benchmark Rigor (2 pts)

Requirement: 2+ test cases, clear metrics, edge cases, and insightful failure analysis

Check	Evidence	Status
2+ test cases	4 test cases: SQL Performance, Data Cleanliness, WAL Protocol (edge), Vague Input (ambiguous)	✓
Baseline comparison	Single-prompt GPT-4 run documented with actual output excerpts. Agentic: 4.5/5.0 vs. Baseline: 2.0/5.0	✓
Clear metrics defined	5 metrics with 1-5 scale anchors: Actionability, Voice Consistency, Strategic Depth, Narrative Cohesion, LinkedIn Optimization	✓
Edge case included	WAL Protocol — highly technical niche topic. Score: 3.4/5.0, Voice Consistency failed (2.5)	✓
Ambiguous case included	Vague Input ("Improve your relationship with data") — Score: 3.6/5.0, Skill 1 interview rescued it	✓
Success criteria defined before testing	Rubric frozen before running test cases (documented in Benchmark Appendix §1)	✓
Prompts frozen across test cases	Benchmark Appendix §9: reproducibility notes, all settings documented	✓
Aggregate results reported	Benchmark Appendix §7: aggregate results table per metric	✓

Worst failure reported	Benchmark Appendix §8: WAL Protocol Week 5, Voice Consistency 2.0 — root cause and recommendation documented	✓
Inputs and outputs shown	Baseline output excerpts and agentic excerpts shown side-by-side in Benchmark Appendix §3	✓

Self-Score: Excellent (2/2)

Bonus: MCP Tool (+1 pt)

Requirement: Working MCP tool that meaningfully extends system capabilities

Check	Evidence	Status
MCP tool included	@ldraney/mcp-linkedin installed and configured in ~/.claude/settings.json	✓
Meaningfully extends capabilities	Direct LinkedIn API posting — Skill 8 documents linkedin_create_post, scheduling, and image posting tools	✓
Working (end-to-end)	Week 1 of AI Chatbot to Dev Tool series published live via LinkedIn API (direct Node.js call) — URN: urn:li:share:7430329696413843458	✓
Documented	Skill 8 documents full OAuth flow, MCP setup, direct API fallback, error handling	✓
Caveat	MCP server tools did not load into Claude Code session (known issue). Fallback to direct LinkedIn API call used. Post was published successfully.	⚠

Self-Score: Satisfactory → Excellent edge case. MCP published successfully via direct API; Claude Code MCP loading issue is a platform limitation, not a skill design issue.

Bonus: Demo Video (+1 pt)

Requirement: 2-5 min demo showing system end-to-end

Check	Evidence	Status
Demo recorded	Not recorded	✗
Input shown	—	—
Workflow/orchestration shown	—	—
Final output shown	—	—
Evaluation criterion mentioned	—	—

Self-Score: 0/1 — Demo not submitted.

Recommendation: Record a 3-minute walkthrough showing:

- 1. The GitHub repo structure (30 sec)
- 2. The test run archive file — show all 6 posts (60 sec)
- 3. The live LinkedIn post for Week 1 (30 sec)
- 4. The benchmark scorecard and explain one finding (60 sec)

Point Summary

Category	Max	Self-Assessed	Evidence Quality
Problem Selection	1	1	Strong — LinkedIn posts explicitly in assignment examples
System Design	4	4	Strong — 9 distinct skills, clear roles, architecture diagram
Implementation Quality	3	3	Strong — 3 runs, documented iterations, honest critique
Real Usage + Iteration	2	2	Strong — SQL + Data Cleanliness + live AI Dev Tool run
Tutorial Writing Quality	3	3	Strong — replicable, figures, benchmarking, bottlenecks
Benchmark Rigor	2	2	Strong — 4 test cases, baseline, edge/ambiguous, failure analysis
Subtotal	15	15	
Bonus: MCP Tool	1	0.5-1	Partial — published live but MCP server had loading issues
Bonus: Demo Video	1	0	Not submitted
TOTAL	17	15.5-16	

Gaps to Address Before Final Submission

Priority	Gap	Recommendation
HIGH	Demo video not recorded	Record 3-minute walkthrough of repo + live post + benchmark (30-60 min effort)
MEDIUM	Tutorial does not include exact prompts for all 9 skills inline	Skills are in repo; could add appendix to Tutorial with all 9 skill files printed
LOW	Single rater bias in benchmark	Acknowledged in Tutorial §6.3 and Benchmark Appendix §6 — no action needed for submission
LOW	Week 5 vision weakness	Acknowledged as known limitation in Tutorial §7 — no action needed

Path A Skill Format Compliance

Each skill must include: name, purpose, inputs/outputs, step-by-step behavior, constraints/failure modes, example usage.

Skill	Name & Purpose	Inputs/Outputs	Step-by-Step	Constraints	Failure Modes	Example
skill_0	✓	✓	✓	✓	✓	✓
skill_1	✓	✓	✓	✓	✓	✓
skill_2	✓	✓	✓	✓	✓	✓
skill_3	✓	✓	✓	✓	✓	✓
skill_4	✓	✓	✓	✓	✓	✓
skill_5	✓	✓	✓	✓	✓	✓
skill_6	✓	✓	✓	✓	✓	✓
skill_7	✓	✓	✓	✓	✓	✓

skill_8						
---------	--	--	--	--	--	--

All 9 skills are fully compliant with Path A format requirements.

Canvas Submission Checklist

Item	File	Status
System artifact link	https://github.com/taashim-eng/linkedin-thought-leadership-agent	Ready
Tutorial Write-up PDF	TUTORIAL_WRITEUP.pdf	Ready
Benchmark Appendix	benchmark/BENCHMARK_APPENDIX.md (or as PDF)	Ready
Demo Recording URL	Not available	Optional

Self-assessment completed: February 2026. This checklist is for personal reference only — not a required submission.