

LinkedIn Thought Leadership Agent

An agentic AI skills pack that transforms a single strategic intent into a 6-week LinkedIn content series — complete with narrative arcs, voice consistency, human-in-the-loop checkpoints, and automated posting via MCP.

MSIS 549 B — Machine Learning & AI for Business Applications University of Washington, Winter 2025-2026 | Homework 2: Agentic AI for Real-World Impact

The Problem

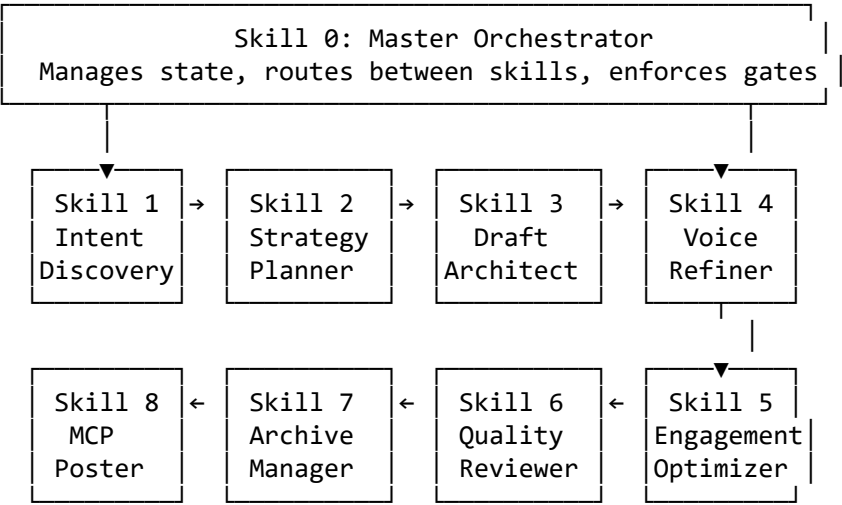
Creating consistent, high-quality LinkedIn thought leadership is time-consuming and cognitively demanding. Most professionals default to one of two failure modes:

- 1. **Generic AI output** — "In today's data-driven world..." posts that all sound the same
- 2. **Sporadic posting** — starting strong, then going silent for weeks

This system solves both by decomposing content creation into specialized skills that enforce quality, consistency, and strategic coherence across a 6-week publishing cadence.

Architecture

The system follows **Path A** (Agentic Skills Pack) with **9 specialized skills** orchestrated by a master controller:



HITL Checkpoints: — 5a (Content Review) — 6a (Benchmark Gate) —

Skills Summary

| # | Skill | Purpose |
|---|---------------------|---|
| 0 | Master Orchestrator | Pipeline control, state management, error recovery |
| 1 | Intent Discovery | 5-question interview to extract strategic intent |
| 2 | Content Strategist | 6-week narrative arc with Hook→Framework→Story→Tactics→Vision→CTA structure |
| 3 | Draft Architect | First drafts with anti-AI-ism rules (no "In today's...", no "Let's dive in...") |

| | | |
|---|----------------------|---|
| 4 | Voice & Tone Refiner | Detects and replaces 15+ AI patterns with human voice |
| 5 | Engagement Optimizer | LinkedIn-specific optimization (hooks, CTAs, hashtags, visuals) + HITL 5a |
| 6 | Quality Reviewer | 5-metric scoring rubric (1-5 scale) + HITL 6a benchmark gate |
| 7 | Archive Manager | Structured archival with full metadata for reproducibility |
| 8 | Poster & Reviewer | LiGo MCP integration for automated LinkedIn posting |

Key Design Decisions

- 1. **Anti-AI-ism as a first-class concern** — Skills 3-4 contain explicit banned-pattern lists and replacement tables
- 2. **Two mandatory HITL checkpoints** — Content never publishes without human approval
- 3. **Narrative cohesion by design** — Skill 2 plans the full 6-week arc before any drafting begins
- 4. **MCP integration** — Skill 8 uses LiGo MCP for direct LinkedIn posting (with manual fallback)

Benchmark Results

| Test Case | Actionability | Voice | Depth | Cohesion | LinkedIn | Overall |
|----------------------------|---------------|-------|-------|----------|----------|---------|
| SQL Performance (Agentic) | 4.2 | 4.7 | 4.3 | 4.8 | 4.3 | 4.5 |
| Data Cleanliness (Agentic) | 4.2 | 4.2 | 4.0 | 4.5 | 4.0 | 4.2 |
| WAL Protocol (Edge Case) | 3.5 | 2.5 | 4.0 | 3.8 | 3.2 | 3.4 |
| Vague Input (Ambiguous) | 3.5 | 3.8 | 3.2 | 4.0 | 3.5 | 3.6 |
| Single-Prompt Baseline | 2.5 | 2.2 | 1.8 | 1.0 | 2.5 | 2.0 |

Key finding: The agentic system outperformed the single-prompt baseline by **+2.5 points** on the primary test case (4.5 vs 2.0). Biggest improvement was in Narrative Cohesion (+3.8), which a single prompt fundamentally cannot achieve.

See [benchmark/BENCHMARK_APPENDIX.md](#) for full methodology, scoring rubrics, and failure analysis.

Repository Structure

```
├── README.md # This file
├── TUTORIAL_WRITEUP.md # Full tutorial (assignment submission)
├── skills/
│   ├── skill_0_master_orchestrator.md # Pipeline orchestration
│   ├── skill_1_intent_discovery.md # Strategic intent interview
│   ├── skill_2_content_strategist.md # 6-week arc planning
│   ├── skill_3_draft_architect.md # First draft generation
│   ├── skill_4_voice_tone_refiner.md # AI pattern detection & replacement
│   ├── skill_5_engagement_optimizer.md # LinkedIn optimization + HITL 5a
│   ├── skill_6_quality_reviewer.md # 5-metric scoring + HITL 6a
│   ├── skill_7_archive_manager.md # Structured archival
│   └── skill_8_poster_reviewer.md # LiGo MCP integration
├── benchmark/
│   └── BENCHMARK_APPENDIX.md # Full benchmark methodology & results
├── outputs/
│   ├── Q1_2026_SQL_Performance_Series.md # Generated 6-week series
│   └── 2026_Roadmap_Plan.md # Annual content roadmap
└── diagrams/
```

```
|   | architecture_infographic.html # A4 landscape architecture diagram
|   | archive/                      # Session archives (generated at runtime)
```

How to Use

Prerequisites

- Claude Code (or any Claude-based IDE with skill file support)
- Optional: [LiGo MCP](#) for automated LinkedIn posting

Quick Start

1. Clone this repository
2. Copy the `skills/` directory to your Claude Code skills folder (`~/.claude/skills/`)
3. Start a conversation and say: *"I want to create a LinkedIn thought leadership series about [your topic]"*
4. The orchestrator will guide you through:
 - **Intent Discovery** — 5-question interview to clarify your topic, audience, and message
 - **Content Strategy** — A 6-week narrative arc tailored to your intent
 - **Draft Generation** — 6 polished posts with anti-AI-ism enforcement
 - **Human Review** — Two checkpoints where you approve or revise content
 - **Publishing** — Direct to LinkedIn via MCP or manual copy/paste

Example Input

Topic: Optimizing SQL Query Performance

Audience: Data Engineers & Business Stakeholders

Core Message: Continuous improvement in SQL is essential for AI-readiness

Anecdote: Business team frustrated when data wasn't in sync with AI models

Tone: Provocative + Educational

Example Output

See [outputs/Q1_2026_SQL_Performance_Series.md](#) for the full 6-week series generated from the input above.

Tech Stack

- **LLM:** Claude (via Manus AI agentic platform)
- **MCP Integration:** LiGo MCP for LinkedIn API access
- **Evaluation:** Human rubric scoring (5-metric, 1-5 scale)
- **Architecture:** Path A — Agentic Skills Pack (markdown skill files)

Known Limitations

1. **Voice Consistency on technical topics** — Skill 4's voice refiner is optimized for Manager/Director audiences. Deeply technical audiences (e.g., database kernel engineers) score lower (2.5/5) because the refiner over-simplifies jargon.
2. **Single-rater evaluation** — Benchmark scores are from one human evaluator. Inter-rater reliability with 2+ evaluators would strengthen the results.
3. **No A/B testing** — Posts haven't been tested against actual LinkedIn engagement metrics yet.

License

MIT

Author

Taashi Manyanga — University of Washington, MSIS 549 B (Winter 2025-2026)