Complete Testing Resources Guide

NLP + LLM Insurance Decision System

📋 Required Document Types for Testing

🏥 1. HEALTH INSURANCE DOCUMENTS

Primary Documents (Must Have)

1. **Sample Health Insurance Policy PDF**
   - CMS Summary of Benefits and Coverage Sample
   - **Link**: https://www.cms.gov/cciio/resources/forms-reports-and-other-resources/downloads/sbc-sample-completed-mm-508-fixed-4-12-16.pdf
   - **Content**: Medical coverage, deductibles, copayments, exclusions

2. **Personal Health Insurance Policy Sample**
   - Scribd Personal Health Insurance Policy
   - **Link**: https://www.scribd.com/document/467931285/slf-personal-health-insurance-enhanced-sample-en
   - **Content**: Drug coverage, extended health benefits, vision, dental

3. **Health Insurance Coverage Summary**
   - **Link**: https://www.cms.gov/CCIIO/Resources/Regulations-and-Guidance/Downloads/SBC-Sample-Completed.pdf
   - **Content**: Plan details, coverage limits, network providers

Additional Health Documents

4. **Patient Advocate Health Insurance Guide**
   - **Link**: https://www.patientadvocate.org/wp-content/uploads/Understanding-Your-Health-Insurance-Policy-Documents-Tip-Sheet.pdf
   - **Purpose**: Understanding policy documents, claim procedures

5. **Health Insurance Forms Collection**
   - Sample Forms - Health Insurance Forms
   - **Link**: https://www.sampleforms.com/health-insurance-form.html
   - **Content**: Claim forms, coverage applications

🚑 2. VEHICLE INSURANCE DOCUMENTS

Primary Documents

6. **Vehicle Insurance Policy Format**
   - Scribd Vehicle Insurance Policy
   - **Link**: https://www.scribd.com/document/184902535/Vehicle-Insurance-Policy-Format
   - **Content**: Policy details, coverage limits, liability information

7. **Car Insurance Certificate Template**
   - **Link**: https://certificate-of-motor-insurance.pdffiller.com/
   - **Purpose**: Standard vehicle insurance format

🏠 3. GENERAL INSURANCE DOCUMENTS

Multi-Purpose Templates

8. **Insurance Policy Examples Collection**
   - Examples.com - Insurance Policy Templates
   - **Link**: https://www.examples.com/docs/insurance-policy.html
   - **Content**: Various insurance types and formats

9. **Insurance PDF Templates (JotForm)**
   - JotForm Insurance Templates
   - **Link**: https://www.jotform.com/pdf-templates/insurance
   - **Content**: Certificates, quotes, waivers, policies

10. **Life Insurance Policy Sample**
    - **Link**: https://healthsourceri.com/wp-content/uploads/2016/11/Principal-Sample-Life-Insurance-Policy.pdf
    - **Content**: Life insurance terms, benefits, exclusions

---

🧪 4. TESTING DATASETS & SCENARIOS

A. Sample Test Queries (Create these scenarios)

Scenario 1: Basic Medical Claim

json

```json
{
  "query": "I'm 45 years old, had knee surgery in Mumbai, policy is 8 months old. Am I covered?",
  "expected_entities": {
    "age": 45,
    "procedure": "knee surgery",
    "location": "Mumbai",
    "policy_duration": 8
  },
  "expected_decision": "APPROVED"
}
```

Scenario 2: Waiting Period Violation

```json
{
  "query": "32-year-old female, appendectomy, 2-month-old policy",
  "expected_entities": {
    "age": 32,
    "gender": "female",
    "procedure": "appendectomy",
    "policy_duration": 2
  },
  "expected_decision": "REJECTED"
}
```

Scenario 3: Age Limit Check

```json
{
  "query": "I'm 67 years old, need cataract surgery, had policy for 3 years",
  "expected_entities": {
    "age": 67,
    "procedure": "cataract surgery",
    "policy_duration": 36
  },
  "expected_decision": "REQUIRES_REVIEW"
}
```

Scenario 4: Complex Query

json

```json
{
  "query": "My husband is 52, had cardiac bypass in Delhi hospital, our family policy POL12345 is 18 months old, claim a
  "expected_entities": {
    "age": 52,
    "gender": "male",
    "procedure": "cardiac bypass",
    "location": "Delhi",
    "policy_number": "POL12345",
    "policy_duration": 18,
    "claim_amount": 450000
  },
  "expected_decision": "APPROVED"
}
```

📑 5. REFERENCE DOCUMENTS

Insurance Regulation & Compliance

11. NIOS Insurance Services Documentation

- NIOS Insurance Documents
- **Link**: https://nios.ac.in/media/documents/VocInsServices/m3-f2.pdf
- **Purpose**: Understanding insurance regulations

12. Cornell University Insurance Certificate

- **Link**: https://scl.cornell.edu/sites/scl/files/documents/OSFL-coi-%20sample.pdf
- **Purpose**: Certificate formats and requirements

13. Title Insurance Sample

- **Link**: https://www.wieb.uscourts.gov/sites/default/files/chambers/svk/LouJones/Sample-Owner's Policy of Title Insurance.pdf
- **Purpose**: Understanding policy structures

🔧 6. DEVELOPMENT RESOURCES

NLP Libraries & Models

```
python
```

```
# Required Python packages
pip install spacy
pip install transformers
pip install sentence-transformers
pip install pandas
pip install nltk
pip install re
pip install dateutil

# Download spaCy models
python -m spacy download en_core_web_sm
python -m spacy download en_core_web_lg
```

Sample Training Data (Create CSV files)

entities_training.csv

```csv
text,age,gender,location,procedure,policy_duration,amount
"45-year-old male knee surgery Mumbai 8 months",45,male,Mumbai,knee surgery,8,
"32 female appendectomy 2-month policy",32,female,,appendectomy,2,
"67 years old cataract surgery 3 years policy",67,,cataract surgery,36,
```

business_rules.csv

```csv
procedure_type,waiting_period_months,age_min,age_max,coverage_limit
medical_procedures,6,18,65,500000
dental,12,18,65,100000
cosmetic,24,21,60,200000
emergency,0,0,100,1000000
```

Configuration Files

config.json

```json

```

```json
{
  "chunk_size": 1000,
  "chunk_overlap": 100,
  "max_retrieved_docs": 8,
  "similarity_threshold": 0.1,
  "llm_model": "llama-3.3-70b-versatile",
  "embedding_model": "sentence-transformers/all-MiniLM-L6-v2",
  "waiting_periods": {
    "medical": 6,
    "dental": 12,
    "pre_existing": 24
  },
  "age_limits": {
    "min": 18,
    "max": 65,
    "senior_review": 60
  }
}
```

🧮 7. EVALUATION DATASETS

Create Test Cases (test_cases.json)

json

```json
[
  {
    "id": "TC001",
    "query": "I'm 46 years old and I just had knee surgery in Pune. My insurance policy is only 3 months old—am I covere
    "expected_result": {
      "decision": "REJECTED",
      "reason": "Waiting period not met",
      "entities_extracted": {
        "age": 46,
        "procedure": "knee surgery",
        "location": "Pune",
        "policy_duration": 3
      }
    }
  },
  {
    "id": "TC002",
    "query": "28-year-old woman, dental cleaning, 15-month policy",
    "expected_result": {
      "decision": "APPROVED",
      "reason": "All criteria met",
      "entities_extracted": {
        "age": 28,
        "gender": "female",
        "procedure": "dental cleaning",
        "policy_duration": 15
      }
    }
  }
]
```

📊 8. PERFORMANCE BENCHMARKS

Accuracy Metrics to Track

- **Entity Extraction Accuracy**: >85%

- **Decision Accuracy**: >90%

- **Processing Speed**: <3 seconds per query

- **Document Retrieval Precision**: >80%

- **Business Rules Compliance**: 100%

Test Coverage Requirements

- ✅ 50+ different query variations
- ✅ All entity types (age, gender, location, procedure)
- ✅ All decision types (approved, rejected, pending)
- ✅ Edge cases (missing info, unclear queries)
- ✅ Multiple document types (health, auto, life)

🎯 9. DEMONSTRATION SETUP

Demo Script Examples

```python
# Demo Query 1: Clear Approval
demo_queries = [
    "I'm 35, had eye surgery in Chennai, 12-month-old policy",
    "45-year-old male, cardiac procedure, policy 2 years old, claim ₹3,00,000",
    "My daughter is 16, dental treatment, family policy 8 months old",
    "67-year-old senior, hip replacement, Medicare supplement"
]
```

Expected Demo Flow

1. **Upload Documents** → Show policy PDFs being processed
2. **Enter Query** → Natural language input
3. **Show Processing** → Entity extraction, document retrieval
4. **Display Decision** → JSON output with reasoning
5. **Explain Results** → Why approved/rejected

⚡ 10. QUICK START CHECKLIST

Phase 1: Document Collection (Week 1)

☐ Download all PDF documents from links above
☐ Create test query dataset (50+ examples)
☐ Set up document processing pipeline

Phase 2: NLP Implementation (Week 2-3)

☐ Implement entity extraction
☐ Set up semantic search
☐ Test document chunking and retrieval

Phase 3: Business Logic (Week 4-5)

☐ Implement business rules engine
☐ Create decision making logic
☐ Add structured JSON output

Phase 4: Integration & Testing (Week 6-7)

☐ Integrate all components
☐ Run test cases and measure accuracy
☐ Debug and optimize performance

Phase 5: Demo Preparation (Week 8)

☐ Create web interface
☐ Prepare demonstration scenarios
☐ Document results and create presentation

🎓 Academic Requirements Coverage

NLP Components Demonstrated

✅ **Named Entity Recognition** - Extract age, location, procedures ✅ **Text Processing** - Document chunking, cleaning, tokenization
✅ **Semantic Search** - Vector embeddings, similarity matching ✅ **Information Extraction** - Structure data from unstructured text

LLM Integration Shown

✅ **Query Understanding** - Interpret natural language queries ✅ **Contextual Reasoning** - Apply business logic with context ✅ **Decision Making** - Generate approve/reject decisions ✅ **Explanation Generation** - Provide human-readable justifications

Technical Innovation

✅ **Hybrid AI System** - Combines rule-based and ML approaches ✅ **Domain-Specific Application** - Insurance industry focus ✅ **Structured Output** - JSON responses for system integration ✅ **Explainable AI** - Clear reasoning and source attribution

---

Final Note: This comprehensive resource list provides everything needed to build, test, and demonstrate your NLP + LLM insurance decision system. Start with downloading the core insurance policy documents, then systematically implement each component while testing against the provided scenarios.