NASA MCP Agent — System Blueprint (v0.1)

#0) Purpose
Design and deploy a continuous intelligence agent that ingests
NASA/partner publications, standards, and datasets, then surfaces
alignment actions and compliance deltas for our programs (SDC & COMMS,
Space Data Center; Autonomous House energy R&D where relevant to space
power; and future lunar relay initiatives). The agent normalizes
sources into a shared ontology, maps them to requirements (e.g., LNIS,
ICSIS, SCaN policies), and produces actionable briefs and ICD-ready
references.

#1) Scope & Success Criteria
In-scope: NASA/partner public docs, specs, selected APIs;
automated retrieval; semantic parsing; requirement mapping; change
detection; weekly briefs; dashboard; exportable citations.

Out-of-scope (v0.1): non-public portals; ITAR/EAR material; human-
in-the-loop proposal drafting (provided later as add-on module);
automated bid submission.

Success (90 days)
- Daily crawls of core sources; <24h ingestion lag.
- LNIS/ICSIS/SCaN requirement map covering ≥80% of clauses that affect
comms/PNT/routing.
- Change alerts <12h from new LNIS/ICSIS release.
- Alignment brief auto-generated weekly with citations;
red/amber/green per subsystem.

#2) Source Catalog (Phase 1)
1. **NASA NTRS / STI**: Technical reports, conference papers, program
updates (JSON/HTML scraping + PDF parse).
2. **SCaN & LunaNet**: LNIS (v5 baseline), drafts/errata, LCRNS
overview pages and testbed notes.
3. **IOAG/UNOOSA/ESA LunaNet**: PNT frameworks, ICSIS, Moonlight.
4. **NASA Open Source & APIs**: code.nasa.gov, API endpoints
(metadata, MDPI imagery when needed), Open MCT for context.
5. **Partner Standards**: CCSDS, SFCG references; public spectrum
policy notes.

> Phase 2 adds: JPL/GSFC public GitHub, relevant SBIR topics, AFRL/UNP
Mission Concept Program outputs, peer-reviewed comms/optical papers.

#3) Ontology & Schema
Create a **NASA Alignment Ontology (NAO)** with these top-level
classes:
- **Program/Standard** (e.g., LNIS, ICSIS, SCaN-MOCS) → hasVersion,
hasSection, supersedes, dependsOn
- **Service** (Comms, PNT/Timing, DTN, Network Mgmt, Optical, RF)
- **Interface/ICD** (signal-in-space, API, protocol stack, timing

reference)
- **Requirement** (id, text, rationale, status, verification method)
- **Artifact** (pdf, html, code, dataset) $\rightarrow$ sourceURL, checksum, date
- **Subsystem** (Gateway node, Orion user terminal, Relay, Ground segment, Optical terminal)
- **Design Element** (antenna, modem, laser terminal, pointing, routing policy, DTN node, security)
- **Impact** (aligns, conflicts, extends, TBD)

Mapping tables link NAO entities to our project components (e.g., Optical Mesh Router $\rightarrow$ Service: Optical Comms; Interface: Acquisition/Tracking; Protocol: DTN BPv7; Requirement refs: LNIS §5.x, ICSIS §4.x).

#4) Pipeline Architecture
A. Ingestion
- Web fetchers (RSS/HTML/robots-aware), PDF downloader, rate-limited.
- Format detectors (PDF, HTML, PPT/PDF exports) $\rightarrow$ queue to parsing.

B. Parsing & Normalization
- PDF text extraction (layout-aware, figure/table capture), OCR fallback.
- Metadata enrichment (title, authors, date, org, standard version, section IDs).
- Chunking to ~2–5k tokens with structural anchors (headings, numbered clauses).

C. Intelligence Layer
- Embeddings index (hybrid BM25 + vector) per corpus.
- Clause classifier: requirement vs narrative vs background.
- Rule tagger for LNIS/ICSIS/SCaN sections; regex + ML few-shot.
- Knowledge graph: NAO triples stored in graph DB.

D. Alignment Engine
- Crosswalk rules: map requirement clauses $\rightarrow$ project subsystems.
- Delta detector: new/changed clauses vs last baseline; semantic diff.
- Risk scorer: impact x confidence; flags for human review.

E. Outputs
- Weekly Alignment Brief (PDF/MD) with RAG status per subsystem.
- Change Log with diffs and citations.
- ICD Reference Pack: curated excerpts with section IDs for engineering.
- JSON export for MECSAI and ARCHITECT agents.

#5) Tech Stack
- **Storage**: Postgres (artifacts & metadata), MinIO/S3 (docs), Neo4j or GraphDB (NAO), OpenSearch (BM25), vector DB (FAISS/pgvector).
- **Ingest/ETL**: Airflow/N8n jobs; retry & audit logs.
- **Parsing**: GROBID + pdfminer + layout-parser; Tesseract OCR fallback.
- **NLP**: Embeddings model (local OSS-20B embed or OpenAI text-

embedding-3-large); lightweight zero-shot/few-shot classifier.
- **Services**: FastAPI for query; Model Context Protocol (MCP) server shim to expose tools to agents; auth via API keys.
- **UI**: Open MCT (optional) for timeline context; React dashboard for alerts.

#6) Security & Compliance
- Public-source only; block ITAR/EAR via allowlist of domains.
- Immutable artifact store with checksums and provenance.
- License tags; citation preservation; robots.txt compliance.

#7) Alignment Rules (Initial Set)
- **LNIS v5**: Services, Interfaces, LNSP-to-LNSP, Security, PNT/Augmented Forward Signal (AFS), DTN usage.
- **ICSIS Rev B**: Interoperability constraints across agencies; messaging, timing, and signal alignment.
- **SCaN-MOCS**: Service definitions, performance tiers, scheduling policies, spectrum references.
- **LCRNS**: Testbed constraints and validation targets for lunar relay services.

Each rule encodes: `source`, `section`, `pattern(s)`, `mapping`, `verification`, `impact`.

#8) Deliverables
1. **MCP Agent Service** (containerized) with REST + MCP tools.
2. **Baseline Crawl & Index**: LNIS v5, ICSIS Rev B, SCaN-MOCS, LCRNS pages.
3. **NAO Graph** populated with top 500 requirements.
4. **Weekly Brief Template** and **Change-Detection Alerts**.
5. **ICD Reference Pack** export (MD/PDF).

#9) Milestones & Timeline (first 8 weeks)
- **Week 1–2**: Ingestion scaffolding; storage; baselines seeded; NAO v0.1.
- **Week 3–4**: Parsers + classifiers; first requirement maps; dashboard v0.
- **Week 5–6**: Delta detection; RAG scoring; MCP tool endpoints.
- **Week 7–8**: QA, dry-run briefs; tune precision/recall; export packs.

#10) Interfaces to Our Agent Swarm
- **ARCHITECT/MECHWORK**: Pull ICD references; subscribe to clause deltas.
- **MECSAI**: Consume JSON rule updates; trigger engineering tasks on spec change.

- **ATB (SYS-SAFE, ICS-SEC)**: Review high-impact deltas; annotate mitigations.

#11) Future Upgrades
- SBIR/STTR opportunity watcher with eligibility matching.
- Ground-station availability + optical weather feed for ops planning.
- Human-in-the-loop summarization for proposals; citation autostitching.
- Spectrum coordination assistant (NTIA/ITU/SFCG trackers).

#12) Open Questions / Risks
- Variability of PDF structure; need robust parsers.
- Version tracking for evolving standards (mirror + semantic diff).
- Disambiguating similarly named artifacts; require strong metadata.

#13) Next Actions
- Approve source list & ontology v0.1.
- Stand up storage + ingest.
- Seed baselines (LNIS v5, ICSIS Rev B, SCaN-MOCS, LCRNS).
- Ship first weekly brief (dry run) with 10 sample requirements mapped to SDC & COMMS optical mesh.