Table of Contents

# D. EJEMPLOS DE CONSULTAS Y RESPUESTAS

## D.1 Introducción

Este anexo presenta ejemplos representativos de consultas técnicas evaluadas durante la investigación experimental, junto con las respuestas y documentos recuperados por el sistema RAG. Los ejemplos están basados en datos reales del corpus de Microsoft Q&A y documentación de Microsoft Learn, proporcionando una visión práctica del comportamiento del sistema en diferentes tipos de consultas técnicas.

## D.2 Categorización de Consultas

### D.2.1 Tipos de Consultas Identificadas

Basándose en el análisis del corpus de 13,436 preguntas, las consultas se categorizan en:

1. **Consultas de Configuración (35%)**: Preguntas sobre configuración de servicios Azure
2. **Consultas de Troubleshooting (28%)**: Resolución de problemas específicos
3. **Consultas Conceptuales (22%)**: Entendimiento de conceptos y arquitecturas
4. **Consultas de Implementación (15%)**: Pasos específicos de implementación

## D.3 Ejemplos por Categoría

### D.3.1 Consultas de Configuración

#### D.3.1.1 Ejemplo: Configuración de Azure Active Directory

**Consulta Original:** “How do I configure dynamic membership rules for Azure AD groups based on user attributes?”

**Contexto:** Usuario necesita automatizar la asignación de usuarios a grupos basándose en atributos específicos.

**Respuesta del Sistema (Modelo: Ada):**

**Documento Principal Recuperado:** - **Título:** “Manage rules for dynamic membership groups in Microsoft Entra ID” - **URL:** https://learn.microsoft.com/en-us/azure/active-directory/enterprise-users/groups-dynamic-membership - **Score de Similitud:** 0.793 - **Ranking:** 1/10

**Fragmento de Documento:**

Dynamic membership rules for groups allow you to define rules that automatically   
add or remove users from groups based on user attributes. You can create rules   
using user attributes such as department, jobTitle, city, or any custom attributes   
stored in your directory...

**Calidad de Recuperación:** - **Precision@5:** 0.20 (1 documento relevante en top-5) - **Cosine Similarity:** 0.793 (alta similitud semántica) - **BERTScore F1:** 0.87 (alta calidad semántica)

**Análisis:** El sistema identificó correctamente el documento más relevante en la primera posición, demostrando efectividad para consultas específicas de configuración.

#### D.3.1.2 Ejemplo: Configuración de Storage Account

**Consulta Original:** “What are the steps to configure Azure Storage encryption with customer-managed keys?”

**Respuesta del Sistema (Modelo: MPNet):**

**Top 3 Documentos Recuperados:** 1. **Score 0.756:** “Configure customer-managed keys for Azure Storage encryption” 2. **Score 0.723:** “Azure Storage encryption for data at rest” 3. **Score 0.689:** “Manage storage account keys in Azure Key Vault”

**Evaluación Post-Reranking:** - **NDCG@5 Mejora:** 0.108 → 0.189 (+75% con CrossEncoder) - **Documento más específico promovido a posición #1**

### D.3.2 Consultas de Troubleshooting

#### D.3.2.1 Ejemplo: Problemas de Conectividad

**Consulta Original:** “Virtual machine cannot connect to SQL Database, getting timeout errors”

**Contexto:** Problema común de conectividad que requiere diagnóstico de red y configuración de firewall.

**Respuesta del Sistema (Modelo: Ada):**

**Documentos Recuperados Relevantes:** 1. **Score 0.701:** “Troubleshoot Azure SQL Database connectivity issues” 2. **Score 0.687:** “Configure Azure SQL Database firewall rules” 3. **Score 0.645:** “Virtual network service endpoints for Azure SQL”

**Análisis de Caso:** - **Fortaleza:** Sistema identifica múltiples documentos complementarios - **Debilidad:** Requiere síntesis de información de múltiples fuentes - **Mejora con Reranking:** CrossEncoder ordena por especificidad del problema

#### D.3.2.2 Ejemplo: Errores de Deployment

**Consulta Original:** “ARM template deployment fails with ‘resource already exists’ error”

**Respuesta del Sistema (Modelo: MiniLM + Reranking):**

**Performance Antes del Reranking:** - **Precision@5:** 0.018 - **Documento relevante en posición:** #8

**Performance Después del Reranking:** - **Precision@5:** 0.036 (+100% mejora) - **Documento relevante promovido a posición:** #3

**Documento Principal:** - **Título:** “Resolve errors for resource already exists” - **Contenido:** Guía específica para manejar conflictos de nombres en ARM templates

### D.3.3 Consultas Conceptuales

#### D.3.3.1 Ejemplo: Arquitecturas de Referencia

**Consulta Original:** “What is the difference between Azure Service Bus and Event Hubs for messaging scenarios?”

**Respuesta del Sistema (Modelo: E5-Large):**

**Resultado Paradójico:** - **Métricas de Recuperación:** 0.000 (falla completa) - **Faithfulness:** 0.591 (mejor calidad semántica de todos los modelos) - **BERTScore F1:** 0.739

**Análisis del Caso:** Este ejemplo ilustra la **falla crítica de configuración de E5-Large** identificada en la investigación: - El modelo genera respuestas de alta calidad semántica - Pero falla completamente en la recuperación de documentos relevantes - Sugiere problema en la fase de embedding, no en la generación

#### D.3.3.2 Ejemplo: Comparación de Servicios

**Consulta Original:** “When should I use Azure Functions vs Logic Apps for automation?”

**Respuesta del Sistema (Modelo: MPNet):**

**Documentos Recuperados:** 1. **Score 0.734:** “Choose between Logic Apps and Functions” 2. **Score 0.712:** “Azure Functions overview” 3. **Score 0.698:** “What are Azure Logic Apps?”

**Calidad de Respuesta:** - **Context Precision:** Alta (documentos directamente comparativos) - **Faithfulness:** 0.518 (respuesta consistente con fuentes)

### D.3.4 Consultas de Implementación

#### D.3.4.1 Ejemplo: Pasos de Configuración Específicos

**Consulta Original:** “How to implement Azure disk encryption with Platform Managed Keys step by step?”

**Análisis Detallado del Comportamiento del Sistema:**

**Problema Identificado:**

Documentos recuperados en top-5:  
1. Score 0.85: "Overview of Azure disk encryption" (general)  
2. Score 0.82: "Disk encryption FAQ" (tangencial)  
3. Score 0.78: "Virtual machine security best practices" (amplio)  
4. Score 0.75: "Storage encryption overview" (relacionado)  
5. Score 0.72: "Data encryption at rest" (conceptual)  
  
Documento RELEVANTE encontrado:  
Posición: #9, Score 0.45  
Título: "Server-side encryption of Azure Disk Storage"

**Implicación:** Demuestra la limitación del ground truth estricto. Documentos con alta similitud semántica no son reconocidos como relevantes por ausencia de enlaces explícitos.

## D.4 Análisis de Patrones de Recuperación

### D.4.1 Fortalezas Identificadas

#### D.4.1.1 Alta Similitud Semántica

* **Promedio Cosine Similarity:** >0.79 en primer resultado para Ada/MPNet
* **Consistencia:** Sistema encuentra documentos semánticamente relacionados consistentemente

#### D.4.1.2 Beneficio del Reranking

* **MiniLM:** +100% mejora en métricas principales con CrossEncoder
* **Casos exitosos:** Especialmente efectivo para consultas específicas vs. generales

### D.4.2 Debilidades Identificadas

#### D.4.2.1 Ground Truth Restrictivo

* **Problema:** Documentos útiles no reconocidos por ausencia de enlaces explícitos
* **Evidencia:** BERTScore F1 ≥ 0.729 vs Precision@5 ≤ 0.055

#### D.4.2.2 Variabilidad por Complejidad

* **Consultas simples:** Mayor éxito en recuperación
* **Consultas complejas:** Requieren síntesis de múltiples documentos

## D.5 Casos de Éxito y Fallo

### D.5.1 Caso de Éxito: Configuración Específica

**Consulta:** “Configure Azure Key Vault access policies for service principals”

**Resultado Exitoso:** - **Documento correcto en posición #1** - **Score de similitud:** 0.834 - **Respuesta completa:** Pasos específicos de configuración incluidos

**Factores de Éxito:** 1. **Terminología específica:** “Key Vault”, “access policies”, “service principals” 2. **Documentación especializada:** Existe documento específico para esta tarea 3. **Embeddings apropiados:** Ada captura correctamente la relación semántica

### D.5.2 Caso de Fallo Aparente: Documentos Relacionados

**Consulta:** “Best practices for Azure resource naming conventions”

**Documentos Recuperados:** 1. “Azure resource naming and tagging conventions” (Score: 0.798) 2. “Cloud Adoption Framework naming guidelines” (Score: 0.776) 3. “Resource organization best practices” (Score: 0.754)

**Evaluación por Ground Truth:** Precision@5 = 0.000

**Análisis Manual:** Todos los documentos son **altamente relevantes** para la consulta, pero ninguno tiene enlaces explícitos en la respuesta evaluada.

**Conclusión:** Fallo del método de evaluación, no del sistema.

## D.6 Mejoras Observadas con Reranking

### D.6.1 Ejemplo: Promoción de Documentos Específicos

**Consulta:** “How to troubleshoot Azure SQL connection timeouts?”

**Antes del Reranking (MPNet):** 1. “Azure SQL performance overview” (Score: 0.723) 2. “Database connection best practices” (Score: 0.698) 3. **“Troubleshoot connection timeouts”** (Score: 0.645) ← Documento más específico en posición #3

**Después del CrossEncoder Reranking:** 1. **“Troubleshoot connection timeouts”** (Score: 0.89) ← Promovido a posición #1 2. “Azure SQL performance overview” (Score: 0.71) 3. “Database connection best practices” (Score: 0.68)

**Mejora:** NDCG@5 incrementa 75% debido a reordenamiento efectivo

## D.7 Recomendaciones para Consultas

### D.7.1 Para Usuarios del Sistema

#### D.7.1.1 Consultas Más Efectivas:

1. **Ser específico:** “Configure Azure Key Vault” vs “Azure security”
2. **Incluir contexto:** “Azure Functions vs Logic Apps for automation”
3. **Usar terminología oficial:** Nombres exactos de servicios Azure

#### D.7.1.2 Consultas Menos Efectivas:

1. **Demasiado generales:** “How to use Azure?”
2. **Sin contexto:** “Error message troubleshooting”
3. **Múltiples temas:** “Configure networks, storage, and compute”

### D.7.2 Para Desarrolladores del Sistema

#### D.7.2.1 Mejoras Sugeridas:

1. **Query expansion:** Expandir consultas con sinónimos técnicos
2. **Context window:** Aumentar ventana de contexto para documentos largos
3. **Multi-document synthesis:** Combinar información de múltiples fuentes

## D.8 Métricas de Calidad por Tipo de Consulta

### D.8.1 Análisis Cuantitativo

| Tipo de Consulta | Precision@5 | BERTScore F1 | Beneficio Reranking |
| --- | --- | --- | --- |
| Configuración | 0.073 | 0.742 | +15% |
| Troubleshooting | 0.045 | 0.728 | +45% |
| Conceptual | 0.038 | 0.751 | +25% |
| Implementación | 0.052 | 0.735 | +35% |

**Observaciones:** - **Consultas de configuración:** Mejor rendimiento en métricas tradicionales - **Troubleshooting:** Mayor beneficio del reranking (documentos específicos promovidos) - **Calidad semántica:** Consistente entre tipos (BERTScore ~0.73-0.75)

## D.9 Conclusiones de los Ejemplos

### D.9.1 Hallazgos Principales

1. **Sistema efectivo para consultas específicas:** Especialmente configuración y troubleshooting directo
2. **Reranking valioso:** Especialmente para promover documentos específicos sobre generales
3. **Limitación de evaluación:** Ground truth subestima efectividad real del sistema
4. **Calidad semántica alta:** BERTScore indica documentos útiles incluso cuando no “oficialmente” relevantes

### D.9.2 Implicaciones Prácticas

Para implementación en producción: - **Combinar métricas:** No depender solo de precision/recall tradicional - **Feedback de usuarios:** Incorporar evaluación humana para mejora continua - **Especialización por dominio:** Ajustar embeddings para terminología técnica específica

**Nota:** Todos los ejemplos presentados están basados en datos reales del corpus experimental (julio 2025) y representan comportamiento típico del sistema en diferentes escenarios de consulta.