

Unified Schema dla AquaRAG - Zoptymalizowany pod LLM Reasoning

Jedna struktura dla wszystkiego

Dlaczego to będzie lepsze dla reasoning:

1. Single Source of Truth

- `full_content` = wszystko co LLM potrzebuje w jednym polu
- Nie trzeba łączyć informacji z różnych części JSON
- Model dostaje kompletny kontekst od razu

2. Query-Optimized Fields

- `query_text` = zoptymalizowane pod wyszukiwanie (keywords, synonimy)
- `intent` = co użytkownik chce osiągnąć
- `difficulty` = dopasowanie odpowiedzi do poziomu

3. Reasoning-Friendly

- Consistent fields napříč wszystkimi typami treści
- Łatwe filtrowanie w retry logic
- Clear categorization dla refined searches

```
json
{
  "id": "af-gel-fix-001",
  "content_type": "product|faq|guide",
  "lang": "pl",
  "title": "AF Gel Fix - klej żelowy do koralowców",
  "query_text": "AF Gel Fix klej żelowy koralowce 10 sekund wiązanie nietoksyczny aquascaping",
  "full_content": "WSZYSTKO w jednym miejscu - pełny opis, instrukcje, dawkowanie, ceny, kompat",
  "domain": "seawater|freshwater|universal",
  "category": "aquascaping|parameters|bacteria|problems|education",
  "intent": ["repair", "maintenance", "startup", "troubleshooting"],
  "difficulty": "beginner|intermediate|expert",
  "url": "https://aquaforest.eu/pl/produkty/...",
  "tags": ["klej", "10_sekund", "koralowce", "nietoksyczny"],
  "related_items": ["af-frag-rocks", "af-mini-rocks"],
  "updated_at": "2025-05-28"
}
```

Przykłady dla różnych typów treści

PRODUKT:

```
json

{
  "id": "af-gel-fix-001",
  "content_type": "product",
  "lang": "pl",
  "title": "AF Gel Fix – klej żelowy do koralowców",
  "query_text": "AF Gel Fix klej żelowy koralowce aquascaping mocowanie szczepki 10 sekund wiąza",
  "full_content": "AF Gel Fix to szybkowiążący klej żelowy do precyzyjnego mocowania szczepki k",
  "domain": "seawater",
  "category": "aquascaping",
  "intent": ["coral_mounting", "aquascaping", "repair"],
  "difficulty": "beginner",
  "url": "https://aquaforest.eu/pl/produkty/seawater/aquascaping/af-gel-fix/",
  "tags": ["klej", "10_sekund", "żel", "nietoksyczny", "precyzyjny", "koralowce"],
  "related_items": ["af-frag-rocks", "af-mini-rocks", "af-plug-rocks"],
  "updated_at": "2025-05-28"
}
```

FAQ:

```
json

{
  "id": "faq-pl-001",
  "content_type": "faq",
  "lang": "pl",
  "title": "Jak obniżyć azotany (NO3) w akwarium rafowym?",
  "query_text": "azotany NO3 obniżyć zmniejszyć akwarium rafowe koralowce problem wysokie",
  "full_content": "Wysokie azotany (NO3) to częsty problem w akwariach rafowych. Oto skuteczne",
  "domain": "seawater",
  "category": "problems",
  "intent": ["troubleshooting", "water_parameters", "maintenance"],
  "difficulty": "intermediate",
  "url": "https://aquaforest.eu/pl/faq/azotany-obnizenie",
  "tags": ["NO3", "azotany", "problemy", "-NP_Pro", "Pro_Bio_S"],
  "related_items": ["af-np-pro", "pro-bio-s"],
  "updated_at": "2025-05-28"
}
```

KNOWLEDGE BASE:

json

```
{
  "id": "guide-pl-001",
  "content_type": "guide",
  "lang": "pl",
  "title": "Kompletny przewodnik po cyklu azotowym w akwarium morskim",
  "query_text": "cykl azotowy akwarium morskie bakterie nitryfikacyjne NH3 NO2 NO3 startup dojr",
  "full_content": "Cykl azotowy to fundament każdego zdrowego akwarium morskiego. To proces bic",
  "domain": "seawater",
  "category": "education",
  "intent": ["learning", "startup", "understanding"],
  "difficulty": "beginner",
  "url": "https://aquaforest.eu/pl/poradniki/cykl-azotowy",
  "tags": ["cykl_azotowy", "bakterie", "NH3", "NO2", "NO3", "startup", "Bio_S"],
  "related_items": ["bio-s", "af-bio-sand", "pro-bio-s"],
  "updated_at": "2025-05-28"
}
```

 **Dlaczego ten schemat jest IDEALNY dla Twojego reasoning workflow:**

Reasoning Step 1: Initial Search

```
python

# Pinecone search na query_text + filtry
results = pinecone.query(
    vector=embed(user_query),
    filter={
        "domain": detected_domain,
        "intent": {"$in": detected_intents}
    },
    top_k=5
)
```

Reasoning Step 2: LLM Evaluation

```
python

# LLM dostaje full_content - wszystko co potrzeba do odpowiedzi
confidence = llm.evaluate(
    query=user_query,
    context=results[0].metadata['full_content'],
    url=results[0].metadata['url'] # 🚀 NOWE!
)
```

Reasoning Step 3: Refined Search (jeśli potrzeba)

python

```
# Refined search z lepszymi filtrami
refined_results = pinecone.query(
    vector=embed(refined_query),
    filter={
        "category": specific_category,
        "difficulty": user_level
    }
)
```

💡 Dlaczego URL jest kluczowe:

1. Better User Experience

- LLM może odpowiedzieć: "Więcej szczegółów znajdziesz tutaj: [link]"
- User ma bezpośredni dostęp do źródła
- Zwiększona wiarygodność odpowiedzi

2. Tracking & Analytics

- Monitoring które produkty są najpopularniejsze
- Click-through rates z RAG do strony
- A/B testing różnych strategii linkowania

3. Legal & Compliance

- Jasne źródła informacji
- Możliwość weryfikacji przez użytkownika
- Attribution do oryginalnych treści

4. Enhanced RAG Responses

python

```
def generate_response(query, context, url):  
    prompt = f"""  
    Kontekst: {context}  
  
    Użytkownik pyta: {query}  
  
    Odpowiedz na pytanie używając kontekstu.  
    Na końcu dodaj: "Więcej informacji: {url}"  
    """  
    return llm.generate(prompt)
```

Migration Strategy:

Week 1: Schema Design

- Finalize unified schema **z URL**
- Create validation rules
- Test with 2-3 products

Week 2: Data Migration

- Convert existing 16 products + **add URLs**
- Add 5-10 FAQ items
- Add 2-3 knowledge base articles

Week 3: RAG Testing

- Test reasoning workflow **z URL integration**
- Compare with old structure
- Iterate on schema if needed





Week 4: Scale Up

- Convert remaining products
- Add more FAQ/guides
- Production deployment

Rekomendacja: **ZRÓB TO!**

Twoja intuicja jest 100% słuszna. Jeden zoptymalizowany schemat **z URL** będzie:

-  **Szybszy w development**

-  Łatwiejszy w maintenance
-  Lepszy dla LLM reasoning
-  Bardziej skalowalny
-  Better user experience (URLs!)

16 produktów to 1-2 dni pracy na conversion. Zyskasz miesiące czasu w przyszłości!