Assignment 1: QA System and Knowledge Graph Specification

Task 1: Scenario and Domain Selection

Selected Domain: Hospital Resource Management

For this assignment, I've chosen **Hospital Resource Management** as my domain. The focus here is on managing and optimizing hospital resources - everything from staff and equipment to facilities and supplies.

Why This Domain Makes Sense

I think hospital resource management works really well for a GraphRAG system for several reasons:

High Business Value
Hospitals are always under pressure to use their resources efficiently while keeping patient care quality high. When resources aren't allocated properly, you see increased costs, burned-out staff, treatment delays, and unhappy patients. This is a real problem that needs solving.

Naturally Interconnected Data
What's interesting about hospitals is how everything connects. You have staff (doctors, nurses, technicians) who need specific equipment (MRI machines, ventilators) in particular facilities (operating rooms, ICU units) across different departments, all on various schedules. These relationships are exactly what knowledge graphs are good at representing.

Available Data Sources
Hospitals already collect tons of structured data across different systems. There are HR databases with staff credentials and schedules, asset management systems tracking equipment, EHRs with patient volumes, facility management systems showing room availability, and supply chain databases. All this data exists - it just needs to be connected properly.

Established Standards
Healthcare has mature ontologies we can build on, like ICD-10/11 for procedures, SNOMED CT for clinical terms, HL7 standards, and various equipment taxonomies. This means I won't be starting from scratch.

Clear User Need
Hospital administrators and operations managers are constantly trying to answer complex

questions about resource availability, utilization patterns, and where bottlenecks are occurring. Right now, they have to manually pull data from multiple systems, which is time-consuming and error-prone.

Task 2: System Description

What the System Does

The system I'm proposing is basically an intelligent Q&A platform for hospital resource management. Hospital administrators and operations managers would be able to ask questions in plain English, and the system would pull from a knowledge graph containing all the hospital's resource data - staff, equipment, facilities, departments, schedules, and metrics. An LLM would interpret the questions and generate accurate answers.

Specifically, users could get:

- Real-time info on what's available right now (which staff are on duty, equipment status, bed availability)
- Historical analytics (how departments have been performing, equipment usage over time, staffing trends)
- Relationship insights (who's qualified to use what equipment, how departments depend on each other)
- Optimization suggestions (where resources are being underused, where bottlenecks exist, scheduling conflicts)
- Predictive insights (forecasting resource needs, upcoming maintenance)

Who Would Use It

I'm targeting several user groups:

Hospital Administrators
These are the senior managers dealing with strategic planning, budgets, and overall efficiency. They need high-level insights for big decisions.

Operations Managers
The people running day-to-day operations, coordinating between departments, and solving immediate resource problems. They need quick, actionable information.

Department Heads
Leaders of Emergency, Surgery, Radiology, etc. who need to understand their own resource situation and compare themselves to other departments.

Facility Managers
The folks responsible for maintaining the physical stuff - buildings, equipment, facilities.

HR Managers
People handling workforce planning, scheduling, and making sure there's adequate staffing everywhere.

Why This Would Be Valuable

I see several ways this system would create real value:

Better Decision Making
Right now, managers either go with their gut or spend hours pulling data from different systems. With this system, they'd get accurate answers to complex questions quickly, leading to smarter decisions.

Less Waste
By identifying underutilized equipment, overstaffed areas, or inefficient schedules, hospitals could cut unnecessary costs significantly.

Faster Response in Crises
When something goes wrong - an equipment failure, a staffing shortage, an emergency surge - managers could instantly query alternatives and reallocate resources instead of scrambling.

Preventing Problems Before They Happen
By seeing how things connect (like when equipment maintenance conflicts with surgery schedules), managers could prevent bottlenecks proactively.

Long-term Planning
Historical data and trends would help with strategic decisions about buying new equipment, hiring staff, expanding facilities, and budget planning.

Staying Compliant
The system could track certifications, maintenance schedules, and regulatory requirements automatically, reducing compliance risks.

Making Data Accessible
Non-technical managers could access complex operational data without needing to learn SQL or navigate multiple software systems.

Task 3: Question Space Definition

I've designed the system to handle questions at different complexity levels:

Simple Factual Queries (Direct Lookups)

1. How many ICU beds are currently available in the hospital?
2. Which MRI machines are operational today?

3. What is the capacity of the Emergency Department?
4. How many nurses are scheduled for the night shift in Ward 3?
5. When is the next maintenance scheduled for Ventilator-A12?

Relational Questions (Connections Between Entities)

6. Which staff members are certified to operate the new CT scanner?
7. What equipment is assigned to the Cardiology department?
8. Which operating rooms have access to robotic surgery equipment?
9. Which departments share resources with the Emergency Department?
10. What backup equipment is available if MRI-2 goes down?

Aggregation and Statistical Questions

11. What is the average occupancy rate of ICU beds over the past 3 months?
12. Which department has the highest equipment maintenance costs this year?
13. How many staff members have certifications expiring in the next 6 months?
14. What is the total cost of unused medical equipment in storage?

Analytical and Comparative Questions

15. Which departments are consistently understaffed compared to their patient volume?
16. What is the utilization rate of our surgical equipment compared to industry benchmarks?
17. Which pieces of equipment have the highest downtime-to-usage ratio?
18. Are there patterns in equipment failures that correlate with specific departments or time periods?

Optimization and Recommendation Questions

19. What resources could be reallocated from Department X to Department Y to improve overall efficiency?
20. Which staff scheduling changes would reduce overtime costs while maintaining coverage?

Coverage Analysis

These 20 questions cover the range I'm aiming for:

- Questions 1-5 are straightforward lookups of entity properties
- Questions 6-10 require following relationships between entities in the graph
- Questions 11-14 need aggregation - counting, averaging, summing across multiple records
- Questions 15-18 involve analysis - comparing things, finding patterns, identifying trends
- Questions 19-20 are the hardest, requiring reasoning about multiple constraints to provide recommendations

---

Task 4: Knowledge Graph Requirements

To answer the questions I've defined, here's what the knowledge graph needs to contain:

Core Entities and Properties

1. Staff

- Unique identifier
- Name and role (doctor, nurse, technician, administrator)
- Department assignment(s)
- Specializations and qualifications
- Certifications (type, expiry date, issuing body)
- Employment status (full-time, part-time, contract)
- Shift schedules (current and historical)
- Equipment operation certifications
- Contact information

2. Equipment

- Unique identifier and name
- Type/category (diagnostic, therapeutic, surgical, support)
- Manufacturer and model
- Operational status (operational, under maintenance, broken, in storage)
- Location (department, room)
- Purchase date and cost
- Maintenance schedule and history
- Usage logs (frequency, duration, department)
- Required certifications to operate
- Warranty and lifecycle information

3. Facilities

- Unique identifier
- Type (operating room, ICU bed, ward, examination room)
- Capacity (number of beds, maximum occupancy)
- Current occupancy/availability status
- Department ownership
- Equipment contained/assigned
- Scheduled usage (surgeries, procedures, maintenance)
- Physical characteristics (size, specialized capabilities)

4. Departments

- Unique identifier and name
- Type (Emergency, Surgery, Radiology, Cardiology, etc.)

- Staff roster (current assignments)
- Equipment inventory
- Facility allocations
- Budget information
- Patient volume (current and historical)
- Operating hours
- Resource sharing agreements with other departments

5. Schedules

- Staff work schedules (shifts, rotations, time-off)
- Equipment maintenance schedules
- Facility usage schedules (OR bookings, room reservations)
- Department operational schedules

6. Resource Metrics

- Utilization rates (equipment, staff, facilities)
- Occupancy rates (beds, rooms)
- Costs (equipment maintenance, staff overtime, procurement)
- Performance indicators (patient wait times, equipment downtime)
- Historical trends and patterns

Key Relationships

The graph needs to capture these important connections:

1. **Staff** ↔ **Department:** Staff members assigned to departments (sometimes multiple)
2. **Staff** ↔ **Equipment:** Staff certified to operate specific equipment
3. **Staff** ↔ **Schedule:** Staff assigned to specific shifts and time slots
4. **Equipment** ↔ **Department:** Equipment allocated to departments
5. **Equipment** ↔ **Facility:** Equipment located in or assigned to facilities
6. **Equipment** ↔ **Maintenance:** Equipment linked to maintenance records and schedules
7. **Facility** ↔ **Department:** Facilities owned or managed by departments
8. **Department** ↔ **Department:** Resource sharing relationships, dependencies
9. **Department** ↔ **Budget:** Financial allocation and spending
10. **Equipment** ↔ **Usage:** Equipment usage events with timestamps and departments

Temporal Information

Time is crucial here. The graph needs to handle:

- Historical states (what staffing levels were like last month, past equipment status)
- Current real-time status (what's happening right now)
- Future schedules and planned events (upcoming surgeries, scheduled maintenance)
- Trend data over different time periods (daily patterns, weekly cycles, monthly trends, yearly comparisons)

Derived/Computed Information

Some information needs to be calculated on the fly:

- Availability (considering real-time status, schedules, and current usage)
- Utilization rates (actual usage time divided by available time)
- Cost metrics (totals by department, by equipment type, etc.)
- Certification compliance (valid certifications vs. those expiring soon)
- Capacity metrics (current occupancy as a percentage of total capacity)
- Performance indicators (downtime percentages, overtime hours)

Data Quality Requirements

For this to work reliably, the data needs to meet certain standards:

- **Completeness:** Critical entities can't have missing information
- **Currency:** Status information must be up-to-date (real-time or near-real-time)
- **Consistency:** Relationships must work both ways (if Equipment X is in Department Y, then Department Y must show Equipment X in its inventory)
- **Validation:** The system needs to enforce constraints (like staff can't be in two places at once, equipment can't be in multiple locations)

Integration Requirements

The knowledge graph would need to pull data from several existing hospital systems:

- HR management systems (staff info, schedules)
- Asset management systems (equipment inventory, maintenance records)
- Facility management systems (room availability, bookings)
- Financial systems (budgets, costs, procurement data)
- Operational systems (usage logs, patient volumes)

Summary

What I've outlined here is a GraphRAG-based QA system for hospital resource management that tackles a real operational challenge in healthcare. By connecting a comprehensive knowledge graph of hospital resources with natural language processing, the system would let administrators and managers make data-driven decisions quickly and easily.

The knowledge graph's detailed representation of entities (staff, equipment, facilities, departments), how they relate to each other, temporal information, and various metrics gives us the foundation to answer everything from simple lookups to complex analytical questions. If implemented well, this system could meaningfully improve how hospitals use their resources, cut costs, prevent operational problems, and ultimately help deliver better patient care through smarter resource allocation.